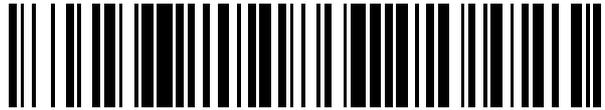


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 628 345**

51 Int. Cl.:

**C12N 9/42** (2006.01)

**C07K 14/435** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **23.06.2011 PCT/EP2011/060577**

87 Fecha y número de publicación internacional: **05.01.2012 WO12000892**

96 Fecha de presentación y número de la solicitud europea: **23.06.2011 E 11726853 (2)**

97 Fecha y número de publicación de la concesión europea: **22.03.2017 EP 2588603**

54 Título: **Polipéptido que tiene o que ayuda en la actividad de degradación de materiales glucídicos y usos del mismo**

30 Prioridad:

**29.06.2010 US 359563 P**  
**29.06.2010 EP 10167771**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**02.08.2017**

73 Titular/es:

**DSM IP ASSETS B.V. (100.0%)**  
**Het Overloon, 1**  
**6411 TE Heerlen, NL**

72 Inventor/es:

**SCHOONEVELD-BERGMANS, MARGOT,**  
**ELISABETH, FRANCOISE;**  
**HEIJNE, WILBERT HERMAN MARIE y**  
**LOS, ALRIK PIETER**

74 Agente/Representante:

**LEHMANN NOVO, María Isabel**

**ES 2 628 345 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

Polipéptido que tiene o que ayuda en la actividad de degradación de materiales glucídicos y usos del mismo

### Campo de la invención

5 La invención se refiere a secuencias que comprenden genes que codifican polipéptidos que tienen o ayudan en la actividad de degradación de materiales glucídicos. La invención la secuencia codificante de longitud completa del nuevo gen así como la secuencia de aminoácidos de la proteína funcional de longitud completa, y variantes y fragmentos del gen o la secuencia de aminoácidos. La invención también se refiere a métodos para usar estas proteínas en procedimientos industriales. También se incluyen en la invención células transformadas con un polinucleótido según la invención adecuadas para producir estas proteínas. Además, la invención se refiere a la expresión satisfactoria de los genes que codifican polipéptidos que tienen actividad de degradación de materiales glucídicos en un hospedador. El hospedador puede ser cualquier hospedador adecuado, a modo de ejemplo *Aspergillus*, p. ej. *Aspergillus niger* o *Talaromyces*, p. ej. *Talaromyces emersonii*.

### Antecedentes de la invención

15 Los carbohidratos constituyen los compuestos orgánicos más abundantes de la tierra. Sin embargo, gran parte de este carbohidrato está secuestrado en polímeros complejos incluyendo almidón (el principal carbohidrato de almacenamiento en semillas y cereales), y un conjunto de carbohidratos y lignina conocido como lignocelulosa. Los principales componentes de carbohidrato de la lignocelulosa son celulosa, hemicelulosa y pectinas. Estos polímeros complejos se denominan a menudo colectivamente lignocelulosa.

20 La bioconversión de biomasa lignocelulósica renovable en un azúcar fermentable que posteriormente se fermenta para producir alcohol (p. ej., etanol) como una alternativa a los combustibles líquidos ha atraído una atención intensa de los investigadores desde los 1970, cuando estalló la crisis del petróleo debido a una disminución de la producción de petróleo por la OPEC. El etanol se ha usado ampliamente como una combinación al 10% con gasolina en los EE. UU. de A. o como un combustible puro para vehículos en Brasil en las últimas dos décadas. Más recientemente, el uso de E85, una combinación de etanol al 85% se ha puesto en práctica especialmente para aplicaciones en ciudades limpias. La importancia del combustible bioetanol se incrementará en paralelo con los incrementos de los precios del petróleo y el agotamiento gradual de sus fuentes. Adicionalmente, se están usando azúcares fermentables para producir plásticos, polímeros y otros productos basados biológicamente y se espera que esta industria crezca sustancialmente, incrementado por lo tanto la demanda de azúcares fermentables abundantes de bajo coste que se puedan usar como una materia prima en lugar de materias primas basadas en petróleo.

30 El secuestro de estas grandes cantidades de carbohidratos en biomasa vegetal proporciona una fuente abundante de energía potencial en la forma de azúcares, azúcares tanto de cinco carbonos como de seis carbonos que se podrían utilizar para numerosos procedimientos industriales y agrícolas. Sin embargo, el enorme potencial de energía de estos carbohidratos actualmente está infrutilizado debido a que los azúcares están bloqueados en polímeros complejos, y de ahí que no sean fácilmente accesibles para la fermentación. Métodos que generen azúcares a partir de biomasa vegetal proporcionarían materias primas económicamente competitivas abundantes para la fermentación en productos químicos, plásticos, tales como, as a modo de ejemplo, ácido succínico y (bio)combustibles, incluyendo combustibles líquidos sintéticos de etanol, metanol y butanol y biogás.

40 Independientemente del tipo de materia prima celulósica, el coste y la eficacia hidrolítica de las enzimas son factores importantes que restringen la comercialización de los procedimientos de conversión de biomasa. Los costes de producción de enzimas producidas microbianamente están estrechamente conectados con una productividad de la cepa productora de enzimas y el rendimiento de actividad final en el caldo de fermentación.

45 A pesar de la investigación continuada en las últimas décadas para entender la degradación enzimática de biomasa lignocelulósica y la producción de celulosa, sigue siendo deseable descubrir o manipular nuevas celulasas y hemicelulasas muy activas. También sería muy deseable construir composiciones enzimáticas muy eficaces capaces de realizar una biodegradación rápida y eficaz de materiales lignocelulósicos, en particular tales estas celulasas y hemicelulasas que tengan termoestabilidad incrementada.

50 Estas enzimas se pueden usar para producir azúcares para la fermentación en productos químicos, plásticos, tales como, a modo de ejemplo, ácido succínico y (bio)combustibles, incluyendo etanol, metanol, butanol, combustibles líquidos sintéticos y biogás, para ensilaje, y también como enzima en otros procedimientos industriales, por ejemplo en las industrias alimentaria o de piensos, textil, de pasta papelera o papel o detergentes y otras industrias.

**Sumario de la invención**

La presente invención proporciona polinucleótidos que codifican polipéptidos que tienen la capacidad de degradar (es decir ayudar en la degradación de) un carbohidrato (por ejemplo polisacáridos), en particular, lignocelulosa. Los polinucleótidos de la invención típicamente codifican un polipéptido que tiene o que ayuda en la actividad de degradación de carbohidratos.

La invención también proporciona polipéptidos producidos naturalmente y recombinantemente que tienen esta actividad, así como líneas celulares recombinantes que producen estas enzimas. Además, se proporcionan métodos para elaborar y usar los polinucleótidos y polipéptidos de la invención.

Según la invención, se proporciona un polipéptido que comprende la secuencia de aminoácidos indicada en SEQ ID N°: 2 o una secuencia de aminoácidos codificada por la secuencia nucleotídica de SEQ ID N°: 1, o un polipéptido variante o polinucleótido variante del mismo, en donde el polipéptido variante tiene al menos 95% de identidad de secuencia con la secuencia indicada en SEQ ID N°: 2 o el polinucleótido variante codifica un polipéptido que tiene al menos 95% de identidad de secuencia con la secuencia indicada en SEQ ID N°: 2 y en donde el polipéptido tiene o ayuda en la actividad de degradación de materiales glucídicos .

Los polipéptidos según la invención tienen propiedades favorables, en particular la propiedad de tener o ayudar en la actividad de degradación de materiales glucídicos . En una realización, el polipéptido según la invención tiene actividad de oxidohidrolasa. En una realización, el polipéptido tiene actividad de GH61.

En una realización, el polipéptido variante tiene un residuo His27 (posición como en SEQ ID N°: 2). Se supone en la presente que His27 representa un papel en la unión de iones metálicos ( $Ni^{2+}$  y  $Mg^{2+}$  y potencialmente otros cationes). En una realización, el polipéptido variante tiene al menos 95% de identidad de secuencia con la secuencia indicada en SEQ ID N°: 2, un residuo catalítico His27 y actividad de oxidohidrolasa y/o GH61.

Además, los polipéptidos tienen alta termoestabilidad. Los polipéptidos según la invención retienen una alta actividad relativa (% de la actividad inicial) de, como una función del tiempo de incubación (h), p. ej. 2 horas, 3 horas, 4 horas, cinco horas, seis horas, ocho horas, nueve horas, 10 h o más, 20 h o más, 30 h o más, en particular a altas temperaturas, a modo de ejemplo a 60°C o más, a 65°C o más, o a 70°C o más, p. ej. 8 horas a 65°C o 72 horas a 60°C.

La invención también proporciona un polinucleótido que comprende:

(a) la secuencia nucleotídica indicada en SEQ ID N°: 1; o

(b) una secuencia nucleotídica que se hibrida selectivamente con un polinucleótido que es el complemento inverso de SEQ ID N°: 1; o

(c) una secuencia nucleotídica que tiene al menos aproximadamente 95% de identidad de secuencia con la secuencia nucleotídica de SEQ ID N°: 1; o

(d) una secuencia que está degenerada como resultado del código genético en una secuencia según se define en uno cualquiera de (a), (b) o (c);o

(e) una secuencia nucleotídica que es el complemento inverso de una secuencia nucleotídica según se define en (a), (b), (c) o (d).

También se proporciona según la invención un vector, tal como un vector de expresión, que incorpora una secuencia polinucleotídica de la invención y una célula que comprende un polipéptido, un polinucleótido o un vector de la invención.

La invención también proporciona:

un método para la preparación de un polipéptido que tiene o que ayuda en la actividad de degradación de carbohidratos, método que comprende cultivar una célula de la invención bajo condiciones que permitan la expresión de dicho polipéptido y, opcionalmente, recuperar el polipéptido expresado;

un polipéptido obtenible mediante tal método; y

una composición que comprende: (i) un polipéptido de la invención y; (ii) a celulasa y/o a hemicelulasa y/o una pectinasa.

5 Los polipéptidos de la invención que tienen o ayudan en la actividad de degradación de carbohidratos se pueden usar en procedimientos industriales. Así, la invención proporciona un método para el tratamiento de un sustrato que comprende material glucídico, método que comprende poner en contacto el sustrato con un polipéptido o una composición de la invención.

10 En particular, la invención proporciona un método para producir un azúcar o azúcares a partir de material lignocelulósico, método que comprende poner en contacto el material lignocelulósico con un polipéptido o una composición de la invención.

15 Los azúcares producidos de este modo se pueden usar en un procedimiento de fermentación. Según esto, la invención proporciona un método para producir un producto de fermentación, método que comprende: producir un azúcar fermentable usando lo descrito anteriormente; y fermentar el azúcar fermentable resultante, para producir de ese modo un producto de fermentación.

20 Un polipéptido o una composición de la invención también se puede usar, por ejemplo, en la preparación de un producto alimenticio, en la preparación de un detergente, en la preparación de un pienso, en el tratamiento de pasta papelera o en la fabricación de papel o en la preparación de una tela o material textil o en la limpieza de los mismos.

Se divulga en la presente:

un material procesado obtenible al poner en contacto un material vegetal o material lignocelulósico con un polipéptido o una composición de la invención;

un alimento o pienso que comprende un polipéptido o una composición de la invención; y

25 una planta o parte de la misma que comprende un polinucleótido, un polipéptido, un vector o una célula según la invención.

### Breve descripción de los dibujos

30 Fig 1: Mapa de pGBTOP para la expresión de genes en *A. niger*. Se representan el gen de interés (GOI) expresado a partir del promotor de glucoamilasa (PglA). Además, se representa el flanco de glucoamilasa (3'glA) del casete de expresión. En esta solicitud, un gen de interés es la secuencia codificante de TEMER07589 según se define posteriormente en la presente.

35 Fig 2: Gráfica de liberación de glucosa, expresada en mmol/l, en el tiempo, a partir de paja de trigo pretratada con 2% de dm, incubada con 15 mg de proteína como equivalente de BSA por gramo de materia seca de la paja de trigo pretratada, mediante una mezcla de 4 enzimas, que contiene 1 BG, 1 CBHI, 1 CBHII y 1 EG o que ayuda en la actividad de degradación de carbohidratos (--●-- = CEA; --x-- = CEB; -■- = TEMER07589) o mediante un producto de celulasa clásico (--▲-- = Filtrase® NL).

### Breve descripción de la lista de secuencias

SEQ ID N°: 1 indica la secuencia codificante de TEMER07589;

SEQ ID N°: 2 indica la secuencia de aminoácidos de TEMER07589;

40 SEQ ID N°: 3 indica la secuencia de señal de TEMER07589.

### Descripción detallada de la invención

45 A lo largo de la presente memoria descriptiva y las reivindicaciones adjuntas, las palabras "comprender" e "incluir" y variación tales como "comprende", "que comprende", "incluye" y "que incluye" se han de interpretar inclusivamente. Esto es, estas palabras están destinadas a comunicar la posible inclusión de otros elementos o números enteros no citados específicamente, cuando el contexto lo permita.

Los artículos "un" y "uno/a" se usan en la presente para referirse a uno o más de uno (es decir a uno o al menos uno) del objeto gramatical del artículo. A modo de ejemplo, "un elemento" puede significar un elemento o más de un elemento.

- 5 La presente invención proporciona polinucleótidos que codifican polipéptidos, p. ej. enzimas que tienen la capacidad de modificar, por ejemplo degradar, un material glucídico. A material glucídico es un material que comprende, consiste en o consiste sustancialmente en uno o más carbohidratos. Las enzimas son en la presente una subclase de los polipéptidos.
- 10 Sustrato (también llamado materia prima) se usa en la presente para referirse a una sustancia que comprende material glucídico, que se puede tratar con enzimas según la invención, de modo que el material glucídico del mismo se modifique. Además del material glucídico, el sustrato puede contener cualquier otro componente, incluyendo, pero no limitado a, un material no glucídico y almidón.
- 15 La presente invención proporciona polinucleótidos que codifican polipéptidos, p. ej. enzimas que tienen la capacidad de modificar, por ejemplo degradar, un material glucídico. Un material glucídico es un material que comprende, consiste en o consiste sustancialmente en uno o más carbohidratos. Las enzimas son en la presente una subclase de los polipéptidos.
- 20 Típicamente, un polipéptido de la invención codifica un polipéptido que tiene al menos o que ayuda en la actividad de degradación de carbohidratos, llamado tentativamente TEMER07589, que tiene una secuencia de aminoácidos según SEQ ID N°: 2, o una secuencia que es una variante del mismo, típicamente funcionalmente equivalente al polipéptido que tiene la secuencia de SEQ ID N°: 2, o una secuencia que es un fragmento de cualquiera de los mismos.
- 25 "Que tiene o que ayuda en la actividad de degradación de carbohidratos" se define en la presente como que el polipéptido tiene actividad de degradación de carbohidratos o que el polipéptido ayuda en la degradación de carbohidratos o ambos. En una realización, el polipéptido potencia la actividad de al menos una celulasa. En esa realización, cuando el polipéptido de la invención está presente en una mezcla con una o más celulasas, a modo de ejemplo, en una mezcla con celobiohidrolasa (CBH) y beta-glucosidasa (BG), potenciará la actividad de estas celulasas, lo que dará como resultado una actividad superior de la mezcla para degradar celulosa. Se cree que TEMER07589 pertenece a la familia GH61 de enzimas. Las enzimas de esta familia GH61 se clasificaron originalmente como una familia de glucósido hidrolasas basándose en la medida de la actividad de endo-1,4-b-D-glucanasa muy débil en un miembro de la familia (endoglucanasa (EC 3.2.1.4)). La estructura y el modo de acción de estas enzimas son ciertamente no canónicos y no se pueden considerar glucosidasas auténticas. Sin embargo,
- 30 se mantienen en la clasificación CAZy basándose en su capacidad para potenciar la descomposición de lignocelulosa cuando se usan junto con una celulasa o una mezcla de celulasas. Una visión general de miembros conocidos de la familia GH61 se da en la figura 5 de Harris, PV y cols., Biochemistry 2010, 49, 3305-3316.
- 35 En una realización, además de la actividad potenciadora de celulasa, la proteína potenciadora de celulasa según la invención tiene actividad de endoglucanasa (EG). En esta realización, se puede evitar la adición de una endoglucanasa (distinta a la proteína que potencia la celulasa según la invención) a una mezcla de celulasas, que habitualmente es esencial para la degradación eficaz de celulosa.
- 40 En la presente, una endo- $\beta$ -1,4-glucanasa (EC 3.2.1.4) es cualquier polipéptido que sea capaz de catalizar la endohidrólisis de enlaces 1,4- $\beta$ -D-glucosídicos en celulosa, liquenina o  $\beta$ -D-glucano de cereales. Este polipéptido también puede ser capaz de hidrolizar enlaces 1,4 en  $\beta$ -D-glucanos que también contienen enlaces 1,3. Esta enzima también se puede denominar celulasa, avicelasa,  $\beta$ -1,4-endoglucano hidrolasa,  $\beta$ -1,4-glucanasa, carboximetil celulasa, celudextrinasa, endo-1,4- $\beta$ -D-glucanasa, endo-1,4- $\beta$ -D-glucanohidrolasa, endo-1,4- $\beta$ -glucanasa o endoglucanasa.
- 50 En una realización, un polipéptido de la invención puede tener una o más actividades alternativas y/o adicionales distintas a la actividad potenciadora de celulasa y la actividad de endoglucanasa que se mencionan anteriormente, por ejemplo una de las otras actividades de degradación de carbohidratos y/o hidrolización de carbohidratos mencionadas en la presente.
- 55 Carbohidrato en este contexto incluye todos los sacáridos, por ejemplo polisacáridos, oligosacáridos, disacáridos o monosacáridos.
- 60 Un polipéptido según la invención puede modificar y/o degradar un material glucídico al degradar químicamente o degradar físicamente este material o hidrolizar el carbohidrato. Física incluye, p. ej., la interrupción de la interacción entre microfibrillas de celulosa y/o la apertura de la estructura de las fibras de celulosa. La modificación química del material glucídico puede dar como resultado la degradación de este material, por ejemplo mediante hidrólisis, oxidación u otra modificación química tal como mediante la acción de una liasa. La modificación física puede estar acompañada o no por modificación química.
- 65

## Materiales glucídicos adecuados

- Un carbohidrato no amiláceo adecuado para la modificación por un polipéptido de la invención es la lignocelulosa. Los principales polisacáridos que comprenden diferentes residuos lignocelulósicos, que se pueden considerar una materia prima renovable potencial, son la celulosa (glucanos), hemicelulosas (xilanos, heteroxilanos y xiloglucanos).
- 5 Además, alguna hemicelulosa puede estar presente como glucomananos, por ejemplo en materias primas derivadas de madera. La hidrólisis enzimática de estos polisacáridos hasta azúcares solubles, por ejemplo glucosa, xilosa, arabinosa, galactosa, fructosa, manosa, ramnosa, ribosa, ácido D-galacturónico y otras hexosas y pentosas, se produce bajo la acción de diferentes enzimas que actúan en conjunto.
- 10 Además, pectinas y otras sustancias pécticas tales como arabinanos pueden constituir una proporción considerable de la masa seca, típicamente de paredes celulares de tejidos vegetales no leñosos (de aproximadamente un cuarto a la mitad de la masa seca pueden ser pectinas).
- 15 La celulosa es un polisacárido lineal compuesto por residuos de glucosa ligados por enlaces  $\beta$ -1,4. La naturaleza lineal de las fibras celulósicas, así como la estequiometría de la glucosa liada en  $\beta$  (con relación a  $\alpha$ ) genera estructuras más propensas a entrelazar el enlace de hidrógeno que las estructuras de almidón reticuladas en  $\alpha$  altamente ramificadas. Así, los polímeros celulósicos son generalmente menos solubles, y forman fibras más fuertemente unidas que las fibras encontradas en el almidón.
- 20 La hemicelulosa es un polímero complejo, y su composición a menudo varía ampliamente de organismo a organismo, y de un tipo de tejido a otro. En general, un componente principal de la hemicelulosa es xilosa conectada por  $\beta$ -1,4, un azúcar de cinco carbonos. Sin embargo, esta xilosa está ramificada a menudo en O-3 y/u O-2 y puede estar sustituida con enlaces con arabinosa, galactosa, manosa, ácido glucurónico, ácido galacturónico o por esterificación a ácido acético (y esterificación de ácido ferúlico a arabinosa). La hemicelulosa también puede
- 25 contener glucano, que es un término general para azúcares de seis carbonos enlazados en  $\beta$  (tales como los glucanos y heteroglucanos  $\beta$ -(1,3)(1,4) mencionados previamente) y adicionalmente glucomananos (en los que están presentes tanto glucosa como manosa en el esqueleto lineal, enlazadas entre sí por enlaces  $\beta$ ).
- 30 La composición, la naturaleza de la sustitución y el grado de ramificación de la hemicelulosa son muy diferentes en plantas dicotiledóneas (es decir, plantas cuyas semillas tienen dos cotiledones u hojas primordiales tales como habas, cacahuetes, almendras, guisantes, habichuelas) en comparación con las plantas monocotiledóneas (es decir, plantas que tienen un solo cotiledón u hoja primordial tales como maíz, trigo, arroz, gramíneas, cebada). En las dicotiledóneas, la hemicelulosa está comprendida principalmente por xiloglucanos que son cadenas de glucosa con enlaces 1,4- $\beta$  con cadenas laterales de xilosilo con enlaces 1,6- $\beta$ . En las monocotiledóneas, incluyendo la mayoría
- 35 de los cultivos de cereales, los principales componentes de la hemicelulosa son heteroxilanos. Estos están comprendidos principalmente por polímeros de esqueleto de xilosa con enlaces 1,4- $\beta$  con enlaces 1,3- $\alpha$  con arabinosa, galactosa, manosa y ácido glucurónico o ácido 4-O-metil-glucurónico así como xilosa modificada por ácidos acéticos con enlaces éster. También se presentan  $\beta$ -glucanos comprendidos por cadenas de glucosilo con enlaces 1,3- y 1,4- $\beta$ . En las monocotiledóneas, pueden estar presentes celulosa, heteroxilanos y  $\beta$ -glucanos en cantidades aproximadamente iguales, comprendiendo cada uno aproximadamente 15-25% de la materia seca de las
- 40 paredes celulares. Además, diferentes plantas pueden comprender diferentes cantidades de, y diferentes composiciones de, sustancias pécticas. Por ejemplo, la remolacha azucarera contiene aproximadamente 19% de pectina y aproximadamente 21% de arabinano en una base en peso seco.
- 45 Según esto, una composición de la invención se puede adaptar en vista de la materia prima (también llamada sustrato) particular que se vaya a usar. Es decir, el espectro de actividades en una composición de la invención puede variar dependiendo de la materia prima en cuestión.
- 50 Las combinaciones de enzimas o los tratamientos físicos se pueden administrar simultáneamente o secuencialmente. Las enzimas pueden producirse exógenamente en microorganismos, levadoras, hongos, bacterias o plantas, a continuación aislarse y añadirse a la materia prima lignocelulósica. Alternativamente, las enzimas se producen, pero no se aíslan, y se añaden a la materia prima caldo de fermentación de la masa celular en bruto, o material vegetal (tal como forraje de maíz) y similares. Alternativamente, la masa celular en bruto o el medio de
- 55 producción de enzimas o el material vegetal puede tratarse para prevenir el crecimiento microbiano adicional (por ejemplo, mediante calentamiento o adición de agentes antimicrobianos), y a continuación añadirse a la materia prima. Estas mezclas de enzimas en bruto pueden incluir el organismo que produce la enzima. Alternativamente, la enzima se puede producir en una fermentación que usa materia prima (tal como forraje de maíz) para proporcionar nutrición a un organismo que produce una enzima o enzimas. De este modo, plantas que producen las enzimas pueden servir como la materia prima lignocelulósica y añadirse a la materia prima lignocelulósica.
- 60

## Actividad enzimática

Las endo-1,4- $\beta$ -glucanasas (EG) y exo-celobiohidrolasas (CBH) catalizan la hidrólisis de celulosa insoluble hasta celooligosacáridos (celobiosa como producto principal), mientras que las  $\beta$ -glucosidasas (BGL) convierten los oligosacáridos, principalmente celobiosa y celotriosa, en glucosa.

5 Las xilanasas junto con enzimas auxiliares, por ejemplo  $\alpha$ -L-arabinofuranosidasas, feruloil y acetilxilano esterasas, glucuronidasas y  $\beta$ -xilosidasas) catalizan la hidrólisis de parte de las hemicelulosas.

10 Las sustancias pécticas incluyen pectinas, arabinanos, galactanos y arabinogalactanos. Las pectinas son los polisacáridos más complejos de la pared celular de las plantas. Se acumulan alrededor de una cadena central de unidades de ácido D-galacturónico con enlaces  $\alpha$ (1,4) intercaladas en algún grado con L-ramnosa. En una pared celular cualquiera hay un número de unidades estructurales que se ajusta a esta descripción y generalmente se ha considerado que en una molécula péctica individual, las cadenas centrales de diferentes unidades estructurales son contiguas entre sí.

15 Las pectinasas incluyen, por ejemplo, una endopoligalacturonasa, una pectina metil esterasa, una endogalactanasa, una betagalactosidasa, una pectina acetil esterasa, una endopectina liasa, pectato liasa, alfa-ramnosidasa, una exogalacturonasa, una expoligalacturonato liasa, una ramnogalacturonano hidrolasa, una ramnogalacturonano liasa, una ramnogalacturonano acetil esterasa, una ramnogalacturonano galacturonohidrolasa, una xilogalacturonasa, una  $\alpha$ -arabinofuranosidasa.

20 Los principales tipos de unidad estructural son: galacturonano (homogalacturonano), que puede estar sustituido con metanol en el grupo carboxilo y acetato en O-2 y O-3; ramnogalacturonano I (RGI), en el que las unidades de ácido galacturónico se alternan con unidades de ramnosa que soportan cadenas laterales de galactano con enlaces (1,4) y arabinano con enlaces (1,5). Las cadenas laterales de arabinano pueden estar unidas directamente a la ramnosa o indirectamente a través de las cadenas de galactano; xilogalacturonano, con unidades de xilosilo individuales en O-3 de ácido galacturónico (estrechamente asociadas con RGI); y ramnogalacturonano II (RGII), una unidad secundaria particularmente compleja que contiene azúcares inhabituales, por ejemplo apiosa. Una unidad de RGII puede contener dos residuos apiosilo que, bajo condiciones iónicas adecuadas, pueden formar reversiblemente ésteres con borato.

25 Como se indica anteriormente, un polipéptido de la invención tendrá típicamente actividad potenciadora de celulasa. Sin embargo, un polipéptido de la invención puede tener una o más de las actividades indicadas anteriormente además de o alternativamente a esa actividad. Además, una composición de la invención según se describe en la presente puede tener una o más de las actividades mencionadas anteriormente además de la proporcionada por un polipéptido de la invención que tiene actividad potenciadora de celulasa.

## Secuencia polinucleotídica

40 La invención proporciona secuencias polinucleotídicas genómicas que comprende el gen que codifica el TEMER07589 así como su secuencia codificante. Según esto, la invención se refiere a un polinucleótido aislado que comprende la secuencia nucleotídica genómica según la secuencia nucleotídica codificante según SEQ ID N°: 1 y a variantes, tales como equivalentes funcionales, de cualquiera de los mismos.

45 En particular, la invención se refiere a un polinucleótido aislado que es capaz de hibridarse selectivamente, por ejemplo bajo condiciones rigurosas, preferiblemente bajo condiciones muy rigurosas, con el complemento inverso de un polinucleótido que comprende la secuencia indicada en SEQ ID N°: 1.

Más específicamente, la invención se refiere a un polinucleótido que comprende o que consiste esencialmente en una secuencia nucleotídica según SEQ ID N°: 1.

50 La invención también se refiere a un polinucleótido aislado que comprende o que consiste esencialmente en una secuencia que codifica al menos un dominio funcional de un polipéptido según SEQ ID N°: 2 o una variante del mismo, tal como un equivalente funcional, o un fragmento de cualquiera de los mismos.

55 Según se usan en la presente, los términos "gen" y "gen recombinante" se refieren a moléculas de ácido nucleico que se pueden aislar de ADN cromosómico, que incluyen un marco de lectura abierto que codifica una proteína, p. ej. la proteína que potencia celulasa según la presente invención.

60 Un gen puede incluir secuencias codificante, secuencias no codificantes, intrones y/o secuencias reguladoras. Por otra parte, el término "gen" se puede referir a una molécula de ácido nucleico aislada según se define en la presente.

- Una molécula de ácido nucleico de la presente invención, tal como una molécula de ácido nucleico que tiene la secuencia nucleotídica de SEQ ID N°: 1 o una variante de la misma, tal como un equivalente funcional, se puede aislar usando técnicas de biología molecular estándar y la información de secuencia proporcionada en la presente. Por ejemplo, usando la totalidad o una porción de la secuencia de ácido nucleico de SEQ ID N°: 1 como una sonda de hibridación, moléculas de ácido nucleico según la invención se pueden aislar usando técnicas de hibridación y clonación estándar (p. ej., según se describe en Sambrook, J., Fritsh, E. F., y Maniatis, T. *Molecular Cloning: A Laboratory Manual*. 2ª ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).
- Por otra parte, una molécula de ácido nucleico que abarca la totalidad o una porción de SEQ ID N°: 1 se puede aislar mediante la reacción en cadena de la polimerasa (PCR) usando cebadores oligonucleotídicos sintéticos diseñados basándose en la información de secuencia contenida en SEQ ID N°: 1.
- Un ácido nucleico de la invención se puede amplificar usando ADNc, ARNm o, alternativamente, ADN genómico, como una plantilla y cebadores oligonucleotídicos apropiados según técnicas de amplificación por PCR estándar. El ácido nucleico así amplificado se puede clonar en un vector apropiado y caracterizar mediante el análisis de la secuencia de ADN.
- Por otra parte, oligonucleótidos correspondientes a o hibridables con una secuencia nucleotídica según la invención se pueden preparar mediante técnicas sintéticas estándar, p. ej., usando un sintetizador de ADN automatizado.
- En una realización preferida, una molécula de ácido nucleico aislada de la invención comprende la secuencia nucleotídica mostrada en SEQ ID N°: 1.
- En otra realización preferida, una molécula de ácido nucleico aislada de la invención comprende una molécula de ácido nucleico que es el complemento inverso de la secuencia nucleotídica mostrada en SEQ ID N°: 1 o una variante, tal como un equivalente funcional, de cualquiera de tales secuencias nucleotídicas.
- Una molécula de ácido nucleico que es complementaria a otra secuencia nucleotídica es una que es suficientemente complementaria a la otra secuencia nucleotídica de modo que se pueda hibridar a la otra secuencia nucleotídica formando de ese modo un dúplex estable.
- Un aspecto de la invención trata de moléculas de ácido nucleico aisladas que codifican un polipéptido de la invención o una variante, tal como un equivalente funcional del mismo, por ejemplo un fragmento o dominio biológicamente activo, así como moléculas de ácido nucleico suficientes para el uso como sondas de hibridación para identificar moléculas de ácido nucleico que codifican un polipéptido de la invención y fragmentos de tales moléculas de ácido nucleico adecuados para el uso como cebadores de PCR para la amplificación o mutación de moléculas de ácido nucleico.
- Un polinucleótido según la invención se puede "aislar". En el contexto de esta invención, un "polinucleótido aislado" o "ácido nucleico aislado" en un ADN o ARN que no es inmediatamente contiguo con una o ambas de las secuencias codificantes con las que es inmediatamente contiguo (una en el extremo 5' y una en el extremo 3') en el genoma presente en la naturaleza del organismo del que se deriva. Así, en una realización, un ácido nucleico aislado incluye algunas o todas las secuencias no codificantes (p. ej. promotoras) 5' que son inmediatamente contiguas a la secuencia codificante. Por lo tanto, el término incluye, por ejemplo, un ADN recombinante que se incorpora en un vector, en un plásmido o virus que se replica autónomamente, o en el ADN genómico de un procarionte o eucariote, o que existe como una molécula separada (p. ej., un ADNc o un fragmento de ADN genómico producido por PCR o tratamiento con endonucleasas de restricción) independiente de otras secuencias. También incluye un ADN recombinante que es parte de un gen híbrido que codifica un polipéptido adicional que está sustancialmente libre de material celular, material viral o medio de cultivo (cuando se produce mediante técnicas de DNA recombinante), o precursores químicos u otros productos químicos (cuando se sintetiza químicamente). Por otra parte, un "fragmento de ácido nucleico aislado" es un fragmento de ácido nucleico que no está presente en la naturaleza como un fragmento y no se encontraría en estado natural.
- Según se usan en la presente, los términos "polinucleótido" o "molécula de ácido nucleico" pretenden incluir moléculas de ADN (p. ej., ADNc o ADN genómico) y moléculas de ARN (p. ej., ARNm) y análogos del ADN o ARN generados usando análogos nucleotídicos. La molécula de ácido nucleico puede ser de una sola hebra o de doble hebra, pero preferiblemente es ADN de doble hebra. El ácido nucleico se puede sintetizar usando análogos o derivados oligonucleotídicos (p. ej., nucleótidos de inosina o fosforotioato). Estos oligonucleótidos se pueden usar, por ejemplo, para preparar ácidos nucleicos que tienen capacidades de apareamiento de bases alteradas o resistencia a nucleasas incrementada.
- Otra realización de la invención proporciona una molécula de ácido nucleico aislada que es antisentido con respecto a una molécula de ácido nucleico de TEMER07589, p. ej., la hebra codificante de una molécula de ácido nucleico de

TEMER07589. También se incluyen dentro del alcance de la invención las hebras complementarias de las moléculas de ácido nucleico descritas en la presente.

5 A menos que se indique otra cosa, todas las secuencias nucleotídicas determinadas al someter a secuenciación a una molécula de ADN en la presente se determinaron usando un secuenciador de ADN automático y todas las secuencias de aminoácidos de los polipéptidos codificados por moléculas de ADN determinadas en la presente se predijeron mediante la traducción de una secuencia de ADN determinada como anteriormente. Por lo tanto, como se sabe en la técnica para cualquier secuencia de ADN determinada mediante este enfoque automático, cualquier secuencia nucleotídica determinada en la presente puede contener algunos errores. Las secuencias nucleotídicas  
10 determinadas automáticamente son típicamente al menos aproximadamente 90% idénticas, más típicamente de al menos aproximadamente 95% a al menos aproximadamente 99,9% idénticas a la secuencia nucleotídica real de la molécula de ADN sometida a secuenciación.

15 La secuencia real se puede determinar más precisamente mediante otros enfoques incluyendo métodos de secuenciación de ADN manuales muy conocidos en la técnica. Como también se conoce en la técnica, una sola inserción o eliminación en una secuencia nucleotídica determinada comparada con la secuencia real provocará un cambio de marco en la traducción de la secuencia nucleotídica de modo que la secuencia de aminoácidos predicha codificada por una secuencia nucleotídica determinada será completamente diferente de la secuencia de aminoácidos realmente codificada por la molécula de ADN sometida a secuenciación, empezando en el punto de  
20 esta inserción o eliminación.

El experto en la técnica es capaz de identificar estas bases erróneamente identificadas y sabe cómo corregir estos errores.

25 Una molécula de ácido nucleico según la invención puede comprender solo una porción o un fragmento de la secuencia de ácido nucleico mostrada en SEQ ID N°: 1 (o una variante de cualquiera de los mismos), por ejemplo un fragmento que se puede usar como una sonda o cebador o un fragmento que codifica una porción de un polipéptido TEMER07589.

30 La secuencia nucleotídica determinada a partir de la clonación del gen TEMER07589 y ADNc permite la generación de sondas y cebadores diseñados para el uso en la identificación y/o clonación de otros miembros de la familia de TEMER07589, así como homólogos de TEMER07589 procedentes de otras especies.

35 La sonda/el cebador comprende típicamente un oligonucleótido sustancialmente purificado que comprende típicamente una región de secuencia nucleotídica que se hibrida preferiblemente bajo condiciones muy rigurosas a al menos de aproximadamente 12 a aproximadamente 15, preferiblemente de aproximadamente 18 a aproximadamente 20, preferiblemente de aproximadamente 22 a aproximadamente 25, más preferiblemente aproximadamente 30, aproximadamente 35, aproximadamente 40, aproximadamente 45, aproximadamente 50, aproximadamente 55, aproximadamente 60, aproximadamente 65 o aproximadamente 75 o más nucleótidos consecutivos de una secuencia nucleotídica mostrada en SEQ ID N°: 1 o de una variante, tal como un equivalente funcional, de cualquiera de los mismos.  
40

Se pueden usar sondas basadas en las secuencias nucleotídicas de TEMER07589 para detectar transcritos o secuencias de TEMER07589 genómicas que codifican polipéptidos iguales u homólogos, a modo de ejemplo, en otros organismos. En realizaciones preferidas, la sonda comprende además un grupo marcador ligado a la misma, p. ej., el grupo marcador puede ser un radioisótopo, un compuesto fluorescente, una enzima o un cofactor enzimático. Estas sondas también se pueden usar como parte de un estuche para pruebas diagnósticas para identificar células que expresan un polipéptido TEMER07589.  
45

50 Los polinucleótidos de la presente pueden ser polinucleótidos sintéticos. Los polinucleótidos sintéticos se pueden optimizar en el uso de codones, preferiblemente según los métodos descritos en los documentos WO2006/077258 y/o PCT/EP2007/055943. El documento PCT/EP2007/055943 se dirige a la optimización de pares de codones. La optimización de pares de codones es un método en el que las secuencias nucleotídicas que codifican un polipéptido se han modificado con respecto a su utilización de codones, en particular los pares de codones que se usan, para  
55 obtener una expresión mejorada de la secuencia nucleotídica que codifica el polipéptido y/o una producción mejorada del polipéptido codificado. Los pares de codones se definen como un grupo de dos tripletes (codones) subsiguientes en una secuencia codificante.

60 La invención se refiere además a una construcción de ácido nucleico que comprende el polinucleótido que se describe anteriormente. "Construcción de ácido nucleico" se define en la presente como una molécula de ácido nucleico, bien de una sola hebra o bien de doble hebra, que se aísla de un gen presente en la naturaleza o que se ha modificado para contener segmentos de ácido nucleico que se combinan y yuxtaponen de un modo que de otra forma no existiría en la naturaleza. El término construcción de ácido nucleico es sinónimo del término "casete de expresión" cuando la construcción de ácido nucleico contiene todas las secuencias de control requeridas para la  
65 expresión de una secuencia codificante. El término "secuencia codificante", según se define en la presente, es una

- 5 secuencia que se transcribe en ARNm y se traduce en un activador transcripcional de un promotor de proteasa de la invención. Los límites de la secuencia codificante generalmente están determinados por el codón de iniciación ATG en el extremo 5' del ARNm y una secuencia codónica de terminación de la traducción que termina el marco de lectura abierto en el extremo 3' del ARNm. Una secuencia codificante puede incluir, pero no se limita a, ADN, ADNc, y secuencias de ácido nucleico recombinantes. Preferiblemente, el ácido nucleico tiene un alto contenido de GC. El contenido de GC en la presente indica el número de nucleótidos de G y C en la construcción, dividido por el número total de nucleótidos, expresado en %. El contenido de GC es preferiblemente 56% o mayor, 57% o mayor, 58% o mayor, 59% o mayor, 60% o mayor, o en el intervalo de 56-70% o el intervalo de 58-65%.
- 10 Preferiblemente, la construcción de ADN comprende una secuencia de ADN promotora, una secuencia codificante en asociación operativa con dicha secuencia de ADN promotora y secuencias de control tales como:
- una secuencia de terminación de la traducción orientada en la dirección 5' hacia 3' seleccionada de la siguiente lista de secuencias: TAAG, TAGA y TAAA, preferiblemente TAAA, y/o
- 15 una secuencia codificante iniciadora de la traducción orientada en la dirección 5' hacia 3' seleccionada de la siguiente lista de secuencias: GCTACCCCC; GCTACCTCC; GCTACCCTC; GCTACCTTC; GCTCCCCCC; GCTCCCTCC; GCTCCCCTC; GCTCCCTTC; GCTGCCCC; GCTGCCTCC; GCTGCCCTC; GCTGCCTTC; GCTTCCCC; GCTTCTCC; GCTTCCCTC; y GCTTCCTTC, preferiblemente GCT TCC TTC, y/o
- 20 una secuencia iniciadora de la traducción seleccionada de la siguiente lista de secuencias: 5'-mwChkyCAAA-3'; 5'-mwChkyCACA-3' o 5'-mwChkyCAAG-3', que usan códigos de ambigüedad para los nucleótidos: m (A/C); w (A/T); y (C/T); k (G/T); h (A/C/T), preferiblemente 5'-CACCGTCAAA-3' o 5'-CGCAGTCAAG-3'.
- 25 En el contexto de esta invención, el término "secuencia codificante iniciadora de la traducción" se define como los nueve nucleótidos inmediatamente aguas abajo del codón iniciador o de iniciación del marco de lectura abierto de una secuencia codificante de ADN. El codón iniciador o de iniciación codifica el aminoácido metionina. El codón iniciador es típicamente ATG, pero también puede ser cualquier codón de iniciación funcional tal como GTG.
- En el contexto de esta invención, el término "secuencia de terminación de la traducción" se define como los cuatro nucleótidos partiendo desde el codón de terminación de la traducción en el extremo 3' del marco de lectura abierto o la secuencia nucleotídica codificante y orientados en la dirección 5' hacia 3'.
- 30 En el contexto de esta invención, el término "secuencia iniciadora de la traducción" se define como los diez nucleótidos inmediatamente aguas arriba del codón iniciador o de iniciación del marco de lectura abierto de una secuencia de ADN que codifica un polipéptido. El codón iniciador o de iniciación codifica el aminoácido metionina. El codón iniciador es típicamente ATG, pero también puede ser cualquier codón de iniciación funcional tal como GTG. Se sabe bien en la técnica que el uracilo, U, reemplaza al desoxinucleótidos timina, T, en el ARN.
- 35 Homología e identidad
- Se dice que las secuencias de aminoácidos o nucleótidos son homólogas cuando exhiben un cierto nivel de similitud. Dos secuencias que son homólogas indican un origen evolutivo común. Que dos secuencias homólogas estén estrechamente relacionadas o más distantemente relacionadas se indica por el "porcentaje de identidad" o el "porcentaje de similitud", que es alto o bajo, respectivamente. Aunque es discutido, para indicar "porcentaje de identidad" o "porcentaje de similitud", frecuentemente se usan intercambiamente "nivel de homología" o "porcentaje de homología".
- 40 Los términos "homología", "porcentaje de homología", "porcentaje de identidad" o "porcentaje de similitud" se usan intercambiamente en la presente. Para el propósito de esta invención, se define aquí que a fin de determinar el porcentaje de identidad de dos secuencias de aminoácidos o de dos secuencias de ácido nucleico, las secuencias completas se alinean con propósitos de comparación óptimos. A fin de optimizar el alineamiento entre las dos secuencias se pueden introducir huecos en cualquiera de las dos secuencias que se comparan. Tal alineamiento se lleva a cabo a lo largo de toda la longitud de las secuencias que se comparan. Alternativamente, el alineamiento se puede llevar a cabo a lo largo de una longitud más corta, por ejemplo a lo largo de aproximadamente 20, aproximadamente 50, aproximadamente 100 o más ácidos nucleicos/bases o aminoácidos. La identidad es el porcentaje de emparejamiento idénticos entre las dos secuencias a lo largo de la región alineada presentada.
- 45 Una comparación de secuencias y una determinación del porcentaje de identidad entre dos secuencias se pueden efectuar usando un algoritmo matemático. El experto estará al tanto del hecho de que están disponibles varios programas informáticos diferentes para alinear dos secuencias y determinar la homología entre dos secuencias (Kruskal, J. B. (1983) An overview of squence comparison In D. Sankoff and J. B. Kruskal, (ed.), Time warps, string
- 55

5 edits and macromolecules: the theory and practice of sequence comparison, pp. 1-44 Addison Wesley). El porcentaje de identidad entre dos secuencias de aminoácidos se puede determinar usando el algoritmo de Needleman y Wunsch para el alineamiento de dos secuencias. (Needleman, S. B. y Wunsch, C. D. (1970) J. Mol. Biol. 48, 443-453). El algoritmo alinea secuencias de aminoácidos así como secuencias nucleotídicas. El algoritmo de Needleman-Wunsch se ha aplicado en el programa informático NEEDLE. Para el propósito de esta invención, se usó el programa NEEDLE del paquete EMBOSS (versión 2.8.0 o superior, EMBOSS: The European Molecular Biology Open Software Suite (2000) Rice, P. Longden, I. y Bleasby, A. Trends in Genetics 16, (6) pp276-277, <http://emboss.bioinformatics.nl/>). Para secuencias polipeptídicas, se usa EBLOSUM62 para la matriz de sustitución. Para secuencias nucleotídicas, se usa EADNFULL. Se pueden especificar otras matrices. Los parámetros opcionales usados para el alineamiento de secuencias de aminoácidos son una penalización por hueco-abertura de 10 y una penalización por extensión del hueco de 0,5. El experto apreciará que todos estos parámetros diferentes darán resultados ligeramente diferentes pero que el porcentaje de identidad global de dos secuencias no se altera significativamente cuando se usan diferentes algoritmos.

#### Definición de homología global

15 La homología o identidad es el porcentaje de emparejamientos idénticos entre las dos secuencias completas a lo largo de la región alineada total incluyendo cualesquiera huecos o extensiones. La homología o identidad entre las dos secuencias alineadas se calcula como sigue: Número de posiciones correspondientes en el alineamiento que muestran un aminoácido idéntico en ambas secuencias dividido por la longitud total del alineamiento incluyendo los huecos. La identidad definida como en la presente se puede obtener a partir de NEEDLE y se marca en los resultados del programa como "IDENTIDAD".

#### Definición de la identidad más larga

25 La homología o identidad entre las dos secuencias alineadas se calcula como sigue: Número de posiciones correspondientes en el alineamiento que muestran un aminoácido idéntico en ambas secuencias dividido por la longitud total del alineamiento después de la sustracción del número total de huecos en el alineamiento. La identidad definida como en la presente se puede obtener a partir de NEEDLE al usar la opción NOBRIEF y se marca en los resultados del programa como "identidad más larga". Para los propósitos de la invención, el nivel de identidad (homología) entre dos secuencias (aminoácido o nucleótido) se calcula según la definición de "identidad más larga" que se puede llevar a cabo al usar el programa NEEDLE.

30 Las secuencias polipeptídicas de la presente invención se pueden usar además como una "secuencia de consulta" para realizar una búsqueda frente a bases de datos de secuencias, por ejemplo para identificar otros miembros de la familia o secuencias relacionadas. Tales búsquedas se puede realizar usando los programas BLAST. El software para realizar análisis BLAST está disponible públicamente a través de the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov>). BLASTP se usa para secuencias de aminoácidos y BLASTN para secuencias nucleotídicas. El programa BLAST usa como defectos:

Coste para abrir un hueco: defecto = 5 para nucleótidos/11 para polipéptidos

Coste para extender un hueco: defecto = 2 para nucleótidos/1 para polipéptidos

Penalización por desemparejamiento de nucleótidos: defecto = -3

Recompensa por emparejamiento de nucleótidos: defecto = 1

40 Valor esperado: defecto = 10

Tamaño de la palabra: defecto = 11 para nucleótidos/ 28 para megablast/ 3 para polipéptidos

45 Por otra parte, el grado de identidad (homología) local entre la consulta de la secuencia de aminoácidos o la consulta de la secuencia de ácido nucleico y las secuencias homólogas recuperadas se determina mediante el programa BLAST. Sin embargo, solamente se comparan los segmentos de secuencia que dan un emparejamiento por encima de un cierto umbral. Según esto, el programa calcula la identidad solamente para estos segmentos emparejados. Por lo tanto, la identidad calculada de este modo se denomina identidad local.

#### Hibridación

- Según se usa en la presente, el término "hibridar selectivamente", "se hibrida selectivamente" y términos similares están destinados a describir condiciones para la hibridación y el lavado bajo las que secuencias nucleotídicas al menos aproximadamente 50%, al menos aproximadamente 55%, al menos aproximadamente 60%, al menos aproximadamente 65%, al menos aproximadamente 70%, al menos aproximadamente 75%, al menos aproximadamente 80%, más preferiblemente al menos aproximadamente 85%, aún más preferiblemente al menos aproximadamente 90%, preferiblemente al menos 95%, más preferiblemente al menos aproximadamente 98% o más preferiblemente al menos aproximadamente 99% homólogas entre sí típicamente permanecen hibridadas entre sí. Es decir, tales secuencias que se hibridan pueden compartir al menos aproximadamente 50%, al menos aproximadamente 55%, al menos aproximadamente 60%, al menos aproximadamente 65%, al menos aproximadamente 70%, al menos aproximadamente 75%, al menos aproximadamente 80%, más preferiblemente al menos aproximadamente 85%, aún más preferiblemente al menos aproximadamente 90%, más preferiblemente al menos 95%, más preferiblemente al menos 98% o más preferiblemente al menos aproximadamente 99% de identidad de secuencia.
- Un no limitativo preferidos de estas condiciones de hibridación es la hibridación en 6X cloruro sódico/citrato sódico (SSC) a aproximadamente 45°C, seguido por uno o más lavados en 1 X SSC, 0,1% de SDS a aproximadamente 50°C, preferiblemente a aproximadamente 55°C, preferiblemente a aproximadamente 60°C e incluso más preferiblemente a aproximadamente 65°C.
- Condiciones muy rigurosas incluyen, por ejemplo, hibridación a aproximadamente 68°C en 5x SSC/5x solución de Denhardt / 1,0% de SDS y lavado en 0,2x SSC/0,1% de SDS a temperatura ambiente. Alternativamente, el lavado se puede realizar a 42°C.
- El experto en la técnica sabrá qué condiciones aplicar para condiciones de hibridación rigurosas y muy rigurosas. Directrices adicionales relativas a estas condiciones están fácilmente disponibles en la técnica, por ejemplo, en Sambrook y cols., 1989, *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Press, N.Y.; y Ausubel y cols. (eds.), 1995, *Current Protocols in Molecular Biology*, (John Wiley & Sons, N.Y.).
- Por supuesto, un polinucleótido que se hibrida solamente a una secuencia de poli A (tal como el trecho de poli(A) terminal 3' de ARNm), o a un tramo complementario de residuos de T (o U), no se incluiría en un polinucleótido de la invención usado para hibridarse específicamente a una porción de un ácido nucleico de la invención, puesto que este polinucleótido se hibridaría a cualquier molécula de ácido nucleico que contuviera un tramo de poli (A) o el complemento del mismo (p. ej., prácticamente cualquier clon de ADNc de doble hebra).
- En un enfoque típico, se pueden cribar bibliotecas de ADNs construidas a partir de otros organismos, p. ej. un hongo filamentoso, en particular a partir de la familia de microorganismos Trichomaceae, por ejemplo del género *Penicillium*, tal como *Penicillium decumbens*.
- Por ejemplo, las cepas de *Penicillium* se pueden cribar con respecto a polinucleótidos TEMER07589 homólogos mediante análisis de transferencia Northern. Al detectar transcritos homólogos a polinucleótidos según la invención, se pueden construir bibliotecas de ADNc a partir de ARN aislado de la cepa apropiada, utilizando técnicas estándar muy conocidas por los expertos en la técnica. Alternativamente, una biblioteca de ADN genómico total se puede rastrear usando una sonda capaz de hibridarse a un polinucleótido TEMER07589 según la invención.
- Secuencias génicas homólogas se pueden aislar, por ejemplo, al realizar PCR usando dos conjuntos de cebadores oligonucleotídicos degenerados diseñados sobre la base de secuencias nucleotídicas como las mostradas en la presente.
- La plantilla para la reacción puede ser ADNc obtenido mediante transcripción inversa de ARNm preparado a partir de cepas que se sabe o se sospecha que expresan un polinucleótido según la invención. El producto de PCR se puede subclonar y someter a secuenciación para asegurar que las secuencias amplificadas representen las secuencias de una nueva secuencia de ácido nucleico de TEMER07589, o un equivalente funcional de la misma.
- A continuación, el fragmento de PCR se puede usar para aislar un clon de ADNc de longitud completa mediante una variedad de métodos conocidos. Por ejemplo, el fragmento amplificado se puede marcar y usar para cribar una biblioteca de ADNc de bacteriófagos o cósmidos. Alternativamente, el fragmento marcado se puede usar para cribar una biblioteca genómica.
- La tecnología de PCR también se puede usar para aislar secuencias de ADNc de longitud completa de otros organismos. Por ejemplo, el ARN se puede aislar, siguiendo procedimientos estándar, de una fuente celular o tisular apropiada. Una reacción de transcripción inversa se puede realizar sobre el ARN usando un cebador oligonucleotídico específico para el extremo más 5' del fragmento amplificado para el cebado de la síntesis de la primera hebra.

A continuación, al híbrido de ARN/ADN resultante se le puede "añadir una cola" (p. ej., con guaninas) usando una reacción de transferasa terminal estándar, el híbrido se puede digerir con RNasa H, y a continuación la síntesis de la segunda hebra se puede cebar (p. ej., con un cebador de poli-C). Así, las secuencias de ADNc aguas arriba del fragmento amplificado se pueden aislar fácilmente. Para una revisión de estrategias de clonación útiles, véanse, p. ej., Sambrook y cols., anteriormente; y Ausubel y cols., supra.

## Vectores

Otros aspectos de la invención trata de vectores, incluyendo vectores de clonación y expresión, que comprenden un polinucleótido de la invención que codifica un polipéptido TEMER07589 o un equivalente funcional del mismo y métodos de crecimiento, transformación o transfección de estos vectores en una célula hospedadora adecuada, por ejemplo bajo condiciones en las que se produce la expresión de un polipéptido de la invención. Según se usa en la presente, el término "vector" se refiere a una molécula de ácido nucleico capaz de transportar otro ácido nucleico al que se ha enlazado.

Los polinucleótidos de la invención se pueden incorporar en un vector replicable recombinante, por ejemplo un vector de clonación o expresión. El vector se puede usar para replicar el ácido nucleico en una célula hospedadora compatible. Así, en una realización adicional, la invención proporciona un método para elaborar polinucleótidos de la invención al introducir un polinucleótido de la invención en un vector replicable, introducir el vector en una célula hospedadora compatible y hacer crecer la célula hospedadora bajo condiciones que realizan la replicación del vector. El vector se puede recuperar de la célula hospedadora. Células hospedadoras adecuadas se describen posteriormente.

El vector en el que se inserta el casete de expresión o el polinucleótido de la invención puede ser cualquier vector que se pueda someter convenientemente a procedimientos de DNA recombinante, y la elección del vector dependerá a menudo de la célula hospedadora en la que se introduzca.

Un vector según la invención puede ser un vector que se replica autónomamente, es decir un vector que existe como una entidad extracromosómica, cuya replicación es independiente de la replicación cromosómica, p. ej. un plásmido. Alternativamente, el vector puede ser una que, cuando se introduzca en una célula hospedadora, se integre en el genoma de la célula hospedadora y se replique junto con el cromosoma o los cromosomas en los que se ha integrado.

Un tipo de vector es un "plásmido", que se refiere a un bucle de ADN circula de doble hebra en el que se pueden ligar segmentos de ADN adicionales. Otro tipo de vector es un vector viral, en el que segmentos de ADN adicionales se pueden ligar en el genoma viral. Ciertos vectores son capaces de replicación autónoma en una célula hospedadora en la que se introducen (p. ej., vectores bacterianos que tienen origen de replicación bacteriano y vectores mamíferos episómicos). Otros vectores (p. ej., vectores mamíferos no episómicos) se integran en el genoma de una célula hospedadora al introducirse en la célula hospedadora, y de ese modo se replican junto con el genoma del hospedador. Por otra parte, ciertos vectores son capaces de dirigir la expresión de genes a los que están ligados operativamente. Tales vectores se denominan en la presente "vectores de expresión". En general, los vectores de expresión de utilidad en técnicas de ADN recombinante están a menudo en la forma de plásmidos. Los términos "plásmido" y "vector" se pueden usar intercambiamente en la presente ya que el plásmido es la forma de vector más comúnmente usada. Sin embargo, la invención está destinada a incluir estas otras formas de vectores de expresión, tales como un cósmido, vectores virales (p. ej., retrovirus, adenovirus y virus adenoasociados defectuosos en la replicación) y vectores fágicos que realizan funciones equivalentes.

Para la mayoría de los hongos filamentosos y las levaduras, el vector o la construcción de expresión está integrado preferiblemente en el genoma de la célula hospedadora a fin de obtener transformantes estables. Sin embargo, para ciertas levaduras, también están disponibles vectores episómicos adecuados en los que se puede incorporar la construcción de expresión para una expresión estable y de alto nivel, ejemplos de los mismos incluyen vectores derivados de los plásmidos 2 $\mu$  y pKD1 de *Saccharomyces* y *Kluyveromyces*, respectivamente, o vectores que contienen una secuencia AMA (p. ej. AMA1 de *Aspergillus*). En caso de que las construcciones de expresión estén integradas en el genoma de las células hospedadoras, las construcciones se integran bien en loci aleatorios en el genoma o bien en loci elegidos predeterminados usando recombinación homóloga, en cuyo caso los loci elegidos comprenden preferiblemente un gen altamente expresado.

Según esto, vectores de expresión útiles en la presente invención incluyen vectores cromosómicos, episómicos y derivados de virus, p. ej., vectores derivados de plásmidos bacterianos, un bacteriófago, un episoma de levadura, elementos cromosómicos de levadura, virus tales como baculovirus, papovavirus, virus variolovacunales, adenovirus, virus de viruela aviar, virus de seudorrabias y retrovirus, y vectores derivados de combinaciones de los mismos, tales como los derivados de elementos genéticos de un plásmido y un bacteriófago, tales como cósmidos y fagémidos.

Los vectores según la invención se pueden usar in vitro, por ejemplo para la producción de ARN o se pueden usar para transfectar o transformar una célula hospedadora. El vector de expresión recombinante se puede transcribir y traducir in vitro, por ejemplo usando secuencias reguladoras del promotor T7 y T7 polimerasa.

- 5 Un vector de la invención puede comprender dos o más, por ejemplo tres, cuatro o cinco, polinucleótidos de la invención, por ejemplo para la sobreexpresión.

10 Los vectores de expresión recombinantes de la invención comprenden un ácido nucleico de la invención en una forma adecuada para la expresión del ácido nucleico en una célula hospedadora, lo que significa que el vector de expresión recombinante incluye una o más secuencias reguladoras, seleccionadas sobre la base de las células hospedadoras que se van a usar para la expresión, que está ligado operativamente a la secuencia de ácido nucleico que se va a expresar.

15 Dentro de un vector, tal como un vector de expresión, "ligado operativamente" está destinado a significa que la secuencia nucleotídica de interés está ligada a la secuencia o secuencias reguladoras de un modo que permita la expresión de la secuencia nucleotídica (p. ej., en un sistema de transcripción/traducción in vitro o en una célula hospedadora cuando el vector se introduce en la célula hospedadora), es decir, el término "ligado operativamente" se refiere a una yuxtaposición en las que los componentes descritos en la presente están en una relación que los permite funcional de su modo pretendido. Una secuencia reguladora tal como un promotor, un potenciador u otra  
20 señal de regulación de la expresión "ligada operativamente" a una secuencia codificante está situada de tal modo que la expresión de la secuencia codificante se alcance bajo una condición compatible con las secuencias de control o las secuencias estén dispuestas de modo que funcionen de forma concertada para su propósito pretendido, por ejemplo la transcripción se inicia en un promotor y avanza a través de la secuencia de ADN que codifica el polipéptido.

25 Un vector o construcción de expresión para una célula hospedadora dada puede comprender así los siguientes elementos ligados operativamente entre sí en un orden consecutivo desde el extremo 5' hasta el extremo 3' con relación a la hebra codificante de la secuencia que codifica el polipéptido de la primar invención: (1) una secuencia promotora capaz de dirigir la transcripción de la secuencia nucleotídica que codifica el polipéptido en la célula hospedadora dada; (2) opcionalmente, una secuencia de señal capaz de dirigir la secreción del polipéptido desde la célula hospedadora dada a un medio de cultivo; (3) una secuencia de ADN de la invención que codifica una forma madura y preferiblemente activa de un polipéptido que tiene actividad potenciadora de celulasa; y preferiblemente también (4) una región de terminación de la transcripción (terminador) capaz de terminar la transcripción aguas  
30 abajo de la secuencia nucleotídica que codifica el polipéptido.

35 Aguas abajo de la secuencia nucleotídica según la invención puede haber una región no traducida 3' que contiene uno o más sitios de terminación de la transcripción (p. ej. un terminador). El origen del terminador es menos crítico. El terminador, por ejemplo, puede ser natural a la secuencia de ADN que codifica el polipéptido. Sin embargo, preferiblemente, se usa un terminador de levadura en células hospedadoras de levadura y se usa un terminador de hongo filamentoso en células hospedadoras de hongos filamentosos. Más preferiblemente, el terminador es endógeno a la célula hospedadora (en la que se va a expresar la secuencia nucleotídica que codifica el polipéptido). En la región transcrita, puede estar presente un sitio de unión al ribosoma para la traducción. La porción codificante de los transcritos maduros expresados por las construcciones incluirá un AUG de iniciación de la traducción al principio y un codón de terminación situado apropiadamente al final del polipéptido que se va a traducir.

40 También se puede alcanzar una expresión potenciada del polinucleótido de la invención mediante la selección de regiones reguladoras heterólogas, p. ej. regiones promotora, líder de secreción y/o terminadora, que pueden servir para incrementar la expresión y, si se desea, los niveles de secreción del polipéptido de interés desde el hospedador de expresión y/o proporcionar el control inducible de la expresión de un polipéptido de la invención.

45 Será apreciado por los expertos en la técnica que el diseño del vector de expresión puede depender de factores tales como la elección de la célula hospedadora que se va a transformar, el nivel de expresión de polipéptido deseado, etc. Los vectores, tales como vectores de expresión, de la invención se pueden introducir en células hospedadoras para producir de ese modo proteínas o péptidos, codificados por ácidos nucleicos según se describe en la presente (p. ej. polipéptidos TEMER07589, formas mutantes de polipéptidos TEMER07589, fragmentos,  
50 variantes o equivalentes funcionales de los mismos. Los vectores, tales como vectores de expresión recombinantes, de la invención se pueden diseñar para la expresión de polipéptidos TEMER07589 en células procarióticas o eucarióticas.

60 Por ejemplo, polipéptidos TEMER07589 se pueden expresar en células bacterianas tales como E. coli, células de insecto (usando vectores de expresión baculovirales), hongos filamentosos, células de levadura o células de mamífero. Células hospedadoras adecuadas se analizan adicionalmente en Goeddel, Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, CA (1990). Ejemplos representativos de hospedadores apropiados se describen posteriormente en la presente.

65

Medios de cultivo y condiciones apropiados para las células hospedadoras descritas anteriormente se conocen en la técnica.

5 El término "secuencias de control" o "secuencias reguladoras" se define en la presente para incluir al menos un componente que pueda ser necesario y/o ventajoso para la expresión de un polipéptido. Cualquier secuencia de control puede ser natural o extraña a la secuencia de ácido nucleico de la invención que codifica un polipéptido. Estas secuencias de control pueden incluir, pero no se limitan a, un promotor, un líder, secuencias de iniciación de la traducción óptimas (como las descritas en Kozak, 1991, J. Biol. Chem. 266:19867-19870), una secuencia señalizadora de secreción, una secuencia propeptídica, una secuencia de poliadenilación, un terminador de la transcripción. Como mínimo, las secuencias de control incluyen típicamente un promotor, y señales de terminación de la transcripción y la traducción. Según se indica anteriormente, el término "ligado operativamente" se define en la presente como una configuración en la que una secuencia de control está situada apropiadamente en una posición con relación a la secuencia codificante de la secuencia de ADN de modo que la secuencia de control dirija la producción de un polipéptido.

15 Las secuencias de control pueden estar provistas de conectores con el propósito de introducir sitios de restricción específicos que facilitan la ligación de las secuencias de control con la región codificante de la secuencia de ácido nucleico que codifica un polipéptido. El término "ligado operativamente" se define en la presente como una configuración en la que una secuencia de control está situada apropiadamente en una posición con relación a la secuencia codificante de la secuencia de ADN de modo que la secuencia de control dirija la producción de un polipéptido.

20 La secuencia de control puede ser cualquier secuencia promotora apropiada, una secuencia de ácido nucleico, que sea reconocida por una célula hospedadora para la expresión de la secuencia de ácido nucleico. La secuencia promotora contiene secuencias de control de la transcripción, que median en la expresión del polipéptido. El promotor puede ser cualquier secuencia de ácido nucleico, que muestra actividad de transcripción en la célula, incluyendo promotores mutantes, truncados e híbridos, y se pueden obtener de genes que codifican polipéptidos extracelulares o intracelulares bien homólogos o bien heterólogos para la célula.

25 La secuencia de control también puede ser una secuencia terminadora de la transcripción adecuada, una secuencia reconocida por una célula de hongo filamentoso para terminar la transcripción. La secuencia terminadora está ligada operativamente al término 3' de la secuencia de ácido nucleico que codifica el polipéptido. Cualquier terminador, que sea funcional en la célula, se puede usar en la presente invención.

30 La secuencia de control también puede ser un terminador. Terminadores preferidos para células de hongos filamentosos se obtienen a partir de los genes que codifican TAKA amilasa de *A. oryzae*, glucoamilasa (glaA) de *A. niger*, antranilato sintasa de *A. nidulans*, alfa-glucosidasa de *A. niger*, gen *trpC* y proteasa tripsinoide de *Fusarium oxysporum*.

35 La secuencia de control también puede ser una secuencia líder adecuada, una región no traducida de un ARNm que es importante para la traducción de la célula de hongo filamentoso. La secuencia líder está conectada operativamente al término 5' de la secuencia de ácido nucleico que codifica el polipéptido. Cualquier secuencia líder, que sea funcional en la célula, se puede usar en la presente invención. Líderes preferidos para células de hongos filamentosos se obtienen de los genes que codifican TAKA amilasa de *A. oryzae* y triosa fosfato isomerasa de *A. nidulans* y glaA de *A. niger*.

Otras secuencias de control se pueden aislar del gen IPNS de *Penicillium*, o gen *pcbC*, el gen de beta tubulina. Secuencias de control se citan en el documento WO 01/21779.

40 La secuencia de control también puede ser una secuencia de poliadenilación, una secuencia que está conectada operativamente al término 3' de la secuencia de ácido nucleico y que, cuando se transcribe, es reconocida por la célula de hongo filamentoso como una señal para añadir residuos de poliadenosina a ARNm transcrito. Cualquier secuencia de poliadenilación, que sea funcional en la célula, se puede usar en la presente invención. Secuencias de poliadenilación preferidas para células de hongos filamentosos se obtienen de los genes que codifican TAKA amilasa de *A. oryzae*, glucoamilasa de *A. niger*, antranilato sintasa de *A. nidulans*, proteasa tripsinoide de *Fusarium* y alfa-glucosidasa de *A. niger*.

45 Cuando el polipéptido según la invención se va a secretar desde la célula hospedadora al medio de cultivo, una secuencia de señal apropiada se puede añadir al polipéptido a fin de dirigir el polipéptido sintetizado de novo a la ruta de secreción de la célula hospedadora. El experto en la técnica conoce la selección de una secuencia de señal apropiada para un hospedador específico. La secuencia de señal puede ser natural a la célula hospedadora, o puede ser extraña a la célula hospedadora. Como un ejemplo, se puede usar una secuencia de señal procedente de un polipéptido natural para la célula hospedadora. Preferiblemente, dicho polipéptido natural es un polipéptido altamente secretado, es decir un polipéptido que se secreta en cantidades superiores al 10% de la cantidad total de

polipéptido que se secreta. Las secuencias de señal usadas preferiblemente según la invención son, por ejemplo: pmeA.

5 Como una alternativa para una secuencia de señal, el polipéptido de la invención se puede fusionar a un polipéptido portador secretado, o parte del mismos. Tal construcción quimérica se dirige a la ruta de secreción por medio de la secuencia de señal del polipéptido portador, o parte del mismo. Además, el polipéptido portador proporcionará un efecto estabilizante al polipéptido según la invención y/o puede potenciar la solubilidad. Tal polipéptido portador puede ser cualquier polipéptido. Preferiblemente, un polipéptido altamente secretado se usa como un polipéptido portador. El polipéptido portador puede ser natural o extraño al polipéptido según la invención. El polipéptido portador puede ser natural o puede ser extraño a la célula hospedadora. Ejemplos de tales polipéptidos portadores son glucoamilasa, preprosecuencia de factor de emparejamiento alfa, dominio de unión a celulosa de proteína A que se une a celulosa de *Clostridium cellulovorans*, glutationa S-transferasa, dominio de unión a quitina de quitinasa A1 de *Bacillus circulans*, dominio de unión a maltosa codificado por el gen malE de *E. coli* K12, beta-galactosidasa y fosfatasa alcalina. Un polipéptido portador preferido para la expresión de esta construcción quimérica en células de *Aspergillus* es la glucoamilasa. La proteína y el polipéptido portadores pueden contener un motivo de aminoácido específico para facilitar el aislamiento del polipéptido; el polipéptido según la invención se puede liberar mediante un agente de liberación especial. El agente de liberación puede ser una enzima proteolítica o un agente químico. Un ejemplo de este motivo de aminoácido es el sitio de escisión de proteasa KEX, que es muy conocido por los expertos en la técnica.

20 Una secuencia de señal se puede usar para facilitar la secreción y el aislamiento de una proteína o un polipéptido de la invención. Las secuencias de señal se caracterizan típicamente por un núcleo de aminoácidos hidrófobos, que are generalmente se escinden de la proteína madura durante la secreción en uno o más episodios de escisión. Estos péptidos de señal contienen sitios de procesamiento que permiten la escisión de la secuencia de señal de las proteínas maduras a medida que pasan a través de la ruta secretora. La secuencia de señal dirige la secreción de la proteína, tal como desde un hospedador eucariótico en el que se transforma el vector de expresión, y la secuencia de señal se escinde posteriormente o simultáneamente. A continuación, la proteína se puede purificar fácilmente del medio extracelular mediante métodos conocidos. Alternativamente, la secuencia de señal se puede conectar a la proteína de interés usando una secuencia, lo que facilita la purificación, tal como con un dominio GST. Así, a modo de ejemplo, la secuencia que codifica el polipéptido se puede fusionar a una secuencia marcadora, tal como una secuencia que codifica un péptido, lo que facilita la purificación del polipéptido fusionado. En ciertas realizaciones preferidas de este aspecto de la invención, la secuencia marcadora es un péptido de hexahistidina, tal como la etiqueta tag proporcionada en un vector pQE (Qiagen, Inc.), entre otras, muchas de las cuales están disponibles comercialmente. Según se describe en Gentz y cols, Proc. Natl. Acad. Sci. USA 86:821-824 (1989), a modo de ejemplo, la hexahistidina proporciona una purificación conveniente de la proteína de fusión. La etiqueta HA es otro péptido útil para la purificación que corresponde a un epítipo derivado de proteína de hemaglutinina de la gripe, que se ha descrito por Wilson y cols., Cell 37:767 (1984), a modo de ejemplo.

40 Preferiblemente, una proteína de fusión de TEMER07589 de la invención se produce mediante técnicas de DNA recombinante estándar. Por ejemplo, fragmentos de ADN que codifican las diferentes secuencias polipeptídicas se ligan entre sí en el marco según técnicas convencionales, por ejemplo al emplear términos de extremos romos o extremos escalonados para la ligación, digestión con enzimas de restricción para proporcionar términos apropiados, relleno de extremos cohesivos según sea apropiado, tratamiento con fosfatasa alcalina para evitar un ensamblaje no deseable, y ligación enzimática. En otra realización, el gen de fusión se puede sintetizar mediante técnicas convencionales incluyendo sintetizadores de ADN automáticos. Alternativamente, la amplificación por PCR de fragmentos génicos se puede llevar a cabo usando cebadores de anclaje, que dan lugar a protuberancias complementarias entre dos fragmentos génicos consecutivos que posteriormente se puedan renaturalizar y reamplificar para generar una secuencia génica quimérica (véase, por ejemplo, Current Protocols in Molecular Biology, eds. Ausubel y cols. John Wiley & Sons: 1992). Por otra parte, están disponibles muchos vectores de expresión que ya codifican un resto de fusión (p. ej., un polipéptido GST). Un ácido nucleico que codifica TEMER07589 se puede clonar en este vector de expresión de modo que el resto de fusión esté conectado en el marco a la proteína de TEMER07589.

55 Preferiblemente, la eficacia de la integración dirigida en el genoma de la célula hospedadora, es decir la integración en un locus elegido predeterminado, es incrementada por las capacidades de recombinación homóloga aumentadas de la célula hospedadora. Tal fenotipo de la célula implica preferiblemente un gen hdfA o hdfB deficiente según se describe en el documento WO2005/095624. El documento WO2005/095624 divulga un método preferido para obtener una célula de hongo filamentoso que comprende una eficacia incrementada de integración dirigida.

60 Opcionalmente, la célula hospedadora comprende una respuesta de proteína desplegada (UPR) elevada en comparación con la célula silvestre para potenciar las capacidades de producción de un polipéptido de interés. La UPR se puede incrementar mediante las técnicas descritas en los documentos US2004/0186070A1 y/o US2001/0034045A1 y/o WO01/72783A2 y/o WO2005/123763. Más específicamente, el nivel proteínico de HAC1 y/o IRE1 y/o PTC2 se ha modulado, y/o la proteína SEC61 se ha manipulado para obtener una célula hospedadora que tiene una UPR elevada.

Alternativamente, o en combinación con una UPR elevada, la célula hospedadora se modifica genéticamente para obtener un fenotipo que exhibe menor expresión de proteasa y/o secreción de proteasa en comparación con la célula silvestre a fin de potenciar las capacidades de producción de un polipéptido de interés. Este fenotipo se puede obtener mediante eliminación y/o modificación y/o inactivación de un regulador transcripcional de la expresión de proteasas. Este regulador transcripcional es, p. ej., prtT. La reducción de la expresión de proteasas mediante la modulación de prtT se puede realizar mediante las técnicas descritas en el documento US2004/0191864A1.

Alternativamente, o en combinación con una UPR elevada y/o un fenotipo que exhibe menor expresión de proteasa y/o secreción de proteasa, la célula hospedadora exhibe un fenotipo deficiente en oxalato a fin de potenciar el rendimiento de producción de un polipéptido de interés. Un fenotipo deficiente en oxalato se puede obtener mediante las técnicas descritas en el documento WO2004/070022A2.

Alternativamente, o en combinación con una UPR elevada y/o un fenotipo que exhibe menor expresión de proteasa y/o secreción de proteasa y/o deficiencia de oxalato, la célula hospedadora exhibe una combinación de diferencias fenotípicas en comparación con la célula silvestre para potenciar el rendimiento de producción del polipéptido de interés. Estas diferencias pueden incluir, pero no se limitan a, expresión disminuida de glucoamilasa y/o alfa-amilasa neutra A y/o alfa-amilasa neutra B, proteasa y ácido oxálico hidrolasa. Dichas diferencias fenotípicas exhibidas por la célula hospedadora se pueden obtener mediante modificación genética según las técnicas descritas en el documento US2004/0191864A1.

Alternativamente, o en combinación con una UPR elevada y/o un fenotipo que exhibe menor expresión de proteasa y/o secreción de proteasa y/o deficiencia de oxalato y una combinación de diferencias fenotípicas en comparación con la célula silvestre para potenciar el rendimiento de producción del polipéptido de interés, la célula hospedadora exhibe una deficiencia en genes de toxina, inhabilitando la capacidad de la célula hospedadora de hongo filamentoso para expresar toxinas. Estas toxinas incluyen, pero no se limitan a, ocratoxinas, fumonisinas, ácido ciclapiazónico, ácido 3-nitropropiónico, emodina, malformina, aflatoxinas y ácidos secalónicos. Esta deficiencia es preferiblemente tal como se describe en el documento WO2000/039322.

(Sobre)expresión

En una realización preferida, los polinucleótidos de la presente invención que se describen en la presente se pueden sobreexpresar en una cepa microbiana de la invención en comparación con la cepa microbiana original en la que dicho gen no se sobreexpresa. La sobreexpresión de una secuencia polinucleotídica se define en la presente como la expresión de dicho gen de secuencia que da como resultado una actividad de la enzima codificada por dicha secuencia en una cepa microbiana que es al menos aproximadamente 1,5 veces la actividad de la enzima en la cepa microbiana original; preferiblemente la actividad de dicha enzima es al menos aproximadamente 2 veces, más preferiblemente al menos aproximadamente 3 veces, más preferiblemente al menos aproximadamente 4 veces, más preferiblemente al menos aproximadamente 5 veces, aún más preferiblemente al menos aproximadamente 10 veces y lo más preferiblemente al menos aproximadamente 20 veces la actividad de la enzima en la cepa microbiana original.

El vector puede incluir además secuencias que flanquean el polinucleótido dando lugar a ARN que comprenden secuencias homólogas a secuencias genómicas eucarióticas o secuencias genómicas virales. Esto permite la introducción de los polinucleótidos de la invención en el genoma de una célula hospedadora.

Un vector de clonación integrador se puede integrar aleatoriamente o en un locus elegido predeterminado en el cromosoma o los cromosomas de la célula hospedadora en la que se va a integrar. En una realización preferida de la invención, un vector de clonación integrador puede comprender un fragmento de ADN que es homólogo a una secuencia de ADN en un locus elegido predeterminado en el genoma de la célula hospedadora para dirigir la integración del vector de clonación en este locus predeterminado. A fin de promover la integración dirigida, el vector de clonación preferiblemente se puede linealizar antes de la transformación de la célula hospedadora. La linealización se puede realizar preferiblemente de modo que al menos uno pero preferiblemente cualquier extremo del vector de clonación esté flanqueado por secuencias homólogas al locus elegido. La longitud de las secuencias homólogas que flanquean el locus elegido es preferiblemente al menos aproximadamente 0,1 kb, tal como aproximadamente al menos 0,2 kb, más preferiblemente al menos aproximadamente 0,5 kb, aún más preferiblemente al menos aproximadamente 1 kb, lo más preferiblemente al menos aproximadamente 2 kb. Preferiblemente, las cepas hospedadoras originales se pueden modificar para una frecuencia mejorada de la integración de ADN elegido, según se describe en los documentos WO05/095624 y/o WO2007/115886.

El ejemplo de eliminación proporcionado en la presente invención usa el promotor del gen como flanco 5' y el gen como el flanco 3' para insertar un marcador de selección entre el promotor y el gen, alterando (es decir inactivando funcionalmente) de ese modo la transcripción génica. Las secuencias génicas dadas anteriormente se pueden usar para elaborar genes funcionalmente inactivados similares. Los genes se pueden dividir en dos, dando un flanco 5' y

un flanco 3', pero el gen también se puede usar para clonar un trozo más grande de ADN genómico que contiene las regiones promotora y terminadora del gen, que pueden funcionar como flanco 5' y flanco 3'.

5 El sistema vectorial puede ser un solo vector, tal como un solo plásmido, o dos o más vectores, tales como dos o más plásmidos, que juntos contiene el ADN total que se va a introducir en el genoma de la célula hospedadora.

El vector puede contener un polinucleótido de la invención orientado en una dirección antisentido para proporcionar la producción de ARN antisentido.

10 El ADN del vector se puede introducir en células procarióticas o eucarióticas a través de técnicas de transformación o transfección convencionales. Según se usa en la presente, los términos "transformación" y "transfección" pretenden referirse a una variedad de técnicas reconocidas en la especialidad para introducir un ácido nucleico (p. ej., ADN) extraño en una célula hospedadora, incluyendo coprecipitación con fosfato cálcico o cloruro cálcico, transfección mediada por DEAE-dextrano, transducción, infección, lipofección, transfección mediada por lípidos catiónicos o electroporación. Métodos adecuados para transformar o transfectar células hospedadoras se pueden encontrar en Sambrook y cols. (Molecular Cloning: A Laboratory Manual, 2ª ed. Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989), Davis y cols., Basic Methods in Molecular Biology (1986) y otros manuales de laboratorio.

20 Para la transfección estable de células de mamífero, se sabe que, dependiendo del vector de expresión y la técnica de transfección usados, solo una pequeña fracción de células puede integrar el ADN extraño en su genoma. A fin de identificar y seleccionar estos integrantes, un gen que codifica un marcador seleccionable (p. ej., resistencia a antibióticos) se introduce generalmente en las células hospedadoras junto con el gen de interés. Marcadores seleccionables preferidos incluyen, pero no se limitan a, los que confieren resistencia a fármacos o que complementan un defecto en la célula hospedadora. Incluyen, p. ej., genes marcadores versátiles que se pueden usar para la transformación de la mayoría de los hongos filamentosos y las levaduras tales como genes de acetamidasa o ADNc (los genes amdS, niaD, facA o ADNcs de *A. nidulans*, *A. oryzae* o *A. niger*), o genes que proporcionan resistencia a antibióticos como resistencia a G418, higromicina, bleomicina, kanamicina, metotrexato, fleomicina o benomilo (benA). Alternativamente, se pueden usar marcadores de selección específicos tales como marcadores auxotróficos que requieren cepas hospedadoras mutantes correspondientes: p. ej. URA3 (procedente de *S. cerevisiae* o genes análogos procedentes de otras levaduras), pyrG o pyrA (procedentes de *A. nidulans* o *A. niger*), argB (procedente de *A. nidulans* o *A. niger*) o trpC. En una realización preferida, el marcador de selección se elimina de la célula hospedadora transformada después de la introducción de la construcción de expresión a fin de obtener células hospedadoras transformadas capaces de producir el polipéptido que estén libres de genes marcadores de selección.

35 Otros marcadores incluyen ATP sintetasa, subunidad 9 (oliC), orotidina-5'-fosfatodescarboxilasa (pvrA), el gen de resistencia a G418 bacteriano (este también se pueden usar en levaduras, pero no en hongos), el gen de resistencia a ampicilina (*E. coli*), el gen de resistencia a neomicina (*Bacillus*), el gen de resistencia a nourseotricina nat1 procedente de *Streptomyces nursei*, el gen de resistencia a piritiamina ptrA de *Aspergillus oryzae* y el gen uidA de *E. coli*, que codifica  $\beta$ -glucuronidasa (GUS). Los vectores de pueden usar in vitro, por ejemplo para la producción de ARN o se pueden usar para transfectar o transformar una célula hospedadora.

45 La expresión de proteínas en procariotas se lleva a cabo a menudo en *E. coli* con vectores que contienen promotores constitutivos o inducibles que dirigen la expresión de proteínas bien de fusión o bien no de fusión. Los vectores de fusión añaden un número de aminoácidos a una proteína codificada en los mismos, p. ej. al término amino de la proteína recombinante. Tales vectores de fusión consiguen típicamente tres propósitos: 1) incrementar la expresión de proteína recombinante; 2) incrementar la solubilidad de la proteína recombinante; y 3) ayudar en la purificación de la proteína recombinante al actuar como un ligando en la purificación por afinidad. A menudo, en vectores de expresión de fusión, se introduce un sitio de escisión proteolítica en la unión de la proteína de fusión y la proteína recombinante para permitir la separación de la proteína recombinante del resto de fusión después de la purificación de la proteína de fusión.

50 Según se indica, los vectores de expresión preferiblemente contendrán marcadores seleccionables. Estos marcadores incluyen dihidrofolato reductasa o resistencia a neomicina para un cultivo de células eucarióticas y resistencia a tetraciclina o ampicilina para cultivar en *E. coli* y otras bacterias.

55 Vectores preferidos para el uso en bacterias se divulgan, por ejemplo, en el documento WO-A1-2004/074468. Otros vectores adecuados serán fácilmente evidentes para el experto.

60 Para la secreción de la proteína traducida en la luz del retículo endoplásmico, en el espacio periplásmico o en el ambiente extracelular, una señal de secreción apropiada se puede incorporar en el polipéptido expresado. Las señales pueden ser endógenas al polipéptido o pueden ser señales heterólogas.

El polipéptido TEMER07589 se puede expresar en una forma modificada, tal como una proteína de fusión, y puede incluir no solo señales de secreción sino también regiones funcionales heterólogas adicionales. Así, a modo de ejemplo, una región de aminoácidos, particularmente aminoácidos cargados, adicionales se puede añadir al término N del polipéptido para mejorar la estabilidad y la persistencia en la célula hospedadora, durante la purificación o durante el manejo y el almacenamiento posteriores. Además, se pueden añadir restos peptídicos al polipéptido para facilitar la purificación

La invención proporciona un polipéptido aislado que tiene la secuencia de aminoácidos según SEQ ID N°: 2 y una secuencia de aminoácidos obtenible al expresar el polinucleótido de SEQ ID N°: 1 en un hospedador apropiado. Además, un péptido o polipéptido que comprende una variante de los polipéptidos anteriores, tal como un equivalente funcional, está comprendido dentro de la presente invención. Los polipéptidos anteriores están comprendidos colectivamente en el término "polipéptidos según la invención"

El término "péptido variante" o "polipéptido variante" se define en la presente como un péptido o polipéptido, respectivamente, que comprende una o más alteraciones, tales como sustituciones, inserciones, eliminaciones y/o truncamientos de uno o más residuos de aminoácido específicos en una o más posiciones en el péptido o polipéptido, respectivamente. Según esto, un péptido de señal variante es un péptido de señal que comprende una o más alteraciones, tales como sustituciones, inserciones, eliminaciones y/o truncamientos de uno o más residuos de aminoácido específicos en el péptido de señal.

El término "polinucleótido" es idéntico al término "molécula de ácido nucleico" y se puede usar en la presente intercambiamente. El término se refiere a una molécula de polinucleótido, que es una molécula de ácido ribonucleico (ARN) o ácido desoxirribonucleico (ADN), bien de una sola hebra o de doble hebra. Un polinucleótido bien puede estar presente en forma aislada o bien estar comprendido en moléculas o vectores de ácido nucleico recombinantes, o bien estar comprendido en una célula hospedadora.

El término "polinucleótido variante" se define en la presente como un polinucleótido que comprende una o más alteraciones, tales como sustituciones, inserciones, eliminaciones y/o truncamientos de uno o más nucleótidos en una o más posiciones específicas en el polinucleótido.

Los términos "péptido" y "oligopéptido" se consideran sinónimos (según se reconoce comúnmente) y cada uno de los términos se puede usar intercambiamente, según requiera el contexto, para indicar una cadena de al menos dos aminoácidos acoplados por enlaces peptídicos. La palabra "polipéptido" se usa en la presente para cadenas que contienen más de siete residuos de aminoácido. Todas las fórmulas o secuencias de oligopéptidos y polipéptidos de la presente se escriben de izquierda a derecha y en la dirección del término aminico al término carboxílico. El código de aminoácidos de una letra usado en la presente se conoce comúnmente en la técnica y se puede encontrar en Sambrook, y cols. (Molecular Cloning: A Laboratory Manual, 2ª ed. Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989)

Por polipéptido o proteína "aislados" se entiende un polipéptido o una proteína retirados de su ambiente natural. Por ejemplo, los polipéptidos y las proteínas producidos recombinantemente expresados en células hospedadoras se consideran aislados para el propósito de la invención ya que son polipéptidos naturales o recombinantes que se han purificado sustancialmente mediante cualquier técnica adecuada tal como, por ejemplo, el método de purificación en una sola etapa divulgado en Smith y Johnson, Gene 67:31-40 (1988).

El polipéptido potenciador de TEMER07589 celulasa según la invención se puede recuperar y purificar de cultivos celulares recombinantes mediante métodos conocidos en la técnica. Lo más preferiblemente, se emplea para la purificación cromatografía de líquidos de alto rendimiento ("HPLC").

Los polipéptidos de la presente invención incluyen productos purificados naturalmente, productos de procedimientos químicos sintéticos y productos producidos mediante técnicas recombinantes a partir de un hospedador procariótico o eucariótico, incluyendo, por ejemplo, células de bacterias, levaduras, plantas superiores, insectos y mamíferos. Dependiendo del hospedador empleado en un procedimiento de producción recombinante, los polipéptidos de la presente invención pueden ser glicosilados o pueden ser no glicosilados. Además, los polipéptidos de la invención también pueden incluir un residuo de metionina modificado inicial, en algunos casos como resultado de procesos mediados por el hospedador.

La invención también ofrece fragmentos biológicamente activos de los polipéptidos según la invención.

Fragmentos biológicamente activos de un polipéptido de la invención incluyen polipéptidos que comprenden secuencias de aminoácidos suficientemente idénticas a o derivadas de la secuencia de aminoácidos del polipéptido TEMER07589 (p. ej., la secuencia de aminoácidos de SEQ ID N°: 2), que incluyen menos aminoácidos que el polipéptido de longitud completa pero que exhiben al menos una actividad biológica del correspondiente polipéptido de longitud completa. Típicamente, los fragmentos biológicamente activos comprenden un dominio o motivo con al menos una actividad del polipéptido TEMER07589.

5 Un fragmento biológicamente activo de un polipéptido de la invención puede ser un polipéptido que tiene, por ejemplo, aproximadamente 10, aproximadamente 25, aproximadamente 50, aproximadamente 100 o más aminoácidos de longitud o al menos aproximadamente 100 aminoácidos, al menos 150, 200, 250, 300, 350, 400 aminoácidos de longitud, o de una longitud hasta el número total de aminoácidos del polipéptido de la invención.

10 Por otra parte, otras porciones biológicamente activas, en las que se eliminan otras regiones del polipéptido, se pueden preparar mediante técnicas recombinantes y evaluar con respecto a una o más de las actividades biológicas de la forma natural de un polipéptido de la invención.

10 La invención también presente fragmentos de ácido nucleico que codifican los fragmentos biológicamente activos anteriores del polipéptido TEMER07589.

#### Polipéptidos

15 En otro aspecto de la invención, se proporcionan polipéptidos TEMER07589 mejorados. Los polipéptidos TEMER07589 mejorados son polipéptidos en los que al menos una actividad biológica se mejora. Estos polipéptidos se pueden obtener al introducir aleatoriamente mutaciones a lo largo de la totalidad o parte de la secuencia codificante de TEMER07589, tal como mediante mutagénesis de saturación, y los mutantes resultantes se pueden expresar recombinantemente y cribar con respecto a una actividad biológica.

20 Variantes mejoradas de las secuencias de aminoácidos de la presente invención que conducen a una función de potenciación de celulasa mejorada se pueden obtener mediante los correspondientes genes de la presente invención. Entre estas modificaciones se incluyen:

1. PCR tendente a erros para introducir mutaciones aleatorias, seguido por un cribado de variantes obtenidas y aislamiento de variantes con propiedades cinética mejoradas

25 2. Barajado familiar de variantes relacionadas de los genes que codifican la enzima potenciadora de celulasa, seguido por un cribado de las variantes obtenidas y aislamiento de variantes con propiedades cinéticas mejoradas

Variantes de los genes de la presente invención que conducen a un nivel incrementado de ARNm y/o polipéptido, que dan como resultado más actividad potenciadora de celulasa, se pueden obtener mediante las secuencias polinucleotídicas de dichos genes. Entre estas modificaciones, se incluyen:

30 1. Mejorar la utilización de codones de tal modo que los codones se adapten (óptimamente) al hospedador microbiano original.

2. Mejorar la utilización de pares de codones de tal modo que los codones se adapten (óptimamente) al hospedador microbiano original

35 3. Adición de secuencias estabilizantes a la información genómica que codifica el polipéptido potenciador de celulasa dando como resultado moléculas de ARNm con una semivida incrementada

Métodos preferidos para aislar variantes con propiedades catalíticas mejoradas o niveles incrementados de ARNm o polipéptido se describen en los documentos WO03/010183 y WO03/01311. Métodos preferidos para optimizar la utilización de codones en cepas microbianas originales se describen en el documento PCT/EP2007/05594. Métodos preferidos para la adición de elementos estabilizantes a los genes que codifican el polipéptido potenciador de celulasa de la invención se describen en el documento WO2005/059149.

40 En una realización preferida, el polipéptido TEMER07589 tiene una secuencia de aminoácidos según SEQ ID N°: 2. En otra realización, el polipéptido TEMER07589 es sustancialmente homólogo a la secuencia de aminoácidos según SEQ ID N°: 2 y retiene al menos una actividad biológica de un polipéptido según SEQ ID N°: 2, y sin embargo difiere en la secuencia de aminoácidos debido a variación natural o mutagénesis según se describe.

50 En una realización preferida adicional, el polipéptido TEMER07589 tiene una secuencia de aminoácidos codificada por un fragmento de ácido nucleico aislado capaz de hibridarse a un ácido nucleico según SEQ ID N°: 1, preferiblemente bajo condiciones de hibridación muy rigurosas.

Según esto, el polipéptido TEMER07589 es preferiblemente un polipéptido que comprende una secuencia de aminoácidos al menos 76%, al menos 77%, al menos 78%, al menos 79%, al menos 80%, al menos 81%, al menos

82%, al menos 83%, al menos 84%, al menos 85%, al menos 86%, al menos 87%, al menos 88%, al menos 89%, al menos 90%, al menos 91%, 92%, 93%, 94%, 95%, 96%, 96%, 97%, 98%, 99% o más homóloga a la secuencia de aminoácidos mostrada en SEQ ID N°: 2 y, típicamente, retiene al menos una actividad funcional del polipéptido según SEQ ID N°: 2.

5 También se pueden identificar equivalentes funcionales de un polipéptido según la invención, p. ej., mediante bibliotecas combinatorias de mutantes, p. ej. mutantes de truncamiento, del polipéptido de la invención para la actividad potenciadora de celulasa. En una realización, se genera una biblioteca variegada de variantes mediante mutagénesis combinatoria a nivel de ácidos nucleicos. Una biblioteca variegada de variantes se puede producir, por ejemplo, al ligar enzimáticamente una mezcla de oligonucleótidos sintéticos en secuencias génicas de modo que un grupo degenerado de secuencias polipeptídicas potenciales sea expresable como polipéptidos individuales, o, alternativamente, como un grupo de proteínas de fusión mayores (p. ej. para la exhibición de fagos). Existe una variedad de métodos que se pueden usar para producir bibliotecas de variantes potenciales de los polipéptidos de la invención a partir de una secuencia oligonucleotídica degenerada. Se conocen en la técnica métodos para sintetizar oligonucleótidos degenerados (véanse, p. ej., Narang (1983) *Tetrahedron* 39:3; Itakura y cols. (1984) *Annu. Rev. Biochem.* 53:323; Itakura y cols. (1984) *Science* 198:1056; Ike y cols. (1983) *Nucleic Acid Res.* 11:477).

Además, se pueden usar bibliotecas de fragmentos de la secuencia codificante de un polipéptido de la invención para generar una población variegada de polipéptidos para secretar una selección posterior de variantes. Por ejemplo, se puede generar una biblioteca de fragmentos de secuencias codificantes al tratar un fragmento de PCR de doble hebra de la secuencia codificante de interés con una nucleasa bajo condiciones en las que se produce melladura solo aproximadamente una vez por molécula, desnaturalizar el ADN de doble hebra, renaturalizar el ADN para formar ADN de doble hebra que puede incluir pares de sentido/antisentido de diferentes productos mellados, retirar porciones de una sola hebra de dúplex reformados mediante tratamiento con S1 nucleasa y ligar la biblioteca de fragmentos resultante en un vector de expresión. Mediante este método, se puede derivar una biblioteca de expresión que codifica fragmentos N-terminales e internos de diversos tamaños de la proteína de interés.

Se conocen varias técnicas en la especialidad para cribar productos génicos de bibliotecas combinatorias mediante mutaciones puntuales de truncamiento, y para cribar bibliotecas de ADNc con respecto a productos génicos que tengan una propiedad seleccionada. Las técnicas más ampliamente usadas, que están disponibles para un análisis de alto rendimiento, para cribar bibliotecas génicas grandes incluyen típicamente clonar la biblioteca génica en vectores de expresión replicables, transformar células apropiadas con la biblioteca de vectores resultante y expresar los genes combinatorios bajo condiciones en las que la detección de una actividad deseada facilite el aislamiento del vector que codifica el gen cuyo producto se detectaba. La mutagénesis recursiva de conjunto (REM), una técnica que potencia la frecuencia de mutantes funcionales en las bibliotecas, se puede usar en combinación con los ensayos de cribado para identificar variantes de un polipéptido de la invención (Arkin y Yourvan (1992) *Proc. Natl. Acad. Sci. USA* 89:7811-7815; Delgrave y cols. (1993) *Protein Engineering* 6(3): 327-331).

Además de la secuencia génica de TEMER07589 mostrada en SEQ ID N°: 1, será evidente para el experto que pueden existir polimorfismos de secuencia de ADN dentro de una población dada, que pueden conducir a cambios en la secuencia de aminoácidos del polipéptido TEMER07589. Tales polimorfismos genéticos pueden existir en células procedentes de diferentes poblaciones o dentro de una población debido a variación alélica. Las variantes alélicas también pueden incluir equivalentes funcionales.

Los fragmentos de un polinucleótido según la invención también pueden comprender polinucleótidos que codifiquen polipéptidos funcionales. Estos polinucleótidos pueden funcionar como sonda o cebadores para una reacción PCR.

Los ácidos nucleicos según la invención independientemente de si codifican polipéptidos funcionales o no funcionales se pueden usar como sondas de hibridación o cebadores de la reacción en cadena de la polimerasa (PCR). Usos de las moléculas de ácido nucleico de la presente invención que no codifican un polipéptido que tiene una actividad de TEMER07589 incluyen, entre otros, (1) aislar el gen que codifica el polipéptido TEMER07589 o variantes alélicas del mismo de una biblioteca de ADNc, p. ej. de microorganismos adecuados; (2) hibridación in situ (p. ej. FISH) a extensiones cromosómicas metafásicas para proporcionar una localización cromosómica precisa del gen TEMER07589 según se describe en Verma y cols., *Human Chromosomes: a Manual of Basic Techniques*, Pergamon Press, Nueva York (1988); (3) análisis por transferencia Northern para detectar la expresión de ARNm de TEMER07589 en tejidos y/o células específicos y 4) sondas y cebadores que se pueden usar como una herramienta de diagnóstico para analizar la presencia de un ácido nucleico hibridable a la sonda de TEMER07589 en una muestra biológica (p. ej. tejido).

También es abarcado por la invención un método para obtener un equivalente funcional de un gen TEMER07589. Este método implica obtener una sonda marcada que incluye un ácido nucleico aislado que codifica la totalidad o una porción de la secuencia polipeptídica según SEQ ID N°: 2 o una variante del mismo; cribar una biblioteca de fragmentos de ácido nucleico con la sonda marcada bajo condiciones que permitan la hibridación de la sonda a fragmentos de ácido nucleico de la biblioteca, formado de ese modo dúplex de ácido nucleico, y preparar una

secuencia génica de longitud completa a partir de los fragmentos de ácido nucleico en cualquier dúplex marcado para obtener un gen relacionado con el gen TEMER07589.

5 En una realización, un ácido nucleico de TEMER07589 de la invención es al menos 80%, al menos 81%, al menos 82%, al menos 83%, al menos 84%, al menos 85%, al menos 86%, al menos 87%, al menos 88%, al menos 89%, al menos 90%, al menos 91%, al menos 92%, al menos 93%, al menos 94%, al menos 95%, al menos 96%, al menos 97%, al menos 98%, al menos 99% o más homólogo a una secuencia de ácido nucleico mostrada en SEQ ID N°: 1 o el complemento de la misma.

10 Se proporcionan además células hospedadoras que comprenden un polinucleótido o vector de la invención. El polinucleótido puede ser heterólogo al genoma de la célula hospedadora. El término "heterólogo", habitualmente con respecto a la célula hospedadora, significa que el polinucleótido no está presente naturalmente en el genoma de la célula hospedadora o que el polipéptido no es producido naturalmente por esa célula.

15 En otra realización, la invención presenta células, p. ej., células hospedadoras transformadas o células hospedadoras recombinantes que contienen un ácido nucleico abarcado por la invención. Una "célula transformada" o "célula recombinante" es una célula en la que (o en un progenitor de la cual) se ha introducido, por medio de técnicas de ADN recombinante, un ácido nucleico según la invención. Se incluyen células tanto procarióticas como eucarióticas, p. ej., bacterias, hongos, levaduras y similares, se prefieren especialmente células procedentes de  
20 hongos filamentosos, tales como *Aspergillus niger* o *Talaromyces emersonii*.

Se puede elegir una célula hospedadora que module la expresión de las secuencias insertadas o modifique y procese el producto génico de un modo deseado específico. Estas modificaciones (p. ej., glicosilación) y el procesamiento (p. ej., escisión) de productos proteínicos pueden facilitar el funcionamiento óptimo de la proteína.

25 Diversas células hospedadoras tienen mecanismos característicos y específicos para el procesamiento y la modificación postraduccionales de proteínas y productos génicos. Líneas celulares o sistemas hospedadores apropiados familiares para los expertos en la técnica de la biología molecular y/o la microbiología se pueden elegir para asegurar la modificación y el procesamiento deseados y correctos de la proteína extraña expresada. A este fin,  
30 se pueden usar células hospedadoras eucarióticas que poseen la maquinaria celular para un procesamiento apropiado del transcrito primario, glicosilación y fosforilación del producto génico. Estas células hospedadoras son muy conocidas en la técnica.

Si se desea, una célula como la descrita anteriormente se puede usar en la preparación de un polipéptido según la invención. Este método comprende típicamente cultivar una célula hospedadora (p. ej. transformada o transfectada con un vector de expresión como el descrito anteriormente) bajo condiciones que proporcionen la expresión (mediante el vector) de una secuencia codificante que codifica el polipéptido, y opcionalmente recuperar el polipéptido expresado. Los polinucleótidos de la invención se pueden incorporar en un vector replicable recombinante, p. ej. un vector de expresión. El vector se puede usar para replicar en ácido nucleico en una célula hospedadora compatible. Así, en una realización adicional, la invención proporciona un método para elaborar un polinucleótido de la invención al introducir un polinucleótido de la invención en un vector replicable, introducir el vector en una célula hospedadora compatible y hacer crecer la célula hospedadora bajo condiciones que lleven a cabo la replicación del vector. El vector se puede recuperar de la célula hospedadora.

45 Preferiblemente, el polipéptido se produce como una proteína secretada, en cuyo caso la secuencia nucleotídica que codifica una forma madura del polipéptido en la construcción de expresión se conecta operativamente a una secuencia nucleotídica que codifica una secuencia de señal. Preferiblemente, la secuencia de señal es natural (homóloga) a la secuencia nucleotídica que codifica el polipéptido. Alternativamente, la secuencia de señal es extraña (heteróloga) a la secuencia nucleotídica que codifica el polipéptido, en cuyo caso la secuencia de señal es preferiblemente endógena a la célula hospedadora en la que se expresa la secuencia nucleotídica según la invención. Ejemplos de secuencias de señal adecuadas para células hospedadoras de levadura son las secuencias de señal derivadas de genes de factor  $\alpha$  de levadura. De forma similar, una secuencia de señal adecuada para células hospedadoras de hongos filamentosos es, p. ej., una secuencia de señal derivada de un gen de amiloglucosidasa (AG) de hongo filamentosos, p. ej. el gen *glaA* de *A. niger*. Este se puede usar en combinación con el propio promotor de amiloglucosidasa (también llamada (gluco)amilasa), así como en combinación con otros promotores. También se pueden usar secuencias de señal híbridas con el contexto de la presente invención.

60 Secuencias líder de secreción heterólogas preferidas son las que se originan a partir del gen de amiloglucosidasa (AG) fúngica (*glaA*-las versiones tanto de 18 como de 24 aminoácidos, p. ej., procedentes de *Aspergillus*), el gen de factor  $\alpha$  (levaduras, p. ej. *Saccharomyces* y *Kluyveromyces*) o el gen de  $\alpha$ -amilasa (*Bacillus*).

Los vectores se pueden transformar o transfectar en una célula hospedadora adecuada según se describe anteriormente para proporcionar la expresión de un polipéptido de la invención. Este procedimiento puede comprender cultivar una célula hospedadora transformada con un vector de expresión según se describe

anteriormente bajo condiciones que proporcionen la expresión mediante el vector de una secuencia codificante que codifica el polipéptido.

#### Células hospedadoras

5 La invención proporciona así células hospedadoras transformadas o transfectadas con o que comprenden un polinucleótido o vector de la invención. Preferiblemente, el polinucleótido es portado en un vector para la replicación y la expresión del polinucleótido. Las células se elegirán para que sean compatibles con dicho vector y pueden ser, por ejemplo, células procarióticas (por ejemplo bacterianas), fúngicas, de levadura o vegetales.

10 También se puede elegir un huésped heterólogo en el que el polipéptido de la invención se produce en una forma que está sustancialmente libre de otras enzimas de degradación de celulosa o de degradación de hemicelulosa. Esto se puede conseguir al elegir un huésped que no produzca normalmente estas enzimas.

15 La invención abarca procedimientos para la producción del polipéptido de la invención por medio de la expresión recombinante de una secuencia de ADN que codifica el polipéptido. Con este propósito, la secuencia de ADN de la invención se puede usar para la amplificación génica y/o el intercambio de señales de expresión, tales como promotores, secuencias de señal de secreción, a fin de permitir la producción económica del polipéptido en una célula hospedadora homóloga o heteróloga adecuada. Una célula hospedadora homóloga es una célula hospedadora que es de la misma especie o que es una variante dentro de la misma especie que la especie de la que se deriva la secuencia de ADN.

20 Células hospedadoras adecuadas con preferiblemente microorganismos procarióticos tales como bacterias, o más preferiblemente organismos eucarióticos, por ejemplo hongos, tales como levaduras u hongos filamentosos, o células vegetales. En general, las células de levadura se prefieren sobre las células fúngicas debido a que son más fáciles de manipular. Sin embargo, algunas proteínas bien son secretadas escasamente de las levaduras o bien en algunos casos no se procesan apropiadamente (p. ej. hiperglicosilación en levadura). En estos casos, se debe seleccionar un organismo hospedador fúngico.

30 La célula hospedadora puede sobreexpresar el polipéptido, y se conocen bien las técnicas para manipular la sobreexpresión. El hospedador puede tener así dos o más copias del polinucleótido codificante (y según esto el vector puede tener así dos o más copias).

35 Las bacterias del género *Bacillus* son muy adecuadas como hospedadores heterólogos debido a su capacidad para secretar proteínas en el medio de cultivo. Otras bacterias adecuadas como hospedadores son las de los géneros *Streptomyces* y *Pseudomonas*. Una célula hospedadora de levadura preferida para la expresión de la secuencia de ADN que codifica el polipéptido es de los géneros *Saccharomyces*, *Kluyveromyces*, *Hansenula*, *Pichia*, *Yarrowia*, y *Schizosaccharomyces*.

40 Más preferiblemente, una célula hospedadora de levadura se selecciona del grupo que consiste en las especies *Saccharomyces cerevisiae*, *Kluyveromyces lactis* (también conocida como *Kluyveromyces marxianus* var. *lactis*), *Hansenula polymorpha*, *Pichia pastoris*, *Yarrowia lipolytica* y *Schizosaccharomyces pombe*.

45 Sin embargo, las más preferidas son las células hospedadoras de hongos (p. ej. filamentosos). Células hospedadoras de hongos filamentosos preferidas se seleccionan del grupo que consiste en los géneros *Aspergillus*, *Trichoderma/Hypocrea*, *Fusarium*, *Disporotrichum*, *Penicillium*, *Acremonium*, *Neurospora*, *Thermoascus*, *Myceliophthora*, *Sporotrichum*, *Thielavia*, *Chryso sporium*, *Fusarium*, *Humicola*, *Neurospora* y *Talaromyces*.

50 Más preferiblemente, una célula hospedadora de hongo filamentosos es de las especies que incluyen, pero no se limitan a, *Aspergillus niger*, *Aspergillus awamori*, *Aspergillus tubingensis*, *Aspergillus aculeatus*, *Aspergillus foetidus*, *Aspergillus nidulans*, *Aspergillus japonicus*, *Aspergillus oryzae* and *Aspergillus ficuum*, *Trichoderma reesei/Hypocrea jecorina*, *Fusarium graminearum*, *Talaromyces emersonii*, *Penicillium decumbens*, *Acremonium alabamense*, *Neurospora crassa*, *Myceliophthora thernaophilurri*, *Sporotrichum cellulophilum*, *Disporotrichum dimorphosporum*, *Talaromyces emersonii*, *Talaromyces stipitatus* y *Thielavia terrestris*.

55 Las células hospedadoras según la invención incluyen células vegetales, y la invención se extiende por lo tanto a organismos transgénicos, tales como plantas y partes de las mismas, que contienen una o más células de la invención. Las células pueden expresar heterológamente el polipéptido de la invención o pueden contener heterológamente uno o más de los polinucleótidos de la invención. Por lo tanto, la planta transgénica (o genéticamente modificada) puede tener insertado (p. ej. establemente) en su genoma una secuencia que codifica uno o más de los polipéptidos de la invención. La transformación de células vegetales se puede realizar usando técnicas conocidas, por ejemplo usando un plásmido Ti o Ri procedente de *Agrobacterium tumefaciens*. El plásmido (o vector) puede contener así secuencias necesarias para infectar una planta, y se pueden emplear derivados de los plásmidos Ti y/o Ri.

Alternativamente, se puede efectuar la infección directa de una planta, tal como una hoja, raíz o tallo. En esta técnica la planta que se va a infectar se puede herir, por ejemplo, al cortar la planta con una cuchilla o perforar la planta con una aguja o frotar la planta con un abrasivo. A continuación, la herida se inocula con la *Agrobacterium*. A continuación, la planta o parte de planta se puede hacer crecer sobre un medio de cultivo adecuado y dejar desarrollar hasta una planta madura. La regeneración de plantas transformadas en plantas genéticamente modificadas se puede conseguir al usar técnicas conocidas, por ejemplo al seleccionar brotes transformados usando un antibiótico y subcultivar los brotes sobre un medio que contienen los nutrientes, hormonas de plantas y similares apropiados.

La invención también incluye células que se han modificado para expresar el polipéptido mejorador de celulasa de la invención o una variante del mismo. Estas células incluyen líneas celulares eucarióticas superiores transitorias, o preferiblemente estables, tales como células de mamíferos o células de insectos, células eucarióticas inferiores, tales como células de levaduras y hongos (p. ej. filamentosos) o células procarióticas tales como células bacterianas.

También es posible que los polipéptidos de la invención se expresen transitoriamente en una línea celular o sobre una membrana, tal como, por ejemplo, en un sistema de expresión de baculovirus. Estos sistemas, que se adaptan para expresar los polipéptidos según la invención, también se incluyen dentro del alcance de la presente invención.

Según la presente invención, la producción del polipéptido de la invención se puede efectuar mediante el cultivo de hospedadores de expresión microbianos, que se han transformado con uno o más polinucleótidos de la presente invención, en un medio de fermentación con nutrientes convencional.

#### Producción de polipéptidos/enzimas

Las células hospedadoras recombinantes según la invención se pueden cultivar usando procedimientos conocidos en la técnica. Para cada combinación de un promotor y una célula hospedadora, están disponibles condiciones de cultivo que conducen a la expresión de la secuencia de ADN que codifica el polipéptido. Después de alcanzar la densidad o concentración celular deseada del polipéptido, el cultivo se detiene y el polipéptido se recupera usando procedimientos conocidos.

El medio de fermentación puede comprender un medio de cultivo conocido que contiene una fuente de carbono (p. ej. glucosa, maltosa, melazas, almidón, celulosa, xilano, pectina, hidrolizado de biomasa lignocelulósica, etc.), una fuente de nitrógeno (p. ej. sulfato amónico, nitrato amónico, cloruro amónico, etc.), una fuente de nitrógeno orgánico (p. ej. extracto de levadura, extracto de malta, peptona, etc.) y fuentes de nutrientes inorgánicos (p. ej. fosfato, magnesio, potasio, cinc, hierro, etc.). Opcionalmente, se puede incluir un inductor (p. ej. celulosa, pectina, xilano, maltosa, maltodextrina o xilogalacturonano).

La selección del medio apropiado se puede basar en la elección del hospedador de expresión y/o basar en los requisitos reguladores de la construcción de expresión. Estos medios son conocidos por los expertos en la técnica. Si se desea, el medio puede contener componentes adicionales que favorecen los hospedadores de expresión transformados sobre otros microorganismos potencialmente contaminantes.

La producción de polipéptido por el hospedador transformado (fermentación) se puede realizar según cualquier procedimiento conocido. El tiempo de producción se puede prolongar durante un período de aproximadamente 0,5 a aproximadamente 30 días. Puede ser un procedimiento discontinuo, continuo o alimentado por lotes, adecuadamente a una temperatura en el intervalo de 0-100°C o 0-80°C, por ejemplo, de aproximadamente 0 a aproximadamente 60°C y/o a un pH, por ejemplo, de aproximadamente 2 a aproximadamente 10, o de aproximadamente 3 a aproximadamente 9. Condiciones de fermentación preferidas son una temperatura en el intervalo de aproximadamente 20 a aproximadamente 55°C y/o a un pH de aproximadamente 3 a aproximadamente 5. Las condiciones apropiadas se seleccionan habitualmente basándose en la elección del hospedador de expresión y el polipéptido que se va a expresar.

Después de la fermentación, si es necesario, las células se pueden retirar del caldo de fermentación por medio de centrifugación o filtración. Después de que la fermentación se haya detenido o después de la retirada de las células, el polipéptido de la invención se puede recuperar y, si se desea, purificar y aislar por medios convencionales.

#### Composiciones de polipéptido/enzima

La invención proporciona una composición que comprende un polipéptido de la invención y una celulasa y/o una hemicelulasa y/o una pectinasa.

Cuando el polipéptido de la invención es una celulasa, una composición de la invención comprenderá típicamente una hemicelulasa y/o una pectinasa además del polipéptido de la invención.

5 Cuando el polipéptido de la invención es una hemicelulasa, una composición de la invención típicamente comprenderá una celulasa y/o una pectinasa además del polipéptido de la invención.

Cuando el polipéptido de la invención es una pectinasa, una composición de la invención comprenderá típicamente una celulasa y/o una hemicelulasa además del polipéptido de la invención.

10 Una composición de la invención comprenderá una, dos o tres o más clases de celulasa, por ejemplo una, dos o la totalidad de una endo-1,4- $\beta$ -glucanasa (EG), una exo-celobiohidrolasa (CBH) y una  $\beta$ -glucosidasa (BGL).

15 Una composición de la invención puede comprender un polipéptido que tiene la misma actividad enzimática, por ejemplo el mismo tipo de actividad de celulasa, hemicelulasa y/o pectinasa que el proporcionado por un polipéptido de la invención.

20 Una composición de la invención puede comprender un polipéptido que tiene un tipo diferente de actividad de celulasa y/o actividad de hemicelulasa y/o actividad de pectinasa que el proporcionado por un polipéptido de la invención. Por ejemplo, una composición de la invención puede comprender un tipo de actividad de celulasa y/o hemicelulasa y/o actividad de pectinasa proporcionado por un polipéptido de la invención y un segundo tipo de actividad de celulasa y/o hemicelulasa y/o actividad de pectinasa proporcionado por una hemicelulasa/pectinasa adicional.

25 En la presente, una celulasa es cualquier polipéptido que sea capaz de degradar celulosa. Un polipéptido que es capaz de degradar celulosa es uno que es capaz de catalizar el proceso de descomponer la celulosa en unidades menores, bien parcialmente, por ejemplo en celodextrinas, o bien completamente en monómeros de glucosa. Una celulasa según la invención puede dar lugar a una población mixta de celodextrinas y monómeros de glucosa cuando se pone en contacto con la celulosa. Esta degradación tendrá lugar típicamente por medio de una reacción de hidrólisis.

30 En la presente, una hemicelulasa es cualquier polipéptido que sea capaz de degradar hemicelulosa. Es decir, una hemicelulasa puede ser capaz de degradar uno o más de *xy*lano, glucuronoxilano, arabinoxilano, glucomanano y xiloglucano. Un polipéptido que es capaz de degradar una hemicelulosa es uno que es capaz de catalizar el proceso de descomposición de la hemicelulosa en polisacáridos menores, bien parcialmente, por ejemplo en oligosacáridos, o bien completamente en monómeros sacáricos, por ejemplo monómeros sacáricos de hexosa o pentosa. Una hemicelulasa según la invención puede dar lugar a una población mixta de oligosacáridos y monómeros sacáricos cuando se pone en contacto con la hemicelulosa. Tal degradación tendrá lugar típicamente por medio de una reacción de hidrólisis.

35 En la presente, una pectinasa es cualquier polipéptido que sea capaz de degradar pectina. Un polipéptido que sea capaz de degradar pectina es uno que es capaz de catalizar el proceso de descomposición de pectina en unidades menores, bien parcialmente, por ejemplo en oligosacáridos, o bien completamente en monómeros sacáricos. Una pectinasa según la invención puede dar lugar a una población mixta de oligosacáridos y monómeros sacáricos cuando se pone en contacto con la pectinasa. Tal degradación tendrá lugar típicamente por medio de una reacción de hidrólisis.

Según esto, una composición de la invención puede comprender cualquier celulasa, por ejemplo, una celobiohidrolasa, una endo- $\beta$ -1,4-glucanasa, una  $\beta$ -glucosidasa o una  $\beta$ -(1,3)(1,4)-glucanasa.

40 En la presente, una celobiohidrolasa es cualquier polipéptido que sea capaz de catalizar la hidrólisis de enlaces 1,4- $\beta$ -D-glucosídicos en celulosa o celotetraosa, liberando celobiosa de los extremos de las cadenas. Esta enzima también se puede denominar celulasa, 1,4- $\beta$ -celobiosidasa, 1,4- $\beta$ -celobiohidrolasa, 1,4- $\beta$ -D-glucano celobiohidrolasa, avicelasa, exo-1,4- $\beta$ -D-glucanasa, potenciador de exocelulasa o exoglucanasa. Puede tener el código EC EC 3.2.1.91. Las celobiohidrolasas se pueden subdividir en celobiohidrolasa I (CBH I) y celobiohidrolasa II (CBH II). CBH I se define como celobiohidrolasas que hidrolizan celulosa predominantemente desde los extremos reductores, fragmentando la celobiosa. CBH II se define como celobiohidrolasas que hidrolizan celulosa predominantemente desde los extremos no reductores, fragmentando la celobiosa.

45 En la presente, una endo- $\beta$ -1,4-glucanasa (EC 3.2.1.4) es cualquier polipéptido que sea capaz de catalizar la endohidrólisis de enlaces 1,4- $\beta$ -D-glucosídicos en celulosa, liquenina o  $\beta$ -D-glucano de cereales. Este polipéptido también puede ser capaz de hidrolizar enlaces 1,4 en  $\beta$ -D-glucano que también contiene enlaces 1,3. Esta enzima también se puede denominar celulasa, avicelasa,  $\beta$ -1,4-endoglucano hidrolasa,  $\beta$ -1,4-glucanasa, carboximetilcelulasa, celudextrinasa, endo-1,4- $\beta$ -D-glucanasa, endo-1,4- $\beta$ -D-glucanohidrolasa, endo-1,4- $\beta$ -glucanasa o endoglucanasa. CEA es en la presente una endoglucanasa EC 3.2.1.4 que basándose en su estructura tridimensional se clasifica bajo la familia de glicosil hidrolasas 5 (GH5). Actividades conocidas para miembros de la

familia GH5 incluyen quitosanasa (EC 3.2.1.132);  $\beta$ -manosidasa (EC 3.2.1.25); celulosa (EC 3.2.1.4); glucano 1,3- $\beta$ -glucosidasa (EC 3.2.1.58); liqueninasa (EC 3.2.1.73); glucano endo-1,6- $\beta$ -glucosidasa (EC 3.2.1.75); manano endo- $\beta$ -1,4-manosidasa (EC 3.2.1.78); endo- $\beta$ -1,4-xilanasa (EC 3.2.1.8); celulosa  $\beta$ -1,4-celobiosidasa (EC 3.2.1.91); endo- $\beta$ -1,6-galactanasa (EC 3.2.1.-);  $\beta$ -1,3-mananasa (EC 3.2.1.-); endo- $\beta$ -1,4-glucanasa específica de xiloglucano (EC 3.2.1.151); manano transglicosilasa (EC 2.4.1.-). CEB en la presente una endoglucanasa EC 3.2.1.4 que basándose en su estructura tridimensional se clasifica bajo la familia de glicosil hidrolasa 7 (GH7). Actividades conocidas para miembros de la familia GH7 incluyen endo- $\beta$ -1,4-glucanasa (EC 3.2.1.4); celobiohidrolasa [que actúa sobre extremos reductores] (EC 3.2.1.-); quitosanasa (EC 3.2.1.132); endo- $\beta$ -1,3-1,4-glucanasa (EC 3.2.1.73)

En la presente, una  $\beta$ -glucosidasa (abreviada BG) (EC 3.2.1.21) es cualquier polipéptido que sea capaz de catalizar la hidrólisis de residuos de  $\beta$ -D-glucosa no reductores terminales con liberación de  $\beta$ -D-glucosa. Este polipéptido puede tener una amplia especificidad para  $\beta$ -D-glucósidos y también puede hidrolizar uno o más de los siguientes: a  $\beta$ -D-galactósido, un  $\alpha$ -L-arabinósido, un  $\beta$ -D-xilósido o un  $\beta$ -D-fucósido. Esta enzima también se puede denominar amigdalasa,  $\beta$ -D-glucósido glucohidrolasa, celobiasa o gentiobiasa.

En la presente, una  $\beta$ -(1,3)(1,4)-glucanasa (EC 3.2.1.73) es cualquier polipéptido que sea capaz de catalizar la hidrólisis de enlaces 1,4- $\beta$ -D-glucosídicos en  $\beta$ -D-glucanos que contienen uniones 1,3 y 1,4. Este polipéptido puede actuar sobre liquenina y  $\beta$ -D-glucanos de cereales, pero no sobre  $\beta$ -D-glucanos que contienen solamente uniones 1,3 o 1,4. Esta enzima también se puede denominar liqueninasa, 1,3-1,4- $\beta$ -D-glucano 4-glucanohidrolasa,  $\beta$ -glucanasa, endo- $\beta$ -1,3-1,4 glucanasa, liquenasa o  $\beta$ -glucanasa con enlaces mixtos. Una alternativa para este tipo de enzima es EC 3.2.1.6, que se describe como endo-1,3(4)-beta-glucanasa. Este tipo de enzima hidroliza enlaces 1,3 o 1,4 en beta-D-glucanos cuando el residuo de glucosa cuyo grupo reductor está implicado en el enlace que se va a hidrolizar está él mismo sustituido en C-3. Nombres alternativos incluyen endo-1,3-beta-glucanasa, laminarinasa, 1,3-(1,3;1,4)-beta-D-glucano 3 (4) glucanohidrolasa; los sustratos incluyen laminarina, liquenina y beta-D-glucanos de cereales.

Una composición de la invención puede comprender cualquier hemicelulasa, por ejemplo, una endoxilanasa, una  $\beta$ -xilosidasa, una  $\alpha$ -L-arabinofuranosidasa, una  $\alpha$ -D-glucuronidasa, una acetil xilano esterasa, una feruloil esterasa, una cumaroil esterasa, una  $\alpha$ -galactosidasa, una  $\beta$ -galactosidasa, una  $\beta$ -mananasa o una  $\beta$ -manosidasa.

En la presente, una endoxilanasa (EC 3.2.1.8) es cualquier polipéptido que sea capaz de catalizar la endohidrólisis de enlaces 1,4- $\beta$ -D-xilosídicos en xilanos. Esta enzima también se puede denominar endo-1,4- $\beta$ -xilanasa o 1,4- $\beta$ -D-xilano xilanhidrolasa. Una alternativa es EC 3.2.1.136, una glucuronoarabinoxilano endoxilanasa, una enzima que es capaz de hidrolizar enlaces 1,4 xilosídicos en glucuronoarabinoxilanos.

En la presente, una  $\beta$ -xilosidasa (EC 3.2.1.37) es cualquier polipéptido que sea capaz de catalizar la hidrólisis de 1,4- $\beta$ -D-xilanos, para retirar residuos de D-xilosa sucesivos de términos no reductores. Estas enzimas también pueden hidrolizar xilobiosa. Esta enzima también se puede denominar xilano 1,4- $\beta$ -xilosidasa, 1,4- $\beta$ -D-xilano xilohidrolasa, exo-1,4- $\beta$ -xilosidasa o xilobiasa.

En la presente, una  $\alpha$ -L-arabinofuranosidasa (EC 3.2.1.55) es cualquier polipéptido que sea capaz de actuar sobre  $\alpha$ -L-arabinofuranósidos,  $\alpha$ -L-arabinanos que contienen enlaces (1,2) y/o (1,3) y/o (1,5), arabinoxilanos y arabinogalactanos. Esta enzima también se puede denominar  $\alpha$ -N-arabinofuranosidasa, arabinofuranosidasa o arabinosidasa.

En la presente, una  $\alpha$ -D-glucuronidasa (EC 3.2.1.139) es cualquier polipéptido que sea capaz de catalizar una reacción de la siguiente forma: alfa-D-glucuronósido + H(2)O = un alcohol + D-glucuronato. Esta enzima también se puede denominar alfa-glucuronidasa o alfa-glucosiduronasa. Estas enzimas también pueden hidrolizar ácido glucorónico metilado en 4-O, que también puede estar presente como un sustituyente en xilanos. Una alternativa es EC 3.2.1.131: xilano alfa-1,2-glucuronosidasa, que cataliza la hidrólisis de enlaces alfa-1,2-(4-O-metil)glucuronosilo.

En la presente, una acetil xilano esterasa (EC 3.1.1.72) es cualquier polipéptido que sea capaz de catalizar la desacetilación de xilanos y xilo-oligosacáridos. Este polipéptido puede catalizar la hidrólisis de grupos acetilo de xilano polimérico, xilosa acetilada, glucosa acetilada, acetato de naftilo alfa o acetato de p-nitrofenilo, pero, típicamente, no de triacilglicerol. Este polipéptido típicamente no actúa sobre manano o pectina acetilados.

En la presente, una feruloil esterasa (EC 3.1.1.73) es cualquier polipéptido que sea capaz de catalizar una reacción de la forma: feruloil-sacárido + H(2)O = ferulato + sacárido. El sacárido puede ser, por ejemplo, un oligosacárido o un polisacárido. Típicamente, puede catalizar la hidrólisis del grupo 4-hidroxi-3-metoxicinamoilo (feruloilo) de un azúcar esterificado, que habitualmente es arabinosa en sustratos 'naturales'. El acetato de p-nitrofenol y el ferulato de metilo son sustratos típicamente más pobres. Esta enzima también se puede denominar éster cinamofílico hidrolasa, ácido ferúlico esterasa o hidroxicinamoil esterasa. También se puede denominar una enzima accesoria de hemicelulasa, ya que puede ayudar a las xilanasas y pectinasas a descomponer la hemicelulosa y la pectina de la pared de células vegetales.

En la presente, una cumaroil esterasa (EC 3.1.1.73) es cualquier polipéptido que sea capaz de catalizar una reacción de la forma: cumaroil-sacárido + H(2)O = cumarato + sacárido. El sacárido puede ser, por ejemplo, un oligosacárido o un polisacárido. Esta enzima también se puede denominar trans-4-cumaroil esterasa, trans-p-cumaroil esterasa, p-cumaroil esterasa o ácido p-cumárico esterasa. Esta enzima también está dentro de EC 3.1.1.73 de modo que también se puede denominar una feruloil esterasa.

En la presente, una  $\alpha$ -galactosidasa (EC 3.2.1.22) es cualquier polipéptido que sea capaz de catalizar la hidrólisis de residuos de  $\alpha$ -D-galactosa no reductores terminales en  $\alpha$ -D-galactósidos, incluyendo oligosacáridos de galactosa, galactomananos, galactanos y arabinogalactanos. Este polipéptido también puede ser capaz de hidrolizar  $\alpha$ -D-fucósidos. Esta enzima también se puede denominar melibiasa.

En la presente, una  $\beta$ -galactosidasa, (EC 3.2.1.23) es cualquier polipéptido que sea capaz de catalizar la hidrólisis de residuos de  $\beta$ -D-galactosa no reductores terminales en  $\beta$ -D-galactósidos. Este polipéptido también puede ser capaz de hidrolizar  $\alpha$ -L-arabinósidos. Esta enzima también se puede denominar exo-(1 $\rightarrow$ 4)- $\beta$ -D-galactanasa o lactasa.

En la presente, una  $\beta$ -mananasa (EC 3.2.1.78) es cualquier polipéptido que sea capaz de catalizar la hidrólisis aleatoria de enlaces 1,4- $\beta$ -D-manosídicos en mananos, galactomananos y glucomananos. Esta enzima también se puede denominar manano endo-1,4- $\beta$ -manosidasa o endo-1,4-mananasa.

En la presente, una  $\beta$ -manosidasa (EC 3.2.1.25) es cualquier polipéptido que sea capaz de catalizar la hidrólisis de residuos de  $\beta$ -D-manosa no reductores terminales en  $\beta$ -D-manósidos. Esta enzima también se puede denominar mananasa o manasa.

Una composición de la invención puede comprender cualquier pectinasa, por ejemplo una endopoligalacturonasa, una pectina metil esterasa, una endo-galactanasa, una beta galactosidasa, una pectina acetil esterasa, una endo-pectina liasa, pectato liasa, alfa-ramnosidasa, una exo-galacturonasa, una exo-poligalacturonato liasa, una ramnogalacturonano hidrolasa, una ramnogalacturonano liasa, una ramnogalacturonano acetil esterasa, una ramnogalacturonano galacturonohidrolasa o una xilogalacturonasa.

En la presente, una endo-poligalacturonasa (EC 3.2.1.15) es cualquier polipéptido que sea capaz de catalizar la hidrólisis aleatoria de enlaces 1,4- $\alpha$ -D-galactosidurónicos en pectato y otros galacturonanos. Esta enzima también se puede denominar poligalacturonasa pectina despolimerasa, pectinasa, endopoligalacturonasa, pectolasa, pectina hidrolasa, pectina poligalacturonasa, poli- $\alpha$ -1,4-galacturónido glicanohidrolasa, endogalacturonasa; endo-D-galacturonasa o poli(1,4- $\alpha$ -D-galacturónido) glicanohidrolasa.

En la presente, una pectina metil esterasa (EC 3.1.1.11) es cualquier enzima que sea capaz de catalizar la reacción: pectina + n H<sub>2</sub>O = n metanol + pectato. La enzima también se ha conocido como pectinesterasa, pectina desmetoxilasa, pectina metoxilasa, pectina metilesterasa, pectasa, pectinoesterasa o pectina pectilhidrolasa.

En la presente, una endo-galactanasa (EC 3.2.1.89) es cualquier enzima capaz de catalizar la endohidrólisis de enlaces 1,4- $\beta$ -D-galactosídicos en arabinogalactanos. La enzima también se puede conocer como arabinogalactano endo-1,4- $\beta$ -galactosidasa, endo-1,4- $\beta$ -galactanasa, galactanasa, arabinogalactanasa o arabinogalactano 4- $\beta$ -D-galactanohidrolasa.

En la presente, una pectina acetil esterasa se define en la presente como cualquier enzima que tenga actividad de acetil esterasa que catalice la desacetilación de los grupos acetilo en los grupos hidroxilo de residuos de GalUA de pectina.

En la presente, un endo-pectina liasa (EC 4.2.2.10) es cualquier enzima capaz de catalizar la escisión eliminativa de éster metílico de (1 $\rightarrow$ 4)- $\alpha$ -D-galacturonano para dar oligosacáridos con grupos 4-desoxi-6-O-metil- $\alpha$ -D-galact-4-enuronosilo en sus extremos no reductores. La enzima también se puede conocer como pectina liasa, pectina trans-eliminasa; endo-pectina liasa, polimetilgalacturónico transeliminasa, pectina metiltranseliminasa, pectoliasa, PL, PNL o PMGL o (1 $\rightarrow$ 4)-6-O-metil- $\alpha$ -D-galacturonano liasa.

En la presente, una pectato liasa (EC 4.2.2.2) es cualquier enzima capaz de catalizar la escisión eliminativa de (1 $\rightarrow$ 4)- $\alpha$ -D-galacturonano para dar oligosacáridos con grupos 4-desoxi- $\alpha$ -D-galact-4-enuronosilo en sus extremos no reductores. La enzima también se puede conocer como poligalacturónico transeliminasa, ácido péctico transeliminasa, poligalacturonato liasa, endopectina metiltranseliminasa, pectato transeliminasa, endogalacturonato transeliminasa, ácido péctico liasa, liasa péctica, ácido  $\alpha$ -1,4-D-endopoligalacturónico liasa, PGA liasa, PPasa-N, ácido endo- $\alpha$ -1,4-poligalacturónico liasa, ácido poligalacturónico liasa, pectina trans-eliminasa, ácido poligalacturónico trans-eliminasa o (1 $\rightarrow$ 4)- $\alpha$ -D-galacturonano liasa.

En la presente, una alfa ramnosidasa (EC 3.2.1.40) es cualquier polipéptido que sea capaz de catalizar la hidrólisis de residuos de  $\alpha$ -L-ramnosa no reductores terminales en  $\alpha$ -L-ramnósidos o alternativamente en

ramnogalacturonano. Esta enzima también se puede conocer como  $\alpha$ -L-ramnosidasa T,  $\alpha$ -L-ramnosidasa N o  $\alpha$ -L-ramnósido ramnohidrolasa.

5 En la presente, exo-galacturonasa (EC 3.2.1.82) es cualquier polipéptido capaz de la hidrólisis de ácido péctico del digalacturonato del extremo no reductor, liberando digalacturonato. La enzima también se puede conocer como exo-poli- $\alpha$ -galacturonosidasa, exopoligalacturonosidasa o exopoligalacturonosidasa.

10 En la presente, exo-galacturonasa (EC 3.2.1.67) es cualquier polipéptido capaz de catalizar:  $(1,4\text{-}\alpha\text{-D-galacturónido})_n + \text{H}_2\text{O} = (1,4\text{-}\alpha\text{-D-galacturónido})_{n-1} + \text{D-galacturonato}$ . La enzima también se puede conocer como galacturano 1,4- $\alpha$ -galacturonidasa, exopoligalacturonasa, poli(galacturonato) hidrolasa, exo-D-galacturonasa, exo-D-galacturonanasa, exopoli-D-galacturonasa o poli(1,4- $\alpha$ -D-galacturónido) galacturonohidrolasa.

15 En la presente, exopoligalacturonato liasa (EC 4.2.2.9) es cualquier polipéptido capaz de catalizar la escisión eliminativa de 4-(4-desoxi- $\alpha$ -D-galact-4-enuronosil)-D-galacturonato del extremo reductor de pectato, es decir pectina desesterificada. Esta enzima se puede conocer como pectato disacárido-liasa, pectato exo-liasa, ácido exopéctico transeliminasa, exopectato liasa, ácido exopoligalacturónico trans-eliminasa, PATE, exo-PATE, exo-PGL o endo-disacárido-liasa reductora de (1 $\rightarrow$ 4)- $\alpha$ -D-galacturonano.

20 En la presente, ramnogalacturonano hidrolasa es cualquier polipéptido que sea capaz de hidrolizar el enlace entre ácido galactosilurónico y ramnopiranosilo de un modo endo en estructuras de ramnogalacturonano estrictamente alternas, que consisten en el disacárido [ácido (1,2- $\alpha$ -L-ramnoil-(1,4)- $\alpha$ -galactosilurónico)].

25 En la presente, ramnogalacturonano liasa es cualquier polipéptido que sea capaz de escindir enlaces  $\alpha$ -L-Rhap-(1 $\rightarrow$ 4)- $\alpha$ -D-GalpA de un modo endo en ramnogalacturonano mediante beta-eliminación.

En la presente, ramnogalacturonano acetil esterasa es cualquier polipéptido que catalice la desacetilación del esqueleto de residuos de ramnosa y ácido galacturónico alternativos en ramnogalacturonano.

30 En la presente, ramnogalacturonano galacturonohidrolasa es cualquier polipéptido que sea capaz de hidrolizar ácido galacturónico del extremo no reductor de estructuras de ramnogalacturonano estrictamente alternas de un modo exo.

35 En la presente, xilogalacturonasa es cualquier polipéptido que actúa sobre xilogalacturonano al escindir el esqueleto de ácido galacturónico sustituido con  $\beta$ -xilosa de un modo endo. Esta enzima también se puede conocer como xilogalacturonano hidrolasa.

40 En la presente, una  $\alpha$ -L-arabinofuranosidasa (EC 3.2.1.55) es cualquier polipéptido que sea capaz de actuar sobre  $\alpha$ -L-arabinofuranósidos,  $\alpha$ -L-arabinanos que contienen enlaces (1,2) y/o (1,3) y/o (1,5), arabinoxilanos y arabinogalactanos. Esta enzima también se puede denominar  $\alpha$ -N-arabinofuranosidasa, arabinofuranosidasa o arabinosidasa.

45 En la presente, endo-arabinanasa (EC 3.2.1.99) es cualquier polipéptido que sea capaz de catalizar la endohidrólisis de enlaces 1,5- $\alpha$ -arabinofuranosídicos en 1,5-arabinanos. La enzima también se puede conocer como endo-arabinasa, arabinano endo-1,5- $\alpha$ -L-arabinosidasa, endo-1,5- $\alpha$ -L-arabinanasa, endo- $\alpha$ -1,5-arabanasa; endo-arabanasa o 1,5- $\alpha$ -L-arabinano 1,5- $\alpha$ -L-arabinanohidrolasa.

50 Una composición de la invención comprenderá típicamente al menos una celulasa y/o al menos una hemicelulasa y/o al menos una pectinasa (una de las cuales es un polipéptido según la invención). Una composición de la invención puede comprender una celobiohidrolasa, una endoglucanasa y/o una  $\beta$ -glucosidasa. Esta composición también puede comprender una o más hemicelulasas y/o una o más pectinasas.

Una o más (por ejemplo dos, tres, cuatro o la totalidad) de una amilasa, una proteasa, una lipasa, una ligninasa, una hexosiltransferasa o una glucuronidasa pueden estar presentes en una composición de la invención.

55 "Proteasa" incluye enzimas que hidrolizan uniones peptídicas (peptidasas), así como enzimas que hidrolizan uniones entre péptidos y otros restos, tales como azúcares (glicopeptidasas). Muchas proteasas se caracterizan bajo EC 3.4, y son adecuadas para el uso en la invención. Algunos tipos específicos de proteasas incluyen cisteína proteasas incluyendo pepsina, papaína y serina proteasas incluyendo quimotripsinas, carboxipeptidasas y metaloendopeptidasas.

60 "Lipasa" incluye enzimas que hidrolizan lípidos, ácidos grasos y acilglicéridos, incluyendo fosfoglicéridos, lipoproteínas, diacilgliceroles y similares. En plantas, se usan lípidos como componentes estructurales para limitar la pérdida de agua y la infección por patógenos. Estos lípidos incluyen ceras derivadas de ácidos grasos, así como cutina y suberina.

65

"Ligninasa" incluye enzimas que pueden hidrolizar o descomponer la estructura de polímeros de lignina. Enzimas que puede descomponer la lignina incluyen lignina peroxidasas, manganeso peroxidasas, lacasas y feruloil esterases, y otras enzimas descritas en la técnica conocidas por despolimerizar o romper de otro modo polímeros de lignina. También se incluyen enzimas capaces de hidrolizar uniones formadas entre azúcares hemicelulósicos (notablemente arabinosa) y lignina. Las ligninasas incluyen, pero no se limitan a, el siguiente grupo de enzimas: lignina peroxidasas (EC 1.11.14), manganeso peroxidasas (EC 1.11.1.13), lacasas (EC 1.10.3.2) y feruloil esterases (EC 3.1.1.73).

"Hexosiltransferasa" (2.4.1-) incluyen enzimas que son capaces de transferir grupos glicosilo, más específicamente grupos hexosilo. Además de la transferencia de un grupo glicosilo desde un donante que contiene glicosilo a otro compuesto que contiene glicosilo, el aceptor, las enzimas también pueden transferir el grupo glicosilo a agua como un aceptor. Esta reacción también se conoce como una reacción de hidrólisis, en lugar de una reacción de transferencia. Un ejemplo de una hexosiltransferasa que se puede usar en la invención es una  $\beta$ -glucanosiltransferasa. Esta enzima puede ser capaz de catalizar la degradación de (1,3)(1,4)glucano y/o celulosa y/o un producto de degradación de celulosa.

"Glucuronidasa" incluye enzimas que catalizan la hidrólisis de un glucuronósido, por ejemplo  $\beta$ -glucuronósido para dar un alcohol. Se han caracterizado muchas glucuronidasas y pueden ser adecuadas para el uso en la invención, por ejemplo  $\beta$ -glucuronidasa (EC 3.2.1.31), hialurono-glucuronidasa (EC 3.2.1.36), glucuronosil-disulfoglucosamina glucuronidasa (3.2.1.56), glicirrizinato  $\beta$ -glucuronidasa (3.2.1.128) o  $\alpha$ -D-glucuronidasa (EC 3.2.1.139).

Una composición de la invención puede comprender una expansina o una proteína similar a expansina, tal como una swollenina (véase Salheimo y cols., Eur. J. Biochem. 269, 4202-4211, 2002) o una proteína similar a swollenina.

Las expansinas están implicadas en el aflojamiento de la estructura de la pared celular durante el crecimiento de células vegetales. Se ha propuesto que las expansinas rompen la unión de hidrógeno entre celulosa y otros polisacáridos de la pared celular sin tener actividad hidrolítica. De este modo, se cree que permiten el deslizamiento de fibras de celulosa y el engrosamiento de la pared celular. La swollenina, una proteína similar a expansina, contienen un dominio de la familia del módulo de unión a carbohidrato N-terminal 1 (CBD) y un dominio similar a expansina C-terminal. Para los propósitos de esta invención, una proteína similar a expansina o una proteína similar a swollenina puede comprender uno o ambos de tales dominios y/o puede romper la estructura de las paredes celulares (tal como rompiendo la estructura de la celulosa), opcionalmente sin producir cantidades detectables de azúcares reductores.

Una composición de la invención puede comprender el producto polipeptídico de una proteína integrante de celulosa, escafoldina o una proteína similar a escafoldina, por ejemplo CipA o CipC de *Clostridium thermocellum* o *Clostridium cellulolyticum*, respectivamente.

Las escafoldinas y las proteínas integrantes de celulosa son subunidades integradoras multifuncionales que pueden organizar subunidades celolíticas en un complejo multienzimático. Esto se efectúa mediante la interacción de dos clases complementarias de dominio, es decir un dominio de cohesión sobre escafoldina y un dominio de dockerina sobre cada unidad enzimática. La unidad de escafoldina también soporta un módulo de unión a celulosa (CBM) que media en la ligación del celulosoma a su sustrato. Una escafoldina o proteína integradora de celulosa para los propósitos de esta invención puede comprender uno o ambos de tales dominios.

Una composición de la invención puede comprender una proteína inducida por celulosa o una proteína moduladora, por ejemplo como la codificada por el gen *cip1* o *cip2* o genes similares procedentes de *Trichoderma reesei* / *Hypocrea jecorina* (véase Foreman y cols., J. Biol. Chem. 278(34), 31988-31997, 2003). El producto polipeptídico de estos genes son proteínas bimodulares, que contienen un módulo que se une a celulosa y un dominio cuya función o actividad no se pueden relacionar con familias de glicosil hidrolasa conocidas. Sin embargo, la presencia de un módulo que se une a celulosa y la coregulación de la expresión de estos genes con componentes de celulasas indica actividades previamente no reconocidas con un papel potencial en la degradación de biomasa.

Una composición de la invención puede estar compuesta de un miembro de cada una de las clases de los polipéptidos mencionados anteriormente, varios miembros de una clase de polipéptido, o cualquier combinación de estas clases de polipéptidos.

Una composición de la invención puede estar compuesta por polipéptidos, por ejemplo enzimas, de (1) proveedores comerciales; (2) genes clonados que expresan polipéptidos, por ejemplo enzimas; (3) caldo complejo (tal como el resultante del crecimiento de una cepa microbiana en un medio, en donde las cepas secretan proteínas y enzimas al medio; (4) lisados celulares de cepas desarrolladas como en (3); y/o (5) material vegetal que expresa polipéptidos, por ejemplo enzimas. Diferentes polipéptidos, por ejemplo enzimas en una composición de la invención, se pueden obtener de diferentes fuentes.

## Uso de los polipéptidos

Los polipéptidos y las composiciones polipeptídicas según la invención se pueden usar en muchas aplicaciones diferentes. A modo de ejemplo, se pueden usar para producir azúcares fermentables. Los azúcares fermentables se pueden convertir a continuación, como parte de un procedimiento para biocombustible, en biogas o etanol, butanol, isobutanol, 2-butanol u otras sustancias adecuadas. Alternativamente, los polipéptidos y sus composiciones se pueden usar como enzima, a modo de ejemplo en la producción de productos alimenticios, en composiciones detergentes, en la industria del papel y la pasta papelera y en formulaciones antibacterianas, en productos farmacéuticos tales como pastillas para la garganta, pastas de dientes y colutorios. Algunos de los usos se ilustrarán con más detalle posteriormente.

En los usos y los métodos descritos posteriormente, los componentes de las composiciones descritas anteriormente se pueden proporcionar concomitantemente (es decir como una sola composición de por sí) o separadamente o secuencialmente.

La invención también se refiere al uso del polipéptido potenciador de celulasa según la invención y composiciones que comprenden tan enzima en procedimientos industriales.

A pesar de la experiencia a largo plazo obtenida con estos procedimientos, el polipéptido potenciador de celulasa según la invención puede presentar un número de ventajas significativas sobre enzimas usadas normalmente. Dependiendo de la aplicación específica, estas ventajas pueden incluir aspectos tales como menores costes de producción, especificidad superior hacia el sustrato, antigenicidad reducida, menos actividades secundarias no deseables, rendimientos superiores cuando se producen en un microorganismo adecuado, intervalos de pH y temperatura más adecuados, falta de inhibición por productos derivados de lignina hidrófobos o menos inhibición de productos, o, en el caso de la industria alimentaria, un sabor o una textura mejores de un producto final así como calidad alimentaria y aspectos de conformidad con la ley judía.

En principio, un polipéptido o una composición potenciadores de celulasa de la invención se puede usar en cualquier procedimiento que requiera el tratamiento de un material que comprende polisacárido. Así, un polipéptido o una composición de la invención se puede usar en el tratamiento de un material polisacárico. En la presente, un material polisacárico es un material que comprende o consiste esencialmente en uno o, más típicamente, más de un polisacárido.

Típicamente, plantas y material derivado de las mismas comprende cantidades significativas de material polisacárico no amiláceo. Según esto, un polipéptido de la invención se puede usar en el tratamiento de un material vegetal o fúngico o un material derivado del mismo.

## Lignocelulosa

Un importante componente del material polisacárico vegetal no amiláceo es la lignocelulosa (también denominada en la presente biomasa lignocelulolítica). La lignocelulosa es un material vegetal que comprende celulosa y hemicelulosa y lignina. Los polímeros glucídicos (celulosa y hemicelulosas) están fuertemente unidos a la lignina por uniones de hidrógeno y covalentes. Según esto, un polipéptido de la invención se puede usar en el tratamiento de material lignocelulolítico. En la presente, el material lignocelulolítico es un material que comprende o consiste esencialmente en lignocelulosa. Así, en un método de la invención para el tratamiento de un polisacárido no amiláceo, el polisacárido no amiláceo puede ser un material/biomasa lignocelulósicos.

Según esto, la invención proporciona un método para tratar un sustrato que comprende polisacárido no amiláceo en el que el tratamiento comprende la degradación y/o hidrólisis y/o modificación de celulosa y/o hemicelulosa y/o una sustancia péctica.

Degradación en este contexto indica que el tratamiento da como resultado la generación de productos de hidrólisis de celulosa y/o hemicelulosa y/o una sustancia péctica, es decir sacáridos de longitud más corta están presentes como resultado del tratamiento que los que están presentes en un polisacárido no amiláceo no tratado similar. Así, la degradación en este contexto puede dar como resultado la liberación de oligosacáridos y/o monómeros sacáricos.

Todas las plantas contienen polisacárido no amiláceo como virtualmente todos los materiales polisacáricos derivados de plantas. Según esto, en un método de la invención para el tratamiento de un sustrato que comprende un polisacárido no amiláceo, dicho sustrato se puede proveer en la forma de una planta o un material derivado de planta o un material que comprende una planta o material derivado de planta, por ejemplo una pulpa vegetal, un extracto vegetal, un producto alimenticio o ingrediente para el mismos, una tela, un producto textil o una prenda de.

Biomasa lignocelulolítica adecuada para el uso en la invención incluye biomasa y puede incluir biomasa virgen y/o biomasa no virgen tal como biomasa agrícola, productos orgánicos comerciales, desechos de construcción y

demolición, desechos sólidos municipales, papel residual y desechos de jardín. Formas comunes de biomasa incluyen árboles, arbustos y hierbas, trigo, paja de trigo, bagazo de caña de azúcar, maíz, cáscaras de maíz, mazorcas de maíz, grano de maíz incluyendo fibra procedentes de los granos, productos y subproductos de la molienda de granos tales como maíz, trigo y cebada (incluyendo molienda en húmedo y molienda en seco) a menudo llamados "salvado o fibra" así como desechos sólidos municipales, papel residual y desechos de jardín. La biomasa también puede ser, pero no se limita a, material herbáceo, residuos agrícolas, residuos forestales, desechos sólidos municipales, papel residual y residuos de fábricas de papel y pasta papelera. "Biomasa agrícola" incluye ramas, matorrales, cañas, maíz y cáscaras de maíz, cultivos energéticos, bosques, frutos, flores, granos, hierbas, cultivos herbáceos, hojas, corteza, agujas, troncos, raíces, pimpollos, cultivos leñosos de rotación breve, arbustos, pastos varilla, árboles, hortalizas, cáscaras de fruta, vides, pulpa de remolacha azucarera, harinillas de trigo, cáscaras de avena y maderas duras y blandas (sin incluir maderas con materiales nocivos). Además, la biomasa agrícola incluye materiales residuales orgánicos generados de procedimientos agrícolas incluyendo actividades de labranza y forestales, específicamente desechos de madera de bosques. La biomasa agrícola puede ser cualquiera de las mencionadas singularmente o en cualquier combinación de las mismas. Ejemplos adicionales de biomasa adecuada son preparaciones de huertos, un chaparral, desechos de molienda, desechos de madera urbanos, desechos municipales, desechos de maderos, aclaramientos forestales, cultivos leñosos de rotación breve, desechos industriales, paja de trigo, paja de avena, paja de arroz, paja de cebada, paja de centeno, paja de lino, cáscaras de soja, cáscaras de arroz, paja de arroz, piensos de gluten de trigo, cáscaras de avena, caña de azúcar, harina de maíz, tallos de maíz, mazorcas de maíz, cáscaras de maíz, bromo, maicillo oriental, cola de zorra; pulpa de remolacha azucarera, pulpa de cítricos, cáscaras de semillas, desechos animales celulósicos, cortes de césped, algodón, algas marinas, árboles, arbustos, hierbas, trigo, paja de trigo, bagazo de caña de azúcar, maíz, cáscaras de maíz, mazorcas de maíz, grano de maíz, fibra procedente de granos, productos y subproductos procedentes de molindas en húmedo o en seco de granos, desechos sólidos municipales, papel residual, desechos de jardín, material herbáceo, residuos agrícolas, residuos forestales, desechos sólidos municipales, papel residual, pasta papelera, residuos de fábricas de papel, ramas, matorrales, cañas, cáscaras de maíz, un cultivo energético, un bosque, una fruta, una flor, un grano, una hierba, un cultivo herbáceo, una hoja, corteza, una aguja, un tronco, una raíz, un pimpollo, un arbusto, pasto varilla, un árbol, una hortaliza, cáscara de fruta, una vid, pulpa de remolacha azucarera, harinillas de trigo, cáscaras de avena, madera dura o blanda, material residual orgánico generado a partir de un procedimiento agrícola, desechos de madera de bosques, o una combinación de cualesquiera dos o más de los mismos.

Aparte de la biomasa virgen o materias primas ya procesadas en las industrias de los alimentos y los piensos o el papel y la pasta papelera, la biomasa/materia prima se puede pretratar adicionalmente con calor, modificación mecánica y/o química o cualquier combinación de tales métodos a fin de potenciar la degradación enzimática.

### 35 Pretratamiento

Antes del tratamiento enzimático, la materia prima se puede pretratar opcionalmente con calor, modificación mecánica y/o química o cualquier combinación de estos métodos a fin de potenciar la accesibilidad del sustrato a la hidrólisis enzimática y/o hidrolizar la hemicelulosa y/o solubilizar la hemicelulosa y/o celulosa y/o lignina, de cualquier modo conocido en la técnica. El pretratamiento puede comprender exponer al material lignocelulósico a agua (caliente), vapor de agua (explosión de vapor de agua), un ácido, una base, un disolvente, calor, un peróxido, ozono, corte mecánico, trituración, molienda o despresurización rápida, o una combinación de dos o más de los mismos. Un pretratamiento químico a menudo se combina con pretratamiento térmico, p. ej. entre 150-220°C durante de 1 a 30 minutos.

### Presacarificación

Después de la etapa de pretratamiento, se puede utilizar una etapa de licuefacción/hidrólisis o presacarificación que implica la incubación con una enzima o mezcla de enzimas. La etapa de presacarificación se puede realizar a muchas temperaturas diferentes pero se prefiere que la etapa de presacarificación se produzca a la temperatura más adecuada para la mezcla de enzimas que se aplique, o el óptimo enzimático predicho de las enzimas que se van a aplicar. La temperatura de la etapa de presacarificación puede variar de aproximadamente 10°C a aproximadamente 95°C, de aproximadamente 20°C a aproximadamente 85°C, de aproximadamente 30°C a aproximadamente 70°C, de aproximadamente 40°C a aproximadamente 60°C, de aproximadamente 37°C a aproximadamente 50°C, preferiblemente de aproximadamente 37°C a aproximadamente 80°C, más preferiblemente aproximadamente 60-70°C e incluso más preferiblemente alrededor de 65°C. El pH de la mezcla de presacarificación puede variar de aproximadamente 2,0 a aproximadamente 10,0, pero preferiblemente es de aproximadamente 3,0 a aproximadamente 7,0, más preferiblemente de aproximadamente 4,0 a aproximadamente 6,0, aún más preferiblemente de aproximadamente 4,0 a aproximadamente 5,0. De nuevo, el pH se puede ajustar para maximizar la actividad enzimática y se puede ajustar con la adición de la enzima.

La reacción de la etapa de licuefacción/hidrólisis o presacarificación se puede producir de varios minutos a varias horas, tal como de aproximadamente 1 hora a aproximadamente 120 horas, preferiblemente de aproximadamente 2 horas a aproximadamente 48 horas, más preferiblemente de aproximadamente 2 a aproximadamente 24 horas, lo más preferiblemente durante de aproximadamente 2 a aproximadamente 6 horas. El tratamiento con celulasa se puede producir de varios minutos a varias horas, tal como de aproximadamente 6 horas a aproximadamente 120 horas, preferiblemente de aproximadamente 12 horas a aproximadamente 72 horas, más preferiblemente aproximadamente de 24 a 48 horas.

#### Sacarificación

La invención proporciona un método para producir un azúcar a partir de un material lignocelulósico, método que comprende poner en contacto un polipéptido de la invención o una composición de la invención con el material lignocelulósico.

Este método permite que se generen azúcares libres (monómeros) y/o oligosacáridos a partir de biomasa lignocelulósica. Estos métodos implican convertir biomasa lignocelulósica en azúcares libres y oligosacáridos pequeños con un polipéptido o una composición de la invención.

El procedimiento de convertir un carbohidrato complejo tal como lignocelulosa en azúcares permite preferiblemente la conversión en azúcares fermentables. Este procedimiento se puede denominar "sacarificación". Según esto, un método de la invención puede dar como resultado la liberación de uno o más azúcares de hexosa y/o pentosa, tales como una o más de glucosa, xilosa, arabinosa, galactosa, ácido galacturónico, ácido glucurónico, manosa, ramnosa, ribosa y fructosa.

Según esto, otro aspecto de la invención incluye métodos que utilizan el polipéptido o la composición de la invención descritos anteriormente junto con enzimas o tratamientos físicos tales como temperatura y pH para convertir la biomasa vegetal lignocelulósica en azúcares y oligosacáridos.

Aunque la composición se ha analizado como una sola mezcla, se sabe que las enzimas se pueden añadir secuencialmente, donde la temperatura, el pH y otras condiciones se pueden alterar para incrementar la actividad de cada enzima individual. Alternativamente, se pueden determinar un pH y una temperatura óptimos para la mezcla de enzimas.

Las enzimas se hacen reaccionar con un sustrato bajo condiciones apropiadas. Por ejemplo, las enzimas se pueden incubar a aproximadamente 25°C, aproximadamente 30°C, aproximadamente 35°C, aproximadamente 37°C, aproximadamente 40°C, aproximadamente 45°C, aproximadamente 50°C, aproximadamente 55°C, aproximadamente 60°C, aproximadamente 65°C, aproximadamente 70°C, aproximadamente 75°C, aproximadamente 80°C, aproximadamente 85°C, aproximadamente 90°C o más. Esto es, se pueden incubar a una temperatura de aproximadamente 20°C a aproximadamente 95°C, por ejemplo en tampones de fuerza iónica de baja a media y/o desde pH bajo a neutro. Por "fuerza iónica media" se entiende que el tampón tiene una concentración iónica de aproximadamente 200 milimolar (mM) o menos para cualquier componente iónico simple. El pH puede variar de aproximadamente pH 2,5, aproximadamente pH 3,0, aproximadamente pH 3,5, aproximadamente pH 4,0, aproximadamente pH 4,5, aproximadamente pH 5, aproximadamente pH 5,5, aproximadamente pH 6, aproximadamente pH 6,5, aproximadamente pH 7, aproximadamente pH 7,5, aproximadamente pH 8,0, a aproximadamente pH 8,5. Generalmente, el intervalo de pH será de aproximadamente pH 3,0 a aproximadamente pH 7. Para la producción de etanol se prefiere un medio ácido, p. ej. pH=4, mientras que para la producción de biogas es deseable un pH neutro, p. ej. pH=7. La incubación de combinaciones de enzimas bajo estas condiciones da como resultado la emisión o liberación de cantidades sustanciales del azúcar desde la lignocelulosa. Por cantidad sustancial se entiende al menos 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95% o más de azúcar disponible.

Los polipéptidos, tales como enzimas, pueden producirse exógenamente en microorganismos, levaduras, hongos, bacterias o plantas, a continuación aislarse y añadirse, por ejemplo, a materia prima lignocelulósica. Alternativamente, las enzimas se producen, pero no se aíslan, y se pueden añadir caldo de fermentación de la masa celular en bruto, o material vegetal (tal como harina de maíz) y similares a, por ejemplo, la materia prima. Alternativamente la masa celular en bruto o el medio de producción de enzimas o el material vegetal pueden tratarse para evitar un crecimiento microbiano adicional (por ejemplo, mediante calentamiento o adición de agentes antimicrobianos), a continuación añadirse a, por ejemplo, una materia prima. Estas mezclas de enzimas en bruto pueden incluir pueden incluir el organismo que produce la enzima. Alternativamente, la enzima puede ser producida en una fermentación que usa materia prima (tal como harina de maíz) para proporcionar nutrición a un organismo que produce una enzima o enzimas. De este modo, plantas que producen las enzimas pueden servir ellas mismas como una materia prima lignocelulósica y ser añadidas a materia prima lignocelulósica.

#### Fermentación de azúcares

5 Los azúcares fermentables se pueden convertir en productos de fermentación útiles con valor añadido, ejemplos no limitativos de los cuales incluyen aminoácidos, vitaminas, productos farmacéuticos, suplementos para piensos, productos químicos especializados, materias primas químicas, plásticos, disolventes, combustibles, u otros polímeros orgánicos, ácido láctico y etanol, incluyendo etanol combustible. En particular, los azúcares se pueden usar como materias primas para fermentación en productos químicos, plásticos, tales como, a modo de ejemplo, ácido succínico y (bio)combustibles, incluyendo los combustibles líquidos sintéticos etanol, metanol, butanol y biogas.

10 A modo de ejemplo, en el método de la invención, una enzima o combinación de enzimas actúa sobre un sustrato lignocelulósico o biomasa vegetal, sirviendo como la materia prima, a fin de convertir este sustrato complejo en azúcares simples y oligosacáridos para la producción de etanol u otros productos de fermentación útiles.

15 Los azúcares liberados de la biomasa se pueden convertir en productos de fermentación útiles tales como uno de los que incluyen, pero no limitados a, aminoácidos, vitaminas, productos farmacéuticos, suplementos para piensos, productos químicos especializados, materias primas químicas, plásticos y etanol, incluyendo etanol combustible.

Según esto, la invención proporciona un método para la preparación de un producto de fermentación, método que comprende:

a. degradar lignocelulosa usando un método como en descrito en la presente; y

20 b. fermentación del material resultante,

para preparar de ese modo un producto de fermentación.

25 La fermentación se puede llevar a cabo bajo condiciones aerobias o anaerobias. Preferiblemente, el procedimiento se lleva a cabo bajo condiciones microaerófilas o de oxígeno limitado.

30 Un procedimiento de fermentación anaerobia se define en la presente como un procedimiento de fermentación realizado en ausencia de oxígeno o en el que sustancialmente no se consuma oxígeno, preferiblemente aproximadamente 5 o menos, aproximadamente 2,5 o menos o aproximadamente 1 mmol/l/h o menos, y en donde las moléculas orgánicas sirven de donante de electrones y de aceptores de electrones.

35 Un procedimiento de fermentación con oxígeno limitado es un procedimiento en el que el consumo de oxígeno está limitado por la transferencia de oxígeno del gas al líquido. El grado de limitación de oxígeno está determinado por la cantidad y la composición del flujo de gas entrante así como las propiedades de mezcla/transferencia de masas reales del equipo de fermentación usado. Preferiblemente, en un procedimiento bajo condiciones con oxígeno limitado, la velocidad de consumo de oxígeno es al menos aproximadamente 5,5, más preferiblemente al menos aproximadamente 6 y aún más preferiblemente al menos aproximadamente 7 mmol/l/h.

Un método para la preparación de un producto de fermentación puede comprender opcionalmente la recuperación del producto de fermentación.

40 SSF

La fermentación u la sacarificación también se pueden ejecutar en modo de sacarificación y fermentación simultáneas (SSF). Una de las ventajas de este modo es la reducción de la inhibición sacárica sobre la hidrólisis enzimática (la inhibición sacárica sobre celulasas es descrita por Caminal B&B Vol XXVII Pp 1282-1290).

Productos de fermentación

45 Productos de fermentación que se pueden producir según la invención incluyen aminoácidos, vitaminas, productos farmacéuticos, suplementos para piensos, productos químicos especializados, materias primas químicas, plásticos, disolventes, combustibles, u otros polímeros orgánicos, ácido láctico, y etanol, incluyendo etanol (entendiéndose que el término "etanol" incluye alcohol etílico o mezclas de alcohol etílico y agua).

50 Productos de valor añadido específicos que se pueden producir mediante los métodos de la invención incluyen, pero no se limitan a, biocombustibles (incluyendo etanol y butanol y un biogas); ácido láctico; un plástico; un producto químico especializado; un ácido orgánico, incluyendo ácido cítrico, ácido succínico, ácido fumárico, ácido itacónico y ácido maleico; ácido 3-hidoxi-propiónico, ácido acrílico; ácido acético; 1,3-propano-diol; etileno, glicerol; un disolvente; un suplemento para piensos; un producto farmacéutico, tal como un antibiótico de  $\beta$ -lactama o una cefalosporina; vitaminas; un aminoácido, tal como lisina, metionina, triptófano, treonina y ácido aspártico; una

55

enzima industrial, tal como una proteasa, una celulasa, una amilasa, una glucanasa, una lactasa, una lipasa, una liasa, una oxidorreductasa, una transferasa o una xilanasas; y una materia prima química.

#### Biogás

5 La invención también proporciona el uso de un polipéptido o una composición como los descritos en la presente en un método para la preparación de biogas. Biogas se refiere típicamente a un gas producido mediante la descomposición biológica de materia orgánica, por ejemplo un material que contiene carbohidratos no amiláceos, en ausencia de oxígeno. El biogas se origina a partir de material biogénico y es un tipo de biocombustible. Un tipo de biogas se produce mediante digestión o fermentación anaerobia de materiales biodegradables tales como biomasa, estiércol o aguas residuales, desechos municipales y cultivos energéticos. Este tipo de biogas está comprendido  
10 principalmente por metano y dióxido de carbono. El metano gaseoso se puede quemar u oxidar con oxígeno. El aire contienen 21% de oxígeno. Esta liberación de energía permite que el biogás se use como un combustible. El biogás se puede usar como un combustible de bajo coste en cualquier país con cualquier propósito de calentamiento, tal como cocinar. También se puede utilizar en instalaciones modernas de gestión de residuos en las que se puede usar para poner en marcha cualquier tipo de motor térmico, para generar energía bien mecánica o bien eléctrica.

15 La primera etapa en la producción microbiana de biogás consiste en la degradación enzimática de polímeros y sustratos complejos (por ejemplo un carbohidrato no amiláceo). Según esto, la invención proporciona un método para la preparación de un biogás en el que un sustrato que comprende un carbohidrato no amiláceo se pone en contacto con un polipéptido o una composición de la invención, para dar de ese modo un material fermentable que  
20 puede ser convertido en un biogás por un organismo tal como un microorganismo. En este método, un polipéptido de la invención se puede proporcionar por medio de un organismo, por ejemplo un microorganismo que expresa tal polipéptido.

#### Uso de enzimas en productos alimenticios

25 Los polipéptidos y las composiciones de la invención se pueden usar en un método para procesar material vegetal para degradar o modificar los constituyentes de celulosa o hemicelulosa o sustancia péctica de las paredes celulares del material vegetal o fúngico. Estos métodos pueden ser útiles en la preparación de un producto alimenticio. Según esto, la invención proporciona un método para preparar un producto alimenticio, método que comprende incorporar un polipéptido o una composición de la invención durante la preparación del producto alimenticio.

30 La invención también proporciona un método para procesar un material vegetal, método que comprende poner en contacto el material vegetal con un polipéptido o una composición de la invención para degradar o modificar la celulosa en el material (vegetal). Preferiblemente, el material vegetal es una pulpa vegetal o un extracto vegetal, tal como zumos.

35 Los materiales vegetales y que contienen celulosa/hemicelulosa/sustancia péctica incluyen pulpa vegetal, partes de plantas y extractos vegetales. En el contexto de esta invención, un extracto procedente de un material vegetal es cualquier sustancia que se pueda derivar de material vegetal mediante extracción (mecánica y/o química), procesamiento o mediante otras técnicas de separación. El extracto puede ser zumo, néctar, base o concentrados hechos de los mismos. El material vegetal puede comprender o derivarse de hortalizas, p. ej., zanahorias, apio, cebollas, legumbres o plantas leguminosas (soja, habas de soja, guisantes) o fruta, p. ej., pomos o frutas con pepitas  
40 (manzanas, peras, membrillo, etc.), uvas, tomates, cítricos (naranja, limón, lima, mandarina), melones, ciruelas, cerezas, grosellas negras, grosellas rojas, frambuesas, fresas, arándanos, piña y otras frutas tropicales, árboles y partes de los mismos (p. ej. polen, de pinos), o cereales (avena, cebada, trigo, maíz, arroz). El material (que va a ser hidrolizado) también pueden ser residuos agrícolas, tal como pulpa de remolacha azucarera, mazorcas de maíz,  
45 paja de trigo, cáscaras de nuez (trituradas) o materiales reciclables, p. ej. papel (residual).

Los polipéptidos de la invención se pueden usar así para tratar material vegetal incluyendo pulpa vegetal y extractos vegetales. También se pueden usar para tratar productos alimenticios líquidos o sólidos o ingredientes de productos alimenticios comestibles, o se pueden usar en la extracción de aceites vegetales, almidón o como un espesante en  
50 alimentos.

Típicamente, los polipéptidos de la invención se usan como una composición/preparación enzimática según se describe anteriormente. La composición se añadirá generalmente a pulpa vegetal obtenible, por ejemplo, mediante procesamiento mecánico tal como triturando o moliendo material vegetal. La incubación de la composición con la planta se llevará a cabo típicamente durante un tiempo de 10 minutos a 5 horas, tal como de 30 minutos a 2 horas, preferiblemente durante aproximadamente 1 hora. La temperatura de procesamiento es preferiblemente de aproximadamente 10°C a aproximadamente 55°C, p. ej. de aproximadamente 15°C a aproximadamente 25°C, óptimamente aproximadamente 20°C y se pueden usar de aproximadamente 10 g a aproximadamente 300 g,  
55

preferiblemente de aproximadamente 30 g a aproximadamente 70 g, óptimamente aproximadamente 50 g de enzima por tonelada de material que se va a tratar.

5 Toda la enzima o las enzimas o sus composiciones usadas se pueden añadir secuencialmente o al mismo tiempo a la pulpa vegetal. Dependiendo de la composición de la preparación enzimática, el material vegetal en primer lugar se puede macerar (p. ej. hasta un puré) o licuar. Usando los polipéptidos de la invención, se pueden mejorar parámetros de procesamiento tales como el rendimiento de la extracción, la viscosidad del extracto y/o la calidad del extracto.

10 Alternativamente, o además de lo anterior, un polipéptido de la invención se puede añadir a zumo en crudo obtenido del prensado o la licuefacción de la pulpa vegetal. El tratamiento del zumo en crudo se llevará a cabo de un modo similar a la pulpa vegetal con respecto a la dosificación, la temperatura y el tiempo de espera. De nuevo, se pueden incluir otras enzimas tales como las analizadas previamente. Condiciones de incubación típicas son como las descritas en el párrafo previo.

15 Una vez que el zumo en crudo se ha incubado con los polipéptidos de la invención, a continuación el zumo se centrifuga o se (ultra)filtra para producir el producto final.

20 Después del tratamiento con el polipéptido de la invención, el producto (final) se puede tratar térmicamente, p. ej. a aproximadamente 100°C durante un tiempo de aproximadamente 1 minuto a aproximadamente 1 hora, bajo condiciones para inactivar parcialmente o totalmente el polipéptido o los polipéptidos de la invención.

25 Una composición que contiene un polipéptido de la invención también se puede usar durante la preparación de purés de frutas u hortalizas.

En el horneado, el polipéptido puede mejorar la estructura de la masa, modificar su adherencia o flexibilidad, mejorar el volumen de la barra y/o la estructura de la miga o impartir mejores características de textura tales como rotura, desmenuzamiento o calidad de la miga.

30 La presente invención se refiere así a métodos para preparar una masa o un producto alimenticio basado en cereales que comprender incorporar en la masa un polipéptido o una composición de la presente invención. Esto puede mejorar una más propiedades de la masa o el producto alimenticio basado en cereales obtenido a partir de la masa con relación a una masa o un producto alimenticio basado en cereales en los que no se incorpore el polipéptido.

35 La preparación del producto alimenticio basado en cereales según la invención puede comprender además etapas conocidas en la técnica tales como ebullición, secado, fritura, tratamiento al vapor u horneado de la masa obtenida.

40 Productos que se elaboran a partir de una masa que se hierva son, por ejemplo, tallarines hervidos, empanadillas, productos que se elaboran a partir de masa frita son, por ejemplo, rosquillas, buñuelos, tallarines fritos, productos que se elaboran a partir de masa tratada al vapor son, por ejemplo bollos al vapor o tallarines al vapor, ejemplos de productos elaborados a partir de masa desecada son pasta y tallarines desecados y ejemplos de productos elaborados a partir de masa horneada son pan, galletas, torta.

45 El término "propiedad mejorada" se define en la presente como cualquier propiedad de una masa y/o un producto obtenido a partir de la masa, particularmente un producto alimenticio basado en cereales, que se mejore mediante la acción del polipéptido según la invención con relación a una masa o producto en los que no se incorpore el polipéptido según la invención. La propiedad mejorada puede incluir, pero no se limita a, resistencia incrementada de la masa, elasticidad incrementada de la masa, estabilidad incrementada de la masa, capacidad de trabajado a máquina mejorada de la masa, impermeabilización mejorada de la masa, adherencia reducida de la masa, extensibilidad mejorada de la masa, volumen incrementado del producto alimenticio basado en cereales, ampollamiento reducido producto alimenticio basado en cereales, estructura de la miga mejorada del producto horneado, suavidad mejorada del producto alimenticio basado en cereales, sabor mejorado del producto alimenticio basado en cereales, antienranciamiento mejorado del producto alimenticio basado en cereales. Propiedades mejoradas relacionadas con el tipo de pasta y tallarín de los productos basados en cereales son, por ejemplo, firmeza mejorada, adherencia reducida, cohesividad mejorada y pérdida por cocción reducida.

60 La propiedad mejorada se puede determinar por comparación con una masa y/o un producto alimenticio basado en cereales preparados con y sin la adición de un polipéptido de la presente invención. Las cualidades organolépticas se pueden evaluar usando procedimientos bien establecidos en la industria, y pueden incluir, por ejemplo, el uso de un grupo de degustadores entrenados.

65 El término "masa" se define en la presente como una mezcla de harina de cereal y otros ingredientes suficientemente firme para amasar o tratar con rodillo. Ejemplos de cereales son trigo, centeno, maíz, cebada, arroz, sémola, trigo sarraceno y avena. Se pretende ahora y posteriormente que el trigo abarque todas las especies

conocidas del género *Triticum*, por ejemplo *aestivum*, *durum* y/o *spelta*. Ejemplos de otros ingredientes adecuados son: el polipéptido potenciador de celulasa según la presente invención, enzimas adicionales, aditivos químicos y/o adyuvantes de procesamiento. La masa puede ser fresca, congelada, preparada o preheada. La preparación de una masa a partir de los ingredientes descritos anteriormente es muy conocida en la técnica y comprende la mezcla de dichos ingredientes y adyuvantes de procesamiento y una o más etapas de moldeo y opcionalmente fermentación. La preparación de masa congelada es descrita por Kulp y Lorenz en *Frozen and Refrigerated Doughs and Batters*.

El término "producto alimenticio basado en cereales" se define en la presente como cualquier producto preparado a partir de una masa, de carácter bien blando o bien crujiente. Ejemplos de productos alimenticios basados en cereales, ya sea de un tipo blanco, claro u oscuro, que se pueden producir ventajosamente mediante la presente invención son pan (en particular pan blanco, integral o de centeno), típicamente en la forma de barras o rollos, pan francés de tipo baguette, pasta, tallarines, rosquillas, bagels, torta, pan de pita, tortillas, tacos, tortas, panqueques, bizcochos, galletas, masas quebradas, pan al vapor y pan crujiente, y similares.

El término "producto horneado" se define en la presente como cualquier producto alimenticio basado en cereales preparado al hornear la masa.

Los polisacáridos no amiláceos (NSP) pueden incrementar la viscosidad de la masa digerida, lo que, a su vez, puede disminuir la disponibilidad de nutrientes y el rendimiento animal. El uso del polipéptido potenciador de celulasa de la presente invención puede mejorar la utilización de fósforo así como minerales catiónicos y polipéptido durante la digestión del animal.

Añadir nutrientes específicos al pienso mejora la digestión del animal y de ese modo reduce los costes del pienso. Se está usando actualmente una gama de aditivos para piensos y se están desarrollando continuamente nuevos conceptos. El uso de enzimas específicas como enzimas de degradación de carbohidratos no amiláceos podría descomponer la fibra liberando energía así como incrementar la digeribilidad de proteína debido a una mejor accesibilidad de la proteína cuando la fibra se descompone. De este modo, se podría disminuir el coste del pienso así como también se podrían reducir los niveles de proteína en el pienso.

Los polisacáridos no amiláceos (NSP) también están presente virtualmente en todos los ingredientes de piensos de origen vegetal. Los NSP se utilizan escasamente y, cuando se solubilizan, pueden ejercer efectos adversos sobre la digestión. Las enzimas exógenas pueden contribuir a una mejor utilización de estos NSP y como consecuencia reducir cualesquiera efectos antinutricionales. Las enzimas de degradación de carbohidratos no amiláceos de la presente invención se pueden usar con este propósito en dietas basadas en cereales para aves de corral y, en un grado menor, para cerdos y otras especies.

Un polipéptido/una enzima de degradación de carbohidratos no amiláceos de la invención (o una composición que comprende el polipéptido/la enzima de la invención) se puede usar en la industria de los detergentes, por ejemplo para la eliminación de la ropa de manchas basadas en carbohidratos. Una composición detergente puede comprender un polipéptido/una enzima de la invención y, además, una o más de una celulasa, una hemicelulasa, una pectinasa, una proteasa, una lipasa, una cutinasa, una amilasa o una carbohidrasa.

#### Uso de enzimas en composiciones detergentes

Una composición detergente que comprende un polipéptido o una composición de la invención puede estar en cualquier forma conveniente, por ejemplo una pasta, un gel, un polvo o un líquido. Un detergente líquido puede ser acuoso, conteniendo típicamente hasta aproximadamente 70% de agua y de aproximadamente 0 a aproximadamente 30% de disolvente orgánico o material no acuoso.

Esta composición detergente se puede formular, por ejemplo, como una composición detergente para lavar a mano o a máquina adecuada para el pretratamiento de tejidos manchados y una composición suavizante de tejidos añadida en el enjuague, o se puede formular como una composición detergente para uso en operaciones de limpieza de superficies duras domésticas generales, o se puede formular para operaciones de lavado de vajillas a mano o a máquina.

En general, las propiedades de la enzima deben ser compatibles con el detergente seleccionado (por ejemplo, óptimo de pH, compatibilidad con otros ingredientes enzimáticos y/o no enzimáticos, etc.) y la enzima o las enzimas deben estar presentes en una cantidad eficaz.

Una composición detergente puede comprender un tensioactivo, por ejemplo un tensioactivo aniónico o no iónico, un mejorador del detergente o un agente complejante, uno o más polímeros, un sistema blanqueador (por ejemplo una fuente de H<sub>2</sub>O<sub>2</sub>) o un estabilizante de enzimas. Una composición detergente también puede comprender cualquier otro ingrediente de detergentes convencional tal como, por ejemplo, un acondicionador incluyendo una arcilla, un

estimulante de la espuma, un supresor de jabonaduras, un agente anticorrosión, un agente de suspensión de manchas, un agente antiirredeposición de manchas, un tinte, un bactericida, un abrillantador óptico, un hidrótopo, un inhibidor del deslustre o un perfume.

Uso de enzimas en procesamiento de papel y pasta papelera

- 5 Un polipéptido o una composición de la presente invención se puede usar en la industria del papel y la pasta papelera, entre otras cosas en el procedimiento de blanqueo para mejorar el brillo de pastas papeleras blanqueadas, con lo que se puede reducir la cantidad de cloro usada en las fases de blanqueo, y para incrementar la libertad de las pastas papeleras en el procedimiento de reciclaje de papel (Eriksson, K. E. L., Wood Science and Technology 24 (1990):79-101; Paice, y cols., Biotechnol. and Bioeng. 32 (1988):235-239 y Pommier y cols., Tappi Journal (1989):187-191). Por otra parte, un polipéptido o una composición de la invención se puede usar para el tratamiento de pasta papelera lignocelulósica a fin de mejorar la capacidad de blanqueo de la misma. De ese modo, se puede reducir la cantidad de cloro necesaria para obtener un blanqueo satisfactorio de la pasta papelera.

- 15 Un polipéptido o una composición de la invención se puede usar en un método para reducir la velocidad a la que las telas que contienen celulosa se hacen duras o para reducir la dureza de telas que contienen celulosa, comprendiendo el método tratar telas que contienen celulosa con un polipéptido o una composición como los descritos anteriormente. La presente invención se refiere además a un método que proporciona aclaramiento del color de telas que contienen celulosa coloreadas, comprendiendo el método tratar telas que contienen celulosa coloreadas con un polipéptido o una composición como los descritos anteriormente, y un método para proporcionar una variación localizada en el color de telas que contienen celulosa coloreadas, comprendiendo el método tratar telas que contienen celulosa coloreadas con un polipéptido o una composición como los descritos anteriormente. Los métodos de la invención se pueden llevar a cabo al tratar telas que contienen celulosa durante el lavado. Sin embargo, si se desea, el tratamiento de las telas también se puede llevar a cabo durante el remojo o el enjuague o simplemente al añadir el polipéptido o la composición que se describe anteriormente a agua en la que están sumergidas las telas.

Otros usos de las enzimas

Además, un polipéptido o una composición de la presente invención también se puede usar en una formulación antibacteriana así como en productos farmacéuticos tales como pastillas para la garganta, pastas de dientes y un colutorio.

- 30 Los siguientes Ejemplos ilustran la invención:

### Ejemplos

Materiales y métodos

Procedimientos con ADN

- 35 Se llevaron a cabo procedimientos con ADN estándar según se describe en otras partes (Sambrook y cols., 1989, Molecular cloning: a laboratory manual, 2ª Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, Nueva York) a menos que se indique otra cosa. El ADN se amplificó usando la enzima a prueba de lectura enzima polimerasa Phision (Finnzymes). Las enzimas de restricción eran de Invitrogen o New England Biolabs.

Preparación de muestras de celulasa

- 40 Celulasas originarias de *Talaromyces emersonii* se expresaron en *Aspergillus niger*. Se produjeron filtrados concentrados de las enzimas según se describe en el documento WO2004/030468. Después de hacer crecer *Aspergillus niger* que contenía los plásmidos de expresión apropiados, se prepararon sobrenadantes libres de células mediante la centrifugación del caldo de fermentación a 5.000 x g durante 30 minutos a 4°C. Opcionalmente, el sobrenadante se puede ajustar hasta pH=5 con KOH 4 N y filtrarse estérilmente sobre un filtro de 2 µm (tapón de botella) con succión para retirar cualquier material fúngico. Además, los sobrenadantes se pueden filtrar adicionalmente sobre un filtro de microfibras de vidrio GF/A Whatman (150 mm Ø) para retirar cualquier sólido. Los sobrenadantes se ultrafiltraron, se concentraron y se almacenaron hasta el uso a 4°C o se congelaron a -20°C.

Método para la determinación de proteínas totales

El método era una combinación de precipitación de proteína usando ácido tricloroacético (TCA) para retirar sustancias molestas y permite la determinación de la concentración de proteína con la reacción colorimétrica del biuret. En la reacción del biuret, un ion cobre (II) se reduce hasta cobre (I), que forma un complejo con los nitrógenos y los carbonos de las uniones peptídicas en una solución alcalina. Un color violeta indica la presencia de proteínas. La intensidad del color, y de ahí la absorción a 546 nm, es directamente proporcional a la concentración de proteína, según la ley de Beer-Lambert. La estandarización se realizó usando BSA (albúmina de suero bovino) y el contenido de proteína se expresó en g de proteína como equivalente de BSA/l o mg de proteína como equivalente de BSA/ml. El contenido de proteína se calculó usando protocolos de cálculo estándar conocidos en la técnica, al representar la OD<sub>546</sub> frente a la concentración de muestras con concentración conocida, seguido por el cálculo de la concentración de las muestras desconocidas usando la ecuación generada a partir de la curva de calibración.

Preparación de sustrato de paja de trigo pretratado lavado

Se obtuvo paja de trigo pretratada en ácido diluido como se describe en Linde, M. y cols, Biomass and Bioenergy 32 (2008), 326-332 y se puede usar el equipo que se describe en Schell, D.J., Applied Biochemistry and Biotechnology (2003), vol. 105-108, pp 69-85. La paja de trigo pretratada se lavó con agua hasta que la solución con paja de trigo tenía pH 6,0 o superior. El agua de lavado se filtró.

Cribado con respecto a la hidrólisis de celulosa

Una solución de sustrato de materia seca (dm) de paja de trigo pretratada (PWS) ácida lavada al 2% se elabora en tampón de acetato sódico 50 mM, pH 4,5. Para la incubación, se usa una microplaca de 96 pocillos hondos. Se pipetea a cada pocillo 1 ml de sustrato.

Se prepararon mezclas de enzimas a diversas relaciones de las celulasas, y se añadieron a la solución de sustrato a una dosis de proteína fija por gramo de materia seca de sustrato.

La placa se incubó a 65°C durante 20 horas. Después de la incubación, la reacción enzimática se detiene al añadir 50 µl de una solución de hidróxido sódico 1 M. A continuación, la placa se centrifuga durante 15 min. a 3220 rcf y el sobrenadante se diluye hasta una concentración de glucosa entre 40 y 100 µM.

Se pipetea 50 µl de muestra diluida en una placa de PCR y se añaden 150 µl de reactivo BCA. El reactivo BCA se elabora recientemente al mezclar dos soluciones de reserva A y B 1:1 v/v. La solución A consistía en 54,3 g de Na<sub>2</sub>CO<sub>3</sub>, 24,2 de NaHCO<sub>3</sub> y 1,9 g de Na<sub>2</sub>BCA (Sigma D8284) por litro de agua (MQ). La solución B contenía 31,24 ml de solución al 4% de CuSO<sub>4</sub>.5H<sub>2</sub>O (Pierce 185 9078) y 1,26 g de L-lisina (Sigma L5501) por litro de agua (MQ). (Concentración final: 2,5 mM de BCA, 2,5 mM de Cu, 4,3 mM de L-lisina y 400 mM de carbonato). La placa se calienta hasta 80°C durante 60 minutos. Después de enfriar, se pipetea 150 µl en una segunda placa y la absorbancia se mide a 560 nm.

En cada placa se analiza el fondo del sustrato, al incubar el sustrato sin adición de enzimas. Para cada medida, se analiza un estándar de glucosa de 55 µM frente a tampón de acetato para calcular el coeficiente de extinción molar de la glucosa.

Las incubaciones se realizan usando un bloque de PCR termociclador Peltier (PTC200) y las medidas se realizan usando un lector de placas de microvaloración (TECAN Sunrise).

Se calculó la glucosa liberada del sustrato por la acción de las enzimas, como se da posteriormente:

$$* \text{Glucosa liberada (mmol/l)} = \frac{(A_s - A_{bl}) * D_{\text{ensayo}} * V_{\text{ensayo}} * D_{\text{incubados}}}{l * \epsilon * V_{\text{sustrato}}}$$

A<sub>s</sub> = absorbancia de la muestra a 560 nm

A<sub>bl</sub> = absorbancia del blanco del sustrato a 560 nm

D<sub>ensayo</sub> = dilución en el ensayo

V<sub>ensayo</sub> = volumen total del ensayo (µl)

D<sub>incubados</sub> = dilución de los incubados

l = longitud de recorrido (cm)

ε = coeficiente de extinción molar de la glucosa-BCA (mmol<sup>-1</sup>\*cm<sup>-1</sup>)

V<sub>sustrato</sub> = volumen total del sustrato (µl)

Este valor también se corrigió para el contenido de azúcares reductores/proteínas en la muestra de enzimas (analizada mediante el ensayo de BCA)

#### Hidrólisis ampliada de celulosa

- 5 Las incubaciones de combinaciones de enzimas sobre solución de sustrato de materia seca (dm) de paja de trigo pretratada (PWS) ácida lavada al 2% en tampón de acetato sódico 50 mM, pH 4,5, se realizaron a una escala de 10 ml. Las combinaciones de enzimas se añadieron a dosis de proteína fijas por gramo de materia seca de sustrato. Las muestras se recogieron a tiempo, hasta 72 horas de incubación a 65°C. Las reacciones se terminaron en un tiempo dado, al centrifugar el residuo, pipetear el sobrenadante y congelar las muestras hasta el análisis.
- 10 El análisis de la cantidad de glucosa liberada se realizó usando NMR de flujo. Los espectros de <sup>1</sup>H NMR se registraron en un sistema de NMR BrukerAVANCE II BEST que funcionaba a una frecuencia protónica de 500 MHz y una temperatura de la sonda de 27°C.

#### Ejemplo 1

##### 1.1. Construcción de plásmidos de expresión

- 15 La secuencia que tenía SEQ ID N°: 1 se clonó en el vector pGBTOP (Fig. 1) usando los sitios EcoRI y SnaBI, que comprenden la secuencia promotora y terminadora de glucoamilasa. La parte de E.coli se recuperó mediante digestión con NotI antes de la transformación de A. niger CBS 513.88.

##### 1.2. Transformación de A. niger

- 20 A. niger WT-1: Esta cepa de A. niger es CBS513.88 que comprende eliminaciones de los genes que codifican glucoamilasa (glaA), amilasa fúngica y amilasa ácida. A. niger WT 1 se construye al usar el enfoque "libre de gen marcador" que se describe en el documento EP 0 635 574 B1.

- 25 Las construcciones de expresión se cotransformaron hasta la cepa A. niger WT-1 según el método descrito por Tilburn, J. y cols. (1983) Gene 26, 205-221 y Kelly, J. & Hynes, M. (1985) EMBO J., 4, 475-479 con las siguientes modificaciones:

- 30 - Las esporas se germinan y se cultivan durante 16 horas a 30 grados Celsius en un matraz agitado colocado en un agitador giratorio a 300 rpm en medio mínimo de Aspergillus (100 ml). El medio mínimo de Aspergillus contiene por litro: 6 g de NaNO<sub>3</sub>, 0,52 g de KCl, 1,52 g de KH<sub>2</sub>PO<sub>4</sub>, 1,12 ml de KOH 4 M, 0,52 g de MgSO<sub>4</sub>·7H<sub>2</sub>O, 10 g de glucosa, 1 g de casaminoácidos, 22 mg de ZnSO<sub>4</sub>·7H<sub>2</sub>O, 11 mg de H<sub>3</sub>BO<sub>3</sub>, 5 mg de FeSO<sub>4</sub>·7H<sub>2</sub>O, 1,7 mg de CoCl<sub>2</sub>·6H<sub>2</sub>O, 1,6 mg de CuSO<sub>4</sub>·5H<sub>2</sub>O, 5 mg de MnCl<sub>2</sub>·2H<sub>2</sub>O, 1,5 mg de Na<sub>2</sub>MoO<sub>4</sub>·2H<sub>2</sub>O, 50 mg de EDTA, 2 mg de riboflavina, 2 mg de HCl de tiamina, 2 mg de nicotinamida, 1 mg de HCl de piridoxina, 0,2 mg de ácido pantoténico, 4 g de biotina, 10 ml de solución de penicilina (5000 UI/ml) y estreptomina (5000 UG/ml) (Gibco).

- Se usa Novozym 234™ (Novo Industries) en lugar de helicasa para la preparación de protoplastos;

- 35 - Después de la formación de protoplastos (60-90 minutos), se añadió tampón de KC (0,8 M de KCl, 9,5 mM de ácido cítrico, pH 6,2) hasta un volumen final de 45 ml, la suspensión de protoplastos se centrifuga durante 10 minutos a 3.000 rpm a 4 grados Celsius en un rotor de álabes oscilantes. Los protoplastos se resuspenden en 20 ml de tampón de KC y posteriormente se añaden 25 ml de tampón de STC (1,2 M de sorbitol, 10 mM de Tris-HCl pH 7,5, 50 mM de CaCl<sub>2</sub>). La suspensión de protoplastos se centrifuga durante 10 minutos a 3000 rpm a 4 grados Celsius en un rotor de álabes oscilantes, se lava en tampón de STC y se resuspende en tampón de STC a una concentración de 10E8 protoplastos/ml;
- 40

- Se añade a 200 microlitros de la suspensión de protoplastos el fragmento de ADN, disuelto en 10 microlitros de tampón de TE (10 mM de Tris-HCl pH 7,5, 0,1 mM de EDTA) y 100 microlitros de soluciones de PEG (20% de PEG 4000 (Merck), 0,8 M de sorbitol, 10 mM de Tris-HCl pH 7,5, 50 mM de CaCl<sub>2</sub>);

- 45 - Después de la incubación de la suspensión de protoplastos de ADN durante 10 minutos a temperatura ambiente, se añadieron lentamente 1,5 ml de solución de PEG (60% de PEG 4000 (Merck), 10 mM de Tris-HCl pH 7,5, 50 mM de CaCl<sub>2</sub>), con mezcladura repetida de los tubos. Después de la incubación durante 20 minutos a temperatura ambiente, las suspensiones se diluyen con 5 ml de sorbitol 1,2 M, se mezclaron mediante inversión y se centrifugaron durante 10 minutos a 4000 rpm a temperatura ambiente. Los protoplastos se resuspenden suavemente

en 1 ml de sorbitol 1,2 M y se siembran sobre medio de regeneración sólido selectivo que consiste en medio mínimo de *Aspergillus* sin riboflavina, HCl de tiamina, nicotinamida, piridoxina, ácido pantoténico, biotina, casaminoácidos y glucosa. En el caso de la selección de acetamida, el medio contiene 10 mM de acetamida como la única fuente de nitrógeno y 1 M de sacarosa como fuente osmótica y de C. Alternativamente, los protoplastos se siembran sobre PDA (agar de dextrosa de patata, Oxoid) complementado con 1-50 microgramos/ml de fleomicina y 1 M de sacarosa como osmótico. Las placas de regeneración se solidifican usando 2% de agar (agar N° 1, Oxoid L11). Después de la incubación durante 6-10 días a 30 grados Celsius, las conidiosporas de los transformantes se transfieren a placas que consisten en medio selectivo de *Aspergillus* (medio mínimo que contiene acetamida como única fuente de nitrógeno en el caso de la selección de acetamida o PDA complementado con 1-50 microgramos/ml de fleomicina en el caso de la selección de fleomicina) con 2% de glucosa y 1,5% de agarosa (Invitrogen) y se incuban durante 5-10 días a 30 grados Celsius. Se aislaron transformantes individuales y esta etapa de purificación selectiva se repite una vez tras lo cual los transformantes purificados se almacenaron.

Después de la transformación, los transformantes se seleccionaron sobre medios que comprendían acetamida como la única fuente de nitrógeno y las colonias se purificaron. Los números de copias se estimaron mediante PCR cuantitativa y se seleccionaron los transformantes de bajo y alto número de copias. Los transformantes de alto número de copias se cultivaron en matraces agitados en 100 ml de medio CSM-MES según se describe en el documento EP 635 574 a 34°C a 170 rpm en un agitador incubador usando un matraz agitador de 500 ml con deflectores. Después de 3 y 4 días de fermentación, las muestras de sobrenadante se recogieron para determinar la expresión mediante SDS-PAGE.

### 1.3 Contenido de proteína

Se analizó en contenido de proteína de los sobrenadantes ultrafiltrados y concentraron de las fermentaciones en matraz agitado de los transformantes que expresan polipéptido potenciador de celulasa (TEMER07589). Se determinó que el contenido de proteína era 24 mg de proteína como equivalentes de BSA/ml.

### Ejemplo 2

#### 2.1 Preparación de muestras de celulasa

Dos endoglucanasas procedentes de *Talaromyces emersonii*, como las descritas en la patente EP1621628, como CEA (SEQ 1) y CEB (SEQ 3) se prepararon como se describe en el Ejemplo 1 para TEMER07589. Adicionalmente, se prepararon de forma similar dos exoglucanasas, que son CBHI según se describe en la solicitud de patente EP09158739.4 y CBHII (según se describe en la solicitud de patente en tramitación junto con la presente DSM Caso 27829, presentada el mismo día que esta solicitud). Finalmente, una beta-glucosidasa procedente de *Talaromyces emersonii*, como se conoce de Murray y cols., Protein expression and Purification, 2004, 38, 248-257, también se sobreexpresó en *Aspergillus niger*. Los contenidos de proteína de estas muestras se determinaron y clasificaron de 20 a 60 mg de proteína como equivalentes de BSA/ml.

#### 2.2 Hidrólisis de celulosa mediante mezclas de enzimas en 20 h

Se prepararon combinaciones de 4 celulasas. Cada combinación consistía en 1 BG, 1 CBHI, 1 CBHII y 1 EG. Sin embargo, para la combinación de celulasas que contenía actividad potenciadora de celulasa, se omitió la EG y se reemplazó por TEMER07589. Aparte de una actividad potenciadora de endoglucanasa o celulasa diferente en las mezclas de enzimas, también se probaron una pocas relaciones diferentes de las 4 enzimas. Todas estas combinaciones de 4 enzimas se añadieron en 5 mg de proteína como equivalentes de BSA por g de paja de trigo pretratada lavada, y se cribaron con respecto a su capacidad de hidrólisis de celulosa en incubaciones de 20 h a 65°C, según se describe anteriormente. Además de estas mezclas de enzimas, un producto de celulasa de *Talaromyces emersonii* clásico, conocido como Filtrase® NL, también se incubó a una dosis de proteína similar.

Se determinó la liberación de glucosa, y se muestra en la Tabla 1 para las mezclas de 4 enzimas. Para Filtrase® NL, se determinó que la glucosa liberada era 30,1 mmol/l.

Tabla 1: Composición de 4 mezclas compuestas por combinaciones de 4 celulasas, con cantidades relativas variables de las 4 celulasas (dadas como porcentaje de la dosis de proteína total), y la glucosa liberada (expresada como mmol/l) después de 20 h de incubación a pH 4,5 y 65°C, para cada mezcla, cuando contienen CEA, CEB o TEMER07589 como EG.

	Porcentaje de cada una de las 4 enzimas en la mezcla				Glucosa liberada (mmol/l)		
	BG	CBHI	CBHII	EG*	CEA	CEB	TEMER07589
Mezcla 1	4%	14%	25%	57%	17,3	13,6	31,2
Mezcla 2	4%	18%	10%	68%	17,0	15,7	21,6
Mezcla 3	8%	14%	25%	53%	26,4	13,8	28,1
Mezcla 4	8%	20%	29%	43%	18,4	12,0	31,9

\* EG indica el porcentaje de endoglucanasa, en el caso de CEA y CEB, e indica el porcentaje de actividad potenciadora de celulasa en el caso de TEMER07589, mezcla en la que se omite una endoglucanasa verdadera.

5 A partir de estos resultados, esta claro que para las 4 relaciones de enzimas probadas, la actividad potenciadora de celulasa superaba a las dos endoglucanasas. Sin embargo, en el caso de la relación de enzimas que se da en la Mezcla 1 y la Mezcla 4, la mezcla de 4 enzimas que contiene actividad potenciadora de celulasa, TEMER07589, en lugar de endoglucanasa, incluso superaba a Filtrase® NL, que contiene varias endoglucanasas, beta-glucosidasas y las dos exoglucanasas, CBHI y CBHII.

## 10 Ejemplo 3

### 3.1 Optimización de la relación de mezclas de 4 enzimas

15 Las 3 mezclas de 4 enzimas diferentes que se usan en el Ejemplo 2 se usaron en un diseño de mezcla que consistía en 10 vértices, y un número total de 55 combinaciones de relaciones diferentes para encontrar la relación óptima de las 4 enzimas para cada una de las 3 mezclas. Los intervalos de las diferentes enzimas probadas en este diseño eran BG 4-12%, y cada una de las otras 3 enzima en los intervalos de 10-70%. En cada serie de 55 incubaciones, se efectuaron 3 duplicados. El cribado de estas mezclas se realizó sobre paja de trigo pretratada lavada a 2% de DM en tampón de acetato 50 mM de pH 4,5. La incubación se realizó a una dosis de proteína total de 5 mg de proteína como equivalentes de BSA por g de materia seca de paja de trigo pretratada lavada, durante 20 h a 65°C, en placas de microvaloración de 96 pocillos hondos. La capacidad de hidrólisis de celulosa se determinó mediante la glucosa liberada según se determina mediante el ensayo de azúcares reductores con BCA. La evaluación estadística de todos los datos para cada mezcla daba como resultado la relación optimizada de las 4 enzimas en cada mezcla, según se da en la Tabla 2.

25 Tabla 2. Composición óptima de 3 mezclas de 4 enzimas diferentes, que contenía cada una 1 BG, 1 CBHI, 1 CBHII y 1 EG o actividad potenciadora de celulasa, expresada como porcentaje de la proteína total, para cada enzima individual.

	BG	CBHI	CBHII	EG*
Mezcla CEA	12%	36%	34%	18%
Mezcla CEB	4%	13%	15%	68%
Mezcla TEMER07589	9%	30%	24%	37%

\* EG indica el porcentaje de endoglucanasa, en el caso de CEA y CEB, e indica el porcentaje de actividad potenciadora de celulasa en el caso de TEMER07589, mezcla en la que se omite una endoglucanasa verdadera.

## 30 3.2 Hidrólisis ampliada de mezclas de 4 enzimas optimizadas

35 Las 3 mezclas optimizadas se aplicaron en la hidrólisis ampliada a una escala de 10 ml con 2% de DM de paja de trigo pretratada, a pH 4,5 y 65°C. Como comparación, la Filtrase® NL previamente mencionada, un producto de celulasa de *Talaromyces emersonii* clásico, también se incubó a una concentración de sustrato, un pH y una temperatura iguales. La dosis de proteína total en cada una de las incubaciones era 15 mg de proteína como equivalentes de BSA por gramo de materia seca de paja de trigo pretratada. Las incubaciones duraban 72 horas y se tomaron muestras a diversos intervalos de tiempo. Las muestras se centrifugaron y el sobrenadante se congeló hasta el análisis mediante NMR. La liberación de glucosa con el tiempo se muestra en la Figura 2.

40 A partir de la Fig. 2, está claro que la mezcla que contiene actividad potenciadora de celulasa, TEMER07589, supera a las mezclas de dos endoglucanasas que contienen CEA o CEB. La mezcla que contiene actividad potenciadora de celulasa tiene una liberación de glucosa incluso ligeramente superior el final de la incubación que la Filtrase® NL.

LISTA DE SECUENCIAS

<110> DSM IP Assets B.V.

<120> POLIPÉPTIDO QUE TIENE O QUE AYUDA EN LA ACTIVIDAD DE DEGRADACIÓN DE MATERIALES GLUCÍDICOS Y USOS DEL MISMO

5 <130> 27666-WO-PCT

<160> 3

<170> Patent In versión 3.5

<210> 1

< 211> 765

10 < 212> ADN

< 213> T. emersonii

<400> 1

atgctgtctt cgaaggctcc tgtcaccctt gcctttgcag gcctcgctgg ccttctgtcc	60
gccccactgg tcaaggccca tggttttgtc cagggcattg tcatcggtga ccaattctac	120
agcgggtaca tcgtcaacga gttcccctac gaatccaacc cacccccctg catcggctgg	180
gccacgacag ccaccgacct gggcttcgtc gacggcactg aataccaagg accagacatc	240
atctgccacc ggaatgcgac gcccgcgctg ctgacagccc ccgtggccgc cggcggcacc	300
gtcgagctgc agtggacgcc ctggccgtcc agccaccacg ggccgggtcat cacgtacctg	360
gccaactgca acggcaactg ctcgaccgtc gacaagacgc agctggagtt cttcaagatc	420
gaccagtgg gcctgatcaa cgacactgac ccgccgggca cctgggctgc cgacaacctc	480
atcgccaaca acaacagctg gaccgtgacc atccccagca ccctcgagcc gggcaactac	540
gtgctgcgcc acgagatcat cgccctgcac tcggcgggca acaaagacgg cgcccagaac	600
tacccccagt gcatcaacat cgaggtcacg ggcggcggct cggtcgagcc gacgggcacg	660
ctgggcgagg atctctacca cgacacggac ccgggcattc tgatcgacat ttacgagccg	720
attgcgacgt ataccattcc aggaccgcct gagccgacgt tctag	765

<210> 2

15 < 211> 254

< 212> PRT

< 213> T. emersonii

<400> 2

ES 2 628 345 T3

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

Gly Leu Leu Ser Ala Pro Leu Val Lys Ala His Gly Phe Val Gln Gly  
20 25 30

Ile Val Ile Gly Asp Gln Phe Tyr Ser Gly Tyr Ile Val Asn Glu Phe  
35 40 45

Pro Tyr Glu Ser Asn Pro Pro Pro Val Ile Gly Trp Ala Thr Thr Ala  
50 55 60

Thr Asp Leu Gly Phe Val Asp Gly Thr Glu Tyr Gln Gly Pro Asp Ile  
65 70 75 80

Ile Cys His Arg Asn Ala Thr Pro Ala Leu Leu Thr Ala Pro Val Ala  
85 90 95

Ala Gly Gly Thr Val Glu Leu Gln Trp Thr Pro Trp Pro Ser Ser His  
100 105 110

His Gly Pro Val Ile Thr Tyr Leu Ala Asn Cys Asn Gly Asn Cys Ser  
115 120 125

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

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

Ile Ala Asn Asn Asn Ser Trp Thr Val Thr Ile Pro Ser Thr Leu Glu  
165 170 175

Pro Gly Asn Tyr Val Leu Arg His Glu Ile Ile Ala Leu His Ser Ala  
180 185 190

Gly Asn Lys Asp Gly Ala Gln Asn Tyr Pro Gln Cys Ile Asn Ile Glu  
195 200 205

Val Thr Gly Gly Gly Ser Val Glu Pro Thr Gly Thr Leu Gly Glu Asp  
210 215 220

Leu Tyr His Asp Thr Asp Pro Gly Ile Leu Ile Asp Ile Tyr Glu Pro  
225 230 235 240

Ile Ala Thr Tyr Thr Ile Pro Gly Pro Pro Glu Pro Thr Phe  
245 250

<210> 3

5 <211> 26

ES 2 628 345 T3

< 212> PRT

< 213> T. emersonii

<400> 3

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

Gly Leu Leu Ser Ala Pro Leu Val Lys Ala  
          20                   25

5

**REIVINDICACIONES**

1. Un polipéptido que comprende la secuencia de aminoácidos indicada en SEQ ID N°: 2 o una secuencia de aminoácidos codificada por la secuencia nucleotídica de SEQ ID N°: 1, o un polipéptido variante o un polinucleótido variante del mismo, en donde el polipéptido variante tiene al menos 95% de identidad de secuencia con la secuencia indicada en SEQ ID N°: 2 o el polinucleótido variante codifica un polipéptido que tiene al menos 95% de identidad de secuencia con la secuencia indicada en SEQ ID N°: 2 y en donde el polipéptido tiene o consiste en una actividad de degradación de materiales glucídicos .
2. Un polinucleótido que comprende:
- (a) la secuencia nucleotídica indicada en SEQ ID N°: 1; o
- (b) una secuencia nucleotídica que se hibrida selectivamente con un polinucleótido que es el complemento inverso de SEQ ID N°: 1; o
- (c) una secuencia nucleotídica que tiene al menos 95% de identidad de secuencia con la secuencia nucleotídica de SEQ ID N°: 1; o
- (d) una secuencia que está degenerada como resultado del código genético hasta una secuencia como la definida en uno cualquiera de (a), (b) o (c); o
- (e) una secuencia nucleotídica que es el complemento inverso de una secuencia nucleotídica como la definida en (a), (b), (c) o (d).
3. Un polinucleótido según la reivindicación 2, que codifica un polipéptido según la reivindicación 1.
4. Un polinucleótido según la reivindicación 2 o 3, que es una secuencia de ADN.
5. Una construcción de ácido nucleico que comprende el polinucleótido según una cualquiera de las reivindicaciones 2 a 4.
6. Un vector que incorpora una secuencia polinucleotídica según una cualquiera de las reivindicaciones 2 a 4.
7. Una célula recombinante transformada con un polinucleótido según una cualquiera de las reivindicaciones 2 a 4, una construcción de ácido nucleico según la reivindicación 5 o un vector según la reivindicación 6.
8. Una célula según la reivindicación 7, en donde la célula es una célula fúngica.
9. Una célula según la reivindicación 8, en donde la célula fúngica se selecciona del grupo que consiste en los géneros *Aspergillus*, *Trichoderma*/*Hypocrea*, *Fusarium*, *Disporotrichum*, *Penicillium*, *Acremonium*, *Neurospora*, *Thermoascus*, *Myceliophthora*, *Sporotrichum*, *Thielavia*, *Chryosporium*, *Fusarium*, *Humicola*, *Neurospora* y *Talaromyces*.
10. Una célula según una cualquiera de las reivindicaciones 7 a 9, en la que uno o más genes están eliminados, inactivados o alterados totalmente o en parte, en donde opcionalmente el gen codifica una proteasa.
11. Un método para la preparación de un polipéptido según la reivindicación 1, que tiene actividad de degradación de materiales glucídicos y/o actividad de hidrolización de carbohidratos, método que comprende cultivar una célula según cualquiera de las reivindicaciones 7 a 10 bajo condiciones que permiten la expresión de dicho polipéptido y, opcionalmente, recuperar el polipéptido expresado.
12. Una composición que comprende: (i) un polipéptido según la reivindicación 1, y (ii) una celulasa y/o una hemicelulasa y/o una pectinasa.
13. Una composición según la reivindicación 12, en la que la celulasa es una celobiohidrolasa, celobiohidrolasa I, celobiohidrolasa II, una endo- $\beta$ -1,4-glucanasa, una  $\beta$ -glucosidasa o una  $\beta$ -(1,3)(1,4)-glucanasa.
14. Una composición según la reivindicación 12 o 13, en la que la hemicelulasa es una endoxilanasasa, una  $\beta$ -xilosidasa, una  $\alpha$ -L-arabinofuranosidasa, una  $\alpha$ -D-glucuronidasa, una acetil xilano esterasa, una feruloil esterasa, una cumaroil esterasa, una  $\alpha$ -galactosidasa, una  $\beta$ -galactosidasa, una  $\beta$ -mananasa o una  $\beta$ -manosidasa.

- 5 15. Una composición según una cualquiera de las reivindicaciones 12 a 14, en la que la pectinasa es una endopoligalacturonasa, una pectina metil esterasa, una endo-galactanasa, una betagalactosidasa, una pectina acetil esterasa, una endo-pectina liasa, pectato liasa, alfa-ramnosidasa, una exo-galacturonasa, una exopoligalacturonato liasa, una ramnogalacturonano hidrolasa, una ramnogalacturonano liasa, una ramnogalacturonano acetil esterasa, una ramnogalacturonano galacturonohidrolasa, una xilogalacturonasa, una alfa-arabinofuranosidasa o una endo-arabinanasa.
- 10 16. Una composición según una cualquiera de las reivindicaciones 12 a 15, que comprende una ligninasa, una expansina, un polipéptido similar a expansina, una swollenina o el producto polipeptídico de un gen que codifica una proteína integradora de celulosa o una proteína inducida por celulosa.
- 15 17. Un método para el tratamiento de un sustrato que comprende material glucídico, opcionalmente un material vegetal, método que comprende poner en contacto el sustrato con un polipéptido según la reivindicación 1 y/o una composición según una cualquiera de las reivindicaciones 12 a 16.
- 20 18. Un método según la reivindicación 17, en el que el sustrato es un material vegetal y el material vegetal se proporciona en la forma de una planta, una pulpa vegetal, un extracto vegetal, un producto alimenticio o un ingrediente derivado del mismos o un tejido, un material textil o una prenda de vestir que comprende un material vegetal.
19. Uso de un polipéptido según la reivindicación 1 y/o una composición según una cualquiera de las reivindicaciones 12 a 16, para producir azúcar a partir de un material lignocelulósico.
20. Un procedimiento para la preparación de un producto de fermentación, procedimiento que comprende:
- 25 a) degradar lignocelulosa al poner en contacto la lignocelulosa con un polipéptido según la reivindicación 1 y/o una composición según una cualquiera de las reivindicaciones 12 a 16,
- b) fermentar el material resultante para preparar un producto de fermentación, y
- c) opcionalmente, recuperar el producto de fermentación.

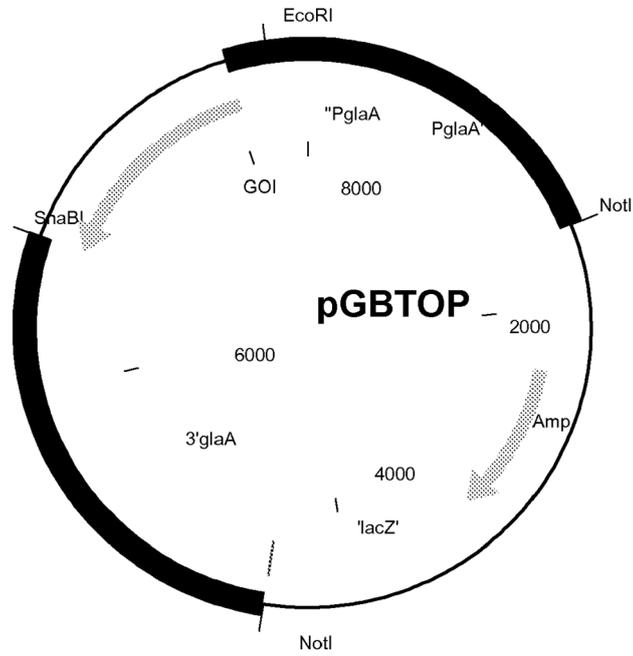


Fig. 1

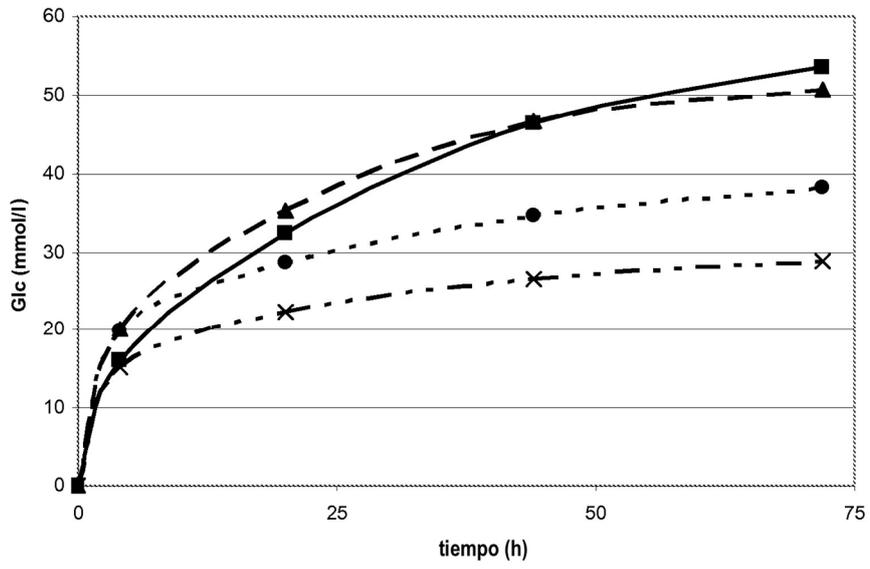


Fig. 2