

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 802 807**

51 Int. Cl.:

**C12N 9/24** (2006.01)

**C12R 1/645** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **03.02.2014 E 18193710 (3)**

97 Fecha y número de publicación de la concesión europea: **13.05.2020 EP 3447136**

54 Título: **Polipéptido degradador de carbohidratos y sus usos**

30 Prioridad:

04.02.2013 EP 13153824	04.02.2013 EP 13153829
04.02.2013 EP 13153831	04.02.2013 EP 13153834
04.02.2013 EP 13153835	04.02.2013 EP 13153836
04.02.2013 EP 13153837	04.02.2013 EP 13153839
04.02.2013 EP 13153840	04.02.2013 EP 13153841
04.02.2013 EP 13153828	04.02.2013 EP 13153825
04.02.2013 EP 13153823	04.02.2013 EP 13153821
04.02.2013 EP 13153833	26.02.2013 EP 13156692
26.02.2013 EP 13156684	26.02.2013 EP 13156693
26.02.2013 EP 13156679	26.02.2013 EP 13156694
26.02.2013 EP 13156685	26.02.2013 EP 13156696
26.02.2013 EP 13156678	26.02.2013 EP 13156698
26.02.2013 EP 13156688	26.02.2013 EP 13156701
26.02.2013 EP 13156682	26.02.2013 EP 13156702
26.02.2013 EP 13156690	26.02.2013 EP 13156703

73 Titular/es:

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

72 Inventor/es:

**LOS, ALRIK PIETER;  
DE JONG, RENÉ MARCEL y  
APPELDOORN, MAAIKE**

74 Agente/Representante:

**LEHMANN NOVO, María Isabel**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**21.01.2021**

**ES 2 802 807 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 degradador de carbohidratos y sus usos

**Campo de la invención**

5 La invención se refiere a secuencias que comprenden genes que codifican polipéptidos que tienen actividad degradadora de materiales lignocelulósicos. La invención presenta 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 degradadora de materiales lignocelulósicos en un organismo hospedador tal como *Aspergillus niger* y/o *Rasamsonia emersonii*.

**Antecedentes de la invención**

15 Los carbohidratos constituyen los compuestos orgánicos más abundantes de la tierra. Sin embargo, mucho 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 glucídicos de la lignocelulosa son celulosa, hemicelulosa y pectinas. A menudo, estos polímeros complejos se denominan 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 gran atención de los investigadores desde los 70, cuando estalló la crisis del petróleo debido a la 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 dos últimas décadas. Más recientemente, el uso de E85, una combinación de etanol al 85%, se ha establecido especialmente para aplicaciones a ciudades limpias. La importancia del bioetanol combustible se incrementará en paralelo con los incrementos en los precios para el crudo y el agotamiento gradual de sus fuentes. Adicionalmente, se están usando azúcares fermentables para producir plásticos, polímeros y otros productos de base biológica y se espera que esta industria crezca sustancialmente incrementando por lo tanto la demanda de azúcares fermentables de bajo coste abundantes que se puedan usar como una materia prima en lugar de materias primas de base petrolífera.

30 El secuestro de estas grandes cantidades de carbohidratos en biomasa vegetal proporciona una fuente copiosa de energía potencial en 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 energético de estos carbohidratos está actualmente infrautilizado debido a que los azúcares están bloqueados en polímeros complejos, y de ahí que no sean fácilmente accesibles para la fermentación. Los métodos que generen azúcares de biomasa vegetal proporcionarán materias primas económicamente competitivas copiosas para la 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 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 bioconversión de biomasa. Los costes de producción de enzimas producidas microbianamente están estrechamente ligados a 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 de las últimas décadas por entender la degradación enzimática de biomasa lignocelulósica y la producción de celulasa, sigue siendo deseable descubrir o manipular nuevas celulasas y hemicelulasas muy activas. También sería muy deseable construir composiciones enzimáticas altamente eficaces capaces de realizar una biodegradación rápida y eficaz de materiales lignocelulósicos, en particular celulasas y hemicelulasas tales 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 ensilado, y también como enzima en otros procedimientos industriales, por ejemplo en las industrias alimenticia o de piensos, textil, de la pasta papelera o el papel o los detergentes y otras industrias.

55 El documento WO 2007/091231 describe enzimas de *Talaromyces mersonii* incluyendo alfa-arabinofuranosidasas.

**Sumario de la invención**

La presente invención proporciona un polipéptido que tiene actividad como hemicelulasa o una actividad según la Tabla 1 que comprende la secuencia de aminoácidos indicada en la SEQ ID NO: 32 o una secuencia de aminoácidos codificada por la secuencia nucleotídica de SEQ ID NO: 31, SEQ ID NO: 34 o uno de sus polipéptidos variantes o polinucleótidos variantes, en donde el polipéptido variante tiene al menos 75% de identidad de secuencia con la secuencia indicada en SEQ ID NO: 32 o el polinucleótido variante codifica un polipéptido que tiene al menos 75% de identidad de secuencia con la secuencia indicada en SEQ ID NO: 32.

Tabla 1: Números de Temer de las proteínas de la invención y su actividad

	Número de Temer	actividad	actividad
1	Temer00088	beta-xilosidasa	GH3
2	Temer09484	beta-xilosidasa	GH3
3	Temer08028	alfa-galactosidasa	GH27
4	Temer02362	alfa-galactosidasa	GH27
5	Temer08862	alfa-galactosidasa	GH27
6	Temer04790	xiloglucanasa	GH12
7	Temer05249	alfa-arabinofuranosidasa	GH51
8	Temer06848	alfa-arabinofuranosidasa	GH51
9	Temer02056	alfa-arabinofuranosidasa	GH51
10	Temer03124	endoxilanasas	GH43
11	Temer09491	manosidasa/xilosidasa	GH31
12	Temer06400	feruloilesterasa	CE1
13	Temer08570	endoxilanasas	GH39
14	Temer08163	endo-exo-xilanasas	GH30
15	Temer07305	alfa-glucuronidasa	GH115

10

El polipéptido de la invención tiene actividad como alfa-arabinofuranosidasa.

Por otra parte, la invención proporciona una secuencia de ácido nucleico que codifica una hemicelulasa, en donde la secuencia de ácido nucleico se selecciona del grupo que consiste en:

15 (a) una secuencia de ácido nucleico que tiene al menos 70% de identidad con la secuencia de ácido nucleico de SEQ ID NO: 31 SEQ ID NO: 34 o SEQ ID NO: 35;

20 (b) una secuencia de ácido nucleico que codifica (i) la secuencia de aminoácidos de SEQ ID NO: 32, (ii) una secuencia de aminoácidos que tiene al menos 70% de identidad con la secuencia de aminoácidos de SEQ ID NO: 32, o (iii) una secuencia de aminoácidos que difiere en 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 o 12 aminoácidos de la secuencia de aminoácidos de SEQ ID NO: 32; o

(c) una secuencia nucleotídica que es el complemento inverso de una secuencia nucleotídica como la definida en (a) o (b).

25 La invención también proporciona un constructo o vector de ácido nucleico que comprende el polinucleótido de la invención y una célula que comprende un polipéptido de la invención o un constructo o vector de ácido nucleico de la invención.

30 Según un aspecto de la invención, la célula es una célula fúngica, preferiblemente una célula fúngica seleccionada del grupo que consiste en los géneros *Acremonium*, *Agaricus*, *Aspergillus*, *Aureobasidium*, *Chrysosporium*, *Coprinus*, *Cryptococcus*, *Filibasidium*, *Fusarium*, *Humicola*, *Magnaporthe*, *Mucor*, *Myceliophthora*, *Neocallimastix*, *Neurospora*, *Paecilomyces*, *Penicillium*, *Piromyces*, *Panerochaete*, *Pleurotus*, *Schizophyllum*, *Talaromyces*, *Rasamsonia*, *Thermoascus*, *Thielavia*, *Tolyocladium* y *Trichoderma*.

Según otro aspecto de la invención, uno o más genes de la célula de la invención se eliminan, inactivan o alteran totalmente o en parte, en donde opcionalmente el gen codifica una proteasa.

La invención también proporciona un método para la preparación de un polipéptido según la invención que tiene actividad como hemicelulasa o una según la Tabla 1, método que comprende cultivar una célula de la invención bajo condiciones que permitan la expresión de dicho polipéptido y, opcionalmente, recuperar el polipéptido expresado. Por otra parte, la invención proporciona una composición que comprende: (i) un polipéptido de la invención y; (ii) una celulosa y/o una hemicelulasa adicional y/o una pectinasa, preferiblemente la celulasa es una GH61, celobiohidrolasa, celobiohidrolasa I, celobiohidrolasa II, endo- $\beta$ -1,4-glucanasa,  $\beta$ -glucosidasa o  $\beta$ -(1,3)(1,4)-glucanasa y/o la hemicelulasa es una endoxilanasas,  $\beta$ -xilosidasa,  $\alpha$ -L-arabinofuranosidasa,  $\alpha$ -D-glucuronidasa, feruloilesterasa, cumaroilesterasa,  $\alpha$ -galactosidasa,  $\beta$ -galactosidasa,  $\beta$ -mananasa o  $\beta$ -manosidasa.

Adicionalmente, la invención proporciona un método para el tratamiento de un sustrato que comprende hemicelulosa, opcionalmente un material vegetal, método que comprende poner en contacto el sustrato con un polipéptido de la invención y/o una composición de la invención.

Otro aspecto de la invención se refiere al uso de un polipéptido de la invención y/o una composición de la invención para producir azúcar a partir de un material lignocelulósico.

La invención también proporciona:

un método para la preparación de un polipéptido que tiene actividad degradadora de material glucídico o hidrolizante 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 este método; y

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

Los polipéptidos de la invención que tienen actividad degradadora de material glucídico o hidrolizante 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.

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.

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.

También se puede usar un polipéptido o una composición de la invención, 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 un papel o en la preparación de una tela o material textil o en su limpieza.

### **Breve descripción de los dibujos**

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 Temer00088, Temer09484, Temer08028, Temer02362, Temer08862, Temer04790, Temer05249, Temer06848, Temer02056, Temer03124, Temer09491, Temer06400, Temer08570, Temer08163 o Temer07305 según se define posteriormente en la presente.

La Fig. 2 muestra un diagrama esquemático del plásmido Te pep.bbn, que es la base para el constructo de sobreexpresión de Temer00088, Temer09484, Temer08028, Temer02362, Temer08862, Temer04790, Temer05249, Temer06848, Temer02056, Temer03124, Temer09491, Temer06400, Temer08570, Temer08163 o Temer07305 en *R. emersonii* que se orienta al locus *RePepA*. El vector comprende una región de flaqueo 5' de 1500 pb, 1,5 kb aguas arriba del ORF de *RePepA* para la orientación en el locus *RePepA*, un sitio lox66, la parte 5' no funcional de la región codificante ble (5'ble) conducida por el promotor *gpdA* de *A. nidulans*, y un gen *ccdB*.

La Fig. 3 muestra un diagrama esquemático del plásmido pEBA1006 que se usaba en el método de orientación génica bipartito en combinación con el vector de expresión pEBA que contiene Temer00088, Temer09484, Temer08028, Temer02362, Temer08862, Temer04790, Temer05249, Temer06848, Temer02056, Temer03124, Temer09491, Temer06400, Temer08570, Temer08163 o Temer07305 con el objetivo de reemplazar el OFR *RePepA* y

aproximadamente 1500 nucleótidos aguas arriba del codón ATG de partida mediante el casete de expresión de Temer00088, Temer09484, Temer08028, Temer02362, Temer08862, Temer04790, Temer05249, Temer06848, Temer02056, Temer03124, Temer09491, Temer06400, Temer08570, Temer08163 o Temer07305 en *Rasamsonia emersonii*. El vector comprende la parte 3' de la región codificante ble, el terminador *trpC* de *A. nidulans*, un sitio lox71, una región de flanqueo 3' de 2500 pb del ORF de *RePepA*, y el esqueleto de pUC19 (Invitrogen, Breda, Países Bajos). El ADN de *E. coli* se retiró mediante digestión con la enzima de restricción *NotI*, antes de la transformación de las cepas de *R. emersonii*.

La Fig. 4 muestra un diagrama esquemático del plásmido de expresión pEBA que contiene Temer00088, Temer09484, Temer08028, Temer02362, Temer08862, Temer04790, Temer05249, Temer06848, Temer02056, Temer03124, Temer09491, Temer06400, Temer08570, Temer08163 o Temer07305 que se usaba en el método de orientación génica bipartito en combinación con el vector pEBA1006 con el objetivo de reemplazar el ORF de *RePepA* y aproximadamente 1500 nucleótidos aguas arriba del codón ATG de partida mediante el casete de expresión de Temer00088, Temer09484, Temer08028, Temer02362, Temer08862, Temer04790, Temer05249, Temer06848, Temer02056, Temer03124, Temer09491, Temer06400, Temer08570, Temer08163 o Temer07305 en *Rasamsonia emersonii*. El vector comprende una región de flanqueo 5' de 1500 pb 1,5 kb aguas arriba del ORF de *RePepA* para la orientación en el locus *RePepA*, consistiendo el casete de expresión de Temer00088, Temer09484, Temer08028, Temer02362, Temer08862, Temer04790, Temer05249, Temer06848, Temer02056, Temer03124, Temer09491, Temer06400, Temer08570, Temer08163 o Temer07305 en el promotor 2 de *R. emersonii*, la región codificante de Temer00088, Temer09484, Temer08028, Temer02362, Temer08862, Temer04790, Temer05249, Temer06848, Temer02056, Temer03124, Temer09491, Temer06400, Temer08570, Temer08163 o Temer07305 en el promotor 2 de *R. emersonii*, la región codificante de *amdS* (*TamdS*) de *A. nidulans*, un sitio lox66, la parte 5' no funcional de la región codificante ble (5' ble) conducida por el promotor *gpdA* de *A. nidulans*. El ADN de *E. coli* se retiró mediante digestión con la enzima de restricción *NotI*, antes de la transformación de las cepas de *R. emersonii*.

Fig. 5 Cromatograma obtenido mediante cromatografía de intercambio aniónico de alta resolución que muestra la formación de oligómeros por Temer04790 de *Rasamsonia emersonii* en comparación con una mezcla de celulasas comercial después de la incubación sobre xiloglucano durante 24 h de incubación a pH 4,5 y 60°C.

**Breve descripción del listado de secuencias**

La Tabla 2 muestra SEQ ID NO's de secuencias codificantes optimizadas para pares de codones, SEQ ID NO's de secuencias de aminoácidos, SEQ ID NO's de secuencias de señalización, SEQ ID NO's de secuencias de ADN genómicas y SEQ ID NO's y de secuencias codificantes silvestres de la presente invención.

	Número de Temer	SEQ ID NO de secuencias codificantes optimizadas para pares de codones:	SEQ ID NO de secuencias de aminoácidos:	SEQ ID NO de secuencias de señalización:	SEQ ID NO de secuencias de ADN genómicas:	SEQ ID NO de secuencias codificantes silvestres:
1	Temer00088	1	2	3	4	5
2	Temer09484	6	7	8	9	10
3	Temer08028	11	12	13	14	15
4	Temer02362	16	17	18	19	20
5	Temer08862	21	22	23	24	25
6	Temer04790	26	27	28	29	30
7	Temer05249	31	32	33	34	35
8	Temer06848	36	37	38	39	40
9	Temer02056	41	42	43	44	45
10	Temer03124	46	47	48	49	50
11	Temer09491	51	52	53	54	55
12	Temer06400	56	57	58	59	60
13	Temer08570	61	62	63	64	65
14	Temer08163	66	67	68	69	70
15	Temer07305	71	72	73	74	75

SEQ ID NO: 76 *RePepA* de *R. emersonii* (secuencia genómica incluyendo flancos)

SEQ ID NO: 77 *RePepA* de *R. emersonii* (ADNc)

5

SEQ ID NO: 78 *RePepA* de *R. emersonii* (proteína)

SEQ ID NO: 79 Promotor *gdpA* de *A. nidulans* y parte 5' de la región codificante *ble*

10 SEQ ID NO: 80 parte 3' de la región codificante *ble* y terminador *TrpC* de *A. nidulans*

SEQ ID NO: 81 Promotor 2 de *R. emersonii*

SEQ ID NO: 82 Terminador *AmdS* de *A. nidulans*

15 **Descripción detallada de la invención**

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

20

Los artículos "un" y "uno(a)" se usan en la presente para referirse a uno o a 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.

25 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 polipéptidos.

30 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 de la misma se modifique. Además del material glucídico, el sustrato puede contener cualquier otro componente, incluyendo, pero no limitado a, material no glucídico y almidón.

35 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 polipéptidos.

40 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 los extremos 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.

45 En la presente, una  $\alpha$ -galactosidasa (EC 3.2.1.22; GH27) 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.

50 En la presente, una xiloglucanasa es una endo- $\beta$ -1,4-glucanasa específica de xiloglucano, que cataliza la escisión de xiloglucano, un esqueleto de residuos de glucosa con enlaces  $\beta$ 1  $\rightarrow$ 4, la mayoría de los cuales sustituidos con cadenas laterales de xilosa con enlaces 1-6.

TEMER05249

55 Típicamente, un polipéptido de la invención codifica un polipéptido que tiene al menos actividad como alfa-arabinofuranosidasa, tentativamente llamado TEMER05249, que tiene una secuencia de aminoácidos según SEQ ID NO: 32, o una secuencia que es una de sus variantes, típicamente funcionalmente equivalente al polipéptido que tiene la secuencia de SEQ ID NO: 32, o una secuencia que es un fragmento de cualquiera de ellas.

60 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.

5 Un polipéptido de la invención puede tener una o más actividades degradadoras de carbohidratos y/o hidrolizadoras de carbohidratos alternativas y/o adicionales distintas a la actividad como alfa-arabinofuranosidasa, por ejemplo una de las otras actividades degradadoras de carbohidratos y/o hidrolizadoras de carbohidratos mencionadas en la presente.

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

15 En la presente, una  $\beta$ -manosidasas (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 mananasas o manasa.

20 En la presente, una feruloilesterasa (EC 3.1.1.73; CE1) es cualquier polipéptido que sea capaz de catalizar una reacción en 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 cinamoílico hidrolasa, ácido ferúlico esterasa o hidroxicinamoilesterasa. También se puede denominar enzima auxiliar de hemicelulosa, ya que ayuda a las xilanasas y pectinasas a romper la hemicelulosa y la pectina de la pared celular vegetal.

25 En la presente, una  $\alpha$ -D-glucuronidasas (EC 3.2.1.139; GH115) 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-glucuronidasas o alfa-glucosiduronasa. Estas enzimas también pueden hidrolizar ácido glucurónico 4-O-metilado, que también puede estar presente como un sustituyente en xilanos. Una alternativa es EC 3.2.1.131: xilano alfa-1,2-glucuronosidasas, que cataliza la hidrólisis de enlaces alfa-1,2-(4-O-metil)glucuronosilo.

35 Carbohidrato en este contexto incluye todos los sacáridos, por ejemplo polisacáridos, oligosacáridos, disacáridos o monosacáridos.

Un polipéptido según la invención puede modificar un material glucídico al degradar químicamente o degradar físicamente este material o hidrolizar el carbohidrato. 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.

#### 40 **Materiales glucídicos adecuados**

Los materiales lignocelulolíticos o lignocelulósicos o la biomasa son abundantes en la naturaleza y tienen gran valor como fuentes de energía alternativas. Los biocombustibles de segunda generación, también conocidos como biocombustibles avanzados, son combustibles que se pueden fabricar a partir de diversos tipos de biomasa. Biomasa es un término amplio que significa cualquier fuente de carbono orgánico que se renueva rápidamente como parte del ciclo del carbono. La biomasa se deriva de materiales vegetales pero también puede incluir materiales animales. La composición de la biomasa lignocelulósica varía, el componente principal es celulosa (35-50%), seguido por xilano (20-35%, un tipo de hemicelulosa) y lignina (10-25%), además de componentes secundarios tales como proteínas, aceites y cenizas que constituyen la fracción restante de la biomasa lignocelulósica. La biomasa lignocelulósica contiene una variedad de carbohidratos. El término carbohidrato es el más común en bioquímica, donde es un sinónimo de sacárido. Los carbohidratos (sacáridos) se dividen en cuatro agrupaciones químicas: monosacáridos, disacáridos, oligosacáridos y polisacáridos. En general, los monosacáridos y disacáridos, que son carbohidratos menores (de peso molecular inferior), se denominan comúnmente azúcares.

55 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 como una materia prima renovable potencial, son celulosa (glucanos), hemicelulosas (xilanos, heteroxilanos y xiloglucanos). Además, algo de 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 conjuntamente.

60 La lignina rellena los espacios de la pared celular entre los componentes de celulosa, hemicelulosa y pectina, especialmente en traqueidas del xilema, elementos vasculares y esclereidas. Está enlazada covalentemente a

hemicelulosa y, por lo tanto, reticula diferentes polisacáridos vegetales, confiriendo resistencia mecánica a la pared celular y por extensión a la planta como un todo. La lignina es un material polimérico aromático reticulado altamente hidrófobo que está formado por diferentes monómeros de monolignol, que pueden estar metoxilados en diversos grados. Existen tres monómeros de monolignol, metoxilados hasta diversos grados: alcohol p-cumarílico, alcohol coniferílico y alcohol sinapílico. Estos lignoles están incorporado en la lignina en la forma de los fenilpropanoides p-hidroxifenilo (H), guayacilo (G) y siringilo (S), respectivamente. La biodegradación de lignina es un requisito previo para procesar biocombustible procedente de materias primas vegetales. La lignina se puede degradar al aplicar diferentes métodos de pretratamiento, o al usar ligninasas o enzimas modificadoras de lignina (LME's). La mejora de la degradación de lignina conduciría al resultado del procesamiento del biocombustible hasta una mejor ganancia o un mejor factor de eficacia, por ejemplo al mejorar la accesibilidad a los componentes (hemi)celulósicos o al retirar enlaces de lignina-(hemi)celulosa en oligosacáridos liberados mediante la acción de (hemi)celulasas.

Además, las pectinas y otras sustancias pécticas tales como los arabinanos pueden constituir una proporción considerable de la masa seca típicamente de paredes celulares de tejidos de plantas no leñosas (alrededor de un cuarto a la mitad de la masa seca pueden ser pectinas).

La celulosa es un polisacárido lineal compuesto por residuos de glucosa enlazados por uniones  $\beta$ -1,4. La naturaleza lineal de las fibras de celulosa, así como la estequiometría de la glucosa con enlazada en  $\beta$  (con relación a  $\alpha$ ), genera estructuras más tendentes a unión por hidrógeno intercatenaria que las estructuras con enlaces  $\alpha$  altamente ramificadas del almidón. Así, los polímeros de celulosa son generalmente menos solubles, y forman fibras más estrechamente unidas que las fibras encontradas en el almidón.

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 con enlaces  $\beta$ -1,4, un azúcar de cinco carbonos. Sin embargo, esta xilosa a menudo está ramificada en O-3 y/u O-2 y puede estar sustituida con enlaces a arabinosa, galactosa, manosa, ácido glucurónico, ácido galacturónico o por esterificación a ácido acético (y esterificación de ácido ferúlico a arabinosa). La hemicelulosa también puede contener glucano, que es un término general para azúcares de seis carbonos con enlaces  $\beta$  (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  $\beta$ ).

La composición, la naturaleza de la sustitución y el grado de ramificación de hemicelulosa es muy diferente en plantas dicotiledóneas (es decir, una planta cuyas semillas tienen dos cotiledones u hojas seminales tales como habas, cacahuètes, almendras, guisantes, frijoles) en comparación con plantas monocotiledóneas (es decir, plantas que tiene un solo cotiledón u hoja seminal tales como maíz, trigo, arroz, hierbas, cebada). En la 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 de los cultivos de cereales, los principales componentes de la hemicelulosa son heteroxilanos. Estos están comprendidos principalmente por polímeros del 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 mediante ácidos acéticos con enlaces éster. También están presentes  $\beta$ -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 paredes celulares. Además, diferentes plantas pueden comprender cantidades diferentes de, y diferentes composiciones de, sustancias pécticas. Por ejemplo, la remolacha azucarera contiene alrededor de 19% de pectina y alrededor de 21% de arabinano sobre una base en peso seco.

Según esto, una composición de la invención se puede ajustar a la 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.

Las combinaciones de enzimas o los tratamientos físicos se pueden administrar concomitantemente o secuencialmente. Las enzimas se pueden producir exógenamente en microorganismos, levaduras, hongos, bacterias o plantas, y 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 cruda, o material vegetal (tal como rastrojo de maíz) y similares. Alternativamente, la masa celular cruda o el medio de producción de enzimas o el material vegetal se pueden tratar para prevenir el crecimiento microbiano adicional (por ejemplo, mediante calentamiento o adición de agentes antimicrobianos), y a continuación se añade a la materia prima. Estas mezclas de enzimas crudas puede incluir el organismo que produce la enzima. Alternativamente, la enzima se puede producir en una fermentación que usa materia prima (tal como rastrojo 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.

**Actividad enzimática**

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

5 Las xilanasas junto con otras enzimas auxiliares, por ejemplo  $\alpha$ -L-arabinofuranosidasas, feruloil y acetilxilano estererasas, 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 en la pared celular vegetal. Se acumulan alrededor de una cadena central de unidades de ácido D-galacturónico con enlaces  $\alpha$ (1,4) intercalados en algún grado con L-ramnosa. En cualquier pared celular existe un número de unidades estructurales que se ajustan a esta descripción y generalmente se ha considerado que en una molécula péctica individual, las cadenas centrales de diferentes unidades estructurales son continuas entre sí.

15 Las pectinasas incluyen, por ejemplo, una endopoligalacturonasa, una pectina metilesterasa, una endogalactanasa, una  $\beta$ -galactosidasa, una pectina acetilesterasa, una endopectina liasa, pectato liasa,  $\alpha$ -ramnosidasa, una exogalacturonasa, un exopoligalacturonato 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 unidades de ácido galacturónico se alternan con unidades de ramnosa que soportan galactano con enlaces (1,4) y arabinano con enlaces (1,5). Las cadenas laterales de arabinano pueden estar ligadas directamente a ramnosa o indirectamente a través de cadenas de galactano; xilogalacturonano, con unidades de xilosilo simples en O-3 del ácido galacturónico (estrechamente asociado con RGI); y ramnogalacturonano II (RGII), una unidad secundaria particularmente compleja que contiene azúcares inusuales, por ejemplo apiosa. Una unidad de RGII puede contener dos residuos de apiosilo que, bajo condiciones iónicas adecuadas, pueden formar reversiblemente ésteres con borato.

25 Según se indica anteriormente, un polipéptido de la invención tendrá típicamente una actividad según la Tabla 1. 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 una actividad según la Tabla 1.

**35 Secuencia polinucleotídica**

La invención proporciona secuencias polinucleotídicas genómicas que comprenden el gen que codifica el Temer05249 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 NO: 31 o SEQ ID NO: 34 o SEQ ID NO: 35 y a variantes, tales como equivalentes funcionales, de cualquiera de las mismas.

40 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 NO: 31 o en SEQ ID NO: 34 o en SEQ ID NO: 35.

45 Más específicamente, la invención se refiere a un polinucleótido que comprende o que consiste esencialmente en una secuencia nucleotídica según SEQ ID NO: 31 o SEQ ID NO: 34 o SEQ ID NO: 35.

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 NO: 32 o una de sus variantes, tal como un equivalente funcional, o un fragmento de cualquiera de las mismas.

50 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 abierta que codifica una proteína, p. ej. la actividad según la presente invención.

55 Un gen puede incluir secuencias codificantes, 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.

60 Una molécula de ácido nucleico de la presente invención, tal como una molécula de ácido nucleico que tiene la secuencia nucleotídica de SEQ ID NO: 31 o SEQ ID NO: 34 o SEQ ID NO: 35 o una de sus variantes, tales como un equivalente funcional, se puede aislar usando técnicas de biología molecular estándar y la información de secuencias proporcionada en la presente. Por ejemplo, usar la totalidad o una porción de la secuencia de ácido nucleico de SEQ

ID NO: 31 como una sonda de hibridación, se pueden aislar moléculas de ácido nucleico según la invención 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 NO: 31 o SEQ ID NO: 34 o SEQ ID NO: 35 se puede aislar mediante la reacción en cadena de la polimerasa (PCR) usando sondas oligonucleotídicas sintéticas diseñadas basándose en la información de secuencia contenida en SEQ ID NO: 31 o en SEQ ID NO: 34 o en SEQ ID NO: 35.

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 análisis de secuencias de ADN.

Por otra parte, oligonucleótidos correspondientes a o hibridables a 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 NO: 31 o en SEQ ID NO: 34 o en SEQ ID NO: 35.

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 NO: 31 o en SEQ ID NO: 34 o en SEQ ID NO: 35 o una variante, tal como un equivalente funcional, de cualquiera de estas 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 uno de sus equivalentes funcionales, 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 estas 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 puede estar "aislado". En el contexto de esta invención, un "polinucleótido aislado" o "ácido nucleico aislado" es 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 alguna 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 está incorporado en un vector, en un plásmido o virus autónomamente replicativo, o en el ADN genómico de un procarionte o eucarionte, o que existe como una molécula separada (p. ej., un ADNc o un fragmento de ADN genómico producido mediante PCR o tratamiento con endonucleasas de restricción) independientemente 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 ADN 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" están destinados a 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 monocatenaria o bicatenaria, pero, preferiblemente, es ADN bicatenario. 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 alteradas de apareamiento de bases o una resistencia incrementada a nucleasas.

Otra realización de la invención proporciona una molécula de ácido nucleico aislada que es antisentido con respecto a una molécula de ácido nucleico de Temer05249, p. ej., la hebra codificante de una molécula de ácido nucleico de Temer05249. También se incluyen dentro del alcance de la invención las hebras complementarias de las moléculas de ácido nucleico descritas en la presente.

A menos que se indique otra cosa, todas las secuencias nucleotídicas determinadas al secuenciar una molécula de ADN se determinaron en la presente usando un secuenciador de ADN automatizado y todas las secuencias de aminoácidos de 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 especialidad para cualquier secuencia de ADN determinada mediante este enfoque automatizado, cualquier secuencia nucleotídica determinada en la presente puede contener algunos errores. Las secuencias nucleotídicas determinadas mediante automatización 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 secuenciada.

La secuencia real se puede determinar más precisamente mediante otros enfoques incluyendo métodos de secuenciación de ADN manuales muy conocidos en la especialidad. Como también se conoce en la especialidad, una sola inserción o eliminación en una secuencia nucleotídica determinada en comparación 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 secuenciada, empezando en el punto de esta inserción o eliminación.

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

Una molécula de ácido nucleico según la invención puede comprender solamente una porción o un fragmento de la secuencia de ácido nucleico mostrada en SEQ ID NO: 31 o en SEQ ID NO: 34 o en SEQ ID NO: 35 (o de una variante de cualquiera de las mismas), por ejemplo un fragmento que se puede usar como una sonda o cebador o un fragmento que codifica una porción de una proteína Temer05249.

La secuencia nucleotídica determinada a partir de la clonación del gen Temer05249 y ADNc permite la generación de sondas y cebadores diseñados para el uso en la identificación y/o la clonación de otros miembros de la familia Temer05249, así como homólogos de Temer05249 procedentes de otras especies.

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 NO: 31 o en SEQ ID NO: 34 o en SEQ ID NO: 35 o de una variante, tal como un equivalente funcional, de cualquiera de las mismas.

Sondas basadas en secuencias nucleotídicas de Temer05249 se pueden usar para detectar transcritos o secuencias de Temer05249 genómicas que codifican proteínas iguales u homólogas, por ejemplo, en otros organismos. En realizaciones preferidas, la sonda comprende además un grupo de etiquetado ligado a la misma, p. ej., el grupo de etiquetado puede ser un radioisótopo, un compuesto fluorescente, una enzima o un cofactor enzimático. Tales sondas también se pueden usar como parte de un estuche de pruebas diagnósticas para identificar células que expresan una proteína TEMER09484.

Los polinucleótidos de la presente pueden ser polinucleótidos sintéticos. Los polinucleótidos sintéticos se pueden optimizar en uso en 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 obtener una expresión mejorada de la secuencia nucleotídica que codifica el polipéptido y/o la producción mejorada del polipéptido codificado. Los pares de codones se definen como un conjunto de dos tripletes (codones) posteriores en una secuencia codificante.

La invención se refiere además a un constructo de ácido nucleico que comprende el polinucleótido que se describe anteriormente. "Constructo de ácido nucleico" se define en la presente como una molécula de ácido nucleico, bien mono- o bien bicatenaria, que se aísla de un gen presente en la naturaleza o que se ha modificado para contener segmentos de ácido nucleico que están combinados y yuxtapuestos de un modo que de otra manera no existiría en la naturaleza. El término constructo de ácido nucleico es sinónimo al término "casete de expresión" cuando el constructo de ácido nucleico contiene todas las secuencias de control requeridas para la expresión de una secuencia codificante. El término "secuencia codificante" según se define en la presente es una 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 se determinan generalmente mediante el codón de partida ATG en el extremo 5' del ARNm y una secuencia de codón de parada 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 alto contenido de GC. El contenido de GC indica en la presente el número de nucleótidos G y C en la construcción, dividido por el número total de nucleótidos, expresado en %. El contenido de GC es preferiblemente 56% o más, 57% o más, 58% o más, 59% o más, 60% o más, o en el intervalo de 56-70% o el intervalo de 58-65%. 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' a 3' seleccionada de la siguiente lista de secuencias: TAAG, TAGA y TAAA, preferiblemente TAAA, y/o

- una secuencia codificante iniciadora de la traducción orientada en la dirección 5' a 3' seleccionada de la siguiente lista de secuencias: GCTACCCCC; GCTACCTCC; GCTACCCTC; GCTACCTTC; GCTCCCCCC; GCTCCCTCC; GCTCCCTC; GCTCCCTTC; GCTGCCCCC; GCTGCCTCC; GCTGCCCTC; GCTGCCTTC; GCTTCCCCC; GCTTCCTCC; GCTTCCTC; y GCTTCCTTC, preferiblemente GCT TCC TTC, y/o

- una secuencia iniciadora de la traducción seleccionada de la siguiente lista de secuencias: 5'-mwChkyCAAA-3'; 5'-mwChkyCACAA-3' o 5'-mwChkyCAAG-3', usando códigos de ambigüedad para 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'.

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 partida del marco de lectura abierto de una secuencia codificante de ADN. El codón iniciador o de partida codifica el AA metionina. El codón iniciador es típicamente ATG, pero también puede ser cualquier codón de partida 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 del codón de parada de la traducción en el extremo 3' del marco de lectura abierto o la secuencia codificante nucleotídica y orientados en la dirección 5' a 3'.

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 partida del marco de lectura abierto de una secuencia de ADN que codifica un polipéptido. El codón iniciador o de partida codifica el AA metionina. El codón iniciador es típicamente ATG, pero también puede ser cualquier codón de partida funcional tal como GTG. Se sabe bien en la especialidad que el uracilo, U, reemplaza al desoxinucleótido timina, T, en el ARN.

## Homología e identidad

Se dice que las secuencias de aminoácidos o nucleotídicas 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 indican por el "porcentaje de identidad" o "porcentaje de similitud", que es alto o bajo, respectivamente. Aunque es discutible, para indicar el "porcentaje de identidad" o el "porcentaje de similitud", se usan frecuentemente de forma intercambiable el "nivel de homología" o "el porcentaje de homología".

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 la alineación entre las dos secuencias, se pueden introducir huecos en cualquiera de las dos secuencias que se comparan. Este 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 coincidencias de identidad entre las dos secuencias sobre la región alineada presentada.

Una comparación de secuencias y una determinación del porcentaje de identidad entre dos secuencias se pueden efectuar usando un algoritmo matemático. El experto conocerá el 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 sequence comparison en D. Sankoff y J. B. Kruskal, (ed.), Time warps, string 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 pueden 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 ejecutado en el programa informático NEEDLE. Para el propósito de esta invención, se usó el programa NEEDLE del lote 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 de proteínas, se usa EBLOSUM62 para la matriz de sustitución. Para secuencias nucleotídicas, se usa EDNAFULL. Se pueden especificar otras matrices. Los parámetros opcionales usados para el alineamiento de secuencias de aminoácidos son una penalización por apertura de hueco 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*

La homología o identidad es el porcentaje de coincidencias idénticas 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 muestra 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 está marcada en la salida del programa como "IDENTITY".

#### *Definición de la Identidad más Larga*

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 del alineamiento. La identidad definida como en la presente se puede obtener a partir de NEEDLE al usar la opción NOBRIEF y se marca en la salida 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.

Las secuencias proteínicas de la presente invención se pueden usar además como "secuencia incógnita" para realizar una búsqueda frente a bases de datos de secuencias, por ejemplo para identificar otros miembros de la familia o secuencias relacionadas. Estas búsquedas se pueden 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 proteínas
- Coste para extender un hueco: defecto = 2 para nucleótidos/ 1 para proteínas
- Penalización para divergencia de nucleótidos: defecto = -3
- Recompensa para coincidencia de nucleótidos: defecto = 1
- Valor esperado: defecto = 10
- Tamaño de palabra: defecto = 11 para nucleótidos/ 28 para megablast/ 3 para proteínas

Por otra parte, el grado de identidad (homología) local entre la incógnita de secuencia de aminoácidos o la incógnita de 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 una coincidencia por encima de un cierto umbral. Según esto, el programa calcula la identidad solamente para estos segmentos coincidentes. Por lo tanto, la identidad calculada de este modo se denomina identidad local.

#### **Vectores**

Otro aspecto de la invención trata de vectores, incluyendo vectores de clonación y expresión, que comprenden un polinucleótido de la invención que codifica una proteína TEMER09484 o uno de sus equivalentes funcionales 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 a 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 lleven a cabo 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 polinucleótido de la invención puede ser cualquier vector que se pueda someter convenientemente a procedimientos de ADN recombinante, y la elección del vector dependerá a menudo de la célula hospedadora en la que se vaya a introducir.

5 Un vector según la invención puede ser un vector autónomamente replicativo, 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 uno que, cuando se introduce en una célula hospedadora, se integra en el genoma de la célula hospedadora y se replica junto con el cromosoma o los cromosomas en los que se ha integrado.

10 Un tipo de vector es un "plásmido", que se refiere a un ciclo de ADN bicatenario circular al que se pueden ligar segmentos de ADN adicionales. Otro tipo de vector es un vector viral, en el que se pueden ligar segmentos de ADN adicionales 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 un origen de replicación bacteriano y vectores de mamífero episómicos). Otros vectores (p. ej., vectores de mamífero no episómicos) se integran en el genoma de una célula  
15 hospedadora tras la introducción 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 enlazados operativamente. Estos 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 forma de plásmidos. Los términos "plásmido" y "vector" se pueden usar intercambiamente en la presente ya que el plásmido es la forma de vector  
20 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 cuanto a la replicación) y vectores fágicos que realizan funciones equivalentes.

25 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.

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.

30 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 basándose en las células hospedadoras que se van a usar para la expresión, que están enlazadas operativamente a la secuencia de ácido nucleico que se va a expresar.

35 Dentro de un vector, tal como un vector de expresión, "enlazado operativamente" se pretende indicar que la secuencia nucleotídica de interés está enlazada a la secuencia o las 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 "enlazado operativamente" se refiere a una yuxtaposición en la que los componentes descritos están en una relación que los permita funcionar en su modo  
40 pretendido. Una secuencia reguladora tal como un promotor, potenciador u otra señal de regulación de la expresión "enlazado operativamente" a una secuencia codificante está situado de tal modo que la expresión de la secuencia codificante se alcance bajo una condición compatible con las secuencias de control o las secuencias están dispuestas de modo que funcionen conjuntamente para su propósito pretendido, por ejemplo la transcripción se inicia en un promotor y avanza a través de secuencia de ADN que codifica el polipéptido.  
45

Un vector o constructo de expresión para una célula hospedadora dada puede comprender así los siguientes elementos enlazados operativamente entre sí en un orden consecutivo del extremo 5' al extremo 3' con relación a la hebra codificante de la secuencia que codifica el polipéptido de la primera 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 señalización capaz de dirigir la secreción del polipéptido procedente de la célula  
50 hospedadora dada en 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 como celobiohidrolasa; y preferiblemente también (4) una región de terminación de la transcripción (terminador) capaz de terminar la transcripción aguas abajo de la secuencia nucleotídica que codifica el polipéptido.  
55

Aguas abajo de la secuencia nucleotídica según la invención puede existir 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. Por ejemplo, el terminador puede ser natural a la secuencia de ADN que codifica el polipéptido. Sin embargo,  
60 preferiblemente, se usa un terminador de levadura en células hospedadoras de levadura y se usa un terminador de hongo filamento en células hospedadoras de hongo 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 los constructos incluirán un AUG iniciador de la traducción al principio y un codón  
65 de terminación situado apropiadamente en el extremo del polipéptido que se va a traducir.

Una expresión potenciada del polinucleótido de la invención también se puede alcanzar mediante la selección de regiones reguladoras heterólogas, p. ej. regiones promotoras, líderes de secreción y/o terminadoras, que pueden servir para incrementar los niveles de expresión y, si se desea, secreción de la proteína de interés desde el hospedador de expresión y/o para proporcionar el control inducible de la expresión de un polipéptido de la invención.

Será apreciado por los expertos en la especialidad 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 proteína 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. proteínas TEMER09484, formas mutantes de proteínas TEMER09484, sus fragmentos, variantes o equivalentes funcionales. Los vectores, tales como vectores de expresión recombinantes, de la invención se pueden diseñar para la expresión de proteínas TEMER09484 en células procarióticas o eucarióticas.

Por ejemplo, las proteínas TEMER09484 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.

Medios y condiciones de cultivo apropiados para las células hospedadoras descritas anteriormente son conocidos en la especialidad.

El vector de expresión recombinante se puede transcribir y traducir *in vitro*, por ejemplo usando secuencias reguladoras promotoras de T7 y polimerasa de T7.

Para la mayoría de los hongos filamentosos y las levaduras, el vector o constructo de expresión preferiblemente se integra 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 el constructo 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 los constructos de expresión estén integrados en el genoma de la célula hospedadora, los constructos bien están integrados en locus aleatorios del genoma, o bien en locus diana predeterminados usando recombinación homóloga, en cuyo caso los locus diana 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, bacteriófago, episoma de levadura, elementos cromosómicos de levadura, virus tales como baculovirus, papovavirus, virus vacunales, adenovirus, poxvirus aviáres, virus de pseudorrabia y retrovirus, y vectores derivados de sus combinaciones, tales como los derivados de elementos genéticos plasmídicos y bacteriofágicos, tales como cósmidos y fagémidos.

El término "secuencias de control" o "secuencias reguladoras" se define en la presente para incluir al menos cualquier 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 (según se describe en Kozak, 1991, *J. Biol. Chem.* 266:19867-19870), una secuencia de señalización de la 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 parada de la transcripción y la traducción. Como se indica anteriormente, el término "enlazado operativamente" se define en la presente como una configuración en la que una secuencia de control está situada apropiadamente en una posición relativa a la secuencia codificante de la secuencia de ADN de modo que la secuencia de control dirija la producción de un polipéptido.

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

La secuencia de control puede ser una secuencia promotora apropiada, una secuencia de ácido nucleico, que es 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 muestre actividad de transcripción en la célula, incluyendo promotores mutantes, truncados e híbridos, y se puede obtener a partir de genes que codifican polipéptidos extracelulares o intracelulares bien homólogos o bien heterólogos a la célula.

El término "promotor" se define en la presente como una secuencia de ADN que se une a ARN polimerasa y dirige la polimerasa al sitio de inicio de la transcripción aguas abajo correcto de una secuencia de ácido nucleico que codifica un compuesto biológico para iniciar la transcripción. La ARN polimerasa cataliza eficazmente el ensamblaje de ARN mensajero complementario a la hebra de ADN apropiada de una región codificante. Se entenderá además que el término "promotor" incluye la región no codificante 5' (entre el promotor y el inicio de la traducción) para la traducción después de la transcripción en ARNm, elementos de control de la transcripción de acción cis tales como potenciadores, y otras secuencias nucleotídicas capaces de interactuar con factores de transcripción. El promotor puede ser cualquier secuencia promotora apropiada adecuada para una célula hospedadora eucariótica o procarionótica, que muestre actividad de transcripción, incluyendo promotores mutantes, truncados e híbridos, y se puede obtener de polinucleótidos que codifican polipéptidos extracelulares o intracelulares bien homólogos (naturales) o bien heterólogos (extraños) a la célula. El promotor puede ser un promotor constitutivo o inducible.

Preferiblemente, el promotor es un promotor inducible. Más preferiblemente, el promotor es un promotor inducible por carbohidrato. Promotores inducibles por carbohidrato que se usan preferiblemente se seleccionan de un promotor inducible por almidón (es decir un promotor inducible mediante almidón, uno de sus monómeros, uno de sus dímeros, uno de sus oligómeros, tal como, por ejemplo, un promotor inducible por maltosa, un promotor inducible por isomaltosa), un promotor inducible por celulosa (es decir un promotor inducible mediante celulosa, uno de sus monómeros, uno de sus dímeros y/o uno de sus oligómeros, tal como, por ejemplo, un promotor inducible por celobiosa, un promotor inducible por soforosa), un promotor inducible por hemicelulosa (es decir un promotor inducible mediante hemicelulosa, uno de sus monómeros, uno de sus dímeros y/o uno de sus oligómeros, tal como, p. ej., un promotor inducible por xilano, un promotor inducible por arabinosa, un promotor inducible por xilosa), un promotor inducible por pectina (es decir un promotor inducible mediante pectina, uno de sus monómeros, uno de sus dímeros y/o uno de sus oligómeros tal como, por ejemplo, un promotor inducible por ácido galacturónico, un promotor inducible por ramnosa), un promotor inducible por arabinano (es decir un promotor inducible mediante arabinano, uno de sus monómeros, uno de sus dímeros y/o uno de sus oligómeros tal como, por ejemplo, un promotor inducible por arabinosa), un promotor inducible por glucosa, un promotor inducible por lactosa, un promotor inducible por galactosa. Otros promotores inducibles son promotores inducibles por cobre, ácido oleico.

Promotores adecuados en hongos filamentosos son promotores que se pueden seleccionar del grupo que incluye, pero no se limita a, promotores obtenidos de los polinucleótidos que codifican TAKA amilasa de *A. oryzae*, proteinasa aspártica de *Rhizomucor miehei*, promotor *gpdA* de *Aspergillus*, alfa-amilasa neutra de *A. niger*, alfa-amilasa estable a los ácidos de *A. niger*, glucoamilasa (*glaA*) de *A. niger* o *A. awamori*, endoxilanasas (*xlnA*) o beta-xilosidasas (*xlnD*) de *A. niger* o *A. awamori*, celobiohidrolasa I (*CBHI*) de *T. reesei*, lipasa de *R. miehei*, proteasa alcalina de *A. oryzae*, fosfato de triosa isomerasa de *A. oryzae*, acetamidasa de *A. nidulans*, amiloglucosidasa de *Fusarium venenatum* (documento WO 00/56900), Dania de *Fusarium venenatum* (documento WO 00/56900), Quinn de *Fusarium venenatum* (documento WO 00/56900), proteasa tripsinoide de *Fusarium oxysporum* (documento WO 96/00787), beta-glucosidasa de *Trichoderma reesei*, celobiohidrolasa I de *Trichoderma reesei*, celobiohidrolasa II de *Trichoderma reesei*, endoglucanasa I de *Trichoderma reesei*, endoglucanasa II de *Trichoderma reesei*, endoglucanasa III de *Trichoderma reesei*, endoglucanasa IV de *Trichoderma reesei*, endoglucanasa V de *Trichoderma reesei*, xilanasas I de *Trichoderma reesei*, xilanasas II de *Trichoderma reesei*, beta-xilosidasas de *Trichoderma reesei*, así como el promotor NA2-tpi (un híbrido de los promotores procedentes de los polinucleótidos que codifican alfa-amilasa neutra de *A. niger* y fosfato de triosa isomerasa de *A. oryzae*), y sus promotores mutantes, truncados e híbridos. Otros ejemplos de promotores son los promotores descritos en los documentos WO2006/092396 y WO2005/100573. Un ejemplo más del uso de promotores se describe en el documento WO2008/098933. Promotores inducibles por carbohidrato preferidos que se pueden usar en hongos filamentosos son la TAKA amilasa de *A. oryzae*, alfa-amilasa neutra de *A. niger*, alfa-amilasa estable a los ácidos de *A. niger*, glucoamilasa (*glaA*) de *A. niger* o *A. awamori*, endoxilanasas (*xlnA*) o beta-xilosidasas (*xlnD*) de *A. niger* o *A. awamori*, beta-glucosidasa de *Trichoderma reesei*, celobiohidrolasa I de *Trichoderma reesei*, celobiohidrolasa II de *Trichoderma reesei*, endoglucanasa I de *Trichoderma reesei*, endoglucanasa II de *Trichoderma reesei*, endoglucanasa III de *Trichoderma reesei*, endoglucanasa IV de *Trichoderma reesei*, endoglucanasa V de *Trichoderma reesei*, xilanasas I de *Trichoderma reesei*, xilanasas II de *Trichoderma reesei*, así como el promotor NA2-tpi (un híbrido de los promotores procedentes de los polinucleótidos que codifican alfa-amilasa neutra de *A. niger* y fosfato de triosa isomerasa de *A. oryzae*) según se define anteriormente.

Ejemplos de estos promotores procedentes de microorganismos grampositivos incluyen, pero no se limitan a, *gnt* (promotor de operón de gluconato); *penP* de *Bacillus licheniformis*; *glnA* (glutamina sintetasa); *xilAB* (operón de xilosa); *araABD* (operón de L-arabinosa) y promotor *Pspac*, un promotor *SPO1/lac* híbrido que puede estar controlado por inductores tales como isopropil- $\beta$ -D-tiogalactopiranosido [IPTG] ((Yansura D.G., Henner D.J. Proc Natl Acad Sci U S A. 1984 81(2):439-443). Los activadores también son proteínas que se unen a ADN específicas de una secuencia que inducen actividad promotora. Ejemplos de estos promotores procedentes de microorganismos grampositivos incluyen, pero no se limitan a, sistemas de dos componentes (*PhoP-PhoR*, *DegU-DegS*, *Spo0A-Phosphorelay*), *LevR*, *Mry* y *GltC*. (ii) La producción de factores sigma secundarios puede ser responsable principalmente de la transcripción a partir de promotores específicos. Ejemplos de microorganismos grampositivos incluyen, pero no se limitan a, los promotores activados mediante factores sigma específicos de la esporulación:  $\sigma^F$ ,  $\sigma^E$ ,  $\sigma^G$  y  $\sigma^K$  y factor sigma de estrés general,  $\sigma^B$ . La respuesta mediada por  $\sigma^B$  es inducida por limitación de energía y estreses ambientales (Hecker M, Völker U. Mol Microbiol. 1998; 29(5):1129-1136.). (iii) La atenuación y la antiterminación también regulan la

transcripción. Ejemplos de microorganismos grampositivos incluyen, pero no se limitan a, operón trp y gen sacB. (iv) Otros promotores regulados en vectores de expresión se basan en el sistema regulador sacR que confiere inducibilidad por sacarosa (Klier AF, Rapoport G. Annu Rev Microbiol. 1988; 42:65-95).

5 Promotores inducibles adecuados útiles en bacterias, tales como *Bacilli*, incluyen: promotores procedentes de microorganismos grampositivos tales como, pero no limitados a, SP01-26, SP01-15, veg, pyc (promotor de piruvato carboxilasa) y amyE. Ejemplos de promotores procedentes de microorganismos gramnegativos incluyen, pero no se limitan a, tac, tet, trp-tet, lpp, lac, lpp-lac, lacIq, T7, T5, T3, gal, trc, ara, SP6,  $\lambda$ -PR y  $\lambda$ -PL.

10 Ejemplos adicionales de promotores útiles en células bacterianas, tales como *Bacilli*, incluyen los promotores de  $\alpha$ -amilasa y SPo2 así como promotores procedentes de genes de proteasa extracelulares.

Otro ejemplo de un promotor adecuado es el promotor obtenido del operón lac de *E. coli*. Otro ejemplo es el promotor del gen de agarasa de *Streptomyces coelicolor* (dagA). Otro ejemplo es el promotor del gen de proteasa alcalina de *Bacillus lentus* (aprH). Otro ejemplo es el promotor del gen de proteasa alcalina de *Bacillus licheniformis* (gen de Carlsberg de subtilisina). Otro ejemplo es el promotor del gen de levansucrasa de *Bacillus subtilis* (sacB). Otro ejemplo es el promotor del gen de alfa-amilasa de *Bacillus subtilis* (amyF). Otro ejemplo es el promotor del gen de alfa-amilasa de *Bacillus licheniformis* (amyL). Otro ejemplo es el promotor del gen de amilasa maltogénica de *Bacillus stearothermophilus* (amyM). Otro ejemplo es el promotor del gen de alfa-amilasa de *Bacillus amyloliquefaciens* (amyQ). Otro ejemplo es un promotor de "consenso" que tiene la secuencia TTGACA para la región "-35" y TATAAT para la región "-10". Otro ejemplo es el promotor del gen de penicilinas de *Bacillus licheniformis* (penP). Otro ejemplo son los promotores de los genes xilA y xilB de *Bacillus subtilis*.

Preferiblemente, la secuencia promotora es de un gen altamente expresado. Ejemplos de genes altamente expresados preferidos de los que se pueden seleccionar promotores y/o que están comprendidos en locus diana predeterminados preferidos para la integración de constructos de expresión incluyen pero no se limitan a genes que codifican enzimas glucolíticas tales como fosfato de triosa isomerasas (TPI), fosfato de gliceraldehído deshidrogenasas (GAPDH), fosfoglicerato cinasas (PGK), piruvato cinasas (PYK o PKI), alcohol deshidrogenasas (ADH), así como genes que codifican amilasas, glucoamilasas, proteasas, xilanasas, celobiohidrolasas,  $\beta$ -galactosidasas, alcohol (metanol) oxidasas, factores de elongación y proteínas ribosómicas. Ejemplos específicos de genes altamente expresados adecuados incluyen, p. ej., el gen LAC4 procedente de *Kluyveromyces* sp., los genes de metanol oxidasa (AOX y MOX) procedentes de *Hansenula* y *Pichia*, respectivamente, los genes de glucoamilasa (glaA) procedentes de *A. niger* y *A. awamori*, el gen de TAKA-amilasa de *A. oryzae*, el gen gdpA de *A. nidulans* y los genes de celobiohidrolasa de *T. reesei*.

Promotores que se pueden usar en levadura incluyen, p. ej., promotores procedentes de genes glucolíticos, tales como los promotores de fosfofructocinasa (PFK), fosfato de triosa isomerasa (TPI), 3-fosfato de gliceraldehído deshidrogenasa (GPD, TDH3 o GAPDH), piruvato cinasa (PYK), promotores de fosfoglicerato cinasa (PGK) procedentes de levaduras u hongos filamentosos; más detalles acerca de estos promotores procedentes de levadura se pueden encontrar en (el documento WO 93/03159). Otros promotores útiles son promotores de genes que codifican proteínas ribosómicas, promotor del gen de lactasa (LAC4), promotores de alcohol deshidrogenasa (ADHI, ADH4 y similares) y el promotor de enolasa (ENO). Otros promotores, tanto constitutivos como inducibles, y potenciadores o secuencias activadoras aguas arriba serán conocidos por los expertos en la especialidad. Los promotores usados en las células hospedadoras de la invención se pueden modificar, si se desea, para afectar a sus características de control. Promotores adecuados en este contexto incluyen promotores naturales tanto constitutivos como inducibles así como promotores manipulados, que son muy conocidos para el experto en la especialidad. Promotores adecuados en células hospedadoras eucarióticas pueden ser *GAL7*, *GAL10*, o *GAL1*, *CYC1*, *HIS3*, *ADH1*, *PGL*, *PH05*, *GAPDH*, *ADC1*, *TRP1*, *URA3*, *LEU2*, *ENO1*, *TPI1* y *AOX1*. Otros promotores adecuados incluyen *PDC1*, *GPD1*, *PGK1*, *TEF1* y *TDH3*. Ejemplos de promotores inducibles por carbohidratos que se pueden usar son promotores GAL, tales como promotores GAL1 o GAL10.

Todos los susodichos promotores están fácilmente disponibles en la especialidad.

La transcripción del ADN que codifica los polipéptidos de la presente invención por eucariotas superiores se puede incrementar al insertar una secuencia potenciadora en el vector. Los potenciadores son elementos de acción cis de ADN, habitualmente de aproximadamente 10 a 300 pb, que actúan para incrementar la actividad de transcripción de un promotor en un tipo de célula hospedadora dado. Ejemplos de potenciadores incluyen el potenciador de SV40, que está situado en el último lado del origen de replicación en los pb 100 a 270, el potenciador del promotor temprano de citomegalovirus, el potenciador de polioma en el último lado del origen de replicación y potenciadores de adenovirus.

La secuencia de control también puede ser una secuencia terminadora de la transcripción adecuada, una secuencia reconocida por una célula de hongo filamentosos para terminar la transcripción. La secuencia terminadora está enlazada operativamente al extremo 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.

65

La secuencia de control también puede ser un terminador. Terminadores preferidos para células de hongos filamentosos se obtienen 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*.

La secuencia de control también puede incluir una secuencia líder adecuada, una región no traducida de un ARNm que es importante para la traducción por la célula de hongo filamentosos. La secuencia líder está enlazada operativamente al extremo 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 fosfato de triosa isomerasa de *A. nidulans* y glaA de *A. niger*. Otras secuencias preferidas se aíslan y/o se divulgan en el documento WO2006/077258.

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

La secuencia de control también puede ser una secuencia de poliadenilación, una secuencia que está enlazada operativamente al extremo 3' de la secuencia de ácido nucleico y que, cuando se transcribe, es reconocida por la célula de hongo filamentosos como una señal para añadir residuos de poliadenosina a RNAm 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 a partir de los genes que codifican TAKA amilasa de *A. oryzae*, glucoamilasa de *A. niger*, antranilato sintasa de *A. nidulans*, proteasa tripsinoide de *Fusarium oxysporum* y alfa-glucosidasa de *A. niger*.

Cuando el polipéptido según la invención se va a secretar de la célula hospedadora en el medio de cultivo, se puede añadir al polipéptido una secuencia de señalización apropiada a fin de dirigir el polipéptido sintetizado de novo a la ruta de secreción de la célula hospedadora. El experto en la especialidad conoce la selección de una secuencia de señalización apropiada para un hospedador específico. La secuencia de señalización puede ser natural a la célula hospedadora o puede ser extraña a la célula hospedadora. Como un ejemplo, se puede usar una secuencia de señalización procedente de una proteína natural a la célula hospedadora. Preferiblemente, dicha proteína natural es una proteína altamente secretada, es decir una proteína que se secreta en cantidades superiores al 10% de la cantidad total de proteína que se secreta. Las secuencias de señalización usadas preferiblemente según la invención son, por ejemplo: pmeA.

Como una alternativa para una secuencia de señalización, el polipéptido de la invención se puede fusionar a una proteína portadora secretada, o una de sus partes. Este constructo quimérico se dirige a la ruta de secreción por medio de la secuencia de señalización de la proteína portadora, o una de sus partes. Además, la proteína portadora proporcionará un efecto estabilizante al polipéptido según la invención y/o puede potenciar la solubilidad. Esta proteína portadora puede ser cualquier proteína. Preferiblemente, se usa una proteína altamente secretada como una proteína portadora. La proteína portadora puede ser natural o extraña al polipéptido según la invención. La proteína portadora puede ser natural o puede ser extraña a la célula hospedadora. Ejemplos de estas proteínas portadoras son glucoamilasa, secuencia prepro de factor alfa-Mating, dominio de unión a celulosa de proteína A de unión a celulosa de *Clostridium cellulovorans*, glutatión 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. Una proteína portadora preferida para la expresión de este constructo quimérico en células de *Aspergillus* es la glucoamilasa. La proteína portadora y el polipéptido según la invención 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 para el experto en la especialidad.

Se puede usar una secuencia de señalización para facilitar la secreción y el aislamiento de una proteína o un polipéptido de la invención. Las secuencias de señalización se caracterizan típicamente por un núcleo de aminoácidos hidrófobos, que generalmente se escinden de la proteína madura durante la secreción en uno o más episodios de escisión. Estos péptidos de señalización contienen sitios de procesamiento que permiten la escisión de la secuencia de señalización de las proteínas maduras a medida que pasan a través de la vía secretora. La secuencia de señalización dirige la secreción de la proteína, tal como a partir de un hospedador eucariota en el que se transforma el vector de expresión, y la secuencia de señalización se escinde posteriormente o simultáneamente. A continuación, la proteína se puede purificar fácilmente del medio extracelular mediante métodos conocidos. Alternativamente, la secuencia de señalización se puede enlazar a la proteína de interés usando una secuencia que facilite 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 hexa-histidina, tal como la etiqueta proporcionada en un vector pQE (Qiagen, Inc.), entre otros, muchos de los 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 hexa-histidina proporciona una purificación conveniente de la proteína de fusión. La etiqueta de

HA es otro péptido útil para la purificación que corresponde a un epítipo derivado de proteína de hemaglutinina de gripe, que ha sido descrito por Wilson y cols., Cell 37:767 (1984), a modo de ejemplo.

Preferiblemente, una proteína de fusión TEMER09484 de la invención se produce mediante técnicas de ADN recombinante estándar. Por ejemplo, fragmentos de ADN que codifican las diferentes secuencias polipeptídicas se ligan conjuntamente en el marco según técnicas convencionales, por ejemplo al emplear terminaciones de extremos romos o extremos escalonados para la ligación, digestión con enzimas de restricción para proporcionar terminaciones apropiadas, relleno de extremos cohesivos según sea apropiado, tratamiento con fosfatasa alcalina para evitar un empalme 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 pueden 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 comercialmente muchos vectores de expresión que ya codifican un resto de fusión (p. ej., un polipéptido de GST). Un ácido nucleico que codifica TEMER09484 se puede clonar en uno de estos vectores de expresión de modo que la proteína de fusión esté enlazada en el marco a la proteína TEMER09484.

### **(Sobre)expresión**

En una realización preferida, los polinucleótidos de la presente invención según 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 comprende secuencias homólogas a secuencias genómicas eucarióticas o secuencias genómicas virales. Esto permitirá 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 diana 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 diana predeterminado en el genoma de la célula hospedadora para orientar la integración del vector de clonación hacia este locus predeterminado. A fin de promover la integración orientada, preferiblemente, el vector de clonación 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 diana. La longitud de las secuencias homólogas que flanquean el locus diana 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 orientada, 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 inactivados funcionalmente 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 una sección mayor de ADN genómico que contiene las regiones promotora y terminadora del gen, que a continuación puede funcionar como flanco 5' y flancos 3'.

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 contienen conjuntamente 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.

El ADN vectorial se puede introducir en células procarióticas o eucarióticas a través de técnicas de transformación o transfección convencionales. Según se usan en la presente, los términos "transformación" y "transfección" están destinados a referirse a una variedad de técnicas reconocidas en la especialidad para introducir 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.

El experto en la especialidad sabe como transformar células con el uno o más casetes de expresión y el marcador seleccionable. Por ejemplo, el experto puede usar uno o más vectores de expresión, en donde el uno o más vectores de clonación comprenden los casetes de expresión y el marcador seleccionable.

La transformación de la célula hospedadora microbiana mutante se puede efectuar mediante cualesquiera métodos conocidos adecuados, incluyendo, p. ej., métodos de electroporación, bombardeo de partículas o bombardeo de microproyectiles, métodos protoplásticos y transformación mediada por *Agrobacterium* (AMT). Preferiblemente, se usa el método protoplástico. Procedimientos para la transformación son descritos por J.R.S. Fincham, Transformation in fungi. 1989, Microbiological reviews. 53, 148-170.

La transformación puede implicar un procedimiento que consiste en la formación de protoplastos, la transformación de los protoplastos y la regeneración de la pared celular de un modo conocido de por sí. Procedimientos para la transformación de células de *Aspergillus* se describen en el documento EP 238 023 y Yelton y cols., 1984, Proceedings of the National Academy of Sciences USA 81:1470-1474. Procedimientos adecuados para la transformación de *Aspergillus* y otras células hospedadoras de hongos filamentosos usando *Agrobacterium tumefaciens* se describen, p. ej., en De Groot y cols., *Agrobacterium tumefaciens*-mediated transformation of filamentous fungi. Nat Biotechnol. 1998, 16:839-842. Erratum en: Nat Biotechnol 1998 16:1074. Un método adecuado para transformar especies de *Fusarium* es descrito por Malardier y cols., 1989, Gene 78:147156 o en el documento WO 96/00787. Se pueden aplicar otros métodos tales como un método que usa transformación biolística como el descrito en: Christiansen y cols., *Biolistic transformation of the obligate plant pathogenic fungus, Erysiphe graminis f.sp. hordei*. 1995, Curr Genet. 29:100-102. Una levadura se puede transformar usando los procedimientos descritos por Becker y Guarente, En Abelson, J. N. y Simon, M. I., editores, Guide to Yeast Genetics and Molecular Biology, Methods in Enzymology, Volumen 194, pp 182-187, Academic Press, Inc., Nueva York; Ito y cols., 1983, Journal of Bacteriology 153: 163; y Hinnen y cols., 1978, Proceedings of the National Academy of Sciences USA 75: 1920.

A fin de aumentar la cantidad de copias del polinucleótido que codifica el compuesto de interés o que codifica un compuesto implicado en la producción por la célula del compuesto de interés (el gen) en la célula hospedadora microbiana mutada, se pueden requerir múltiples transformaciones. De este modo, se puede influir en las relaciones de las diferentes enzimas producidas por la célula hospedadora. Además, un vector de expresión puede comprender múltiples casetes de expresión para incrementar la cantidad de copias del polinucleótido o los polinucleótidos que se van a transformar.

Otro modo podría ser elegir diferentes secuencias de control para los diferentes polinucleótidos, que – dependiendo de la elección – pueden provocar una producción superior o inferior del polipéptido o los polipéptidos deseados.

Las células transformadas con el marcador seleccionable se pueden seleccionar basándose en la presencia del marcador seleccionable. Habitualmente, en el caso de la transformación de células (de *Aspergillus*), cuando la célula se transforma con todo el material de ácido nucleico al mismo tiempo, cuando el marcador seleccionable está presente, también están presentes el polinucleótido o los polinucleótidos que codifican el polipéptido o los polipéptidos deseados.

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 de 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 ADNcs (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.

Otros marcadores incluyen ATP sintetasa, subunidad 9 (oliC), 5'-fosfato de orotidina descarboxilasa (pvrA), el gen de resistencia a G418 bacteriano (este también se puede usar en levadura, pero no en hongos), el gen de resistencia a ampicilina (*E. coli*), el gen de resistencia a neomicina (*Bacillus*) y el gen uidA de *E. coli*, que codifica  $\beta$ -glucuronidasa

(GUS). Los vectores se pueden usar in vitro, por ejemplo para la producción de ARN o se pueden usar para transfectar o transformar una célula hospedadora.

La expresión de proteínas en procariontes 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 que no son de fusión. Los vectores de fusión añaden un número de aminoácidos a una proteína codificada en los mismos, p. ej. a la terminación amino de la proteína recombinante. Estos vectores de fusión tienen típicamente tres propósitos: 1) incrementar la expresión de proteína recombinante; 2) incrementar la solubilidad de la proteína recombinante; y 3) ayudar a 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 del resto de fusión y la proteína recombinante para potenciar la separación de la proteína recombinante del resto de fusión después de la purificación de la proteína de fusión.

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

Vectores preferidos para el uso en bacterias se describen, por ejemplo, en el documento WO-A1-2004/074468, que de ese modo son incluidos por la presente mediante referencia. Otros vectores adecuados serán fácilmente evidentes para el experto.

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 TEMER09484 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 a la terminación 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 NO: 32 y una secuencia de aminoácidos obtenible al expresar el polinucleótido de SEQ ID NO: 31 o SEQ ID NO: 34 o SEQ ID NO: 35 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 específicas en el péptido o polipéptido, respectivamente. Según esto, un péptido de señalización variante es un péptido de señalización 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 específicas en el péptido de señalización.

El término "polinucleótido" es idéntico al término "molécula de ácido nucleico" y se pueden leer intercambiamente en la presente. El término se refiere a una molécula de polinucleótido, que es una molécula de ácido ribonucleico (ARN) o ácido desoxirribonucleico (ADN), bien monocatenaria o bien bicatenaria. Un polinucleótido bien puede estar presente en forma aislada o bien estar comprendido en moléculas de ácido nucleico o vectores 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 del polinucleótido.

Los términos "péptido" y "oligopéptido" se consideran sinónimos (como se reconoce comúnmente) y cada término 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 de la terminación amino a la terminación carboxi. El código de una letra usado en la presente es comúnmente conocido en la especialidad 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, polipéptidos y 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).

5 La proteína Temer05249 según la invención se puede recuperar y purificar de cultivos celulares recombinantes mediante métodos conocidos en la especialidad. Lo más preferiblemente, se emplea para la purificación cromatografía de líquidos de alta resolución ("HPLC").

10 Polipéptidos de la presente invención incluyen productos purificados naturalmente, productos de procedimientos sintéticos químicos 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 estar glicosilados o pueden estar 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 presenta fragmentos biológicamente activos de los polipéptidos según la invención.

20 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 de la proteína Temer05249 (p. ej., la secuencia de aminoácidos de SEQ ID NO: 32), que incluyen menos aminoácidos que la proteína de longitud completa pero que exhiben al menos una actividad biológica de la proteína de longitud completa correspondiente. Típicamente, los fragmentos biológicamente activos comprenden un dominio o motivo con al menos una actividad de la proteína Temer05249.

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

30 Por otra parte, otras porciones biológicamente activas, en las que se eliminan otras regiones de la proteína, 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. La invención también presenta fragmentos de ácido nucleico que codifican los fragmentos biológicamente activos anteriores de la proteína Temer05249.

### 35 **Proteínas**

En otro aspecto de la invención, se proporcionan proteínas Temer05249 mejoradas. Proteínas Temer05249 mejoradas son proteínas en las que se mejora al menos una actividad biológica. Estas proteínas se pueden obtener al introducir aleatoriamente mutaciones a lo largo de la totalidad o parte de la secuencia codificante de Temer05249, tal como mediante mutagénesis de saturación, y los mutantes resultantes se pueden expresar recombinantemente y cribar con respecto a una actividad biológica. A modo de ejemplo, la especialidad proporciona ensayos estándar para medir la actividad enzimática de la proteína de la invención y así se pueden seleccionar fácilmente proteínas mejoradas.

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

1. PCR propensa a errores para introducir mutaciones aleatorias, seguido por un cribado de las variantes obtenidas y un aislamiento de las variantes con propiedades cinéticas mejoradas

2. Redistribución familiar de variantes relacionadas de los genes que codifican la celobiohidrolasa, seguida por un cribado de las variantes obtenidas y un aislamiento de las variantes con propiedades cinéticas mejoradas

50 Variantes de los genes de la presente invención que conducen a un nivel incrementado de ARNm y/o proteína, dando como resultado más de una actividad según la Tabla 1, se pueden obtener mediante las secuencias polinucleotídicas de dichos genes. Entre estas modificaciones se incluyen:

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

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

3. Adición de secuencias estabilizantes de información genómica que codifican la celobiohidrolasa dando como resultado moléculas de ARNm con una semivida incrementada

5 Métodos preferidos para aislar variantes con propiedades catalíticas mejoradas o niveles incrementados de ARNm o proteína 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 la celobiohidrolasa de la invención se describen en el documento WO2005/059149.

10 En una realización preferida, la proteína de la invención tiene una secuencia de aminoácidos según SEQ ID NO: 32. En otra realización, el polipéptido de la invención es sustancialmente homólogo a la secuencia de aminoácidos según SEQ ID NO: 32 y retiene al menos una actividad biológica de un polipéptido según SEQ ID NO: 32 y sin embargo difiere en la secuencia de aminoácidos debido a variación natural o mutagénesis según se describe.

15 En una realización preferida adicional, la proteína de la invención tiene una secuencia de aminoácidos codificada por un fragmento de ácido nucleico aislado capaz de hibridarse a un ácido nucleico según SEQ ID NO: 31 o SEQ ID NO: 34 o SEQ ID NO: 35, preferiblemente bajo condiciones de hibridación muy rigurosas.

20 Según esto, la proteína Temer05249 o la proteína de la invención es preferiblemente una proteína que comprende una secuencia de aminoácidos al menos 50%, al menos 60%, al menos 65%, al menos 70%, al menos 75%, al menos 80%, al menos 85%, al menos 90%, al menos 91%, 92%, 93%, 94%, 95%, 96%, 95%, 96%, 97%, 98%, 97%, 98%, 99%, 99,8%, 99,9% o más homóloga a la secuencia de aminoácidos mostrada en SEQ ID NO: 32 y, típicamente, retiene al menos una actividad funcional del polipéptido según SEQ ID NO: 32.

25 Según un aspecto de la invención, el polipéptido de la invención puede comprender la secuencia de aminoácidos indicada en SEQ ID NO: 32 o una secuencia de aminoácidos que difiere en 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 o 12 aminoácidos de la secuencia de aminoácidos indicada en SEQ ID NO: 32 y de ese modo el polipéptido todavía tiene la actividad o función del polipéptido de la invención. El experto apreciará que estos pequeños cambios de aminoácidos en el polipéptido de la invención pueden estar presentes (por ejemplo mutaciones presentes en la naturaleza) o se pueden elaborar (por ejemplo usando tecnología de ADNr) sin pérdida de la función o actividad de la proteína. En caso de que estas mutaciones estén presentes en un dominio de unión, un sitio activo u otro dominio funcional del polipéptido, una propiedad del polipéptido puede cambiar (por ejemplo su termoestabilidad) pero el polipéptido puede mantener su actividad como hemicelulasa. En caso de que esté presente una mutación que no esté cerca del sitio activo, el dominio de unión u otro dominio funcional, se puede esperar menos efecto.

35 También se pueden identificar equivalentes funcionales de una proteína según la invención, p. ej. al cribar bibliotecas combinatorias de mutantes, p. ej. mutantes de truncamiento, de la proteína de la invención con respecto a una actividad según la Tabla 1. En una realización, se genera una biblioteca variegada de variantes mediante mutagénesis combinatoria a nivel de ácido nucleico. 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 conjunto degenerado de secuencias proteínicas potenciales sea expresable como polipéptidos individuales o, alternativamente, como un conjunto de proteínas de fusión mayores (p. ej. para presentación en 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. Métodos para sintetizar oligonucleótidos degenerados son conocidos en la especialidad (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).

50 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 cribar una selección posterior de variantes. Por ejemplo, se puede generar una biblioteca de fragmentos de secuencias codificantes al tratar un fragmento de PCR bicatenario 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 bicatenario, renaturalizar el ADN para formar ADN bicatenario que puede incluir pares de sentido/antisentido procedentes de diferentes productos mellados, retirar porciones monocatenarias de dúplex reformados mediante tratamiento con nucleasa S1, 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.

60 Se conocen en la especialidad varias técnicas para cribar productos génicos de bibliotecas combinatorias elaboradas mediante mutaciones puntuales de truncamiento, y para cribar bibliotecas de ADNc para productos génicos que tienen una propiedad seleccionada. Las técnicas más ampliamente usadas, que son propensas al análisis de alto rendimiento, para cribar grandes bibliotecas génicas 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 en 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 una proteína 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).

5 Además de la secuencia génica de Temer05249 mostrada en SEQ ID NO: 31 o en SEQ ID NO: 34 o en SEQ ID NO: 35, será evidente para el experto en la especialidad que pueden existir polimorfismos de secuencias de ADN dentro de una población dada, lo que puede conducir a cambios en la secuencia de aminoácidos de la proteína Temer05249. Estos polimorfismos genéticos pueden existir en células procedentes de diferentes poblaciones o dentro de una población debido a variación alélica natural. Las variantes alélicas también pueden incluir equivalentes funcionales.

10 Los fragmentos de un polinucleótido según la invención también pueden comprender polinucleótidos que no codifican polipéptidos funcionales. Estos polinucleótidos pueden funcionar como sondas o cebadores para una reacción de PCR.

15 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 tienen una actividad de Temer05249 incluyen, entre otros, (1) aislar el gen que codifica la proteína Temer05249, o sus variantes alélicas, 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 Temer05249 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 Temer05249 en tejidos y/o células específicos y 4) sondas y cebadores que se pueden usar como una herramienta diagnóstica para analizar la presencia de un ácido nucleico hibridable a la sonda de Temer05249 en una muestra biológica (p. ej. tejido) dada.

25 También es abarcado por la invención un método para obtener un equivalente funcional de un gen Temer05249. Este método implica obtener una sonda etiquetada que incluye un ácido nucleico aislado que codifica la totalidad o una porción de la secuencia proteínica según SEQ ID NO: 32 o una de sus variantes; cribar una biblioteca de fragmentos de ácido nucleico con la sonda etiquetada bajo condiciones que permitan la hibridación de la sonda a fragmentos de ácido nucleico en la biblioteca, formando de ese modo dúplex de ácido nucleico, y preparar una secuencia génica de longitud completa procedente de los fragmentos de ácido nucleico en cualquier dúplex etiquetado para obtener un gen relacionado con el gen Temer05249.

35 En una realización, un ácido nucleico de Temer05249 de la invención es al menos 50%, al menos 60%, al menos 65%, al menos 70%, al menos 75%, al menos 80%, al menos 85%, al menos 90%, al menos 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o más homólogo a una secuencia de ácido nucleico mostrada en SEQ ID NO: 31 o en SEQ ID NO: 34 o en SEQ ID NO: 35 o su complemento.

40 También se proporcionan 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 se presenta en la naturaleza en el genoma de la célula hospedadora o que el polipéptido no es producido naturalmente por esa célula.

45 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 procaríóticas como eucarióticas, p. ej., bacterias, hongos, levaduras y similares, de forma especialmente preferible son células de hongos filamentosos, tales como *Aspergillus niger*.

50 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 procesamiento (p. ej., escisión) de productos proteínicos pueden facilitar el funcionamiento opcional de la proteína.

55 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 especialidad 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, se pueden usar células hospedadoras eucarióticas que poseen la maquinaria celular para el procesamiento apropiado del transcrito primario, la glicosilación y la fosforilación del producto génico. Estas células hospedadoras son muy conocidas en la especialidad.

60 Si se desea, una célula según se describe 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 según se describe anteriormente) bajo condiciones que proporcionen la expresión (por 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 el 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 efectúen la replicación del vector. El vector se puede recuperar de la célula hospedadora.

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 está enlazada operativamente a una secuencia nucleotídica que codifica una secuencia de señalización. Preferiblemente, la secuencia de señalización es natural (homóloga) a la secuencia nucleotídica que codifica el polipéptido. Alternativamente, la secuencia de señalización es extraña (heteróloga) a la secuencia nucleotídica que codifica el polipéptido, en cuyo caso la secuencia de señalización preferiblemente es endógena a la célula hospedadora en la que se expresa la secuencia nucleotídica según la invención. Ejemplos de secuencias de señalización adecuadas para células hospedadoras de levadura son las secuencias de señalización derivadas de genes de factor  $\alpha$  de levadura. De forma similar, una secuencia de señalización adecuada para células hospedadoras de hongo filamentoso es, p. ej., una secuencia de señalización 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 señalización híbridas con el contexto de la presente invención.

Secuencias líder de secreción heterólogas preferidas son las que se originan a partir del gen de amiloglucosidasa (AG) fúngica (*glaA*-ambas versiones de 18 y 24 aminoácidos, p. ej., procedentes de *Aspergillus*), el gen del 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 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 por el vector de una secuencia codificante que codifica el polipéptido.

### Células hospedadoras

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

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

La invención abarca procedimientos para la producción del polipéptido de la invención por medio de 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 señalización 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.

Células hospedadoras adecuadas son 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 escasamente secretadas desde las levaduras o bien en algunos casos no son procesadas apropiadamente (p. ej. hiperglicosilación en levadura). En estos casos, se debe seleccionar un organismo hospedador fúngico.

La célula hospedadora puede sobreexpresar el polipéptido, y se conocen bien 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).

En el contexto de la presente invención, la "célula hospedadora microbiana parental" y la "célula hospedadora microbiana mutante" puede ser cualquier tipo de célula hospedadora. Las realizaciones específicas de la célula hospedadora microbiana mutante se describen posteriormente en la presente. Estará claro para los expertos en la técnica que las realizaciones aplicables a la célula hospedadora microbiana mutante son asimismo aplicables a la célula hospedadora microbiana parental a menos que se indique otra cosa.

La célula hospedadora microbiana mutante según la presente invención puede ser una célula procariótica. Preferiblemente, la célula hospedadora procariótica es una célula bacteriana. El término "célula bacteriana" incluye microorganismos tanto gramnegativos como grampositivos. Bacterias adecuadas se pueden seleccionar, p. ej., de *Escherichia*, *Anabaena*, *Caulobacter*, *Gluconobacter*, *Rhodobacter*, *Pseudomonas*, *Paracoccus*, *Bacillus*, *Brevibacterium*, *Corynebacterium*, *Rhizobium* (*Sinorhizobium*), *Flavobacterium*, *Klebsiella*, *Enterobacter*, *Lactobacillus*, *Lactococcus*, *Methylobacterium*, *Staphylococcus* o *Streptomyces*. Preferiblemente, la célula bacteriana se selecciona del grupo que consiste en *B. subtilis*, *B. amyloliquefaciens*, *B. licheniformis*, *B. pumilus*, *B. megaterium*, *B. halodurans*, *B. pumilus*, *G. oxydans*, *Caulobacter crescentus* CB 15, *Methylobacterium extorquens*, *Rhodobacter sphaeroides*, *Pseudomonas zeaxanthinifaciens*, *Paracoccus denitrificans*, *E. coli*, *C. glutamicum*, *Staphylococcus carnosus*, *Streptomyces lividans*, *Sinorhizobium melioli* y *Rhizobium radiobacter*.

Según una realización, la célula hospedadora microbiana mutante según la invención es una célula hospedadora eucariótica. Preferiblemente, la célula eucariótica es una célula de mamífero, insecto, planta, hongo o alga. Células de mamífero preferidas incluyen, p. ej., células de ovario de hámster chino (CHO), células COS, células 293, células PerC6 e hibridomas. Células de insecto preferidas incluyen, p. ej., células Sf9 y Sf21 y sus derivados. Más preferiblemente, la célula eucariótica es una célula fúngica, es decir una célula de levadura, tal como una cepa de *Candida*, *Hansenula*, *Kluyveromyces*, *Pichia*, *Saccharomyces*, *Schizosaccharomyces* o *Yarrowia*. Más preferiblemente, la célula hospedadora eucariótica es una *Kluyveromyces lactis*, *S. cerevisiae*, *Hansenula polymorpha*, *Yarrowia lipolytica* o *Pichia pastoris* o una célula de hongo filamentoso. Lo más preferiblemente, la célula eucariótica es una célula de hongo filamentoso.

Hongos filamentosos incluyen todas las formas filamentosas de la subdivisión *Eumycota* y *Oomycota* (según se define por Hawksworth y cols., En Ainsworth and Bisby's Dictionary of The Fungi, 8ª edición, 1995, CAB International, University Press, Cambridge, Reino Unido). Los hongos filamentosos se caracterizan por una pared micelar compuesta por quitina, celulosa, glucano, quitosano, manano y otros polisacáridos complejos. El crecimiento vegetativo es mediante elongación de las hifas y el catabolismo del carbono es obligadamente aeróbico. Cepas de hongos filamentosos incluyen, pero no se limitan a, cepas de *Acremonium*, *Agaricus*, *Aspergillus*, *Aureobasidium*, *Chrysosporium*, *Coprinus*, *Cryptococcus*, *Filibasidium*, *Fusarium*, *Humicola*, *Magnaporthe*, *Mucor*, *Myceliophthora*, *Neocallimastix*, *Neurospora*, *Paecilomyces*, *Penicillium*, *Piromyces*, *Panerochaete*, *Pleurotus*, *Schizophyllum*, *Talaromyces*, *Rasamsonia*, *Thermoascus*, *Thielavia*, *Tolyocladium* y *Trichoderma*.

Células de hongos filamentosos preferidas pertenecen a una especie de un género *Acremonium*, *Aspergillus*, *Chrysosporium*, *Myceliophthora*, *Penicillium*, *Talaromyces*, *Rasamsonia*, *Thielavia*, *Fusarium* o *Trichoderma*, y lo más preferiblemente una especie de *Aspergillus niger*, *Acremonium alabamense*, *Aspergillus awamori*, *Aspergillus foetidus*, *Aspergillus sojae*, *Aspergillus fumigatus*, *Talaromyces emersonii*, *Rasamsonia emersonii*, *Aspergillus oryzae*, *Chrysosporium lucknowense*, *Fusarium oxysporum*, *Myceliophthora thermophila*, *Trichoderma reesei*, *Thielavia terrestris* o *Penicillium chrysogenum*. Una célula hospedadora más preferida pertenece al género *Aspergillus* o *Rasamsonia*, más preferiblemente la célula hospedadora pertenece a la especie *Aspergillus niger* o *Rasamsonia emersonii*. Cuando la célula hospedadora según la invención es una célula hospedadora de *Aspergillus niger*, la célula hospedadora es preferiblemente CBS 513.88, CBS124.903 o uno de sus derivados.

Varias cepas de hongos filamentosos son fácilmente accesibles al público en un número de colecciones de cultivo, tales como the American Type Culture Collection (ATCC), Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ), Centraalbureau voor Schimmelcultures (CBS), Agricultural Research Service Patent Culture Collection, Northern Regional Research Center (NRRL) y All-Russian Collection of Microorganisms of Russian Academy of Sciences, (abreviatura en ruso - VKM, abreviatura en inglés - RCM), Moscú, Rusia. Cepas útiles en el contexto de la presente invención pueden ser CBS 513.88 de *Aspergillus niger*, CBS124.903, ATCC 20423 de *Aspergillus oryzae*, IFO 4177, ATCC 1011, CBS205.89, ATCC 9576, ATCC14488-14491, ATCC 11601, ATCC12892, CBS 455.95 de *P. chrysogenum*, Wisconsin54-1255 (ATCC28089) de *P. chrysogenum*, ATCC 38065 de *Penicillium citrinum*, P2 de *Penicillium chrysogenum*, NRRL8126 de *Thielavia terrestris*, CBS 124.902 de *Talaromyces emersonii*, ATCC 36225 o ATCC 48272 de *Acremonium chrysogenum*, ATCC 26921 o ATCC 56765 o ATCC 26921 de *Trichoderma reesei*, ATCC11906 de *Aspergillus sojae*, C1 de *Myceliophthora thermophila*, Garg 27K, VKM-F 3500 D, C1 de *Chrysosporium lucknowense*, Garg 27K, VKM-F 3500 D, ATCC44006 y sus derivados.

Según una realización de la invención, cuando la célula hospedadora microbiana mutante según la invención es una célula hospedadora de hongo filamentoso, la célula hospedadora microbiana mutante puede comprender una o más modificaciones en su genoma de modo que la célula hospedadora microbiana mutante sea deficiente en la producción de al menos un producto seleccionado de glucoamilasa (glaA), alfa-amilasa estable a los ácidos (amyA), alfa-amilasa neutra (amyBI y amyBII), ácido oxálico hidrolasa (oahA), una toxina, preferiblemente ocratoxina y/o fumonisina, un regulador transcripcional de proteasa prtT, PepA, un producto codificado por el gen hdfA y/o hdfB, una péptido sintasa no ribosómica *npsE*, si se compara con una célula hospedadora original y se mide bajo las mismas condiciones.

Por lo tanto, cuando la célula hospedadora microbiana mutante según la invención es una célula hospedadora de hongo filamentoso, la célula hospedadora puede comprender una o más modificaciones en su genoma para dar como resultado una deficiencia en la producción de la proteasa aspártica extracelular principal PepA. Por ejemplo, la célula

hospedadora según la invención puede comprender además una alteración del gen *pepA* que codifica la proteasa aspártica extracelular principal *PepA*.

5 Cuando la célula hospedadora microbiana mutante según la invención es una célula hospedadora de hongo filamentoso, la célula hospedadora según la invención puede comprender adicionalmente una o más modificaciones en su genoma para dar como resultado una deficiencia en la producción del producto codificado por el gen *hdfA* y/o *hdfB*. Por ejemplo, la célula hospedadora según la invención puede comprender además una alteración del gen *hdfA* y/o *hdfB*. Células hospedadoras de hongo filamentoso que son deficientes en un producto codificado por el gen *hdfA* y/o *hdfB* se han descrito en el documento WO 2005/095624.

10 Cuando la célula hospedadora microbiana mutante según la invención es una célula hospedadora de hongo filamentoso, la célula hospedadora según la invención puede comprender adicionalmente una modificación en su genoma que da como resultado la deficiencia en la producción de la péptido sintasa no ribosómica *npsE*. Estas células hospedadoras deficientes en la producción de péptido sintasa no ribosómica *npsE* se han descrito en el documento WO 2012/001169 (*npsE* tiene una secuencia genómica como la representada en SEQ ID NO: 35, una secuencia codificante representada en SEQ ID NO: 36, el ARNm representado en SEQ ID NO: 37 y la proteína *npsE* en SEQ ID NO: 38 del documento WO 2012/001169).

20 Cuando la célula hospedadora microbiana mutante según la invención es una célula hospedadora de hongo filamentoso, la célula hospedadora puede comprender adicionalmente al menos dos dominios de ADN sustancialmente homólogos adecuados para la integración de una o más copias de un polinucleótido que codifica un compuesto de interés en el que al menos uno de los al menos dos dominios de ADN sustancialmente homólogos está adaptado para tener una preferencia de integración potenciada para el polinucleótido que codifica un compuesto de interés en comparación con el dominio de ADN sustancialmente homólogo a partir del que se origina, y en donde el dominio de ADN sustancialmente homólogo, donde el dominio de ADN sustancialmente homólogo adaptado a partir del que se origina tiene una frecuencia de conversión génica que es al menos 10% superior que uno de los otros de los al menos dos dominios de ADN sustancialmente homólogos. Estas células se han descrito en el documento WO2011/009700. Las cepas que contienen dos o más copias de estos dominios de ADN sustancialmente homólogos también se denominan en lo sucesivo en la presente cepa que contiene dos o más amplicones. Ejemplos de células hospedadoras que comprenden estos amplicones se describen, p. ej., en van Dijck y cols, 2003, *Regulatory Toxicology and Pharmacology* 28; 27-35: On the safety of a new generation of DSM *Aspergillus niger* enzyme production strains. En van Dijck y cols, se describe una cepa de *Aspergillus niger* que comprende 7 locus génicos de glucoamilasa amplificados, es decir 7 amplicones. Células hospedadoras preferidas dentro de este contexto son células hospedadoras de hongo filamentoso, preferiblemente células hospedadoras de *A. niger*, que comprenden dos o más amplicones, preferiblemente dos o más amplicones  $\Delta g/aA$  (que comprenden preferiblemente 3, 4, 5, 6, 7 amplicones  $\Delta g/aA$ ) en donde el amplicón que tiene la frecuencia de conversión génica más alta se ha adaptado para tener una preferencia de integración potenciada para el polinucleótido que codifica un compuesto de interés en comparación con el amplicón a partir del que se origina. La adaptación del amplicón se puede realizar según uno cualquiera de los métodos descritos en el documento WO 2011/009700. Un ejemplo de estas células hospedadoras, descrito en el documento WO2011/009700, son células hospedadoras que comprenden tres amplicones  $\Delta g/aA$  que son un amplicón truncado por BamHI, un amplicón truncado por SaI y un amplicón truncado por BglII y en donde el amplicón de BamHI se ha adaptado para tener una preferencia de integración potenciada para un polinucleótido que codifica un compuesto de interés en comparación con el amplicón de BamHI a partir del que se origina. Las células hospedadoras que comprenden dos o más amplicones en las que un amplicón se ha adaptado para tener una preferencia de integración potenciada para un polinucleótido que codifica un compuesto de interés en comparación con el amplicón a partir del que se origina se denominan en la presente posteriormente células hospedadoras que comprenden un amplicón adaptado.

50 Cuando la célula hospedadora microbiana mutante según la invención es una célula hospedadora de hongo filamentoso, la célula hospedadora según la invención puede comprender adicionalmente una modificación de *Sec61*. Una modificación de *SEC61* preferida es una modificación que da como resultado un mutante unidireccional de *SEC61*; es decir un mutante en el que la proteína sintetizada de novo puede entrar en el ER a través de *SEC61*, pero la proteína no puede abandonar el ER a través de *SEC61*. Estas modificaciones se describen a fondo en el documento WO 2005/123763. Lo más preferiblemente, la modificación de *SEC 61* es la mutación S376W en la que la Serina 376 se reemplaza por triptófano.

60 Las células hospedadoras según la invención incluyen células vegetales, y por lo tanto la invención se extiende a organismos transgénicos, tales como plantas y sus partes, 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.

65

Alternativamente, se puede afectar a la infección directa de una parte 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 pinchar 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 células transformadas en plantas genéticamente modificadas se puede conseguir usando técnicas conocidas, por ejemplo al seleccionar brotes transformados usando un antibiótico y al subcultivar los brotes en un medio que contiene los nutrientes apropiados, hormonas vegetales y similares.

La invención también incluye células que se han modificado para expresar la celobiohidrolasa de la invención o una de sus variantes. Estas células incluyen líneas celulares eucarióticas superiores transitorias, o preferiblemente estables, tales como células de mamífero o células de insecto, células eucarióticas inferiores, tales como células de levadura y hongos (p. ej. filamentosos) o células procarióticas tales como células bacterianas.

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

Según la presente invención, la producción del polipéptido de la invención puede estar afectada por 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 especialidad. 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 celular o el título deseados 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, melaza, almidón, celulosa, xilano, pectina, hidrolizado de biomasa lignocelulítica, 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 especialidad. Si se desea, el medio puede contener componentes adicionales que favorezcan los hospedadores de expresión transformados sobre otros microorganismos potencialmente contaminantes.

La fermentación se puede realizar a lo largo de un período de aproximadamente 0,5 a aproximadamente 30 días. Puede ser un procedimiento discontinuo, continuo o semicontinuo, adecuadamente a una temperatura en el intervalo de 0-100°C o 0-80°C, por ejemplo, de aproximadamente 0 a aproximadamente 50°C y/o a un pH, por ejemplo, de aproximadamente 2 a aproximadamente 10. Condiciones de fermentación preferidas son una temperatura en el intervalo de aproximadamente 20 a aproximadamente 45°C y/o a un pH de aproximadamente 3 a aproximadamente 9. Las condiciones apropiadas se seleccionan habitualmente basándose en la elección del hospedador de expresión y la proteína que se vaya 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 retirar a continuación 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 y/o ligninasa o una enzima modificadora de lignina.

Cuando el polipéptido de la invención sea una celulasa, una composición de la invención comprenderá típicamente una hemicelulasa y/o una pectinasa y/o ligninasa o una enzima modificadora de lignina además del polipéptido de la invención.

Cuando el polipéptido de la invención sea una hemicelulasa, una composición de la invención comprenderá típicamente una celulasa y/o una pectinasa y/o ligninasa o una enzima modificadora de lignina además del polipéptido de la invención.

5 Cuando el polipéptido de la invención sea una pectinasa, una composición de la invención comprenderá típicamente una celulasa y/o una hemicelulasa y/o ligninasa o una enzima modificadora de lignina además del polipéptido de la invención.

10 Cuando el polipéptido de la invención sea una ligninasa o una enzima modificadora de lignina, una composición de la invención comprenderá típicamente una celulasa y/o una hemicelulasa y/o una pectinasa además del polipéptido de la invención.

15 Una composición de la invención puede comprender una, dos, tres o más clases de celulasa, por ejemplo una, dos o todas de una GH61, una endo-1,4-3-glucanasa (EG), una exocelobiohidrolasa (CBH) y una  $\beta$ -glucosidasa (BGL).

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

25 En la presente, una celulasa es cualquier polipéptido que sea capaz de degradar o celulosa. Un polipéptido que sea capaz de degradar celulosa es uno que sea capaz de catalizar el proceso de descomposición de 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 ponga 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 o hemicelulosa. Es decir, una hemicelulasa puede ser capaz de degradar uno o más de xilano, glucuronoxilano, arabinoxilano, glucomanano y xiloglucano. Un polipéptido que sea capaz de degradar una hemicelulosa es uno que sea 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 los monómeros sacáricos hexosa o pentosa. Una hemicelulasa según la invención puede dar lugar a una población mixta de oligosacáridos y monómeros sacáricos cuando se ponga en contacto con la hemicelulosa. Esta 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 o pectina. Un polipéptido que sea capaz de degradar pectina es uno que sea 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 ponga en contacto con la pectinasa. Esta degradación tendrá lugar típicamente por medio de una reacción de hidrólisis.

40 En la presente, una ligninasa o una enzima modificadora de lignina es cualquier polipéptido que sea capaz de degradar o modificar lignina o sus componentes de degradación. Un polipéptido que sea capaz de degradar o modificar lignina es uno que sea capaz de catalizar el proceso de descomposición de lignina en unidades menores, parcialmente, por ejemplo en compuestos monofenólicos. Una ligninasa o una enzima modificadora de lignina según la invención puede dar lugar a una población mixta de compuestos fenólicos cuando se pone en contacto con la lignina. Esta degradación tendrá lugar típicamente por medio de una reacción de oxidación. En la presente, una ligninasa o una enzima modificadora de lignina también puede ser cualquier polipéptido que sea capaz de degradar productos de degradación fenólicos de lignina. Un polipéptido que sea capaz de degradar productos de degradación fenólicos de lignina es uno que sea capaz de catalizar el proceso de descomposición de productos de degradación fenólicos de lignina en unidades todavía menores, por ejemplo al catalizar una reacción de apertura de anillo del anillo fenólico. Una ligninasa o una enzima modificadora de lignina según la invención puede dar lugar a una población mixta de productos de degradación de anillo abierto de compuestos fenólicos cuando se ponga en contacto con los productos de degradación fenólicos de lignina. Esta degradación tendrá lugar típicamente por medio de una reacción de oxidación. La ligninasa o una enzima modificadora de lignina puede ser capaz además de romper enlaces entre celulosa o hemicelulosa y la lignina o sus productos de degradación. Enzimas que pueden descomponer incluyen lignina peroxidasas, manganeso peroxidasas, lacasas y feruloilesterasas, y otras enzimas conocidas en la técnica para 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 feruloilesterasas (EC 3.1.1.73).

- 5 Según esto, una composición de la invención puede comprender cualquier celulasa, por ejemplo, una GH61, una celobiohidrolasa, una endo- $\beta$ -1,4-glucanasa, una  $\beta$ -glucosidasa o una  $\beta$ -(1,3)(1,4)-glucanasa.

10 Las proteínas GH61 (familia de glucósido hidrolasas 61 o a veces denominadas EGIV) son polisacárido monooxigenasas (PMO's) dependientes de oxígeno según la última bibliografía. A menudo, en la bibliografía, se menciona que estas proteínas potencian la acción de celulasas sobre sustratos de lignocelulosa. GH61 se clasificó originalmente como endoglucanasa basándose en la medida de una actividad como endo-1,4- $\beta$ -d-glucanasa muy débil en un miembro de la familia. El término "GH61", según se usa en la presente, se ha de entender como una familia de enzimas, que comparten porciones de secuencia conservadas y pliegues comunes que se van a clasificar en la familia 61 del sistema de clasificación CAZY GH bien establecido (<http://www.cazy.org/GH61.html>). La familia 61 de glucósido hidrolasas es un miembro de la familia de glucósido hidrolasas EC 3.2.1. GH61 se usa en la presente como si fuera parte de las celulasas.

20 En la presente, una celobiohidrolasa (EC 3.2.1.91) 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 no reductores 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, exocelobiohidrolasa o exoglucanasa. Puede tener el código EC EC 3.2.1.91.

25 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-glucanos 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. La endo-glucanasa también puede catalizar la escisión de xiloglucano, un esqueleto de residuos de glucosa con enlaces  $\beta$ 1 $\rightarrow$ 4, la mayoría de los cuales sustituidos con cadenas laterales de xilosa con enlaces 1-6, y la enzima se denomina entonces una endo- $\beta$ -1,4-glucanasa específica de xiloglucano o una xiloglucanasa.

35 En la presente, una  $\beta$ -glucosidasa (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-glucosidos y también puede hidrolizar uno o más de los siguientes: un  $\beta$ -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 gentobiasa.

40 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 contengan uniones 1,3 y 1,4. Tal 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 de 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 esté 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; sustratos incluyen laminarina, liquenina y beta-D-glucanos de cereales.

50 Una composición de la invención puede comprender cualquier hemicelulasa, por ejemplo, una endo-xilanasa, una  $\beta$ -xilosidasa, una  $\alpha$ -L-arabinofuranosidasa, una  $\alpha$ -D-glucuronidasa, una celobiohidrolasa, una feruloilesterasa, una cumarolesterasa, una  $\alpha$ -galactosidasa, una  $\beta$ -galactosidasa, una  $\beta$ -mananasa o una  $\beta$ -manosidasa.

55 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 xilanohidrolasa. Una alternativa es EC 3.2.1.136, una glucuronoarabinoxilano endoxilanasa, una enzima que es capaz de hidrolizar enlaces 1,4 xilosídicos en glucuronoarabinoxilanos.

60 En la presente, una  $\beta$ -xilosidasa (EC 3.2.1.37; GH3) es cualquier polipéptido que sea capaz de catalizar la hidrólisis de 1,4- $\beta$ -D-xilanos, para retirar residuos de D-xilosa sucesivos de las terminaciones no reductoras. 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.

65 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:  $\alpha$ -D-glucuronósido + H(2)O = un alcohol + D-glucuronato. Esta enzima también se puede denominar  $\alpha$ -glucuronidasa o  $\alpha$ -glucosiduronasa. Estas enzimas también pueden hidrolizar ácido glucurónico metilado en 4-O, que también puede estar presente como un sustituyente en xilanos. Una alternativa es EC 3.2.1.131: xilano  $\alpha$ -1,2-glucuronosidasa, que cataliza la hidrólisis de enlaces  $\alpha$ -1,2-(4-O-metil)glucuronosílicos.

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 xilooligosacáridos. Este polipéptido puede catalizar la hidrólisis de grupos acetilo de xilano polimérico, xilosa acetilada, glucosa acetilada, acetato de  $\alpha$ -naftilo o acetato de p-nitrofenilo, pero, típicamente, no de triacetilglicerol. Típicamente, este polipéptido no actúa sobre manano acetilado o pectina.

En la presente, una feruloilesterasa (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) desde un azúcar esterificado, que habitualmente es arabinosa en sustratos 'naturales'. Típicamente, el acetato de p-nitrofenol y el ferulato de metilo son sustratos más pobres. Esta enzima también se puede denominar éster cinamoílico hidrolasa, ácido ferúlico esterasa o hidroxicinamoilesterasa. También se puede denominar una enzima auxiliar de hemicelulasa, puesto que puede ayudar a la xilanasas y pectinasas a descomponer la hemicelulosa y la pectina de la pared celular de la planta.

En la presente, una cumaroilesterasa (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-cumaroilesterasa, trans-p-cumaroilesterasa, p-cumaroilesterasa o ácido p-cumárico esterasa. Esta enzima también se encuentra dentro de EC 3.1.1.73, de modo que también se puede denominar una feruloilesterasa.

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->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 exogalacturonasa, una exopoligalacturonato 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 glucanohidrolasa, endogalacturonasa; endo-D-galacturonasa o poli(1,4- $\alpha$ -D-galacturónido) glucanohidrolasa.

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 puede conocer 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 una enzima que tiene una actividad como acetil esterasa que cataliza la desacetilación de los grupos acetilo en los grupos hidroxilo de residuos de GalUA de pectina.

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

10 En la presente, una pectato liasa (EC 4.2.2.2) es cualquier enzima capaz de catalizar la escisión eliminativa de (1→4)-α-D-galacturonano para dar oligosacáridos con grupos 4-desoxi-α-D-galact-4-enuronosilo en sus extremos no reductores. La enzima también se puede conocer como transeliminasa poligalacturónica, ácido péctico transeliminasa, poligalacturonato liasa, endopectina metiltranseliminasa, pectato transeliminasa, endogalacturonato transeliminasa, ácido péctico liasa, liasa péctica, α-1,4-D-endopoli(ácido galacturónico) liasa, PGA liasa, PPasa-N, endo-α-1,4-poli(ácido galacturónico) liasa, poli(ácido galacturónico) liasa, pectina transeliminasa, poli(ácido galacturónico) transeliminasa o (1→4)-α-D-galacturonano liasa.

20 En la presente, una alfa ramnosidasa (EC 3.2.1.40) es cualquier polipéptido que sea capaz de catalizar la hidrólisis de residuos de α-L-ramnosa no reductores terminales en α-L-ramnósidos o alternativamente en ramnogalacturonano. Esta enzima también se puede conocer como α-L-ramnosidasa T, α-L-ramnosidasa N o α-L-ramnósido ramnohidrolasa.

25 En la presente, exogalacturonasa (EC 3.2.1.82) es cualquier polipéptido capaz de hidrólisis de ácido péctico desde el extremo no reductor, liberando digalacturonato. La enzima también se puede conocer como exo-poli-α-galacturonosidasa, exopoligalacturonosidasa o exopoligalacturanosidasa.

30 En la presente, exogalacturonasa (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-α-galacturonidasa, exopoligalacturonasa, poli(galacturonato) hidrolasa, exo-D-galacturonasa, exo-D-galacturonanasa, exo-poli-D-galacturonasa o poli(1,4-α-D-galacturónido) galacturonohidrolasa.

35 En la presente, exo-poligalacturonato liasa (EC 4.2.2.9) es cualquier polipéptido capaz de catalizar la escisión eliminativa de 4-(4-desoxi-α-D-galact-4-enuronosil)-D-galacturonato desde el extremo reductor de pectato, es decir pectina desesterificada. Esta enzima se puede conocer como pectato disacárido-liasa, pectato exo-liasa, ácido exopéctico transeliminasa, exo-pectato liasa, exopoli(ácido galacturónico)-trans-eliminasa, PATE, exo-PATE, exo-PGL o extremo reductor de (1→4)-α-D-galacturonano disacárido-liasa.

40 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-alfa-L-ramnoil-(1,4)-alfa-galactosilurónico].

45 En la presente, ramnogalacturonano liasa es cualquier polipéptido que sea capaz de escindir enlaces α-L-Rhap-(1→4)-α-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 alternos en ramnogalacturonano.

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

55 En la presente, xilogalacturonasa es cualquier polipéptido que actúe sobre xilogalacturonano al escindir el ácido galacturónico sustituido con β-xilosa de un modo endo. Esta enzima también se puede conocer como xilogalacturonano hidrolasa.

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

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

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.

5 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.

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

15 "Lipasa" incluye enzimas que hidrolizan lípidos, ácidos grasos y acilglicéridos, incluyendo fosfoglicéridos, lipoproteínas, diacilglicérols 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.

20 "Hexosiltransferasa" (2.4.1-) incluye enzimas que sean capaces de transferir grupos glicosilo, más específicamente grupos hexosilo. Además, para la transferencia de un grupo glicosilo de un donante que contiene grupo 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  
25 celulosa.

"Glucuronidasa" incluye enzimas que catalizan la hidrólisis de un glucoronósido, 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), glicirricinato  $\beta$ -glucuronidasa (3.2.1.128) o  $\alpha$ -D-glucuronidasa (EC 3.2.1.139).

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

35 Las expansinas están implicadas en la pérdida de la estructura de la pared celular durante el crecimiento de la célula vegetal. Se ha propuesto que las expansinas alteran la unión por hidrógeno entre la 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 de tipo expansina, contiene un dominio de la familia 1 de módulos de unión glucídicos (CBD) N-terminal y un dominio de tipo expansina C-terminal. Para los  
40 propósitos de esta invención, una proteína de tipo expansina o una proteína de tipo swollenina puede comprender uno o ambos de estos dominios y/o puede alterar la estructura de paredes celulares (tal como alterar la estructura de la celulosa), opcionalmente sin producir cantidades detectables de azúcares reductores.

45 Una composición de la invención puede comprender el producto polipeptídico de una proteína integradora de celulosa, un andamiaje o una proteína de tipo andamiaje, por ejemplo CipA o CipC procedentes de *Clostridium thermocellum* o *Clostridium cellulolyticum*, respectivamente.

50 Los andamiajes y las proteínas integradoras de celulosa son subunidades integradoras multifuncionales que pueden organizar subunidades celulolíticas en un complejo multienzimático. Esto se efectúa mediante la interacción de dos clases de dominio complementarias, es decir un dominio de cohesión sobre un andamiaje y un dominio de dockerina sobre cada unidad enzimática. La subunidad de andamiaje también soporta un módulo de unión a celulosa (CBM) que media en la ligazón del celulosoma a su sustrato. Un andamiaje o una proteína integradora de celulosa, para los propósitos de esta invención, puede comprender uno o ambos de estos dominios.

55 Una composición de la invención puede comprender una proteína inducida por celulosa o proteína moduladora, por ejemplo como la codificada por el gen cip1 o cip2 o genes similares procedentes de *Trichoderma reesei* / *Hypocrea jacobina* (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 de unión a celulosa y un dominio cuya función o actividad no se puede relacionar con familias de glicosil hidrolasa conocidas. No obstante, la presencia de un módulo que se une a  
60 celulosa y la correulación de la expresión de estos genes con componentes de celulasas indica actividades previamente no reconocidas con un papel potencial en la degradación de biomasa.

65 Una composición de la invención puede estar compuesta por 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, procedentes 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 medio, en donde las cepas secretan proteínas y enzimas en el 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. A continuación, los azúcares fermentables, como parte de un procedimiento para un biocombustible, se pueden convertir en biogás 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 dentífricas 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 de la celobiohidrolasa según la invención y composiciones que comprenden esta enzima en procedimientos industriales.

A pesar de la experiencia prolongada obtenida con estos procedimientos, la celobiohidrolasa según la invención puede presentar un número de ventajas significativas sobre enzimas actualmente usadas. 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 produzcan en un microorganismo adecuado, intervalos de pH y temperatura más adecuados, falta de inhibición por productos hidrófobos derivados de lignina o menos inhibición de producto o, en el caso de la industria alimentaria, un mejor sabor o textura de un producto final así como calidad alimentaria y aspectos relacionados con la ley judía.

En principio, una celobiohidrolasa o una composición 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 material polisacárido. En la presente, material polisacárido es un material que comprende o consiste esencialmente en uno o, más típicamente, más de un polisacárido.

Típicamente, las plantas y el material derivado de las mismas comprenden cantidades significativas de material polisacárido 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 componente importante del material polisacárido no amiláceo vegetal es la lignocelulosa (también denominada en la presente biomasa lignocelulolítica). La lignocelulosa es material vegetal que comprende celulosa y hemicelulosa y lignina. Los polímeros glucídicos (celulosa y hemicelulosas) están estrechamente unidos a la lignina mediante uniones por hidrógeno y covalentes. Según esto, un polipéptido de la invención se puede usar en el tratamiento de material lignocelulolítico. En la presente, 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 lignocelulósico/una biomasa.

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 la hidrólisis y/o la 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 emisión de oligosacáridos y/o monómeros sacáricos.

Todas las plantas contienen polisacárido no amiláceo como virtualmente todos los materiales polisacáridos 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 proporcionar en la forma de una planta o un material derivado de planta o un material que comprende una planta o un material derivado de planta, por ejemplo una pasta papelera vegetal, un extracto vegetal, un producto alimenticio o un ingrediente, por lo tanto, una tela, un producto textil o una prenda de vestir.

5 Biomasa lignocelulolítica para el uso en la invención incluye biomasa y puede incluir biomasa virgen y/o biomasa no virgen tal como biomasa agrícola, materiales orgánicos comerciales, residuos de construcción y demolición, residuos sólidos municipales, papel residual y residuos 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, granos de maíz  
10 incluyendo fibra procedente de granos, productos y subproductos de molienda de cereales tales como maíz, trigo y cebada (incluyendo molienda en húmedo y molienda en seco) a menudo llamados "salvado o fibra" así como residuos sólidos municipales, papel residual y residuos de jardines. La biomasa también puede ser, pero no se limita a, material herbáceo, residuos agrícolas, residuos forestales, residuos sólidos municipales, papel residual y residuos de pasta papelera y molienda de papel. "Biomasa agrícola" incluye ramas, matorrales, cañas, maíz y cáscaras de maíz, cultivos energéticos, bosques, frutos, flores, cereales, hierbas, cultivos herbáceos, hojas, corteza, agujas, leños, raíces, plántulas, cultivos leñosos de rotación corta, arbustos, pastos varilla, árboles, hortalizas, cáscaras de fruta, vides, pulpa de remolacha azucarera, moyuelo 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 agrícolas y forestales, incluyendo específicamente residuos  
15 madereros forestales. La biomasa agrícola puede ser cualquiera de las susodichas individualmente o en cualquiera de sus combinaciones o mezclas. Ejemplos adicionales de biomasa adecuada son preparación de huertos, chaparrales, residuo de molienda, residuos madereros urbanos, residuos municipales, residuos de explotación forestal, aclarados forestales, cultivos leñosos de rotación corta, residuos 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, alimentación de gluten de maíz, cáscaras de avena, caña de azúcar, rastrojo de maíz, tallos de maíz, mazorcas de  
20 maíz, cáscaras de maíz, hierba de pradera, zacate maicero, cola de zorra; pulpa de remolacha azucarera, pulpa de cítricos, cáscaras de semillas, residuos 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 de granos, productos y subproductos procedentes de molienda en húmedo o en seco de cereales, residuos sólidos municipales, papel residual, residuos de jardines, material herbáceo, residuos agrícolas, residuos forestales, residuos sólidos municipales, papel residual, pasta papelera, residuos de molienda de papel, ramas, matorrales, cañas, maíz, cáscaras de maíz, un cultivo energético, bosques, una fruta, una flor, un cereal, una hierba, un cultivo herbáceo, una hoja, corteza, una aguja, un leño, una raíz, una plántula, un arbusto, pasto varilla, un árbol, una hortaliza, piel de fruta, una vid, pulpa de remolacha azucarera, moyuelo de trigo, cáscaras de avena, madera dura  
25 o blanda, material residual orgánico generado a partir de un procedimiento agrícola, residuos madereros forestales, o una combinación de cualesquiera dos o más de los mismos.

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

### Pretratamiento

45 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 la celulosa y/o la lignina, de cualquier modo conocido en la especialidad. El pretratamiento puede comprender exponer el material lignocelulósico a agua (caliente), vapor de agua (explosión de vapor de agua), un ácido, una base, un disolvente, calor, un peróxido, ozono, desfibración mecánica, trituración, molienda o despresurización rápida, o una combinación de cualesquiera dos o más de los mismos. Este pretratamiento químico a menudo se combina con pretratamiento térmico, p. ej. entre 150-220°C durante de 1 a 30 minutos.

### 50 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 prueba, o el óptimo de enzima predicho de las enzimas que se van a probar. La temperatura de la etapa de presacarificación puede variar de aproximadamente 10°C a aproximadamente 95°C, aproximadamente 20°C a aproximadamente 85°C, aproximadamente 30°C a aproximadamente 70°C, aproximadamente 40°C a aproximadamente 60°C, aproximadamente 37°C a aproximadamente 50°C, preferiblemente de aproximadamente 37°C a aproximadamente 80°C, más preferiblemente aproximadamente 60-70°C, aún 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 comparación de los resultados del ensayo procedentes de esta prueba permitirá modificar el método para adaptarse mejor a las enzimas que se prueban.

5 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 de celulosa 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

15 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 para una composición de la invención con el material lignocelulósico.

20 Este método permite que se generen azúcares libres (monómeros) y/u 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.

25 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 emisión de uno o más azúcares de hexosa y/o pentosa, tales como uno o más de glucosa, xilosa, arabinosa, galactosa, ácido galacturónico, ácido glucurónico, manosa, ramnosa, ribosa y fructosa.

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

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

40 Las enzimas se hacen reaccionar con sustrato bajo cualesquiera 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 baja a media y/o de 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 individual. 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 biogás es deseable pH neutro, p. ej. pH=7. La incubación de combinaciones de enzimas bajo estas condiciones da como resultado la liberación o la emisión 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.

55 Los polipéptidos, tales como enzimas, se pueden producir exógenamente en microorganismos, levaduras, hongos, bacterias o plantas, y a continuación aislarse y añadirse, por ejemplo, a materia prima lignocelulósica. Alternativamente, las enzimas se producen, pero no se aíslan, y caldo de fermentación de masa celular cruda o material vegetal (tal como rastrojo de maíz) y similares se pueden añadir a, por ejemplo, la materia prima. Alternativamente, la masa celular cruda o el medio de producción de enzimas o el material vegetal se puede tratar para prevenir un crecimiento microbiano adicional (por ejemplo, mediante calentamiento o adición de agentes antimicrobianos) y a continuación añadir a, por ejemplo, una materia prima. Estas mezclas de enzimas brutas pueden incluir el organismo que produce la enzima. Alternativamente, la enzima se puede producir en una fermentación que usa materia prima (tal como rastrojo 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 se pueden añadir 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, complementos 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 etanol, metanol, butanol, combustibles líquidos sintéticos y biogás.

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 biomasa se pueden convertir en productos de fermentación útiles tales como uno de los que incluyen, pero no se limitan a, aminoácidos, vitaminas, productos farmacéuticos, complementos para piensos, productos químicos especializados, materias primas químicas, plásticos y etanol, incluyendo etanol combustible

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

a. degradar lignocelulosa usando un método según se describe en la presente; y

b. fermentación del material resultante,

para preparar de ese modo un producto de fermentación.

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

30 Un procedimiento de fermentación anaeróbica se define en la presente como un procedimiento de fermentación realizado en ausencia de oxígeno o en el que sustancialmente no se consume oxígeno, preferiblemente aproximadamente 5 o menos, aproximadamente 2,5 o menos o aproximadamente 1 mmol/l/h o menos, y en donde moléculas orgánicas sirven tanto como donantes de electrones como 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 masa reales del equipo de fermentación usado. Preferiblemente, en un procedimiento bajo condiciones de 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.

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

### SSF

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

### Productos de fermentación

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

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 biogás); 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-hidroxipropiónico, ácido acrílico; ácido acético; 1,3-propanodiol; etileno, glicerol; un disolvente; un complemento 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 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

La invención también proporciona el uso de un polipéptido o una composición según se describen en la presente en un método para la preparación de biogás. Biogás se refiere típicamente a un gas producido por la descomposición biológica de materia orgánica, por ejemplo material que contiene carbohidratos no amiláceos, en ausencia de oxígeno. El biogás se origina a partir de material biogénico y es un tipo de biocombustible. Un tipo de biogás se produce mediante digestión o fermentación anaeróbica de materiales biodegradables tales como biomasa, estiércol o aguas residuales, residuos municipales y cultivos energéticos. Este tipo de biogás está comprendido principalmente por metano y dióxido de carbono. El metano gaseoso se puede quemar u oxidar con oxígeno. El aire contiene 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 para cualquier propósito de calentamiento, tal como cocinar. También se puede utilizar en instalaciones modernas de tratamiento de residuos en las que se puede usar para poner en marcha cualquier tipo de motor calentador, para generar energía bien mecánica o bien eléctrica.

La primera etapa en la producción de biogás microbiano consiste en la degradación enzimática de polímeros y sustratos complejos (por ejemplo carbohidrato no amiláceo). Según esto, la invención proporciona un método para la preparación de un biogás en el que un sustrato que comprende carbohidrato no amiláceo se pone en contacto con un polipéptido o una composición de la invención, para dar de ese modo material fermentable que 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 exprese este polipéptido.

### 25 Uso de enzimas en productos alimenticios

Los polipéptidos y las composiciones de la invención se pueden usar en un método de procesamiento de material vegetal para degradar o modificar la celulosa o hemicelulosa o constituyentes de sustancias pécticas 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.

La invención también proporciona un método de procesamiento de 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 del material (vegetal). Preferiblemente, el material vegetal es una pulpa vegetal o un extracto vegetal, tal como jugos.

La presente invención también proporciona un método para reducir la viscosidad, la transparencia y/o la capacidad de filtración de un extracto vegetal, método que comprende poner en contacto en extracto vegetal con un polipéptido o una composición de la invención en una cantidad eficaz en la degradación de celulosa o hemicelulosa o sustancias pécticas contenidas en el extracto vegetal.

Materiales vegetales y que contienen celulosa/hemicelulosa/sustancia péctica incluyen pulpa vegetal, partes de plantas y extractos de plantas. 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 jugo, néctar, base o concentrados elaborados de los mismos. El material vegetal puede comprender o derivarse de hortalizas, p. ej., zanahorias, apio, cebollas, legumbres o plantas leguminosas (soja, haba de soja, guisantes) o fruta, p. ej., pomos o frutas de semilla (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 sus partes (p. ej. polen, procedente de pinos), o cereales (avena, cebada, trigo, maíz, arroz). El material (que se va a hidrolizar) también pueden ser residuos agrícolas, tales como pulpa de remolacha azucarera, mazorcas de maíz, 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 comestibles de productos alimenticios, o usar en la extracción de café, aceites vegetales, almidón o como un espesante en 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 trituración o molienda de material vegetal. La incubación de la composición con la

5 planta típicamente se llevará a cabo 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, preferiblemente de aproximadamente 30 g a aproximadamente 70 g, óptimamente aproximadamente 50 g de enzima por tonelada de material que se va a tratar.

10 La totalidad de la enzima o las enzimas o sus composiciones usada se puede añadir secuencialmente o al mismo tiempo a la pulpa vegetal. Dependiendo de la composición de la preparación enzimática, el material vegetal se puede macerar (p. ej. hasta un puré) o licuar en primer lugar. 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.

15 Alternativamente, o además de lo anterior, un polipéptido de la invención se puede añadir al jugo en bruto obtenido a partir del prensado o licuado de la pulpa vegetal. El tratamiento del jugo en bruto se puede llevar 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.

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

25 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 que inactiven parcialmente o totalmente el polipéptido o los polipéptidos de la invención.

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

30 El polipéptido de la invención también se puede usar en fabricación de cerveza, elaboración de vino, destilación o la panificación. Por lo tanto, se puede usar en la preparación de bebidas alcohólicas tales como vino y cerveza. Por ejemplo, puede mejorar la capacidad de filtración o la transparencia, por ejemplo de cervezas, mosto (p. ej. que contiene malta de cebada y/o sorgo) o vino.

35 Por otra parte, un polipéptido o una composición de la invención se puede usar para el tratamiento de orujo de cerveza, es decir residuos de la producción de mosto de cerveza que contienen cebada o cebada malteada y otros cereales, a fin de mejorar la utilización de los residuos para, p. ej., piensos.

40 La proteína puede ayudar en la retirada de sustancias orgánicas disueltas de caldo o medio de cultivo, por ejemplo cuando residuo de destilería de origen orgánico se bioconvierte en biomasa microbiana. El polipéptido de la invención puede mejorar la capacidad de filtración y/o reducir la viscosidad en jarabes de glucosa, tales como los procedentes de cereales producidos por licuefacción (p. ej. con  $\alpha$ -amilasa).

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

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

55 La preparación del producto alimenticio a base de cereales según la invención puede comprender además etapas conocidas en la especialidad tales como ebullición, secado, fritura, tratamiento al vapor o cocción de la masa obtenida.

60 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 al vapor son por ejemplo bollos al vapor y tallarines al vapor, ejemplos de productos elaborados a partir de masa deshidratada son pasta y tallarines deshidratados y ejemplos de productos elaborados a partir de masa cocida son pan, galletas y bizcochos.

65 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 a base de 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

5 mejorada de la masa, impermeabilización mejorada de la masa, pegajosidad reducida de la masa, extensibilidad mejorada de la masa, volumen incrementado del producto alimenticio a base de cereales, vesiculación reducida del producto alimenticio a base de cereales, estructura de la miga mejorada del producto cocido, blandura mejorada del producto alimenticio a base de cereales, aroma mejorado del producto alimenticio a base de cereales, antienranciamiento mejorado del producto alimenticio a base de cereales. Propiedades mejoradas relacionadas con productos a base de cereales de tipo pasta y tallarín son por ejemplo firmeza mejorada, pegajosidad reducida, cohesión mejorada y pérdida por cocción reducida.

10 La propiedad mejorada se puede determinar mediante la comparación de una masa y/o un producto alimenticio a base de 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 de fabricación de pan y pueden incluir, por ejemplo, el uso de un grupo de catadores entrenados.

15 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. Trigo pretende abarcar ahora y posteriormente en la presente todas las especies conocidas del género Triticum, por ejemplo harinero, duro y/o espelta. Ejemplos de otros ingredientes adecuados son: la celobiohidrolasa según la presente invención, enzimas adicionales, aditivos químicos y/o adyuvantes de procesamiento. La masa puede ser fresca, congelada, preparada o precocida. La preparación de una masa a partir de los ingredientes descritos anteriormente es muy conocida en la especialidad y comprende la mezcladura de dichos ingredientes y adyuvantes de procesamiento y una o más etapas de moldeo y opcionalmente fermentación. La preparación de una masa congelada es descrita por Kulp y Lorenz en Frozen and Refrigerated Doughs and Batters.

25 El término "producto alimenticio a base de 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 a base de cereales, ya sean de tipo blanco, claro u oscuro, que pueden ser producidos 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, bizcocho, pan de pita, tortillas, tacos, bizcochos, panqueques, pasteles, galletas, tartaletas, pan al vapor y pan crujiente y similares.

30 El término "producto cocido" se define en la presente como cualquier producto alimenticio a base de cereales preparado al cocer la masa.

35 Los polisacáridos no amiláceos (NSP) pueden incrementar la viscosidad de los digestos lo que, a su vez, puede disminuir la disponibilidad de nutrientes y el rendimiento de los animales. El uso de la celobiohidrolasa de la presente invención puede mejorar la utilización de fósforo así como minerales catiónicos y proteína durante los digestos animales.

40 Añadir nutrientes específicos al pienso mejora la digestión de los animales y de ese modo reduce los costes del pienso. Se está usando actualmente una gran cantidad de aditivos de piensos y se desarrollan continuamente nuevos conceptos. El uso de enzimas específicas como enzimas degradadoras de carbohidratos no amiláceos podría descomponer la fibra liberando energía así como incrementando la digeribilidad de proteínas debido a la mejor accesibilidad de la proteína cuando la fibra se descompone. De este modo, el coste del pienso disminuiría así como también se podrían reducir los niveles de proteína en el pienso.

45 También están presentes polisacáridos no amiláceos (NSPs) virtualmente en todos los ingredientes de piensos de origen vegetal. Los NSPs son poco utilizados y, cuando están solubilizados, pueden ejercer efectos adversos sobre la digestión. Las enzimas exógenas pueden contribuir a una mejor utilización de estos NSPs y, como consecuencia, reducen cualesquiera efectos antinutricionales. Las enzimas degradadoras 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 una menor extensión, para cerdos y otras especies.

50 Un polipéptido/enzima degradadores de carbohidratos no amiláceos de la invención (de 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 manchas glucídicas de la ropa. 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**

60 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 la ropa a mano o a máquina que incluye una composición aditiva para la ropa para el pretratamiento de telas manchadas y una composición suavizante para ropa añadida por enjuague, o se puede formular como una composición detergente para el uso en operaciones de limpieza de superficies duras domésticas generales, o se puede formular para operaciones de lavado de platos 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 iniónico, 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 la enzima. Una composición detergente también puede comprender cualquier otro ingrediente detergente convencional tal como, por ejemplo, un acondicionador incluyendo una arcilla, un reforzador de la espuma, un supresor de las jabonaduras, un agente anticorrosivo, un agente de suspensión de la suciedad, un agente contra la redeposición de la suciedad, un colorante, un bactericida, un abrillantador óptico, un hidrótrofo, un inhibidor del deslustre o un perfume.

### Uso de enzimas en el procesamiento de papel y pasta papelera

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 potenciar el brillo de pastas papeleras blanqueadas con lo que la cantidad de cloro usada en las fases de blanqueo se puede reducir, y para incrementar la frescura de las pastas papeleras en el procedimiento de reciclado 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 su capacidad de blanqueo. De ese modo, se puede reducir la cantidad de cloro necesaria para obtener un blanqueo satisfactorio de la pasta papelera.

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 vuelven ásperas o para reducir la aspereza de las telas que contienen celulosa, comprendiendo el método tratar las 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 clarificación cromática de telas que contienen celulosa, 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 la imbibición o el enjuague o simplemente al añadir el polipéptido o la composición que se describen anteriormente a agua en la que las telas están sumergidas o se van a sumergir.

### Otros usos de la enzima

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 dentífricas y colutorios.

Los siguientes Ejemplos ilustran la invención:

### **EJEMPLOS**

#### **Información experimental**

#### **Cepas y composiciones enzimáticas**

La cepa de *Aspergillus niger* se deposita en the CBS Institute bajo el número de depósito CBS 513.88.

La cepa de *Rasamsonia (Talaromyces) emersonii* TEC-142 se deposita en CENTRAAL BUREAU VOOR SCHIMMELCULTURES, Uppsalalaan 8, P.O. Box 85167, NL-3508 AD Utrecht, Países Bajos, el 1 de julio de 2009 teniendo el Número de Registro CBS 124902. TEC-142S es un aislado individual de TEC-142.

La cepa de *Rasamsonia (Talaromyces) emersonii* se depositó en CENTRAAL BUREAU VOOR SCHIMMELCULTURES, Uppsalalaan 8, P.O. Box 85167, NL-3508 AD Utrecht, Países Bajos, en diciembre de 1964 teniendo el Número de Registro CBS 393.64. Otras cepas adecuadas se pueden usar igualmente en los presentes ejemplos para mostrar el efecto y las ventajas de la invención. Por ejemplo, TEC-101, TEC-147, TEC-192, TEC-201 o TEC-210 son cepas de *Rasamsonia* adecuadas que se describen en el documento WO2011/000949.

La composición que contenía celulosa TEC-210 se produce según los procedimientos tales como inoculación y fermentación que se describen en el documento WO2011/000949.

La beta-glucosidasa (BG) se produce mediante sobreexpresión de EBA4 en *Aspergillus niger* según se describe en el documento WO2011/098577, seguido por fermentación del transformante de *Aspergillus niger*. EBA4 es una BG de *Rasamsonia emersonii (Talaromyces emersonii)* y se identifica en el documento WO2011/098577 como beta-glucosidasa (BG) de *T. emersonii* y se representa mediante SEQ ID NO: 5 en el documento WO2011/098577.

Celluclast (*Trichoderma* celulasa) se obtuvo de Sigma

### Técnicas de biología molecular

En estas cepas, usando técnicas de biología molecular conocidas por los expertos (véase: Sambrook & Russell, Molecular Cloning: A Laboratory Manual, 3ª Ed., CSHL Press, Cold Spring Harbor, NY, 2001), varios genes se sobreexpresaron y otros se regularon a la baja según se describe posteriormente. Ejemplos del diseño general de vectores de expresión para la sobreexpresión génica y vectores de alteración para la regulación a la baja, la transformación, el uso de marcadores y medios selectivos se pueden encontrar en los documentos WO199846772, WO199932617, WO2001121779, WO2005095624, WO2006040312, EP635574B, WO2005100573, WO2011009700, WO2012001169 y WO2011054899. Todos los vectores de sustitución génica comprenden regiones de flaqueo de aproximadamente 1 - 2 kb de las secuencias de ORF respectivas, para orientarse para la recombinación homóloga en los locus genómicos predestinados. Además, los vectores de *A. niger* contienen el marcador de selección *amdS* bidireccional de *A. nidulans* para la transformación, en repeticiones directas intermedias. El método aplicado para la eliminación génica en todos los ejemplos presentes usa ADN lineal, que se integra en el genoma en el locus homólogo de las secuencias de flaqueo mediante un cruce doble, sustituyendo así el gen que se va a eliminar por el gen *amdS*. Después de la transformación, las repeticiones directas permiten la retirada del marcador de selección mediante un (segundo) episodio de recombinación homóloga. La retirada del marcador *amdS* se puede realizar al sembrar en placa sobre medio de fluoroacetamida, dando como resultado la selección de cepas libres de gen marcador. Usando esta estrategia de transformación y la contraselección posterior, que también se describe como el enfoque "LIBRE DE GEN MARCADOR" en el documento EP 0 635 574, el marcador *amdS* se puede usar indefinidamente en programas de modificación de cepas.

### Medios y soluciones:

Agar de dextrosa de patata, PDA, (Fluka, N° Cat. 70139): por litro: Extracto de patata 4 g; Dextrosa 20 g; Bactoagar 15 g; pH 5,4; Esterilícese 20 min a 120°C.

Medio de agar de Rasamsonia: por litro: Fracción salina n° 3 15 g; Celulosa 30 g; Bactopeptona 7,5 g; Harina de cereal 15 g; KH<sub>2</sub>PO<sub>4</sub> 5 g; CaCl<sub>2</sub>.2agua 1 g; Bactoagar 20 g; pH 6,0; Esterilícese 20 min a 120°C.

Composición de la fracción salina: La "fracción salina n° 3" se ajustaba a la divulgación del documento WO98/37179, Tabla 1. Las desviaciones de la composición de esta tabla eran CaCl<sub>2</sub>.2agua 1,0 g/l, KCl 1,8 g/l, ácido cítrico 1ac 0,45 g/l (agente quelante).

### Medios de matraz agitador para *Rasamsonia*

Medio para Rasamsonia 1: por litro: Glucosa 20 g; Extracto de levadura (Difco) 20 g; Clerol FBA3107 (AF) 4 gotas; pH 6.0; Esterilícese 20 min a 120°C.

Medio para Rasamsonia 2: por litro: Fracción salina n° 3 15 g; Celulosa 20 g; Bactopeptona 4 g; Harina de cereal 7,5 g; KH<sub>2</sub>PO<sub>4</sub> 10 g; CaCl<sub>2</sub>.2H<sub>2</sub>O 0,5 g; Clerol FBA3107 (AF) 0,4 ml; pH 5; Esterilícese 20 min a 120°C.

Medio para Rasamsonia 3: por litro: Fracción salina n° 3 15 g; glucosa 50 g; Bactopeptona 7,5 g; KH<sub>2</sub>PO<sub>4</sub> 10 g; CaCl<sub>2</sub>.2H<sub>2</sub>O 0,5 g; Clerol FBA3107 (AF) 0,4 ml; pH 5; Esterilícese 20 min a 120°C.

**Preparación discontinua de esporas para *Rasamsonia***

5 Las cepas se hicieron crecer a partir de soluciones madre sobre medio de agar para *Rasamsonia* en placas de Petri de 10 cm de diámetro durante 5-7 días a 40°C. Para fermentaciones en MTP, las cepas se hicieron crecer en placas de 96 pocillos que contenían medio de agar para *Rasamsonia*. Las soluciones madre de las cepas se almacenaron a -80°C en glicerol al 10%.

**Aislamiento de ADN cromosómico**

Las cepas se hicieron crecer en medio YGG (por litro: 8 g de KCl, 16 g de glucosa.H<sub>2</sub>O, 20 ml de extracto de levadura al 10%, 10 ml de 100xpen/estrep, 6,66 g de YNB+aminoácidos, 1,5 g de ácido cítrico y 6 g de K<sub>2</sub>HPO<sub>4</sub>) durante 16 horas a 42°C, 250 rpm, y el ADN cromosómico se aisló usando el DNeasy plant mini kit (Qiagen, Hilden, Alemania).

**10 Fermentación en MTP de *Rasamsonia***

15 Se usaron placas de microvaloración (MTP) de 96 pocillos con cepas de *R. emersonii* esporuladas para recolectar esporas para fermentaciones en MTP. Para hacer esto, se añadieron a cada pocillo 200 µl de medio para *Rasamsonia* 1 diluido 10 veces y, después de resuspender la mezcla, se incubaron 100 µl de suspensión de esporas en agitadores higrométricos (Infors) para 44°C a 550 rpm y 80% de humedad durante 16 horas. Posteriormente, se usaron 50 µl del precultivo para inocular 250 µl de medio para *Rasamsonia* 2 en placas MTP. Las placas de 96 pocillos se incubaron en agitadores higrométricos (Infors) para 44°C a 550 rpm, y 80% de humedad durante 6 días. Las placas se centrifugaron y los sobrenadantes se recolectaron.

**Protocolo de crecimiento de *Rasamsonia* en matraces agitadores**

20 Las esporas se inocularon directamente en matraces agitadores de 500 ml que contenían 100 ml de medio para *Rasamsonia* bien 2 o bien 3 y se incubaron a 45°C a 250 rpm en un agitador incubador durante 3-4 días. Alternativamente, las esporas se inocularon en matraces agitadores de 100 ml que contenían medio para *Rasamsonia* 1 y se incubaron a 45°C a 250 rpm en un agitador incubador durante 1 día (precultivo) y, posteriormente, 5 o 10 ml de biomasa procedente del precultivo se transfirieron a matraces agitadores de 500 ml que contenían 100 ml de medio para *Rasamsonia* 2 o 3 y se hicieron crecer bajo condiciones como las descritas anteriormente.

**25 Análisis de proteínas**

Se separaron muestras de proteínas bajo condiciones reductoras sobre NuPAGE 4-12% Bis-Tris gel (Invitrogen, Breda, Países Bajos) y se tiñeron. Los geles se tiñeron bien con InstantBlue (Expedeon, Cambridge, Reino Unido), bien con SimplyBlue safestain (Invitrogen, Breda, Países Bajos) o bien con Sypro Ruby (Invitrogen, Breda, Países Bajos) según las instrucciones del fabricante.

**30 Contenido de proteína total**

35 El contenido de proteínas del sobrenadante recuperado se determinó según el método de Bradford. La cantidad de proteína en las muestras de enzimas se determinó con el ensayo de proteínas de Bradford, usando reactivo para proteínas de Coomassie. Se mezclaron 25 µl de muestra de enzima apropiadamente diluida con 1,2 ml de reactivo de Coomassie. Después de 10 minutos a temperatura ambiente, la absorbancia de la mezcla a 595 nm se determinó usando un espectrofotómetro (Uvikon XL). El contenido de proteína se calculó en comparación con un estándar de BSA.

**Ensayo de actividad de liberación de azúcar desde materia prima de rastrojo de maíz pretratada con ácido**

40 Para cada condición de ensayo de (hemi-)celulasa, el sobrenadante del cultivo enzimático se analizó por duplicado según el siguiente procedimiento: 5 mg de proteína/g de materia prima seca del sobrenadante del cultivo enzimático se transfirieron a un vial adecuado que contenía 800 µl de materia seca al 2,5 % (p/p) de un sustrato de rastrojo de maíz suavemente pretratado con ácido en un tampón de citrato 50 mM, se tamponaron a pH 3,5 o pH 4,5 o 5,0. Adicionalmente, como una muestra en blanco, se añadió a otro vial la misma cantidad de sobrenadante de cultivo enzimático, donde los 800 µl de materia seca al 2,5% (p/p) de un sustrato de rastrojo de maíz suavemente pretratado con ácido en un tampón de citrato 50 mM se reemplazaron por 800 µl de tampón de citrato 50 mM, tamponado a pH 4,5. Las muestras de ensayo tamponadas a pH 3,5 se incubaron a 65°C durante 72 horas. Las muestras de ensayo tamponadas a pH 5,0 se incubaron a 50°C durante 72 horas. Las muestras de ensayo tamponadas a pH 4,5 y las muestras en blanco para la corrección del contenido de azúcares monoméricos en los sobrenadantes enzimáticos se

incubaron a 65°C durante 72 horas. Además, muestras de ensayo tamponadas a pH 4,5 se incubaron a 75°C durante 72 horas.

Además de las incubaciones individuales que se describen anteriormente, el sobrenadante de cultivo enzimático también se probó en combinación con dos mezclas de hemicelulasas diferentes; TEC-210 (*Rasamsonia emersonii*) a la que se añadía beta-glucosidasa (BG) adicional (cepa de *Aspergillus niger* que expresa una BG procedente de *Rasamsonia emersonii*) (0,08 mg/g de materia seca) y Celluclast (*Trichoderma reesei*) a la que se añadía BG (Novozym-188) adicional (0,08 mg/g de materia seca). Las mezclas se añadieron hasta una concentración de 1 mg de proteína/g de materia seca de la materia prima. Estas incubaciones se realizaron en las mismas condiciones que se describen anteriormente.

Para cada procedimiento, se realizó un ensayo en el que el sobrenadante enzimático se reemplazaba por agua desmineralizada, a fin de corregir posibles azúcares monoméricos presentes en la materia prima después de la incubación.

Después de la incubación de las muestras de ensayo, se añadió a cada vial un volumen fijado de un patrón interno, ácido maleico (20 g/l), EDTA (40 g/l) y DSS (5-sulfonato de 2,2-dimetil-2-silapentano) (0,5 g/l). Después de la centrifugación, se transfirieron 650 µl del sobrenadante a un nuevo vial.

El sobrenadante de las muestras se liofiliza durante la noche, posteriormente, se añaden 50 µl de D2O al residuo secado y se liofilizan una vez más. El residuo secado se disuelve en 600 µl de D2O. Se registran espectros de 1D 1H NMR unidimensionales en un Bruker Avance III HD 400 MHz, equipado con una criosonda enfriada con N<sub>2</sub>, usando un programa pulsátil sin supresión de agua a una temperatura de 17°C con un impulso de excitación de 90 grados, un tiempo de adquisición de 2,0 s y un retardo de relajación de 10 s. Las concentraciones de analito se calculan basándose en las siguientes señales ( $\delta$  relativa a DSS (ácido 4,4-dimetil-4-silapentano-1-sulfónico)): 1/2 del pico de  $\beta$ -glucosa a 4,63 ppm (d, 0,31 H, J = 8 Hz), 1/2 del pico de  $\beta$ -xilosa a 4,56 ppm (d, 0,315 H, J = 8 Hz), el pico de xiloligo a 4,45 ppm (d, 1H, J=8Hz), 1/2 del anómero  $\beta$  del extremo reductor del pico de celobiosa a 4,66 ppm (d, 0,31 H, J=8 Hz). El usuario de señal para el pico de patrón:ácido maleico a 6,26 ppm (s, 2H)

La solución de enzima (hemi)-celulasa puede contener azúcares residuales. Por lo tanto, los resultados del ensayo se corrigen con respecto al contenido de azúcar medido después de la incubación de la solución enzimática.

### **Medida de la actividad como $\beta$ -xilosidasa**

Este ensayo mide la liberación de p-nitrofenol mediante la acción de  $\beta$ -xilosidasa sobre p-nitrofenil- $\beta$ -D-xilopiranosido (PNPX). Una unidad de actividad como  $\beta$ -xilosidasa es la cantidad de enzima que libera 1 micromol de p-nitrofenol en un minuto a 60°C y pH 4,5. Se prepara tampón de acetato (0,1 M, pH 4,5) como sigue: se disuelven 8,2 g de acetato sódico anhidro en agua destilada de modo que el volumen final de la solución sea 1000 ml (Solución A). En un matraz separado, se mezclan 6,0 g (5,72 ml) de ácido acético glacial con agua destilada para formar el volumen total de 1000 ml (Solución B). El tampón de acetato 0,1 M final, pH 4,5, se prepara al mezclar la Solución A con la Solución B hasta que el pH de la solución resultante sea igual a 4,5. Se añade una gota (~ 25 µl) de Triton X-100 / l de solución tamponadora. Se usa PNPX (Sigma) como el sustrato de ensayo.

Se disuelven 100 mg de PNPX en 84 ml de tampón de acetato 0,1 M para obtener una solución madre 4,4 mM. El reactivo de parada (solución de carbonato sódico 1 M) se prepara como sigue: se disuelven 10,6 g de carbonato sódico anhidro en 50 ml de agua destilada y el volumen de la solución se ajusta hasta 100 ml. Este reactivo se usa para terminar la reacción enzimática.

Para la incubación con enzima, se mezclan 0,1 ml de solución madre de PNPX 4,4 mM con 0,1 ml de la muestra de enzima diluida apropiada y se incuban a 60°C durante 60 minutos. Después de 60 minutos de incubación, se mezclan 0,1 ml de la mezcla de reacción con 0,1 ml de solución de carbonato sódico 1 M y la absorbancia se mide a 405 nm en placas de microvaloración como A<sub>s</sub>.

Para el blanco del sustrato, se mezclan 0,1 ml de solución madre de PNPX 4,4 mM con 0,1 ml de tampón de acetato 0,1 M, pH 4,5 y se tratan igual que las muestras: se incuban a 60°C durante 60 minutos después de lo cual se mezclan 0,1 ml de la mezcla de reacción con 0,1 ml de solución de carbonato sódico 1 M y la absorbancia a 405 nm se mide en placas de microvaloración como A<sub>SB</sub>.

Los blancos de enzima (sin adición de sustrato) se miden para corregir el color de fondo que se origina a partir de las enzimas. Se mezclan 0,1 ml de la enzima disuelta apropiada con 0,1 ml de tampón de acetato 0,1 M, pH 4,5, y se incuban a 60°C durante 60 minutos. Después de 60 minutos de incubación, se mezclan 0,1 ml de la mezcla de reacción con 0,1 ml de solución de carbonato sódico 1 M y la absorbancia se mide a 405 nm en placas de microvaloración como A<sub>EB</sub>.

Una curva de calibración de p-nitrofenol (diluido apropiadamente en tampón de acetato 0,1 M, pH 4,5) mezclado en una relación de 1:1 con solución de carbonato sódico 1 M se usa para cuantificar su liberación del PNPX mediante la acción de la enzima.

- 5 Después de la incubación de la enzima con sustrato, se usa la absorbancia corregida ( $= A_S - A_{EB} - A_{SB}$ ), para calcular la cantidad de p-nitrofenol liberada por la enzima.

La actividad se expresa como la cantidad de enzima requerida para liberar p-nitrofenol 1  $\mu\text{M}/\text{min}$  bajo las condiciones de ensayo.

- 10 Este ensayo se puede usar para probar la actividad de enzimas tales como, pero no limitadas a, enzimas GH3, GH30, GH39, GH43, GH52, y GH54.

### **Ensayo de actividad como $\beta$ -xilosidasa 2**

Este ensayo mide la liberación de xilosa mediante la acción de  $\beta$ -xilosidasa sobre xilobiosa.

- 15 Se preparó tampón de acetato sódico (0,05 M, pH 4,5) como sigue. Se disolvieron 4,1 g de acetato sódico anhidro o 6,8 g de acetato sódico  $\cdot 3\text{H}_2\text{O}$  en agua destilada hasta un volumen final de 1000 ml (Solución A). En un matraz separado, se mezclaron 3,0 g (2,86 ml) de ácido acético glacial con agua destilada para formar el volumen total de 1000 ml (Solución B). El tampón de acetato sódico 0,05 M final, pH 4,5, se preparó al mezclar la Solución A con la Solución B hasta que el pH de la solución resultante fuera igual a 4,5. La xilobiosa se adquirió de Sigma y se disolvió en tampón de acetato sódico pH 4,5 hasta una concentración de 100  $\mu\text{g}/\text{ml}$ .
- 20

El ensayo se realizó como se detalla posteriormente.

- 25 El sobrenadante de cultivo enzimático se añadió al sustrato en una dosificación de 1 y 5 mg de proteína/g de sustrato que a continuación se incubó a 62°C durante 24 horas. La reacción se detuvo al calentar las muestras durante 10 minutos a 100°C. La liberación de xilosa se analizó mediante cromatografía de intercambio aniónico de alta resolución.

### **Blanco del sustrato**

- 30 En lugar de sobrenadante de cultivo enzimático, la misma cantidad de tampón se añadió a la solución de sustrato que a continuación se incubó a 62°C durante 24 horas. La reacción se detuvo al calentar las muestras durante 10 minutos a 100°C. La muestra se analizó mediante cromatografía de intercambio aniónico de alta resolución. El análisis se realizó usando un sistema de HPLC Dionex equipado con una columna Dionex CarboPac PA-1 (2 mm de diámetro interno x 250 mm) en combinación con una columna de seguridad CarboPac PA (2 mm de diámetro interno x 50 mm) y un detector de PAD Dionex (Dionex Co. Sunnyvale). Se usó un caudal de 0,3 ml/min con el siguiente gradiente de acetato sódico en NaOH 0,1 M: 0-20 min, 0-180 mM. Cada elución era seguida por una etapa de lavado de 5 min acetato sódico 1000 mM en NaOH 0,1 M y una etapa de equilibración de 15 min. NaOH 0,1 M.
- 35

- 40 En el caso de que estuvieran presentes compuestos interferentes que compliquen la identificación de xilosa, el análisis se realizó mediante marcha isocrática sobre  $\text{H}_2\text{O}$  durante un gradiente de 30 min (se añadió NaOH 0,5 M después de la columna en 0,1 ml/min para la detección) seguido por una etapa de lavado de 5 min acetato sódico 1000 mM en NaOH 0,1 M y una etapa de equilibración de 15 min  $\text{H}_2\text{O}$ .

Se usaron patrones de xilosa y xilobiosa (Sigma) para la identificación del sustrato y el producto formado por la enzima.

Este ensayo se puede usar para probar la actividad de enzimas tales como, pero no limitadas a, enzimas GH3, GH30, GH39, GH43, GH52 y GH54.

### **Ensayo de actividad como $\beta$ -xilosidasa 3**

Se realizó el mismo ensayo que se describe anteriormente con sustratos de xilano como arabinoxilano de avena, xilano de madera de haya y xilano de madera de abedul (Sigma) en lugar de xilobiosa para medir la actividad como xilosidasa sobre sustratos poliméricos.

- 50 Las condiciones de ensayo eran las mismas con la excepción de que todos los sustratos se disolvían hasta una concentración de 2 mg/ml. La incubación se realizó a 60°C durante 24 h a una dosificación de 10 mg/g. Después de la xilosa y la xilobiosa, también se cuantificaron xilotriosa y xilotetraosa.

**Medida de actividad como  $\alpha$ -galactosidasa**

Este ensayo mide la liberación de *p*-nitrofenol mediante la acción de  $\alpha$ -galactosidasa sobre *p*-nitrofenil- $\alpha$ -D-galactopiranosido (PNPG). Una unidad de actividad como  $\alpha$ -galactosidasa es la cantidad de enzima que emite 1 micromol de *p*-nitrofenol en un minuto a 60°C y pH 4,5. Se prepara tampón de acetato (0,1 M, pH 4,5) como sigue: se disuelven 8,2 g de acetato sódico anhidro en agua destilada de modo que el volumen final de la solución sea 1000 ml (Solución A). En un matraz separado, se mezclan 6,0 g (5,72 ml) de ácido acético glacial con agua destilada para formar el volumen total de 1000 ml (Solución B). El tampón de acetato 0,1 M final, pH 4,5, se prepara al mezclar la Solución A con la Solución B hasta que el pH de la solución resultante sea igual a 4,5. Se añade una gota (~ 25  $\mu$ l) de Triton X-100 / l de solución tamponadora. Se usa PNPG (Sigma) como el sustrato de ensayo.

Se elabora una solución madre de PNPG 4,4 mM en tampón de acetato 0,1 M. El reactivo de parada (solución de carbonato sódico 1 M) se prepara como sigue: se disuelven 10,6 g de carbonato sódico anhidro en 50 ml de agua destilada y el volumen de la solución se ajusta hasta 100 ml. Este reactivo se usa para terminar la reacción enzimática.

Para la incubación con enzima, se mezclan 0,1 ml de solución madre de PNPG 4,4 mM con 0,1 ml de la muestra de enzima diluida apropiada y se incuban a 60°C durante 60 minutos. Después de 60 minutos de incubación, se mezclan 0,1 ml de la mezcla de reacción con 0,1 ml de solución de carbonato sódico 1 M y la absorbancia se mide a 405 nm en placas de microvaloración como  $A_s$ .

Para el blanco del sustrato, se mezclan 0,1 ml de solución madre de PNPG 4,4 mM con 0,1 ml de tampón de acetato 0,1 M, pH 4,5, y se tratan igual que las muestras: se incuban a 60°C durante 60 minutos, después de lo cual se mezclan 0,1 ml de la mezcla de reacción con 0,1 ml de solución de carbonato sódico 1 M y la absorbancia a 405 nm se mide en placas de microvaloración como  $A_{SB}$ .

Se miden los blancos de la enzima (sin adición de sustrato) para corregir el color de fondo que se origina a partir de las enzimas. Se mezclan 0,1 ml de la muestra de enzima fluida apropiada con 0,1 ml de tampón de acetato 0,1 M, pH 4,5, y se incuban a 60°C durante 60 minutos. Después de 60 minutos de incubación, se mezclan 0,1 ml de la mezcla de reacción con 0,1 ml de solución de carbonato sódico 1 M y la absorbancia se mide a 405 nm en placas de microvaloración como  $A_{EB}$ .

Se usa una curva de calibración de *p*-nitrofenol (diluido apropiadamente en tampón de acetato 0,1 M, pH 4,5) mezclado en una relación de 1:1 con solución de carbonato sódico 1 M para cuantificar su liberación desde PNPG mediante la acción de la enzima.

Después de la incubación de enzima con sustrato, la absorbancia corregida ( $= A_s - A_{EB} - A_{SB}$ ) se usa para calcular la cantidad de *p*-nitrofenol liberada por la enzima.

La actividad se expresa como la cantidad de enzima requerida para liberar *p*-nitrofenol 1  $\mu$ M/min bajo las condiciones de ensayo.

Este ensayo se puede usar para probar la actividad de enzimas tales como, pero no limitadas a, enzimas GH4, GH27 y GH36.

**Ensayo de actividad como xiloglucanasa 1**

Se preparó tampón de acetato sódico (0,05 M, pH 4,5) como sigue. Se disolvieron 4,1 g de acetato sódico anhidro en agua destilada hasta un volumen final de 1000 ml (Solución A). En un matraz separado, se mezclaron 3,0 g (2,86 ml) de ácido acético glacial con agua destilada para formar el volumen total de 1000 ml (Solución B). El tampón de acetato sódico 0,05 M final, pH 4,5, se preparó al mezclar la Solución A con la Solución B hasta que el pH de la solución resultante fuera 4,5.

Se disolvió xiloglucano de tamarindo en tampón de acetato sódico para obtener 2,0 mg/ml. El sobrenadante de cultivo enzimático se añadió al sustrato en una dosificación de 10 mg de proteína/ g de sustrato que a continuación se incubó a 60°C durante 24 horas. La reacción se detuvo al calentar las muestras durante 10 minutos a 100°C. La liberación de oligosacáridos se analizó mediante cromatografía de intercambio aniónico de alta resolución.

Como una muestra en blanco, el sustrato se trató y se incubó del mismo modo pero entonces sin la adición de enzima.

Como una referencia, el sustrato también se incubó bajo las mismas condiciones con una preparación de celulasa comercial procedente de *Trichoderma Reesei* (Celluclast; Sigma) que se diluyó 50 veces, después de lo cual se añadieron a la incubación 20  $\mu$ l.

El análisis se realizó usando un sistema de HPLC Dionex equipado con una columna Dionex CarboPac PA-1 (2 mm de diámetro interno x 250 mm) en combinación con una columna de seguridad CarboPac PA (2 mm de diámetro

interno x 50 mm) y un detector de PAD Dionex (Dionex Co. Sunnyvale). Se usó un caudal de 0,3 ml/min con el siguiente gradiente de acetato sódico en NaOH 0,1 M: 0-40 min, 0-150 mM. Cada elución fue seguida por una etapa de lavado de 5 min acetato sódico 1000 mM en NaOH 0,1 M y una etapa de equilibración de 15 min NaOH 0,1 M.

- 5 Este ensayo se usó para probar la actividad de enzimas tales como, pero no limitadas a, enzimas GH5, GH12, GH16, GH44 y GH74.

### **Ensayo de actividad como xiloglucanasa 2**

- 10 El siguiente ejemplo ilustra el ensayo para medir la actividad como xiloglucanasa. Esta actividad se demostró al usar xiloglucano como sustrato y un ensayo de azúcares reductores (PAHBAH) como método de detección. Los valores se compararon con un patrón, que se preparó usando una preparación de celulasa comercial procedente de *Trichoderma Reesei* (Celluclast; Sigma).

- 15 Reactivo A: se suspendieron 5 g de hidrazida de ácido p-hidroxibenzoico (PAHBAH) en 60 ml de agua, se añadieron 4,1 ml de ácido clorhídrico concentrado y el volumen se ajustó hasta 100 ml. Reactivo B: se disolvieron 24,9 g de citrato trisódico en 500 ml de agua. Se añadieron a esta solución 2,2 g de cloruro cálcico y 40 g de hidróxido sódico. El volumen se ajustó hasta 2 l con agua. Ambos reactivos se almacenaron a temperatura ambiente. Reactivo de trabajo: se añadieron 10 ml de Reactivo A a 40 ml de Reactivo B. Esta solución se preparó recientemente todos los días y se almacenó sobre hielo entre usos. Usando los reactivos anteriores, el ensayo se realizó según se detalla posteriormente

- 20 Junto con el xiloglucano, también se usó carboximetilcelulosa como un sustrato para determinar la especificidad de la enzima.

- 25 Después de la incubación, 10 µl de cada pocillo se mezclaron con 200 µl de reactivo de trabajo. Estas soluciones se calentaron a 70°C durante 30. Después de enfriar, las muestras se analizaron al medir la absorbancia a 405 nm. Se usó glucosa como un patrón para cuantificar extremos reductores formados como equivalentes de glucosa.

- 30 Como controles, los sustratos también se incubaron sin la adición de sobrenadante de cultivo enzimático y los sobrenadantes de cultivo enzimático se incubaron sin sustrato.

Este ensayo se puede usar para probar la actividad de enzimas tales como, pero no limitadas a, enzimas GH5, GH12, GH16, GH44 y GH74.

### **Ensayo de actividad como xiloglucanasa 3**

- 35 Se prepara tampón de acetato sódico (0,05 M, pH 4,5) como sigue. Se disolvieron 4,1 g de acetato sódico anhidro en agua destilada hasta un volumen final de 1000 ml (Solución A). En un matraz separado, se mezclan 3,0 g (2,86 ml) de ácido acético glacial con agua destilada para formar el volumen total de 1000 ml (Solución B). El tampón de acetato sódico 0,05 M final, pH 4,5, se prepara al mezclar la Solución A con la Solución B hasta que el pH de la solución resultante sea 4,5.

- 40 Se disuelve xiloglucano de tamarindo en tampón de acetato sódico para obtener 2,0 mg/ml. La enzima se añade al sustrato en una dosificación de 10 mg de proteína / g de sustrato que a continuación se incuba a 60°C durante 24 horas. La reacción se detiene al calentar las muestras durante 10 minutos a 100°C. La formación de oligosacáridos de peso molecular inferior se analiza mediante cromatografía de exclusión por tamaño de alta resolución.

Como una muestra en blanco, el sustrato se trata y se incuba del mismo modo pero entonces sin la adición de enzima.

- 45 Como una referencia, el sustrato también se incuba bajo las mismas condiciones con una preparación de celulasa comercial procedente de, p. ej., *Aspergillus niger* o *Trichoderma Reesei* (el patrón de celulasa a su propia temperatura óptima en caso de inactividad a 60°C).

- 50 El análisis se realiza usando cromatografía de exclusión por tamaño de alta resolución (HPSEC) realizada en tres columnas de gel de TSK (6,0 mm x 15,0 cm por columna) en serie SuperAW4000, SuperAW3000, SuperAW2500; Tosoh Bioscience), en combinación con una columna de seguridad PWX (Tosoh Bioscience). La elución se realiza a 55°C con nitrato sódico 0,2 M a 0,6 ml/min. El eluido se comprobó usando un detector del índice de refracción (RI) Shodex RI-101 (Kawasaki). La calibración se realizó al usar pululanos (Associated Polymer Labs Inc., Nueva York, EE. UU. de A.) con un peso molecular en el intervalo de 0,18-788 kDa.

- 55 Este ensayo se puede usar para probar la actividad de enzimas tales como, pero no limitadas a, enzimas GH5, GH12, GH16, GH44 y GH74.

**Ensayo de actividad como  $\alpha$ -arabinofuranosidasa**

El siguiente ejemplo ilustra un ensayo para medir la capacidad de  $\alpha$ -arabinofuranosidasas para retirar los residuos de  $\alpha$ -L-arabinofuranosilo desde residuos de xilosa sustituidos.

- 5 Para la degradación completa de arabinoxilanos hasta arabinosa y xilosa, se necesitan varias actividades enzimáticas, incluyendo endoxilanasas y arabinofuranosidasas. La molécula de arabinoxilano procedente de trigo está altamente sustituida con residuos de arabinosilo. Estos pueden estar sustituidos bien en la posición C2 o bien la C3 del residuo de xilosilo (sustitución individual), o tanto en la posición C2 como la C3 de la xilosa (sustitución doble).
- 10 Se prepararon oligosacáridos sustituidos individualmente y doblemente al incubar arabinoxilano de trigo (WAX; 10 mg/ml; Megazyme, Bray, Irlanda) en tampón de acetato 50 mM pH 4,5 con una cantidad apropiada de endoxilanasas (procedente de *Aspergillus awamori*, Kormelink F. y cols; Journal of Biotechnology (1993) 27: 249-265) 48 horas a 40°C para producir una cantidad suficiente de arabinoxilooligosacáridos. La reacción se detuvo al calentar las muestras a 100°C durante 10 minutos. Las muestras se centrifugaron durante 5 minutos a 10.000 x g. El sobrenadante se usó
- 15 para experimentos adicionales. La degradación del arabinoxilano estaba seguida por análisis de los azúcares reductores formados y cromatografía de intercambio aniónico de alta resolución (HPAEC).

- El sobrenadante de cultivo enzimático se añadió a los arabinoxilooligosacáridos sustituidos individualmente y doblemente (WAX tratado con endoxilanasas; 2 mg/ml) en una dosificación de 10 mg de proteína/g de sustrato en tampón de acetato sódico 50 mM que a continuación se incubó a 65°C durante 24 horas. La reacción se detuvo al calentar las muestras a 100°C durante 10 minutos. Las muestras se centrifugaron durante 5 minutos a 10.000 x g. La liberación de arabinosa fue seguida por análisis de HPAEC.
- 20

- El análisis se realizó usando un sistema de HPLC Dionex equipado con una columna Dionex CarboPac PA-1 (2 mm de diámetro interno x 250 mm) en combinación con una columna de seguridad CarboPac PA (2 mm de diámetro interno x 50 mm) y un detector de PAD Dionex (Dionex Co. Sunnyvale). Se usó un caudal de 0,3 ml/min con el siguiente gradiente de acetato sódico en NaOH 0,1 M: 0-40 min, 0-400 mM. Cada elución fue seguida por una etapa de lavado de 5 min acetato sódico 1000 mM en NaOH 0,1 M y una etapa de equilibración de 15 min NaOH 0,1 M. La liberación de arabinosa se identificó y se cuantificó mediante un patrón (Sigma).
- 25

- 30 Este ensayo se puede usar para probar la actividad de enzimas tales como, pero no limitadas a, enzimas GH3, GH43, GH51, GH54 y GH62.

**Ensayo de actividad como endoxilanasas**

- 35 Las endoxilanasas son enzimas capaces de hidrolizar una unión  $\beta$ -1,4 en el esqueleto de xilano, produciendo xilooligosacáridos cortos. Este ensayo mide la liberación de xilosa y xilooligosacáridos mediante la acción de xilanasas sobre arabinoxilano de trigo (WAX) (Megazyme, Viscosidad media 29 cSt), arabinoxilano de avena, xilano de madera de haya y xilano de madera de abedul (Sigma).

- 40 Se preparó tampón de acetato sódico (0,05 M, pH 4,5) como sigue; se disolvieron 4,1 g de acetato sódico anhidro en agua destilada hasta un volumen final de 1000 ml (Solución A). En un matraz separado, se mezclaron 3,0 g (2,86 ml) de ácido acético glacial con agua destilada para formar el volumen total de 1000 ml (Solución B). El tampón de acetato sódico 0,05 M final, pH 4,5, se preparó al mezclar la Solución A con la Solución B hasta que el pH de la solución resultante fuera 4,5. Cada sustrato se disolvió en tampón de acetato sódico para obtener 2,0 mg/ml. El sobrenadante de cultivo enzimático se añadió al sustrato en una dosificación de 10 mg de proteína/g de sustrato que a continuación se incubó a 60°C durante 20 horas. La reacción se detuvo al calentar las muestras durante 10 minutos a 100°C. La liberación de xilosa y xilooligosacáridos se analizó mediante cromatografía de intercambio aniónico de alta resolución.
- 45

Como una muestra en blanco, el sustrato se trató y se incubó del mismo modo pero entonces sin la adición de enzima.

- 50 El análisis se realizó usando un sistema de HPLC Dionex equipado con una columna Dionex CarboPac PA-1 (2 mm de diámetro interno x 250 mm) en combinación con una columna de seguridad CarboPac PA (2 mm de diámetro interno x 50 mm) y un detector de PAD Dionex (Dionex Co. Sunnyvale). Se usó un caudal de 0,3 ml/min con el siguiente gradiente de acetato sódico en NaOH 0,1 M: 0-40 min, 0-400 mM. Cada elución fue seguida por una etapa de lavado de 5 min acetato sódico 1000 mM en NaOH 0,1 M y una etapa de equilibración de 15 min NaOH 0,1 M. Se usaron patrones de xilosa, xilobiosa y xilotriosa (Sigma) para identificar estos oligómeros liberados por la acción de la enzima.
- 55

Este ensayo se puede usar para probar la actividad de enzimas tales como, pero no limitadas a, enzimas GH5, GH8, GH10 y GH11.

**Medida de la actividad como  $\alpha/\beta$ -xilosidasa**

Este ensayo mide la liberación de *p*-nitrofenol mediante la acción de  $\alpha/\beta$ -xilosidasa sobre *p*-nitrofenil- $\alpha/\beta$ -D-xilopiranosido (PNPX). Una unidad de actividad como  $\beta$ -xilosidasa es la cantidad de enzima que emite 1 micromol de *p*-nitrofenol en un minuto a 60°C y pH 4,5. Se prepara tampón de acetato (0,1 M, pH 4,5) como sigue: se disuelven 8,2 g de acetato sódico anhidro en agua destilada de modo que el volumen final de la solución sea 1000 ml (Solución A). En un matraz separado, se mezclan 6,0 g (5,72 ml) de ácido acético glacial con agua destilada para formar el volumen total de 1000 ml (Solución B). El tampón de acetato 0,1 M final, pH 4,5, se prepara al mezclar la Solución A con la Solución B hasta que el pH de la solución resultante sea igual a 4,5. Se añade una gota (~ 25  $\mu$ l) de Triton X-100/l de solución tamponadora. Se usa PNPX (Sigma) como el sustrato de ensayo.

5 Se disuelven 100 mg de PNPX en 84 ml de tampón de acetato 0,1 M para obtener una solución madre 4,4 mM. El reactivo de parada (solución de carbonato sódico 1 M) se prepara como sigue: se disuelven 10,6 g de carbonato sódico anhidro en 50 ml de agua destilada y el volumen de la solución se ajusta hasta 100 ml. Este reactivo se usa para terminar la reacción enzimática.

10 Para la incubación con enzima, se mezclan 0,1 ml de solución madre de PNPX 4,4 mM con 0,1 ml de la muestra de enzima diluida apropiada y se incuban a 60°C durante 60 minutos. Después de 60 minutos de incubación, se mezclan 0,1 ml de la mezcla de reacción con 0,1 ml de solución de carbonato sódico 1 M y la absorbancia se mide a 405 nm en placas de microvaloración como  $A_s$ .

15 Para el blanco del sustrato, se mezclan 0,1 ml de solución madre de PNPX 4,4 mM con 0,1 ml de tampón de acetato 0,1 M, pH 4,5 y los mismos se tratan como las muestras: se incuban a 60°C durante 60 minutos después de lo cual se mezclan 0,1 ml de la mezcla de reacción con 0,1 ml de solución de carbonato sódico 1 M y la absorbancia a 405 nm se mide en placas de microvaloración como  $A_{SB}$ .

20 Los blancos de la enzima (sin adición de sustrato) se miden para corregir el color de fondo que se origina a partir de las enzimas. Se mezclan 0,1 ml de la muestra de enzima diluida apropiada con 0,1 ml de tampón de acetato 0,1 M, pH 4,5, y se incuban a 60°C durante 60 minutos. Después de 60 minutos de incubación, se mezclan 0,1 ml de la mezcla de reacción con 0,1 ml de solución de carbonato sódico 1 M y la absorbancia se mide a 405 nm en placas de microvaloración como  $A_{EB}$ .

25 Se usa una curva de calibración de *p*-nitrofenol (diluido apropiadamente en tampón de acetato 0,1 M, pH 4,5) mezclado en una relación de 1:1 con solución de carbonato sódico 1 M para cuantificar su liberación desde PNPX por la acción de la enzima.

30 Después de la incubación de enzima con sustrato, se usa la absorbancia corregida (=  $A_s - A_{EB} - A_{SB}$ ) para calcular la cantidad de *p*-nitrofenol liberada por la enzima.

35 La actividad se expresa como la cantidad de enzima requerida para liberar *p*-nitrofenol 1  $\mu$ M/min bajo las condiciones de ensayo.

40 Este ensayo se puede usar para probar la actividad de enzimas tales como, pero no limitadas a, enzimas GH3, GH30, GH31, GH39, GH43, GH52 y GH54.

**Medida de la actividad como  $\alpha/\beta$ -manosidasa**

Este ensayo mide la liberación de *p*-nitrofenol mediante la acción de  $\alpha/\beta$ -manosidasa sobre *p*-nitrofenil- $\alpha/\beta$ -D-manopiranosido (PNPM). Una unidad de actividad como  $\alpha/\beta$ -manosidasa es la cantidad de enzima que libera 1 micromol de *p*-nitrofenol en un minuto a 60°C y pH 4,5. Se prepara tampón de acetato (0,1 M, pH 4,5) como sigue: se disuelven 8,2 g de acetato sódico anhidro en agua destilada de modo que el volumen final de la solución sea 1000 ml (Solución A). En un matraz separado, se mezclan 6,0 g (5,72 ml) de ácido acético glacial con agua destilada para formar el volumen total de 1000 ml (Solución B). El tampón de acetato 0,1 M final, pH 4,5 se prepara al mezclar la Solución A con la Solución B hasta que el pH de la solución resultante sea igual a 4,5. Se añade una gota (~ 25  $\mu$ l) de Triton X-100/l de solución tamponadora. Se usa PNPM (Sigma) como el sustrato de ensayo.

Se elabora una solución madre de PNPM 4,4 mM en tampón de acetato 0,1 M. El reactivo de parada (solución de carbonato sódico 1 M) se prepara como sigue: se disuelven 10,6 g de carbonato sódico anhidro en 50 ml de agua destilada y el volumen de la solución se ajusta hasta 100 ml. Este reactivo se usa para terminar la reacción enzimática.

Para la incubación con enzima, se mezclan 0,1 ml de solución madre de PNPM 4,4 mM con 0,1 ml de la muestra de enzima diluida apropiada y se incuban a 60°C durante 60 minutos. Después de 60 minutos de incubación, se mezclan 0,1 ml de la mezcla de reacción con 0,1 ml de solución de carbonato sódico 1 M y la absorbancia se mide a 405 nm en placas de microvaloración como  $A_s$ .

Para el blanco del sustrato, se mezclan 0,1 ml de solución madre de PNPM 4,4 mM con 0,1 ml de tampón de acetato 0,1 M, pH 4,5 y se tratan igual que las muestras: se incuban a 60°C durante 60 minutos después de lo cual se mezclan 0,1 ml de la mezcla de reacción con 0,1 ml de solución de carbonato sódico 1 M y la absorbancia a 405 nm se mide en placas de microvaloración como  $A_{SB}$ .

Se miden los blancos de enzima (sin adición de sustrato) para corregir el color de fondo que se origina a partir de las enzimas. Se mezclan 0,1 ml de la muestra de enzima diluida apropiada con 0,1 ml de tampón de acetato 0,1 M, pH 4,5, y se incubaron a 60°C durante 60 minutos. Después de 60 minutos de incubación, se mezclan 0,1 ml de la mezcla de reacción con 0,1 ml de solución de carbonato sódico 1 M y la absorbancia se mide a 405 nm en placas de microvaloración como  $A_{EB}$ .

Se usa una curva de calibración de *p*-nitrofenol (diluido apropiadamente en tampón de acetato 0,1 M, pH 4,5) mezclado en una relación de 1:1 con solución de carbonato sódico 1 M para cuantificar su liberación desde PNPM por la acción de la enzima.

Después de la incubación de enzima con sustrato, se usa la absorbancia corregida ( $= A_S - A_{EB} - A_{SB}$ ) para calcular la cantidad de *p*-nitrofenol liberada por la enzima.

La actividad se expresa como la cantidad de enzima requerida para liberar *p*-nitrofenol 1  $\mu$ M/min bajo las condiciones de ensayo.

Este ensayo se puede usar para probar la actividad de enzimas tales como, pero no limitadas a, enzimas GH1, GH2, GH5, GH38, GH47, GH92 y GH125.

#### **Medida de la actividad como feruloilesterasa**

Sustratos sintéticos: El cafeato de metilo, el cumarato de metilo, el sinapinato de metilo y el ferulato de metilo se obtienen de Apin Chemicals. La actividad hacia estos sustratos sintéticos se determina al incubar la enzima con el sustrato en una dosificación de aproximadamente 5 mg/g de DM a un pH de 5,0 (tampón de acetato sódico 50 mM). La reacción se realizará a 60°C durante hasta 24 h.

Al final de la incubación, las muestras se hierven durante 5 minutos para inactivar las enzimas y se centrifugan a temperatura ambiente (10 min, 10.000  $\times$  g). La liberación de ácido hidroxicinámico desde el sustrato se mide mediante análisis por RP-UHPLC-MS en modo de ionización negativo según se describe previamente (Appeldoorn y cols., 2010) en un sistema de UHPLC Accela (Thermo Scientific) equipado con una columna Hypersyl GOLD (2,1 mm  $\times$  150 mm, 1,9  $\mu$ m de tamaño de partícula; Thermo Scientific). La fase móvil está compuesta por (A) H<sub>2</sub>O + 1% (v/v) de acetonitrilo + 0,2% (v/v) de ácido acético y (B) acetonitrilo + 0,2% (v/v) de ácido acético. El caudal es 0,4 ml/min y la temperatura de la columna es 30°C. El perfil de elución es como sigue: primeros 5 min, isocrático 0% de B; 5-23 min, lineal de 0 a 50% de B; 23-24 min, lineal de 50 a 100% de B; 24-27 min, isocrático a 100% de B; 27-28 min, lineal de 100 a 0% de B, seguido por reacondicionamiento de la columna durante 7 min. Los datos espectrales se recogen de 200 a 600 nm y la cuantificación se realiza a 320 nm. Los contenidos de ácidos ferúlico, cafeico, sinápico y cumárico se identifican y se cuantifican basándose en patrones.

Se recogen datos de MS en el modo negativo con un voltaje de pulverización iónica de 3,5 kV, un voltaje capilar de -20 V y una temperatura capilar de 350°C. Se realizan barridos de MS completos dentro del intervalo m/z 150-1500, y se obtienen los datos de MS<sub>2</sub> de los iones más intensos.

Este ensayo se puede usar para probar la actividad de enzimas tales como, pero no limitadas a, enzimas CE1.

#### **Medida de la actividad como feruloilesterasa**

Sustrato presente en la naturaleza: Oligómeros de arabinosilano purificados a partir de fibra de maíz (CF) pretratada (1 mg/ml cada uno) (Appeldoorn y cols 2010) se incuban con esterases de ácido ferúlico en una dosificación de aproximadamente 5 mg/g de DM a un pH de 5,0 (tampón de acetato sódico 50 mM). La reacción se realizará a 60°C durante hasta 24 h.

Al final de la incubación, las muestras se hierven durante 5 minutos para inactivar las enzimas y se centrifugan a temperatura ambiente (10 min, 10.000  $\times$  g). La liberación de ácido hidroxicinámico del sustrato se mide mediante análisis por RP-UHPLC-MS en modo de ionización negativo según se describe previamente (Appeldoorn y cols., 2010) en un sistema de UHPLC Accela (Thermo Scientific) equipado con una columna Hypersyl GOLD (2,1 mm  $\times$  150 mm, 1,9  $\mu$ m de tamaño de partícula; Thermo Scientific). La fase móvil está compuesta por (A) H<sub>2</sub>O + 1% (v/v) de acetonitrilo + 0,2% (v/v) de ácido acético y (B) acetonitrilo + 0,2% (v/v) de ácido acético. El caudal es 0,4 ml/min y la temperatura de la columna es 30°C. El perfil de elución es como sigue: primeros 5 min, isocrático 0% de B; 5-23 min, lineal de 0 a

50% de B; 23-24 min, lineal de 50 a 100% de B; 24-27 min, isocrático a 100% de B; 27-28 min, lineal de 100 a 0% de B, seguido por reacondicionamiento de la columna durante 7 min. Se recogen datos espectrales de 200 a 600 nm, y la cuantificación se realiza a 320 nm. Los contenidos de ácidos ferúlico y cumárico se identifican y cuantifican basándose en patrones.

5 Se recogen datos de MS en el modo negativo con un voltaje de pulverización iónica de 3,5 kV, un voltaje capilar de -20 V y una temperatura capilar de 350°C. Se realizaron barridos de MS completos dentro del intervalo m/z 150-1500, y se obtienen datos de MS2 de los iones más intensos.

10 La cantidad total de ácido ferúlico con enlaces éster en oligómeros de maíz se determinó después de hidrólisis alcalina y extracción con éter etílico usando el método de UHPLC descrito anteriormente.

Referencia: MAAIKE M. APPELDOORN y cols, J. Agric. Food Chem. 2010, 58, 11294-11301

15 Este ensayo se puede usar para probar la actividad de enzimas tales como, pero no limitadas a, enzimas CE1.

#### **Ensayo de actividad como $\alpha$ -glucuronidasa**

El siguiente ejemplo ilustra el ensayo para medir la actividad como  $\alpha$ -glucuronidasa hacia ácidos aldourónicos (Megazyme). Este ensayo mide la liberación de xilosa y xilooligómeros por la acción de la  $\alpha$ -glucuronidasa sobre los oligosacáridos de glucuronoxilano.

20 Se preparó tampón de acetato sódico (0,05 M, pH 4,5) como sigue. Se disolvieron 4,1 g de acetato sódico anhidro en agua destilada hasta un volumen final de 1000 ml (Solución A). En un matraz separado, se mezclaron 3,0 g (2,86 ml) de ácido acético glacial con agua destilada para formar el volumen total de 1000 ml (Solución B). El tampón de acetato sódico 0,05 M final, pH 4,5, se preparó al mezclar la Solución A con la Solución B hasta que el pH de la solución resultante fuera 4,5.

30 Para determinar la actividad sobre oligómeros pequeños, los ácidos aldourónicos se disuelven en tampón de acetato sódico para obtener 1,0 mg/ml. El sobrenadante de cultivo enzimático se añadió al sustrato en una dosificación de 1 y 10 mg de proteína/g de sustrato que a continuación se incubó a 60°C durante 24 horas. La reacción se detuvo al calentar las muestras durante 10 minutos a 100°C. La liberación de xilooligómeros como resultado de la retirada de ácido 4-O-metilglucurónico se analizó mediante cromatografía de intercambio aniónico de alta resolución

Como una muestra en blanco, el sustrato se trató y se incubó del mismo modo pero entonces sin la adición de enzima.

35 El análisis se realizó usando un sistema de HPLC Dionex equipado con una columna Dionex CarboPac PA-1 (2 mm de diámetro interno x 250 mm) en combinación con una columna de seguridad CarboPac PA (2 mm de diámetro interno x 50 mm) y un detector de PAD Dionex (Dionex Co. Sunnyvale). Se usó un caudal de 0,3 ml/min con el siguiente gradiente de acetato sódico en NaOH 0,1 M: 0-40 min, 0-400 mM. Cada elución fue seguida por una etapa de lavado de 5 min acetato sódico 1000 mM en NaOH 0,1 M y una etapa de equilibración de 15 min NaOH 0,1 M.

40 Se usaron patrones de xilosa, xilobiosa y xilotriosa (Sigma) para identificar los xilooligómeros liberados por la acción de la enzima que retira el ácido 4-O-metilglucurónico de estos oligómeros.

45 Este ensayo se puede usar para probar la actividad de enzimas tales como, pero no limitadas a, enzimas GH67 y GH115.

#### **Ejemplo 1: Construcción de vectores de expresión de *A. niger***

50 Este Ejemplo describe la construcción de un constructo de expresión para la sobreexpresión de Temer00088, Temer09484, Temer08028, Temer02362, Temer08862, Temer04790, Temer05249, Temer06848, Temer02056, Temer03124, Temer09491, Temer06400, Temer08570, Temer08163 o Temer07305 en *A. niger*. ADN genómico de la cepa de *Rasamsonia emersonii* CBS393.64 se secuenció y analizó. Se identificaba el gen con proteína traducida anotado como actividad según la Tabla 1. Secuencias del gen Temer00088, Temer09484, Temer08028, Temer02362, Temer08862, Temer04790, Temer05249, Temer06848, Temer02056, Temer03124, Temer09491, Temer06400, Temer08570, Temer08163 y Temer07305 de *R. emersonii*, que comprenden la secuencia de ORF optimizada para el par de codones, la secuencia proteínica, la secuencia de señalización, la secuencia genómica y la secuencia de ADNc silvestre se muestran en los listados de secuencias SEQ ID NO: 1 a 75.

#### **Construcción de plásmidos de expresión**

La secuencia que tiene SEQ ID NO: 1, 6, 11, 16, 21, 26, 31, 36, 41, 46, 51, 56, 61, 66 o 71 se clona en el vector pGBTOP (Fig. 1) usando los sitios *EcoRI* y *PacI*, que comprenden la secuencia promotora y terminadora de

glucoamilasa. La parte de *E.coli* se retiró mediante digestión con *NotI* antes de la transformación de CBS 513.88 de *A. niger*.

#### Transformación de *A. niger* y fermentaciones en matraces agitadores

5 La cepa de *A. niger* CBS513.88 se cotransforma con los constructos de expresión y un plásmido que contiene  
 10 marcador de selección apropiado (*amdS* o oleomicina) según el método descrito en la sección de información  
 experimental. De cepas de *A.niger* recombinantes y de control se genera una gran partida de esporas al sembrar  
 esporas o micelios en placas de PDA (agar de dextrosa de patata, Oxoid), preparadas según las instrucciones del  
 fabricante. Después del crecimiento durante 3-7 días a 30 grados Celsius, las esporas se recogen después de añadir  
 15 a las placas 0,01% de Triton X-100. Después de lavar con agua estéril, aproximadamente  $10^7$  esporas de  
 transformantes seleccionados y cepas de control se inoculan en matraces agitadores de 100 ml con pantallas que  
 contienen 20 ml de medio de precultivo líquido que consiste en, por litro: 30 g de maltosa.H<sub>2</sub>O; 5 g de extracto de  
 levadura; 10 g de caseína hidrolizada; 1 g de KH<sub>2</sub>PO<sub>4</sub>; 0,5 g de MgSO<sub>4</sub>.7H<sub>2</sub>O; 0,03 g de ZnCl<sub>2</sub>; 0,02 g de CaCl<sub>2</sub>; 0,01  
 g de MnSO<sub>4</sub>.4H<sub>2</sub>O; 0,3 g de FeSO<sub>4</sub>.7H<sub>2</sub>O; 3 g de Tween 80; 10 ml de penicilina (5000 IU/ml)/estreptomocina (5000  
 UG/ml); pH 5,5. Estos cultivos se desarrollaron a 34 grados Celsius durante 16-24 horas. Se inocularon 10 ml de este  
 cultivo a matraces agitadores de 500 ml con pantallas que contienen 100 ml de medio de fermentación que consiste  
 20 en, por litro: 70 g de glucosa.H<sub>2</sub>O; 25 g de caseína hidrolizada; 12,5 g de extracto de levadura; 1 g de KH<sub>2</sub>PO<sub>4</sub>; 2 g de  
 K<sub>2</sub>SO<sub>4</sub>; 0,5 g de MgSO<sub>4</sub>.7H<sub>2</sub>O; 0,03 g de ZnCl<sub>2</sub>; 0,02 g de CaCl<sub>2</sub>; 0,01 g de MnSO<sub>4</sub>.4H<sub>2</sub>O; 0,3 g de FeSO<sub>4</sub>.7H<sub>2</sub>O; 10 ml  
 de penicilina (5000 IU/ml)/estreptomocina (5000 UG/ml); ajustados hasta pH 5,6. Estos cultivos se desarrollan a 34  
 grados Celsius hasta que toda la glucosa se agotaba (habitualmente después de 4-7 días). Las muestras tomadas del  
 caldo de fermentación se centrifugan (10 min a 5000 x g) en una centrífuga de cubetas basculantes y los  
 sobrenadantes se recogieron y se filtraron sobre un filtro de 0,2 µm (Nalgene)

25 Los sobrenadantes se analizaron con respecto a la expresión de Temer00088, Temer09484, Temer08028,  
 Temer02362, Temer08862, Temer04790, Temer05249, Temer06848, Temer02056, Temer03124, Temer09491,  
 Temer06400, Temer08570, Temer08163 y Temer07305 mediante SDS-PAGE y medidas de proteínas totales.

#### **Ejemplo 2: Construcción de vectores de expresión de *R. emersonii*.**

Este Ejemplo describe la construcción de un constructo de expresión para la sobreexpresión de Temer00088,  
 Temer09484, Temer08028, Temer02362, Temer08862, Temer04790, Temer05249, Temer06848, Temer02056,  
 Temer03124, Temer09491, Temer06400, Temer08570, Temer08163 o Temer07305 en *R. emersonii*. El casete de  
 30 expresión se orientaba integrado en el locus *RePepA*.

A fin de orientar los constructos de promotor-indicador en el locus *pepA*, se clonaron vectores de expresión para la  
 orientación. El gen con proteína traducida anotado como proteasa *pepA* se identificó en el genoma. Secuencias de  
 35 *pepA* de *Rasamsonia emersonii* (*RePepA*), que comprenden la secuencia genómica del ORF y aproximadamente  
 3000 pb de la región 5' y 2500 pb de las regiones de flanqueo 3', secuencia de ADNc y proteínica secuencia, se  
 muestran en el listado de secuencias 76, 77 y 78, respectivamente.

Se construyeron dos vectores según procedimientos de clonación habituales para orientar al locus *RePepA*. Los  
 40 fragmentos insertados de ambos vectores juntos se pueden aplicar en el llamado método de "orientación génica  
 bipartito" (Nielsen y cols., 2006, 43: 54-64). Este método está usando dos fragmentos de ADN no funcionales de un  
 marcador de selección que están solapados (véase además el documento WO2008113847 para detalles adicionales  
 del método bipartito) junto con secuencias de orientación génica. Tras la recombinación homóloga correcta, el  
 marcador de selección se vuelve funcional mediante la integración en un locus diana homólogo. Como también se  
 45 detalla en el documento WO 2008113847, dos vectores de eliminación diferentes, Te *pep*.bbn y pEBA1006, se  
 diseñaron y se construyeron para que fueran capaces de proporcionar las dos moléculas de ADN solapadas para la  
 orientación génica bipartita. El primer vector Te *pep*.bbn (disposición general como en la Fig. 2) comprende una región  
 de flanqueo 5' de 1500 pb aproximadamente 1,5 kb aguas arriba del ORF de *RePepA* en el locus *RePepA* (ORF y  
 aproximadamente 1500 pb del promotor de *RePepA*), un sitio *lox66* y la parte 5' no funcional de la región codificante  
 50 ble conducida por el promotor *gpdA* de *A.nidulans* (secuencia *PgpdA*-ble que carece de las últimas 104 bases de la  
 secuencia codificante en el extremo 3' de ble, SEQ ID NO: 79). Para permitir la clonación eficaz de casetes de  
 promotor-indicador en *E.coli*, se insertó un gen *ccdB* entre la región de flanqueo *RePepA* 5' y el sitio *lox66*. El segundo  
 vector pEBA1006 (disposición general como en la Fig. 3) comprende la parte 3' no funcional de la región codificante  
 55 ble y el terminador *trpC* de *A.nidulans* (secuencia ble-TtrpC que carece de las primeras 12 bases de la secuencia  
 codificante en el extremo 5' de ble, SEQ ID NO: 80), un sitio *lox71* y una región de flanqueo 3' de 2500 pb del OFR de  
*RePepA* para la orientación en el locus *RePepA*. Tras la recombinación homóloga, los fragmentos no funcionales  
 primero y segundo se hacen funcionales produciendo un casete de ble funcional. Las regiones de flanqueo génicas  
 tanto aguas arriba como aguas abajo de *RePepA* orientan para la recombinación homóloga de los fragmentos  
 bipartitos en el locus genómico de *RePepA* predestinado.

El gen *ccdB* en el vector Te pep.bbn se reemplaza por casetes de expresión de Temer00088, Temer09484, Temer08028, Temer02362, Temer08862, Temer04790, Temer05249, Temer06848, Temer02056, Temer03124, Temer09491, Temer06400, Temer08570, Temer08163 o Temer07305 según procedimientos de clonación habituales. El promotor 2 de *R. emersonii*, representado por SEQ ID NO: 81, se clona aguas arriba de la región codificante de Temer00088, Temer09484, Temer08028, Temer02362, Temer08862, Temer04790, Temer05249, Temer06848, Temer02056, Temer03124, Temer09491, Temer06400, Temer08570, Temer08163 o Temer07305 de *R. emersonii* con el terminador *amdS* de *A. nidulans*, generando el constructo pEBA. La secuencia terminadora *amdS* de *A. nidulans* está representada por SEQ ID NO: 82. Una representación esquemática de pEBA para la sobreexpresión del gen de interés (GOI) que es Temer00088, Temer09484, Temer08028, Temer02362, Temer08862, Temer04790, Temer05249, Temer06848, Temer02056, Temer03124, Temer09491, Temer06400, Temer08570, Temer08163 o Temer07305 se muestra en la Figura 4.

**Ejemplo 3: Sobreexpresión del gen Temer00088, Temer09484, Temer08028, Temer02362, Temer08862, Temer04790, Temer05249, Temer06848, Temer02056, Temer03124, Temer09491, Temer06400, Temer08570, Temer08163 o Temer07305 en *Rasamsonia emersonii***

ADN lineales de pEBA y pEBA1006 se aíslan y se usan para transformar *Rasamsonia emersonii* usando un método como el descrito previamente en el documento WO2011/054899. Los ADN lineales se pueden integrar conjuntamente en el genoma en el locus *RePepA*, sustituyendo así el gen *RePepA* por el gen Temer00088, Temer09484, Temer08028, Temer02362, Temer08862, Temer04790, Temer05249, Temer06848, Temer02056, Temer03124, Temer09491, Temer06400, Temer08570, Temer08163 o Temer07305 y *ble*. Los transformantes se seleccionan en medio de fleomicina y las colonias se purifican y se prueban según procedimientos como los descritos en el documento WO2011/054899. Las colonias crecientes se diagnostican mediante PCR con respecto a la integración en el locus *RePepA* usando un cebador en el promotor *gpdA* del casete de eliminación y un cebador dirigido contra la secuencia genómica directamente aguas arriba de la región de orientación 5'. Se obtienen posibles transformantes en los que *RePepA* se reemplaza por casetes de Temer00088, Temer09484, Temer08028, Temer02362, Temer08862, Temer04790, Temer05249, Temer06848, Temer02056, Temer03124, Temer09491, Temer06400, Temer08570, Temer08163 o Temer07305/*ble*.

**Ejemplo 4: actividad enzimática en cepas de *Rasamsonia emersonii* que sobreexpresan Temer00088, Temer09484, Temer08028, Temer02362, Temer08862, Temer04790, Temer05249, Temer06848, Temer02056, Temer03124, Temer09491, Temer06400, Temer08570, Temer08163 o Temer07305**

Cepas que sobreexpresan Temer00088, Temer09484, Temer08028, Temer02362, Temer08862, Temer04790, Temer05249, Temer06848, Temer02056, Temer03124, Temer09491, Temer06400, Temer08570, Temer08163 o Temer07305 se fermentan en un matraz agitador en medio para *Rasamsonia* 3 y los sobrenadantes se analizan con respecto a la actividad según la Tabla 1 en un ensayo adecuado. Se observa un incremento en la actividad en sobrenadantes de cepas que sobreexpresan Temer00088, Temer09484, Temer08028, Temer02362, Temer08862, Temer04790, Temer05249, Temer06848, Temer02056, Temer03124, Temer09491, Temer06400, Temer08570, Temer08163 o Temer07305 en comparación con la cepa silvestre, indicando que la sobreexpresión de Temer00088, Temer09484, Temer08028, Temer02362, Temer08862, Temer04790, Temer05249, Temer06848, Temer02056, Temer03124, Temer09491, Temer06400, Temer08570, Temer08163 o Temer07305 mejora la actividad en *R. emersonii*.

**Ejemplo 5: Fermentación en matraz agitador de *Aspergillus niger***

Aproximadamente 10<sup>7</sup> esporas de transformantes seleccionados y cepas de control se inocularon en matraces agitadores de 100 ml con pantallas que contenían 20 ml de medio de cultivo líquido que consistía en, por litro: 30 g de maltosa.H<sub>2</sub>O; 5 g de extracto de levadura; 10 g de caseína hidrolizada; 1 g de KH<sub>2</sub>PO<sub>4</sub>; 0,5 g de MgSO<sub>4</sub>.7H<sub>2</sub>O; 0,03 g de ZnCl<sub>2</sub>; 0,02 g de CaCl<sub>2</sub>; 0,01 g de MnSO<sub>4</sub>.4H<sub>2</sub>O; 0,3 g de FeSO<sub>4</sub>.7H<sub>2</sub>O; 3 g de Tween 80; 10 ml de penicilina (5000 IU/ml)/estreptomicina (5000 UG/ml); pH 5,5. Estos cultivos se desarrollaron a 34 grados Celsius durante 16-24 horas. Se inocularon 10 ml de este cultivo en matraces agitadores de 500 ml con pantallas que contenían 100 ml de medio de fermentación que consistía en, por litro: 70 g de glucosa.H<sub>2</sub>O; 25 g de caseína hidrolizada; 12,5 g de extracto de levadura; 1 g de KH<sub>2</sub>PO<sub>4</sub>; 2 g de K<sub>2</sub>SO<sub>4</sub>; 0,5 g de MgSO<sub>4</sub>.7H<sub>2</sub>O; 0,03 g de ZnCl<sub>2</sub>; 0,02 g de CaCl<sub>2</sub>; 0,01 g de MnSO<sub>4</sub>.4H<sub>2</sub>O; 0,3 g de FeSO<sub>4</sub>.7H<sub>2</sub>O; 10 ml de penicilina (5000 IU/ml)/estreptomicina (5000 UG/ml); ajustados hasta pH 5,6. Estos cultivos se desarrollaron a 34 grados Celsius hasta que se agotaba toda la glucosa (habitualmente después de 4-7 días). Muestras tomadas del caldo de fermentación se centrifugaron (10 min at 5000 x g) en una centrífuga de cubetas basculantes y los sobrenadantes se recogieron y se filtraron sobre un filtro de 0,2 µm (Nalgene)

**Determinación de la concentración en matraces agitadores y la concentración de proteína con el método de TCA-biuret**

5 A fin de obtener mayores cantidades de material para una prueba adicional, los sobrenadantes de fermentación obtenidos como se describe anteriormente (volumen entre 75 y 100 ml) se concentraron usando un filtro giratorio de 10 kDa hasta un volumen de aproximadamente 5 ml. Posteriormente, se determinó la concentración de proteína en el sobrenadante concentrado a través de un método de TCA-biuret.

10 Muestras de proteína concentradas (sobrenadantes) se diluyeron con agua hasta una concentración entre 2 y 8 mg/ml. Se realizaron diluciones de albúmina sérica bovina (BSA) (0, 1, 2, 5, 8 y 10 mg/ml) y se incluían como muestras para generar una curva de calibración. De cada muestra de proteína diluida, 270 µl se transfirieron a un tubo de 10 ml que contenía 830 µl de una solución de ácido tricloroacético al 12% (p/v) en acetona y se mezclaron a fondo. Posteriormente, los tubos se incubaron sobre agua de hielo durante una hora y se centrifugaron durante 30 minutos, a 4°C y 6000 rpm. El sobrenadante se descartó y las pellas se secaron al invertir los tubos sobre un tejido y dejarlas reposar durante 30 minutos a temperatura ambiente. Posteriormente, se añadieron a la pella del tubo 3 ml de mezcla de reactivos de biuret BioQuant y la pella se solubilizó tras la mezcladura seguido por la adición de 1 ml de agua. El tubo se mezcló a fondo y se incubó a temperatura ambiente durante 30 minutos. La absorción de la mezcla se midió a 546 nm con una muestra de agua usada como una medida del blanco y la concentración de proteína se calculó a través de la línea de calibración de BSA.

**Ejemplo 6: Identificación de la actividad como beta-xilosidasa de Temer09484 de *Rasamsonia emersonii* termófila sobre xilobiosa**

15 La actividad como beta-xilosidasa de Temer09484 de *Rasamsonia emersonii* se analizó según se describe anteriormente. El sobrenadante de fermentación en matraz agitador de Temer09484 *A. niger* se concentró y se ensayó en dos dosificaciones con respecto a la liberación de xilosa desde xilobiosa después de la incubación durante 24 horas a pH 4,5 y 62°C. La enzima mostraba una liberación de xilosa significativa desde xilobiosa según se muestra en la Tabla 3. Esta muestra que Temer09484 tiene actividad como beta-xilosidasa.

20 Tabla 3: Efecto de Temer09484 de *Rasamsonia emersonii* sobre la liberación de xilosa desde xilobiosa (100 µg/ml) después de 24 h de incubación a pH 4,5 y 62°C.

ID Proteína	Dosificación (mg/g de DM)	Xilosa obtenida como producto (µg/ml)
Sin enzima	0	0
Temer09484	1	100
Temer09484	5	100

**Ejemplo 7: Identificación de la actividad como beta-xilosidasa de Temer09484 de *Rasamsonia emersonii* termófila sobre sustratos de xilano poliméricos**

30 Como un segundo experimento, también se analizó la actividad como beta-xilosidasa de Temer09484 de *Rasamsonia emersonii* sobre sustratos de xilano poliméricos. El sobrenadante del matraz de Temer09484 *A. niger* se dosificó en 10 mg/g a tres sustratos de xilano poliméricos diferentes. Se liberaba xilosa desde los tres sustratos (Tabla 4) mientras que no se formaban xilooligómeros. Esta muestra que Temer09484 también tiene actividad como beta-xilosidasa sobre sustratos poliméricos junto con oligómeros pequeños según se muestra en el Ejemplo 6.

35 Tabla 4: Efecto de Temer09484 de *Rasamsonia emersonii* sobre la liberación de xilosa desde varios sustratos de xilano después de la incubación durante 24 h a pH 4,5 y 60°C con una dosificación 10 mg/g de DM.

Sustrato (2 mg/ml)	µg/ml*
Xilano de madera de haya	320
Xilano de madera de abedul	250
Arabinoxilano de avena	334

40 \* Todos los sustratos contienen < 3,0 µg/ml de xilosa cuando no se añadía enzima

**Ejemplo 8. Mejora de dos mezclas de celulosa diferentes mediante la adición de Temer09484 para la hidrólisis de materias primas lignocelulósicas**

5 El sobrenadante de la fermentación en matraz agitador de Temer09484 *A. niger* se concentró y se agregó una materia prima de rastrojo de maíz pretratada con ácido suave según se describe anteriormente. La enzima mostraba una liberación de xilosa significativa desde esta materia prima en un amplio intervalo de temperaturas (50, 65 y 75°C) y valores de pH (3,5 – 4,5 – 5,0) usados durante las 72 horas de incubación según se muestra en la Tabla 5. Esta muestra que Temer09484 es importante para la hidrólisis de materias primas lignocelulósicas.

10 Tabla 5: Efecto de Temer 09484 de *Rasamsonia emersonii* sobre la liberación de xilosa (g/l) desde materia prima de rastrojo de maíz pretratada suavemente con ácido después de 72 h de incubación en diferentes condiciones de temperatura/pH.

ID Proteína	pH 5,0-50°C	pH 3,5-65°C	pH 4,5-65°C	pH 4,5-75°C
Materia prima sola - sin enzima	0,14	0,14	0,14	0,13
Temer09484	0,29	0,22	0,23	0,19

15 El sobrenadante de la fermentación en matraz agitador de Temer09484 *A. niger* también se probó en combinación con 2 mezclas de celulosas diferentes: TEC-210 y Celluclast, ambas con BG adicional añadida. La liberación de xilosa desde rastrojo de maíz pretratado suavemente con ácido se mejoraba para ambas mezclas de celulosas mediante la adición de Temer09484 en un amplio intervalo de temperaturas (50, 65 y 75°C) y valores de pH (3,5 – 4,5 – 5,0) usado durante las 72 horas de incubación según se muestra en la Tabla 6. Esta muestra que Temer09484 se puede usar para mejorar mezclas de celulosas en un amplio intervalo de temperaturas y valores de pH usados para la hidrólisis de materias primas lignocelulósicas.

20 Tabla 6: Efecto de Temer09484 de *Rasamsonia emersonii* cuando se han agregado dos mezclas de celulosas diferentes sobre la liberación de xilosa (g/l) desde materia prima de rastrojo de maíz pretratada suavemente con ácido después de 72 h de incubación a diferentes condiciones de temperatura/pH.

ID Proteína	pH 5,0-50°C	pH 3,5-65°C	pH 4,5-65°C	pH 4,5-75°C
Materia prima sola - sin enzima	0,14	0,14	0,14	0,13
TEC-210 +8% de BG	0,51	0,45	0,57	0,22
TEC-210 + 8% de BG + Temer09484	0,63	0,53	0,63	0,33
Celluclast + 8% de BG	0,49	0,23	0,27	0,19
Celluclast + 8% de BG + Temer09484	0,61	0,24	0,32	0,24

**Ejemplo 9: Identificación de la actividad como beta-xilosidasa de Temer00088 de *Rasamsonia emersonii* termófila sobre xilobiosa**

30 La actividad como beta-xilosidasa de Temer00088 de *Rasamsonia emersonii* se analizó según se describe anteriormente. El sobrenadante de la fermentación en matraz agitador de Temer00088 *A. niger* se concentró y se ensayó en dos dosificaciones para la liberación de xilosa desde xilobiosa después de la incubación durante 24 horas a pH 4,5 y 62°C. La enzima mostraba una liberación de xilosa significativa desde xilobiosa según se muestra en la Tabla 7. Esta muestra que Temer00088 tiene actividad como beta-xilosidasa.

35 Tabla 7: Efecto de Temer00088 de *Rasamsonia emersonii* sobre la liberación de xilosa desde xilobiosa (100 ug/ml) después de 24 h de incubación a pH 4,5 y 62°C.

ID Proteína	Dosificación (mg/g de DM)	Xilosa obtenida como producto (ug/ml)
Sin enzima	0	0
Temer00088	1	76
Temer00088	5	99

**Ejemplo 10: Identificación de actividad como beta-xilosidasa de Temer00088 de *Rasamsonia emersonii* termófila sobre sustratos de xilano poliméricos**

Como un segundo experimento, también se analizó la actividad como beta-xilosidasa de Temer00088 de *Rasamsonia emersonii* sobre sustratos de xilano poliméricos. El sobrenadante del matraz agitador de Temer00088 *A. niger* se dosificó en 10 mg/g a los tres sustratos de xilano poliméricos diferentes. Se liberaba xilosa de los tres sustratos (Tabla 8) mientras que no se formaban xilooligómeros. Esta muestra que Temer00088 también tiene actividad como beta-xilosidasa sobre sustratos poliméricos junto con oligómeros pequeños según se muestra en el Ejemplo 9.

Tabla 8: Efecto de Temer 00088 de *Rasamsonia emersonii* sobre la liberación de xilosa desde varios sustratos de xilano después de la incubación durante 20 h a pH 4,5 y 60°C con una dosificación de 10 mg/g de DM.

	ug/ml*
Sustrato (2 mg/m)	xilosa
Xilano de madera de haya	373
Xilano de madera de abedul	469
Arabinoxilano de avena	298

\* Todos los sustratos contienen < 3,0 ug/ml de xilosa cuando no se añadía enzima

**Ejemplo 11. Mejora de dos mezclas de celulosa diferentes mediante la adición de Temer00088 para la hidrólisis de materias primas lignocelulósicas**

El sobrenadante de la fermentación en matraz agitador de Temer00088 *A. niger* se concentró y se agregó una materia prima de rastrojo de ácido pretratada con ácido suave según se describe anteriormente. La enzima mostraba una liberación de xilosa significativa desde esta materia prima en un amplio intervalo de temperaturas (50, 65 y 75°C) y valores de pH (3,5 – 4,5 – 5,0) usado durante las 72 horas de incubación según se muestra en la Tabla 3. Esta muestra que Temer00088 es importante para la hidrólisis de materias primas lignocelulósicas.

Tabla 9: Efecto de Temer00088 de *Rasamsonia emersonii* sobre la liberación de xilosa (g/l) desde materia prima de rastrojo de maíz pretratada suavemente con ácido después de 72 h de incubación a diferentes condiciones de temperatura/pH.

ID Proteína	pH 5,0-50°C	pH 3,5-65°C	pH 4,5-65°C	pH 4,5-75°C
Materia prima sola - sin enzima	0,14	0,14	0,14	0,13
Temer00088	0,29	0,26	0,27	0,22

El sobrenadante de la fermentación en matraz agitador de Temer00088 *A. niger* también se probó en combinación con 2 mezclas de celulosas diferentes: TEC-210 y Celluclast, ambas con BG adicional añadida. La liberación de xilosa desde rastrojo de maíz pretratado suavemente con ácido se mejoraba para ambas mezclas de celulosas mediante la adición de Temer00088 en un amplio intervalo de temperaturas (50, 65 y 75°C) y valores de pH (3,5 – 4,5 – 5,0) usados durante las 72 horas de incubación según se muestra en la Tabla 10. Esta muestra que Temer00088 se puede usar para mejorar las mezclas de celulosas en un amplio intervalo de temperaturas y valores de pH usados para la hidrólisis de materias primas lignocelulósicas.

Tabla 10 Efecto de Temer00088 de *Rasamsonia emersonii* cuando se agrega a dos mezclas de celulosas diferentes durante la liberación de xilosa (g/l) desde materia prima de rastrojo de maíz pretratada suavemente con ácido después de 72 h de incubación a diferentes condiciones de temperatura/pH.

ID Proteína	pH 5,0-50°C	pH 3,5-65°C	pH 4,5-65°C	pH 4,5-75°C
Materia prima sola - sin enzima	0,14	0,14	0,14	0,13
TEC-210 + 8% de BG	0,51	0,45	0,57	0,22
TEC-210 + 8% de BG + Temer09484	0,62	0,59	0,67	0,39
Celluclast + 8% de BG	0,49	0,23	0,27	0,19
Celluclast + 8% de BG + Temer09484	0,66	0,29	0,35	0,27

**Ejemplo 12: Identificación de endoglucanasa específica de xiloglucano de *Rasamsonia emersonii* termófila**

La actividad como xiloglucanasa de Temer04790 de *Rasamsonia emersonii* se analizó según se describe anteriormente. El sobrenadante de la fermentación en matraz agitador de Temer04790 *A. niger* se concentró, se añadió al sustrato xiloglucano y se incubó durante 24 horas a pH 4,5 y 60°C. La enzima era capaz de liberar varios oligómeros según se muestra en la Figura 6. Esta muestra que Temer04790 es activo sobre xiloglucano y libera oligómeros similares como la mezcla de celulasas comercial Celluclast de *Trichoderma reesei*.

Para cuantificar la cantidad de oligómeros, se midieron los extremos reductores formados después de la incubación tanto de Temer04790 como de la mezcla de celulasas. También se usó carboximetilcelulosa como sustrato para determinar la especificidad de las enzimas y una temperatura superior, 75°C, después de 60°C. Temer04790 es específico hacia xiloglucano ya que apenas se observó actividad sobre CMC en contraste con la mezcla de celulasas (Tabla 11). Por otra parte, Temer04790 todavía era activo a 75°C mientras que la mezcla de celulasas era casi inactiva sobre xiloglucano a 75°C.

Tabla 11: Efecto de Temer04790 de *Rasamsonia emersonii* sobre la hidrólisis de xiloglucano (tamarindo) y carboximetilcelulosa (CMC) (Sigma) medido mediante la formación de extremos reductores expresados como equivalentes de glucosa (ug/ml) después de una incubación de 24 h a pH 4,5 a 60°C y 75°C .

	60°C pH 4,5		75°C pH 4,5	
	xiloglucano	CMC	xiloglucano	CMC
sin enzima	-17	-18	-18	-18
Temer04790	92	8	69	-9
mezcla de celulasas*	64	132	5	47

\* Celluclast procedente de *Trichoderma reesei* (Sigma)

**Ejemplo 13: Identificación de actividad como arabinofuranosidasa de *Rasamsonia emersonii* termófila**

La actividad como arabinofuranosidasa de Temer05249 de *Rasamsonia emersonii* se analizó según se describe anteriormente. El sobrenadante de la fermentación en matraz agitador de Temer05249 *A. niger* se concentró y se añadió a arabinoxilooligómeros en 10 mg/g seguido por incubación durante 24 horas a pH 4,5 y 65°C. La enzima mostraba una liberación de arabinosa significativa desde arabinoxilooligómeros según se muestra en la Tabla 12. Esta muestra que Temer05249 tiene actividad como arabinofuranosidasa.

Tabla 12: Efecto de Temer05249 de *Rasamsonia emersonii* sobre la liberación de arabinosa desde arabinoxilano de trigo, que se preincubaba con una endoxilanasas, después de la incubación durante 24 h a pH 4,5 y 65°C en una dosificación de 10 mg/g de DM.

ID Proteína	Arabinosa (ug/ml)
Sin enzima	4
Temer05249	111

**Ejemplo 14: Identificación de actividad como endoxilanasas de *Rasamsonia emersonii* termófila**

La actividad como endoxilanasas de Temer03124 de *Rasamsonia emersonii* se analizó como se describe anteriormente. El sobrenadante de la fermentación en matraz agitador de Temer03124 *A. niger* se concentró y se añadió a varios sustratos de xilano en 10 mg/g seguido por incubación durante 20 horas a pH 4,5 y 60°C. La enzima mostraba una liberación significativa de xilosa y una gama de xilooligómeros según se muestra en la Tabla 13. Esta muestra que Temer03124 tiene actividad como endoxilanasas.

Tabla 13: Efecto de Temer03124 de *Rasamsonia emersonii* sobre la liberación de xilosa y oligómeros de xilosa desde varios sustratos de xilano después de la incubación durante 20 h a pH 4,5 y 60°C en una dosificación de 10 mg/g de DM.

Sustrato (2 mg/ml)	ug/ml*			
	xilosa	xilobiosa	xilotriosa	xilotetraosa
Xilano de madera de haya	34.3	12.6	10.4	11.4
Xilano de madera de abedul	30.6	17.8	16.1	16.1
Arabinoxilano de avena	27.3	17.2	12.5	10.9
Arabinoxilano de trigo	33.4	36.9	15.0	5.2

\* Todos los sustratos contienen < 3,5 ug/ml de cada producto medido si no se añade enzima

#### Ejemplo 15: Identificación de actividad como endoxilanasas de *Rasamsonia emersonii* termófila

5 La actividad como endoxilanasas de Temer08570 de *Rasamsonia emersonii* se analizó según se describe anteriormente. El sobrenadante de la fermentación en matraz agitador de Temer08570 *A. niger* se concentró y se añadió a varios sustratos de xilano en 10 mg/g seguido por incubación durante 20 horas a pH 4,5 y 60°C. La enzima mostraba una liberación significativa de xilosa y una gama de xilooligómeros según se muestra en la Tabla 14. Esta muestra que Temer08570 tiene endo-xilanasas con xilobiosa, xilotriosa y xilotetraosa como productos.

10 Tabla 14: Efecto de Temer08570 de *Rasamsonia emersonii* sobre la liberación de xilosa y oligómeros de xilosa desde varios sustratos de xilano después de la incubación durante 20 h a pH 4,5 y 60°C en una dosificación de 10 mg/g de DM.

Sustrato (2 mg/ml)	ug/ml*			
	xilosa	xilobiosa	xilotriosa	xilotetraosa
Xilano de madera de haya	5	19	25	25
Xilano de madera de abedul	4	14	17	17
Arabinoxilano de avena	0	16	18	15

15 \* Todos los sustratos contienen < 3,5 ug/ml de cada producto medido si no se añade enzima

#### Ejemplo 16: Identificación de actividad como endoxilanasas de *Rasamsonia emersonii* termófila

20 La actividad como endoxilanasas de Temer08163 de *Rasamsonia emersonii* se analizó según se describe anteriormente. El sobrenadante de la fermentación en matraz agitador de Temer08163 *A. niger* se concentró y se añadió a varios sustratos de xilano en 10 mg/g seguido por incubación durante 20 horas a pH 4,5 y 60°C. La enzima mostraba una liberación significativa de xilobiosa y xilosa según se muestra en la Tabla 15. Esta muestra que Temer08570 tiene actividad como endoxilanasas con xilobiosa como producto principal que era 12 - 25 veces superior que la cantidad de xilosa liberada.

25 Tabla 15: Efecto de Temer08163 de *Rasamsonia emersonii* sobre la liberación de xilosa y oligómeros de xilosa desde varios sustratos de xilano después de la incubación durante 20 h a pH 4,5 y 60°C en una dosificación de 10 mg/g de DM.

Sustrato (2 mg/ml)	ug/mL*			
	xilosa	xilobiosa	xilotriosa	xilotetraosa
Xilano de madera de haya	22,4	581,4	0	0
Xilano de madera de abedul	30,9	527,3	0	0
Arabinoxilano de avena	13,6	205,1	0	0
Arabinoxilano de trigo	5,4	65,9	0	0

30 \* Todos los sustratos contienen < 3,5 ug/ml de cada producto medido.

**Ejemplo 17: Identificación de actividad como alfa-glucuronidasa de *Rasamsonia emersonii* termófila**

La actividad como alfa-glucuronidasa de Temer07305 de *Rasamsonia emersonii* se analizó según se describe anteriormente. El sobrenadante de la fermentación en matraz agitador de Temer07305 *A. niger* se concentró y se añadió a ácidos aldourónicos tanto 1 como 10 mg/g seguido por incubación durante 24 horas a pH 4,5 y 60°C. La enzima era capaz de retirar ácido 4-O-metilglucurónico de los xilooligómeros dando como resultado la liberación simultánea de xilosa, xilobiosa, xilotriosa y xilotetraosa según se muestra en la Tabla 16. Esta muestra que Temer07305 tiene actividad como alfa-glucuronidasa.

Tabla 16: La liberación de xilosa y oligómeros de xilosa por Temer07305 de *Rasamsonia emersonii* desde ácidos aldourónicos como resultado de la hidrólisis de ácido 4-O-metilglucurónico desde estos xilooligómeros, después de la incubación durante 24 h a pH 4,5 y 60°C en una dosificación de 1 y 10 mg/g de DM.

ID Proteína	Dosificación (mg/g de DM)	Superficie / (mg/ml) sustrato			
		xilosa	xilobiosa	xilotriosa	xilotetraosa
Sin enzima	x	25	6	0	0
Temer07305	1	45	180	58	19
Temer07305	10	120	180	55	9

## Listado de secuencias

15 <110> DSM IP Assets B.V.  
 <120> POLIPÉPTIDO DEGRADADOR DE CARBOHIDRATOS Y SUS USOS  
 <130> 29207-EP-ETD  
 20 <160> 82  
 <170> PatentIn versión 3.5  
 25 <210> 1  
 <211> 2286  
 <212> ADN  
 30 <213> *Rasamsonia emersonii*  
 <220>  
 35 <221> fuente  
 <222> (1)..(2286)  
 <223> /organismo="*Rasamsonia emersonii*" /tipo\_molar="ADN no asignado"  
 40 <400> 1

# ES 2 802 807 T3

atggtccttg gtgtctccct ggtccttctg gccactgctg tctccgccac cttccccgac	60
tgctctcagc ctccctctcaa ggacaacgcc gtctgcgaca ccagccttga ccctgtctcc	120
cgcgctgctg ccctcgttgc tgctttcacc ctggaggaga agatcaaaa cactcagaac	180
ggctcccccg gtgttccccg tctgggtctg cctccctacc agtgggtgtc tgagggtctg	240
cacggtgttg ccatctcccc cggcgtcaac ttctctgctc acggtgactt ctccctacgcc	300
accagcttcc ctcagcccat cctgatgtcc gctgccttcg atgatgacct catccgccag	360
gtcggctccg tcgtcagcac tgaggctcgt gctttctcca acgccaaccg cgctggtctg	420
gactactgga ctcccaacat caacccttc aaggacctc gctggggtcg tggccaggag	480
actcccgggtg aagatgcctt ccacatccag cgctacgtgt actctttgat tgatggcctc	540
cagaacggca ttggacctgc caacccaag atcatggcca cctgcaagca cttcgtgcc	600
tacgacctcg aggactggca cggcaacgag cgctacggtt tcaacgccgt tgtgagcact	660
caggacctgg cgaataacta ccttctccc ttcaagtctt gcgctcgtga tgccaaggtc	720
gatgccctca tgtgctcgta caacgctgtc aacggtgtgc cctcttgccg ggactcctac	780
ctcttgaag atctcctccg tgaccactgg ggctggaaca ccaccggtca ctgggtgacc	840
tccgactgcg atgccgtcca gaacatctac gccaacacc actacacttc caccgctccc	900
caggctgctg ccgatgcctt gggtgctggt actgatcttg actgcggtac cacctacccc	960
gacaaccttg gtgccgccta caccagggt ctgttccaga accagacctt ggacaccgct	1020
ctgatccgtc tgtactcttc tctggtaag ctggatact tcgaccctcc cgagaaccag	1080
ccctaccgca gcattggctg gtcagatgtc agcactcctg ctgctcagca gcttgccccg	1140
actgctgctg ctgagggcat tgtcctctc aagaacgatg agaagaaggt cctccccctc	1200

ES 2 802 807 T3

agccgtgagg gtcagaccct cgccgtcatc ggtcccttcg ccaacgccac caccagctg 1260  
cagggtaaact accaggggtg tgctccttac atctggactg ttggtgctgc tgcggagcag 1320  
ctgggtaca aggtcaacta cgccgacggt actgccatca acgctaccaa caccactggc 1380  
ttcggcgagg ccgttgctgc cgccaagtcc tccgatgtg tcatctacgc tgggtgtatc 1440  
gacaacagca ttgaggctga aggccacgac cgtgacacca ttgtgtggcc cggcaaccag 1500  
ctccagctca tctccgagct gggcgagacc ggcaagcctc tggttgtcat ccagttcggg 1560  
gggtgtcagg ttgatgactc ctccctcctt gccaacgact ccggtgtcaa cgccctcctc 1620  
tgggctggct acccctccca ggctgggtgt gctgccatct tcgagatcct gaccggtgcc 1680  
actgcccccg ccggtcgtct gcctaccacc cagtaccctg ctcagtacgt ggatgaggtt 1740  
cccatgaccg acatgaccct ccgcccctct gccaccaacc ccggtcgcac ctacagatgg 1800  
tacgacaagg ccgtcatccc ctctcggttc ggcttgact acaccacctt cgatgtcacc 1860  
tggagcaagg cccagcttgg accttacgaa attggctctc ttgtcaagaa cgcgtcgtcg 1920  
tcggatgctc ctgccgacac cgcccccttc gacaccttca ccatccacgt ccgcaacact 1980  
ggcaagacca cctcggacta cgtggctctc ctcttctct ccaccgcaa cgccggtcct 2040  
gctccctacc ccctcaagac cctagttgga tacactcgtg ccggtgccat ccagcccggg 2100  
gagactcgtg ccgtogacat tgccgtcacc gttggcagcg ttgcccgta cgcagagcgt 2160  
ggtgacctcg tcctctacc cggcacctac accctcgagg ttgacgtcaa cggccagtac 2220  
cccactgctg gcttogaagt caccgggtgag gctgctgtcc tggatgagtt ccccagccc 2280  
ccggcg 2286

<210> 2

5 <211> 746

<212> PRT

<213> *Rasamsonia emersonii*

10

<400> 2

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

ES 2 802 807 T3

His Gly Val Ala Ile Ser Pro Gly Val Asn Phe Ser Ala His Gly Asp  
65 70 75 80

Phe Ser Tyr Ala Thr Ser Phe Pro Gln Pro Ile Leu Met Ser Ala Ala  
85 90 95

Phe Asp Asp Asp Leu Ile Arg Gln Val Gly Ser Val Val Ser Thr Glu  
100 105 110

Ala Arg Ala Phe Ser Asn Ala Asn Arg Ala Gly Leu Asp Tyr Trp Thr  
115 120 125

Pro Asn Ile Asn Pro Phe Lys Asp Pro Arg Trp Gly Arg Gly Gln Glu  
130 135 140

Thr Pro Gly Glu Asp Ala Phe His Ile Gln Arg Tyr Val Tyr Ser Leu  
145 150 155 160

Ile Asp Gly Leu Gln Asn Gly Ile Gly Pro Ala Asn Pro Lys Ile Met  
165 170 175

Ala Thr Cys Lys His Phe Ala Ala Tyr Asp Leu Glu Asp Trp His Gly  
180 185 190

Asn Glu Arg Tyr Gly Phe Asn Ala Val Val Ser Thr Gln Asp Leu Ala  
195 200 205

Glu Tyr Tyr Leu Pro Pro Phe Lys Ser Cys Ala Arg Asp Ala Lys Val  
210 215 220

Asp Ala Leu Met Cys Ser Tyr Asn Ala Val Asn Gly Val Pro Ser Cys  
225 230 235 240

Ala Asp Ser Tyr Leu Leu Glu Asp Ile Leu Arg Asp His Trp Gly Trp  
245 250 255

Asn Thr Thr Gly His Trp Val Thr Ser Asp Cys Asp Ala Val Gln Asn  
260 265 270

Ile Tyr Ala Asn His His Tyr Thr Ser Thr Ala Pro Gln Ala Ala Ala  
275 280 285

Asp Ala Leu Gly Ala Gly Thr Asp Leu Asp Cys Gly Thr Thr Tyr Pro  
290 295 300

Asp Asn Leu Gly Ala Ala Tyr Thr Gln Gly Leu Phe Gln Asn Gln Thr  
305 310 315 320

ES 2 802 807 T3

Leu Asp Thr Ala Leu Ile Arg Leu Tyr Ser Ser Leu Val Lys Leu Gly  
 325 330 335  
 Tyr Phe Asp Pro Pro Glu Asn Gln Pro Tyr Arg Ser Ile Gly Trp Ser  
 340 345 350  
 Asp Val Ser Thr Pro Ala Ala Gln Gln Leu Ala Arg Thr Ala Ala Ala  
 355 360 365  
 Glu Gly Ile Val Leu Leu Lys Asn Asp Glu Lys Lys Val Leu Pro Leu  
 370 375 380  
 Ser Arg Glu Gly Gln Thr Leu Ala Val Ile Gly Pro Phe Ala Asn Ala  
 385 390 395 400  
 Thr Thr Gln Leu Gln Gly Asn Tyr Gln Gly Val Ala Pro Tyr Ile Trp  
 405 410 415  
 Thr Val Val Ala Ala Ala Glu Gln Leu Gly Tyr Lys Val Asn Tyr Ala  
 420 425 430  
 Asp Gly Thr Ala Ile Asn Ala Thr Asn Thr Thr Gly Phe Ala Glu Ala  
 435 440 445  
 Val Ala Ala Ala Lys Ser Ser Asp Val Val Ile Tyr Ala Gly Gly Ile  
 450 455 460  
 Asp Asn Ser Ile Glu Ala Glu Gly His Asp Arg Asp Thr Ile Val Trp  
 465 470 475 480  
 Pro Gly Asn Gln Leu Gln Leu Ile Ser Glu Leu Ala Gln Thr Gly Lys  
 485 490 495  
 Pro Leu Val Val Ile Gln Phe Gly Gly Gln Val Asp Asp Ser Ser  
 500 505 510  
 Leu Leu Ala Asn Asp Ser Gly Val Asn Ala Leu Leu Trp Ala Gly Tyr  
 515 520 525  
 Pro Ser Gln Ala Gly Gly Ala Ala Ile Phe Glu Ile Leu Thr Gly Ala  
 530 535 540  
 Thr Ala Pro Ala Gly Arg Leu Pro Thr Thr Gln Tyr Pro Ala Gln Tyr  
 545 550 555 560  
 Val Asp Glu Val Pro Met Thr Asp Met Thr Leu Arg Pro Ser Ala Thr  
 565 570 575

ES 2 802 807 T3

Asn Pro Gly Arg Thr Tyr Arg Trp Tyr Asp Lys Ala Val Ile Pro Phe  
580 585 590

Gly Phe Gly Leu His Tyr Thr Thr Phe Asp Val Thr Trp Ser Lys Ala  
595 600 605

Gln Leu Gly Pro Tyr Glu Ile Gly Ser Leu Val Lys Asn Ala Ser Ser  
610 615 620

Ser Asp Ala Pro Ala Asp Thr Ala Pro Phe Asp Thr Phe Thr Ile His  
625 630 635 640

Val Arg Asn Thr Gly Lys Thr Thr Ser Asp Tyr Val Ala Leu Leu Phe  
645 650 655

Leu Ser Thr Arg Asn Ala Gly Pro Ala Pro Tyr Pro Leu Lys Thr Leu  
660 665 670

Val Gly Tyr Thr Arg Ala Arg Ala Ile Gln Pro Gly Glu Thr Arg Ala  
675 680 685

Val Asp Ile Ala Val Thr Val Gly Ser Val Ala Arg Thr Asp Glu Arg  
690 695 700

Gly Asp Leu Val Leu Tyr Pro Gly Thr Tyr Thr Leu Glu Val Asp Val  
705 710 715 720

Asn Gly Gln Tyr Pro Thr Ala Gly Phe Glu Val Thr Gly Glu Ala Ala  
725 730 735

Val Leu Asp Glu Phe Pro Gln Pro Pro Ala  
740 745

<210> 3

5 <211> 16

<212> PRT

10 <213> Rasamsonia emersonii

<400> 3

Met Val Leu Gly Val Ser Leu Val Leu Leu Ala Thr Ala Val Ser Ala  
1 5 10 15

15 <210> 4

<211> 2289

<212> ADN

20 <213> Rasamsonia emersonii

<220>

25 <221> fuente

<222> (1)..(2289)

ES 2 802 807 T3

<223> /organismo="Rasamsonia emersonii" /tipo\_molar="ADN no asignado"

<400> 4

5

atggttctg	gtgtcagtct	cgctcctcctg	gcgacggccg	tctctgccac	gttccccgac	60
tgcagccagc	ctccgctgaa	ggacaacgcc	gtgtgcgata	cctcgctgga	ccccgtatcg	120
cgtgctgctg	cgctcgtcgc	ggccttcacg	ctcgaggaga	agatcaacaa	cacgcagaat	180
ggatcgctg	gctgccccg	gctgggcctg	ccgccgtatc	aatggtggag	cgaaggcctg	240
cacggcgtgg	ccatctcgcc	gggcgtcaac	ttctccgcgc	acggcgactt	cagctacgca	300
acctccttcc	cccagcccat	cctcatgagc	gccgcctttg	acgacgacct	catcaggcag	360
gtcggcagtg	tggtcagcac	cgaggcgcgc	gccttcagca	acgccaaccg	cgctggcctc	420
gactactgga	cgcccaacat	caacccttc	aaggaccgcg	gctggggtcg	cgggcaggag	480
acgcccggcg	aagacgcatt	ccacatccag	cgctacgtct	acagcctcat	cgacggcctg	540
caaaacggca	ttgggccggc	caaccccaag	atcatggcca	cctgcaagca	cttcgccgcc	600
tacgacctcg	aggactggca	cggcaacgag	cgctatggct	tcaacgccgt	cgtgtccacg	660
caggacctgg	ccgagtacta	cctgccgccc	ttcaagagct	gcgcccgca	cgccaaagtc	720
gacgcgctca	tgtgcagcta	caacgccgtc	aacgggggtg	cctcgtgcgc	cgactcgtat	780
ctgctggagg	acatcctgcg	cgaccactgg	ggctggaaca	cgaccggcca	ctgggtgacg	840
tccgactgcg	atgccgtgca	gaacatctac	gccaaacctc	actacacgtc	gactgcgccc	900
caggcggccg	cggacgccct	gggcgccggc	accgatctcg	actgcggcac	cacttatccc	960
gacaatctgg	gcgctgccta	cacgcagggg	ctcttcacaga	accagaccct	cgacacggcg	1020
ctcatccgcc	tgtactcgtc	gctcgtcaag	ctgggctact	ttgatccgcc	ggagaaccag	1080
ccctaccggt	cgataggctg	gagtgatgtg	tccacgccgg	ccgcgcagca	gctggcccgc	1140
acggcggcgg	cggaggggat	cgttcttctc	aagaacgacg	agaagaaagt	cctgccgctg	1200
tcgcgcgagg	gacagacgct	cgccgtgatc	gggcccttcg	ccaacgcgac	gaccagctg	1260
cagggcaatt	accaaggcgt	gggcgccgtac	atctggacgg	tggtcgcggc	agcagagcag	1320
ctgggataca	aggtcaacta	cgcggtatgg	acggctatca	acgcgaccaa	cacgactggc	1380
tttgcagagg	cagtggccgc	agccaagtgc	tccgacgtcg	tcatctatgc	gggcggcatc	1440
gacaactoga	tcgaagcgga	gggccacgac	cgcgacacga	tcgtctggcc	gggcaaccag	1500
ctgcagctga	tcagcgagct	agcgcagact	ggaaaaccgc	tggtcgtcat	ccaattcggc	1560
ggcgggcaag	tggacgactc	gtccctgctg	gcaaacgaca	gcggcgtcaa	cgccctgctg	1620
tgggcgggct	atcccagcca	ggctgggggc	gccgccatct	tcgaaatcct	gacgggagcg	1680
acagcgcggg	caggccgtct	cccgacgacg	cagtaccggg	cgcagtacgt	cgacgaggtg	1740

ES 2 802 807 T3

ccgatgacgg acatgacgct ggggcccgagc gcgacgaacc cgggacggac gtaccggtgg 1800  
 tacgacaagg cggatgatccc gttcggattc gggctgcact acacgacgct tgacgtgacg 1860  
 tggagcaagg cgcagctggg accgtacgag atcggctctc tcgtgaaaaa cgcctcttct 1920  
 tccgatgcmc ctgccgacac ggctccgttc gacaccttca ccatccacgt ccgaaacacg 1980  
 ggcaagacga cctcagacta cgtcgcgctg ctcttcctct cgacgcgcaa cgcgggacca 2040  
 gcaccgtatc cgctcaagac gctggtgggc tacacgcggg cgcgggcgat ccagccgggc 2100  
 gagacgcgcg cggtcgacat cgcggtgacg gtgggatcgg tggcgcggac agacgagcgc 2160  
 ggggatctgg tactgtatcc ggggacgtac acgctggagg tggatgtgaa cgggcagtat 2220  
 ccaactgcag ggtttgaggt gaccggggag gcggcggctc tggatgagtt tccgcagccg 2280  
 ccggcataa 2289

<210> 5

5 <211> 2289

<212> ADN

<213> Rasamsonia emersonii

10 <220>

<221> fuente

15 <222> (1)..(2289)

<223> /organismo="Rasamsonia emersonii" /tipo\_molar="ADN no asignado"

20 <400> 5

atggttctgg gtgtcagtct cgtcctcctg gcgacggccg tctctgccac gttccccgac 60  
 tgcagccagc ctccgctgaa ggacaacgcc gtgtgcgata cctcgtgga ccccgatcgc 120  
 cgtgctgctg cgctcgtcgc ggccttcacg ctcgaggaga agatcaaaa cacgcagaat 180  
 ggatcgcctg gcgtgccccg gctgggcctg ccgccgtatc aatggtggag cgaagcctg 240  
 cacggcgtgg ccatctcgcc gggcgtcaac ttctccgcgc acggcgactt cagctacgca 300  
 acctccttcc cccagcccat cctcatgagc gccgcctttg acgacgacct catcaggcag 360  
 gtcggcagtg tggtcagcac cgaggcgcgc gccttcagca acgccaaccg cgctggcctc 420  
 gactactgga cgcccaacat caacccttc aaggaccgcg gctggggtcg cgggcaggag 480  
 acgcccggcg aagacgcatt ccacatccag cgctacgtct acagcctcat cgacggcctg 540  
 caaaacggca ttgggcccgc caaccceaag atcatggcca cctgcaagca cttcgccgcc 600  
 tacgacctcg aggactggca cggcaacgag cgctatggct tcaacgccgt cgtgtccacg 660  
 caggacctgg ccgagtacta cctgccgccc ttcaagagct gcgccgcga cgccaaagt 720  
 gacgcgctca tgtcagcta caacgccgtc aacggggtgc cctcgtgcmc cgactcgtat 780  
 ctgctggagg acatcctgcm cgaccactgg ggctggaaca cgaccggcca ctgggtgacg 840  
 tccgactgcm atgccgtgca gaacatctac gccaacctac actacacgctc gactgcgccc 900

ES 2 802 807 T3

caggcggccg cggacgccct gggcgccggc accgatctcg actgcggcac cacttatccc 960  
gacaatctgg gcgctgccta cacgcagggg ctcttccaga accagaccct cgacacggcg 1020  
ctcatccgcc tgtactcgtc gctcgtcaag ctgggctact ttgatccgcc ggagaaccag 1080  
ccctaccggt cgataggctg gagtcatgtg tccacgccgg ccgcgagca gctggcccgc 1140  
acggcggcgg cggaggggat cgttcttctc aagaacgacg agaagaaagt cctgccgctg 1200  
tcgcgcgagg gacagacgct cgccgtgatc gggcccttcg ccaacgcgac gaccagctg 1260  
cagggcaatt accaaggcgt ggcgccgtac atctggacgg tggtcgcggc agcagagcag 1320  
ctgggataca aggtcaacta cgcggatggc acggctatca acgcgaccaa cacgactggc 1380  
tttgacagagg cagtggccgc agccaagtgc tccgacgtcg tcatctatgc gggcggcatc 1440  
gacaactcga tcgaagcggg gggccacgac cgcgacacga tcgtctggcc gggcaaccag 1500  
ctgcagctga tcagcgagct agcgcagact ggaaaaccgc tggtcgtcat ccaattcggc 1560  
ggcgggcaag tggacgactc gtccctgctg gcaaacgaca gcggcgtcaa cgccctgctg 1620  
tgggcgggct atcccagcca ggctgggggc gccgccatct tcgaaatcct gacgggagcg 1680  
acagcgcggc caggccgtct cccgacgacg cagtaccggt cgcagtacgt cgacgaggtg 1740  
ccgatgacgg acatgacgct gcggccgagc gcgacgaacc cgggacggac gtaccggtgg 1800  
tacgacaagg cggatgatcc gtccggttc gggctgcact acacgacgct tgacgtgacg 1860  
tggagcaagg cgcagctggg accgtacgag atcggctctc tcgtgaaaaa cgcctcttct 1920  
tccgatgccc ctgcgcacac ggctccgttc gacaccttca ccatccacgt ccgaaacacg 1980  
ggcaagacga cctcagacta cgtcgcgctg ctcttctctc cgacgcgcaa cgcgggacca 2040  
gcaccgtatc cgtcaagac gctggtgggc tacacgcggg cgcgggagat ccagccgggc 2100  
gagacgcgcg cggtcgacat cgcggtgacg gtgggatcgg tggcgcggac agacgagcgc 2160  
ggggatctgg tactgtatcc ggggacgtac acgctggagg tggatgtgaa cgggcagtat 2220  
ccaactgcag ggtttgaggt gaccggggag gcgcggttc tggatgagtt tccgcagccg 2280  
ccggcataa 2289

<210> 6

5 <211> 2388

<212> ADN

10 <213> Rasamsonia emersonii

<220>

<221> fuente

15 <222> (1)..(2388)

<223> /organismo="Rasamsonia emersonii" /tipo\_molar="ADN no asignado"

<400> 6

20

atggtgacct gtgctgcat tttgactgct gttgctgctc tcctccccac tgccacctgg 60

ES 2 802 807 T3

gcgcaggaca accagaccta cgccaactac ttttcccaga gccagcccga cctcttcccc 120  
 cgcaccggtg ccaccatcga tctttctttc cctgactgcg agaacggacc tctgagcacc 180  
 aacctcgtct gcaacaagtc ggcggacccc tgggctcgtg ctgaggctct gatctccctc 240  
 ttcaactttg aggagctcat caacaacacc cagaacactg ctcccgggtg tcctcgtctg 300  
 ggcctcccc agtaccaggt ctggaacgag gccttgacg gtcttgaccg tgccaacttc 360  
 tctcactccg gtgaatactc ttgggccacc agcttcccca tgccatcct gagcatggcc 420  
 tccttcaacc gcaccctcat caaccagatt gcctccatca ttgccactca ggcccgtgct 480  
 ttcaacaacg ctggccgcta cggcttgat tcctacgctc ccaacatcaa cggtttccgc 540  
 tctcctctct ggggtcgtg tcaaggaaacc cccggtgagg atgccttctt cctgagctca 600  
 acatacgct acgaatacat caccggcttg cagggtggtg ttgacctga gcacgtcaag 660  
 atcgtcgcca ctgccaagca cttcgtgga tacgacctg agaactggg taacgtgtct 720  
 cgcttggtt tcaacgccat catcaccag caggatctt ctgaacta cactccccag 780  
 ttctggcct ccgctcgtc cgccaagacc cgcagcatca tgtgtccta caacgcgctc 840  
 aacggtgttc cctcctgcgc gaactcgttc ttctgcaga ctctcctcc cgagaacttc 900  
 gacttcgtcg acgatggata cgtgagctcc gactgcgat cgtctaca cgtgttcaac 960  
 cccacggat acgctctcaa ccagtctggt gctgctgccc actctcttct ggctgttacc 1020  
 gacatcgact gcggtcagac cctcccctgg cacctgaacg agagcttctt ggaaggatat 1080  
 gtctcccgtg gtgacattga gaagtccctc acccgtctgt actccaacct ggtccgccta 1140  
 ggatacttcg atggaaacaa cagcgaatac cgcaacctca actggaacga cgttgtcacc 1200  
 accgatgctt ggaacatctc ctacgaagcc gccgttgagg gtatcaccct gctcaagaac 1260  
 gatggcacc ttctctttc caagaaggtc cgctccatcg ccttgattgg accttgggcc 1320  
 aacgccactg tccagatgca gggcaactac tacggcactc ctccctacct catctcccct 1380  
 ctggaggctg ccaaggcctc cggcttcacc gtcaactacg ctttcggcac caacatctcc 1440  
 accgacagca cccagtgggt cgctgaggcc atcgctgctg ccaagaagtc ggatgtcatc 1500  
 atctacgctg gtggcattga caacaccatt gaggctgagg gccaggaccg taccgacttg 1560  
 aatggcccc gcaaccagct cgacctcatt gagcagctgt cccaggctcg caagcctctg 1620  
 gtcgtcctcc agatgggtg tggccaggtc gactcttct ccctcaaggc caacaagaac 1680  
 gtcaacgccc ttgtctggg tggataccct ggccagtcg gcggtgctgc tcttttcgac 1740  
 atcctgaccg gcaagcgtgc tcctgctggt cgtcttgtct ccaccagta ccccgctgaa 1800  
 tacgccacc agttccccgc caacgacatg aacctgcgcc ccaacggctc caccccggc 1860  
 cagacttaca tctggtacac cgttactccc gtctacgagt tcggccacgg tctgttctac 1920

ES 2 802 807 T3

actgagttcc aggagtccgc tgctgctggt accaacaaga ccagcacttt tgacatcctg 1980  
gatctgttct ccaactcctca ccctgggttac gagtacatcg agcaggttcc cttcatcaac 2040  
gtcaccgttg atgtcaagaa cggttgccac accccttctc cctacaccgg cctcctcttc 2100  
gccaacacca ctgctggtcc caagccctac cccaacaagt ggctggtcgg tttcgaccgt 2160  
ctgcccacca ttcaaccagg cgaaaccgcc aagctcacca tccccgttcc tcttggtgcc 2220  
attgcccggtg ccgatgagaa cggcaacaag gttgtgttcc ccggttaacta cgagcttgct 2280  
ctcaacaacg agcgctccgt tgttgtctct ttcaccctca ccggtgatgc tgccaccctc 2340  
gagaagtggc ccctctggga gcaggctggt cccggtgtca ctcagcag 2388

<210> 7

5 <211> 775

<212> PRT

<213> Rasamsonia emersonii

10

<400> 7

Gln Asp Asn Gln Thr Tyr Ala Asn Tyr Ser Ser Gln Ser Gln Pro Asp  
1 5 10 15

Leu Phe Pro Arg Thr Val Ala Thr Ile Asp Leu Ser Phe Pro Asp Cys  
20 25 30

Glu Asn Gly Pro Leu Ser Thr Asn Leu Val Cys Asn Lys Ser Ala Asp  
35 40 45

Pro Trp Ala Arg Ala Glu Ala Leu Ile Ser Leu Phe Thr Leu Glu Glu  
50 55 60

Leu Ile Asn Asn Thr Gln Asn Thr Ala Pro Gly Val Pro Arg Leu Gly  
65 70 75 80

Leu Pro Gln Tyr Gln Val Trp Asn Glu Ala Leu His Gly Leu Asp Arg  
85 90 95

Ala Asn Phe Ser His Ser Gly Glu Tyr Ser Trp Ala Thr Ser Phe Pro  
100 105 110

Met Pro Ile Leu Ser Met Ala Ser Phe Asn Arg Thr Leu Ile Asn Gln  
115 120 125

Ile Ala Ser Ile Ile Ala Thr Gln Ala Arg Ala Phe Asn Asn Ala Gly  
130 135 140

Arg Tyr Gly Leu Asp Ser Tyr Ala Pro Asn Ile Asn Gly Phe Arg Ser  
145 150 155 160

ES 2 802 807 T3

Pro Leu Trp Gly Arg Gly Gln Glu Thr Pro Gly Glu Asp Ala Phe Phe  
 165 170 175

Leu Ser Ser Thr Tyr Ala Tyr Glu Tyr Ile Thr Gly Leu Gln Gly Gly  
 180 185 190

Val Asp Pro Glu His Val Lys Ile Val Ala Thr Ala Lys His Phe Ala  
 195 200 205

Gly Tyr Asp Leu Glu Asn Trp Gly Asn Val Ser Arg Leu Gly Phe Asn  
 210 215 220

Ala Ile Ile Thr Gln Gln Asp Leu Ser Glu Tyr Tyr Thr Pro Gln Phe  
 225 230 235 240

Leu Ala Ser Ala Arg Tyr Ala Lys Thr Arg Ser Ile Met Cys Ser Tyr  
 245 250 255

Asn Ala Val Asn Gly Val Pro Ser Cys Ala Asn Ser Phe Phe Leu Gln  
 260 265 270

Thr Leu Leu Arg Glu Asn Phe Asp Phe Val Asp Asp Gly Tyr Val Ser  
 275 280 285

Ser Asp Cys Asp Ala Val Tyr Asn Val Phe Asn Pro His Gly Tyr Ala  
 290 295 300

Leu Asn Gln Ser Gly Ala Ala Ala Asp Ser Leu Leu Ala Gly Thr Asp  
 305 310 315 320

Ile Asp Cys Gly Gln Thr Leu Pro Trp His Leu Asn Glu Ser Phe Val  
 325 330 335

Glu Gly Tyr Val Ser Arg Gly Asp Ile Glu Lys Ser Leu Thr Arg Leu  
 340 345 350

Tyr Ser Asn Leu Val Arg Leu Gly Tyr Phe Asp Gly Asn Asn Ser Glu  
 355 360 365

Tyr Arg Asn Leu Asn Trp Asn Asp Val Val Thr Thr Asp Ala Trp Asn  
 370 375 380

Ile Ser Tyr Glu Ala Ala Val Glu Gly Ile Thr Leu Leu Lys Asn Asp  
 385 390 395 400

Gly Thr Leu Pro Leu Ser Lys Lys Val Arg Ser Ile Ala Leu Ile Gly  
 405 410 415

ES 2 802 807 T3

Pro Trp Ala Asn Ala Thr Val Gln Met Gln Gly Asn Tyr Tyr Gly Thr  
 420 425 430

Pro Pro Tyr Leu Ile Ser Pro Leu Glu Ala Ala Lys Ala Ser Gly Phe  
 435 440 445

Thr Val Asn Tyr Ala Phe Gly Thr Asn Ile Ser Thr Asp Ser Thr Gln  
 450 455 460

Trp Phe Ala Glu Ala Ile Ala Ala Ala Lys Lys Ser Asp Val Ile Ile  
 465 470 475 480

Tyr Ala Gly Gly Ile Asp Asn Thr Ile Glu Ala Glu Gly Gln Asp Arg  
 485 490 495

Thr Asp Leu Lys Trp Pro Gly Asn Gln Leu Asp Leu Ile Glu Gln Leu  
 500 505 510

Ser Gln Val Gly Lys Pro Leu Val Val Leu Gln Met Gly Gly Gly Gln  
 515 520 525

Val Asp Ser Ser Ser Leu Lys Ala Asn Lys Asn Val Asn Ala Leu Val  
 530 535 540

Trp Gly Gly Tyr Pro Gly Gln Ser Gly Gly Ala Ala Leu Phe Asp Ile  
 545 550 555 560

Leu Thr Gly Lys Arg Ala Pro Ala Gly Arg Leu Val Ser Thr Gln Tyr  
 565 570 575

Pro Ala Glu Tyr Ala Thr Gln Phe Pro Ala Asn Asp Met Asn Leu Arg  
 580 585 590

Pro Asn Gly Ser Asn Pro Gly Gln Thr Tyr Ile Trp Tyr Thr Gly Thr  
 595 600 605

Pro Val Tyr Glu Phe Gly His Gly Leu Phe Tyr Thr Glu Phe Gln Glu  
 610 615 620

Ser Ala Ala Ala Gly Thr Asn Lys Thr Ser Thr Phe Asp Ile Leu Asp  
 625 630 635 640

Leu Phe Ser Thr Pro His Pro Gly Tyr Glu Tyr Ile Glu Gln Val Pro  
 645 650 655

Phe Ile Asn Val Thr Val Asp Val Lys Asn Val Gly His Thr Pro Ser

ES 2 802 807 T3

	660		665		670	
	Pro Tyr Thr Gly Leu Leu Phe Ala Asn Thr Thr Ala Gly Pro Lys Pro					
	675		680		685	
	Tyr Pro Asn Lys Trp Leu Val Gly Phe Asp Arg Leu Pro Thr Ile Gln					
	690		695		700	
	Pro Gly Glu Thr Ala Lys Leu Thr Ile Pro Val Pro Leu Gly Ala Ile					
	705		710		715	720
	Ala Arg Ala Asp Glu Asn Gly Asn Lys Val Val Phe Pro Gly Asn Tyr					
		725		730		735
	Glu Leu Ala Leu Asn Asn Glu Arg Ser Val Val Val Ser Phe Thr Leu					
		740		745		750
	Thr Gly Asp Ala Ala Thr Leu Glu Lys Trp Pro Leu Trp Glu Gln Ala					
		755		760		765
	Val Pro Gly Val Thr Gln Gln					
	770		775			
	<210> 8					
5	<211> 21					
	<212> PRT					
	<213> Rasamsonia emersonii					
10	<400> 8					
	Met Val Thr Arg Ala Ala Ile Leu Thr Ala Val Ala Ala Leu Leu Pro					
	1	5		10		15
	Thr Ala Thr Trp Ala					
		20				
15	<210> 9					
	<211> 2391					
	<212> ADN					
20	<213> Rasamsonia emersonii					
	<220>					
25	<221> fuente					
	<222> (1)..(2391)					
	<223> /organismo="Rasamsonia emersonii" /tipo_molar="ADN no asignado"					
30	<400> 9					

ES 2 802 807 T3

atggtgactc gcgcgcgcat tctcaccgca gtggcggcgc tcctgcccac cgcgacatgg 60  
 gcacaggata accaaaccta tgccaattac tcgtcgcagt ctgagccgga cctgtttccc 120  
 cggaccgtcg cgaccatcga cctgtccttc cccgactgtg agaatggccc gctcagcacg 180  
 aacctggtgt gcaacaaatc ggccgatccc tgggcccagag ctgaggccct catctcgctc 240  
 tttaccctcg aagagctgat taacaacacc cagaacaccg ctcttgccgt gccccgtttg 300  
 ggtctgcccc agtatcaggt gtggaatgaa gctctgcacg gactggaccg cgccaatttc 360  
 tcccattcgg gcgaatacag ctgggccacg tccttcccca tgcccatcct gtcgatggcg 420  
 tccttcaacc ggaccctcat caaccagatt gcctccatca ttgcaacgca agcccgtgcc 480  
 ttcaacaacg ccggccgtta cggccttgac agctatgctc ccaacatcaa tggcttccgc 540  
 agtcccctct ggggccgtgg acaggagacg cctggtgagg atgcttctt cttgagttcc 600  
 acctatgctg acgagtacat cacaggcctg cagggcggtg tcgaccaga gcatgtcaag 660  
 atcgtcgcga cggcgaagca cttcgcggc tatgatctgg agaactggg caacgtctct 720  
 cggctggggg tcaatgctat catcacgcag caggatctct ccgagtacta caccctcag 780  
 ttcttgccgt ctgctcgata cgccaagacg cgcagcatca tgtgctccta caatgcagtg 840  
 aatggagtcc caagctgtgc caactccttc ttctccaga cgcttctcc agaaaacttt 900  
 gacttcgttg acgacgggta cgtctcgtcg gattgcgacg ccgtctaca cgtcttcaac 960  
 ccacacgggt acgcccttaa ccagtcggga gccgctgcgg actcgtcctc agcaggtacc 1020  
 gatatcgact gtggtcagac cttgccgtgg cacctgaatg agtccttctg agaaggatac 1080  
 gtctcccgcg gtgatatcga gaaatccctc acccgtctct actcaaacct ggtgcgtctc 1140  
 ggctactttg acggcaacaa cagcagtagc cgcaacctca actggaacga cgtcgtgact 1200  
 acggagcctt ggaacatctc gtacgaggcc cgggtggaag gtatcaccct gctcaagaac 1260  
 gacggaacgc tgccgctgtc caagaaggtc cgcagcattg cgctcatcgg tccttgggccc 1320  
 aatgccacgg tgcagatgca gggtaactac tatggaacgc caccgtatct gatcagtcgg 1380  
 ctggaagccg ccaaggccag tgggttcacg gtcaactatg cattcggtag caacatctcg 1440  
 accgattcta cccagtgtt cgcggaagcc atcgcggcgg cgaagaagtc ggacgtgatc 1500  
 atctacgccg gtggtattga caacacgatc gaggcagagg gacaggaccg cacggatctc 1560  
 aagtggccgg ggaaccagct ggatctgatc gagcagctca gccagggtgg caagcccttg 1620  
 gtcgtcctgc agatgggcgg tggccaggtg gattcgtcgt cactcaaggc caacaagaat 1680  
 gtcaacgctc tgggtgtggg tggctatccc ggacagtcgg gtggtgcggc cctgtttgac 1740  
 atccttacgg gcaagcgtgc gccggccggt cgtctggtga gcacgcagta cccggccgag 1800  
 tatgcgacgc agttcccggc caacgacatg aacctgcgtc cgaacggcag caaccggga 1860  
 cagacataca tctggtacac gggcacgccc gtgtatgagt tcggccacgg tctgttctac 1920  
 acggagtcc aggagtcggc tgcggcgggc acgaacaaga cgtcgacttt cgacattctg 1980  
 gacctttct ccaccctca tccgggatac gagtacatcg agcaggttcc gttcatcaac 2040

ES 2 802 807 T3

gtgactgtgg acgtgaagaa cgctggccac acgccatcgc cgtacacggg tctgttgttc 2100  
gcgaacacga cagccgggcc caagccgtac ccgaacaaat ggctcgtcgg gttcgaccgg 2160  
ctgccgacga tccagccggg cgagactgcc aagttgacga tcccggtgcc gttgggcgcg 2220  
attgctcggg cggacgagaa cggcaacaag gtggtcttcc cgggcaacta cgaattggca 2280  
ctgaacaatg agcgatcggg agtgggtgtcg ttcacgctga cgggcgatgc ggcgactcta 2340  
gagaaatggc ctttgtggga gcaggcggg cgggggtga cccagcagta g 2391

<210> 10

5 <211> 2391

<212> ADN

<213> Rasamsonia emersonii

10 <220>

<221> fuente

15 <222> (1)..(2391)

<223> /organismo="Rasamsonia emersonii" /tipo\_molar="ADN no asignado"

<400> 10

20 atggtgactc gcgcggcgat tctcaccgca gtggcggcgc tcctgcccac cgcgacatgg 60  
gcacaggata accaaaccta tgccaattac tcgtcgcagt ctccagccgga cctggttccc 120  
cggaccgtcg cgaccatcga cctgtccttc cccgactgtg agaatggccc gctcagcacg 180  
aacctggtgt gcaacaaatc ggccgatccc tgggcccagag ctgaggccct catctcgctc 240  
tttaccctcg aagagctgat taacaacacc cagaacaccg ctctggcgt gccccgtttg 300  
ggctgcccc agtatcagg gtggaatgaa gctctgcacg gactggaccg cgccaatttc 360  
tcccattcgg gcgaatacag ctgggccaag tccttcccga tgcccatcct gtcgatggcg 420  
tccttcaacc ggaccctcat caaccagatt gcctccatca ttgcaacgca agcccgtgcc 480  
ttcaacaacg ccggccgcta cggccttgac agctatgcgc ccaacatcaa tggcttccgc 540  
agtcccctct ggggcccgtg acaggagacg cctggtgagg atgogttctt cttgagttcc 600  
acctatgcgt acgagtacat cacaggcctg caggcgggtg tcgaccaga gcatgtcaag 660  
atcgtcgcga cggcgaagca cttcgcgggc tatgatctgg agaactgggg caacgtctct 720  
cggctggggt tcaatgctat catcacgcag caggatctct ccgagtacta caccctcag 780  
ttcctggcgt ctgctcgata cgccaagacg cgcagcatca tgtgctccta caatgcagtg 840  
aatggagtcc caagctgtgc caactccttc ttcctccaga cgcttctccg agaaaacttt 900  
gacttcgttg acgacgggta cgtctcgtcg gattgcgacg ccgtctaaa cgtcttcaac 960  
ccacacggtt acgcccttaa ccagtcggga gccgctgagg actcgtcctc agcaggtacc 1020  
gatatcgact gtggtcagac cttgcccgtg cacctgaatg agtccttctg agaaggatac 1080  
gtctcccgcg gtgatatcga gaaatccctc acccgtctct actcaaacct ggtgcgtctc 1140

ES 2 802 807 T3

ggctactttg acggaacaa cagcgagtac cgcaacctca actggaacga cgtcgtgact 1200  
 acggacgcct ggaacatctc gtacgaggcc gcggtggaag gtatcaccct gctcaagaac 1260  
 gacggaacgc tgccgctgtc caagaaggtc cgcagcattg cgctcatcgg tccttgggcc 1320  
 aatgccacgg tgcagatgca gggtaactac tatggaacgc caccgtatct gatcagtcgg 1380  
 ctggaagccg ccaaggccag tgggttcacg gtcaactatg cattcggtag caacatctcg 1440  
 accgattcta cccagtgggt cgcggaagcc atcgcggcgg cgaagaagtc ggacgtgatc 1500  
 atctacgccg gtggtattga caacacgatc gaggcagagg gacaggaccg cacggatctc 1560  
 aagtggccgg ggaaccagct ggatctgatc gagcagctca gccagggtgg caagcccttg 1620  
 gtcgtcctgc agatgggagg tggccagggt gattcgtcgt cactcaaggc caacaagaat 1680  
 gtcaacgctc tgggtggtgg tggctatccc ggacagtcgg gtggtgcggc cctgtttgac 1740  
 atccttacgg gcaagcgtgc gccggccggt cgtctggtga gcacgcagta cccggccgag 1800  
 tatgcgacgc agttcccggc caacgacatg aacctgcgtc cgaacggcag caaccggga 1860  
 cagacataca tctggtacac gggcacgccc gtgtatgagt tcggccacgg tctgtttctac 1920  
 acggagtacc aggagtcggc tgcggcgggc acgaacaaga cgtcgacttt cgacattctg 1980  
 gaccttttct ccaccctca tccgggatac gagtacatcg agcaggttcc gttcatcaac 2040  
 gtgactgtgg acgtgaagaa cgtcggccac acgccatcgc cgtacacggg tctgttgttc 2100  
 gcgaacacga cagccgggccc caagccgtac ccgaacaaat ggctcgtcgg gttcgaccgg 2160  
 ctgccgacga tccagccggg cgagactgcc aagttgacga tcccgggtgcc gttgggcggg 2220  
 attgcccggg cggacgagaa cggcaacaag gtggtcttcc cgggcaacta cgaattggca 2280  
 ctgaacaatg agcgatcggg agtgggtgctg ttcacgctga cgggcgatgc ggcgactcta 2340  
 gagaaatggc ctttgtggga gcaggcgggt ccgggggtga cccagcagta g 2391

<210> 11

5 <211> 1473

<212> ADN

10 <213> *Rasamsonia emersonii*

<220>

<221> fuente

15 <222> (1)..(1473)

<223> /organismo="Rasamsonia emersonii" /tipo\_molar="ADN no asignado"

<400> 11

20

atgggtcgcc gcattgccac tgctgctctg ctctgttctg gtctgaacgc cctcaagggt 60  
 gttgttgctc tcgacaacgg tcttgccatc actccccaga tgggctggaa cacctggaac 120  
 tcgttcggct gctcgtgtaa cgaaaccgtc atcctggatg ctgctgagaa gcttgtgtcg 180  
 ctaggattca aggacctagg atacgagtac gtggtgctgg atgactgctg gtccgctggc 240

ES 2 802 807 T3

cgtaacgcct ctggctacct ggttcctgac cctgccaagt tccccaacgg cattgatggt 300  
 cttgccgaga agatccacgc tcttggcctc aagatcggta tctactcttc tgctggtacc 360  
 atgacctgcg ctcgttacgc tggctctctt ggatacagag agaaggatgc cgccctctgg 420  
 gcctcctggg gtatcgacta cctcaagtac gacaactgct acaaccaggg tcaagaggg 480  
 actccaagc tcagctacga cgcctacaac gccatgggcc aggctctgaa caagaccggc 540  
 cgtcccattc tctactccct ctgcaactgg ggtgttgacg gccctggaa cttcgcctcc 600  
 accattgcca acagctggcg cacctccggt gacctgctca acacctggga ccgtgatgat 660  
 gccaaactgcc cctgctctga gctcgaaggc ttggactgca agacccccgg ctacaagtgc 720  
 tcgatcctga acgcatcaa caaggccgtc tactaccctt ccaaggcctt ccctggtgcc 780  
 tggaacgacc ttgacatgct ccaggtcggc aacggtggat tgaccgacga tgaggccgtt 840  
 gctcacatga gcctctgggc tgctttcaag tctcctctcc tcatgaccaa cgtcctcagc 900  
 aacattgacc ctcccaccct gtccatcctg cagaaccccg ccgttctagc ggtgtctcag 960  
 gaccccgctg gtcctccgt caaccgtatc tggcgtact acgtggatga tgtcgacgcc 1020  
 aacggatacg gcgagatcca gctgttctcc ggtggtcttg ctggtggtga ccagctcgtc 1080  
 ctccttctga acgctggcag caaggaccgt accatgaacg ccaccctgga ggacatcttc 1140  
 tgggaagatg gacctggtg tactgcctcc caggtcacgc agtctggga tgtctacgat 1200  
 ctctgggcca accgcatgag caacgagact gctgctgcca tcatcaacgc cgccaacagc 1260  
 actggaagcg ctgctcctgc tgctcccatc aacatgactg ctcttgggtg tgccaagcac 1320  
 gtctactctc aggttctctc ctccgactcc aaggctctga tgggtaccaa ggtcggctcc 1380  
 gtccagccct ccggcaccgt caaggcttcc gtcaaggccc acggtggtgc catgctccgc 1440  
 ctccgtcagc agtcccagaa gaaggacgaa cta 1473

<210> 12

5 <211> 468

<212> PRT

<213> Rasamsonia emersonii

10

<400> 12

Leu Asp Asn Gly Leu Ala Ile Thr Pro Gln Met Gly Trp Asn Thr Trp  
 1 5 10 15

Asn Ser Phe Gly Cys Ser Leu Asn Glu Thr Val Ile Leu Asp Ala Ala  
 20 25 30

Glu Lys Leu Val Ser Leu Gly Phe Lys Asp Leu Gly Tyr Glu Tyr Val  
 35 40 45

Val Leu Asp Asp Cys Trp Ser Ala Gly Arg Asn Ala Ser Gly Tyr Leu



ES 2 802 807 T3

Asn Arg Ile Trp Arg Tyr Tyr Val Asp Asp Val Asp Ala Asn Gly Tyr  
305 310 315 320

Gly Glu Ile Gln Leu Phe Ser Gly Gly Leu Ala Gly Gly Asp Gln Leu  
325 330 335

Val Leu Leu Leu Asn Ala Gly Ser Lys Asp Arg Thr Met Asn Ala Thr  
340 345 350

Leu Glu Asp Ile Phe Trp Glu Asp Gly Pro Gly Gly Thr Ala Ser Gln  
355 360 365

Val Gln Gln Ser Trp Asp Val Tyr Asp Leu Trp Ala Asn Arg Met Ser  
370 375 380

Asn Glu Thr Ala Ala Ala Ile Ile Asn Ala Ala Asn Ser Thr Gly Ser  
385 390 395 400

Ala Ala Pro Ala Ala Pro Ile Asn Met Thr Ala Leu Gly Gly Ala Lys  
405 410 415

His Val Tyr Ser Gln Val Pro Pro Ser Asp Ser Lys Ala Leu Met Gly  
420 425 430

Thr Lys Val Gly Ser Val Gln Pro Ser Gly Thr Val Lys Ala Phe Val  
435 440 445

Lys Ala His Gly Val Ala Met Leu Arg Leu Arg Gln Gln Ser Gln Lys  
450 455 460

Lys Asp Glu Leu  
465

<210> 13

5 <211> 23

<212> PRT

<213> Rasamsonia emersonii

10

<400> 13

Met Gly Arg Arg Ile Ala Thr Ala Ala Leu Leu Leu Phe Gly Leu Asn  
1 5 10 15

Ala Leu Lys Gly Val Val Ala  
20

15 <210> 14

<211> 1670

<212> ADN

20

<213> Rasamsonia emersonii

<220>

ES 2 802 807 T3

<221> fuente

<222> (1)..(1670)

5 <223> /organismo="Rasamsonia emersonii" /tipo\_molar="ADN no asignado"

<400> 14

```

atggggagaa gaattgctac tgctgcactt ctcttgctcg ggctcaatgc cttgaaagga      60
gtggtggcgc tggacaatgg cctggccatc actccgcaga tgggatgggt aagaatatat      120
ccaattgaat ggccaaactt gaaagttgca attggaattt gaataatcct gctcatgcat      180
aagtcaacag aacacatgga actccttcgg atgctcgcta aacgagaccg tcatcctaga      240
cgcgccgaa aagctcgtct cgctcggatt caaggatctc ggctatgagt acgtcgtcct      300
ggacgactgc tggctgccc gccggaacgc gtctggctat ctggtgcccg atccagccaa      360
gtttcccaat ggcatcgatg gcctggcggg gaagatccat gcgctgggct tgaagattgg      420
catttattcc agtgcgggaa ccatgacttg cgcccgtat gctgggtcgt tgggatatga      480
agagaaggat gcagcgcttt gggctagtgt ggggtaagt acgcctaaag ggcataaaac      540
ctgcaaagaa aataacttaa aaaaaacgta aatagatcga ctacctcaa tacgacaact      600
gctacaacca aggccaagaa ggcacgcca agctctogta cgaccgctac aacgccatgg      660
gccaagccct caacaagaca ggcgcccga ttctctactc gctctgcaac tggggcgtcg      720
acggcccgtg gaactttgcc togaccatcg ccaactcgtg gcgcacgtcg ggcgatctgc      780
tcaacacgtg ggaccgac gacgcaaact gtccgtgcag tgagctcgag ggcctcgact      840
gcaagactcc cgggtacaag tgttcgatcc tgaacgtcat caacaaggcg gtgtattatc      900
cgtctaaggc gttccctggg gcgtggaatg acctggacat gctacgtatg tctccatcta      960
tctatcattc tattacatga tctaacacgc ctgcagaggt tggcaacgga ggcctcaccg     1020
acgacgaagc cgtcgccac atgagtctct gggcagcctt caagtccccg ctgctgatga     1080
ccaacgtcct gagcaacatc gaccctccta cgctgtccat cctgcagaac cggccgtgc      1140
tggccgtctc gcaggaccg gtggctcca gcgtcaaccg catctggcgg tactacgtcg     1200
acgacgtgga cgccaacggg tacggcgaga tccagctgtt cagcggcggc ctggcaggag     1260
gcgaccagct ggtcgtgctg ctgaacgcgg ggtcgaagga ccgcacatg aacgcgacgc     1320
tggaggacat cttctgggaa gacgggcccg gtggcacggc cagccaggta cagcagagct     1380
gggacgtgta cgacctgtgg gogaaccgga tgagcaacga gacggcggcg gcgatcatca     1440
acgccgcaa ctgcaccggt tctgctgccc ctgctgccc gatcaacatg accgccctcg     1500
ggggtgcaa gcacgtgtac tcgcaggtgc cggctccga ctggaaggcg ttgatgggca     1560
ccaaggttg gagcgtccag ccgagtggga cggcgaaggc gtttgtcaag gcgcacggcg     1620
tggcgtggt gcgctgccc cagcagtcgc agaagaagga tgagttgtaa     1670

```

10

<210> 15

<211> 1476

15

<212> ADN

ES 2 802 807 T3

<213> Rasamsonia emersonii

<220>

5 <221> fuente

<222> (1)..(1476)

<223> /organismo="Rasamsonia emersonii" /tipo\_molar="ADN no asignado"

10

<400> 15

```

atggggagaa gaattgctac tgctgcactt ctcttgctcg ggctcaatgc cttgaaagga      60
gtgggtggcgc tggacaatgg cctggccatc actccgcaga tgggatggaa cacatggaac      120
tccttcggat gctcgctaaa cgagaccgtc atcctagacg cggccgaaaa gctcgtctcg      180
ctcggattca aggatctcgg ctatgagtac gtcgtcctgg acgactgctg gtctgccggc      240
cggaaacgcgt ctggctatct ggtgccggat ccagccaagt ttccaatgg catcgatggc      300
ctggcggaga agatccatgc gctgggcttg aagattggca tttattccag tgcggaacc      360
atgacttgcg cccgctatgc tgggtcgttg ggatatgaag agaaggatgc agcgctttgg      420
gctagtggg ggatcgacta cctcaaatac gacaactgct acaaccaagg ccaagaaggc      480
acgcccgaagc tctcgtacga ccgctacaac gccatgggcc aagccctcaa caagacaggc      540
cgcccgatcc tctactcgct ctgcaactgg ggcgtcgacg gcccgtgga ctttgctcg      600
accatcgcca actcgtggcg cacgtcgggc gatctgctca acacgtggga ccgcgacgac      660
gcaaactgtc cgtgcagtga gctcgagggc ctcgactgca agactcccgg gtacaagtgt      720
tcgatcctga acgtcatcaa caaggcggtg tattatccgt ctaaggcgtt ccctggggcg      780
tggaatgacc tggacatgct acaggttggc aacggaggcc tcaccgacga cgaagccgtc      840
gcccacatga gtctctgggc agccttcaag tccccgctgc tgatgaccaa cgtcctgagc      900
aacatcgacc ctctacgct gtccatcctg cagaaccggg ccgtgctggc cgtctcgag      960
gacccggtcg gctccagcgt caaccgcacg tggcggact acgtcgacga cgtggacgcc      1020
aacgggtacg gcgagatcca gctgttcagc ggcggcctgg caggaggcga ccagctggtg      1080
ctgctgctga acgcggggtc gaaggaccgc accatgaacg cgacgctgga ggacatcttc      1140
tgggaagacg ggcccggtg caccggccagc caggtacagc agagctggga cgtgtacgac      1200
ctgtggggca accggatgag caacgagacg gcggcggcga tcatcaacgc cgccaactcg      1260
accggttctg ctgcgcctgc tgcgccgatc aacatgaccg ccctcggggg tgccaagcac      1320
gtgtactcgc aggtgccgcc gtccgactcg aaggcgttga tgggcaccaa ggttgggagc      1380
gtccagccga gtgggacggt caaggcgttt gtcaaggcgc acggcgtggc gatgttcggg      1440
ctgcgccagc agtcgcagaa gaaggatgag ttgtaa      1476

```

15 <210> 16

<211> 2046

<212> ADN

20

<213> Rasamsonia emersonii

ES 2 802 807 T3

<220>

<221> fuente

5 <222> (1)..(2046)

<223> /organismo="Rasamsonia emersonii" /tipo\_molar="ADN no asignado"

<400> 16

10

```

atgactcgcct acccctccct ctccactactac gctctcagca gctaccttct cttcacctcc      60
tgcggtgctc ttgccagcac togtctggag ccccgctctgg acaacggatt ggccatcact      120
cctcccattgg gctggaactc ctacaaccac tactcttgct cccccaacga gtccatcact      180
cagtccaacg ccaaggcctt ggttgacttc ggtctggaca ccctaggata ccgctacgct      240
accaccgact gcggtggac tgtcccctac cgtctgcca acggctccct cacctggaac      300
gagactctct tcccctccgg tttccctgct cttggccagt acatccacga ccttggaactc      360
ctcttcggtg tctaccagga tgctggatc aagacctgcg gtggtcccc tgaccaggctc      420
ggttccctgt tccacgagca gcaggatgcc gagactttcg ctgcctggaa ggccgatgcc      480
ctcaagtacg acaactgcta ctcgatgct gctgctggct accccgatgc cgactacact      540
cctagcacct ccccagctt ccgttacgcc aacatgacca aggccttgc ctccgttgac      600
cgcaagatcc tcttccagat ctgcgactgg ggtgttgact tccccgcct ctgggtcct      660
tctctgggaa acacctggcg catctccaac gacattattc ctgcctggcg caccatcttc      720
cgtatcctga accaggcggg gcctcagact tctttcgctg gtcccggca ctggcccgcac      780
ctcgacatgc ttgaggttg caacaacgtg ttcaccatcc ccgaggagca gaccacttc      840
tctctgtggg ccattctcaa gtcgcctctg accattggtg ctgccctcaa cgacaccctg      900
accaccatcc gcgatgcctc cctccagatc ctcaagcaga aggatgtcat cagctacaac      960
caggaccctt tgggtgtctc tgctccctg aagcgtcgtt ggaccgaaga aggattcgag     1020
gtctggagcg gtcccatctc cggtggaag accgttgctg ctctcatcaa ctgggcggac     1080
gagtcccga acctgacctt ggatcttctt gttgttgcc tccagcacgc tcagaccctc     1140
cgcaacatct gggatgagtc tgctgccacc aacgtccgca cctoctacac tgccaacgctc     1200
gctgctcacg gcaccatggt ggttgagctt gctggtacca ctgaggctgg aaagtacccc     1260
gccgacatct tcgccacctc ggatggccac tccaccactt tcgagaacat ctacgctgag     1320
actacttctt cccagtacca gctcaccatt gctttcgctc ccggttcttc tcactcttcc     1380
gagatcacca tccgtacctc gctaggcggc ttcaagacct cggccaaggt ccagcccact     1440
gagagccaga tctccgtcaa catctccctg tctgctggca gctccaacac catcaccatc     1500

```

ES 2 802 807 T3

tctcccagcc ccatcatctc ctacatcaac gtcacctcgc ctagcggtag ctactacccc 1560  
 tgcacctcgt tcactcctgt cggttccgcc aagcccagaga cttgcgatgc tggtttctgc 1620  
 ctgcccgtcg gctccaagat tggttacatc tccccaccg gcaacgccag catcaccatt 1680  
 cccgccaccg tcgtcagcgg ctccggtag aacaccactg ccgccaccac caccgctggc 1740  
 aagtacctcg aaattgatta catcaacaac gacatcgctt tctccacctc ctggaccact 1800  
 ggcagcaact ctgcgaacct caccatttcc gtcaacggtag gtccctctac ccgtattgag 1860  
 gttcctcttt ctggcaagca cagcgaactc ttcggtcctg gtcgtggctg gtgggattct 1920  
 gccaccttcg gtgtgcttgt tgacggctgg cgcaacggtag agaacactgt tgcattggc 1980  
 aaccagggtag gtgacgaggg tgtgcagccc tacggtgccg acttcgtagc tctgcgtctg 2040  
 tacgac 2046

<210> 17

5 <211> 657

<212> PRT

<213> *Rasamsonia emersonii*

10

<400> 17

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

ES 2 802 807 T3

Ala Ala Ala Gly Tyr Pro Asp Ala Asp Tyr Thr Pro Ser Thr Ser Pro  
145 150 155 160

Ser Phe Arg Tyr Ala Asn Met Thr Lys Ala Leu Ala Ser Val Asp Arg  
165 170 175

Lys Ile Leu Phe Gln Ile Cys Asp Trp Gly Val Asp Phe Pro Ala Leu  
180 185 190

Trp Ala Pro Ser Leu Gly Asn Thr Trp Arg Ile Ser Asn Asp Ile Ile  
195 200 205

Pro Ala Trp Arg Thr Ile Phe Arg Ile Leu Asn Gln Ala Val Pro Gln  
210 215 220

Thr Ser Phe Ala Gly Pro Gly His Trp Pro Asp Leu Asp Met Leu Glu  
225 230 235 240

Val Gly Asn Asn Val Phe Thr Ile Pro Glu Glu Gln Thr His Phe Ser  
245 250 255

Leu Trp Ala Ile Leu Lys Ser Pro Leu Thr Ile Gly Ala Ala Leu Asn  
260 265 270

Asp Thr Leu Thr Thr Ile Arg Asp Ala Ser Leu Gln Ile Leu Lys Gln  
275 280 285

Lys Asp Val Ile Ser Tyr Asn Gln Asp Pro Leu Gly Val Ser Ala Ser  
290 295 300

Leu Lys Arg Arg Trp Thr Glu Glu Gly Phe Glu Val Trp Ser Gly Pro  
305 310 315 320

Ile Ser Gly Gly Lys Thr Val Ala Ala Leu Ile Asn Trp Ala Asp Glu  
325 330 335

Ser Arg Asn Leu Thr Leu Asp Leu Pro Val Val Gly Leu Gln His Ala  
340 345 350

Gln Thr Leu Arg Asn Ile Trp Asp Glu Ser Ala Ala Thr Asn Val Arg  
355 360 365

Thr Ser Tyr Thr Ala Asn Val Ala Ala His Gly Thr Met Leu Val Glu  
370 375 380

Leu Ala Gly Thr Thr Glu Ala Gly Lys Tyr Pro Ala Asp Ile Phe Ala



ES 2 802 807 T3

<212> PRT

<213> Rasamsonia emersonii

5 <400> 18

Met Thr Arg Tyr Pro Ser Leu Ser Tyr Tyr Ala Leu Ser Ser Tyr Leu  
1 5 10 15

Leu Phe Thr Ser Cys Gly Ala Leu Ala  
20 25

<210> 19

10

<211> 2315

<212> ADN

15 <213> Rasamsonia emersonii

<220>

<221> fuente

20

<222> (1)..(2315)

<223> /organismo="Rasamsonia emersonii" /tipo\_molar="ADN no asignado"

25 <400> 19

atgacaagat atcctagtct tagctactat gcactctctt catatttact atttaccagc	60
tgtggtgcac tggcatccac aagggtggaa cctaggttag acaatggact ggcaatcaca	120
cctccaatgg ggtaagtctc gtcttgtgaa gagttattat tattattatt ctagctagat	180
acagcttggc tcattcatat atgaaccgtg aagatggaac tcctacaacc actactcatg	240
ctccccgaac gaatcgatca tccagtcaaa cgcgaaggct ctcgtggact ttggccttga	300
tacccttggc tatcgatatg tcacaactga ttgcggctgg acggtcccct accgactccc	360
caatgggtcc ttgacctgga atgaaacgct gtttccgagt ggattcccgg cactagggca	420
gtacattcac gatctggggc ttctctttgg cgtctatcaa gatgcgggga ttaagacttg	480
tggaggacc cctgatcaag taggaagttt atgtaagggc tttgactccc cgggatgtg	540
cattggccgg cgtggttaat tcgaatcaag tccatgaaca acaggacgca gagacgttgc	600
ctgcctggaa agcggatgcg ttgaaatgta cgtgctgttg tgtctcttct tgcgcccgtc	660
gcattgattg atactgatct tccttccaga cgacaactgc tattctgatg cagccgctgg	720
gtatccagat gcggattata caccgagcac ctcgccatcg ttccgctacg cgaacatgac	780
aaaggcgtg gccagtgtgg atcggaaaat ccttttccaa atctgtgact ggggcgtcga	840

ES 2 802 807 T3

tttccccgcg ctctgggctc cctctctggg caacacctgg cgaatttcca acgacataat 900  
 cccccgctgg cggacgatct tcaggatcct caaccaggcc gtgccgcaga cttcttttgc 960  
 cggacctggc cactggccag acctcgatat gctcgagggtg gggaacaatg tgtttacgat 1020  
 cccggaagag cagacgcact tttccctttg ggcgatcctg aaaagcccc tgaccatcgg 1080  
 tgccgcgttg aatgatacct tgacgacgat cagagatgcc tcgctccaga tcttgaaca 1140  
 gaaggatgtc atcagctaca accaagaccc gcttggtgtg agtgccagtc tgaagaggag 1200  
 atggaccgaa gagggctttg aagtgtggag cgggccata tcgggtggga agacagtggc 1260  
 tgcattgatc aactgggctg atgaatcgag gaacttgacg ctggatctcc ctgtcgtagg 1320  
 cctacaacac gcgcagacgc tgcggaacat ctgggacgaa agcgcagcca cgaacgtccg 1380  
 tacgtcgtac actgcgaacg tggcagctca tggaacgatg ttggtagaac tggccgggac 1440  
 aacagaagca ggaagtatc ctgctgatat ttttgaact tcagatgggt cagtcctaac 1500  
 ccctcttccc ttgacaggaa aaaaaaata agggacgctg accgcttagc aggcactcga 1560  
 caacttttga aaacatctat gccgagacca ctacgagcca gtaccagttg acaatcgctg 1620  
 ttgctcccgg atcatctcac tcgagcgaaa taacgatccg aacgagcctc gggggattca 1680  
 agacgtctgc caaggtgcaa ccgactgaga gccaaatctc ggtcaacatc tcgctgtcag 1740  
 caggttcacg gaacacgatc acgatttcgc cgtctccaat tatcagttac ataaacgtca 1800  
 cctccccatc ggggacatac tatccctgta catcgtttac acctgtggga tccgccaac 1860  
 ccgagacttg cgatgcggga ttctgcctgc ctgtcgggtc caaaatcggc tatatcagcc 1920  
 cgacaggtaa tgcaagcatc actatccctg ccacgggtgt gtctggcagt ggcgataata 1980  
 ctactgtgc tactacgact gcagggaaat acctcgagat cgactacatc aacaacgaca 2040  
 ttgccttttc aacctcgtgg accacgggct ccaattctcg gaatctgaca atctcgtca 2100  
 acggcggacc accgacgagg atcgaggtgc ctctttctgg gaagcacagc gagctcttcg 2160  
 gaccaggccg tgctggtgg gatagcgcga cctttggcgt tttggtgat ggggtgagga 2220  
 atggagaaaa cacggtggtg atagggaaac agggaggtga cgagggagta cagccgatg 2280  
 gtgcggactt tgtgggcttg cgactgtatg attga 2315

<210> 20

5 <211> 2049

<212> ADN

<213> Rasamsonia emersonii

10

<220>

<221> fuente

15 <222> (1)..(2049)

<223> /organismo="Rasamsonia emersonii" /tipo\_molar="ADN no asignado"

<400> 20

ES 2 802 807 T3

atgacaagat atcctagtct tagctactat gcaactctctt catatcttact atttaccagc 60  
tgtggtgcac tggcatccac aagggttgaa cctaggttag acaatggact ggcaatcaca 120  
cctccaatgg gatggaactc ctacaaccac tactcatgct ccccgaaacga atcgatcatc 180  
cagtcaaacg cgaaggctct cgtggacttt ggccttgata cccttggtta tcgatatgtc 240  
acaactgatt gcggctggac ggtcccctac cgactcccca atgggtcctt gacctggaat 300  
gaaacgctgt ttccgagtgg attcccggca ctagggcagt acattcacga tctggggctt 360  
ctctttggcg tctatcaaga tgcggggatt aagacttggt gaggaccccc tgatcaagta 420  
ggaagtttat tccatgaaca acaggacgca gagacgttgc ctgcctggaa agcggatgcg 480  
ttgaaatagc acaactgcta ttctgatgca gccgctgggt atccagatgc ggattataca 540  
ccgagcacct cggcatcgtt ccgctacgcg aacatgacaa aggcgctggc cagtgtggat 600  
cggaaaatcc ttttccaaat ctgtgactgg ggcgtcgatt tccccgcgct ctgggctccc 660  
tctctgggca acacctggcg aatttccaac gacataatcc ccgctggcg gacgatcttc 720  
aggatcctca accaggccgt gccgcagact tcttttgcg gacctggcca ctggccagac 780  
ctcgatatgc tcgagggtgg gaacaatgtg tttacgatcc cggaaagaca gacgcacttt 840  
tccctttggg cgatcctgaa aagccccctg accatcgggt ccgctgtgaa tgataccttg 900  
acgacgatca gagatgcctc gctccagatc ttgaaacaga aggatgtcat cagctacaac 960  
caagaccgcg ttggtgtgag tgccagtctg aagaggagat ggaccgaaga gggctttgaa 1020  
gtgtggagcg ggccatatac ggggtgggaag acagtggctg cattgatcaa ctgggcgat 1080  
gaatcgagga acttgacgct ggatctccct gtctgtaggc tacaacacgc gcagacgctg 1140  
cggaaacatct gggacgaaag cgcagccacg aacgtccgta cgtctacac tgcgaacgtg 1200  
gcagctcatg gaacgatgtt ggtagaactg gccgggacaa cagaagcagg gaagtatcct 1260  
gctgatattt ttgcaacttc agatgggac tcgacaactt ttgaaaacat ctatgccgag 1320  
accactagca gccagtacca gttgacaatc gcgtttgctc ccggatcatc tcaactcgagc 1380  
gaaataacga tccgaacgag cctcggggga ttcaagacgt ctgccaaagg gcaaccgact 1440  
gagagccaaa tctcgggtcaa catctcgtct tcagcagggt catcgaacac gatcacgatt 1500  
tcgccgtctc caattatcag ttacataaac gtcacctccc catcggggac atactatccc 1560  
tgtacatcgt ttacacctgt gggatccgcc aaacccgaga cttgcgatgc gggattctgc 1620  
ctgcctgtcg ggtccaaaat cggctatata agcccagac gtaatgcaag catcactatc 1680  
cctgccacgg tgggtgtctg cagtggcgat aataactact ctgctactac gactgcaggg 1740  
aaatacctcg agatcgacta catcaacaac gacattgcct tttcaacctc gtggaccacg 1800  
ggctccaatt ctcggaatct gacaatctcg gtcaacggcg gaccaccgac gaggatcgag 1860  
gtgcctcttt ctgggaagca cagcgagctc ttoggaccag gccgtggctg gtgggatagc 1920  
gcgacctttg cgtttttggt tgatgggtgg aggaatggag aaaacacggt ggtgataggg 1980  
aaccagggag gtgacgaggg agtacagccg tatggtgctg actttgtggg cttgcgactg 2040  
tatgattga 2049

5 <210> 21

<211> 1821

ES 2 802 807 T3

<212> ADN

<213> Rasamsonia emersonii

5 <220>

<221> fuente

<222> (1)..(1821)

10

<223> /organismo="Rasamsonia emersonii" /tipo\_molar="ADN no asignado"

<400> 21

atgatctccg	ctgctgttct	ggtgtctttg	ctactcacca	ccatccagcc	cggtgcctac	60
gcccagtggc	agtgcggtgg	ttccacctac	actcccggct	ccgtcaactt	caccgaggag	120
tgctaccagg	ctgctcagga	ctgctgttct	cagttcgggtg	cgaacgcctc	tctggtcaac	180
tgccaggatg	ctgctggtaa	cctcttcatg	cagcagcagt	ccaaccaggg	caagaacaac	240
gactacttga	ttgcctacca	ggatatacctg	gacttctgct	tgctcgatgg	attcaccact	300
ggtacctggt	acgatgactc	tcagtgttac	tggatggctg	ctgagcctgg	ctgctactct	360
cccaacggct	ccatcggtac	caccggctccc	ggtttctgcg	tgcagaaccg	cgatgacacc	420
gtcctcaacg	gctgctaccc	ccagccccag	tccggtgctg	gtcctctcca	ggtcctccgc	480
actgctcgca	ctgccaacgg	cttcacctcg	tctgcccgtg	gctggaacag	ctggggaatc	540
caggctctgg	agaacccccag	caccatcccc	ggctggactg	tgttcaacca	gactgctgtc	600
aagcagcagt	gctccgtcct	tgctcgtctc	gacttcaagg	ctgctggcta	cgacctttgc	660
tctcttgatg	ctggctggtc	gacctcttcc	gaggttgatg	agtacggccg	tatcctgtac	720
aacagcacc	tcttcgacct	tctgagctg	gccgactacc	tccacggta	gggtctgaag	780
ctcgggtgtct	acgtcatccc	cggtgttctc	tgcgttgccg	ccaacaagac	cattgaggg	840
accaacatcc	gcattggtga	tgtcctgaac	ggcaacaacg	acgagctctc	ctactgcgac	900
tgggacttct	ccaaggatgg	tgtccagcag	tggcacgact	cgtgatcaa	cctctgggcc	960
agctggggtg	ttgacatgat	caagctcgac	ttcgtcactc	ctggaagccc	tcagaacggt	1020
gccaacttgg	tgtgcaacaa	ctctgctgcc	gttgaggcct	accacaaggc	cattgccaac	1080
tctggccgtc	agatccgtct	ggacctcagc	tggaagctct	gccgtaacga	aacctacctg	1140
cccgtctggt	ccagccttgc	cgactccatc	cgcactgacc	aggacattga	taactacggc	1200
tacaacacct	tcgtcgcctg	gcaggtcgtc	cagcgtgcca	tcgagaacta	ccgtcagtac	1260
15 atcctgctgc	agaagcagcg	caacgttccc	atcaccctgt	accccgacat	ggacaacctg	1320

ES 2 802 807 T3

ttcgtcggca acgccgcaa cctcacgggt gtctccgatg cccagcgtat caccatcatg 1380  
aaccactggc ttggtgctgc tgccaacctg atcattggct ccgacctcaa ccagctcgat 1440  
gaccttggtta ccaagctcct cacctccaac gagtccatcc aggctgccga cttcttcgct 1500  
cagtacccca tgcagccccc caacctgggt actggtgaca acgtcccccga gcagctccag 1560  
gcctggattg ctggcccctc cgacaacaac gaagcctacg tgcttgttgt caactacggc 1620  
cctgaccagg gccacgggtg tttcgggtacc agcttgactg gtgtgcagaa cgtcaccggt 1680  
tctcttgccg atcttggtat cgctggcaag agctgggaat tctccgatgt ctggaacgga 1740  
aactcctcca ccgtcagcac ctcgttcact gcctacctgg atgaggggtga gtcccagctc 1800  
ctgcacctca ccgttgcgca g 1821

<210> 22

5 <211> 586

<212> PRT

<213> *Rasamsonia emersonii*

10

<400> 22

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

ES 2 802 807 T3

Asn Gly Phe Thr Ser Ser Ala Arg Gly Trp Asn Ser Trp Gly Ile Gln  
 145 150 155 160  
 Ala Leu Glu Asn Pro Ser Thr Ile Pro Gly Trp Thr Val Phe Asn Gln  
 165 170 175  
 Thr Ala Val Lys Gln Gln Cys Ser Val Leu Ala Arg Ser Asp Phe Lys  
 180 185 190  
 Ala Ala Gly Tyr Asp Leu Cys Ser Leu Asp Ala Gly Trp Ser Thr Ser  
 195 200 205  
 Ser Glu Val Asp Glu Tyr Gly Arg Ile Leu Tyr Asn Ser Thr Leu Phe  
 210 215 220  
 Asp Leu Pro Glu Leu Ala Asp Tyr Leu His Gly Gln Gly Leu Lys Leu  
 225 230 235 240  
 Gly Val Tyr Val Ile Pro Gly Val Pro Cys Val Ala Ala Asn Lys Thr  
 245 250 255  
 Ile Glu Gly Thr Asn Ile Arg Ile Gly Asp Val Leu Asn Gly Asn Asn  
 260 265 270  
 Asp Glu Leu Ser Tyr Cys Asp Trp Asp Phe Ser Lys Asp Gly Val Gln  
 275 280 285  
 Gln Trp His Asp Ser Leu Ile Asn Leu Trp Ala Ser Trp Gly Val Asp  
 290 295 300  
 Met Ile Lys Leu Asp Phe Val Thr Pro Gly Ser Pro Gln Asn Gly Ala  
 305 310 315 320  
 Asn Leu Val Cys Asn Asn Ser Ala Ala Val Glu Ala Tyr His Lys Ala  
 325 330 335  
 Ile Ala Asn Ser Gly Arg Gln Ile Arg Leu Asp Leu Ser Trp Lys Leu  
 340 345 350  
 Cys Arg Asn Glu Thr Tyr Leu Pro Val Trp Ser Ser Leu Ala Asp Ser  
 355 360 365  
 Ile Arg Thr Asp Gln Asp Ile Asp Asn Tyr Gly Tyr Asn Thr Phe Val  
 370 375 380  
 Ala Trp Gln Val Val Gln Arg Ala Ile Glu Asn Tyr Arg Gln Tyr Ile  
 385 390 395 400

ES 2 802 807 T3

Leu Leu Gln Lys Gln Arg Asn Val Pro Ile Thr Leu Tyr Pro Asp Met  
 405 410 415

Asp Asn Leu Phe Val Gly Asn Ala Ala Asn Leu Thr Gly Val Ser Asp  
 420 425 430

Ala Gln Arg Ile Thr Ile Met Asn His Trp Leu Gly Ala Ala Ala Asn  
 435 440 445

Leu Ile Ile Gly Ser Asp Leu Asn Gln Leu Asp Asp Leu Gly Thr Lys  
 450 455 460

Leu Leu Thr Ser Asn Glu Ser Ile Gln Ala Ala Asp Phe Phe Ala Gln  
 465 470 475 480

Tyr Pro Met Gln Pro Arg Asn Pro Gly Thr Gly Asp Asn Val Pro Gln  
 485 490 495

Gln Leu Gln Ala Trp Ile Ala Gly Pro Ser Asp Asn Asn Glu Ala Tyr  
 500 505 510

Val Leu Val Val Asn Tyr Gly Pro Asp Gln Gly His Gly Gly Phe Gly  
 515 520 525

Thr Ser Leu Thr Gly Val Gln Asn Val Thr Val Ser Leu Ala Asp Leu  
 530 535 540

Gly Ile Ala Gly Lys Ser Trp Glu Phe Ser Asp Val Trp Asn Gly Asn  
 545 550 555 560

Ser Ser Thr Val Ser Thr Ser Phe Thr Ala Tyr Leu Asp Glu Gly Glu  
 565 570 575

Ser Gln Leu Leu His Leu Thr Val Ala Gln  
 580 585

<210> 23

5 <211> 21

<212> PRT

<213> *Rasamsonia emersonii*

10

<400> 23

Met Ile Ser Ala Ala Val Leu Val Ser Leu Leu Leu Thr Thr Ile Gln  
 1 5 10 15

Pro Gly Ala Tyr Ala  
 20

15 <210> 24

<211> 1824

<212> ADN

20

ES 2 802 807 T3

<213> Rasamsonia emersonii

<220>

5 <221> fuente

<222> (1)..(1824)

<223> /organismo="Rasamsonia emersonii" /tipo\_molar="ADN no asignado"

10

<400> 24

```

atgatctccg ccgcagtcct ggtctccctg cttctgacca ccatccagcc cggggcttac      60
gccagtggc aatgtggcgg atcaacgtac acgccaggct cggtgaactt caccgaagag      120
tgctaccagg ctgctcagga ttgcgtcgcg caattcggcg ctaatgccag cctggtcaac      180
tgtcaagacg ccgcaggcaa tctgttcacg cagcagcaat ccaaccaggg gaagaacaac      240
gactacctga ttgcctacca ggacatcctt gacttctgcc tgctcgatgg cttcacgaca      300
gggacgtggt acgacgacag ccagtggtac tggatggccg ctgaaccggg ctgctactct      360
ccaaacggca gcattggcac cactgggccc ggtttctgcg tccagaaccg cgacgacacc      420
gtcctcaatg gttgctatcc ccagccccag tctggtgcag gccccctgca ggtattgcga      480
accgcccgca cagccaacgg gtttacttcg tctgcgcgag gatggaattc ctggggaatt      540
caggcgttg agaatccgag cacgattccc ggctggacgg tattcaacca gacagccgtg      600
aacagcaaat gctcagtgct ggctcgcctg gatttcaagg cagcaggcta cgatctctgc      660
agcctggacg ccggttggtc cagcagcagc gaagtcgacg aatacgggcg aatcctctac      720
aacagcacct tgttcgacct ccccgagctg gcagattacc tgcacggaca aggcctcaag      780
ctgggtgtgt atgtgatccc gggcgtgccc tgcgttgccg caaataagac gatcgaaggg      840
accaacattc gcatcgggga cgtcctcaat ggcaacaacg atgaactgtc gtactgcgac      900
tgggacttct ccaaggacgg cgtgcagcaa tggcacgatt cacttatcaa cctgtgggcc      960
tcctggggcg tggacatgat caagctggac tttgtcacgc caggcagtcc gcaaaacggc     1020
gcgaatctcg tctgcaacaa cagcgcgccg gtggaggcgt accacaaggc aatcgccaat     1080
tcggggcgcc agatccgact ggatctgtcg tggaaactgt gccgcaatga aacctacctg     1140
ccggtctgga gcagtctggc ggattcgatc cgaacggacc aggacatcga caattatgga     1200
tacaacacgt tcgtcgcgat gcaggtcgtg cagcgggcca tcgagaatta ccgccagtat     1260
atcctgttgc agaagcagcg caacgtgccc atcacgctgt acccggacat ggacaatctc     1320
tttgtgggca atgcagcgaa cctgacgggg gtcagcgacg cgcaacgcat caccatcatg     1380
aaccactggc tcggggcgcc agcgaatctg atcataggct cggatttgaa ccagctcgac     1440
gatctgggca cgaagctgct gacgagcaat gagtccatcc aggctgccga tttctttgcg     1500

```

ES 2 802 807 T3

cagtatccga tgcaaccgcg caaccgggc acgggcgata acgtcccgca acagctgcaa 1560  
gCGTGGATCG cCGGACCGTC cGACAACAAC gaagcgtatg tcctcgtcgt caactacgga 1620  
cCGGATCAGG gGCACGGCGG gTTTGGGACG tcgTtgactg gCGTCCAGAA tGTGACCGTG 1680  
tcgctggcgg atctgggcat cgcaggcaag agctgggagt tctctgatgt ctggaatggg 1740  
aactcgagca cggTgtcgac ttcttttaca gcctatctag atgaagggga gtcgcagctc 1800  
ctgcatttga cCGTGGCACA ataa 1824

<210> 25

5 <211> 1824

<212> ADN

10 <213> Rasamsonia emersonii

<220>

<221> fuente

15 <222> (1)..(1824)

<223> /organismo="Rasamsonia emersonii" /tipo\_molar="ADN no asignado"

20 <400> 25

atgatctccg cgcagtcct ggtctccctg cttctgacca ccatccagcc cggggcttac 60  
gccagtggc aatgtggcgg atcaacgtac acgccaggct cggTgaactt caccgaagag 120  
tgctaccagg ctgctcagga ttgcgtcgcg caattcggcg ctaatgccag cctggTcaac 180  
tgtcaagacg cCGCAGGCAA tctgttcAtg cagcagcaat ccaaccaggg gaagaacaac 240  
gactacctga ttgcctacca ggacatcctt gacttctgcc tgctcgatgg cttcacgaca 300  
gggacgtggt acgacgacag ccagtggtac tggatggccg ctgaaccggg ctgctactct 360  
ccaaacggca gcattggcac cactgggccc ggtttctgcg tccagaaccg cgacgacacc 420  
gtcctcaatg gttgctatcc ccagccccag tctggtgcag gccccctgca ggtattgcga 480  
accgcccgca cagccaacgg gtttactctg tctgcgcgag gatggaattc ctggggaatt 540  
caggcgttg agaatccgag cagcattccc ggctggacgg tattcaacca gacagccgtg 600  
aaacagcaat gctcagtgct ggctcgtcgc gatttcaagg cagcaggcta cgatctctgc 660  
agcctggacg cCGTtggTc cagcagcagc gaagtcgacg aatacgggcg aatcctctac 720  
aacagcacct tGttGacCct ccccgagctg gcagattacc tgcacggaca aggcctcaag 780  
ctgggtgtgt atgtgatccc gggcgtgccc tgcgttgccg caaataagac gatcgaaggg 840  
accaacattc gcacgggga cgtcctcaat ggcaacaacg atgaactgtc gtactgcgac 900  
tgggacttct ccaaggacgg cgtgcagcaa tggcagcatt cacttatcaa cctgtgggcc 960  
tcctggggcg tggacatgat caagctggac tttgtcacgc caggcagtcc gcaaaacggc 1020  
gcgaatctcg tctgcaacaa cagcgcgca gtggaggcgt accacaaggc aatcgccaat 1080  
tcggggcgcc agatccgact ggatctgtcg tggaaactgt gccgcaatga aacctacctg 1140

ES 2 802 807 T3

ccggtctgga gcagtctggc ggattcgcgc cgaacggacc aggacatcga caattatgga 1200  
 tacaacacgt tcgtcgcgatg gcaggtcgtg cagcgggcca tcgagaatta ccgccagtat 1260  
 atcctgttgc agaagcagcg caacgtgccc atcacgctgt acccggacat ggacaatctc 1320  
 tttgtgggca atgcagcgaa cctgacgggg gtcagcgagc cgcaacgcat caccatcatg 1380  
 aaccactggc tcggggcggc agcgaatctg atcataggct cggatttgaa ccagctcgac 1440  
 gatctgggca cgaagctgct gacgagcaat gaggccatcc aggtgcccga tttctttgcg 1500  
 cagtatccga tgcaaccgcg caaccggggc acgggcgata acgtcccga acagctgcaa 1560  
 gcgtggatcg ccggaccgtc cgacaacaac gaagcgtatg tcctcgtcgt caactacgga 1620  
 ccggtcagg ggcacggcgg gtttgggagc tcgttgactg gcgtccagaa tgtgaccgtg 1680  
 tcgctggcgg atctgggcat cgcaggcaag agctgggagt tctctgatgt ctggaatggg 1740  
 aactcgagca cgtgtcgcac ttcttttaca gcctatctag atgaagggga gtcgcagctc 1800  
 ctgcatttga ccgtggcaca ataa 1824

<210> 26

5 <211> 756

<212> ADN

<213> *Rasamsonia emersonii*

10 <220>

<221> fuente

15 <222> (1)..(756)

<223> /organismo="Rasamsonia emersonii" /tipo\_molar="ADN no asignado"

<400> 26

20 atgaagtctg tcctggctgt gaccctcagc ctttctcttg cctccctgac cactgctgct 60  
 cctactccca ccaagaccct ggagaagcgt gccgacttct gcggccagtg ggattcgcac 120  
 gtcactggca gctacaccgt ttacaacaac ctctggggtg agagctctgc ctccctccggc 180  
 agccagtgca ccggtgttga ctgctggaac ggcaacaccc tggcctggca caccagctgg 240  
 acctgggagg gtggtgcctc ttccgtcaag tcgttcgcca acgctgccta cgctttcact 300  
 gccactcagc tctcccagat ctccctccatc cccaccacct ggcagtggag ctacactggt 360  
 acctccatcg atgcgatgt tgcctacgac ctcttcacct cctcctcccc cagcggcagc 420  
 aacgaatacg aagtcgatg ctggctagct gctcttggtg gtgctttccc tatactctcc 480  
 accggtacca cccccattgc cactcctacc attggtggta tctcctggaa cctgtactct 540  
 ggccccaacg gtgccaccac cgtctactct ttctctgctt cccgcgagca gaccgacttc 600  
 tccggtgaca tcaacgactt cctgacctac ctccagcaga acgagggctt ttctctctcc 660  
 cagtacctcc tctccatcca ggctggcacc gagcccttca ctggtgagaa cgctgagctc 720  
 aagaccactg cctactctgt ctccgtcagc acttca 756

<210> 27

ES 2 802 807 T3

<211> 233

<212> PRT

5

<213> *Rasamsonia emersonii*

<400> 27

Ala Pro Thr Pro Thr Lys Thr Leu Glu Lys Arg Ala Asp Phe Cys Gly  
1 5 10 15

Gln Trp Asp Ser Thr Val Thr Gly Ser Tyr Thr Val Tyr Asn Asn Leu  
20 25 30

Trp Gly Glu Ser Ser Ala Ser Ser Gly Ser Gln Cys Thr Gly Val Asp  
35 40 45

Ser Leu Asn Gly Asn Thr Leu Ala Trp His Thr Ser Trp Thr Trp Glu  
50 55 60

Gly Gly Ala Ser Ser Val Lys Ser Phe Ala Asn Ala Ala Tyr Ala Phe  
65 70 75 80

Thr Ala Thr Gln Leu Ser Gln Ile Ser Ser Ile Pro Thr Thr Trp Gln  
85 90 95

Trp Ser Tyr Thr Gly Thr Ser Ile Asp Ala Asp Val Ala Tyr Asp Leu  
100 105 110

Phe Thr Ser Ser Ser Pro Ser Gly Ser Asn Glu Tyr Glu Val Met Ile  
115 120 125

Trp Leu Ala Ala Leu Gly Gly Ala Phe Pro Ile Ser Ser Thr Gly Thr  
130 135 140

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

Ser Gly Pro Asn Gly Ala Thr Thr Val Tyr Ser Phe Val Ala Ser Arg  
165 170 175

Glu Gln Thr Asp Phe Ser Gly Asp Ile Asn Asp Phe Leu Thr Tyr Leu  
180 185 190

Gln Gln Asn Glu Gly Leu Ser Ser Ser Gln Tyr Leu Leu Ser Ile Gln  
195 200 205

Ala Gly Thr Glu Pro Phe Thr Gly Glu Asn Ala Glu Leu Lys Thr Thr  
210 215 220

10

Ala Tyr Ser Val Ser Val Ser Thr Ser  
225 230

<210> 28

15

<211> 19

ES 2 802 807 T3

<212> PRT

<213> Rasamsonia emersonii

5 <400> 28

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

Thr Thr Ala

<210> 29

10

<211> 819

<212> ADN

15 <213> Rasamsonia emersonii

<220>

<221> fuente

20

<222> (1)..(819)

<223> /organismo="Rasamsonia emersonii" /tipo\_molar="ADN no asignado"

25 <400> 29

```

atgaagtctt tcctcgccgt aactctgtcc ctatccctcg caagcctcac caccgccgcc      60
ccaaccccaa ccaaaaccct cgaaaagcgg gccgacttct gcggacaatg ggactcgacc      120
gtcacaggca gctacaccgt ctacaacaac ctctgggggg aatccagcgc gtcgtccgga      180
tcgcagtgca cgggcgctga ctcgctcaac gggaacacgc tcgcctggca cacctcgtgg      240
acctgggagg gcggggcgct cagtgtcaag agcttcgcca atgccgcgta tgcgttcacg      300
gcgacgcagc tgagccagat tagcagtatt ccgacgacgt ggcaagtggag gtatgaatga      360
cacttccttt tctgcctctc tgtctatgta cttgctgctg acacgtacag ctacaccggc      420
acctccatcg acgcccagct cgcgtacgac ctattcacga gttccagccc cagcggcagc      480
aacgagtacg aggttatgat ctggctggcc gcgctcgggg gcgcattccc catctcgtcg      540
acgggcacga cgcccatcgc gacgccacc atcggcggca tctcatgaa tctgtacagc      600
gggccaacg gcgcgacgac ggtgtacagc ttcgtagcgt cgagagagca gacggacttt      660
tcgggcgaca tcaacgattt tctgacctat ctccagcaga acgagggcct cagctcgagc      720
cagtacctgc tgtccatcca ggccggcacg gagccgttta cgggggagaa tgcggagctg      780
aagacgacgg cgtatagtgt cagtgtatca acctcgtga      819

```

<210> 30

30

<211> 759

<212> ADN

35 <213> Rasamsonia emersonii

<220>

<221> fuente

ES 2 802 807 T3

<222> (1)..(759)

5

<223> /organismo="Rasamsonia emersonii" /tipo\_molar="ADN no asignado"

<400> 30

atgaagtctt	tcctcgccgt	aactctgtcc	ctatccctcg	caagcctcac	caccgccgcc	60
ccaaccccaa	ccaaaaccct	cgaaaagcgg	gccgacttct	gctgacaatg	ggactcgacc	120
gtcacaggca	gctacaccgt	ctacaacaac	ctctgggggg	aatccagcgc	gtcgtccgga	180
tcgcagtgca	cgggcgctga	ctcgctcaac	gggaacacgc	tcgcctggca	cacctcgtgg	240
acctgggagg	gcggggcgtc	cagtgtcaag	agcttcgcca	atgccgcgta	tgcgttcacg	300
gcgagcgagc	tgagccagat	tagcagtatt	ccgacgacgt	ggcagtgag	ctacaccggc	360
acctccatcg	acgccgacgt	cgcgctacgac	ctattcacga	gttccagccc	cagcggcagc	420
aacgagtacg	aggttatgat	ctggctggcc	gcgctcgggg	gcgcattccc	catctcgtcg	480
acgggcaoga	cgcccatcgc	gacgccacc	atcggcggca	tctcatgaa	tctgtacagc	540
gggcccacag	gcgcgacgac	ggtgtacagc	ttcgtagcgt	cgagagagca	gacggacttt	600
tcggggcaca	tcaacgattt	tctgacctat	ctccagcaga	acgagggcct	cagctcgagc	660
cagtacctgc	tgtccatcca	ggcgggcacg	gagccgttta	cgggggagaa	tgccggagctg	720
aagacgacgg	cgtatagtgt	cagtgtatca	acctcgtga			759

10

<210> 31

<211> 1923

<212> ADN

15

<213> Rasamsonia emersonii

<220>

20

<221> fuente

<222> (1)..(1923)

25

<223> /organismo="Rasamsonia emersonii" /tipo\_molar="ADN no asignado"

<400> 31

atgtccctcg	ccagcaagaa	cctgcttgct	tgccgcctct	tcctggctgt	tgctcagctt	60
gctcagcccg	ttactggcct	ggaattgact	gtcagcacct	ccggtggcaa	caagtcctcc	120
cccatcctct	acggcttcat	gttcgaagat	atcaaccact	ccggtgatgg	tggtatccac	180
tcccagctcc	tccagaacaa	cggattccag	ggtgccaacc	ccggtctgac	tgccctacgcc	240
cccatcggta	acgtttccct	cgccggtgac	actgagaacc	ctctctcctc	tgccatcacc	300
cgctcgtgta	aggtcaccgt	tcttcccaac	gcctctggca	aggtcggttt	ctccaactcc	360
ggctactggg	ggatacccg	caacgaggcc	aactacacca	actcgttcta	catcaagggt	420

ES 2 802 807 T3

gactacaacg gtgatgtcac cctccgcctc gccggtgcct cctccggctc tctgtacgcc 480  
 tcccagacca tctccgtcaa gtccaacagc agcagctaca cctacgtcga gactcagtac 540  
 cccgtcaagc aggctcctga cggcaacaac atctgggagc tcacctcga tgcctccaag 600  
 gttgctggca agtccctcaa cttcgacctc atccagctgt accctcctac cttccacgac 660  
 cgtcccaacg gtctgaagcc cgagattgcc aacgtccttg ttgacatgaa gggctctttc 720  
 ctccgtttcc ccggtggtaa caacctcgag ggtctgaccc ccgccaccgg ttggaagtgg 780  
 aacgagacta ttggacctct gacctcccgt cctggctgcc agggtaactg gggctacccc 840  
 aacaccgatg ctcttggtct gaacgaatac ttctctgga tccaggacat gggcttgact 900  
 gctgtcctgg gtgtctggga tggcttgacc attggtggtc cttctcccca ggtcatcacc 960  
 ggtgatgctc tcaagcccta cattgatgat gttctcaacg agctggaata catcctgggt 1020  
 gacaaggaca ccacctacgg caagctccgt gcctctcagc gattccccga cccttggcc 1080  
 ctccagtacg tggagattgg caacgaggac aacctgaacg gtggtggacc tagctacgct 1140  
 gagcgttca ctgctttcca cgacgccatc aaggccaagt acccccacct gaccatcatc 1200  
 tccagcactg accagtacct tcccacccc aagcccaagg gtatctggat ggactggcac 1260  
 acctactctc ctgctgggta attggtgagc atgtttaaca agttcgacaa cgtcgaccgc 1320  
 gctttcccct acttogtcgg tgaatacgcc tgcatttctc ttgacaacgg cactgagctt 1380  
 gccagccca tcatgcaggg ctctggtgct gaggtgtgt tcatgatcgg tatggagcgc 1440  
 aactcggatg ttgtcaagat ggccgcctac gctcccttgc tccagcagta cggcggctac 1500  
 acctggaacg accacgtccc caacctgatc atgttcaaca acaaccccaa cggcattgtc 1560  
 cgctccacct cgtactacgt gcagcagatg ttctccgtga accgtggaag caccatcctt 1620  
 cctgttgagt ccgacggctc tttcggctct gtctactacg tggcctccaa gacctccgac 1680  
 aacaagacct actacgtcaa gttcgccaac tacggtccta ctcccagcc cgtcacccgtc 1740  
 aacatcccca agaccgagac tggtcgtctg accaccctct ctggtgctgc cactgctgcc 1800  
 aacacctggt ccagcccaa ccgctcact cctaagacct cgaacgtcag ctccgccgat 1860  
 ggtgctcctg gcaagttcac catcacctc cccgcctggt ccggtgctgt ccttgctgtc 1920  
 gag 1923

<210> 32

5 <211> 615

<212> PRT

<213> *Rasamsonia emersonii*

10

<400> 32

Leu Glu Leu Thr Val Ser Thr Ser Gly Gly Asn Lys Ser Ser Pro Ile  
 1 5 10 15

ES 2 802 807 T3

Leu Tyr Gly Phe Met Phe Glu Asp Ile Asn His Ser Gly Asp Gly Gly  
 20 25 30  
 Ile His Ser Gln Leu Leu Gln Asn Asn Gly Phe Gln Gly Ala Asn Pro  
 35 40 45  
 Gly Leu Thr Ala Tyr Ala Pro Ile Gly Asn Val Ser Leu Ala Val Asp  
 50 55 60  
 Thr Glu Asn Pro Leu Ser Ser Ala Ile Thr Arg Ser Leu Lys Val Thr  
 65 70 75 80  
 Val Pro Pro Asn Ala Ser Gly Lys Val Gly Phe Ser Asn Ser Gly Tyr  
 85 90 95  
 Trp Gly Ile Pro Val Asn Glu Ala Asn Tyr Thr Asn Ser Phe Tyr Ile  
 100 105 110  
 Lys Gly Asp Tyr Asn Gly Asp Val Thr Leu Arg Leu Ala Gly Ala Ser  
 115 120 125  
 Ser Gly Ser Leu Tyr Ala Ser Gln Thr Ile Ser Val Lys Ser Asn Ser  
 130 135 140  
 Ser Ser Tyr Thr Tyr Val Glu Thr Gln Tyr Pro Val Lys Gln Ala Pro  
 145 150 155 160  
 Asp Gly Asn Asn Ile Trp Glu Leu Thr Phe Asp Ala Ser Lys Val Ala  
 165 170 175  
 Gly Lys Ser Leu Asn Phe Asp Leu Ile Gln Leu Tyr Pro Pro Thr Phe  
 180 185 190  
 His Asp Arg Pro Asn Gly Leu Lys Pro Glu Ile Ala Asn Val Leu Val  
 195 200 205  
 Asp Met Lys Gly Ser Phe Leu Arg Phe Pro Gly Gly Asn Asn Leu Glu  
 210 215 220  
 Gly Leu Thr Pro Ala Thr Arg Trp Lys Trp Asn Glu Thr Ile Gly Pro  
 225 230 235 240  
 Leu Thr Ser Arg Pro Gly Arg Gln Gly Asn Trp Gly Tyr Pro Asn Thr  
 245 250 255  
 Asp Ala Leu Gly Leu Asn Glu Tyr Phe Leu Trp Ile Gln Asp Met Gly  
 260 265 270

ES 2 802 807 T3

Leu Thr Ala Val Leu Gly Val Trp Asp Gly Leu Thr Ile Gly Gly Pro  
 275 280 285  
 Ser Pro Gln Val Ile Thr Gly Asp Ala Leu Lys Pro Tyr Ile Asp Asp  
 290 295 300  
 Val Leu Asn Glu Leu Glu Tyr Ile Leu Gly Asp Lys Asp Thr Thr Tyr  
 305 310 315 320  
 Gly Lys Leu Arg Ala Ser His Gly Phe Pro Asp Pro Trp Pro Leu Gln  
 325 330 335  
 Tyr Val Glu Ile Gly Asn Glu Asp Asn Leu Asn Gly Gly Gly Pro Ser  
 340 345 350  
 Tyr Ala Glu Arg Phe Thr Ala Phe His Asp Ala Ile Lys Ala Lys Tyr  
 355 360 365  
 Pro His Leu Thr Ile Ile Ser Ser Thr Asp Gln Tyr Leu Pro Asn Pro  
 370 375 380  
 Lys Pro Lys Gly Ile Trp Met Asp Trp His Thr Tyr Ser Pro Ala Gly  
 385 390 395 400  
 Glu Leu Val Ser Met Phe Asn Lys Phe Asp Asn Val Asp Arg Ala Phe  
 405 410 415  
 Pro Tyr Phe Val Gly Glu Tyr Ala Cys Ile Ser Leu Asp Asn Gly Thr  
 420 425 430  
 Glu Leu Ala Gln Pro Ile Met Gln Gly Ser Val Ala Glu Ala Val Phe  
 435 440 445  
 Met Ile Gly Met Glu Arg Asn Ser Asp Val Val Lys Met Ala Ala Tyr  
 450 455 460  
 Ala Pro Leu Leu Gln Gln Tyr Gly Gly Tyr Thr Trp Asn Asp His Val  
 465 470 475 480  
 Pro Asn Leu Ile Met Phe Asn Asn Asn Pro Asn Gly Ile Val Arg Ser  
 485 490 495  
 Thr Ser Tyr Tyr Val Gln Gln Met Phe Ser Val Asn Arg Gly Ser Thr  
 500 505 510  
 Ile Leu Pro Val Glu Ser Asp Gly Ser Phe Gly Pro Val Tyr Tyr Val

ES 2 802 807 T3

515

520

525

Ala Ser Lys Thr Ser Asp Asn Lys Thr Tyr Tyr Val Lys Phe Ala Asn  
530 535 540

Tyr Gly Pro Thr Pro Gln Pro Val Thr Val Asn Ile Pro Lys Thr Glu  
545 550 555 560

Thr Gly Arg Leu Thr Thr Leu Ser Gly Ala Ala Thr Ala Ala Asn Thr  
565 570 575

Trp Ser Ser Pro Asn Ala Val Thr Pro Lys Thr Ser Asn Val Ser Ser  
580 585 590

Ala Asp Gly Ala Pro Gly Lys Phe Thr Ile Thr Leu Pro Ala Trp Ser  
595 600 605

Val Ala Val Leu Ala Val Glu  
610 615

<210> 33

5 <211> 26

<212> PRT

<213> Rasamsonia emersonii

10

<400> 33

Met Ser Leu Ala Ser Lys Asn Leu Leu Ala Cys Gly Leu Phe Leu Ala  
1 5 10 15

Val Ala Gln Leu Ala Gln Pro Val Thr Gly  
20 25

15 <210> 34

<211> 2381

<212> ADN

20

<213> Rasamsonia emersonii

<220>

25 <221> fuente

<222> (1)..(2381)

<223> /organismo="Rasamsonia emersonii" /tipo\_molar="ADN no asignado"

30

<400> 34

ES 2 802 807 T3

atgtcgttgg cttcgaagaa tctcctggcc tgcgggcttt tcttagcagt cgcccagctt 60  
 gcacagccag tgactggttt agaacttacg gtgtcgacgt ctggcggcaa taaatccagc 120  
 cctatcttgt atgggttcat gttcagagtg agtttttgcc aagctggctg tgaaccgcc 180  
 tgctggtgga gagttacgat aaccatgtcc acgctcacta tagcgtgatt gattatgtgc 240  
 tgacacctgc aggatatcaa ccaactctggt gatggaggga tccacagtca gcttctgcag 300  
 aacaacggct tccagggtgc aaacccggc ttgactgcat acgctccaat cgggaatgtg 360  
 tcgttggccg tggatacggg aaaccctctc tcgagcgcga taacgcggtc tctgaaggtc 420  
 actgtcccgc ccaacgcctc cggaagggtt ggcttctcga attcgggtta ctgggttatt 480  
 cccgtgaacg aggccaacta caccaacagc ttctacatca agggagacta taatggcgac 540  
 gtcacgcttc gccttgccgg ggccagcagc ggcagtctgt atgcatcgca aacgatcagc 600  
 gtcaagagca actcgtccag ctacacctat gtcgagacgc aatatccagt gaagcaggct 660  
 cctgacggca acaacatctg ggaactgacc tttgacgcgt ccaagggtggc gggaaagtca 720  
 ttgaattttg acctgatcca gctctacca ccaacattcc acgacagggtg cgattgcttg 780  
 tattcctcca ggattgtgtc gctgatgctg acctatttta tagacctaat ggcctcaaac 840  
 cggaaattgc aaatgttttg gttgacatga agggttcttt cctgcgcttt cccggaggaa 900  
 acaatctgtg agttgcttcg gtctgtatct ggcgatcctt gctgactttt ggccagcgaa 960  
 ggattaactc cggcaacgcg ctggaagtgg aatgagacca ttgggcctct gacgagccgt 1020  
 cctggcgggc aaggtaactg ggggtatccc aacacagatg cacttggtag gctttattta 1080  
 aaagtgatct gtgtatgaag cattgctaaa tgcgatgaca ggtctgaatg aatatttctt 1140  
 gtggatccag gacatgggtt tgacggcagt tctcgggtgc tgggatggac tgactatcgg 1200  
 agggccgagc cctcaagtta tcaactggga cgctttgaag ccctacatcg acgacgtttt 1260  
 gaacgagctt gaggtaagtg cctctggatg atctttttgt tcggtttcgc taacaattgc 1320  
 agtatatcct tggagacaag gataccacgt acggcaagct tcgtgcttct catggcttcc 1380  
 ctgaccctg gccactcaa tacgtggaaa tcggcaacga ggacaacctc aatggaggcg 1440  
 ggccctcata tgcaagagcg ttcaacggcct tccacgacgc aatcaaggcc aagtaccgc 1500  
 atctgaccat cattagcagc acggaccagt acctcccaa tcccaagcca aaaggaatct 1560  
 ggatggactg gcacacgtac tcgccggcgg gggagcttgt gagcatgttc aacaagtctg 1620  
 acaacgtgga ccgggctttc ccgtactttg tcggcgcaata tgcctgcatc tcctcgata 1680  
 acgggaccga gctcgcgcag ccgatcatgc aggggtctgt ggcggaagct gtctttatga 1740  
 ttggcatgga aaggaacagc gatgttgtca agatggctgc ctatgcaccg ctgctgcagc 1800  
 aatacggcgg gtacacatgg aatgtgagat gcgcaaacga aaaatcaaga agactgatat 1860  
 atgcagactg actgactgac atcaggacca cgtgccaaac ctgattatgt tcaacaacia 1920  
 cccaaacggc attgtccgga gcaccagcta ctacgtacag cagatgttct cagtgaaccg 1980  
 gggcagcacc attctgccag tggaatcaga cgggtcattt ggaccaggta cgtacagacc 2040  
 aacaaacacc agccacagcc agggcagcag gtccgcctcg acttactaat ttcaatgaca 2100  
 atgcagtcta ctacgtcgc tccaaaacct cagacaacia gacatactac gtcaagtctg 2160

ES 2 802 807 T3

ccaactacgg ccctaccccc cagcccgta ccgtcaacat cccaagaca gaaaccgggc 2220  
gactgacgac gctgtccggc gccgctactg cagccaacac ctggagcagc cctaattgctg 2280  
tcacacccaa gaccagcaat gtcagctctg ctgatggtgc gccaggaaag ttactatta 2340  
ctctgcctgc gtggagtgtt gctgtgttgg ctgtggagtg a 2381

<210> 35

5 <211> 1926

<212> ADN

10 <213> Rasamsonia emersonii

<220>

<221> fuente

15 <222> (1)..(1926)

<223> /organismo="Rasamsonia emersonii" /tipo\_molar="ADN no asignado"

20 <400> 35

atgtcgttgg cttcgaagaa tctcctggcc tgcgggcttt tcttagcagt cgcccagctt 60  
gcacagccag tgactggttt agaacttacg gtgtcgacgt ctggcggcaa taaatccagc 120  
cctatcttgt atgggttcat gttcgaggat atcaaccact ctggtgatgg agggatccac 180  
agtcagcttc tgcagaacaa cggcttccag ggtgcaaacc cgggcttgac tgcatacgtc 240  
ccaatcggga atgtgtcgtt ggccgtgat acgaaaacc ctctctcgag cgcgataacg 300  
cggctcttga agtgcactgt cccgcccaac gcctccggaa aggttggctt ctcgaattcg 360  
ggttactggg gtattcccggt gaacgaggcc aactacacca acagcttcta catcaagga 420  
gactataatg gcgacgtcac gcttcgcctt gcgggggcca gcagcggcag tctgtatgca 480  
tcgcaaacga tcagcgtcaa gagcaactcg tccagctaca cctatgtcga gacgcaatat 540  
ccagtgaagc aggctcctga cggcaacaac atctgggaac tgaccttga cgcgtccaag 600  
gtggcgggaa agtcaattgaa ttttgacctg atccagctct accaccaac attccacgac 660  
agacctaatg gcctcaaacc ggaaattgca aatgttttgg ttgacatgaa gggttctttc 720  
ctgcgctttc ccggaggaaa caatctcgaa ggattaactc cggcaacgcg ctggaagtgg 780  
aatgagacca ttgggcctct gacgagccgt cctgggcggc aaggtaactg ggggtatccc 840  
aacacagatg cacttggctt gaatgaatat ttcttgtgga tccaggacat gggtttgacg 900  
gcagttctcg gtgtctggga tggactgact atcggagggc cgagccctca agttatcact 960  
ggggacgctt tgaagcccta catcgacgac gttttgaacg agcttgagta tacccttggga 1020  
gacaaggata ccacgtacgg caagcttctg gcttctcatg gcttccctga ccctggcca 1080  
ctccaatacg tggaaatcgg caacgaggac aacctcaatg gaggcgggccc ctcatatgca 1140  
gagcgcctca cggccttcca cgacgcaatc aaggccaagt acccgcatct gaccatcatt 1200  
agcagcacgg accagtacct cccaaatccc aagccaaaag gaatctgat ggactggcac 1260

ES 2 802 807 T3

acgtactcgc cggcggggga gcttgtgagc atgttcaaca agttcgacaa cgtggaccgg 1320  
gctttcccggt actttgtcgg cgaatatgcc tgcattctccc tcgataacgg gaccgagctc 1380  
gcgagccga tcatgcaggg gtctgtggcg gaagctgtct ttatgattgg catgaaagg 1440  
aacagcgatg ttgtcaagat ggctgcctat gcaccgctgc tgcagcaata cggcgggtac 1500  
acatggaatg accacgtgcc aaacctgatt atgttcaaca acaacccaaa cggcattgtc 1560  
cggagcacca gctactacgt acagcagatg ttctcagtga accggggcag caccattctg 1620  
ccagtggaat cagacgggtc atttggacca gtctactacg tcgcctcaa aacctcagac 1680  
aacaagacat actacgtcaa gttcgccaac tacggcccta ccccgagcc cgtcaccgtc 1740  
aacatcccca agacagaaac cgggcgactg acgacgctgt cggcgccgc tactgcagcc 1800  
aacacctgga gcagccctaa tgctgtcaca cccaagacca gcaatgtcag ctctgctgat 1860  
ggcgcgccag gaaagtttac tattactctg cctgcgtgga gtgttgctgt gttggctgtg 1920  
gagtga 1926

<210> 36

5 <211> 2235

<212> ADN

10 <213> *Rasamsonia emersonii*

<220>

<221> fuente

15 <222> (1)..(2235)

<223> /organismo="Rasamsonia emersonii" /tipo\_molar="ADN no asignado"

<400> 36

20 atgatcctcc gctcgtacc catccacgcc gccgccattg ctgctgcttt gacttctctt 60  
gttgacgcca agcccctccg ccacgacatc cgtcagggtg ctgagactca ccgtatcaac 120  
agcactcact acaacgacac ttctgctggt gtctctgagtg tgtctctgaa ggacgccgcc 180  
ggcgtgaaca acacttctcc tctcctgtac ggctggatgt tcgaggacat ctcccactct 240  
ggcgtggcg gtatctacgc tgagctgatc aagaaccgtg ctttcaggg cagcacttct 300  
actgtcaagc aggtcctcctg tatctccggt gactccgttg ttgctgctca gaacccatt 360  
gtgcccttcg gtcctgtcct cgatggctgg cgcctatag gagatgcaa gctgtctctt 420  
gatgtcctgc accctctgct ggatgccctc cccgtcgtcc tccagatcga tatcccctgg 480  
aacgccactg gcgagggttg tttcctcaac gaaggctggg ggggtattga tgtccgccct 540  
cagacctaca acgcctcctt ctacatgctt gccaacgctc ctcgctacaa caagacctg 600  
acctccatca acctgtcgtc acgcagcaac ctcaccgacg atgtctgggc caccaccacc 660  
atccacgttg accccgacaa ggctcctacc ttcgactacg agcagtacca ggcctccatt 720  
gtgaacaccg tcaaggcccc caactccaac aacaccttcg ccatcacctt caacgccgac 780

ES 2 802 807 T3

gaggttgctg gaagcacctt ctacttcggc ttggtgtctc tgttccctga aaccttcaac 840  
aaccgcccc aacggcctacg caaggacctt gctcagggta tcaaggacat gggtgccaag 900  
ttcctgcggt tccctgggtg aaacaacctc gagggttact ccatcttcca gcggttgaag 960  
tggaacgaaa ccattggtcc cctccgttac cgcaagggtc gtgtcggtaa ctgggaatac 1020  
tacaacacca acggtccttg tcttttgtaa ttcttggat ggactgagga ccttggcatg 1080  
gagcccgtcc ttgctgtcta cgctggcttc tctctggaca tctggggcca ggaaggcacc 1140  
tcctaccccg aggaccgat ggatgagatc gtccaggaca tcctgaacga gcttgagtac 1200  
tgcatgggtg atgtcaacac ccactacggt gctctccgtg cccagcacgg ccaccccgag 1260  
cccttcgaca tcaagtacat tgagatcggg aacgaggact ggttctccag cacctacccc 1320  
taccgtttcc ccatcatcta caaggccatc aagtccgctt accccaacat taccctcatc 1380  
tccactgctt acaacgagaa cgccaactac accatcgaca tccctgctgg tggatgtgg 1440  
gataccacc actacgaaac cccctccttc ttctggaga acttcaacta cttcgacaac 1500  
tggcaggctg ccaactaaca caccgatgtc aagatcttcg tcggtgaata ctccgtctac 1560  
cagatcgaca cccccgatg ctacgtcaac ttctccaacc ccgagggtat ccacatgttc 1620  
ttccccgagc ttgtatctgc tattgctgag gctgtctacc tccttgggtc tgagcgtaac 1680  
cccaacactg tcaccatgac ctctacgct cccagcttcc agaacctgaa ctggtacaac 1740  
tggccccca acctggttgc tttcactgcc aaccctgacg agactgtgtt cagcaccagc 1800  
tactacatgc agaagatggt cgccaaccac cgtggcacc agacctccc tgtcaagaac 1860  
agcaagggtg acttcaacc tctctggtgg gttgccacca ttgacgaggg tgctggtgtt 1920  
gtctacttca agatgttcaa ctccggcaac tcctccatcc ccctcacat taacctggac 1980  
caggcgtaca aggtgtcaa cggcaccatc ctctccgcc gccgtatctg cctccagctc 2040  
cacatctaca ccgtccccga tgagctcacc attgctttcc tccagacca ccccaacctc 2100  
tacggattca actacctgaa caaccagact gccattgtcc ccgtgcctgc caacatcacc 2160  
tccccagca agtcccagaa ccgcttcgac tgggatgttc ccaagtactc cgtcacctgt 2220  
ctccagttcg acatc 2235

<210> 37

5 <211> 722

<212> PRT

<213> Rasamsonia emersonii

10

<400> 37

Lys Pro Leu Arg His Asp Ile Arg Gln Val Ala Glu Thr His Arg Ile  
1 5 10 15

Asn Ser Thr His Tyr Asn Asp Thr Ser Ala Val Val Leu Asp Val Ser



ES 2 802 807 T3

Asp Met Gly Ala Lys Phe Leu Arg Phe Pro Gly Gly Asn Asn Leu Glu  
 275 280 285

Gly Tyr Ser Ile Phe Gln Arg Trp Lys Trp Asn Glu Thr Ile Gly Pro  
 290 295 300

Leu Arg Tyr Arg Lys Gly Arg Val Gly Asn Trp Glu Tyr Tyr Asn Thr  
 305 310 315 320

Asn Gly Leu Gly Leu Leu Glu Phe Leu Glu Trp Thr Glu Asp Leu Gly  
 325 330 335

Met Glu Pro Val Leu Ala Val Tyr Ala Gly Phe Ser Leu Asp Ile Trp  
 340 345 350

Gly Gln Glu Gly Thr Ser Tyr Pro Glu Asp Arg Met Asp Glu Ile Val  
 355 360 365

Gln Asp Ile Leu Asn Glu Leu Glu Tyr Cys Met Gly Asp Val Asn Thr  
 370 375 380

His Tyr Gly Ala Leu Arg Ala Gln His Gly His Pro Glu Pro Phe Asp  
 385 390 395 400

Ile Lys Tyr Ile Glu Ile Gly Asn Glu Asp Trp Phe Ser Ser Thr Tyr  
 405 410 415

Pro Tyr Arg Phe Pro Ile Ile Tyr Lys Ala Ile Lys Ser Ala Tyr Pro  
 420 425 430

Asn Ile Thr Leu Ile Ser Thr Ala Tyr Asn Glu Asn Ala Asn Tyr Thr  
 435 440 445

Ile Asp Ile Pro Ala Gly Gly Met Trp Asp Thr His His Tyr Glu Thr  
 450 455 460

Pro Ser Phe Phe Leu Glu Asn Phe Asn Tyr Phe Asp Asn Trp Gln Ala  
 465 470 475 480

Ala Thr Asn Asn Thr Asp Val Lys Ile Phe Val Gly Glu Tyr Ser Val  
 485 490 495

Tyr Gln Ile Asp Thr Pro Asp Gly Tyr Val Asn Phe Ser Asn Pro Glu  
 500 505 510

Gly Ile His Met Phe Phe Pro Glu Leu Val Ser Ala Ile Ala Glu Ala  
 515 520 525

ES 2 802 807 T3

Val Tyr Leu Leu Gly Ala Glu Arg Asn Pro Asn Thr Val Thr Met Thr  
530 535 540

Ser Tyr Ala Pro Ser Phe Gln Asn Leu Asn Trp Tyr Asn Trp Ser Pro  
545 550 555 560

Asn Leu Val Ala Phe Thr Ala Asn Pro Asp Glu Thr Val Phe Ser Thr  
565 570 575

Ser Tyr Tyr Met Gln Lys Met Phe Ala Asn His Arg Gly Thr Gln Thr  
580 585 590

Leu Pro Val Lys Asn Ser Lys Gly Asp Phe Asn Pro Leu Trp Trp Val  
595 600 605

Ala Thr Ile Asp Glu Gly Ala Gly Val Val Tyr Phe Lys Ile Val Asn  
610 615 620

Ser Gly Asn Ser Ser Ile Pro Leu Thr Ile Asn Leu Asp Gln Ala Tyr  
625 630 635 640

Lys Gly Val Asn Gly Thr Ile Leu Val Arg Arg Arg Ile Cys Leu Gln  
645 650 655

Leu His Ile Tyr Thr Val Pro Asp Glu Leu Thr Ile Ala Phe Leu Gln  
660 665 670

Thr His Pro Asn Leu Tyr Gly Phe Asn Tyr Leu Asn Asn Gln Thr Ala  
675 680 685

Ile Val Pro Val Pro Ala Asn Ile Thr Ser Pro Ser Lys Ser Gln Asn  
690 695 700

Arg Phe Asp Trp Asp Val Pro Lys Tyr Ser Val Thr Val Leu Gln Phe  
705 710 715 720

Asp Ile

<210> 38

5 <211> 23

<212> PRT

10 <213> Rasamsonia emersonii

<400> 38

Met Ile Leu Arg Ser Leu Pro Ile His Ala Ala Ala Ile Ala Ala Ala  
1 5 10 15

15 Leu Thr Ser Leu Val Asp Ala  
20

<210> 39

20 <211> 2483

<212> ADN

<213> Rasamsonia emersonii

5 <220>

<221> fuente

<222> (1)..(2483)

10

<223> /organismo="Rasamsonia emersonii" /tipo\_molar="ADN no asignado"

<400> 39

atgatcctcc gatccctccc tatccatgct gcggccattg cagcagcctt aacttcctc	60
gtcgcagcaa aaccctgcg acatgacatc aggcaggtgg cagaaacgca tcggatcaat	120
agcacgcact acaacgacac ctctgccgtc gtgctggatg tgtcgttgaa ggacgcggcc	180
gggaggaaca atacgtcgcc gctgctgtat gggatggatgt ttgaggatat tagtgtaatt	240
acctccattt gttgaccctt gactattacc gttctcaaga tgcgttgttt gttttgtact	300
ggtgtcattt tatattactg tgcttcagac gaagatggtt tttttttcct atttcctcaa	360
gctaagcctt cttgcagcac tccggcgacg gcggtatata cgcggaattg atcaagaacc	420
gggctttcca gggtgagaac taccttgact ttctttctct tcgtcgttct gacgacctgc	480
caggctcgac atccaccgtc aaacaagcac ccggcatttc aggcgacagc gtggttgctg	540
cacagaacct catcgtcccg tttggccccg tcctcgacgg atggcgacct attggcgatg	600
caaaactctc tctcgacgtg cttcaccctc tgtccgatgc ccttcgggtc gtcctgcaga	660
tcgacatccc ctggaacgcc accggcgaaag ttggcttcct gaatgaaggc tgggtgggca	720
tcgacgtccg gccgcagaca tacaacgcct cgttctacat gctggccaac gccccgcgat	780
acaacaagac gctgaccagc atcaacctgt cgctcgggtc gaacctgacg gacgacgtct	840
gggccacgac gacgatccac gtcgaccggg acaaggtccc cacgttcgac tacgagcagt	900
accaggccag catcgtcaac acggtaagg cgcccaactc gaacaacacc tttgccatca	960
ccttcaatgc ggacgaggtg gctggctcga cattctactt tggcctcgtc agtctgttcc	1020
ctgagacgtt caacaaccgt cccaacggtc tgcgcaagga cctcgcccag ggcacgaagg	1080
acatgggccc gaaattcctc cgcttcccag gcggcaacaa cctcgaggga tactccatct	1140
tccagcgggtg gaagtggaac gagaccatcg gcccggtgag ataccgcaag ggacgcgtcg	1200
gcaactggga atactacaac accaacgggc tgggtctgct ggagttcctg gagtggacgg	1260
aggacctcgg catggagccc gttctggccc tgtacgccgg attctcgtc gacatatggg	1320
15 gccaggaggg cacgtcctac cccgaagacc gcatggacga gattgtgcag gacatcctga	1380

ES 2 802 807 T3

acgagctgga gtactgcatg ggcgacgtca acacgacta cggcgctctg cgcgcgcagc 1440  
 acggccaccc ggagccggtc gatatacaagt acatcgagat cggcaacgag gactggttct 1500  
 ccagcacgta tccctaccgg ttccccatca tctacaaggc catcaagtcc gcgtacccca 1560  
 acatcacctt catctcgacc gcctacaacg agaacgcaa ctacaccatc gacatccccg 1620  
 cgggcggcat gtgggacacg catcactacg agacgcgctc cttcttcctc gagaacttca 1680  
 actacttoga caactggcag gcggcgacga acaacaccga cgtcaagatc ttcgtcggcg 1740  
 agtactccgt ctaccagatc gacacaccgg acggctacgt caacttcagc aaccggagg 1800  
 gcatccacat gttcttcccg gagctcgtgt cggcgatcgc agaggcagtc tatctgctcg 1860  
 gcgcagagcg caatccgaac acggtcacca tgacctoga cgcgcccagc ttccagaacc 1920  
 tgaactggta caactggtcg cccaacctgg tggcctttac tgccaatccc gacgagacgg 1980  
 tcttcagcac cagctactac atgcagaaga tgtttgcaa tcaccggggc acgcagacgc 2040  
 tgccagtga gaattccaag ggagacttta acccgctgtg gtgggttgca acgattgacg 2100  
 agggggctgg tgtgttttac ttcaaggtat tatccattat aacattgctc gttcctggtg 2160  
 tcactaacia caagcagatc gtcaactcgg gaaactcctc catcccgttg acaatcaacc 2220  
 tcgaccaggc atacaaggga gtcaatggga cgatactggt gcgtcgcaga atctgccttc 2280  
 agctacatat atataccggt cctgacgaac tgacaattgc ctttctacag acacacccca 2340  
 acctctacgg attcaactac ctgaacaacc agactgccat cgtgcctgtt cggcggaaca 2400  
 tcacttctcc ctogaagagt cagaaccgct tcgattggga tgtcccgaag tattcgggtca 2460  
 cgggtgttgca gtttgatatt taa 2483

<210> 40

5 <211> 2238

<212> ADN

10 <213> *Rasamsonia emersonii*

<220>

<221> fuente

15 <222> (1)..(2238)

<223> /organismo="Rasamsonia emersonii" /tipo\_molar="ADN no asignado"

<400> 40

20 atgatcctcc gatccctccc tatccatgct gcggccattg cagcagcctt aacttcctc 60  
 gtcgacgcaa aaccctgctg acatgacatc aggcaggtgg cagaaacgca tcggatcaat 120  
 agcacgcact acaacgacac ctctgccgtc gtgctggatg tgtcgttgaa ggacgcggcc 180  
 gggaggaaca atacgtcgcc gctgctgtat ggggtgatgt ttgaggatat tagtcactcc 240  
 ggcgacggcg gtatatacgc ggaattgatc aagaaccggg ctttccaggg ctcgacatcc 300  
 accgtcaaac aagcaccggc catttcaggc gacagcgtgg ttgctgcaca gaaccccatc 360

ES 2 802 807 T3

gtcccgtttg gccccgtcct cgacggatgg cgacccattg gcgatgcaaa actctctctc 420  
gacgtgcttc accccctgtc cgatgccctt ccggctcgtcc tgcagatcga catcccctgg 480  
aacgccaccg gcgaagttgg cttcctgaat gaaggctggg ggggcatcga cgtccggccg 540  
cagacataca acgcctcgtt ctacatgctg gccaacgccc cgcgatacaa caagacgctg 600  
accagcatca acctgtcgtc gcggtcgaac ctgacggacg acgtctgggc cacgacgacg 660  
atccacgtcg acccggacaa ggtccccacg ttcgactacg agcagtacca gcccagcatc 720  
gtcaaacacg tcaaggcgcc caactcgaac aacacctttg ccatcacctt caatgctggac 780  
gagggtggctg gctcgcacatt ctactttggc ctcgtcagtc tgttccctga gacgttcaac 840  
aacgtccca acggtctcgc caaggacctc gcccagggca tcaaggacat gggcgcgaaa 900  
ttcctccgct tcccaggcgg caacaacctc gagggatact ccatcttcca gcggtggaag 960  
tggaacgaga ccatcggccc gttgcgatac cgcaagggac gcgtcggcaa ctgggaatac 1020  
tacaacacca acgggtcggg tctgctggag ttccctggagt ggacggaggga cctcggcatg 1080  
gagcccgctt tggccgtgta cgccggattc tcgctcgaca tatggggcca ggagggcacg 1140  
tcctaccccg aagaccgcat ggacgagatt gtgcaggaca tcctgaacga gctggagtac 1200  
tgcatgggcg acgtcaacac gcaactacggc gctctgcgcg cgcagcacgg ccaccggag 1260  
ccgttcgata tcaagtacat cgagatcggc aacgaggact ggttctccag cacgtatccc 1320  
taccggttcc ccatcatcta caaggccatc aagtccgctg accccaacat caccctcatc 1380  
tcgaccgcct acaacgagaa cgccaactac accatcgaca tccccgcggg cggcatgtgg 1440  
gacacgcac actacgagac gccgtccttc ttccctcgaga acttcaacta cttcgacaac 1500  
tggcaggcgg cgacgaacaa caccgacgtc aagatcttgg tcggcgagta ctccgtctac 1560  
cagatcgaca caccggacgg ctacgtcaac ttcagcaacc cggagggcat ccacatgttc 1620  
ttccccgagc tcgtgtcggc gatcgagag gcagtctatc tgctcggcgc agagcgcaat 1680  
ccgaacacgg tcacatgac ctcgtacggc cccagcttcc agaacctgaa ctggtacaac 1740  
tggtcgocca acctggtggc ctttactgcc aatcccgcag agacggctctt cagcaccagc 1800  
tactacatgc agaagatggt tgccaatcac cggggcacgc agacgctgcc agtgaagaat 1860  
tccaagggag actttaacct gctgtggtgg gttgcaacga ttgacgaggg ggctggtgtt 1920  
gtttacttca agatcgtcaa ctcgggaaac tcctccatcc cgttgacaat caacctcgac 1980  
caggcataca agggagtcaa tgggacgata ctggtgcgtc gcagaatctg ccttcagcta 2040  
catatatata ccgttcctga cgaactgaca attgcctttc tacagacaca ccccaacctc 2100  
tacggattca actacctgaa caaccagact gccatcgtgc ctgttccggc gaacatcact 2160  
tctccctcga agagtcagaa ccgcttcgat tgggatgtcc cgaagtattc ggtcacggtg 2220  
ttgcagtttg atatttaa 2238

<210> 41

5 <211> 1518

<212> ADN

<213> Rasamsonia emersonii

10

<220>

ES 2 802 807 T3

<221> fuente

<222> (1)..(1518)

5 <223> /organismo="Rasamsonia emersonii" /tipo\_molar="ADN no asignado"

<400> 41

```

atgaccacct tcaccaagct cagcgacgag gagactccta ccatctccgt ccacgcttct      60
cgccgcctct ccaagatcaa cccaacatc tactctggct tcaactgagca catgggccgc      120
tgcatctacg gtggtatcta cgaccccggc aaccctctct ccgacgagaa cggattccgc      180
aaggatgttc tcgaagcggt gaaggagctc aacatccccg tcgtccgta ccccggtggc      240
aacttcaactg ccacctacca ctggttggat ggtgttggtc ccaaggacca gcgccccgcc      300
cgtcccgagc tggcctggct aggaaccgaa accaaccagt tcggtaccga cgagttcctc      360
aagtgtgctg agatgcttgg tgctgagccc tacctctgcc tgaacttcgg tactgttacc      420
ctcgacgagg ccatggcctg ggttgaatac tgcaacggca ccggcaacac ctactacgcc      480
aacctccgtc gcaagaacgg ccgtgagaag ccctacaacg tcaagtactg ggctcttggc      540
aacgaaacct ggggtccctg gcagatcgag cagatgacca aggaagccta cgctcacaag      600
gcctaccagt gggccaaggc tctcaagctc ctggatocca acatcatcct gatcctctgc      660
ggtcaggatg gaactgccag ctgggactac tacaccctca agcactgctt gcagcccacc      720
aaggccaacc tcaactcaa ccccgctcct ctgattgaca tgcaactccat ccacatgtac      780
actgcttctt ccaagcacct gcccaacgcc actgcgcctc ttgctgctga gcgtgccatt      840
gagatcacca gcagcttgat tgaccttgcct cgcattgaga acggcatctc ccccgatgag      900
cctcgtccca ccatctgctt cgacgaatgg aacgtctggg accctgtccg tgctgagggc      960
tccaagggtg ctgaggagag ctacactctg tcgcatgccc ttgctggttg tgtgttcctc     1020
aacgtgttcg tccgcaagtc caaggacctt ggaatggcct gcattgcca gtccgtcaac     1080
gtcatctccc ctctgatgac cactcgtgac ggtatcgtca agcagaccac ctggtggcct     1140
ctctggctgt tctcccgctt catgcgtggc tggaccattg gtgctcacgt ttctctcggt     1200
gcctacgagg gtgagacttc tctgcctgg ctgcgtggtg tcaaggacac tccttggtg     1260
gatgtctctg ccgctcttgg tgatgacggc ttctgcaacg ttggtgttgt caacatccac     1320
gaaaccaagg acttcaagac caacgtgaag ggtgtctccg gtgggcccg gaccgtctac     1380
accgtcactg gtgctgatgt gtctgccacc aacatgaagg gcaagcagga agtcggtgtc     1440
caggagtcca cctgggatgg caagggcagc tacgtgttcc ccaagcactc gctgaccctg     1500
ctccgctgga aggccgag

```

<210> 42

15 <211> 506

<212> PRT

<213> Rasamsonia emersonii

20

<400> 42

ES 2 802 807 T3

Met Thr Thr Phe Thr Lys Leu Ser Asp Glu Glu Thr Pro Thr Ile Ser  
 1 5 10 15

Val His Ala Ser Arg Arg Leu Ser Lys Ile Asn Pro Asn Ile Tyr Ser  
 20 25 30

Gly Phe Thr Glu His Met Gly Arg Cys Ile Tyr Gly Gly Ile Tyr Asp  
 35 40 45

Pro Gly Asn Pro Leu Ser Asp Glu Asn Gly Phe Arg Lys Asp Val Leu  
 50 55 60

Glu Ala Leu Lys Glu Leu Asn Ile Pro Val Val Arg Tyr Pro Gly Gly  
 65 70 75 80

Asn Phe Thr Ala Thr Tyr His Trp Leu Asp Gly Val Gly Pro Lys Asp  
 85 90 95

Gln Arg Pro Ala Arg Pro Glu Leu Ala Trp Leu Gly Thr Glu Thr Asn  
 100 105 110

Gln Phe Gly Thr Asp Glu Phe Leu Lys Trp Cys Glu Met Leu Gly Ala  
 115 120 125

Glu Pro Tyr Leu Cys Leu Asn Phe Gly Thr Gly Thr Leu Asp Glu Ala  
 130 135 140

Met Ala Trp Val Glu Tyr Cys Asn Gly Thr Gly Asn Thr Tyr Tyr Ala  
 145 150 155 160

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

Trp Ala Leu Gly Asn Glu Thr Trp Gly Pro Trp Gln Ile Glu Gln Met  
 180 185 190

Thr Lys Glu Ala Tyr Ala His Lys Ala Tyr Gln Trp Ala Lys Ala Leu  
 195 200 205

ES 2 802 807 T3

Lys Leu Leu Asp Pro Asn Ile Ile Leu Ile Leu Cys Gly Gln Asp Gly  
 210 215 220

Thr Ala Ser Trp Asp Tyr Tyr Thr Leu Lys His Cys Leu Gln Pro Thr  
 225 230 235 240

Lys Ala Thr Leu Asn Ser Asn Pro Val Pro Leu Ile Asp Met His Ser  
 245 250 255

Ile His Met Tyr Thr Ala Ser Ser Lys His Leu Pro Asn Ala Thr Ala  
 260 265 270

Pro Leu Ala Ala Glu Arg Ala Ile Glu Ile Thr Ser Ser Leu Ile Asp  
 275 280 285

Leu Ala Arg Ile Glu Asn Gly Ile Ser Pro Asp Glu Pro Arg Pro Thr  
 290 295 300

Ile Cys Phe Asp Glu Trp Asn Val Trp Asp Pro Val Arg Ala Glu Gly  
 305 310 315

Ser Lys Gly Ala Glu Glu Ser Tyr Thr Leu Ser Asp Ala Leu Ala Val  
 325 330 335

Gly Val Phe Leu Asn Val Phe Val Arg Lys Ser Lys Asp Leu Gly Met  
 340 345 350

Ala Cys Ile Ala Gln Ser Val Asn Val Ile Ser Pro Leu Met Thr Thr  
 355 360 365

Arg Asp Gly Ile Val Lys Gln Thr Thr Trp Trp Pro Leu Trp Leu Phe  
 370 375 380

Ser Arg Phe Met Arg Gly Trp Thr Ile Gly Ala His Val Ser Cys Gly  
 385 390 395 400

Ala Tyr Glu Gly Glu Thr Ser Pro Ala Trp Leu Arg Gly Val Lys Asp  
 405 410 415

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

Asn Val Val Val Val Asn Ile His Glu Thr Lys Asp Phe Lys Thr Asn  
 435 440 445

Val Lys Gly Val Ser Gly Gly Pro Val Thr Val Tyr Thr Val Thr Gly  
 450 455 460

Ala Asp Val Ser Ala Thr Asn Met Lys Gly Lys Gln Glu Val Gly Val  
 465 470 475 480

Gln Glu Ser Thr Trp Asp Gly Lys Gly Ser Tyr Val Phe Pro Lys His  
 485 490 495

Ser Leu Thr Leu Leu Arg Trp Lys Ala Glu  
 500 505

<210> 43  
 <211> 4  
 5 <212> PRT  
 <213> Artificial Secuencia  
 <220>  
 10 <223> sin péptido de señalización predicho  
 <400> 43  
**Met Met Met Met**  
 15 **1**  
 <210> 44  
 <211> 2005  
 20 <212> ADN  
 <213> Rasamsonia emersonii  
 25 <220>  
 <221> fuente  
 <222> (1)..(2005)  
 30 <223> /organismo="Rasamsonia emersonii" /tipo\_molar="ADN no asignado"  
 <400> 44  
 atgacgacat tcaccaagct cagcgacgag gagacgcca ccatctctgt ccatgcgctcc 60  
 aggcgtctgt ccaaaatcaa cccaacatc tactcgggggt tcacagagta gggagccgac 120  
 agacgtgtct accggtttta gttcgctagt gagaccctgc tgactgatag ctttcgttac 180  
 caggcatatg ggacgatgca tctacggagg catatatgat cccggaaatc ctctgtcggg 240  
 cgagaatggc tttcgaaagg atgttctgga ggctctgaag gagctcaata tccctgtggt 300  
 tcgctatccc ggcggaaact ttaccgccac gtatcactgg ttagacgggtg ttggcccgaa 360  
 ggatcaacga cccgcgaggt attcaatgac ttcttctctt ggatttcatt tctttcgttt 420  
 gactgtctgt gttttacat ggttctgttt taataccatg aactatttga gagactttag 480  
 ctggtttcca tataatTTTT ttaccatacg cggcacctgt catgctaaca aagggaaca 540  
 atttgaacag accggagctc gcctggctgg ggactgagac caatcagttc ggactgatg 600  
 agttcttaaa atgggtgtgag atgctcggcg cggagccgta cctctgcttg aactttggaa 660  
 35 cagggtcggc gtgccgatat cctgaccctt gtacctgaga acaaacaac taattggatt 720

ES 2 802 807 T3

gatttaggca ctcttgatga aggtgagcat ccttcaattc ctgtcctata atgccagcac 780  
aagatgatgt tgacacttct cagcaatggc atgggtogag tactgcaatg ggacgggaaa 840  
tacctactat gccaacctgc ggaggaagaa tggtcgggag aagccctaca atgtacgatc 900  
cagtcactga ttggacattc tgttttttaa atactaatac cagagtttca ggttaaatac 960  
tgggctctgg gtaacgagac atgggggccg tggcaaattg aacaaatgac caaggaagcc 1020  
tatgtcaca aggcgtatca atggggccaaa ggtacgttga caatttata atctatttat 1080  
ctatacctct aacatacgtc cagctctcaa gctgctcgat ccgaatataa ttctcattct 1140  
ttgcggtcag gacggcactg catcgtggga ctactacacc ctcaaact gtctgcagcc 1200  
aaciaaggct actctaaaca gcaaccgggt tccccttatt gacatgcaca gcattcacat 1260  
gtacacggcg tcgtccaagc atttgccaaa cgccaccgca cccttggctg cggagcgtgc 1320  
gattgagatt acatcctccc tgatcgacct ggctcggatc gaaaacggga tatccccgga 1380  
tgagccgcgc ccgacgatct gctttgacga gtggaacgtc tgggatcccg tcgctgccga 1440  
gggcagcaag ggagccgagg agagctacac cctgtccgac gcgttggcgg tgggagtttt 1500  
cctcaacgtg ttcgtccgga agagcaagga cctgggaatg gcgtgcatcg cgcaaagcgt 1560  
caatgtcatt tcgcccgtca tgacgaccag agacggcatt gtgaagcaaa cgacgtggtg 1620  
gccgctgtgg cttttctccc gattcatgcg gggctggacc atcgggtgcg atgttcctg 1680  
tggcgcctac gagggcgaaa cctcaccgc gtggctgcg gcgctcaagg aactccctg 1740  
gctggacgtc agcggccct tgggtgacga tggctttgtc aacgtagtgt tcgtcaacat 1800  
ccatgaaact aaggacttta agaccaacgt caagggcgtc agtggaggtc ctgtcactgt 1860  
ctacacggtc accggagccc atgtgtctgc tactaacatg aaggggaagc aagaggtggg 1920  
agttcaggag agtacatggg acggcaaagg cagttatgtg ttcccgaagc attctctgac 1980  
gctgctgaga tggaaaggcag agtga 2005

<210> 45

5 <211> 1521

<212> ADN

10 <213> Rasamsonia emersonii

<220>

<221> fuente

15 <222> (1)..(1521)

<223> /organismo="Rasamsonia emersonii" /tipo\_molar="ADN no asignado"

20 <400> 45

atgacgacat tcaccaagct cagcgacgag gagacgccca ccatctctgt ccatgcgtcc 60  
aggcgtctgt ccaaaatcaa cccaacatc tactcgggggt tcacagagca tatgggacga 120  
tgcactctac gaggcataata tgatcccgga aatcctctgt cggacgagaa tggctttcga 180

ES 2 802 807 T3

aaggatgttc tggaggctct gaaggagctc aatatccctg tggttcgcta tcccggcggg 240  
 aactttaccg ccacgtatca ctggttagac ggtgttggcc cgaaggatca acgaccccg 300  
 agaccggagc tcgctggct ggggactgag accaatcagt tcggcactga tgagttctta 360  
 aatggtgtg agatgctcgg cgcgagccg tacctctgct tgaactttgg aacaggcact 420  
 cttgatgaag caatggcatg ggtcgagtac tgcaatggga cgggaaatac ctactatgcc 480  
 aacctgcgga ggaagaatgg tcgggagaag ccctacaatg ttaaatactg ggctctgggt 540  
 aacgagacat ggggtccgtg gcaaattgaa caaatgacca aggaagccta tgctcacaag 600  
 gcgtatcaat gggccaaagc tctcaagctg ctcgatccga atataattct cattctttgc 660  
 ggtcaggacg gcaactgcatc gtgggactac tacaccctca aacactgtct gcagccaaca 720  
 aaggctactc taaacagcaa cccggttccc cttattgaca tgcacagcat tcacatgtac 780  
 acggcgtcgt ccaagcattt gccaaacgcc accgcaccct tggctgcgga gcgtgcgatt 840  
 gagattacat cctccctgat cgacctggct cggatcgaaa acgggatatc cccggatgag 900  
 ccgcgccga cgatctgctt tgacgagtgg aacgtctggg atcccgctgcg tgccgagggc 960  
 agcaagggag ccgaggagag ctacaccctg tccgacgcgt tggcgggtggg agttttcctc 1020  
 aacgtgttcg tccggaagag caaggacctg ggaatggcgt gcatcgcgca aagcgtcaat 1080  
 gtcatttcgc cgctcatgac gaccagagac ggcattgtga agcaaacgac gtggtggccg 1140  
 ctgtggcttt tctcccgatt catgcggggc tggaccatcg gtgocgatgt ttctgtggc 1200  
 gcctacgagg gcgaaacctc acccgctgg ctgctcgggc tcaaggacac tccttgctg 1260  
 gacgtcagcg cggccttggg tgacgatggc tttgtcaacg tagttgtcgt caacatccat 1320  
 gaaactaagg actttaagac caacgtcaag ggcgtcagtg gaggtcctgt cactgtctac 1380  
 acggtcaccg gagccgatgt gtctgctact aacatgaagg ggaagcaaga ggtgggagtt 1440  
 caggagagta catgggacgg caaaggcagt tatgtgttcc cgaagcattc tctgacgctg 1500  
 ctgagatgga aggcagagtg a 1521

<210> 46

5 <211> 966

<212> ADN

<213> Rasamsonia emersonii

10

<220>

<221> fuente

15 <222> (1)..(966)

<223> /organismo="Rasamsonia emersonii" /tipo\_molar="ADN no asignado"

<400> 46

20

atgcgccccg gcttctccg ctacgtgtct cttgctctc tgctctggtc tgctctgcc 60  
 tccgctgagc ctgtccgtgt tctcgacacc gacttccctg acccttgctt gatctccacc 120

ES 2 802 807 T3

aacggcaagt actacgcttt cgccaccact ggcaacggtg tcaacgtcca gattgccag 180  
 tcggatgact tcgtcacctg ggagcgtctt gctggtactg atgcccttcc tggcccttc 240  
 cccagctggg ttgcctcctc cccactgtc tgggctcccg atgtcatcca gcgtctggac 300  
 ggtacctacg tcattgacta ctctgctctt tcttccaccg acccagcaa gcactgcttc 360  
 ggtgctgcca cctcgacctc catcaccggt ccctacaagc ccgaggacaa ctacattgcc 420  
 tgccccctcg accaggggtg tgccatcgac cccgatggat tcatcgatga tgacggcacc 480  
 atgtacgtgg tctacaaggt tgacggcagc aacctggatg gagatggcac catccacccc 540  
 actcccatca tgctccaggc tctggagcct gacggcatca cccccactgg tgaccccatc 600  
 aagctcctcg accgtgatgc ctccgacggc atcttgattg aggctccctc cctcgtccgc 660  
 tccgttgttt ccggcaccta cttcctgagc tactcttctc actactacgc ctccctgcac 720  
 tacgacgttg gatacgccac tggccctgcc gtcaagggtc ccttcaccaa ggcccaggct 780  
 cctcttctcg tcaccgggtga caacaccacc aacgccggtc ctctgggtgg tcccgtggg 840  
 gccgacttct ccgtcgacgg cactcgtatc gtgttccacg ctttcgagaa cggccgcaac 900  
 ctgaccaacg gccgtgctct ctacaccagc ggtatcgtcc tcgaggggtga tgtcattcgc 960  
 ctagtt 966

<210> 47

5 <211> 300

<212> PRT

<213> *Rasamsonia emersonii*

10

<400> 47

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

ES 2 802 807 T3

100

105

110

Glu Asp Asn Tyr Ile Ala Cys Pro Leu Asp Gln Gly Gly Ala Ile Asp  
 115 120 125

Pro Asp Gly Phe Ile Asp Asp Asp Gly Thr Met Tyr Val Val Tyr Lys  
 130 135 140

Val Asp Gly Ser Asn Leu Asp Gly Asp Gly Thr Ile His Pro Thr Pro  
 145 150 155 160

Ile Met Leu Gln Ala Leu Glu Pro Asp Gly Ile Thr Pro Thr Gly Asp  
 165 170 175

Pro Ile Lys Leu Leu Asp Arg Asp Ala Ser Asp Gly Ile Leu Ile Glu  
 180 185 190

Ala Pro Ser Leu Val Arg Ser Val Val Ser Gly Thr Tyr Phe Leu Ser  
 195 200 205

Tyr Ser Ser His Tyr Tyr Ala Ser Leu His Tyr Asp Val Gly Tyr Ala  
 210 215 220

Thr Gly Pro Ala Val Lys Gly Pro Phe Thr Lys Ala Gln Ala Pro Leu  
 225 230 235 240

Leu Val Thr Gly Asp Asn Thr Thr Asn Ala Gly Pro Leu Gly Gly Pro  
 245 250 255

Gly Gly Ala Asp Phe Ser Val Asp Gly Thr Arg Ile Val Phe His Ala  
 260 265 270

Phe Glu Asn Gly Arg Asn Leu Thr Asn Gly Arg Ala Leu Tyr Thr Ser  
 275 280 285

Gly Ile Val Leu Glu Gly Asp Val Ile Arg Leu Val  
 290 295 300

<210> 48

5 <211> 22

<212> PRT

<213> Rasamsonia emersonii

10

<400> 48

Met Arg Pro Ser Phe Leu Arg Tyr Val Ser Leu Ala Pro Leu Leu Trp  
 1 5 10 15

Ser Ala Ala Ala Ser Ala

20

15

<210> 49

<211> 1026  
 <212> ADN  
 5 <213> *Rasamsonia emersonii*  
 <220>  
 <221> fuente  
 10 <222> (1)..(1026)  
 <223> /organismo="Rasamsonia emersonii" /tipo\_molar="ADN no asignado"  
 15 <400> 49  
 atgcgacat ctttccttcg atacgtctct ctagcaccctc tcctgtggag cgcagcagca 60  
 tctgctgagc ccgtccgagt gctggacacc gacttcccag acccctgcct gatctccacc 120  
 aacggcaagt actatgcctt tgcgacgacg ggcaatggcg tcaacgtgca aatcgcccag 180  
 tcggacgact tcgtcacctg ggagcgtctg gctggcactg atgccctgcc gggaccgttt 240  
 ccctcgtggg tggcgtcttc gccgactgta tgggcgccgg acgtgattca acgggtgcgt 300  
 ttaaactata cataggattt ttcacgtca tatctaacga ggtgcatcta gctcgatggc 360  
 acgtacgtga tgtactactc cgccctctcg agcacggacc cgtccaagca ctgctttggc 420  
 gccgccacct cgacctccat cacaggccct tacaagcccg aagacaacta catcgccctgt 480  
 ccgctggacc agggcggcgc catcgaccga gacggcttca tcgacgacga cggcaccatg 540  
 tacgtcgtct acaaggctcg cgggagcaac ctggacggcg acggcacgat ccatccgacg 600  
 cccatcatgc tgcaggcgtt ggagccggac ggcatcactc cgacagggga cccgatcaag 660  
 ctgctagacc gggacgcgtc ggacgggatc ctgattgagg cgcccagtct ggtcagaagc 720  
 gtcgtcagcg ggacgtactt cctctcgtac tcgtcgcact actacgccag tttgcactac 780  
 gacgtcggct acgcgacagc gccggctgtg aaggggccct tcaccaaggc gcaggccccg 840  
 ttgctggtga cgggcgacaa tactaccaat gctggtctct tgggtggacc tggaggagca 900  
 gacttttccg tcgacgggac gaggattgtc tttcatgctt ttgagaatgg aaggaatctt 960  
 actaatggca gggcgtgta tacgtcgggg attgtccttg agggggacgt catacggttg 1020  
 gtgtag 1026  
 <210> 50  
 20 <211> 969  
 <212> ADN  
 25 <213> *Rasamsonia emersonii*  
 <220>  
 <221> fuente  
 30 <222> (1)..(969)  
 <223> /organismo="Rasamsonia emersonii" /tipo\_molar="ADN no asignado"  
 35 <400> 50

ES 2 802 807 T3

atgcgacat ctttccttcg atacgtctct cttagcacc cctctgtggag cgcagcagca 60  
tctgctgagc ccgtccgagt gctggacacc gacttcccag acccctgcct gatctccacc 120  
aacggcaagt actatgcctt tgcgacgacg ggcaatggcg tcaacgtgca aatcgcccag 180  
tcggacgact tcgtcacctg ggagcgtctg gctggcactg atgccctgcc gggaccgttt 240  
ccctcgtggg tggcgtcttc gccgactgta tgggcgccgg acgtgattca acggctcgat 300  
ggcacgtacg tgatgtacta ctccgccctc tcgagcacgg acccgtccaa gcactgcttt 360  
ggcgcgccca cctcgacctc catcacaggc ccttacaagc ccgaagacaa ctacatcgcc 420  
tgtccgctgg accagggcgg cgccatcgac ccagacggct tcatcgacga cgacggcacc 480  
atgtacgtcg tctacaaggt cgacgggagc aacctggagc gcgacggcac gatccatccg 540  
acgccatca tgctgcaggc gctggagccg gacggcatca ctccgacagg ggacccgatc 600  
aagctgctag accgggacgc gtcggacggg atcctgattg aggcgccag tctggtcaga 660  
agcgtcgtca gcgggacgta cttcctctcg tactcgtcgc actactacgc cagtttgca 720  
tacgacgtcg gctacgcgac agggccggt gtgaagggc cttcaccaa ggcgacggcc 780  
ccgttctgg tgacgggcca caatactacc aatgctggtc ctctgggtgg acctggagga 840  
gcagactttt ccgtcgacgg gacgaggatt gtctttcatg cgtttgagaa tggaaggaat 900  
cttactaatg gcagggcgct gtatacgtcg gggattgtcc ttgaggggga cgtcatacgg 960  
ttggtgtag 969

<210> 51

5 <211> 1998

<212> ADN

10 <213> *Rasamsonia emersonii*

<220>

<221> fuente

15 <222> (1)..(1998)

<223> /organismo="Rasamsonia emersonii" /tipo\_molar="ADN no asignado"

<400> 51

20

atgctttctg ctttcctctt ctctctcccc actctcagct ctctcagtc tgcttccaag 60  
accgagaaca gcactgcca cactcctggt atccgcttga tcaacggcca cgaacacgtc 120  
ctgatccagc cctgggggtg cgatggcttc cgcgtctgtg ccaactctca cgccttcccc 180  
gacaactccg tccagcgtca ccaccgctct ccttccggcc gtcccgcga tcttccatt 240  
ggtaccagct actccacat caacgccact cagactgcca ccgtcaccaa cggacacgcc 300  
tccgtcgtcg tcgacaaggg tcagctgtcg ttctaccgtg tcaacggcaa caactccaag 360  
accctcgtcc tccaggagct ctggctccag acccccgtt tccctgctcg ctggtaccag 420  
aagcgtgatg ctgcgccgtg attccgtgct cagttctcct tcgtcgtctc tgactcccc 480

ES 2 802 807 T3

accgagcgtc tgttcggtac cggccaggat ttgactggct ctctgatcaa gaagaaccag 540  
 accatcgatc ttgtcaagtt caacaccctc aaccccatcc ccaccattgt ctccgaccgt 600  
 ggatacctct tcttctggaa cgttccctcc ctccggtcaga tggagctttc tcctaccgc 660  
 acctcgttc tgtccgacca cacctccgtt gtggactact acattgccat ccgtcccag 720  
 gccgactacg acggtctgct ccagcgttac acttccgtca ccggccgag ccccatgatc 780  
 cctgacttgc gtactggata ctggcagtgc aagctccgtt acgccacca ggaagaattg 840  
 ctcaacgtca ccatgggttt cgctgagcgc aagatccccg tcagcatgtt catcatcaac 900  
 tacctgagct ggtcccacga gggtgactgg gctctcaacg cctccgcctg gcctgacccc 960  
 gccaaagtgg ctgctgaggc tgagctgatg gcctccgtct ggccctccat tgaggatgcc 1020  
 agccccaact gggctgagat gcagtctctt ggattcctg cccagacttt cgaccctctt 1080  
 gacattgctg gctactggca gaacaagtac cagaacggta tccgcaactt ctggctagat 1140  
 gaggatgagg gtggccaggt tgctctcgat gtctaccctt ggactgacta ctaccttggg 1200  
 cctggtgacc agtacgccat gttttcccc tacttccacc agatgggtgt tttcgagggc 1260  
 cagatgaaaa ctggtggcag caacaacgtt tctgctgtct ctctgtcccg ctctctgcc 1320  
 aactgggaca gcatgaagat gatgatctcc gctggccagt ccatggccat gtccggccag 1380  
 ggatggtgga ccctcgacat tgggtggttc aagactgatg gccagtcgaa ctccggccaac 1440  
 atctcggacc ccgagtacca ggagctgttc gtccgctggc tgcagtggg taccttcctc 1500  
 cccatcatgc gcaaccacgg tatgctgacc tgctgcct ccgctcagga aggcttcctg 1560  
 acctgcccct ccgagccctg gagctacggt cctgacaacc tgcccatcat tgtcagctac 1620  
 atcaacctcc gttacaagct gcagccctac atcaaggccc tcttccgcat gctgagcgaa 1680  
 tctggccgtg ccatgatgcg ccctctgttc atggacttct ccctctccga cccctacacc 1740  
 ctggaggcca ctgaggacct caagctccag tacatgttcg gtccctcgtt gcttgtctct 1800  
 cctgtcacca cctaccgtgc caccaacgcc accgtctacc ttccaagct tccttcggat 1860  
 gctcccgata ccaagtggac ctactggtgg accaacgagt ctttcgatgg tggtcagtgg 1920  
 gtgaccgttc ctgccaacaa gtccatcatc ccctgttcc gtctgggcag cgaagcggac 1980  
 atcctcaccg gcaacatt 1998

<210> 52

5 <211> 647

<212> PRT

<213> *Rasamsonia emersonii*

10

<400> 52

Lys Thr Glu Asn Ser Thr Ala Asn Thr Pro Gly Ile Arg Leu Ile Asn  
 1 5 10 15

ES 2 802 807 T3

Gly His Glu His Val Leu Ile Gln Pro Trp Gly Ala Asp Gly Phe Arg  
 20 25 30

Val Arg Ala Thr Leu Asn Arg Phe Pro Asp Asn Ser Val Gln Arg His  
 35 40 45

His Arg Ser Pro Ser Gly Arg Pro Ala Asp Leu Pro Ile Gly Thr Ser  
 50 55 60

Tyr Ser Thr Ile Asn Ala Thr Gln Thr Ala Thr Val Thr Asn Gly His  
 65 70 75 80

Ala Ser Val Val Val Asp Lys Gly Gln Leu Ser Phe Tyr Arg Val Asn  
 85 90 95

Gly Asn Asn Ser Lys Thr Leu Val Leu Gln Glu Leu Trp Leu Gln Thr  
 100 105 110

Pro Gly Phe Pro Ala Arg Trp Tyr Gln Lys Arg Asp Ala Arg Arg Gly  
 115 120 125

Phe Arg Ala Gln Phe Ser Phe Val Ala Pro Asp Ser Pro Thr Glu Arg  
 130 135 140

Leu Phe Gly Thr Gly Gln Asp Leu Thr Gly Ser Leu Ile Lys Lys Asn  
 145 150 155 160

Gln Thr Ile Asp Leu Val Lys Phe Asn Thr Leu Asn Pro Ile Pro Thr  
 165 170 175

Ile Val Ser Asp Arg Gly Tyr Leu Phe Phe Trp Asn Val Pro Ser Leu  
 180 185 190

Gly Gln Met Glu Leu Ser Pro Thr Arg Thr Ser Phe Leu Ser Asp His  
 195 200 205

Thr Ser Val Val Asp Tyr Tyr Ile Ala Ile Arg Pro Glu Ala Asp Tyr  
 210 215 220

Asp Gly Leu Leu Gln Arg Tyr Thr Ser Val Thr Gly Arg Ser Pro Met  
 225 230 235 240

Ile Pro Asp Phe Gly Thr Gly Tyr Trp Gln Cys Lys Leu Arg Tyr Ala  
 245 250 255

Thr Gln Glu Glu Leu Leu Asn Val Thr Met Gly Phe Ala Glu Arg Lys



ES 2 802 807 T3

Asp Asn Leu Pro Ile Ile Val Ser Tyr Ile Asn Leu Arg Tyr Lys Leu  
 515 520 525

Gln Pro Tyr Ile Lys Ala Leu Phe Arg Met Leu Ser Glu Ser Gly Arg  
 530 535 540

Ala Met Met Arg Pro Leu Phe Met Asp Phe Ser Leu Ser Asp Pro Tyr  
 545 550 555 560

Thr Leu Glu Ala Thr Glu Asp Leu Lys Leu Gln Tyr Met Phe Gly Pro  
 565 570 575

Arg Leu Leu Val Ser Pro Val Thr Thr Tyr Arg Ala Thr Asn Ala Thr  
 580 585 590

Val Tyr Leu Pro Lys Leu Pro Ser Asp Ala Pro Asp Thr Lys Trp Thr  
 595 600 605

Tyr Trp Trp Thr Asn Glu Ser Phe Asp Gly Gly Gln Trp Val Thr Val  
 610 615 620

Pro Ala Asn Lys Ser Ile Ile Pro Leu Phe Arg Leu Gly Ser Glu Ala  
 625 630 635 640

Asp Ile Leu Thr Arg Asn Ile  
 645

<210> 53

5 <211> 19

<212> PRT

<213> Rasamsonia emersonii

10

<400> 53

Met Leu Ser Ala Phe Leu Phe Ser Ser Pro Thr Leu Ser Ser Pro Gln  
 1 5 10 15

Ser Ala Ser

15 <210> 54

<211> 2684

<212> ADN

20

<213> Rasamsonia emersonii

<220>

25 <221> fuente

<222> (1)..(2684)

30

<223> /organismo="Rasamsonia emersonii" /tipo\_molar="ADN no asignado"

<400> 54

ES 2 802 807 T3

atgctttccg cttttctttt ctcttcacca acgctgtcct caccacaaag cgcgtccaaa 60  
 acagagaaca gcacagcaaa tacaccaggg attcgtctga tcaacggcca tgaacacgtt 120  
 ctgattcagc cttggggagc agatgggttc cgtgtccggg ccaactctgaa cagatttcct 180  
 gacagtgcgt agccacggaa aacgcacott gttcatctga ctaatttctt acctgctgat 240  
 tgacactctg tagatgaagc gtacagcgtc atcatagatc cccctctgga cggcccgcag 300  
 acctcccat cggcaccagc tacagacca tcaatgcgac acaaaccgcc accgtgacca 360  
 acggccatgc ctccgtggtc gtcgacaagg gccagctgag tttctatcgc gtaaatggca 420  
 acaacagtaa aacactcgtc ttgcaggagc tgtggctaca gacaccggga ttcccagcgc 480  
 gatggtatca aaagagagat gcaagacggg gctttcagc gcagttcagc ttcgtcgcgc 540  
 ccgacagtcc cacggagcgg ctgtttggga cgggccagga tctgacggc tctctcatca 600  
 agaagaatca gaccatcgc ttggtcaagt tcaacacgct gaatccgatt ccgacaatcg 660  
 tgtccgatcg agggatatctg tttttctgga acgtgccttc gcttgggcag atggagtgtg 720  
 ctccaacgag gtatgcatca gcaaaatttg aggaacagga tcatttctaa tgtcaatatt 780  
 agaaccagtt tcctgagcga tcacacttca gtggtggact attacattgc cattogcccc 840  
 gaagccgact atgacggcct gctgcagcgg tacaccagtg tcaactgggag atcgcctatg 900  
 attcctgatt ttggcaccgg aactggcag tgcaagctcc gttatgcgac tcaggaggag 960  
 ctctcaatg tgactatggg atttgcagag gtcagaccct ttggaaatga taagactgtt 1020  
 tctgctgact tgtttgcagc gaaaaatacc tgtttcgatg ttcacatca actatctcag 1080  
 ctggtctcac gagggtgact gggctctgaa cgcacggcg tgccagatc cggcgaagat 1140  
 ggccgcagag ggtgagaaga gtcatttcta ccaagggtgt gtagtctgac tggagccgaa 1200  
 ctgatggcat ccgtctggcc cagtatcag gacgccagtc ccaactgggc tgaatgcag 1260  
 agcttgggtt ttgccgcgca gacattcgat cctctcgata ttgctggcta ctggcagaac 1320  
 aaatacgtgc atatgatcga tcctaccaat gcagcagcca gggacttcct ctggagtaac 1380  
 cttaccggc attactagca gaacgggatc cgaaacttct ggctcgacga agacgaggga 1440  
 ggccaggctc cgctcgacgt ctaccctggg acggactatt acctgggtcc gggggatcag 1500  
 tatgccatgc tgttccccta ctttcatcag atgggggttt ttgaaggaca gatgaagaca 1560  
 gtggggtcga ataagtgttc cgctgtgtcc ctgtcgggt cctcgtgggt gggctcccag 1620  
 cgtttttgcc ggccagctgt ggagtggaga tattttgtat gtgttgaca cttttcaata 1680  
 atgtaacctg ctgattctgg tcaacagcgc caactgggat tccatgaaga tgatgatc 1740  
 cgcaggacag agcatggcca tgtcagggca gggttggtg actttggaca ttggaggctt 1800  
 caagacagac gggcagcca actcggcgaa tattagcgac ccggagtacc aggaattgtt 1860  
 tgtacgggtg ctgcagtggt aagcatttga atactccatt tcctgtattg ataagtaata 1920

ES 2 802 807 T3

cgtacagggg aaccttctta cgggtgcgta aatgaatgcc ctggtggcaa ttcaccaact 1980  
 gaacttgacg tagattatgc gcaaccacgg aatgagaacc tgtctccctt cagcacagga 2040  
 aggcttctctg acctgtccta gcggtacaga cccgtgccat atttttataa cccaataag 2100  
 tatctactga cactgtcaaa cttctagaac cctggagcta cggccccgac aacctcccca 2160  
 tcatcgtgtc ctacatcaac ctccgataca aactgcagcc ctacatcaag gcgctcttcc 2220  
 ggatgctctc cgagagtgga cgggcaatga tgcggcccct gtttatggat ttctctctca 2280  
 gcgatcctta cactttggag gcgacggaag atctgaaatt gcagtacatg tttgggccgc 2340  
 gcttgctcgt cagtcccgtc acaacgtatc gcgcgacgaa tgcgaccgtc tacctgccaa 2400  
 aattgccgtc ggatgctccc gatacgaat ggacgtactg gtggacgaac gagagttttg 2460  
 acggtggtca atgggtgagt tgtctctctc ttttttgata gccagttgac tgacattaat 2520  
 ttctcaggtc accgttctctg cgaacaaaag catgtaagta ttcaaatcac gtacacatgc 2580  
 agcttcgaat tgaatatctc ctgtctgacc tgagctgac tgcatgcaga attccgctct 2640  
 tccgactcgg cagcgaagca gacatcttga cgaggaacat ctaa 2684

<210> 55

5 <211> 2001

<212> ADN

<213> Rasamsonia emersonii

10

<220>

<221> fuente

15 <222> (1)..(2001)

<223> /organismo="Rasamsonia emersonii" /tipo\_molar="ADN no asignado"

<400> 55

20

atgctttccg cttttctttt ctcttcacca acgctgtcct caccacaaag cgcgtccaaa 60  
 acagagaaca gcacagcaaa tacaccaggg attcgtctga tcaacggcca tgaacacggt 120  
 ctgattcagc cttggggagc agatgggttc cgtgtccggg ccaactctgaa cagatttctc 180  
 gacaatagcg tacagcgtca tcatagatcc ccctctggac ggcccgcaga cctccccatc 240  
 ggcaccagct acagcaccat caatgcgaca caaaccgcca ccgtgaccaa cggccatgcc 300  
 tcggtggtcg tcgacaaggc ccagctgagt ttctatcgcg taaatggcaa caacagtaaa 360  
 aactcgtct tgcaggagct gtggctacag acaccgggat tcccagcgcg atggtatcaa 420  
 aagagagatg caagacgggg ctttcgagcg cagttcagct tcgtcgcgcc cgacagtccc 480  
 acggagcggc tgtttgggac gggccaggat ctgacgggct ctctcatcaa gaagaatcag 540  
 accatcgact tggtaagtt caacacgctg aatccgattc cgacaatcgt gtccgatoga 600  
 gggatatctgt ttttctggaa cgtgccttcg cttgggcaga tggagttgtc tccaacgaga 660  
 accagtttcc tgagcgatca cacttcagtg gtggactatt acattgccat tcgccccgaa 720

ES 2 802 807 T3

gccgactatg acggcctgct gcagcggtag accagtgtca ctgggagatc gcccatgatt 780  
 cctgattttg gcaccggata ctggcagtgc aagctccggt atgcgactca ggaggagctc 840  
 ctcaatgtga ctatgggatt tgcagagcga aaaataacctg ttogatggt catcatcaac 900  
 tatctcagct ggtctcacga gggtgactgg gctctgaacg catcggcgtg gccagatccg 960  
 gcgaagatgg ccgcagaggc cgaactgatg gcatccgtct ggcccagtat cgaggacgcc 1020  
 agtcccaact gggctgaaat gcagagcttg ggttttgccg cgcagacatt cgatcctctc 1080  
 gatattgctg gctactggca gaacaaatac cagaacggga tccgaaactt ctggctcgac 1140  
 gaagacgagg gaggccaggt cgcgctcgac gtctaccctg ggacggacta ttacctgggt 1200  
 ccgggggatc agtatgccat gctgttcccc tactttcctc agatgggggt ttttgaagga 1260  
 cagatgaaga cagtggggtc gaataatggt tccgctgtgt ccctgtcgcg gtcctccgcc 1320  
 aactgggatt ccatgaagat gatgatatcc gcaggacaga gcatggccat gtcagggcag 1380  
 ggttggtgga ctttggacat tggaggcttc aagacagacg ggcagtcaa ctcggcgaat 1440  
 attagcgacc cggagtacca ggaattgttt gtacggtggc tgcagtgggg aaccttctta 1500  
 ccgattatgc gcaaccacgg aatgagaacc tgtctccctt cagcacagga aggcttctg 1560  
 acctgtccta gcgaaccctg gagctacggc cccgacaacc tccccatcat cgtgtcctac 1620  
 atcaacctcc gatacaaaact gcagccctac atcaaggcgc tcttccggat gctctccgag 1680  
 agtggacggg caatgatgcg gccctgttt atggatttct ctctcagcga tccttacct 1740  
 ttggaggcga cgaagatct gaaattgcag tacatgtttg ggccgcgctt gctcgtcagt 1800  
 cccgtcacia cgtatcgcgc gacgaatgcg accgtctacc tgccaaaatt gccgtcggat 1860  
 gctcccgata cgaatggac gtactggtgg acgaacgaga gttttgacgg tggatcaatgg 1920  
 gtcaccgttc ctgcgaacia aagcataatt ccgctcttcc gactcggcag cgaagcagac 1980  
 atcttgacga ggaacatcta a 2001

<210> 56

5 <211> 840

<212> ADN

10 <213> Rasamsonia emersonii

<220>

<221> fuente

15 <222> (1)..(840)

<223> /organismo="Rasamsonia emersonii" /tipo\_molar="ADN no asignado"

20 <400> 56

atgcagatgc tcctcttggt gctcctcctc cttcctocca gcctttgctg tatcatcaag 60  
 ggtatctccg ctccctctct tgcccgtctg acccagatgg ccaccctctg catgagcacc 120  
 tacctgaacg acctttgcat tgttcccggg ggtatgacca agatctccga catcaccaac 180

ES 2 802 807 T3

tccaccactg atgtccacgg ctggatcctg cgtgatgatg gtgcccgtga gatcctggcg 240  
 gtgttccgty gtaccgagtc cctccagaac tacgccaccg acaccaacta caccctggct 300  
 cccttcgaca tcttccccca gtgcgaaggc tgcgaggtcc acggtggata ctaccttgcc 360  
 tgggtgtcga ttgtcgagca ggtccagget cgcctcgagg agcagaaggc cctcttcccc 420  
 gactacggtg ttgtcttgac tggccactcc ctcggtggct ccctggctgc tcttgctgct 480  
 gcccagttct ctctctctct cgacaacatc accgtctaca ccatgggtga gcctcgcaact 540  
 ggcaacgccg ccttcgcctc cttcattgac cagcgttaca gcacttcttc tcccagagact 600  
 acccgcttct accgctgcac ccacgccgat gacggtatcc ccaacgctcc tcccacctcc 660  
 gacggctacg tccaccacgg tctggaatac tggaacctcg accccacttc tgctgagaac 720  
 acctacgtgt gcactgagga tgggtgccgtc cagtgtctcg aagcgcagaa cggcagcggg 780  
 atcaacgctg cccacctggt ctacttcggc cgccccggtg ttgttggtgg ccagtgtctg 840

<210> 57

5 <211> 233

<212> PRT

<213> Rasamsonia emersonii

10

<400> 57

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

ES 2 802 807 T3

Val Leu Thr Gly His Ser Leu Gly Gly Ser Leu Ala Ala Leu Ala Ala  
 130 135 140

Ala Gln Phe Ser Pro Leu Phe Asp Asn Ile Thr Val Tyr Thr Met Gly  
 145 150 155 160

Glu Pro Arg Thr Gly Asn Ala Ala Phe Ala Ser Phe Ile Asp Gln Arg  
 165 170 175

Tyr Ser Thr Ser Ser Pro Glu Thr Thr Arg Phe Tyr Arg Cys Thr His  
 180 185 190

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

His His Gly Leu Glu Tyr Trp Asn Leu Asp Pro Thr Ser Ala Glu Asn  
 210 215 220

Thr Phe Pro Met Asn Gly Gln Leu Ser  
 225 230

<210> 58

5 <211> 20

<212> PRT

<213> Rasamsonia emersonii

10

<400> 58

Met Gln Met Leu Leu Leu Leu Leu Leu Leu Leu Pro Pro Ser Leu Cys  
 1 5 10 15

Ala Ile Ile Lys  
 20

15 <210> 59

<211> 1100

<212> ADN

20

<213> Rasamsonia emersonii

<220>

25 <221> fuente

<222> (1)..(1100)

<223> /organismo="Rasamsonia emersonii" /tipo\_molar="ADN no asignado"

30

<400> 59

atgcaaacgc aaatgcaaat gtcctacta ctctgctgc tactaccacc ctccctctgc 60

gccatcatca aaggcatctc cgccccctc ctagcccgcc taaccagat ggccacgctc 120

tgcatgtcca cctacctcaa cgacctgtgc atcgtgcccg ggggatgac caagatcagc 180

gacatcacca acagcacaac cgacgtgcac ggctggatcc tgcgcgacga cggcgcgcgc 240

ES 2 802 807 T3

gagatcctcg ccgtgttccg cggcaccgag tcgctgcaga actacgcgac ggatacgaac 300  
 tacacgcttg cgccggttca cattttcccg cagtgtgagg gctgtgaggt ccatggcggg 360  
 tactatctcg cctgggtatc gattgtcgag caggtgcagg ctcgctcgca ggagcagaag 420  
 gcgctgtttc cggattatgg ggttgttttg acggggcata ggtatgcata taccaccta 480  
 acctaggtat aaatatactg tctgtgtata gatactgact tgatataatag tcttgaggc 540  
 tcctcgccg cgctggccgc agcacagttc tctcccctct tcgacaacat caccgtgtac 600  
 acaatgggag agcctcgcac cggcaacgcc gccttcgctt ccttcacgca ccagcgtac 660  
 agcacgtcct cgcccagagc gacgcggttc taccggtgca cgcacgcgga cgacggcatc 720  
 cccaatgccc cgccgacgtc ggacgggtat gtgcaccacg ggctggagta ctggaacttg 780  
 gacccgacga gtgcggagaa tacgtatgtc tgtacggagg acggggcggg gcagtgtgctc 840  
 gaggcgcaga atgggtcggg tatcaatgct gctcatctgg tgtatthttg tcggccggtt 900  
 gtggtggggg ggacgtgtct ttaaatagta ttgagaaggg gaatcaacag gctggatggg 960  
 atcagatttg gatattcacg aggactaaca ttagctgatg ggtgtacata caggacattt 1020  
 atataaatga ccgcagatcg aaatcaatca actaataact acagacgtgt aggtttccta 1080  
 tgaacggtca actatcttag 1100

<210> 60

5

<211> 843

<212> ADN

10 <213> Rasamsonia emersonii

<220>

<221> fuente

15

<222> (1)..(843)

<223> /organismo="Rasamsonia emersonii" /tipo\_molar="ADN no asignado"

20 <400> 60

atgcaaatgc tcctactact cctgctgcta ctaccaccct ccctctgccc catcatcaaa 60  
 ggcattctccg cccccctcct agcccgccta acccagatgg ccacgctctg catgtccacc 120  
 tacctcaacg acctgtgcat cgtgcccggg gggatgacca agatcagcga catcaccaac 180  
 agcacaaccg acgtgcacgg ctggatcctg cgcgacgacg gcgcgcgcca gatcctcgcc 240  
 gtgttccgcg gcaccgagtc gctgcagaac tacgcgacgg atacgaacta cacgcttgcg 300  
 ccgttcgaca ttttcccgca gtgtgagggc tgtgagggtc atggcgggta ctatctcgcc 360  
 tgggtatcga ttgtcgagca ggtgcaggct cgtctcgagg agcagaaggc gctgtttccg 420  
 gattatgggg ttgttttgac ggggcatagt cttggaggct ccctcgcccg gctggcccga 480  
 gcacagttct ctcccctcct cgacaacatc accgtgtaca caatgggcca gcctcgcaac 540  
 ggcaacgccg ccttcgcttc cttcatcgac cagcgtctaca gcacgtcctc gcccagagcg 600

ES 2 802 807 T3

acgcggttct accggtgcac gcacgcggac gacggcatcc ccaatgcgcc gccgacgtcg 660  
 gacgggtatg tgcaccacgg gctggagtac tggaacttgg acccgacgag tgcggagaat 720  
 acgtatgtct gtacggagga cggggcggtg cagtgtctcg aggcgcagaa tgggtcgggt 780  
 atcaatgctg ctcatctggt gtattttggt cggccggttg tgggtggggg gcagtgtctt 840  
 taa 843  
 <210> 61  
 5 <211> 1374  
 <212> ADN  
 <213> Rasamsonia emersonii  
 10 <220>  
 <221> fuente  
 15 <222> (1)..(1374)  
 <223> /organismo="Rasamsonia emersonii" /tipo\_molar="ADN no asignado"  
 20 <400> 61  
 atgaagtgc tggctacgc ttttgcctc gcctccgccc ttgctgcctc caacagcact 60  
 ggtactgcc ctgttgacct caacgtcaag cgtggacctc cctcccacct tgcctccgga 120  
 ttcactctac gcatctctga ccccccaac caggttctctg ctcaactgta caccgacatg 180  
 ggtttccgct acggccgtgc tgggtggtgct cagcttggtg ctctgctcg tggctggatt 240  
 tggggattg atgaatacca gggccgtctg aactcgacct tcagcaacta ccgcacctgc 300  
 cgtcagtacg gtgccgactt catcatcctc cccacgaca tctggggtag cgaccacgcc 360  
 aacgacagca ccactctggc tggtgacaac ggtgactgga ctgactacga caacttcata 420  
 cgccgtctga tggccgatct caaggccaac aacgccttgg atggcctggt ctgggatgta 480  
 tggaacgagc ctgacatctc catcttctgg actcgtctc agcagcagtg gattgatctc 540  
 tacatccgca cccacaagct cctccgccag gatcctgact tcgaccgct gcagatctct 600  
 ggccccacc tagcgttccg tcccttcccc aacaacacct ggtggaccaa ctggttgat 660  
 cagattgctg gcaaccagac cattcccgat cagtacgct accacctga gggtagcccc 720  
 agcagccccg atgatgacct gcagaacacc aacgcctccc tggctgcat gctgcagacc 780  
 taccgctcc cctctcgcca gatcaacatc aacgagtacg ccaccttcgc tgagcagatc 840  
 cccgccggtg ctgcctggtg gatctcccgt cttgagcgt acgaggccta cggcctgctg 900  
 ggcaactggc agtccggcac catgctccac gacctctcg ccaacctcct caccaagaag 960  
 tccgaccct ccaactacac tgccaccgac tacgtgtctg ctctgaata cccctctac 1020  
 cgctactact accgcaacat gactggccag cgtgtgcaga ccaactgatc cgaggacctg 1080  
 ctggttgatg tctacgctac tgttgacaag gacaaggtcc gtatcctggc tgggtgtccgt 1140  
 ctggccactg gtacctggca ggtcaccgtc aacagcctt ctgctggtg attgccctcc 1200

ES 2 802 807 T3

gctggtacc tcaacatcca gacctggggt ttcgatggtg actccgtctg ggaagaggtt 1260  
gaccgccccg aggatcttgg tgtcactgct cacacctact ctggcgactc cgtcaccttc 1320  
cccatctacc agaccgacaa ccacaccgcc tgggcgttcg agttcgatgt agct 1374

<210> 62

5 <211> 442

<212> PRT

<213> *Rasamsonia emersonii*

10

<400> 62

Ser Asn Ser Thr Gly Thr Ala Thr Val Asp Leu Asn Val Lys Arg Gly  
1 5 10 15

Pro Pro Ser His Leu Ala Ser Gly Phe Ile Tyr Gly Ile Pro Asp Thr  
20 25 30

Pro Asn Gln Val Pro Ala His Trp Tyr Thr Asp Met Gly Phe Arg Tyr  
35 40 45

Gly Arg Ala Gly Gly Ala Gln Leu Gly Ala Pro Ala Arg Gly Trp Ile  
50 55 60

Trp Gly Ile Asp Glu Tyr Gln Gly Arg Leu Asn Ser Thr Leu Ser Asn  
65 70 75 80

Tyr Arg Thr Cys Arg Gln Tyr Gly Ala Asp Phe Ile Ile Leu Pro His  
85 90 95

Asp Ile Trp Gly Thr Asp His Ala Asn Asp Ser Thr Ile Trp Pro Gly  
100 105 110

Asp Asn Gly Asp Trp Thr Asp Tyr Asp Asn Phe Ile Arg Arg Leu Met  
115 120 125

Ala Asp Leu Lys Ala Asn Asn Ala Leu Asp Gly Leu Val Trp Asp Val  
130 135 140

Trp Asn Glu Pro Asp Ile Ser Ile Phe Trp Thr Arg Ser Gln Gln Gln  
145 150 155 160

Trp Ile Asp Leu Tyr Ile Arg Thr His Lys Leu Leu Arg Gln Asp Pro  
165 170 175

Asp Phe Asp Arg Val Gln Ile Ser Gly Pro Thr Leu Ala Phe Arg Pro  
180 185 190

Phe Pro Asn Asn Thr Trp Trp Thr Asn Trp Leu Asp Gln Ile Ala Gly



ES 2 802 807 T3

Met Lys Ser Leu Val Tyr Ala Leu Cys Phe Ala Ser Ala Val Ala Ala  
 1                    5                    10                    15

<210> 64

5 <211> 1377

<212> ADN

<213> Rasamsonia emersonii

10

<220>

<221> fuente

15 <222> (1)..(1377)

<223> /organismo="Rasamsonia emersonii" /tipo\_molar="ADN no asignado"

<400> 64

20

```

atgaaaagtc tggctatgct cctgtgcttc gccagcgctg ttgctgctag caacagcaca      60
ggcacggcga cagtgcacct caatgtcaag cgtggcccgc cctctcacct cgcttcaggc      120
ttcatctacg gcatccccga taccccaac caggtccccg cccactggta caccgacatg      180
ggcttccgat acggccgcgc aggcgggtgcg cagctagggg ccccgcgcgc aggctggatc      240
tggggcatcg acgagtacca gggccgtctc aactcgacc cgtggaatta ccgcacctgt      300
cgtcagtacg ggcgcgactt tatcatcttg ccgcacgaca tctggggcac cgaccacgcc      360
aacgactcaa caatctggcc cggcgacaac ggcgactgga ccgactacga caacttcatc      420
cgccgactga tggccgacct gaaggcgaac aacgccctgg acggcctcgt ctgggatggt      480
tggaacgagc ccgacatctc catcttctgg acgcgcagcc agcagcagtg gatcgatctg      540
tacatccgca cgcacaagct tctccggcag gaccccgact ttgaccgtgt tcagatctca      600
ggtcccacc cggccttccg tccttcccc aacaatacct ggtggacgaa ctggctggac      660
cagattgcag gtaaccagac tatcccggac cagtacgcct atcatctgga gggcgatcca      720
tcctcgccgg atgatgacct ccagaacacc aacgcgtccc tggcggcgat gctgcagacg      780
taccggttc cgtcgcgga gatcaacatc aacgagtacg cgacattcgc agagcagatt      840
cccgcaggcg cggcatggtg gatctcccgg ctcgagcgct atgagcgta cgggctgcgc      900
gggaactggc agagcggaac catgctgcat gatctttttg ctaatctcct gaccaagaag      960
tcggaccctt ccaactacac ggcgacagac tacgtatcgg cgccggagta cccggtgtac     1020
aggtactact accggaacat gacaggccag cgggtgcaga cgacgggctc cgaggatcgc     1080
ctgctggacg tgtatgccac tgtcgacaag gacaagggtc gtatcttggc ggggtgtcgc     1140
ctggccactg gcacctggca agtcacggtc aattcactca gcgccgtggg cctaccagtc     1200
gcggggacgt tgaacatcca gacctgggtg tttgatggag actcgggtgtg ggaggaggtc     1260
gaccgtccag aggatcttgg tgttaccgcg catacctaca gcggggattc ggtgacgttt     1320
ccaatctacc agacggacaa tcatacggcg tgggcatttg aattcgacgt tgcttga      1377
  
```

25 <210> 65

<211> 1377

<212> ADN

5 <213> *Rasamsonia emersonii*

<220>

<221> fuente

10 <222> (1)..(1377)

<223> /organismo="Rasamsonia emersonii" /tipo\_molar="ADN no asignado"

15 <400> 65

```

atgaaaagtc tggctctatgc cctgtgcttc gccagcgtg ttgctgctag caacagcaca      60
ggcacggcga cagtcgacct caatgtcaag cgtggccgc cctctcacct cgcttcaggc      120
ttcatctacg gcatccccga tacccecaac caggtccccg cccactggta caccgacatg      180
ggcttccgat acggccgcgc aggcggtgcg cagctagggg cccccgcgcg aggctggatc      240
tggggcatcg acgagtacca gggccgtctc aactcgacct tgtogaatta ccgcacctgt      300
cgtcagtacg gcgccgactt tatcatcttg ccgcacgaca tctggggcac cgaccacgcc      360
aacgactcaa caatctggcc cggcgacaac ggcgactgga ccgactacga caacttcatc      420
cgccgactga tggccgacct gaaggcgaac aacgccctgg acggcctcgt ctgggatggt      480
tggaacgagc ccgacatctc catcttctgg acgcgcagcc agcagcagtg gatcgatctg      540
tacatccgca cgcacaagct tctccggcag gaccccgact ttgaccgtgt tcagatctca      600
ggtcccaccc tggccttccg tcccttcccc aacaatacct ggtggacgaa ctggctggac      660
cagattgcag gtaaccagac tatcccggac cagtacgcct atcatctgga gggcgatcca      720
tcctcgccgg atgatgacct ccagaacacc aacgcgtccc tggcggcgat gctgcagacg      780
taccggcttc cgtcgcgga gatcaacatc aacgagtacg cgacattcgc agagcagatt      840
cccgaggcgc cggcatgggt gatctcccgg ctcgagcgtc atgagggcgt cgggctgcgc      900
gggaactggc agagcggaac catgctgcat gatctttttg ctaatctcct gaccaagaag      960
tcggaccctt ccaactacac ggcgacagac tacgtatcgg cgccggagta cccggtgtac     1020
aggatctact accggaacat gacaggccag cgggtgcaga cgacgggctc cgaggatcgc     1080
ctgctggacg tgtatgccac tgtcgacaag gacaagggtg gtatcttggc ggggtgtcgc     1140
ctggccactg gcacctggca agtcacggtc aattcactca gcgccgtggg cctaccagt     1200
gcggggacgt tgaacatcca gacctgggt tttgatggag actcgggtgt ggaggaggtc     1260
gaccgtccag aggatcttgg tgttaccgcg catacctaca gcggggattc ggtgacgttt     1320
ccaatctacc agacggacaa tcatacggcg tgggcatttg aattcgacgt tgcttga       1377

```

20 <210> 66

<211> 1425

25 <212> ADN

<213> *Rasamsonia emersonii*

ES 2 802 807 T3

<220>

<221> fuente

5

<222> (1)..(1425)

<223> /organismo="Rasamsonia emersonii" /tipo\_molar="ADN no asignado"

10

<400> 66

```

atgatctccc tgctggccct tgctgctggt cccatccccg tcctctctgc tccctcccgc      60
ccccacgctc acatctccgt caacaccagc actcgccttc agcagggtga cggcttcggt      120
ttcagcgaag cgttccagcg tgctgaggat atctacggca aggacggtct ttctcctgag      180
aaccgtacct gtgtcctcga cctcctcttc tccgacacc acggtgctgg tctgaccatt      240
gtccgcaacg gcattggcag cagcaacagc agcatcaagg acttcatgaa ctccattgag      300
cccttctccc ccggtcccc ctctgctcct cctcactacg tgtgggaccg caacgactct      360
ggccaggctc ggctatccca cgaagccgcc tcctacgggtg tcaacacctt ctacgccgat      420
gcctggtccg ctcttgata catgaaaacc aacgatgatg actccaacgg tggatacctc      480
tgccggtgca ccaacacttc ttgcgcttcc ggtgactgga agcaggccta cgccaactac      540
cttggtcagt acgtgcgctt ctaccagcag gtcggtatca aggtcaccca cctcggattc      600
ctgaacgagc cccaggaaga tgtcacctac gcctccatgc tctccgacgg caccaggtc      660
gccgacttca tcaaggtcct tgccccacc gtcaaggccg ccggtctgga tgtcaagctc      720
acctgctgcg atggtggttg ctgggaggag cagcgtgcc a tgcttcctgg tctgcaggct      780
ggtggtcccg agcactccgc tgagagctac ctgtccgtca tcaactgctca cggatacaac      840
agccctccca ccaactcctc cgagacttct ctgcccgtgt ggatgactga atgggcggat      900
ctcaacggca actacactgc tgccctgtac gagaacggcg ctgctggtga aggcttgacc      960
tgggccaacc gtatccagga tgccctcacc cgctccaacg tgtctgcttt cctgcaactgg     1020
attggtgctg agaacggcac cagcaactcg cctctgatca acctcaacgg tgacagctac     1080
gtggccacca agcgtctgtg ggcgttcggc cagttctccc gcttcgtccg tcccgtgccc     1140
gtccgtatcg atgcccgttc ctccgactct ttagtgaccg tttccgcttt ccagaacaag     1200
gacaacgggt ttgttgccac tcaggccatc aacaacggcg acaccgacta cgagggttgag     1260
gttgagttga ctggctacgg tctgtctccc gtcattccagc cctacctcac caacaacgag     1320
aacgacctcg aggctccaa gcccatcatt gctctcggcg tcggtgctgc caccaagttc     1380
acctccatcg tcccgcggcg ctctcttgtc tccttcgtat ctaag                          1425

```

<210> 67

15

<211> 459

<212> PRT

20

<213> Rasamsonia emersonii

<400> 67

ES 2 802 807 T3

Ala Pro Ser Arg Pro His Ala His Ile Ser Val Asn Thr Gln Thr Arg  
 1 5 10 15

Phe Gln Gln Val Asp Gly Phe Gly Phe Ser Glu Ala Phe Gln Arg Ala  
 20 25 30

Glu Asp Ile Tyr Gly Lys Asp Gly Leu Ser Pro Glu Asn Arg Thr Arg  
 35 40 45

Val Leu Asp Leu Leu Phe Ser Asp Thr His Gly Ala Gly Leu Thr Ile  
 50 55 60

Val Arg Asn Gly Ile Gly Ser Ser Asn Ser Ser Ile Lys Asp Phe Met  
 65 70 75 80

Asn Ser Ile Glu Pro Phe Ser Pro Gly Ser Pro Ser Ala Pro Pro His  
 85 90 95

Tyr Val Trp Asp Arg Asn Asp Ser Gly Gln Val Trp Leu Ser His Glu  
 100 105 110

Ala Ala Ser Tyr Gly Val Asn Thr Phe Tyr Ala Asp Ala Trp Ser Ala  
 115 120 125

Pro Gly Tyr Met Lys Thr Asn Asp Asp Asp Ser Asn Gly Gly Tyr Leu  
 130 135 140

Cys Gly Val Thr Asn Thr Ser Cys Ala Ser Gly Asp Trp Lys Gln Ala  
 145 150 155 160

Tyr Ala Asn Tyr Leu Val Gln Tyr Val Arg Phe Tyr Gln Gln Val Gly  
 165 170 175

Ile Lys Val Thr His Leu Gly Phe Leu Asn Glu Pro Gln Glu Asp Val  
 180 185 190

Thr Tyr Ala Ser Met Leu Ser Asp Gly Thr Gln Ala Ala Asp Phe Ile  
 195 200 205

Lys Val Leu Ala Pro Thr Val Lys Ala Ala Gly Leu Asp Val Lys Leu  
 210 215 220

ES 2 802 807 T3

Thr Cys Cys Asp Gly Val Gly Trp Glu Glu Gln Arg Ala Met Leu Pro  
225 230 235 240

Gly Leu Gln Ala Gly Gly Pro Glu His Ser Ala Glu Ser Tyr Leu Ser  
245 250 255

Val Ile Thr Ala His Gly Tyr Asn Ser Pro Pro Thr Thr Pro Leu Glu  
260 265 270

Thr Ser Leu Pro Val Trp Met Thr Glu Trp Ala Asp Leu Asn Gly Asn  
275 280 285

Tyr Thr Ala Ala Trp Tyr Glu Asn Gly Ala Ala Gly Glu Gly Leu Thr  
290 295 300

Trp Ala Asn Arg Ile Gln Asp Ala Phe Thr Arg Ser Asn Val Ser Ala  
305 310 315 320

Phe Leu His Trp Ile Gly Ala Glu Asn Gly Thr Ser Asn Ser Pro Leu  
325 330 335

Ile Asn Leu Asn Gly Asp Ser Tyr Val Ala Thr Lys Arg Leu Trp Ala  
340 345 350

Phe Gly Gln Phe Ser Arg Phe Val Arg Pro Gly Ala Val Arg Ile Asp  
355 360 365

Ala Val Ser Ser Asp Ser Leu Val Thr Val Ser Ala Phe Gln Asn Lys  
370 375 380

Asp Asn Gly Val Val Ala Thr Gln Ala Ile Asn Asn Ala Asp Thr Asp  
385 390 395 400

Tyr Glu Val Glu Val Glu Leu Thr Gly Tyr Gly Pro Val Ser Val Ile  
405 410 415

Gln Pro Tyr Leu Thr Asn Asn Glu Asn Asp Leu Glu Ala Ser Lys Pro  
420 425 430

Ile Ile Ala Leu Ala Val Gly Ala Ala Thr Lys Phe Thr Ser Ile Val  
435 440 445

Pro Ala Arg Ser Leu Val Ser Phe Val Ser Lys  
450 455

<210> 68

5 <211> 16

<212> PRT

<213> Rasamsonia emersonii

10

<400> 68

Met Ile Ser Leu Leu Ala Leu Ala Ala Val Pro Ile Pro Val Leu Ser  
1 5 10 15

ES 2 802 807 T3

<210> 69

<211> 1428

5 <212> ADN

<213> *Rasamsonia emersonii*

10 <220>

<221> fuente

<222> (1)..(1428)

15 <223> /organismo="Rasamsonia emersonii" /tipo\_molar="ADN no asignado"

<400> 69

atgatctctc	tcctcgcggt	ggccgctgtg	ccaattcctg	tcctgtccgc	cccgtctaga	60
cctcacgccc	acatctccgt	caatacacia	acccgcttcc	agcaggtcga	cggtatcggc	120
ttctccgagg	ccttccagcg	cgcagaggac	atctacggca	aggatggctt	gtcgcccag	180
aaccgcacca	gagtgtctga	tctcctgttc	agcgacaccc	atggcgccgg	cctgaccatc	240
gtccgcaatg	gcctcggctc	cagcaatagc	tcctcaagg	acttcatgaa	ctccatcgag	300
cccttcagtc	ccggctcggc	gtctgcgccc	ccgcactacg	tctgggaccg	caacgacagc	360
ggccaggtgt	ggctgtccca	cgaagcggcc	tcgatatggc	tcaacacggt	ctatgccgac	420
gcctggtctg	ctcccgggta	catgaagacc	aacgacgacg	actccaacgg	cgggtacctg	480
tgccgctga	ccaacacgct	ctgcgcctcc	ggggactgga	agcaggccta	cgccaactac	540
ctcgtccagt	acgtccggtt	ctaccaacag	gtcggcatca	aggtcaccca	cctgggtttc	600
ctgaacgagc	cccaggaaga	cgctcacctat	gccagcatgc	tgtcggacgg	caccagggcg	660
gccgacttca	tcaaagtctt	agcaccacc	gtcaaggcgg	caggactcga	cgtcaagctc	720
acctgctgcg	acggcgtcgg	gtgggaggag	cagcgagcga	tgctgccggg	cctccaggct	780
ggaggtcccg	agcacagcgc	ggaaagctat	ctgtctgtga	tcaccgcca	tggctacaac	840
tccccgcca	caacgcctct	tgagacctcg	ttgccagtgt	ggatgaccga	gtgggcccgat	900
ctgaatggga	actatacggc	cgcctgttac	gagaacggcg	ccgccggcga	gggtctgacc	960
tgggcgaacc	gcatccaaga	cgccttcacg	cggagcaacg	tcagcgcctt	tctgcactgg	1020
atcggagcgg	agaatggcac	gagcaacagc	cctctcatca	acctgaacgg	cgattcgtat	1080
gtcgcgacca	agcggctgtg	ggcgttcggc	cagttcagtc	gattcgtccg	gccgggtgcc	1140
gtccgcatcg	acgcggctcag	ttcggactcc	ttggtcacgg	tgagcgcggt	ccagaacaag	1200
20 gataacggcg	tggtgccac	gcaggccatc	aacaatgccg	acacagacta	tgaggtcgag	1260
gtggagtga	cgggttatgg	acccgtgtcc	gttatccagc	cgtatctgac	gaacaatgaa	1320
aacgacctgg	aagcctcaa	gccgatcatc	gcgcttgctg	ttggggcagc	aaccaagttc	1380
acgagtatcg	ttccggcgag	gtccttggtt	agtttcgtaa	gtaaataag		1428

<210> 70

25

ES 2 802 807 T3

<211> 1428

<212> ADN

5 <213> *Rasamsonia emersonii*

<220>

<221> fuente

10 <222> (1)..(1428)

<223> /organismo="Rasamsonia emersonii" /tipo\_molar="ADN no asignado"

15 <400> 70

```

atgatctctc tcctcgcggt ggccgctgtg ccaattcctg tcctgtccgc cccgtctaga      60
cctcacgccc acatctccgt caatacacia acccgcttcc agcaggtcga cggattcggc      120
ttctccgagg ccttcagcgc cgcagaggac atctacggca aggatggctt gtcgcccgag      180
aaccgcacca gagtgctcga tctcctgttc agcgacaccc atggcgccgg cctgaccatc      240
gtccgcaatg gcatcggctc cagcaatagc tccatcaagg acttcatgaa ctccatcgag      300
cccttcagtc ccggtcgcgc gtctgcgccc ccgcactacg tctgggaccg caacgacagc      360
ggccagggtg ggctgtccca cgaagcggcc tcgatggcg tcaacacggt ctatgccgac      420
gcctggtctg ctcccgggta catgaagacc aacgacgacg actccaacgg cgggtacctg      480
tgcggcgtga ccaacacgtc ctgcgccctc ggggactgga agcaggccta cgccaactac      540
ctcgtccagt acgtccggtt ctaccaacag gtcggcatca aggtcaccca cctgggtttc      600
ctgaacgagc cccaggaaga cgtcacctat gccagcatgc tgtcggacgg cacccaggcg      660
gccgacttca tcaaagttct agcaccacc gtcaaggcgg caggactcga cgtcaagctc      720
acctgctgcg acggcgtcgg gtgggaggag cagcgagcga tgctgccggg cctccaggct      780
ggaggtcccg agcacagcgc ggaaagctat ctgtctgtga tcaccgcca tggctacaac      840
tccccgcca caacgcctct tgagacctcg ttgccagtgt ggatgaccga gtgggccgat      900
ctgaatggga actatacggc cgcctgttac gagaacggcg ccgccggcga gggctctgacc      960
tgggcgaacc gcatccaaga cgccttcacg cggagcaacg tcagcgcctt tctgcactgg     1020
atcggagcgg agaatggcac gagcaacagc cctctcatca acctgaacgg cgattcgtat     1080
gtcgcgacca agcggctgtg ggcgttcggc cagttcagtc gattcgtccg gccgggtgcc     1140
gtccgcatcg acgcggctcag ttcggactcc ttggtcacgg tgagcgcggt ccagaacaag     1200
gataacggcg tggtgccac gcaggccatc aacaatgccg acacagacta tgaggtcgag     1260
gtggagtga cgggttatgg acccgtgtcc gttatccagc cgtatctgac gaacaatgaa     1320
aacgacctgg aagcctcaa gccgatcatc gcgcttgctg ttggggcagc aaccaagttc     1380
acgagtatcg ttccggcgag gtccttggtt agtttogtaa gtaaatag                       1428

```

20 <210> 71

<211> 3036

25 <212> ADN

ES 2 802 807 T3

<213> Rasamsonia emersonii

<220>

5 <221> fuente

<222> (1)..(3036)

<223> /organismo="Rasamsonia emersonii" /tipo\_molar="ADN no asignado"

10

<400> 71

```

atgacacctga gctctctcct cgtcggctgc gtatctctct ggggtgctgc ccacgccctt      60
ggccagcagc ccgttggtgc cttcacctcc tccccggat acctcaagct ggctggcccc      120
ggtaactcctc ccggtaccgt tgtcctcgac tctgccgact ggcccgctgt cctgcgtgct      180
gccggtgatc ttgctgtcga cttcggctgt gtcaccggca ccaacctgac caagactgtc      240
atcaacggta ccaccaccag ctccggctctt gctgctgttt ccaacaaggg ccccgtcatc      300
attgctggta ccattggtaa cagcagcttg atcaacgccc ttgccagtc cggcaagatc      360
gatgtctccg ccaactgaggg ccgctgggag gctttccaga ccgagattgt ggacaacccc      420
ttcccagga tatctcgtgc tctggctggt tccggctccg accgtcgtgg aactgtctac      480
ggcttgtagc acatctccga gcagatcggg gtctccccct ggtactggtt cgccgatggt      540
cctcctgctc agcacgagga gatctacgcc ctggacacca agaagatcca gggttcccc      600
tccgtcaagt accgtggtct gttcatcaac gacgaggctc ccgccctcac caactggatc      660
aacgtcaact accctccttg caagtacggg cctggcttca acgcccactt ctacgcccac      720
gttttcgagc tcttgctgcg cctccgtgcc aactacctct ggccctgctga atggagcaac      780
atcttcgccc ttgacgaccc ccgcaacttc cctactgccg atctctacgg cattgtcatt      840
ggcacctccc acaccgagcc cttgatgcgc tggaccctgg aacaatcgtc actactccag      900
ggcccttggg actggctgac caacgagaag aacatctacc agttcctcaa ggaaggtggt      960
gagcgtcca agaactacga agtcatctac accatgggta tgcgtggtct tggtgacact     1020
gcctccccc ccatcaacgc caccactctg gctcagatcg tcgctgctga ggagcagatc     1080
ctctccgagg tgttaacac caccaacatc tcctggatcc cccagatgtg gtgcctctac     1140
aaggaagtgc gtggttacta cgaggatggt ctgctgtgtg ccgatgacat cactctcctc     1200
tgggctgatg acaactgggg taacattgag agacttocca tcggtaacga gactgctcgc     1260
tccggtggtg ctggtgtcta ctaccacttc gactacgtcg gtgaccccca ggactacaag     1320
tggatcaaca ccatctccct ccagaaaacc tgggagcaga tgcacctggc ctacgagcgt     1380

```

ES 2 802 807 T3

caggctcgca acatctggat tgtcaacggt ggtgacctga aggcccttga gatccccacc 1440  
aaccacttct tcgaccttgc ctacgacacc cccagctggt cagatcccga cagcacctcc 1500  
cgctggctca agctctgggc tactcgtgag ttcggtgctg ctggtgccga caaggtcgcc 1560  
gatatcgtca accgctacgg ccagtagcct gcccgctgca agttcgagat gatcaccccc 1620  
tccaccttct ccatcatcaa ctacaacgag gccacaacg tcgtgtctca gtgggaatct 1680  
ctggtcaacg acgcccgtga tgtctaccgt ggcctgtccg aggctgctcg tcccccttc 1740  
ttcgagctgg tcctccagcc ctgcatggct ggtcacatca tcaccagat ccacgtcact 1800  
gctgccaaaga acaacctgta cgcctcccag cgcgcactt ccgccaaggc catggccgat 1860  
aaggccctga agcttttcaa ggaagatcac gacctcaccg tgagctacca caagattctt 1920  
gacggcaagt ggaaccacat catggaccag acccacctgg gctacctcta ctggcagcag 1980  
cccatgcgca acacctccc tcccctccag tacaccagc tgttgaaga ttcgctagcg 2040  
ggacctatgg gtgtttccgt tgagggcagc aacgcctccg ttcccgggtga tgaccagtac 2100  
cacgctctct ccagcaccac ctgactctt cctcccattg acccctacgg cccttctact 2160  
cgttggattg agatctactc tcgctcgacc gagtcttcg agttccagat atctcccaat 2220  
gtcagctggg tgaccgccac tcccagctct ggcactgtcc acgtttctgg caacggcacc 2280  
accaccgatg ttcgcgttca gctctccgtt gactgggaca aggctcccat tggaagctct 2340  
ctcgtcagca tcaacgtttc catccccact ggccacttcc ccaccgggtga ttacgtaac 2400  
tccgccatgc ccgtcgtcca gctccccgtc aacaagactg ttgttctctc ttctttccac 2460  
ggtttcgtgg aatcggatgg tgtcgtttct atcgaagctc ctcacgccac ccgcaacacc 2520  
agcactgcc aacgtcagta cgcctcctc cccggtacg gccgcaccct gtccggtgtc 2580  
accctcctgc ctgtcaacgc cccctcccag aaggctcctt ctcccctcg tctggagtac 2640  
gacatgtacc tcttctctcc cgcgcacaac ggatcgaccg tcaacatcac cctctacctc 2700  
ggtccctccc tgaacgtcaa ccccgaccgc cctctccgtt acgctcttgc tttcgatgac 2760  
gatgacaacc cccaggttgt ccagtagctg cccagcaccg tcctcggcac ctaccacct 2820  
gactggccct ctgctgtctc caacaatgtg tgggcctcca ccacctcta caccatcaag 2880  
ggtttcaacg gcaactccac ttctgccact ggcaccggtt ccgctcacac tctcaagctg 2940  
tgggctctgg agcccgggtg tgttttcgag aagatcgtca ttgaccttgg tggtcagcgt 3000  
gacagctacc ttggtcctcc tgagtccagg atagta 3036

<210> 72

5 <211> 991

<212> PRT

<213> *Rasamsonia emersonii*

10

<400> 72

ES 2 802 807 T3

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

Ala Gly Pro Gly Thr Pro Pro Gly Thr Val Val Leu Asp Ser Ala Asp  
 20 25 30

Trp Pro Ala Val Leu Arg Ala Ala Gly Asp Leu Ala Val Asp Phe Gly  
 35 40 45

Arg Val Thr Gly Thr Asn Leu Thr Lys Thr Val Ile Asn Gly Thr Thr  
 50 55 60

Thr Ser Ser Gly Leu Ala Ala Val Ser Asn Lys Gly Pro Val Ile Ile  
 65 70 75 80

Ala Gly Thr Ile Gly Asn Ser Ser Leu Ile Asn Ala Leu Ala Gln Ser  
 85 90 95

Gly Lys Ile Asp Val Ser Ala Thr Glu Gly Arg Trp Glu Ala Phe Gln  
 100 105 110

Thr Glu Ile Val Asp Asn Pro Phe Pro Gly Ile Ser Arg Ala Leu Val  
 115 120 125

Val Ser Gly Ser Asp Arg Arg Gly Thr Val Tyr Gly Leu Tyr Asp Ile  
 130 135 140

Ser Glu Gln Ile Gly Val Ser Pro Trp Tyr Trp Phe Ala Asp Val Pro  
 145 150 155 160

Pro Ala Gln His Glu Glu Ile Tyr Ala Leu Asp Thr Lys Lys Ile Gln  
 165 170 175

Gly Ser Pro Ser Val Lys Tyr Arg Gly Leu Phe Ile Asn Asp Glu Ala  
 180 185 190

Pro Ala Leu Thr Asn Trp Ile Asn Val Asn Tyr Pro Pro Cys Lys Tyr  
 195 200 205

Gly Pro Gly Phe Asn Ala Asp Phe Tyr Ala His Val Phe Glu Leu Leu  
 210 215 220

Leu Arg Leu Arg Ala Asn Tyr Leu Trp Pro Ala Glu Trp Ser Asn Ile  
 225 230 235 240

Phe Ala Leu Asp Asp Pro Arg Asn Phe Pro Thr Ala Asp Leu Tyr Gly  
 245 250 255

ES 2 802 807 T3

Ile Val Ile Gly Thr Ser His Thr Glu Pro Leu Met Arg Trp Thr Leu  
 260 265 270

Glu Gln Ser Leu Leu Leu Gln Gly Pro Trp Asn Trp Leu Thr Asn Glu  
 275 280 285

Lys Asn Ile Tyr Gln Phe Leu Lys Glu Gly Val Glu Arg Ser Lys Asn  
 290 295 300

Tyr Glu Val Ile Tyr Thr Met Gly Met Arg Gly Leu Gly Asp Thr Ala  
 305 310 315 320

Ser Pro Thr Ile Asn Ala Thr Thr Leu Ala Gln Ile Val Ala Ala Glu  
 325 330 335

Glu Gln Ile Leu Ser Glu Val Phe Asn Thr Thr Asn Ile Ser Trp Ile  
 340 345 350

Pro Gln Met Trp Cys Leu Tyr Lys Glu Val Gly Gly Tyr Tyr Glu Asp  
 355 360 365

Gly Leu Arg Val Ala Asp Asp Ile Thr Leu Leu Trp Ala Asp Asp Asn  
 370 375 380

Trp Gly Asn Ile Glu Arg Leu Pro Ile Gly Asn Glu Thr Ala Arg Ser  
 385 390 395 400

Gly Gly Ala Gly Val Tyr Tyr His Phe Asp Tyr Val Gly Asp Pro Gln  
 405 410 415

Asp Tyr Lys Trp Ile Asn Thr Ile Ser Leu Gln Lys Thr Trp Glu Gln  
 420 425 430

Met His Leu Ala Tyr Glu Arg Gln Ala Arg Asn Ile Trp Ile Val Asn  
 435 440 445

Val Gly Asp Leu Lys Ala Leu Glu Ile Pro Thr Asn His Phe Phe Asp  
 450 455 460

Leu Ala Tyr Asp Thr Pro Ser Trp Ser Asp Pro Asp Ser Thr Ser Arg  
 465 470 475 480

Trp Leu Lys Leu Trp Ala Thr Arg Glu Phe Gly Ala Ala Gly Ala Asp  
 485 490 495

Lys Val Ala Asp Ile Val Asn Arg Tyr Gly Gln Tyr Ala Ala Arg Arg

# ES 2 802 807 T3

500	505	510
Lys Phe Glu Met Ile Thr Pro Ser Thr Phe Ser Ile Ile Asn Tyr Asn 515 520 525		
Glu Ala Asn Asn Val Val Ser Gln Trp Glu Ser Leu Val Asn Asp Ala 530 535 540		
Arg Asp Val Tyr Arg Gly Leu Ser Glu Ala Ala Arg Pro Ala Phe Phe 545 550 555 560		
Glu Leu Val Leu Gln Pro Cys Met Ala Gly His Ile Ile Thr Gln Ile 565 570 575		
His Val Thr Ala Ala Lys Asn Asn Leu Tyr Ala Ser Gln Arg Arg Thr 580 585 590		
Ser Ala Lys Ala Met Ala Asp Lys Ala Leu Lys Leu Phe Lys Glu Asp 595 600 605		
His Asp Leu Thr Val Ser Tyr His Lys Ile Leu Asp Gly Lys Trp Asn 610 615 620		
His Ile Met Asp Gln Thr His Leu Gly Tyr Leu Tyr Trp Gln Gln Pro 625 630 635 640		
Met Arg Asn Thr Leu Pro Pro Leu Gln Tyr Thr Gln Leu Leu Glu Asp 645 650 655		
Ser Leu Ala Gly Pro Met Gly Val Ser Val Glu Gly Ser Asn Ala Ser 660 665 670		
Val Pro Gly Asp Asp Gln Tyr His Ala Leu Ser Ser Thr Thr Leu Thr 675 680 685		
Leu Pro Pro Ile Asp Pro Tyr Gly Pro Ser Thr Arg Trp Ile Glu Ile 690 695 700		
Tyr Ser Arg Ser Thr Glu Ser Phe Glu Phe Gln Ile Ser Pro Asn Val 705 710 715 720		
Ser Trp Val Thr Ala Thr Pro Ser Ser Gly Thr Val His Val Ser Gly 725 730 735		
Asn Gly Thr Thr Thr Asp Val Arg Val Gln Leu Ser Val Asp Trp Asp 740 745 750		

ES 2 802 807 T3

Lys Ala Pro Ile Gly Ser Ser Leu Val Ser Ile Asn Val Ser Ile Pro  
755 760 765

Thr Gly His Phe Pro Thr Gly Asp Tyr Gly Asn Ser Ala Met Pro Val  
770 775 780

Val Gln Leu Pro Val Asn Lys Thr Val Val Pro Ala Ser Phe His Gly  
785 790 795 800

Phe Val Glu Ser Asp Gly Val Val Ser Ile Glu Ala Pro His Ala Thr  
805 810 815

Arg Asn Thr Ser Thr Ala Asn Val Ser Tyr Ala Val Ile Pro Gly Tyr  
820 825 830

Gly Arg Thr Leu Ser Gly Val Thr Leu Leu Pro Val Asn Ala Pro Ser  
835 840 845

Gln Lys Ala Pro Ser Ser Pro Arg Leu Glu Tyr Asp Met Tyr Leu Phe  
850 855 860

Ser Pro Ala Ala Asn Gly Ser Thr Val Asn Ile Thr Leu Tyr Leu Gly  
865 870 875 880

Pro Ser Leu Asn Val Asn Pro Asp Arg Pro Leu Arg Tyr Ala Leu Ala  
885 890 895

Phe Asp Asp Asp Asp Asn Pro Gln Val Val Gln Tyr Val Pro Ser Thr  
900 905 910

Val Leu Gly Thr Tyr Pro Pro Asp Trp Pro Ser Ala Val Ser Asn Asn  
915 920 925

Val Trp Ala Ser Thr Thr Ser Tyr Thr Ile Lys Gly Phe Asn Gly Asn  
930 935 940

Ser Thr Ser Ala Thr Gly Thr Gly Ser Ala His Thr Leu Lys Leu Trp  
945 950 955 960

Ala Leu Glu Pro Gly Val Val Phe Glu Lys Ile Val Ile Asp Leu Gly  
965 970 975

Gly Gln Arg Asp Ser Tyr Leu Gly Pro Pro Glu Ser Arg Ile Val  
980 985 990

<210> 73

5 <211> 21

<212> PRT

<213> Rasamsonia emersonii

10

<400> 73

ES 2 802 807 T3

Met Ile Leu Ser Ser Leu Leu Val Gly Cys Val Ser Leu Trp Gly Ala  
 1 5 10 15

Ala His Ala Leu Gly  
 20

<210> 74

5 <211> 3152

<212> ADN

10 <213> Rasamsonia emersonii

<220>

<221> fuente

15 <222> (1)..(3152)

<223> /organismo="Rasamsonia emersonii" /tipo\_molar="ADN no asignado"

<400> 74

20 atgatcctct ccagtctcct tgtggggtgc gtctccctat ggggagcagc acatgccctt 60  
 ggacaacagc ctgtcgtcgc cttcacctcc agtccaggat atctgaagct tgcaggccca 120  
 ggaactcccc ctggaacggt tgttctcgat tcagccgact ggccggcagt cttgagagca 180  
 gccggtgacc ttgccgtgga ctttgggcgt gttactggaa cgaacctgac caaacgggtg 240  
 atcaatggca ccacaacgtc gtctgggcta gcagcagtat caaacaagg ccccgatgatt 300  
 atcgccggca ccattggcaa ctccagtctc atcaatgctt tggcgcagtc cggcaagatc 360  
 gacgtctccg cgacagaggg cagatgggag gccttcaga ccgaaatcgt ggacaatcct 420  
 ttcccgggga tatcgcgggc gttggtcgtc agtggcagcg acaggcgtgg cactgtctac 480  
 gggctgtacg acatctcaga gcagatcggg gtctcgccgt ggtactggtt cgccgacgtg 540  
 ccgcccgctc agcacgagga gatatacgtc ttggacacga agaaaatcca gggctcgccg 600  
 tcggtcaaat accgcggtt gttcatcaac gacgaggcac ctgccttgac aaactggatc 660  
 aatgtgaact acccacctg caagtatggg cctgggttca acgccgactt ctacgcacac 720  
 gtgttcgaac tcctcctgag gctgagagcc aactacctgt ggccggcaga atggagcaac 780  
 atcttcgctc tggacgatcc gcgcaatttc cccacagcag acctgtacgg gattgtgatt 840  
 ggaaccagcc acacagagcc cctgatgcga tggactctgg agcagagctt gcttctccag 900  
 ggcccgtgga attggcttac gaacgagaaa aacatctacc agttcctgaa agaaggcgtc 960  
 gagaggtcca agaattacga ggtcatatat actatgggca tgagaggtct cggcgacacg 1020  
 gcatcgccca cgatcaacgc cacaactctg gcgcaaatcg tcgccgagga ggagcagatc 1080  
 ctctccgagg ttttcaacac gacgaacatc tcttgatcc cgcagatgtg gtgtctctac 1140  
 aaggaagtgg gcggctacta cgaagacggg ctccgtgtgg cggacgatat cacgctcctc 1200

ES 2 802 807 T3

tgggcagacg acaactgggg gaacatcgag agactgccga ttggcaatga gactgcacgg 1260  
 tctggcgggtg ctggcgtcta ttatcacttt gactacgtcg gggaccccca agactacaag 1320  
 tggatcaaca ctatttcatt gcagaaaacc tgggagcaga tgcatctcgc atacgagcgg 1380  
 caggcaagga atatttggat cgtgaatggt ggcgacttga aagcattggt tagcaagcct 1440  
 cttccgtggg aatatccgtc atatattgac aagatgacta ggaaatcccg actaaccact 1500  
 tcttcgactt ggccctacgat acaccagct ggtccgatcc ggacagcacc agccggtggc 1560  
 tcaagctctg ggccactcgt gaattcgggg cagcgggggc cgacaaagtc gcagacatcg 1620  
 taaaccgcta cggccagtac gcagctagac gcaaattcga gatgatcacc ccgagcacgt 1680  
 tcagcatcat caactacaac gaggctaaca acgtggtgag tcaatgggag agtctggtca 1740  
 acgatgcccc tgatgtgtac cgcggactca gcgaagctgc ccggccggcg ttcttcgagc 1800  
 tggatttgca accttgcatg gctggccaca tcatcactca gatccatggt acggccgcca 1860  
 agaataatct ctatgcgtcc cagagacgga caagcgcgaa ggcaatggca gacaaggcgt 1920  
 tgaaactggt caaagaggac cacgacttga cggtttccta tcacaagatt cttgatggaa 1980  
 aatggaacca tatcatggac cagacgcatc tgggttacct ttactggtga gtatTTTTTT 2040  
 ttactaataa ccttgaccga ggttttgctg aaacggattc gtataggcag cagccgatga 2100  
 gaaacacact cccccctctc caatacactc agctgttaga ggacagtttg gcgggtccaa 2160  
 tgggtgtttc tgtcgagggc agcaatgcct ctgtccctgg cgacgatcaa taccacgctt 2220  
 tgtcgagtac caccctgacc ctgccacca tagatcccta cggaccgtcg acgcgctgga 2280  
 tcgagatcta ctctcgetca actgaaagct tcgagttcca aatctcccc aacgtctcct 2340  
 gggtcaccgc gacaccgtcc tccggcaccg tccatgttag tggaaacgga acgaccaccg 2400  
 acgtcccgct gcagctgtcc gtcgactggg acaaggctcc tateggctcg agcctcgtct 2460  
 cgatcaacgt cagcatccca accggccact tcccgcggg ggactacggg aactcggcca 2520  
 tgccggttgt ccagctcccc gtgaacaaga cggctcgtcc agcgtccttc cacggcttcg 2580  
 tcgagtcgga cggcgtcgtc agcatcgagg cgccgcacgc aacgcgcaac acctcgacgg 2640  
 cgaacgtctc ctacgccgtg atcccgggg acggacgcac actgtcgggc gtcaccttgc 2700  
 tgccggtgaa cgcgccgtca cagaagcgc cttcctcccc gcgtctcgag tacgacatgt 2760  
 acctcttcag ccctgcggcc aacggcagca ccgtcaacat aacgtgtac ctgggaccgt 2820  
 cactcaacgt caaccctgac cgaccgctgc gatacgcgtt ggccttcgac gacgacgaca 2880  
 atccccaggt agtgcagtac gtcccagca cgggtcgtgg cacctacca ccggactggc 2940  
 cctcggccgt gtccaacaac gtgtgggctg cgacgacgtc gtatacaatc aagggcttta 3000  
 acggaacag cactagtgca acaggaacgg gatccgcgca cacgctcaag ctgtgggctc 3060  
 ttgacctgg cgtggtgttt gagaagattg tcattgatct gggcgggcag agggatagtt 3120  
 atctcgttcc gccggagagt cggatttgtg ag 3152

5 <210> 75  
 <211> 3039  
 <212> ADN

10

ES 2 802 807 T3

<213> Rasamsonia emersonii

<220>

5 <221> fuente

<222> (1)..(3039)

<223> /organismo="Rasamsonia emersonii" /tipo\_molar="ADN no asignado"

10

<400> 75

```

atgacacctt ccagctctct tgtgggggtgc gtctccctat ggggagcagc acatgccctt      60
ggacaacagc ctgtcgtcgc cttcacctcc agtccaggat atctgaagct tgcaggccca      120
ggaactcccc ctggaacggt tgttctcgat tcagccgact ggccggcagt cttgagagca      180
gccggtgacc ttgccgtgga ctttgggcgt gttactggaa cgaacctgac caaaacggtg      240
atcaatggca ccacaacgct gtctgggcta gcagcagtat caaacaaggg ccccgtgatt      300
atcgccggca ccattggcaa ctccagtctc atcaatgctt tggcgcagtc cggcaagatc      360
gacgtctccg cgacagaggg cagatgggag gccttcaga ccgaaatcgt ggacaatcct      420
ttcccgggga tatcgggggc gttggtcgtc agtggcagcg acagggcgtg cactgtctac      480
gggctgtacg acatctcaga gcagatcggg gtctcggcgt ggtactggtt cgccgacgtg      540
ccgcccgctc agcacgagga gatatacgtt ttggacacga agaaaatcca gggctcggcg      600
tcggtcaaat accgcggtt gttcatcaac gacgaggcac ctgccttgac aaactggatc      660
aatgtgaact acccaccctg caagtatggg cctgggttca acgccgactt ctacgcacac      720
gtgttcgaac tcctcctgag gctgagagcc aactacctgt ggccggcaga atggagcaac      780
atcttcgctc tggacgatcc gcgcaatttc cccacagcag acctgtacgg gattgtgatt      840
ggaaccagcc acacagagcc cctgatgcga tggactctgg agcagagctt gcttctccag      900
ggcccgtgga attggcttac gaacgagaaa aacatctacc agttcctgaa agaaggcgtc      960
gagaggtcca agaattacga ggtcatatat actatgggca tgagaggtct cggcgacacg     1020
gcatcgccca cgatcaacgc cacaactctg gcgcaaatcg tcgccgcgga ggagcagatc     1080
ctctccgagg ttttcaacac gacgaacatc tcttggatcc cgcagatgtg gtgtctctac     1140
aaggaagtgg gcggctacta cgaagacggg ctccgtgtgg cggacgatat cacgctcctc     1200
tgggcagacg acaactgggg gaacatcgag agactgccga ttggcaatga gactgcacgg     1260
tctggcgggtg ctggcgtcta ttatcacttt gactacgtcg gggacccccca agactacaag     1320
tggatcaaca ctatctcatt gcagaaaacc tgggagcaga tgcattctgc atacgagcgg     1380

```

ES 2 802 807 T3

caggcaagga atatttggat cgtgaatggt ggcgacttga aagcattgga aatcccgact 1440  
aaccacttct tcgacttggc ctacgatata cccagctggc ccgatccgga cagcaccagc 1500  
cgggtggctca agctctgggc cactcgtgaa ttcggggcag cgggggcccga caaagtccga 1560  
gacatcgtaa accgctacgg ccagtagcga gctagacgca aattcgagat gatcaccgcc 1620  
agcacgttca gcatcatcaa ctacaacgag gctaacaacg tggtagtca atgggagagt 1680  
ctggtcaacg atgccctgga tgtgtaccgc ggactcagcg aagctgcccg gccggcgttc 1740  
ttcgagctgg tattgcaacc ttgcatggct ggccacatca tcaactcagat ccatgttacg 1800  
gccgccaaga ataatctcta tgcgtcccag agacggacaa gcggaagc aatggcagac 1860  
aaggcgttga aactgttcaa agaggaccac gacttgacgg tttcctatca caagattctt 1920  
gatgaaaat ggaaccatat catggaccag acgcatctgg gttacctta ctggcagcag 1980  
ccgatgagaa acacactccc cctctccaa tacaactcagc tgtagagga cagtttggcg 2040  
ggtccaatgg gtgtttctgt cgagggcagc aatgcctctg tccctggcga cgatcaatac 2100  
cacgctttgt cgagtaccac cctgaccctg ccaccatag atccctacgg accgtcgagc 2160  
cgctggatcg agatctactc tcgctcaact gaaagcttcg agttccaaat ctccccaac 2220  
gtctcctggg tcaccgcgac accgtcctcc ggcaccgtcc atgttagtgg aaacggaacg 2280  
accaccgacg tccgctgca gctgtccgtc gactgggaca aggctcctat cggctcgagc 2340  
ctcgtctcga tcaacgtcag catccaacc ggccacttcc cgacggggga ctacgggaac 2400  
tcggccatgc cggttgtcca gtcctccgtg aacaagacgg tcgtgccagc gtccttccac 2460  
ggcttcgtcg agtcggacgg cgtcgtcagc atcgaggcgc cgcacgcaac gcgcaacacc 2520  
tcgacggcga acgtctccta cgccgtgatc ccgggtacg gacgcacact gtcgggcgtc 2580  
accttctgctc cggtgaacgc gccgtcacag aaagcgcctt cctccccgcy tctcgagtac 2640  
gacatgtacc tcttcagccc tgcggccaac ggcagcaccg tcaacataac gctgtacctg 2700  
ggaccgtcac tcaacgtcaa cctgaccga ccgctcgat acgctggtgc cttcgacgac 2760  
gacgacaatc cccaggtagt gcagtagtc ccgagcagcg tgctgggcac ctaccaccg 2820  
gactggccct cgccctgtgc caacaacgtg tggcgtcga cgacgtcgta tacaatcaag 2880  
ggctttaacg gaaacagcac tagtgcaaca ggaacgggat ccgcgcacac gctcaagctg 2940  
tgggctcttg agcctggcgt ggtgtttgag aagattgtca ttgatctggg cgggcagagg 3000  
gatagttatc tcggtccgcc ggagagtcgg attgtgtag 3039

<210> 76

5 <211> 6493

<212> ADN

10 <213> *Rasamsonia emersonii*

<220>

<221> fuente

15 <222> (1)..(6493)

ES 2 802 807 T3

<223> /organismo="Rasamsonia emersonii" /tipo\_molar="ADN no asignado"

<400> 76

ggagttgctt	gcccgggcag	gatatcggtt	ggcggcgtgg	ttggacttga	ttgcgaagca	60
gtagagacag	gtattctgct	gatagcatgg	atztatatta	tcaataagca	gaggcctaaa	120
gacattcaac	cagaagaaac	gtccttcatg	caacagcaag	ctagaacata	tacatatgag	180
aacaagcata	gtgccaatca	agcagaccta	tactgtactc	tattacaaca	aactactatc	240
acggtagtcg	tacactggtc	aatgaaataa	tgtgagtaaa	aatgatcatg	attctatgac	300
agaacagcta	gtacgcttga	tttattgggg	tataaatfff	actttattta	ggtgggtaag	360
agccagcaga	ctagatatat	agctcagata	taatataatt	aatagtcaca	gaaaaaata	420
aataaaaata	aaaatagcaa	gatccatgat	atggtataca	caaaaaata	ataatcataa	480
atcacacaat	tccatcctct	ccaaaaccac	ctagccagct	cctaccaaac	gatacatact	540
cagtccaagc	aaatcccccg	ttccgtcctc	ccgtccgtcc	gatcagtccc	gaataccgac	600
caaaaaaaaa	aagagacaaa	atccaaatca	cgggttcatt	cacatcccca	caggataccc	660
atggatcagt	cgtcctactt	actgtgttac	agattagaac	agaaaattag	gtttttacac	720
aactcagggg	ggttgcatcg	cattgcatcg	tgctgtggag	ttagttaact	tagttgtact	780
ccatccagtt	catacgcagt	acattattgg	gcatttgacc	ccatcagaca	agatatctaa	840
ggataagagt	agaaattagc	taataatagt	caaaagagaa	gaagagatac	caaggaact	900
agataactaac	aaccaatcag	gatatgccac	agtgtggaac	agaatggaag	cagacaggat	960
caacataact	ggaataaacc	ttttctttct	ttcttcgtac	agcatcttgg	caggaagtaa	1020
cttgatattg	ttaattaatg	tccatgtcca	tgtccatgtc	tttgttatgt	cttgtcttgt	1080
ctttcttgtg	tgctacaagt	acagtgtaac	agattcatal	ccgctgaaac	agacataaca	1140
ttcgacacat	ggaatacggg	gataaagaaa	taaattatac	tatatacggg	atgatatcaa	1200
taaatatccg	tgttgactct	cttattaaga	aagagtggcg	cttggcgcgt	ctacgtgcat	1260
ggactggtac	tacatatttg	attacttoga	tttttaataa	caacaaccta	gtagacgtat	1320
gtatgtcatg	tgaaactttc	gattgcgtgc	tttcttgtct	acttgtcgac	ttggtacaat	1380
cttgtcgaat	attaataata	ataatccatc	gcactgacat	cttggcaagt	actccgtaca	1440
tcaggttaca	tacatactga	ttctctaaag	ctagataacg	aataggattc	tcgcacagac	1500
agtatgtgtc	tcttgtctgt	cagatgataa	gcagatgaac	aaagaaagta	taactgctta	1560
ctacctacat	gccgacattt	agtcgattcc	tttcggagaa	tttattatgg	attattaata	1620
gcataccccg	ggattggcag	aaggggtaaa	aggtccgact	agacaaggat	atccatacag	1680
5 tacataccgt	tgatacagat	cgaatcacat	gcatactgct	gatggtgtga	tgaatccttg	1740

ES 2 802 807 T3

aattagacaa tcattccagac ctgtctggac agagatcctg gcaactgaaca atccactcat 1800  
tgctatctat cggfactctg tacctgtttc agctgaagct tgccaatcgc agactgccat 1860  
ctgcaactga tcagcgccag gatgcaggtc atgatacccc agcgttgttc ccgagggtgc 1920  
attgcttaa cgcgttaacc agtgtgctaa acgtgctaaa cgtgctaaat gctaaactgc 1980  
tgatgctatg cagctgcatc gccgaatctg gagaatgcag atcacctgcc gacggcgggc 2040  
tccgggcacg tgcacggggg accccgtagg acagaaacgt ccatcgagag tacggagtac 2100  
ggagtattac aagaccctgt ccatcagacc ctgtccatcg tcattgcaa gatctctcat 2160  
tgtttctgt ttcatgctcg gatcaccagt ggacagcaat gccccgtgaa cagcaagcgg 2220  
catgctggtc cgtgtcttgt ccgtgtgccg atgtagtatt gctaacgaga cccagaatgg 2280  
catcaatgac gttgctgatg acagaatgag ggggatcatc agtacgtctg ctatcaggat 2340  
gattatccta cggagtattt actcagctga agacaggaac aagatcgtct gatggatgag 2400  
gcccacggcc agccagcaca gactccgtac tcttcagtct tctggatttg accgttcgac 2460  
ggcgctccg acgtagcatc tcgctagcct gatccttggc tgcgctatc gtcggctcat 2520  
gcccctgttg atgacgggga agtggagcgg cgccgcgata aggttgcctt gctaatntag 2580  
cgcctgcacg ctccagccaa aaagaccaat attgaggtcg atcgtctccc ctggctccgt 2640  
gctgctggcc tgcgatcgcc ggcgcgatca taccctgcaa tcacgcccgc agcctatcac 2700  
agaccatgcg gtccttgac catctgggag ctcgagctct cctgactgcc gtcggggcgt 2760  
caatgcgtcc ggagcctccg acgagggcct ctgctcctcg tctgtcctac tggagcttgt 2820  
ccgtcagacg tcgcatcctg agccgtgtgc tgatatgcc atggctctga cgtgatcgac 2880  
tgcgagcggc cggcgaggct ataagaagcc gcaacttgcg gctcgaagta ccgtctccca 2940  
tccatcgatc agacagtcag cagtcctcac tcagtcatc ctcagtcgtc cttcaccacc 3000  
atgggtctgt ccaaagcctt cgtgtctgca ctctcgtgt gctccgccgt cggcgtggcc 3060  
gccccgaccg ggccagctcc caacgtgcag ttctccctga agcaggtcgc ggtgccccgg 3120  
accaagcctc gtgcgccccc agctgccgac tacgcgcgcg ctctggccaa gtatggcgt 3180  
ccaattccgt cgtctgtgcg gacggccgcg tccggcacgc agagcggctc tgcggccaac 3240  
acgcccgtcg ccggcgacag cttgtatctc acgcccgtta ccatcggcca gagcacgctg 3300  
aacctggact ttgacacggg ctctcgggat ctgtaagtgt cccaactctc gcaagaacaa 3360  
gaacggagca gctgactcgt ccagctgggt cttctccaac gagacgccct ccagcgagcg 3420  
cggcaaccac gccatctaca agcccagctc gacggccaag aagctgaacg gctacacctg 3480  
gagcatctcg tacggcgacg gcagctcggc cggcggcgac gtctaccagg acagcgtctc 3540  
ggtggggggc gtcaacgcct ccaaccaggc ggtcgaggcc gccaccaagg tcagctccga 3600

ES 2 802 807 T3

gttcacgcag gagccgggcg acggcttgct gggcctggcc ttcagcagca tcaacaccgt 3660  
 caagcccaag ccgcagacga ccttcttoga cacggccaag tcctcgctcg ccaagccgct 3720  
 gttcgcgctc accctcaagc acaacgagcc cggcagctac gactttggct acatcgacag 3780  
 ctccaagtac aagggcagca tccagtacac ccaggtcgac aactcgcagg gcttctggca 3840  
 gttcacggcc gacggctact cgattggcgg cagcagcggc agcggtggtt ccatttctgg 3900  
 cattgctggt aagaactccc cctacatcag agttatctag atgctgattt cgcagacacc 3960  
 ggcaccaccc tcctcctgct cgacgaccag atcgtcaacg agtactacca gcaggteccag 4020  
 ggcgcgcaga acgaccagaa cgccggcggc tacaccttcc cgtgcgacgc gcagctgccc 4080  
 gagctgagct tcaccatcgg ccagtacacc gccaccgtgc cggccgagta cctcaacttc 4140  
 cagcccgtgt cgcagggcag ccagacctgc ttccggcggc tgcagtccaa ccagggcatt 4200  
 ggcttctcca tcttcggcga cgtcttctcc aagagccagt acgtcgtctt tgactcggac 4260  
 ggtcctcagc tgggctttgc tgctcaggcg tagaccagtc gtccctccagc ccaggttggt 4320  
 tggtaggaga tgatttttcg atcgatcgat tatcatggtg attgatagga tatgtgcatg 4380  
 agcagttgcc tgtacataca tacataatga tttattgaat caattagtta tgatcaatct 4440  
 cgaatatatt ttcagtgaat tacgtacatg gtcatagcat aacgatatac tccgttttct 4500  
 tcaggtagct agtaaatata cacaaattca tcgttctccc ggtccgtcag gtccaggaag 4560  
 gctttgtctc cgatcgtccc gtcgggatca ctctcgtcgg tatcgtgata gcgctccctc 4620  
 atcagtaat caaccacctt ggccggcttt ccattccgca cgcgccgccg ctctgcatc 4680  
 cggttcagca cgaccaaat cgcccactgc gccagcacga cggccaccag cggcacgaag 4740  
 atggccagac aagcccgcac gcccgccga tacgccggcg cgtccttctt gctgaagagc 4800  
 agcgggccga cgatgttgcc cgccgagctg gccgcgttgt acaggctcat cagcggcgat 4860  
 ttcttcgctg tgccgcccggt gttgcccacg atccacgtca cgatcagtgg gttgccgccg 4920  
 aagagaaacg cgagcaggta gtagcctacc aggagggagg gttcagactga gttttgcttc 4980  
 gtgctgttat tactgcgtgg cacggcgtac agaattgcca ggcccgcgac taccggcagc 5040  
 atgaagccgg ccagcacgac gcccttcac cgcgcccgct gcgccagata gctccccgcc 5100  
 aggatgacca gcagctgcag cgcgccaaac ggcattgtga gcagactcgt cgtgtacgcg 5160  
 tcgtacccca ggccgttgag gatcagcggg ccgaacctgt tgctcacgct ggcgccgacg 5220  
 ttcagcagca tcgccatgcc gatccagagg taggttttgg gctcagagcg tcctcagacg 5280  
 acgtgccgga tcttgaactc cgggctccct gtgcccgtct ggttcgcgcg cagccgctcg 5340  
 atggcctgcg ctttttccgt ctccgtcagg aaccgcgctg aggggatgct gttgtctagt 5400  
 ttccagtaga tgaacggcac tgagatgatg gtcaggaggg cgacggtgag gaagatactg 5460  
 cgccagtcgt cagcagtcag tcagtcagcg ttccgggtag aaggggagta catctgccat 5520

ES 2 802 807 T3

ggcctcagaa caggcgactc gatatggccc aacctgtacg acagggccgc cgcgatgaca 5580  
gtcgccgcgc cgttggtact gtaccaggcc gcaatgcgca gcggctgctc ggcgcgccgg 5640  
taccactggc tggatgatgac gctgaacagc ggagacagc cggcctcga caggccgagg 5700  
aagaagcgcg cggccatcag ggagggcga ctgcgacagc cggccatggc ggctggggcg 5760  
acgccccagc ccagacacag cgggggcatc aggcggcgat gcggcacgcg cacgatcagc 5820  
cacgacgaga acggctgcca gacgagctgg gcgatgggcg cgatcgacc cagcagcgag 5880  
tactggttgc ccgtcaggtg cgtgtcggcc tgcaagccga aggtggcccc gtaccgagc 5940  
accgacttgt ccaggatctg caggaagtac acccacacga ggatggccag gatgacgcgg 6000  
tctgtcttgc gccggatgcg cttgctgtcg gcgtccgtga gtgggattct ctgctggccg 6060  
atcaggcgga gcgccgtgtc gccgtggacg gcgggttgc cttcttcatg ggtgacggtc 6120  
ggtttgatg ccattgtagc gattactaga tgtaataag ttgtaatggg agacaaacga 6180  
ccaagttctc tctcgacgtt ttataccggc ttatatgtct gttcagcagc attgcaagtc 6240  
aagtaatgac atcgaattc ctccggttcc ccgattgcg cggcgatcat cggctggcac 6300  
tagcagtata gctagctcag agtccgtatt actggattct attgcattgc gctgattgca 6360  
gacgttgact gacagcagga gctttgactc tattaccccc acgcttcggc aattccccgc 6420  
gtgctcgggc ctctatgcac cccacgtgg gggaacattc cagagtatgc aggcagtatg 6480  
atgcagcatg gat 6493

<210> 77

5 <211> 1194

<212> ADN

10 <213> *Rasamsonia emersonii*

<220>

<221> fuente

15 <222> (1)..(1194)

<223> /organismo="Rasamsonia emersonii" /tipo\_molar="ADN no asignado"

<400> 77

20 atgggtctgt ccaaagcctt cgtgtctgca ctctcgctgt gctccgccgt cgccgtggcc 60  
gccccgaccg ggccagctcc caacgtgcag ttctccctga agcaggctcg ggtgccccgg 120  
accaagcctc gtgcgcccc agctgcccac tacgcgcgcg ctctggcaa gtatggcgct 180  
ccaattccgt cgtctgtgcg gacggccgcg tccggcacgc agagcggctc tgcggccaac 240  
acgcccgtcg ccggcgacag cttgtatctc acgcccgtta ccatcgcca gagcacgctg 300  
aacctggact ttgacacggg ctctgcggat ctctgggtct tctccaacga gacgccctcc 360  
agcgagcgcg gcaaccacgc catctacaag cccagctcga cggccaagaa gctgaacggc 420  
tacacctgga gcatctcgta cggcgacggc agctcggccg gcggcgacgt ctaccaggac 480

ES 2 802 807 T3

agcgtctcgg tggggggcgt caacgcctcc aaccaggcgg tcgaggccgc caccaaggtc 540  
 agctccgagt tcacgcagga gccggggcag ggcttgctgg gcctggcctt cagcagcatc 600  
 aacaccgtca agcccaagcc gcagacgacc ttcttcgaca cggtaagtc ctcgctcggc 660  
 aagccgctgt tcgccgtcac cctcaagcac aacgagcccg gcagctacga ctttggtac 720  
 atcgacagct ccaagtataa gggcagcatc cagtacacc aggtcgacaa ctcgcagggc 780  
 ttctggcagt tcacggccga cggctactcg attggcggca gcagcggcag cggtggtctc 840  
 atttctggca ttgctgacac cggcaccacc ctctctctgc tcgacgacca gatcgtcaac 900  
 gagtactacc agcaggtcca gggcgcgcag aacgaccaga acgccggcgg ctacaccttc 960  
 ccgtgcgacg cgcagctgcc cgagctgagc ttcaccatcg gccagtacac cgccaccgtg 1020  
 ccggccgagt acctcaactt ccagcccgtg tcgcagggca gccagacctg cttcggcggg 1080  
 ctgcagtcca accagggcat tggcttctcc atcttcggcg acgtcttctt caagagccag 1140  
 tacgtcgtct ttgactcggg cggtcctcag ctgggctttg ctgctcaggc gtag 1194

<210> 78

5 <211> 397

<212> PRT

<213> *Rasamsonia emersonii*

10

<400> 78

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

ES 2 802 807 T3

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

Ile Ser Tyr Gly Asp Gly Ser Ser Ala Gly Gly Asp Val Tyr Gln Asp  
 145 150 155 160

Ser Val Ser Val Gly Gly Val Asn Ala Ser Asn Gln Ala Val Glu Ala  
 165 170 175

Ala Thr Lys Val Ser Ser Glu Phe Thr Gln Glu Pro Gly Asp Gly Leu  
 180 185 190

Leu Gly Leu Ala Phe Ser Ser Ile Asn Thr Val Lys Pro Lys Pro Gln  
 195 200 205

Thr Thr Phe Phe Asp Thr Val Lys Ser Ser Leu Ala Lys Pro Leu Phe  
 210 215 220

Ala Val Thr Leu Lys His Asn Glu Pro Gly Ser Tyr Asp Phe Gly Tyr  
 225 230 235 240

Ile Asp Ser Ser Lys Tyr Lys Gly Ser Ile Gln Tyr Thr Gln Val Asp  
 245 250 255

Asn Ser Gln Gly Phe Trp Gln Phe Thr Ala Asp Gly Tyr Ser Ile Gly  
 260 265 270

Gly Ser Ser Gly Ser Gly Gly Ser Ile Ser Gly Ile Ala Asp Thr Gly  
 275 280 285

Thr Thr Leu Leu Leu Leu Asp Asp Gln Ile Val Asn Glu Tyr Tyr Gln  
 290 295 300

Gln Val Gln Gly Ala Gln Asn Asp Gln Asn Ala Gly Gly Tyr Thr Phe  
 305 310 315 320

Pro Cys Asp Ala Gln Leu Pro Glu Leu Ser Phe Thr Ile Gly Gln Tyr  
 325 330 335

Thr Ala Thr Val Pro Ala Glu Tyr Leu Asn Phe Gln Pro Val Ser Gln  
 340 345 350

Gly Ser Gln Thr Cys Phe Gly Gly Leu Gln Ser Asn Gln Gly Ile Gly  
 355 360 365

Phe Ser Ile Phe Gly Asp Val Phe Leu Lys Ser Gln Tyr Val Val Phe  
 370 375 380

Asp Ser Asp Gly Pro Gln Leu Gly Phe Ala Ala Gln Ala  
 385 390 395

5 <210> 79

<211> 1066

ES 2 802 807 T3

<212> ADN

<213> Aspergillus nidulans

5 <220>

<221> fuente

<222> (1)..(1066)

10

<223> /organismo="Aspergillus nidulans" /tipo\_molar="ADN no asignado"

<400> 79

```

agacagctct gccggctctg aggtgcagtg gatgattatt aatccgggac cggccgcccc      60
tccgccccga agtggaaagg ctggtgtgcc cctcgttgac caagaatcta ttgcatcatc      120
ggagaatatg gagcttcatc gaatcaccgg cagtaagcga aggagaatgt gaagccaggg      180
gtgtatagcc gtcggcgaaa tagcatgcca ttaacctagg tacagaagtc caattgcttc      240
cgatctggta aaagattcac gagatagtac cttctccgaa gtaggtagag cgagtaccog      300
gcgcgtaagc tccctaattg gcccatccgg catctgtagg gcgtccaaat atcgtgcctc      360
tcctgctttg cccggtgtat gaaaccggaa aggccgctca ggagctggcc agcggcgcag      420
accggaaca caagctggca gtcgaccat ccggtgctct gactcgcacc tgctgaggtc      480
cctcagtccc tggtaggcag ctttgccccg tctgtccgcc cgggtgtgtcg gcggggttga      540
caagtcggtt gcgtcagtcc aacatttgtt gccatatttt cctgctctcc ccaccagctg      600
ctcttttctt ttctctttct tttcccatct tcagtatatt catcttccca tccaagaacc      660
tttatttccc ctaagtaagt actttgctac atccatactc catccttccc atcccttatt      720
cctttgaacc tttcagttcg agctttccca cttcatcgca gcttgactaa cagctacccc      780
gcttgagcag acatcaccat ggccaagttg accagtgccg ttccgggtgct caccgcgcgc      840
gacgtcgccg gagcggtcga gttctggacc gaccggctcg ggttctcccg ggacttcgtg      900
gaggacgact tcgccgggtg ggtccgggac gacgtgaccc tgttcatcag cgcggtccag      960
gaccaggtgg tgccggaaa caccctggcc tgggtgtggg tgcgcggcct ggacgagctg     1020
15 tacgccgagt ggtcggaggt cgtgtccacg aacttccggg acgcct                       1066

```

<210> 80

<211> 1062

20

<212> ADN

<213> Aspergillus nidulans

25 <220>

<221> fuente

<222> (1)..(1062)

30

<223> /organismo="Aspergillus nidulans" /tipo\_molar="ADN no asignado"

<400> 80

ES 2 802 807 T3

accagtgccg ttccggtgct caccgcgcgc gacgtcgccg gagcggtcga gttctggacc 60  
gaccggctcg ggttctcccc ggacttcctg gaggacgact tcgccggtgt ggtccgggac 120  
gacgtgaccc tgttcatcag cgcgggtccag gaccaggtgg tgccggacaa caccctggcc 180  
tgggtgtggg tgcgcggcct ggacgagctg tacgccgagt ggtcggaggt cgtgtccacg 240  
aacttccggg acgcctccgg gccggccatg accgagatcg gcgagcagcc gtgggggcgg 300  
gagttcgccc tgcgcgaccc ggccggcaac tgcgtgcact tcgtggccga ggagcaggac 360  
tgaccgacgc cgaccaacac cgcgggtccg acggcggccc acgggtccca ggagcttgag 420  
atccacttaa cgttactgaa atcatcaaac agcttgacga atctggatat aagatcgttg 480  
gtgtcgatgt cagctccgga gttgagacaa atggtgttca ggatctcgat aagatacgtt 540  
catttgtcca agcagcaaag agtgccttct agtgatttaa tagctccatg tcaacaagaa 600  
taaaacgctt ttcggggtt acctcttcca gatacagctc atctgcaatg cattaatgca 660  
ttgactgcaa cctagtaacg ccttcaggct ccggcgaaga gaagaatagc ttagcagagc 720  
tattttcatt ttcggggagac gagatcaagc agatcaacgg tcgtcaagag acctacgaga 780  
ctgaggaatc cgtcttggc tccacgcgac tataatattg tctctaattg tactttgaca 840  
tgctcctctt ctttactctg atagcttgac tatgaaaatt ccgtcaccag ccctgggttc 900  
gcaaagataa ttgcatgttt cttccttgaa ctctcaagcc tacaggacac acattcatcg 960  
taggtataaa cctcgaaatc attcctacta agatggtata caatagtaac catgcatggt 1020  
tgcctagtga atgctccgta acaccaata cgcgggccgg cc 1062

<210> 81

5 <211> 1501

<212> ADN

10 <213> Rasamsonia emersonii

<220>

<221> fuente

15 <222> (1)..(1501)

<223> /organismo="Rasamsonia emersonii" /tipo\_molar="ADN no asignado"

<400> 81

20 ccaattattc aaatcttcaa cctgggtcct ttgcgactag actcggcact ataccacact 60  
gagtcgaatc tccgctggac gatttttttt cttagaagaa aaaaagcttc agaggctaag 120  
gattaggctt ccgtacggac tccatgccct atcagacaga gactccgcaa ctcatctgat 180  
ctcttcgatg gagggaaaat ctgctgtttt tcgcaaattt ccaaccacc agaaacacc 240  
agaactgtca cgactcacac gtcctacggt ctttgtgtga gtataactga ttatattcac 300

ES 2 802 807 T3

agatgggtta ctcaatgcag gtacaaactt tgatcgcgct ctttttggga ccgtctatca 360  
 accggcttcg aaaaatggct cgactagcca atctgacagg aaattgcgat gttgcaaccg 420  
 tgtatacggg gtcctctgta caacctctgc caacctctgc cacctcggta catgtacgga 480  
 gtagctcccc gcagccgcga ttggatgcat taaagtgggt caaccgcagt ggcttgcaagt 540  
 ccgctgcacg agtccgatg caataattct tgacacacac gagtgcacat aataatagga 600  
 aagcagacaa actttgagct gaaggctgtc gagcttgcca aattgcagga tctggctagt 660  
 ttcgaagtgc acttcgcgcg cgcagcagta ttgcattatt gagtgtgacc tgctgctggg 720  
 gattagcgtc gcaccggccc aaagctagtc tcatccaagg ctgagcctga gcgctaatta 780  
 ccccgatca gccaaagcct aatggatcta atgaggtgcc tcctccagca ttcggcctgc 840  
 atggtgcggc gaccctctc tccacgtcca ataattgctg ttgcccctgt cgaaccctgc 900  
 caccgcatct ttgcggtttt actccgagat ctgaaaagcc tgctgtggat ggcagttcgc 960  
 aatatgcact ctcaatcagg tctgtagcat cttttaacta ttattctatt actaattgct 1020  
 tctggaaggc ttgtgggggtg tggtttgtca tcaagtggc tccttgagcg ccgcttgca 1080  
 atctccacgc gcggttgtag ggagtatatt catgcggatc cccggggcag agccgtagt 1140  
 catgtgacac taatgatca tccgctcaat tggatcctgg atttcgacct tggcttgaac 1200  
 atatccaatg atcttcagc gacgaaccga cccggtcatg ctttgttacc tacgtacgga 1260  
 gtagcggcct gggatgatgt tccggaaggt ctgctaaaag gagatcgagt atacccccg 1320  
 gggctccgtct gagacttata aagggctctc tgcaactctc cggccgactt ttttcttcat 1380  
 tcgacagcca tcaactggtt atctggtcga ttctgcagac ttgcccaagg agcaaagagc 1440  
 atcttcatac gcgcatcatc catctccagc tttctctctc caaacataca ccgtcaaaat 1500  
 g 1501

<210> 82

5 <211> 301

<212> ADN

10 <213> Aspergillus nidulans

<220>

<221> fuente

15 <222> (1)..(301)

<223> /organismo="Aspergillus nidulans" /tipo\_molar="ADN no asignado"

<400> 82

ctaataagtg tcagatagca atttgcacaa gaaatcaata ccagcaactg taaataagcg 60  
 ctgaagtgac catgccatgc tacgaaagag cagaaaaaaa cctgccgtag aaccgaagag 120  
 atatgacacg cttccatctc tcaaaggaag aatcccctca gggttgcgct tccagtctag 180  
 acacgtataa cggcacaagt gtctctcacc aatggggtta tatctcaaat gtgatctaag 240

ES 2 802 807 T3

gatgaaaagc ccagaatatt ggctgggttg atggctgctt cgagtgcagt ctcatgctgc 300

c 301

## REIVINDICACIONES

1. Un polipéptido que tiene actividad como hemicelulasa que comprende la secuencia de aminoácidos indicada en SEQ ID NO: 32 o una secuencia de aminoácidos codificada por la secuencia nucleotídica de SEQ ID NO: 31, SEQ ID NO: 34 o uno de sus polipéptidos variantes o polinucleótidos variantes, en donde el polipéptido variante tiene al menos 75% de identidad de secuencia con la secuencia indicada en SEQ ID NO: 32, o el polinucleótido variante codifica un polipéptido que tiene al menos 75% de identidad de secuencia con la secuencia indicada en SEQ ID NO: 32.
2. Un polipéptido según la reivindicación 1, que tiene actividad como alfa-arabinofuranosidasa.
3. Una secuencia de ácido nucleico aislada que codifica una hemicelulasa, en donde la secuencia de ácido nucleico se selecciona del grupo que consiste en:
- (a) una secuencia de ácido nucleico que tiene al menos 70% de identidad con la secuencia de ácido nucleico de SEQ ID NO: 31, SEQ ID NO: 34 o SEQ ID NO: 35;
- (b) una secuencia de ácido nucleico que codifica (i) la secuencia de aminoácidos de SEQ ID NO: 32, (ii) una secuencia de aminoácidos que tiene al menos 75% de identidad con la secuencia de aminoácidos de SEQ ID NO: 32, o (iii) una secuencia de aminoácidos que difiere en 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 o 12 aminoácidos con respecto a la secuencia de aminoácidos de SEQ ID NO: 32; o
- (c) una secuencia de ácido nucleico que es el complemento inverso de una secuencia de ácido nucleico como la definida en (a) o (b).
4. Un constructo o vector de ácido nucleico que comprende la secuencia de ácido nucleico según la reivindicación 3.
5. Una célula recombinante que comprende un polipéptido según la reivindicación 1 o 2, una secuencia de ácido nucleico según la reivindicación 3 o una construcción o vector de ácido nucleico según la reivindicación 4, preferiblemente la célula es una célula fúngica, preferiblemente una célula fúngica seleccionada del grupo que consiste en los géneros *Acremonium*, *Agaricus*, *Aspergillus*, *Aureobasidium*, *Chrysosporium*, *Coprinus*, *Cryptococcus*, *Filibasidium*, *Fusarium*, *Humicola*, *Magnaporthe*, *Mucor*, *Myceliophthora*, *Neocallimastix*, *Neurospora*, *Paecilomyces*, *Penicillium*, *Piromyces*, *Panerochaete*, *Pleurotus*, *Schizophyllum*, *Talaromyces*, *Rasamsonia*, *Thermoascus*, *Thielavia*, *Tolypocladium* y *Trichoderma*.
6. Una célula según la reivindicación 5, en la que uno o más genes se eliminan, inactivan o alteran totalmente o en parte, en donde opcionalmente el gen codifica una proteasa.
7. Un método para la preparación de un polipéptido según la reivindicación 1 o 2, que tiene actividad como hemicelulasa, método que comprende cultivar una célula según la reivindicación 5 o 6 bajo condiciones que permitan la expresión de dicho polipéptido y, opcionalmente, recuperar el polipéptido expresado.
8. Una composición que comprende: (i) un polipéptido según la reivindicación 1 o 2 y; (ii) una celulasa y/o una hemicelulasa adicional y/o una pectinasa.
9. Una composición según la reivindicación 8, en donde la celulasa es una GH61, celobiohidrolasa I, celobiohidrolasa II, endo- $\beta$ -1,4-glucanasa,  $\beta$ -glucosidasa o  $\beta$ -(1,3)(1,4)-glucanasa y/o en donde la hemicelulasa adicional es una endoxilanas,  $\beta$ -xilosidasa,  $\alpha$ -L-arabinofuranosidasa,  $\alpha$ -D-glucuronidasa, feruloilesterasa, cumarolesterasa,  $\alpha$ -galactosidasa,  $\beta$ -galactosidasa,  $\beta$ -mananasa o  $\beta$ -manosidasa.
10. Una composición según la reivindicación 8 o 9, en donde la composición es un caldo de fermentación de masa celular cruda.
11. Un método para el tratamiento de un sustrato que comprende hemicelulosa, opcionalmente un material vegetal, método que comprende poner en contacto el sustrato con un polipéptido según la reivindicación 1 o 2 y/o una composición según una cualquiera de las reivindicaciones 8 a 10.
12. Un método según la reivindicación 11, en donde antes del tratamiento enzimático el sustrato se pretrata con calor, modificación mecánica y/o química o cualquier combinación de estos métodos.
13. Un método para producir un producto de fermentación, método que comprende producir un azúcar fermentable al poner en contacto material lignocelulósico con un polipéptido según la reivindicación 1 o 2 y/o una composición según una cualquiera de las reivindicaciones 8 a 10, y fermentar el azúcar fermentable resultante para producir de ese modo el producto de fermentación.
14. Uso de un polipéptido según la reivindicación 1 o 2 y/o una composición según cualquiera de las reivindicaciones 8 a 10, para producir azúcar a partir de un material lignocelulósico.

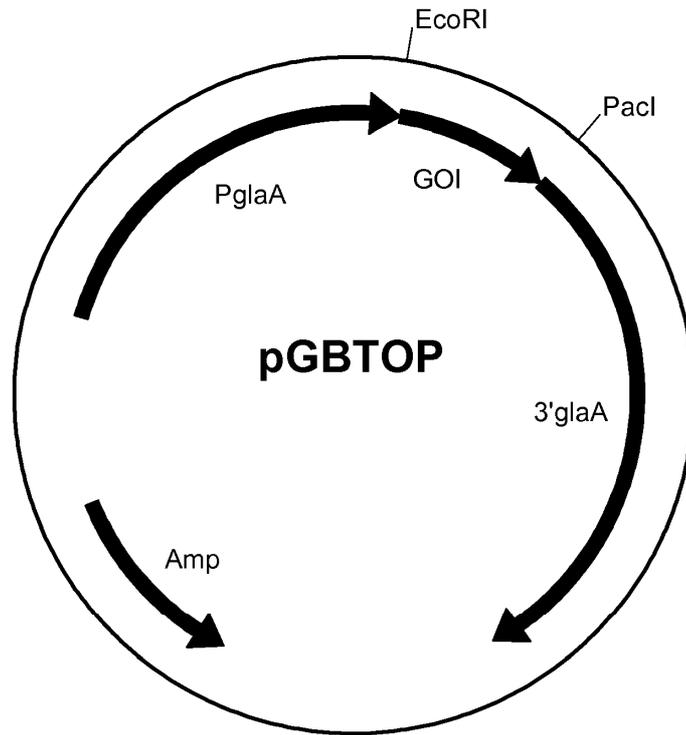


Fig. 1

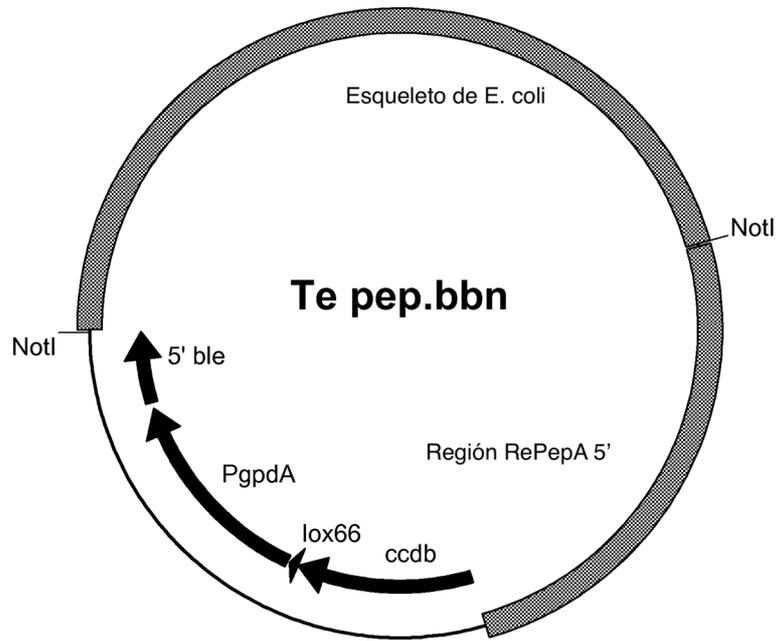


Fig. 2

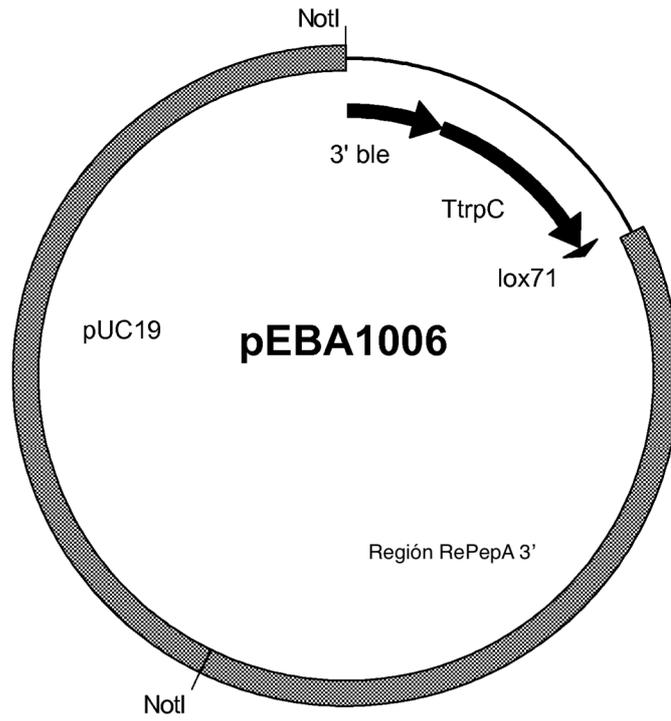


Fig. 3

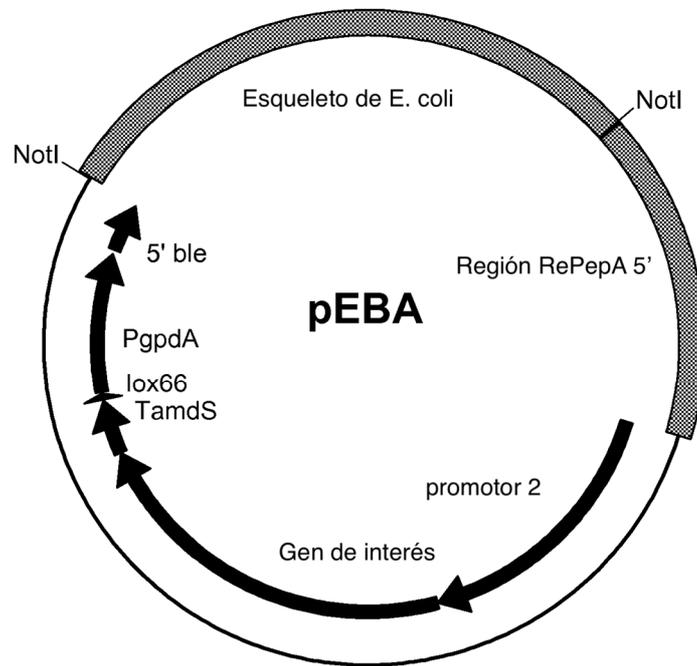


Fig. 4

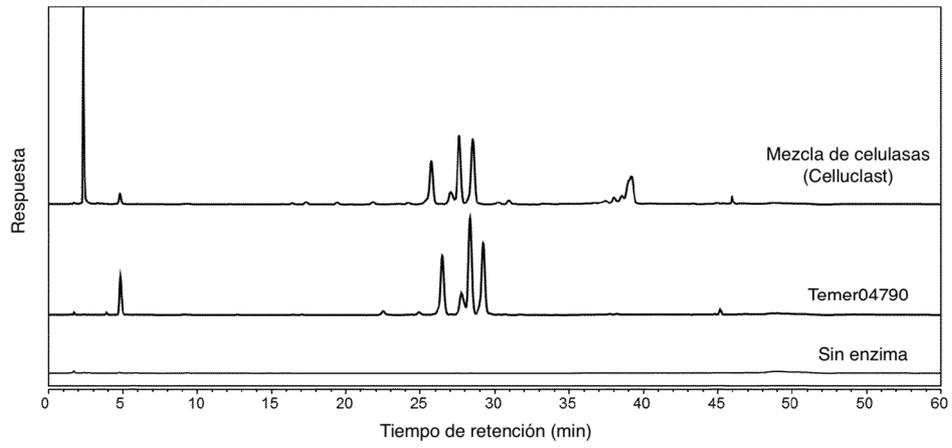


Fig. 5