



19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 356 614**

51 Int. Cl.:

<b>C12N 9/24</b> (2006.01)	<b>C12N 15/56</b> (2006.01)
<b>C12P 19/02</b> (2006.01)	<b>C12P 19/04</b> (2006.01)
<b>C12S 3/10</b> (2006.01)	<b>A21D 8/04</b> (2006.01)
<b>A23L 1/03</b> (2006.01)	<b>A23L 1/29</b> (2006.01)
<b>C12N 1/15</b> (2006.01)	<b>C12N 1/21</b> (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **06722905 .4**

96 Fecha de presentación : **25.04.2006**

97 Número de publicación de la solicitud: **1874926**

97 Fecha de publicación de la solicitud: **09.01.2008**

54

Título: **Arabinofuranosidasas.**

30

Prioridad: **26.04.2005 DK 2005 00609**  
**10.11.2005 DK 2005 01562**

45

Fecha de publicación de la mención BOPI:  
**11.04.2011**

45

Fecha de la publicación del folleto de la patente:  
**11.04.2011**

73

Titular/es: **NOVOZYMES A/S**  
**Krogshøjvej 36**  
**2880 Bagsværd, DK**

72

Inventor/es: **Soerensen, Hanne, Risbjerg;**  
**Joergensen, Christel, Thea;**  
**Christensen, Lars, Hylling;**  
**Joergensen, Christian, Isak;**  
**Hansen, Carsten, Hoerslev y**  
**Kofod, Lene, Venke**

74

Agente: **Tomás Gil, Tesifonte Enrique**

ES 2 356 614 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

Arabinofuranosidasas.

5 **Campo de la invención**

La presente invención se refiere a polipéptidos aislados teniendo actividad de alfa-L-arabinofuranosidasa y secuencias de ácidos nucleicos aisladas codificando los polipéptidos. La invención también se refiere a constructos de ácidos nucleicos, vectores, y células huéspedes comprendiendo las secuencias de ácidos nucleicos al igual que métodos para producir y usar los polipéptidos.

10 **Antecedentes de la invención**

Las arabinofuranosidasas son capaces de hidrolizar residuos de alfa-L-arabinofuranosida terminales no reductores en alfa-L-arabinosidas y se clasifican como EC 3.2.1.55. Las arabinofuranosidasas se han aislado a partir de diferentes organismos incluyendo hongos filamentosos. No obstante, se conocen muy pocas alfa-arabinofuranosidasas capaces de liberar arabinosa a partir de xilosas di-sustituidas. Se ha descrito una enzima intracelular de *Bifidobacterium adolescentis* capaz de liberar arabinosa a partir de C3 de xilosa di-sustituida (también internamente) (Van Laere, 1997, Appl. Microbiol. Biotechnol, 47, 231-235 y Van den Broek, 2005, Applied Microbiology and Biotechnology). A partir de cebada se ha aislado una enzima activa en xilosas mono- y terminalmente di-sustituidas, pero ésta tiene poca actividad en xilosas internamente di-sustituido (Ferre, 2000, Eur. J. Biochem., 267, 6633-6641). Una enzima de *Trichoderma reesei* es posiblemente activa en residuos terminalmente di-sustituidos (actividad vista en 3,5-di-O-alfa-L-arabinofuranosil-alfa-L-arabinofuranosida), pero no tiene actividad hacia un oligo-sustrato con arabinosa C3 internamente sustituida (Nogawa, 1999, Appl. Circundar Environ. Microbiol., 65, 3964-3968).

Una comparación de secuencias del estado anterior de la técnica en toda su longitud muestra que la secuencia de aminoácidos maduros mostrada en la SEC ID NO:2 tiene un 72% de identidad con una secuencia de aminoácidos de *Chaetomium globosum*, y la secuencia de ADN correspondiente en la SEC ID NO:1 muestra un 73% de identidad con los de la secuencia de ADN de *Chaetomium globosum* correspondiente. La identidad entre la secuencia mostrada en la SEC ID NO:2 y la alfa-L-arabinofuranosidasa GH43 bacteriana de *Bifidobacterium sp.* es del 25%.

30 **Resumen de la invención**

Los inventores han aislado una alfa-L-arabinofuranosidasa de una cepa de la *Humicola insolens* de hongos filamentosos (SEC ID NO:2). Los inventores también aislaron el gen que codifica la alfa-L-arabinofuranosidasa nueva. La enzima es extracelular y pertenece a GH43.

La alfa-L-arabinofuranosidasa de *Humicola insolens* es capaz de liberar arabinosa a partir de xilosas di-sustituidas, es decir la alfa-L-arabinofuranosidasa es activa en unidades de xilosa de arabinoxilano de trigo con arabinosa fijada a C2 y C3. La actividad hacia xilosas di-sustituidas es esencial para la hidrólisis total de arabinoxilano a monosacáridos por ejemplo en la producción de etanol a partir de biomasa.

Por consiguiente, en un primer aspecto la invención proporciona una arabinofuranosidasa la cual es: (a) un polipéptido teniendo una secuencia de aminoácidos como el péptido maduro mostrado en la SEC ID n°: 2 o (b) un análogo del polipéptido definido en (a) el cual: (i) tiene al menos un 80% de identidad con dicho polipéptido, (ii) o es una variante alélica de dicho polipéptido. Por consiguiente, en un segundo aspecto la invención proporciona una secuencia de ácidos nucleicos teniendo al menos un 80% de identidad con las secuencias de ADN mostradas en la SEC ID NO:1, o (a) es una variante alélica de la SEC ID NO:1, o (b) una cadena complementaria a a).

Por consiguiente, en un tercer aspecto la invención proporciona un constructo de ácidos nucleicos comprendiendo la secuencia de ácidos nucleicos del segundo aspecto operativamente vinculada a una o más secuencias de control capaces de dirigir la expresión de la arabinofuranosidasa en un huésped de expresión adecuado.

En un cuarto aspecto la invención proporciona un vector de expresión recombinante comprendiendo el constructo de ácidos nucleicos del tercer aspecto.

En un quinto aspecto la invención proporciona una célula huésped recombinante comprendiendo el vector de expresión del cuarto aspecto.

En un sexto aspecto la invención proporciona un método para la producción de una arabinofuranosidasa comprendiendo cultivar la célula huésped del quinto aspecto bajo condiciones propicias para la producción de la arabinofuranosidasa, y recuperar la arabinofuranosidasa.

En un séptimo aspecto la invención proporciona una composición comprendiendo la arabinofuranosidasa según el primer aspecto.

En otros aspectos la invención proporciona usos de las arabinofuranosidasas del primer aspecto ss en una masa, en la producción de etanol combustible a partir de celulosa conteniendo biomasa o en la producción de etanol combustible y/o potable a partir de almidón, en un proceso de maceración, o en un proceso para producir un producto alimenticio.

**Breve descripción de los dibujos**

Fig. 1A-C muestra polímeros de arabinosilano.

5 Fig. 1A muestra arabinosilano intacto.

Fig. 1B muestra arabinosilano di-sustituido.

10 Fig. 1C muestra arabinosilano sustituido individualmente.

Fig. 2A-C muestra arabinosilo-oligosacáridos.

Fig. 2A muestra grupos de arabinosilo vinculados a C-3 interno.

15 Fig. 2B muestra grupos de arabinosilo vinculados a C-3 terminal.

Fig. 2C muestra grupos de arabinosilo vinculados a C-2 interno.

**Descripción detallada de la invención**

20

En una forma de realización del segundo aspecto de la presente invención, el polipéptido aislado tiene una secuencia de aminoácidos la cual tiene al menos un 80% de identidad con la secuencia de aminoácidos mostrada como los aminoácidos 1 a 558 de la SEC ID NO:2. En una forma de realización interesante de la invención el polipéptido tiene al menos un 75%, al menos un 85%, al menos un 90%, al menos un 95%, al menos un 96%, al menos un 97%, al menos un 98%, o al menos un 99% de identidad con la secuencia de aminoácidos mostrada como los aminoácidos 1 a 558 de la SEC ID NO:2.

25

En una forma de realización del segundo aspecto de la invención, el polipéptido aislado es una alfa-L-arabinofuranosidasa capaz de liberar arabinosa a partir de xilosas di-sustituidas, por ejemplo la alfa-L-arabinofuranosidasa es activa en unidades de xilosa de arabinosilano de trigo con arabinosa fijada a C2 y C3.

30

Los alineamientos de secuencias y el cálculo de identidad se pueden determinar adecuadamente mediante programas informáticos conocidos en la técnica tales como GAP proporcionado en el paquete de programas GCG (Program Manual for the Wisconsin Package, Version 8, Agosto 1994, Genetics Computer Group, 575 Science Drive, Madison, Wisconsin, USA 53711) (Needleman, S.B. y Wunsch, C.D., (1970), Journal of Molecular Biology, 48, 443-453. Se usan los siguientes ajustes para la comparación de secuencias de aminoácidos: penalización de creación de GAP de 3.0 y penalización de extensión de GAP de 0.1. La parte relevante de la secuencia de aminoácidos para la determinación de identidad es el polipéptido maduro, es decir sin el péptido señal.

35

Preferiblemente, los polipéptidos de la presente invención comprenden las secuencias de aminoácidos mostradas como los aminoácidos 1 a 558 de la SEC ID NO:2, una variante alélica de lo mismo, o un fragmento de lo mismo que tenga actividad de arabinofuranosidasa. Obviamente, el polipéptido de la invención también puede consistir en las secuencias de aminoácidos mostradas como los aminoácidos 1 a 558 de la SEC ID NO:2.

40

Una variante alélica denota cualquiera de dos o más formas alternativas de un gen ocupando la misma localización cromosómica. Una variación alélica surge naturalmente a través de una mutación, y pueden suponer polimorfismo dentro de poblaciones. Las mutaciones genéticas pueden ser silenciosas (ningún cambio en el polipéptido codificado) o pueden codificar polipéptidos teniendo secuencias de aminoácidos alteradas. Una variante alélica de un polipéptido es un polipéptido codificado por una variante alélica de un gen.

45

Los polipéptidos también pueden ser variantes alélicas o fragmentos de los polipéptidos que tengan actividad de arabinofuranosidasa.

50

Las variantes de los polipéptidos especificados comprendiendo una sustitución, delección, y/o inserción de uno o más aminoácidos. En una forma de realización particular, los polipéptidos son variantes termoestables de los polipéptidos especificados.

55

Las secuencias de aminoácidos de los polipéptidos variantes pueden diferir de la secuencia de aminoácidos especificada en los aminoácidos 1 a 558 de la SEC ID NO:2 por una inserción o delección de uno o más residuos de aminoácidos y/o la sustitución de uno o más residuos de aminoácidos por diferentes residuos de aminoácidos. Preferiblemente, los cambios de aminoácidos son de una naturaleza menor, de la que lo son las sustituciones de aminoácidos conservadoras las cuales no afectan significativamente al plegado y/o la actividad de la proteína; delecciones pequeñas, típicamente de uno a aproximadamente 30 aminoácidos; extensiones amino- o carboxilo-terminales pequeñas, tales como un residuo de metionina aminoterminal; un pequeño péptido de enlace de hasta aproximadamente 20-25 residuos; o una pequeña extensión que facilita la purificación cambiando la carga neta u otra función, tal como un tracto de polihistidina, un epitopo antigénico o un dominio de unión.

65

## ES 2 356 614 T3

Ejemplos de sustituciones conservadoras están en el grupo de aminoácidos básicos (arginina, lisina e histidina), aminoácidos ácidos (ácido glutámico y ácido aspártico), aminoácidos polares (glutamina y asparagina), aminoácidos hidrofóbicos (leucina, isoleucina y valina), aminoácidos aromáticos (fenilalanina, triptófano y tirosina), y aminoácidos pequeños (glicina, alanina, serina, treonina y metionina). Las sustituciones de aminoácidos que generalmente no alteran la actividad específica se conocen en la técnica y están descritas, por ejemplo, por H. Neurath y R.L. Hill, 1979, In, *The Proteins*, Academic Press, Nueva York. Los cambios que ocurren más frecuentemente son Ala/Ser, Val/Ile, Asp/Glu, Thr/Ser, Ala/Gly, Ala/Thr, Ser/Asn, Ala/Val, Ser/Gly, Tyr/Phe, Ala/Pro, Lys/Arg, Asp/Asn, Leu/Ile, Leu/Val, Ala/Glu, y Asp/Gly al igual que a la inversa.

Los polipéptidos denominados en este caso pueden comprender la secuencia de aminoácidos especificada, o pueden ser una variante alélica de la misma; o un fragmento de la misma que tenga la actividad enzimática pertinente. En una forma de realización, los polipéptidos comprenden la secuencia de aminoácidos especificada o una variante alélica de la misma; o un fragmento de la misma que tenga la actividad enzimática pertinente. En otra forma de realización, los polipéptidos consisten en la secuencia de aminoácidos especificados, o una variante alélica de la misma; o un fragmento la misma que tenga la actividad enzimática pertinente.

Un fragmento de una secuencia de aminoácidos específica es un polipéptido teniendo uno o más aminoácidos delecionados del amino y/o carboxilo terminus de esta secuencia de aminoácidos. En una forma de realización, un fragmento contiene al menos 60 residuos de aminoácidos, o al menos 68, o al menos 70, o al menos 75, o al menos 100, o al menos 150, o al menos 160, o al menos 170, o al menos 180, o al menos 190, o al menos 200, o al menos 210, o al menos 220, o al menos 240, o al menos 260, o al menos 280, o al menos 300, o al menos 310, o al menos 320, o al menos 330, o al menos 334, o al menos 350, o al menos 375, o al menos 400, o al menos 425, o al menos 430 residuos de aminoácidos.

Una variante alélica denota cualquiera de dos o más formas alternativas de un gen ocupando la misma localización cromosómica. Una variación alélica surge naturalmente a través de una mutación, y puede suponer un polimorfismo dentro de poblaciones. Las mutaciones genéticas pueden ser silenciosas (ningún cambio en el polipéptido codificado) o pueden codificar polipéptidos teniendo secuencias de aminoácidos alteradas. Una variante alélica de un polipéptido es un polipéptido codificado por una variante alélica de un gen.

Un polipéptido maduro o una secuencia de aminoácidos madura se refiere a la parte de una secuencia de aminoácidos que permanece después de haberse dividido una parte de péptido señal potencial. Y análogamente, un polipéptido maduro codificando una parte de un gen se refiere a la parte de un gen que corresponde a un polipéptido maduro.

La secuencia de ácidos nucleicos de cualquiera de la SEC ID NO:1 o una subsecuencia de la misma, al igual que la secuencia de aminoácidos de cualquiera de la SEC ID NO:2 o un fragmento de la misma, se puede utilizar para diseñar una sonda de ácido nucleico para identificar y clonar ADN codificando polipéptidos teniendo actividad de arabinofuranosidasa a partir de cepas de diferentes géneros o especies según métodos bien conocidos en la técnica. En particular, tales sondas se pueden usar para la hibridación con el genoma o ADNc del género o especie de interés, siguiendo procedimientos de transferencia de Southern estándares, para identificar y aislar el gen correspondiente. Tales sondas pueden ser considerablemente más cortas que la secuencia entera, pero deberían ser de al menos 15, preferiblemente de al menos 25, y más preferiblemente de al menos 35 nucleótidos de longitud. También se pueden usar sondas más largas. Se pueden usar ambas sondas de ADN y ARN. Las sondas están típicamente marcadas para detectar el gen correspondiente (por ejemplo, con  $^{32}\text{P}$ ,  $^3\text{H}$ ,  $^{35}\text{S}$ , biotina, o avidina). Tales sondas están comprendidas por la presente invención.

Así, un ADN genómico o genoteca de ADNc obtenida a partir de tales otros organismos se puede seleccionar para ADN que hibridice con las sondas anteriormente descritas y que codifique un polipéptido teniendo actividad de arabinofuranosidasa. El genoma u otro ADN de tales otros organismos se puede separar por agarosa o electroforesis en gel de poliacrilamida, u otras técnicas de separación conocidas por el experto en la materia. El ADN las genotecas o el ADN separado se puede transferir e inmovilizar en nitrocelulosa u otros materiales portadores adecuados. Para identificar un clon o ADN que sea homólogo a la SEC ID NO:1, o una subsecuencia de la misma, el material portador se usa en una transferencia de Southern. Para objetivos de la presente invención, la hibridación indica que la secuencia de ácidos nucleicos hibridiza a un sonda de ácido nucleico marcada correspondiente a la secuencia de ácidos nucleicos mostrada en la SEC ID NO:1, su cadena complementaria, o una subsecuencia de la misma, bajo de bajas a altas condiciones de astringencia. Moléculas a la cuales hibridiza la sonda de ácidos nucleicos bajo estas condiciones se detectan usando película radiográfica.

En otra forma de realización interesante, la sonda de ácido nucleico es una secuencia de ácidos nucleicos que codifica el polipéptido (maduro) de la SEC ID NO:2, o una subsecuencia de lo mismo. En una tercera forma de realización interesante, la sonda de ácido nucleico es la SEC ID NO:1. En una cuarta forma de realización interesante, la sonda de ácido nucleico es el polipéptido maduro codificando la región de la SEC ID NO:1.

Para sondas largas de al menos 100 nucleótidos de longitud, condiciones de astringencia de bajas a altas se definen como prehibridación e hibridación a 42°C en 5X de SSPE, 0,3% de SDS, 200 µg/ml de ADN de esperma de salmón desnaturalizado y cortado, y o bien un 25% de formamida para astringencia baja, un 35% de formamida para astringencia media, o un 50% de formamida para astringencia alta, siguiendo procedimientos de transferencia de Southern estándares.

## ES 2 356 614 T3

Para sondas largas de al menos 100 nucleótidos de longitud, el material portador se lava finalmente tres veces durante 15 minutos cada vez usando 2 x de SSC, 0,2% de SDS preferiblemente al menos a 50°C (astringencia baja), más preferiblemente al menos a 55°C (astringencia media), incluso más preferiblemente al menos a 65°C (astringencia alta).

5 Para sondas cortas las cuales son de aproximadamente 15 nucleótidos a aproximadamente 70 nucleótidos de longitud, las condiciones de astringencia se definen como prehibridación, hibridación, y post-hibridación de lavado a entre 5°C y 10°C por debajo de la T<sub>m</sub> calculada usando el cálculo según Bolton y McCarthy (1962, Proceedings of the National Academy of Sciences EEUU 48:1390) en 0,9 M de NaCl, 0,09 M de Tris-HCl a pH 7.6, 6 mM de EDTA,  
10 0,5% de NP-40, 1X de solución Denhardt, 1 mM de pirofosfato de sodio, 1 mM de fosfato monobásico de sodio, 0,1 mM de ATP, y 0,2 mg de levadura ARN por ml, siguiendo los procedimientos de transferencia de Southern estándares.

Para sondas cortas las cuales tienen de aproximadamente 15 nucleótidos a aproximadamente 70 nucleótidos en longitud, el material portador se lava una vez en 6X de SSC más un 0,1% de SDS durante 15 minutos y dos veces durante 15 minutos cada una usando 6X de SSC a entre 5°C y 10°C por debajo de la T<sub>m</sub> calculada.

Como se ha indicado anteriormente, el polipéptido de la invención puede ser un polipéptido teniendo la secuencia de aminoácidos de la SEC ID NO:2, o el polipéptido maduro de la misma, donde uno o más aminoácido(s) ha(n) sido sustituido(s) por otro(s) aminoácido(s), donde uno o más aminoácido(s) se ha(n) delecionado, y/o donde se ha insertado un(os) aminoácido(s) más.

Preferiblemente, los cambios de aminoácidos son de una naturaleza menor, de lo que lo son las sustituciones de aminoácidos conservadoras que no afectan significativamente el plegado y/o la actividad de la proteína; deleciones pequeñas, típicamente de uno a aproximadamente 30 aminoácidos; extensiones amino- o carboxilo-terminales pequeñas, tales como un residuo de metionina aminoterminal; un pequeño péptido de enlace de hasta aproximadamente 20-25 residuos; o una pequeña extensión que facilita la purificación cambiando la carga neta u otra función, tal como un tracto de polihistidina, un epítopo antigénico o un dominio de unión.

Ejemplos de sustituciones conservadoras están en el grupo de aminoácidos básicos (arginina, lisina y histidina), aminoácidos ácidos (ácido glutámico y ácido aspártico), aminoácidos polares (glutamina y asparagina), aminoácidos hidrofóbicos (leucina, isoleucina y valina), aminoácidos aromáticos (fenilalanina, triptófano y tirosina), y aminoácidos pequeños (glicina, alanina, serina, treonina y metionina). Sustituciones de aminoácidos que generalmente no alteran la actividad específica se conocen en la técnica y están descritas, por ejemplo, por H. Neurath y R.L. Hill, 1979, en, The Proteins, Academic Press, Nueva York. Los cambios que ocurren más frecuentemente son Ala/Ser, Val/Ile, Asp/Glu, Thr/Ser, Ala/Gly, Ala/Thr, Ser/Asn, Ala/Val, Ser/Gly, Tyr/Phe, Ala/Pro, Lys/Arg, Asp/Asn, Leu/Ile, Leu/Val, Ala/Glu, y Asp/Gly al igual que estos a la inversa.

En general, se prefiere que los polipéptidos de la invención tengan como mínimo un 20% de la actividad de arabinofuranosidasa del polipéptido teniendo la secuencia de aminoácidos mostrada como los aminoácidos del 1 al 558 de la SEC ID NO:2. Particularmente se prefieren polipéptidos que tengan al menos un 30%, tal como al menos un 40%, por ejemplo al menos un 50%, preferiblemente al menos un 60%, tal como al menos un 70%, por ejemplo al menos un 80%, más preferiblemente al menos un 90%, o al menos un 95% de la actividad de arabinofuranosidasa del polipéptido mostrada como los aminoácidos 1 a 558 de la SEC ID NO:2.

Un polipéptido de la presente invención se puede obtener a partir de microorganismos de cualquier género. Para objetivos de la presente invención, el término "obtener a partir de" como se utiliza en este caso en relación a una fuente dada significará que el polipéptido codificado por la secuencia de ácidos nucleicos se produce por la fuente o por una célula en cual la secuencia de ácidos nucleicos de la fuente se ha insertado. En una forma de realización preferida, el polipéptido se segrega extracelularmente.

Un polipéptido de la presente invención puede ser un polipéptido fúngico, y más preferiblemente un polipéptido filamentosamente fúngico tal como un polipéptido de Acremonium, Aspergillus, Aureobasidium, Cryptococcus, Filibasidium, Fusarium, Humicola, Magnaporthe, Meripilus, Mucor, Myceliophthora, Neocallimastix, Neurospora, Paecilomyces, Penicillium, Piromyces, Schizophyllum, Talaromyces, Thermoascus, Thielavia, Tolypocladium, o Trichoderma.

En otra forma de realización preferida, el polipéptido es un polipéptido de *Aspergillus aculeatus*, *Aspergillus awamori*, *Aspergillus foetidus*, *Aspergillus japonicus*, *Aspergillus nidulans*, *Aspergillus niger*, *Aspergillus oryzae*, *Fusarium bac-tridioides*, *Fusarium cerealis*, *Fusarium crookwellense*, *Fusarium culmorum*, *Fusarium graminearum*, *Fusarium graminum*, *Fusarium heterosporum*, *Fusarium negundi*, *Fusarium oxysporum*, *Fusarium reticulatum*, *Fusarium roseum*, *Fusarium sambucinum*, *Fusarium sarcochroum*, *Fusarium sporotrichioides*, *Fusarium sulphureum*, *Fusarium torulosum*, *Fusarium trichothecioides*, *Fusarium venenatum*, *Humicola insolens*, *Humicola lanuginosa*, *Meripilus giganteus*, *Mucor miehei*, *Myceliophthora thermophila*, *Neurospora crassa*, *Penicillium purpurogenum*, *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei*, o *Trichoderma viride*.

En una forma de realización preferida de la invención el polipéptido de la invención se deriva de una cepa de Ascomycota, por ejemplo, del género humicola, tal como de *H. lanuginosa* H. fuscoatra, *H. grisea*, *H. lutea*, *H. nigrescens* y en particular de *H. insolens*, o de una cepa de Basidiomycota, tal como del género Meripilus, tal como del *M. giganteus*.

## ES 2 356 614 T3

Se entenderá que para las especies mencionadas, la invención comprende tanto el estado imperfecto como el perfecto, y otros equivalentes taxonómicos, por ejemplo, anamorfos, sin tener en cuenta el nombre de la especie por la cual estos se conocen. Los expertos en la técnica reconocen fácilmente la identidad de equivalentes apropiados.

5 Cepas de estas especies son fácilmente accesibles al público en un número de colecciones de cultivos, tales como la American Type Culture Collection (ATCC), la Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ), la Centraalbureau voor Schimmelcultures (CBS), y la Agricultural Research Service Patent Culture Collection, Northern Regional Research Center (NRRL).

10 En una forma de realización particularmente preferida de la invención el polipéptido de la invención se deriva de la cepa de *H. insolens* descrita en WO9117243 y depositada el 14 de abril de 1980 en la German Collection of Microorganisms and Cell cultures (DSMZ, Deutsche Sammlung von Mikroorganismen und Zellkulturen) Göttingen, Alemania, bajo el número DSM 1800) conforme a las provisiones del Tratado de Budapest.

15 Además, tales polipéptidos se pueden identificar y obtener de otras fuentes incluyendo microorganismos aislados de la naturaleza (p. ej., tierra, abonos, agua, etc.) usando las sondas mencionadas arriba. Técnicas para aislar microorganismos a partir de hábitats naturales se conocen en la técnica. La secuencia de ácidos nucleicos se puede derivar después seleccionando por similitud un genoma o una genoteca de ADNc de otro microorganismo. Una vez se ha detectado una secuencia de ácidos nucleicos codificando un polipéptido con la(s) sonda(s), la secuencia se puede aislar o clonar utilizando técnicas conocidas por aquellos con un conocimiento común de la técnica (véase, por ejemplo, Sambrook *et al.*, 1989, *supra*).

20 Los polipéptidos codificados por secuencias de ácidos nucleicos de la presente invención también incluyen polipéptidos fusionados o polipéptidos de fusión divisibles en los cuales se fusiona otro polipéptido en el N-término o el C-término del polipéptido o fragmento del mismo. Un polipéptido fusionado se produce por fusión de una secuencia de ácidos nucleicos (o una parte de la misma) codificando otro polipéptido a una secuencia de ácidos nucleicos (o una parte de la misma) de la presente invención. Técnicas para producir polipéptidos de fusión se conocen en la técnica, e incluyen ligar las secuencias de codificación codificando los polipéptidos de modo que éstas están estructurados y que la expresión del polipéptido fusionado está bajo el control de los mismos promotor(es) y terminator.

### 30 *Secuencias de ácidos nucleicos*

La presente invención también se refiere a secuencias de ácidos nucleicos aisladas las cuales codifican un polipéptido de la presente invención.

35 En una forma de realización interesante, la secuencia de ácidos nucleicos tiene al menos un 80% de identidad con la secuencia de ácidos nucleicos mostrada como los nucleótidos 1 a 1677 de la SEC ID NO:1. Preferiblemente, la secuencia de ácidos nucleicos tiene al menos un 85%, al menos un 90%, al menos un 95%, al menos un 96%, al menos un 97%, al menos un 98%, o al menos un 99% de identidad con cualquiera de las secuencias de ácidos nucleicos mostrada como los nucleótidos 1 a 1677 de la SEC ID NO:1. En otra forma de realización interesante de la invención la secuencia de ácidos nucleicos comprende la secuencia de aminoácidos mostrada como los nucleótidos 1 a 1677 de la SEC ID NO:1, una variante alélica de la misma, o un fragmento de la misma capaz de codificar un polipéptido según la invención. Obviamente, la secuencia de ácidos nucleicos puede consistir en la secuencia de aminoácidos mostrada como los nucleótidos 1 a 1677 de la SEC ID NO:1. La presente invención también comprende las secuencias de ácidos nucleicos que codifican un polipéptido teniendo la secuencia de aminoácidos de la SEC ID NO:2, o los polipéptidos maduros de la misma, la cual difiere de la SEC ID NO:1 en virtud de la degeneración del código genético. La presente invención también se refiere a subsecuencias de la SEC ID NO:1 la cual codifica fragmentos de la SEC ID NO:2 teniendo actividad de arabinofuranosidasa. Una subsecuencia de la SEC ID NO:1 es una secuencia de ácidos nucleicos comprendida por los nucleótidos 1 a 1677 de la SEC ID NO:1 excepto porque el o los nucleótidos del extremo 5' y/o 3' se han deletado.

45 La presente invención también se refiere a aislar secuencias de ácidos nucleicos codificando un polipéptido de la presente invención, los cuales hibridan bajo condiciones de astringencia altas, con (i) una cadena complementaria de la secuencia de ácidos nucleicos mostrada como los nucleótidos 1 a 1677 de la SEC ID NO:1, o (ii) una subsecuencia de (i) de al menos 100 nucleótidos. La presente invención también se refiere a cadenas complementarias de (i), y (ii).

50 Las técnicas usadas para aislar o clonar una secuencia de ácidos nucleicos codificando un polipéptido se conocen en la técnica e incluyen el aislamiento a partir de ADN genómico, la preparación de ADNc, o una combinación de las mismas. La clonación de las secuencias de ácidos nucleicos de la presente invención a partir de tal ADN genómico se puede efectuar, por ejemplo, usando la bien conocida reacción en cadena de la polimerasa (PCR) o seleccionando anticuerpos de genotecas de expresión para detectar fragmentos de ADN clonados con características estructurales compartidas. Véase, por ejemplo, Innis *et al.*, 1990, PCR: A Guide to Methods and Application, Academic Press, Nueva York. Se pueden utilizar otros procedimientos de amplificación de ácidos nucleicos tales como la reacción en cadena de la ligasa (LCR), la transcripción activada ligada (LAT) y la amplificación en base a la secuencia de ácidos nucleicos (NASBA). La secuencia de ácidos nucleicos se puede clonar a partir de una cepa de *Humicola insolens* o de una cepa de *Miripilus giganteus*, u otro o un organismo relacionado y puede, por ejemplo, ser una variante alélica o de especie de la región de codificación de polipéptido de la secuencia de ácidos nucleicos.

## ES 2 356 614 T3

Una secuencia de ácidos nucleicos aislada puede, por ejemplo, obtenerse por procedimientos de clonación estándar usados en la ingeniería genética para recolocar la secuencia de ácidos nucleicos de su ubicación natural a un sitio diferente donde ésta se va a reproducir. Los procedimientos de clonación pueden implicar la escisión y el aislamiento de un fragmento de ácidos nucleicos deseado comprendiendo la secuencia de ácidos nucleicos codificando el polipéptido, la inserción del fragmento en una molécula de vector, y la incorporación del vector recombinante en una célula huésped donde se replicarán múltiples copias o clones de la secuencia de ácidos nucleicos. La secuencia de ácidos nucleicos puede ser de origen genómico, de ADNc, de ARN, semisintético, sintético, o cualquier combinación de los mismos.

Para objetivos de la presente invención, el grado de identidad entre dos secuencias de ácidos nucleicos se determina como se ha descrito anteriormente.

La modificación de una secuencia de ácidos nucleicos codificando un polipéptido de la presente invención puede ser necesaria para la síntesis de polipéptidos sustancialmente similares al polipéptido. El término "sustancialmente similar" al polipéptido se refiere a formas del polipéptido ocurriendo de forma no natural. Estos polipéptidos pueden diferir en alguna vía de modificación genética del polipéptido aislado de su fuente nativa, por ejemplo, variantes que difieren en actividad específica, termostabilidad, pH óptimo, o similar. La secuencia variante se puede construir basándose en la secuencia de ácidos nucleicos presentada como la parte de codificación de polipéptido de la SEC ID NO:1, por ejemplo, una subsecuencia de la misma, y/o por introducción de sustituciones de nucleótido que no dan lugar a otra secuencia de aminoácidos del polipéptido codificadas por la secuencia de ácidos nucleicos, pero las cuales corresponden al uso de codón del organismo huésped destinado para la producción de la enzima, o por introducción de sustituciones de nucleótido que pueden dar lugar a una secuencia de aminoácidos diferente. Para una descripción general de la sustitución de nucleótidos, véase, por ejemplo, Ford *et al.*, 1991, *Protein. Expression and Purification* 2: 95-107.

Será aparente para los expertos en la técnica que tales sustituciones se pueden hacer fuera de las regiones críticas para la función de la molécula y todavía resultar en un polipéptido activo. Los residuos de aminoácidos esenciales para la actividad del polipéptido codificado por la secuencia de ácidos nucleicos aislada de la invención, y por lo tanto preferiblemente no sujeta a sustitución, se puede identificar según procedimientos conocidos en la técnica, tales como mutagénesis dirigida o mutagénesis de escaneo por alanina (véase, por ejemplo, Cunningham y Wells, 1989, *Science* 244: 1081-1085). En ésta técnica, se introducen mutaciones en cada residuo positivamente cargado en la molécula, y las moléculas mutantes resultantes se evalúan para conocer la actividad de arabinofuranosidasa para identificar los residuos de aminoácidos que son críticos para la actividad de la molécula. También se pueden determinar sitios de interacción sustrato-enzima por análisis de la estructura tridimensional como se determina por técnicas tales como el análisis de resonancia magnética nuclear, la cristalografía o el marcado por fotoafinidad (véase, por ejemplo, de Vos *et al.*, 1992, *Science* 255: 306-312; Smith *et al.*, 1992, *Journal of Molecular Biology*. 224: 899-904; Wlodaver *et al.*, 1992, *FEBS Letters* 309: 59-64).

### *Constructos de ácidos nucleicos*

La presente invención también se refiere a constructos de ácidos nucleicos comprendiendo una secuencia de ácidos nucleicos de la presente invención operativamente vinculada a una o más secuencias de control capaces de dirigir la expresión del polipéptido en una célula huésped adecuada.

Una secuencia de ácidos nucleicos aislada codificando un polipéptido de la presente invención se puede manipular en una variedad de formas para proporcionar la expresión del polipéptido. La manipulación de la secuencia de ácidos nucleicos antes de su inserción en un vector puede ser deseable o necesaria dependiendo del vector de expresión. Las técnicas para modificar secuencias de ácidos nucleicos utilizando métodos de ADN recombinantes se conocen bien en la técnica.

Las secuencias de control incluyen todos los componentes que son necesarios o ventajosos para la expresión de un polipéptido de la presente invención. Cada secuencia de control puede ser nativa o extraña para la secuencia de ácidos nucleicos codificando el polipéptido. Tales secuencias de control incluyen, pero de forma no limitativa, un líder, secuencia de poliadenilación, secuencia de propéptido, promotor, secuencia de péptido señal, y terminador de transcripción. Como mínimo, las secuencias de control incluyen un promotor, y señales de parada traduccionales y transcripcionales. Las secuencias de control pueden estar provistas de enlazadores con motivo de sitios de restricción específicos de introducción ligación de facilitando de las secuencias de control con la región de codificación de la secuencia de ácidos nucleicos codificando un polipéptido.

La secuencia de control puede ser una secuencia promotora apropiada, una secuencia de ácidos nucleicos que se reconoce por una célula huésped para la expresión de la secuencia de ácidos nucleicos. La secuencia promotora contiene secuencias de control transcripcionales que median la expresión del polipéptido. El promotor puede ser cualquier secuencia de ácidos nucleicos que muestre actividad transcripcional en la célula huésped de elección incluyendo promotores mutantes, truncados, e híbridos, y se pueden obtener de genes codificando polipéptidos intracelulares o extracelulares o bien homólogos o heterólogos a la célula huésped.

Ejemplos de promotores adecuados para dirigir la transcripción de los constructos de ácidos nucleicos de la presente invención, especialmente en una célula huésped bacteriana, son los promotores obtenidos del operón lac de *E. coli*, el gen de agarasa de *Streptomyces coelicolor* (dagA), el gen de levansucrasa de *Bacillus subtilis* (sacB), el gen de

## ES 2 356 614 T3

alfa-amilasa de *Bacillus licheniformis* (amyL), el gen de amilasa maltogénica de *Bacillus stearothermophilus* (amyM), el gen de alfa-amilasa de *Bacillus amyloliquefaciens* (amyQ), el gen de penicilinas de *Bacillus licheniformis* (penP), los genes de *Bacillus subtilis* xylA y xylB, y el gen de beta-lactamasa procariótico (Villa-Kamaroff *et al.*, 1978, Proceedings of the National Academy of Sciences USA 75: 3727-3731), al igual que el promotor tac (DeBoer *et al.*, 1983, Proceedings of the National Academy of Sciences USA 80: 21-25). Otros promotores se describen en “Useful proteins from recombinant bacteria” en Scientific American, 1980, 242: 74-94; y en Sambrook *et al.*, 1989, *supra*.

Ejemplos de promotores adecuados para dirigir la transcripción de los constructos de ácidos nucleicos de la presente invención en una célula huésped filamentosa fúngica son promotores obtenidos de los genes para la amilasa TAKA de *Aspergillus oryzae*, la proteinasa aspártica de *Rhizomucor miehei*, la alfa-amilasa neutra de *Aspergillus niger*, la alfa-amilasa estable en ácido de *Aspergillus niger*, la glucoamilasa de *Aspergillus Niger* o *Aspergillus awamori* (glaA), la lipasa de *Rhizomucor miehei*, la proteasa alcalina de *Aspergillus oryzae*, la triosa fosfato isomerasa de *Aspergillus oryzae*, la acetamidasa de *Aspergillus nidulans*, y la proteasa tipo tripsina de *Fusarium oxysporum* (WO 96/00787), al igual que el promotor NA2-tpi (un híbrido de los promotores de los genes para la alfa-amilasa neutra de *Aspergillus niger* y la triosa fosfato isomerasa de *Aspergillus oryzae*), y promotores mutantes, truncados, e híbridos de los mismos.

En un huésped de levadura, promotores útiles se obtienen a partir de los genes para enolasa de *Saccharomyces cerevisiae* (ENO-1), galactocinasa de *Saccharomyces cerevisiae* (GAL1), alcohol dehidrogenasa de *Saccharomyces cerevisiae*/dehidrogenasa de gliceraldehído-3-fosfato (ADH2/GAP), y 3-fosfoglicerato-quinasa de *Saccharomyces cerevisiae*. Otros promotores útiles para células huéspedes de levadura están descritos por Romanos *et al.*, 1992, Yeast 8: 423-488.

La secuencia de control también puede ser una secuencia del terminador de transcripción adecuada, una secuencia reconocida por una célula huésped para terminar la transcripción. La secuencia del terminador está operativamente enlazada al término 3' de la secuencia de ácidos nucleicos codificando el polipéptido. Cualquier terminador que sea funcional en la célula huésped de elección se puede utilizar en la presente invención.

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

Terminadores preferidos para células huéspedes de levadura se obtienen a partir de los genes para la enolasa de *Saccharomyces cerevisiae*, el citocromo C de *Saccharomyces cerevisiae* (CYC1), y la gliceraldehído-3-fosfato dehidrogenasa de *Saccharomyces cerevisiae*. Otros terminadores útiles para células huéspedes de levadura están descritos por Romanos *et al.*, 1992, *supra*.

La secuencia de control también puede ser una secuencia guía adecuada, una región no traducida de un ARNM el cual es importante para la traducción por la célula huésped. La secuencia guía está operativamente enlazada al término 5' de la secuencia de ácidos nucleicos codificando el polipéptido. Cualquier secuencia guía que sea funcional en la célula huésped de elección se puede utilizar en la presente invención.

Líderes preferidos para células huéspedes fúngicas filamentosas se obtienen a partir de los genes para la amilasa TAKA de *Aspergillus oryzae* y la triosa fosfato isomerasa de *Aspergillus nidulans*.

Líderes adecuados para células huéspedes de levadura se obtienen a partir de los genes para la enolasa de *Saccharomyces cerevisiae* (ENO-1), la fosfoglicerato-quinasa de *Saccharomyces cerevisiae*, el alfa-factor de *Saccharomyces cerevisiae*, y la alcohol dehidrogenasa de *Saccharomyces cerevisiae*/gliceraldehído-3-fosfato dehidrogenasa (ADH2/GAP).

La secuencia de control también puede ser una secuencia de poliadenilación, una secuencia operativamente enlazada al término 3' de la secuencia de ácidos nucleicos y la cual, transcrita, se reconoce por la célula huésped como una señal para añadir residuos de poliadenosina a ARNM transcrito. Cualquier secuencia de poliadenilación que sea funcional en la célula huésped de elección se puede utilizar en la presente invención.

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

Secuencias de poliadenilación útiles para células huéspedes de levadura están descritas por Guo y Sherman, 1995, Molecular Cellular Biology 15: 5983-5990.

La secuencia de control también puede ser una región de codificación de péptido señal que codifica una secuencia de aminoácidos enlazada al amino término de un polipéptido y dirige el polipéptido codificado en la vía secretora de la célula. El extremo 5' de la secuencia codificante de la secuencia de ácidos nucleicos puede contener intrínsecamente una región de codificación de péptido señal naturalmente enlazada en marco de lectura de traducción con el segmento de la región de codificación que codifica el polipéptido segregado. Alternativamente, el extremo 5' de la secuencia codificante puede contener una región de codificación de péptido señal foránea para la secuencia codificante. La región de codificación del péptido señal foránea se puede requerir donde la secuencia codificante no contiene naturalmente



una región de codificación de péptido señal. Alternativamente, la región de codificación de péptido señal foránea puede simplemente reemplazar la región de codificación de péptido señal natural para mejorar la secreción del polipéptido. No obstante, cualquier región de codificación de péptido señal que dirige el polipéptido expresado en la vía secretora de una célula huésped de elección se puede utilizar en la presente invención.

5 Regiones de codificación de péptido señales eficaces para células huéspedes bacterianas son las regiones de codificación de péptido señal obtenidas a partir de los genes para amilasa maltogénica NCIB 11837 de *Bacillus*, alfa-amilasa de *Bacillus stearothermophilus*, subtilisina de *Bacillus licheniformis*, beta-lactamasa de *Bacillus licheniformis*, proteasas neutras de *Bacillus stearothermophilus* (nprT, nprS, nprM), y prsA de *Bacillus subtilis*. Otros péptidos señal  
10 están descritos por Simonen y Palva, 1993, *Microbiological Reviews* 57: 109-137.

Regiones de codificación de péptido señales efectivas para células huéspedes filamentosas fúngicas son las regiones de codificación de péptido señal obtenidas de los genes para amilasa TAKA de *Aspergillus oryzae*, amilasa neutra de *Aspergillus niger*, glucoamilasa de *Aspergillus niger*, proteinasa aspártica de *Rhizomucor miehei*, celulasa de *Humicola insolens*, y lipasa de *Humicola lanuginosa*.  
15

Péptidos señal útiles para células huéspedes de levadura se obtienen de los genes para alfa-factor de *Saccharomyces cerevisiae* e invertasa de *Saccharomyces cerevisiae*. Otras regiones de codificación de péptido señal útiles están descritas por Romanos *et al.*, 1992, *supra*.  
20

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

Donde tanto el péptido señal como regiones de propéptido están presentes en el amino término de un polipéptido,  
30 la región de propéptido está situada junto al amino término de un polipéptido y la región del péptido señal está situada junto al amino término de la región del propéptido.

También puede ser deseable añadir secuencias reguladoras las cuales permiten la regulación de la expresión del polipéptido en relación al crecimiento de la célula huésped. Ejemplos de sistemas reguladores son aquellos que causan la expresión del gen a activar o desactivar en respuesta a un estímulo físico o químico, incluyendo la presencia de un compuesto regulador. Sistemas reguladores en sistemas procarióticos incluyen los sistemas operadores lac, tac, y trp. En levadura se pueden usar el sistema ADH2 o el sistema GAL1. En hongos filamentosos se pueden usar el promotor TAKA de alfa-amilasa, el promotor de glucoamilasa de *Aspergillus niger*, y el promotor de glucoamilasa de *Aspergillus oryzae* como secuencias reguladoras. Otros ejemplos de secuencias reguladoras son los que permiten amplificación  
40 génica. En sistemas eucarióticos, estos incluyen el gen de la dihidrofolato reductasa el cual se amplifica en presencia de metotrexato, y los genes de metalotioneína los cuales se amplifican con metales pesados. En estos casos, la secuencia de ácidos nucleicos codificando el polipéptido se vincularía operativamente con la secuencia reguladora.

#### 55 Vectores de expresión

La presente invención también se refiere a vectores de expresión recombinantes comprendiendo una secuencia de ácidos nucleicos de la presente invención, un promotor, y señales de parada traduccionales y transcripcionales. Los diferentes ácidos nucleicos y secuencias de control anteriormente descritas se pueden unir para producir un vector de expresión recombinante el cual puede incluir uno o más sitios de restricción convenientes para permitir la inserción o  
50 sustitución de la secuencia de ácidos nucleicos codificando el polipéptido en tales sitios. Alternativamente, la secuencia de ácidos nucleicos de la presente invención se puede expresar por inserción de la secuencia de ácidos nucleicos o un constructo de ácidos nucleicos comprendiendo la secuencia en un vector apropiado para la expresión. En la creación el vector de expresión, la secuencia codificante se localiza en el vector de modo que la secuencia codificante está operativamente enlazada con las secuencias de control apropiadas para la expresión.  
55

El vector de expresión recombinante puede ser cualquier vector (p. ej., un plásmido o virus) el cual se puede someter convenientemente a procedimientos de ADN recombinante y puede causar la expresión de la secuencia de ácidos nucleicos. La elección del vector dependerá normalmente de la compatibilidad del vector con la célula huésped en la cual se va a introducir el vector. Los vectores pueden ser plásmidos circulares cerrados o lineales.  
60

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

## ES 2 356 614 T3

Los vectores de la presente invención contienen preferiblemente uno o más marcadores seleccionables los cuales permiten una selección fácil de células transformadas. Un marcador seleccionable es un gen el producto del cual proporciona resistencia biocida o vírica, resistencia a metales pesados, prototrofia a auxótrofos, y similares. Ejemplos de marcadores seleccionables bacterianos son los genes *theda* de *Bacillus subtilis* o *Bacillus licheniformis*, o marcadores que confieren resistencia antibiótica tal como resistencia a ampicilina, canamicina, cloranfenicol o tetraciclina. Marcadores adecuados para células huéspedes de levadura son ADE2, HIS3, LEU2, LiS2, MET3, TRP1, y URA3. Marcadores seleccionables para el uso en una célula huésped filamentosa fúngica incluyen, pero de forma no limitativa, *amdS* (acetamidasa), *argB* (ornitina-carbamoiltransferasa), *bar* (fosfonitricina acetiltransferasa), *hygB* (higromicina fosfotransferasa), *niaD* (nitrato-reductasa), *pyrG* (orotidina-5'-fosfato-descarboxilasa), *sC* (sulfato adeniltransferasa), *trpC* (antranilato sintasa), al igual que equivalentes de los mismos. Se prefieren para el uso en una célula de *Aspergillus* los genes *amdS* y *pyrG* de *Aspergillus nidulans* o *Aspergillus oryzae* y el gen *bar* de *Streptomyces higroscopicus*.

Los vectores de la presente invención contienen preferiblemente un(os) elemento(s) los cuales permiten la integración estable del vector en el genoma de la célula huésped o la replicación autónoma del vector en la célula independiente del genoma.

Para la integración en el genoma de célula huésped, el vector puede depender la secuencia de ácidos nucleicos codificando el polipéptido o cualquier otro elemento del vector para la integración estable del vector en el genoma por recombinación homóloga o no homóloga. Alternativamente, el vector puede contener secuencias de ácidos nucleicos adicionales para dirigir la integración por recombinación homóloga en el genoma de la célula huésped. Las secuencias de ácidos nucleicos adicionales permiten que el vector se integre en el genoma de la célula huésped en una(s) ubicación(es) precisa(s) en el/los cromosoma(s). Para aumentar la probabilidad de integración en una ubicación precisa, los elementos integracionales deberían contener preferiblemente un número suficiente de ácidos nucleicos, tal como de 100 a 1.500 pares de bases, preferiblemente de 400 a 1.500 pares de bases, y de la forma más preferible de 800 a 1.500 pares de bases, los cuales son altamente homólogos a la secuencia objetivo correspondiente para mejorar la probabilidad de recombinación homóloga. Los elementos integracionales pueden ser cualquier secuencia que sea homóloga a la secuencia objetivo en el genoma de la célula huésped.

Además, los elementos integracionales pueden ser secuencias de ácidos nucleicos no codificantes o codificantes. En cambio, el vector se puede integrar en el genoma de la célula huésped por recombinación no homóloga.

Para la replicación autónoma, el vector puede comprender además un origen de replicación permitiendo que vector se replique de manera autónoma en la célula huésped en cuestión. Ejemplos de orígenes bacterianos de replicación son los orígenes de replicación de los plásmidos pBR322; pUC19; pACYC177, y pACYC184 permitiendo la replicación en *E. coli*, y pUB110; pE194,pTA1060, y pAM $\beta$ 1 permitiendo la replicación en *Bacillus*. Ejemplos de orígenes de replicación para el uso en una célula huésped de levadura son el origen de replicación de 2 micrones, ARS1, ARS4, la combinación de ARS1 y CEN3, y la combinación de ARS4 y CEN6. El origen de replicación puede ser uno teniendo una mutación que hace su función de sensibilidad a la temperatura en la célula huésped (véase, por ejemplo, Ehrlich, 1978, Proceedings of the National Academy of Sciences EEUU 75: 1433).

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

Los procedimientos usados para enlazar los elementos anteriormente descritos para construir los vectores de expresión recombinantes de la presente invención son conocidos por un experto en la materia (véase, por ejemplo, Sambrook *et al.*, 1989, *supra*).

### Células huéspedes

La presente invención también se refiere a células huéspedes recombinantes, comprendiendo una secuencia de ácidos nucleicos de la invención, la cual se usa ventajosamente en la producción recombinante de los polipéptidos.

Un vector comprendiendo una secuencia de ácidos nucleicos de la presente invención se introduce en una célula huésped de modo que el vector se mantiene como un integrante cromosómico o como un vector extracromosómico de autoreplicación como se describe anteriormente. La elección de una célula huésped depende en gran parte del gen que codifica el polipéptido y su fuente.

La célula huésped puede ser un microorganismo unicelular, por ejemplo, una procariota, o un microorganismo no unicelular, por ejemplo, una eucariota.

Son células unicelulares útiles células bacterianas tales como bacterias gram positivas incluyendo, pero no limitándose a, una célula de *Bacillus*, por ejemplo, *Bacillus alkalophilus*, *Bacillus amyloliquefaciens*, *Bacillus brevis*, *Bacillus circulans*, *bacillus clausii*, *Bacillus coagulans*, *Bacillus lautus*, *Bacillus lentus*, *Bacillus licheniformis*, *Bacillus me-*

gaterium, *Bacillus stearothermophilus*, *Bacillus subtilis*, y *Bacillus thuringiensis*; o una célula de *Streptomyces*, por ejemplo, *Streptomyces lividans* o *Streptomyces murinus*, o bacterias gram negativas tales como *E. coli* y *Pseudomonas sp.* en una forma de realización preferida, la célula huésped bacteriana es un *Bacillus lentus*, *Bacillus licheniformis*, *Bacillus stearothermophilus*, o célula de *Bacillus subtilis*. En otra forma de realización preferida, la célula de *Bacillus* es un *Bacillus alcalofílico*.

La introducción de un vector en una célula huésped bacteriana puede, por ejemplo, efectuarse por transformación de protoplasto (véase, por ejemplo, Chang y Cohen, 1979, *Molecular General Genetics* 168: 111-115), usando células competentes (véase, por ejemplo, Young y Spizizin, 1961, *Journal of Bacteriology* 81: 823-829, o Dubnau y Davidoff-Abelson, 1971, *Journal of Molecular Biology* 56: 209-221), electroporación (véase, por ejemplo, Shigekawa y Dower, 1988 *Biotechniques* 6: de 742 751), o conjugación (véase, por ejemplo, Koehler y Thorne, 1987, *Journal of Bacteriology* 169: 5771-5278).

La célula huésped puede ser una eucariota, tal como un mamífero, insecto, planta, o célula fúngica. En una forma de realización preferida, la célula huésped es una célula fúngica. "Fungi" como se utiliza en este caso incluye los filos Ascomycota, Basidiomycota, Chitridiomycota, y Zigomycota (como se define por Hawksworth *et al.*, In, Ainsworth and Bisby's Dictionary of The Fungi, 8ª edición, 1995, cabina CAB International, University Press, Cambridge, Reino Unido) al igual que el Oomycota (como se cita en Hawksworth *et al.*, 1995, *supra*, página 171) y todos los hongos mitospóricos (Hawksworth *et al.*, 1995, *supra*).

En una forma de realización más preferida, la célula huésped fúngica es una célula de levadura. "Levadura" como se utiliza en este caso incluye levadura ascoesporógena (Endomycetales), levadura basidiosporógena, y levadura perteneciente a los *Fungi Imperfecti* (Blastomycetes). Ya que la clasificación de la levadura puede cambiar en el futuro, para los objetivos de esta invención, la levadura se debe definir como se describe en *Biology and Activities of Yeast* (Skinner, F.A., Passmore, S.M., y Davenport, R.R., eds., Soc. App. Bacteriol. Symposium Series n° 9, 1980).

En una forma de realización aún más preferida, la célula huésped de levadura es una célula *Candida*, *Hansenula*, *Kluyveromices*, *Pichia*, *Saccharomyces*, *Schizosaccharomyces*, o *Yarrowia*.

En una forma de realización más preferida, la célula huésped de levadura es una célula de *Saccharomyces carlsbergensis*, *Saccharomyces cerevisiae*, *Saccharomyces diastaticus*, *Saccharomyces douglasii*, *Saccharomyces kluyveri*, *Saccharomyces norbensis* o *Saccharomyces oviformis*. En otra forma de realización más preferida, la célula huésped de levadura es una célula de *Kluyveromyces lactis*. En otra forma de realización más preferida, la célula huésped de levadura es una célula de *Yarrowia lipolytica*.

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

En una forma de realización aún más preferida, la célula huésped fúngico filamentosa es una célula de una especie de, pero sin limitarse a, *Acremonium*, *Aspergillus*, *Fusarium*, *Humicola*, *Mucor*, *Myceliophthora*, *Neurospora*, *Penicillium*, *Thielavia*, *Tolyptocladium*, o *Trichoderma*.

En una forma más de realización preferida, la célula huésped fúngica filamentosa es una célula de *Aspergillus awamori*, *Aspergillus foetidus*, *Aspergillus japonicus*, *Aspergillus nidulans*, *Aspergillus niger* o *Aspergillus oryzae*. En otra forma más de realización preferida, la célula huésped fúngica filamentosa es una célula de *Bactridioides fusarium*, *Fusarium cerealis*, *Fusarium crookwellense*, *Fusarium culmorum*, *Fusarium graminearum*, *Fusarium graminum*, *Fusarium heterosporum*, *Fusarium negundi*, *Fusarium oxysporum*, *Fusarium reticulatum*, *Fusarium roseum*, *Fusarium sambucinum*, *Fusarium sarcochromum*, *Fusarium sporotrichioides*, *Fusarium sulphureum*, *Fusarium torulosum*, *Fusarium trichothecioides*, o *Fusarium venenatum*. En una forma de realización incluso más preferida, la célula parental fúngica filamentosa es una célula de *Fusarium Venenatum* (*Nirenberg sp. nov.*). En otra forma de realización más preferida, la célula huésped fúngica filamentosa es una célula de *Humicola insolens*, *Humicola lanuginosa*, *Mucor miehei*, *Myceliophthora thermophila*, *Neurospora crassa*, *Penicillium purpurogenum*, *Thielavia terrestris*, *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei*, o *Trichoderma viride*.

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

## ES 2 356 614 T3

### *Métodos de producción*

La presente invención también se refiere a métodos para producir un polipéptido de la presente invención, el método comprendiendo (a) cultivar una cepa del género *Humicola* para producir un sobrenadante comprendiendo el polipéptido; y (b) recuperar el polipéptido. Preferiblemente, la cepa es de la especie *Humicola insolens*.

La presente invención también se refiere a un método para la producción de un polipéptido de la invención, el método comprendiendo (a) cultivar una célula huésped recombinante como se ha descrito anteriormente bajo condiciones propicias para la producción del polipéptido, y (b) recuperar el polipéptido a partir de las células y/o del medio de cultivo.

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

Los polipéptidos se pueden detectar usando métodos conocidos en la técnica que sean específicos para los polipéptidos. Estos métodos de detección pueden incluir el uso de anticuerpos específicos, la formación de un producto enzimático, o la desaparición de un sustrato enzimático. Por ejemplo, un ensayo enzimático se puede utilizar para determinar la actividad del polipéptido como se describe en este caso.

El polipéptido resultante se puede recuperar por métodos conocidos en la técnica. Por ejemplo, el polipéptido se puede recuperar a partir del medio nutritivo por procedimientos convencionales incluyendo, pero no limitándose a, centrifugado, filtración, extracción, secado por pulverización, evaporación, o precipitación.

Los polipéptidos de la presente invención se pueden purificar por una variedad de procedimientos conocidos en la técnica incluyendo, pero no limitándose a, cromatografía (p. ej., intercambio iónico, afinidad, hidrofobia, cromatografía de exclusión de tamaño), procedimientos electroforéticos (p. ej., isoelectroenfoque preparatorio), solubilidad diferencial (p. ej., precipitación de sulfato amónico), SDS-PAGE, o extracción (véase, por ejemplo, Protein Purification, J.-C. Janson y Lars Ryden, editores, VCH Publishers, Nueva York, 1989).

### *Expresión de las enzimas en plantas*

Una secuencia de ADN que codifica un polipéptido de interés, tal como una arabinofuranosidasa de la presente invención, se puede transformar y expresar en plantas transgénicas como se describe abajo.

La planta transgénica puede ser dicotiledónea o monocotiledónea, para abreviar una dicot o una monocot. Ejemplos de plantas monocot son hierbas, tales como poa de los prados (poa pratense, Poa), hierba forrajera tal como *Festuca*, *Lolium*, césped templado, tal como *Agrostis*, y cereales, por ejemplo trigo, avena, centeno, cebada, arroz, sorgo y maíz (en grano).

Ejemplos de plantas dicot son tabaco, leguminosas, tales como altramuces, patata, remolacha azucarera, guisante, moldura y semilla de soja, y plantas crucíferas (familia Brassicaceae), tal como coliflor, colza de semilla de aceite y el organismo modelo estrechamente relacionado de *Arabidopsis thaliana*.

Ejemplos de partes de planta son vástago, callo, hojas, raíz, frutos, semillas, y tubérculos al igual que los tejidos individuales comprendiendo estas partes, por ejemplo epidermis, mesofilo, parénquima, tejidos vasculares, meristemas. En el presente contexto, también los compartimentos de célula vegetal específicos, tales como cloroplasto, apoplasto, mitocondria, vacuola, peroxisomas y citoplasma se consideran como una parte de planta. Además, cualquier célula vegetal, cualquiera que sea el origen del tejido, se considera como una parte de planta. Asimismo, partes de planta tales como tejidos específicos y células aisladas para facilitar la utilización de la invención también se consideran partes de planta, por ejemplo embriones, endospermas, aleurona y revestimientos de semillas.

También se incluyen dentro del campo de la invención la progenie de tales plantas, partes de planta y células vegetales.

La planta transgénica o la célula vegetal expresando el polipéptido de interés se puede construir conforme a métodos conocidos en la técnica. En resumidas cuentas la planta o la célula vegetal se construye incorporando uno o más constructos de expresión codificando el polipéptido de interés en el genoma de huésped de la planta y propagando la planta modificada resultante o la célula vegetal en una planta transgénica o célula vegetal.

## ES 2 356 614 T3

Convenientemente, el constructo de expresión es un constructo de ADN que comprende un gen codificando el polipéptido de interés en asociación operable con secuencias reguladoras apropiadas requeridas para la expresión del gen en la planta o parte de la planta de elección. Además, el constructo de expresión puede comprender un marcador seleccionable útil para identificar células huéspedes en las cuales el constructo de expresión se ha integrado y secuencias de ADN necesarias para la introducción del constructo en la planta en cuestión (ésta depende del método de introducción de ADN a usar).

La elección de secuencias reguladoras, tales como secuencias de promotor y terminador y opcionalmente secuencias señal o de tránsito se determina, por ejemplo basándose en cuándo, dónde y cómo se desea que se exprese la enzima. Por ejemplo, la expresión del gen codificando la enzima de la invención puede ser constitutiva o inducible, o puede ser desarrollable, fase o tejido específica, y el producto genético se puede prever para un compartimiento celular específico, tejido o parte de planta tal como semillas u hojas. Secuencias reguladoras están, por ejemplo descritas por Tague *et al*, Plant, Phys., 86, 506, 1988.

Para la expresión constitutiva se pueden utilizar el 35S-CaMV, la ubiquitina de maíz 1 y el promotor de actina 1 de arroz (Franck *et al.* 1980. Cell 21: 285-294, Christensen AH, Sharrock RA y Quail 1992. Maize polyubiquitin genes: structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts by electroporation. Plant Mo. Biol. 18, 675-689.; Zhang W, McElroy D. y Wu R 1991, Analysis of rice Act1 5' region activity in transgenic rice plants. Plant Cell 3, 1155-1165). Promotores órgano-específicos pueden ser, por ejemplo un promotor a partir de tejidos sumidero de almacenamiento tales como semillas, tubérculos de patata, y frutos (Edwards & Coruzzi, 1990. Annu. Rev. Genet. 24: 275-303), o de tejidos sumidero metabólicos tales como meristemas (Ito *et al.*, 1994. Plant Mol. Biol. 24: 863-878), un promotor específico de semillas tal como el promotor de glutelina, prolamina, globulina o albúmina de arroz (Wu *et al.*, Plant and Cell Physiology Vol. 39, n° 8 págs. 885-889 (1998), un promotor de *Vicia faba* de la legumina B4 y el gen de proteína de semilla desconocido de promotor de *Vicia faba* descrito por Conrad U. et al, Journal of Plant Physiology Vol. 152, n° 6 págs. 708-711 (1998), un promotor de una proteína corporal de aceite de semilla (Chen *et al.*, Plant and cell physiology vol. 39, n° 9 págs. 935-941 (1998), el promotor napA de la proteína de almacenamiento de *Brassica napus*, o cualquier otro promotor de semilla específico conocido en la técnica, por ejemplo como se describe en WO 91/14772. Además, el promotor puede ser un promotor específico de hoja tal como el promotor de eritrocitos de arroz o tomate (Kyoizuka *et al.*, Plant Physiology Vol. 102, n° 3 págs. 991-1000 (1993), el promotor del gen de metiltransferasa de adenina del virus de Chlorella (Mitra, A. y Higgins, DW, Plant Molecular Biology Vol. 26, n° 1 págs. 85-93 (1994), o el promotor del gen aldP de arroz (Kagaya *et al.*, Molecular and General Genetics Vol. 248, n° 6 págs. 668-674 (1995), o un promotor inducible por herida tal como el promotor pin2 de patata (Xu *et al*, Plant Molecular Biology Vol. 22, n° 4 págs. 573-588 (1993). Asimismo, el promotor se puede inducir por tratamientos abióticos tales como la temperatura, sequía o alteraciones en salinidad o inducir por sustancias aplicadas exógenamente que activan el promotor, por ejemplo etanol, estrógenos, hormonas de planta como etileno, ácido abscísico y ácido giberélico y metales pesados.

Un elemento intensificador de promotor se puede utilizar para conseguir una mayor expresión de la enzima en la planta. Por ejemplo, el elemento intensificador de promotor puede ser un intrón el cual se coloca entre el promotor y la secuencia de nucleótidos codificando la enzima. Por ejemplo, Xu *et al.* op cit revelan el uso del primer intrón del gen de actina de arroz 1 para mejorar la expresión.

El gen marcador seleccionable y cualquier otra parte del constructo de expresión se pueden elegir de aquellas disponibles en la técnica.

El constructo de ADN se incorpora en el genoma de la planta según técnicas convencionales conocidas en la técnica, incluyendo transformación mediada por *Agrobacterium*, transformación mediada por virus, micro inyección, bombardeo de partículas, transformación biolística, y electroporación (Gasser *et al*, Science, 244, 1293; Potrikus, Bio/Techn. 8, 535, 1990; Shimamoto *et Al*, Nature, 338, 274, 1989).

Actualmente, la transferencia genética mediada por *Agrobacterium tumefaciens* es el método de elección para generar dicots (para revisión Hooikas & Schilperoort, 1992. Plant Mol. Biol. 19: 15-38), y también se pueden usar para transformar monocots, aunque frecuentemente se usan otros métodos de transformación para estas plantas. Actualmente, el método de elección para generar monocots transgénicas complementando el enfoque de *Agrobacterium* es el bombardeo de partículas (oro microscópico o partículas de tungsteno revestidas con el ADN de transformación) de callos embrionarios o desarrollando embriones (Christou, 1992. Plant. J. 2: 275-281; Shimamoto, 1994. Curr. Opin. Biotechnol. 5: 158-162; Vasil *et al.*, 1992. Bio/Technology 10: 667-674). Un método alternativo para la transformación de monocots se basa en la transformación de protoplasto como se describe por Omirulleh S, *et al.*, Plant Molecular biology vol. 21, n° 3 págs. 415-428 (1993).

Siguiendo la transformación, los transformantes teniendo incorporado el constructo de expresión se seleccionan y regeneran en la planta entera según métodos bien conocidos en la técnica. frecuentemente el procedimiento de transformación se diseña para la la eliminación selectiva de genes de selección o bien durante la regeneración o en las siguientes generaciones usando por ejemplo la co-transformación con dos constructos de T-ADN separados o la escisión sitio-específica del gen de selección por una recombinasa específica.

*Uso de arabinofuranosidasas*

La presente invención también se refiere al uso de los polipéptidos de la invención, es decir las arabinofuranosidasas, en diferentes aplicaciones industriales, por ejemplo en la conversión de biomasa, tal como en la producción de etanol combustible a partir de celulosa conteniendo biomasa, en la producción de combustible y/o etanol potable a partir de almidón, en la maceración para la producción de cerveza, en una masa para la producción de pan, o en producir un producto de pienso.

La invención además proporciona un proceso donde un arabinoxilano conteniendo sustrato y/o una biomasa se pone en contacto con una arabinofuranosidasa capaz de liberar arabinosa a partir de xilosas di-sustituídas. Preferiblemente la alfa-L-arabinofuranosidasa es una arabinofuranosidasa de GH43. La alfa-L-arabinofuranosidasa de GH43 se deriva preferiblemente de un origen bacteriano, fúngico o vegetal. Preferiblemente el arabinoxilano conteniendo sustrato y/o la biomasa se selecciona a partir de la lista que consiste en brotes de energía boscosos y/o herbáceos, alimentos agrícolas y brotes de alimentos, productos de pienso, tubérculos, raíces, tallos, leguminosas, cáscaras de mandioca, vainas de cacao, cáscaras de arroz y/o cáscaras, salvado de arroz a partir de arroz pulido, mazorcas, paja, cáscaras y/o cáscaras de grano de cereal, palote de caña de azúcar prensada, pulpa de remolacha azucarera, pulpa de semilla de algarroba, manzanas de fruta o vegetales, cereales o residuos de la cosecha agrícola de grano entero, paja, tallos, hojas, salvado de maíz, cáscaras, mazorcas, cáscara, cubiertas, vainas, residuos de madera, corteza, virutas, serrín, pulpa de madera, solución de reducción a pasta, papelón, cartón, construcción y residuos de madera de demolición, sólidos de aguas residuales industriales o municipales o lodo, abono, subproducto a partir de procesos de elaboración y/o fermentación, granos de destilación mojados, granos de destilación secos, grano consumido, vinasse y bagazo.

La presente invención también se refiere a composiciones comprendiendo los polipéptidos de la invención, es decir la arabinofuranosidasa, al igual que a usos de tales composiciones.

**Materiales y Métodos**

La arabinosa y la xilosa se compraron a través de Merck (Darmstadt, Alemania). Las arabinoxilanas de trigo solubles en agua en solubles en agua se obtuvieron a través de Megazyme, (Bray, County Wicklow, Irlanda).

*Enzimas*

Las alfa-L-arabinofuranosidasas se clonaron usando técnicas moleculares básicas (Sambrook *et al.*, 1989, Molecular Cloning, A Laboratory Manual, 2ª edición, Cold Spring Harbor, Nueva York, Christgau *et al.* 1995, Curr. Genet. 27, 135-141, Ausubel *et al.*, 2003, Curr. Prot. Mol. Biol., John Wiley & Sons, Cambridge, EEUU).

Las alfa-L-arabinofuranosidasas de *Bifidobacterium adolescentis* de Van Laere *et al.* (1997) y Van den Broek *et al.* (2005) se obtuvieron de Megazyme (Irlanda).

Shearzyme (GH10) y Pentopan Mono (GH11), preparaciones de endo-1,4-β-xilanasas monocomponente producidas por *Aspergillus aculeatus* y *Thermomyces lanuginosus*, respectivamente, eran productos comerciales de Novozymes A/S (Bagsvaerd, Dinamarca).

*Preparación de polímeros y oligosacáridos de arabinoxilano específicos*

El arabinoxilano doblemente sustituido se preparó incubando arabinoxilano de trigo soluble (1g) en 0,1 M de tampón de acetato (100 ml), a pH 6.0 con 0,167 g de α-L-arabinofuranosidasa de *Meripilus giganteus* (GH51) · kg<sup>-1</sup> de arabinoxilano de trigo soluble en agua durante 48 horas a 30°C. El arabinoxilano individualmente sustituido se preparó incubando arabinoxilano de trigo soluble en agua (1 g) en 0,1 M de tampón de acetato (42 ml), a pH 6.0 con 0,147 g de α-L-arabinofuranosidasa de *Humicola insolens* (GH43) · kg<sup>-1</sup> de arabinoxilano de trigo soluble en agua durante 48 horas a 30°C. Para detener las reacciones enzimáticas las mezclas se calentaron a 100°C durante 10 min. Se precipitaron polímeros de arabinoxilano por adición de etanol (126 ml). Los precipitados se filtraron (Miracloth) y se secaron en vacío.

Oligosacáridos conteniendo grupos de arabinosil enlazados al terminal (1í13) se prepararon incubando el arabinoxilano de trigo insoluble de agua (1g) en 0,1 M de tampón acetato (100 ml), a pH 6.0 con 6,67 g de Shearzyme (xilanasas GH10) · kg<sup>-1</sup> de arabinoxilano de trigo insoluble en agua durante 2 horas a 30°C. Los oligosacáridos conteniendo grupos de arabinosil enlazados al interno (1→3) se prepararon incubando arabinoxilano de trigo insoluble en agua (1g) en 0,1 M de tampón de acetato (100 ml), a pH 6.0 con 0,03 g de Pentopan Mono (xilanasas GH11) · kg<sup>-1</sup> de arabinoxilano de trigo insoluble en agua durante 2 horas a 30°C. Se prepararon oligosacáridos conteniendo grupos de arabinosil vinculados al interno (1→2) incubando arabinoxilano de trigo insoluble en agua (1 g) en 0,1 M de tampón de acetato (100 ml), a pH 6.0 con 0,03 g de Pentopan Mono (xilanasas GH11) · kg<sup>-1</sup> de arabinoxilano de trigo insoluble en agua y alfa-L-arabinofuranosidasa de *H. insolens* (GH43) · kg<sup>-1</sup> de arabinoxilano de trigo soluble en agua durante 2 horas a 30°C. Para detener las reacciones enzimáticas las mezclas se calentaron a 100°C durante 10 min. Los arabinoxilo-oligosacáridos se concentraron en un evaporador giratorio y se evaluaron por <sup>1</sup>H-NMR.

## ES 2 356 614 T3

### Ensayo de actividad hacia la actividad de alfa-L-arabinofuranosidasa

La actividad de alfa-L-arabinofuranosidasa se puede evaluar como se describe por Poutanen *et al.* (Appl. Microbiol. Biotechnol. 1988, 28, 425-432) usando 5 mM de alfa-L-arabinofuranosida de p-nitrofenilo como sustratos. Las reacciones se pueden realizar en 50 mM de tampón de citrato a pH 6.0, a 40°C con un tiempo de reacción total de 30 min. La reacción se detiene añadiendo 0,5 ml de 1 M de carbonato de sodio y el p-nitrofenol liberado se mide a 405 nm. La actividad se expresa en U/ml.

### Ensayo de actividad hacia el xilano C2- y C3-di-sustituido

El arabinoxilano de trigo hidrosoluble de viscosidad media (Megazyme, Bray, Irlanda) se trató con una alfa-arabinofuranosidasa de GH51 de *Meripilus giganteus* (SEC ID NO:2) para eliminar sustituyentes únicos de alfa-arabinofuranosil fijados a la C(O)-3 arabinosa del arabinoxilano para producir un sustrato de arabinoxilano di-sustituido con sustituyentes de arabinofuranosil fijados a ambos C(O)-2,3 de los residuos de xilosa. El sustrato se dializó y liofilizó.

Se preparó un 0,1% de solución del arabinoxilano di-sustituido y la actividad de alfa-arabinofuranosidasa se midió mezclando 0,1 ml de enzima, 0,9 ml de tampón (0,12 M de Ácido succínico, a pH 6.0) y 1,0 ml de solución de sustrato en un tubo Eppendorf. El tubo Eppendorf se incubó a 60°C durante 1 hora con agitación. La cantidad de arabinosa liberada se midió por HPAEC (cromatografía de intercambio aniónico de alto rendimiento).

### HPAEC

Se aplicaron hidrolizados (10  $\mu$ l) sobre un sistema Dionex BioLC ajustado con una columna de protección Dionex CarboPac<sup>TM</sup> PA1 (4 x 250 mm) (Dionex Corporation, Sunnyvale, CA, EEUU) combinado con una precolumna CarboPac<sup>TM</sup> PA1 (4 X 50 mm). Los monosacáridos se separaron isocráticamente con 10 mM de KOH durante 15 min, flujo: 1 mL·min<sup>-1</sup>. Se detectaron monosacáridos mediante un detector electroquímico pulsado en el modo de detección amperiométrico pulsado. El potencial del electrodo se programó para +0,1 V (T = 0-0,4 s) a -2,0 V (T = 0,41-0,42 s) a 0,6 V (T = 0,43 s) y finalmente -0,1 V (T = 0,44-0,50 s), mientras se integra la señal resultante de T = 0,2-0,4 s. Se usó una mezcla de arabinosa y xilosa (concentración de cada componente: 0,0025-0,1 g·L<sup>-1</sup>) como estándar.

### Análisis de <sup>1</sup>H-NMR

Todos los productos de degradación se liofilizaron dos veces a partir de un 99,9% de D2 O y se redisolviéron en un 99,9% de D2 O. Algunos hidrolizados se dializaron (peso molecular de membrana Spectra/Por máximo 1000) para eliminar arabinosa libre antes del análisis espectral. Los espectros de <sup>1</sup>H-NMR se registraron a 30°C en un instrumento Varian Mercury-VX funcionando a 400 MHz y equipado con una sonda auto-modificable de 4 núcleos. Los datos se recogieron de 128-512 sondeos y la señal HD0 se usó como una señal de referencia (4,67 ppm).

### Ejemplos

#### Ejemplo 1

El arabinoxilano de trigo comprende arabinofuranosida como un monosustituyente vinculado a la posición 3 de xilosa interna (A) y arabinofuranosida enlazada a las posiciones 3 (B) y 2 (C) en xilosa di-sustituida, respectivamente. Se produjeron sustratos cada uno comprendiendo sólo uno de los 3 tipos de enlaces de arabinofuranosida. La actividad de arabinofuranosidasas hacia estos sustratos se investigó usando <sup>1</sup>H RMN.

TABLA 1

*Origen, familias, y masa molecular de actividades de alfa-L-arabinofuranosidasa*

Origen	Familia	Masa molar (kDa)
<i>H. insolens</i> (SEC ID NO:2)	GH43	_62
<i>M. giganteus</i>	GH51	_69
<i>B. adolescentis</i>	GH43	_60
<i>H. insolens</i>	GH51	_94

ES 2 356 614 T3

TABLA 2

Actividad en polímeros de arabinoxilano seleccionados, incubación a pH 6, a 40°C durante 2 hrs.

Sustrato	Enlace	Enzima			
		<i>H. insolens</i> (GH43)	<i>B. adolescentis</i> (GH43)	<i>H. insolens</i> (GH51)	<i>M. giganteus</i> (GH51)
Arabinoxilano intacto	Mono-sustituido (1_3)	-	-	x	xx
	Di-sustituido (1_2)	-	-	-	-
	Di-sustituido (1_3)	xx	x	-	-
Arabinoxilano di-sustituido	Di-sustituido (1_2)	-	-	-	-
	Di-sustituido (1_3)	xx	xx	-	-
Arabinoxilano mono-sustituido	Mono-sustituido (1_2)	-	-	xx	xx
	Mono-sustituido (1_3)	-	-	xx	xx

xx se refiere a más del 75% de la hidrólisis, x(x) al 50-75% de la hidrólisis, x al 25-50% de la hidrólisis y (X) al 5-25% de la hidrólisis. - se refiere a la hidrólisis no detectable

Ejemplo 2

Se incubó arabinoxilano de trigo soluble con 0,1 g de proteína enzimática por kg de DM de alfa-L-arabinofuranosidasa a partir de *H. insolens* (GH43), *B. adolescentis* (GH43), *H. insolens* (GH51), y *M. giganteus* (GH51) durante 24 hrs. Se midió la arabinosa liberada. Los resultados se expresan en mg de arabinosa por g de arabinoxilano de trigo soluble en agua, como el promedio de determinaciones triplicadas, coeficiente de variación en una media de < 6.4.

TABLA 3

Arabinosa liberada a partir de arabinoxilano de trigo soluble tratado con alfa-L-arabinofuranosidasa

Origen	Condiciones de reacción	
	pH 6, 40 °C	pH 5, 50 °C
<i>H. insolens</i> (GH43)	128,0 a	147,0 a
<i>M. giganteus</i> (GH51)	48,15 c	121,0 b
<i>B. adolescentis</i> (GH43)	63,43 b	4,833 d
<i>H. insolens</i> (GH51)	20,75 d	18,47 c



## ES 2 356 614 T3

### Ejemplo 3

Se pueden aplicar arabinofuranosidasas en composiciones de pienso para aumentar la digestibilidad. El arabinoxilano de maíz se di-sustituye pesadamente con arabinosa. Para facilitar la degradación de xilano es ventajoso eliminar tanto como sea posible de los sustituyentes de arabinosa. La degradación *in vitro* de arabinoxilanas en una dieta basada en maíz suplementada con la arabinofuranosidasa GH51 de *Humicola insolens* y una xilanasa comercial derivada de *Thermomyces lanuginosus* se estudió en un sistema de digestión *in vitro*.

Condiciones para el sistema de digestión *in vitro*:

Sustrato:	800 mg 30:70 Harina de semillas de soja/Dieta a base de maíz molida y pre-mezclada
pH:	pH 5, 4 y 3 (gástrico) y 4.0 (fase de transición) seguida de una pequeña fase intestinal a pH 6.8
Pepsina:	3000 U/g diet.
Pancreatina:	8 mg/g diet
Temperatura:	40°C
Número de réplicas:	4
Dosis enzimática:	Arabinofuranosidasa y xilanasa 50 y 67 mg de EP/kg de dieta respectivamente

Se precipitaron muestras con un 80% de etanol, se centrifugaron y se descartó el sobrenadante. El granulado se lavó en un 80% de etanol y se centrifugó y el sobrenadante se descartó. El granulado se disolvió en tampón de acetato (pH 5) y el almidón se eliminó antes del análisis de fibra dietética.

TABLA 4

Contenido de arabinosa y xilosa en residuos de polisacáridos sin almidón totales residuales después de la digestión *in vitro*

	Residuos de polisacáridos sin almidón (% de materia seca)	
	Arabinosa	Xilosa
Control	1,78a	2,06a
Arabinofuranosidasa de <i>Humicola insolens</i>	1,71a	1,93a
Xilanasa de <i>Thermomyces lanuginosus</i>	1,68a	1,81b
Xilanasa + arabinofuranosidasa	1,60b	1,78b

Los valores medios de una columna no compartiendo un índice de letra común se diferencian con significación estadística (P<0.05).

### Referencias citadas en la descripción

Esta lista de referencias citada por el solicitante ha sido recopilada exclusivamente para la información del lector. No forma parte del documento de patente europea. La misma ha sido confeccionada con la mayor diligencia; la OEP sin embargo no asume responsabilidad alguna por eventuales errores u omisiones.

**Documentos de patente citados en la descripción**

- WO 9117243 A [0044]
- EP 238023 A [0100]
- 5 • WO 9600787 A [0060] [0100]
- WO 9114772 A [0115]
- WO 9533836 A [0075]

**Bibliografía fuera de la patente citada en la descripción**

- 10 • Van **Laere** *Appl. Microbiol. Biotechnol.*, 1997, vol. 47, 231-235 [0002]
- Van den **Broek** *Applied Microbiology and Biotechnology*, 2005, [0002]
- 15 • **Ferre** *Eur. J. Biochem.*, 2000, vol. 267, 6633-6641 [0002]
- **Nogawa** *Appl. Environ. Microbiol.*, 1999, vol. 65, 3964-3968 [0002]
- 20 • **Needleman, S.B. Wunsch, C.D.** *Journal of Molecular Biology*, 1970, vol. 48, 443-453 [0016]
- **H. Neurath R.L. Hill** *The Proteins Academic Press* 1979. [0022] [0036]
- **Bolton McCarthy** *Proceedings of the National Academy of Sciences USA*, 1962, vol. 48, 1390- [0032]
- 25 • **Innis et al.** *PCR: A Guide to Methods and Application Academic Press* 1990. [0050]
- **Ford et al.** *Protein. Expression and Purification*, 1991, vol. 2, 95-107 [0053]
- **Cunningham Wells** *Science*, 1989, vol. 244, 1081-1085 [0054]
- 30 • **de Vos et al.** *Science*, 1992, vol. 255, 306-312 [0054]
- **Smith et al.** *Journal of Molecular Biology*, 1992, vol. 224, 899-904 [0054]
- 35 • **Wlodaver et al.** *FEBS Letters*, 1992, vol. 309, 59-64 [0054]
- **Villa-Kamaroff et al.** *Proceedings of the National Academy of Sciences USA*, 1978, vol. 75, 3727-3731 [0059]
- 40 • **DeBoer et al.** *Proceedings of the National Academy of Sciences USA*, 1983, vol. 80, 21-25 [0059]
- Useful proteins from recombinant bacteria *Scientific American*, 1980, vol. 242, 74-94 [0059]
- **Romanos et al.** *Yeast*, 1992, vol. 8, 423-488 [0061]
- 45 • **Guo Sherman** *Molecular Cellular Biology*, 1995, vol. 15, 5983-5990 [0070]
- **Simonen Palva** *Microbiological Reviews*, 1993, vol. 57, 109-137 [0072]
- **Ehrlich** *Proceedings of the National Academy of Sciences USA*, 1978, vol. 75, 1433- [0085]
- 50 • **Chang Cohen** *Molecular General Genetics*, 1979, vol. 168, 111-115 [0092]
- **Young Spizizin** *Journal of Bacteriology*, 1961, vol. 81, 823-829 [0092]
- 55 • **Dubnau Davidoff-Abelson** *Journal of Molecular Biology*, 1971, vol. 56, 209-221 [0092]
- **Shigekawa Dower** *Biotechniques*, 1988, vol. 6, 742-751 [0092]
- **Koehler Thorne** *Journal of Bacteriology*, 1987, vol. 169, 5771-5278 [0092]
- 60 • **Hawksworth et al.** *Ainsworth and Bisby's Dictionary of The Fungi University Press* 1995. [0093]
- *Biology and Activities of Yeast Soc. App. Bacteriol. Symposium Series No. 91980.* [0094]
- 65 • **Yelton et al.** *Proceedings of the National Academy of Sciences USA*, 1984, vol. 81, 1470-1474 [0100]
- **Malardier et al.** *Gene*, 1989, vol. 78, 147-156 [0100]

## ES 2 356 614 T3

• Guide to Yeast Genetics and Molecular Biology Becker **Guarente** Methods in Enzymology *Academic Press*, Inc.vol. 194, 182-187 [0100]

5 • **Ito et al.** *Journal of Bacteriology*, 1983, vol. 153, 163- [0100]

• **Hinnen et al.** *Proceedings of the National Academy of Sciences USA*, 1978, vol. 75, 1920- [0100]

• Protein Purification VCH Publishers 1989. [0106]

10 • **Franck et al.** *Cell*, 1980, vol. 21, 285-294 [0115]

• **Christensen AH Sharrock RA Quail** Maize polyubiquitin genes: structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts by electroporation *Plant Mo. Biol.*, 1992, vol. 18, 675-689 [0115]

15 • **Zhang W McElroy D. Wu R** Analysis of rice Act1 5' region activity in transgenic rice plants *Plant Cell*, 1991, vol. 3, 1155-1165 [0115]

20 • Edwards **Coruzzi** *Annu. Rev. Genet.*, 1990, vol. 24, 275-303 [0115]

• **Ito et al.** *Plant Mol. Biol.*, 1994, vol. 24, 863-878 [0115]

• **Wu et al.** *Plant and Cell Physiology*, 1998, vol. 39, no. 8. 885-889 [0115]

25 • **Conrad U. et al.** *Journal of Plant Physiology*, 1998, vol. 152, no. 6. 708-711 [0115]

• **Chen et al.** *Plant and cell physiology*, 1998, vol. 39, no. 9. 935-941 [0115]

• **Kyozuka et al.** *Plant Physiology*, 1993, vol. 102, no. 3. 991-1000 [0115]

30 • **Mitra, A. Higgins, DW** *Plant Molecular Biology*, 1994, vol. 26, no. 1. 85-93 [0115]

• **Kagaya et al.** *Molecular and General Genetics*, 1995, vol. 248, no. 6. 668-674 [0115]

35 • **Xu et al.** *Plant Molecular Biology*, 1993, vol. 22, no. 4. 573-588 [0115]

• **Gasser et al.** *Science*, vol. 244, 1293- [0118]

40 • **Potrykus** *Bio/Techn.*, 1990, vol. 8, 535- [0118]

• **Shimamoto et al.** *Nature*, 1989, vol. 338, 274- [0118]

• Hooykas **Schilperoort** *Plant Mol. Biol.*, 1992, vol. 19, 15-38 [0119]

45 • **Christou** *Plant J.*, 1992, vol. 2, 275-281 [0119]

• **Shimamoto** *Curr. Opin. Biotechnol.*, 1994, vol. 5, 158-162 [0119]

• **Vasil et al.** *Bio/Technology*, 1992, vol. 10, 667-674 [0119]

50 • **Omirulleh S et al.** *Plant Molecular biology*, 1993, vol. 21, no. 3. 415-428 [0119]

• **Sambrook et al.** *Molecular Cloning, A Laboratory Manual Cold Spring Harbor* 1989. [0125]

55 • **Christgau et al.** *Curr. Genet.*, 1995, vol. 27, 135-141 [0125]

• **Ausubel et al.** *Curr. Prot. Mol.Biol. John Wiley & Sons* 2003. [0125]

60 • **Poutanen et al.** *Appl. Microbiol. Biotechnol.*, 1988, vol. 28, 425-432 [0130]

65

# ES 2 356 614 T3

## REIVINDICACIONES

1. Arabinofuranosidasa la cual es:

- 5
- a) un polipéptido teniendo una secuencia de aminoácidos como el péptido maduro mostrado en la SEC ID nº: 2 o
  - b) un análogo del polipéptido definido en (a) el cual:
    - 10 i) tiene al menos un 80% de identidad con dicho polipéptido, o
    - ii) es una variante alélica de dicho polipéptido.

15 2. Arabinofuranosidasa según la reivindicación 1 la cual es nativa a una cepa de *Humicola*, preferiblemente *H. insolens*.

3. Secuencia de ácidos nucleicos la cual tiene al menos un 80% de identidad con las secuencias de ADN mostradas en la SEC ID NO:1, o

- 20
- a) es una variante alélica de la SEC ID NO:1, o
  - b) una cadena complementaria a a).

25 4. Constructo de ácidos nucleicos comprendiendo la secuencia de ácidos nucleicos según la reivindicación 3 operativamente enlazada a una o más secuencias de control capaces de dirigir la expresión de la arabinofuranosidasa en un huésped de expresión adecuado.

5. Vector de expresión recombinante comprendiendo el constructo de ácidos nucleicos según la reivindicación 4.

30 6. Célula huésped recombinante comprendiendo el vector de expresión según la reivindicación 5.

7. Método para la producción de una arabinofuranosidasa comprendiendo cultivar la célula huésped según la reivindicación 6 bajo condiciones propicias para la producción de la arabinofuranosidasa, y recuperar la arabinofuranosidasa.

35 8. Composición comprendiendo la arabinofuranosidasa según cualquiera de las reivindicaciones 1 ó 2.

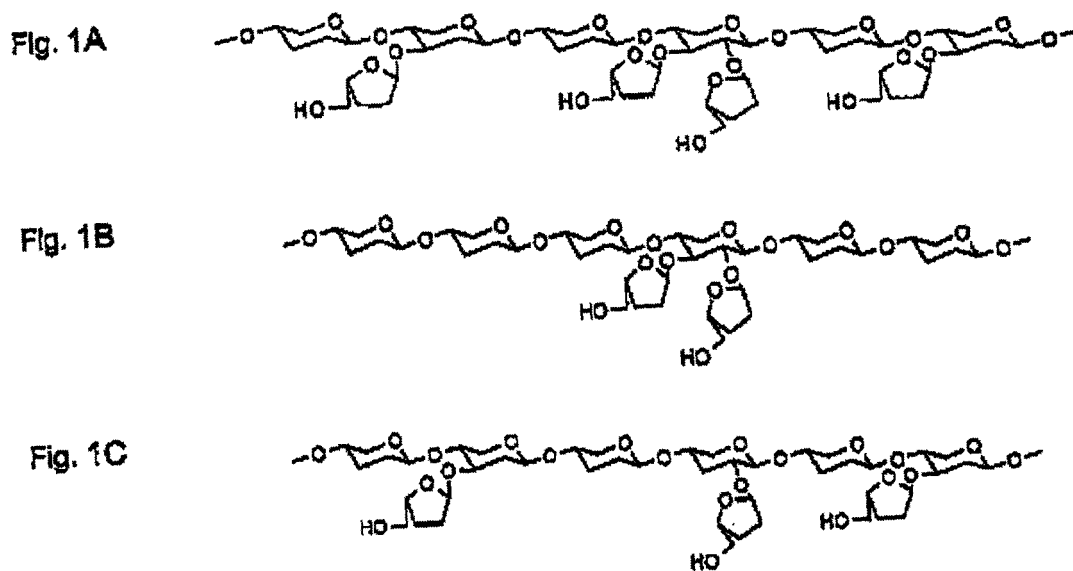
9. Uso de la arabinofuranosidasa según cualquiera de las reivindicaciones 1 ó 2 en una masa.

40 10. Uso de la arabinofuranosidasa según cualquiera de las reivindicaciones 1 ó 2 en la producción de etanol combustible a partir de celulosa conteniendo biomasa o en la producción de etanol combustible y/o potable a partir de almidón

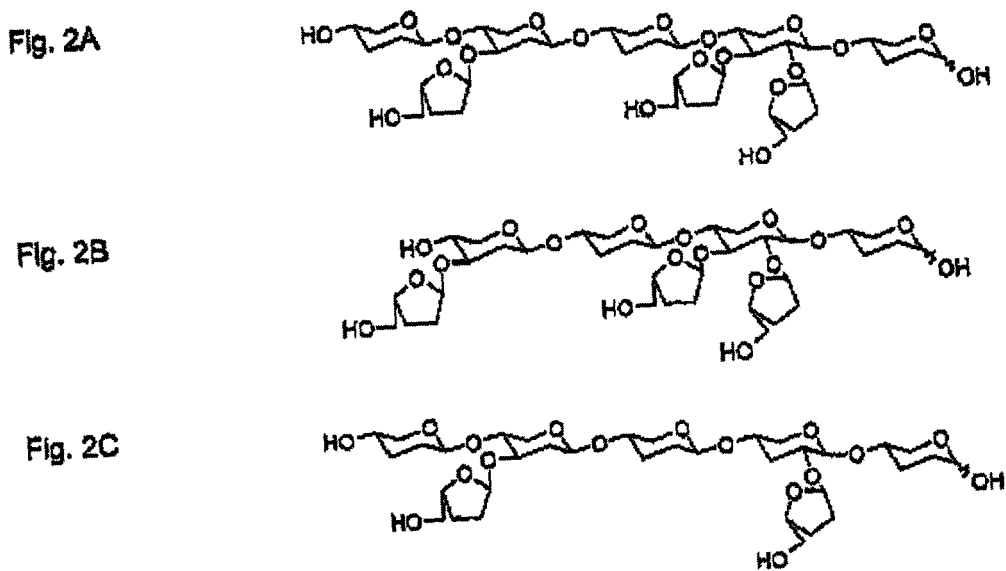
45 11. Uso de la arabinofuranosidasa según cualquiera de las reivindicaciones 1 ó 2 en un proceso de maceración.

12. Uso de la arabinofuranosidasa según cualquiera de las reivindicaciones 1 ó 2 en un proceso para producir un producto alimenticio.

**Fig. 1**



**Fig. 2**



# ES 2 356 614 T3

## LISTA DE SECUENCIAS

	<110> Novozymes A/S	
5	<120> Arabinofuranosidasas	
	<130> 10814.204-WO	
10	<160> 4	
	<170> PatentIn version 3.3	
15	<210> 1	
	<211> 1677	
	<212> ADN	
	<213> <i>Humicola insolens</i>	
20	<220>	
	<221> CDS	
25	<222> (1)..(1677)	
	<220>	
	<221> sig_péptido	
30	<222> (1)..(54)	
	<400> 1	
35	atg cta ggc ttg aag gtc ttg tgt ctc tcc gcc gtc gtg ggg acg gcg 48 Met Leu Gly Leu Lys Val Leu Cys Leu Ser Ala Val Val Gly Thr Ala	
	1 5 10 15	
	gtc tct gtg ccg cac gcg ggc aat ctt ccc cgt cag gcc agc act ttc 96 Val Ser Val Pro His Ala Gly Asn Leu Pro Arg Gln Ala Ser Thr Phe	
	20 25 30	
40	acc aac ccc gtg ctt tgg gaa gat cac cca gat ctc gaa gtg ttc cgc 144 Thr Asn Pro Val Leu Trp Glu Asp His Pro Asp Leu Glu Val Phe Arg	
	35 40 45	
45	gtc ggc tca gta ttc tac tac tcc tcg tcc acc ttc gcc tac tcc ccc 192 Val Gly Ser Val Phe Tyr Tyr Ser Ser Thr Phe Ala Tyr Ser Pro	
	50 55 60	
	ggc gcg ccc gtc ctc aag tcc tac gac ctc gtc cac tgg acg ccc gtc 240 Gly Ala Pro Val Leu Lys Ser Tyr Asp Leu Val His Trp Thr Pro Val	
	65 70 75 80	
50	acc cat tcc gtg ccg cgt ctc aac ttc ggc tcc aac tac gac ctc ccc 288 Thr His Ser Val Pro Arg Leu Asn Phe Gly Ser Asn Tyr Asp Leu Pro	
	85 90 95	
	agc ggc acc ccg ggc gcc tac gtc aag ggc atc tgg gcc tca acc ctc 336 Ser Gly Thr Pro Gly Ala Tyr Val Lys Gly Ile Trp Ala Ser Thr Leu	
55	100 105 110	
	cgc tac cgc cgc tcc aat gac cgc ttc tac tgg tac ggc tgc gtc gaa 384 Arg Tyr Arg Arg Ser Asn Asp Arg Phe Tyr Trp Tyr Gly Cys Val Glu	
	115 120 125	
60	ggg aga acc tac ctc tgg acc agc ccg ggc ggt aac gcg ctc gcc aac 432 Gly Arg Thr Tyr Leu Trp Thr Ser Pro Gly Gly Asn Ala Leu Ala Asn	
	130 135 140	
	aac ggc gag gtg ccc ccc tcc gca tgg aac tgg cag cac acc gcc acc 480 Asn Gly Glu Val Pro Pro Ser Ala Trp Asn Trp Gln His Thr Ala Thr	
65	145 150 155 160	
	atc gac aac tgc tac tac gac gcc ggc ctg ctc atc gac gac gac gac 528	

ES 2 356 614 T3

	Ile	Asp	Asn	Cys	Tyr 165	Tyr	Asp	Ala	Gly	Leu 170	Leu	Ile	Asp	Asp	Asp	Asp	
5	acc	atg	tac	atc	gcg	tac	ggc	aac	ccg	acc	atc	aac	gtc	gcg	cag	ctc	576
	Thr	Met	Tyr	Ile 180	Ala	Tyr	Gly	Asn	Pro 185	Thr	Ile	Asn	Val	Ala 190	Gln	Leu	
10	tcc	ccc	gac	ggc	acc	cgc	cag	gtg	cgc	gtg	cag	cag	cgc	gtc	tac	gcg	624
	Ser	Pro	Asp 195	Gly	Thr	Arg	Gln	Val 200	Arg	Val	Gln	Gln	Arg 205	Val	Tyr	Ala	
15	cac	ccg	cag	ggc	cag	acg	gtc	gag	ggc	gcg	cgc	atg	tac	aag	atc	cgc	672
	His	Pro 210	Gln	Gly	Gln	Thr	Val 215	Glu	Gly	Ala	Arg	Met 220	Tyr	Lys	Ile	Arg	
20	ggc	aac	tac	tac	atc	ctg	gtg	acc	cgc	ccc	gcc	gac	gca	gag	tac	gtg	720
	Gly	Asn	Tyr	Tyr	Ile 225	Leu	Val 230	Thr	Arg	Pro 235	Ala	Asp	Ala	Glu	Tyr	Val 240	
25	ctg	cgg	tcg	acg	acg	ggg	tcg	ccg	ttc	ggc	ccg	tac	gag	gcg	cgc	acg	768
	Leu	Arg	Ser	Thr 245	Thr	Gly	Ser	Pro	Phe	Gly 250	Pro	Tyr	Glu	Ala	Arg 255	Thr	
30	ctg	gtg	tcg	ccg	atc	cag	ggc	ccg	ctg	gcc	aac	gcc	ggg	ttc	gcg	cac	816
	Leu	Val	Ser	Arg 260	Ile	Gln	Gly	Pro	Leu 265	Ala	Asn	Ala	Gly	Phe 270	Ala	His	
35	cag	ggc	ggc	atc	gtc	gac	gcg	ccg	gat	ggg	acg	tgg	cac	tac	gtc	gcg	864
	Gln	Gly	Gly	Ile 275	Val	Asp	Ala	Pro 280	Asp	Gly	Thr	Trp	His 285	Tyr	Val	Ala	
40	ttc	atg	gat	gcg	tat	ccc	ggc	gga	cgc	atc	ccc	gtg	gtg	gcg	ccg	ctg	912
	Phe	Met 290	Asp	Ala	Tyr	Pro	Gly 295	Gly	Arg	Ile	Pro 300	Val	Val	Ala	Pro	Leu	
45	ccg	tgg	acg	gcg	gat	ggg	tgg	ccc	gag	gtg	gtc	acg	gat	tcg	caa	ggg	960
	Arg	Trp 305	Thr	Ala	Asp 310	Gly	Trp	Pro	Glu	Val 315	Val	Thr	Asp	Ser	Gln	Gly 320	
50	agg	tgg	ggg	acg	agc	tat	ccc	att	cca	gtt	cgc	gga	gca	aag	aac	gcg	1008
	Arg	Trp 325	Gly	Thr	Ser	Tyr	Pro	Ile	Pro 330	Val	Arg	Gly	Ala	Lys	Asn 335	Ala	
55	acg	gag	ggg	ctg	gcg	agc	acg	gat	ctg	gac	gag	ttc	cgc	ggg	acg	agg	1056
	Thr	Glu 340	Gly	Leu	Ala	Ser	Thr	Asp	Leu 345	Asp	Glu	Phe	Arg	Gly 350	Thr	Arg	
60	ttc	agc	gag	cat	tgg	gag	tgg	aat	cat	aac	ccg	gac	acg	agc	aag	ttt	1104
	Phe	Ser 355	Glu	His	Trp	Glu	Trp	Asn 360	His	Asn	Pro	Asp	Thr 365	Ser	Lys	Phe	
65	acg	ttg	ctg	ggc	ggt	aac	gag	ggc	ggg	ctc	atc	ctc	cgg	aca	gcg	acc	1152
	Thr	Leu 370	Leu	Gly	Gly	Asn	Glu 375	Gly	Gly	Leu	Ile	Leu 380	Arg	Thr	Ala	Thr	
70	gtg	acg	ggg	gat	ttg	ttt	gcc	gca	agg	aat	acg	ctc	acg	agg	agg	atc	1200
	Val	Thr 385	Gly	Asp	Leu	Phe 390	Ala	Ala	Arg	Asn	Thr 395	Leu	Thr	Arg	Arg	Ile 400	
75	gcg	gga	ccc	aag	gcc	agt	gga	atc	ttc	cgg	ctg	gat	gtg	cgc	ggg	atg	1248
	Ala	Gly 405	Pro	Lys	Ala	Ser	Gly	Ile	Phe	Arg 410	Leu	Asp	Val	Arg	Gly 415	Met	
80	cgc	gac	ggt	gac	cgg	gcc	ggc	gcc	gtg	ctg	ttc	cgg	gat	cgt	gcg	gcg	1296
	Arg	Asp 420	Gly	Asp	Arg	Ala	Gly	Ala	Val 425	Leu	Phe	Arg	Asp	Arg 430	Ala	Ala	
85	tac	atc	ggg	gtg	tgg	aag	cag	ggc	aac	gag	gcg	cgg	att	gtc	atg	gtg	1344

ES 2 356 614 T3

Tyr Ile Gly Val Trp Lys Gln Gly Asn Glu Ala Arg Ile Val Met Val  
 435 440 445  
 5 gac gac ctg cgg ttg aac gag gat ggt tgg agg acg gcg tcc acc ggc 1392  
 Asp Asp Leu Arg Leu Asn Glu Asp Gly Trp Arg Thr Ala Ser Thr Gly  
 450 455 460  
 aga gtg gcc gcc aac ggt ccg gtg atc gac acg aac gct cag cag gat 1440  
 Arg Val Ala Ala Asn Gly Pro Val Ile Asp Thr Asn Ala Gln Gln Asp  
 465 470 475 480  
 10 atc tgg ctg cga att gat gcg gac atc aca ccg gcg ttt ggg acg aac 1488  
 Ile Trp Leu Arg Ile Asp Ala Asp Ile Thr Pro Ala Phe Gly Thr Asn  
 485 490 495  
 15 acg gag cgc acg acg acg ttc tac tac agt att gat ggt ggg agg acg 1536  
 Thr Glu Arg Thr Thr Thr Phe Tyr Tyr Ser Ile Asp Gly Gly Arg Thr  
 500 505 510  
 20 tat acc agg ctg ggc cct gcc ttt gcg atg acc aat tct tgg aga tac 1584  
 Tyr Thr Arg Leu Gly Pro Ala Phe Ala Met Thr Asn Ser Trp Arg Tyr  
 515 520 525  
 25 ttc acc gga tac cgg ttt gga gtg ttc aac ttt tcg acc aag agt ctt 1632  
 Phe Thr Gly Tyr Arg Phe Gly Val Phe Asn Phe Ser Thr Lys Ser Leu  
 530 535 540  
 30 gga ggt gag gtg aag gtt aag ggg ttc aag atg aac atg atc tag 1677  
 Gly Gly Glu Val Lys Val Lys Gly Phe Lys Met Asn Met Ile  
 545 550 555

30 <210> 2  
 <211> 558  
 <212> PRT  
 <213> *Humicola insolens*  
 35  
 <400> 2

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



ES 2 356 614 T3

	115						120						125			
5	Gly	Arg	Thr	Tyr	Leu	Trp	Thr	Ser	Pro	Gly	Gly	Asn	Ala	Leu	Ala	Asn
		130					135					140				
10	Asn	Gly	Glu	Val	Pro	Pro	Ser	Ala	Trp	Asn	Trp	Gln	His	Thr	Ala	Thr
	145					150					155					160
15	Ile	Asp	Asn	Cys	Tyr	Tyr	Asp	Ala	Gly	Leu	Leu	Ile	Asp	Asp	Asp	Asp
				165						170					175	
20	Thr	Met	Tyr	Ile	Ala	Tyr	Gly	Asn	Pro	Thr	Ile	Asn	Val	Ala	Gln	Leu
				180					185					190		
25	Ser	Pro	Asp	Gly	Thr	Arg	Gln	Val	Arg	Val	Gln	Gln	Arg	Val	Tyr	Ala
			195					200					205			
30	His	Pro	Gln	Gly	Gln	Thr	Val	Glu	Gly	Ala	Arg	Met	Tyr	Lys	Ile	Arg
		210					215					220				
35	Gly	Asn	Tyr	Tyr	Ile	Leu	Val	Thr	Arg	Pro	Ala	Asp	Ala	Glu	Tyr	Val
	225					230					235					240
40	Leu	Arg	Ser	Thr	Thr	Gly	Ser	Pro	Phe	Gly	Pro	Tyr	Glu	Ala	Arg	Thr
					245					250					255	
45	Leu	Val	Ser	Arg	Ile	Gln	Gly	Pro	Leu	Ala	Asn	Ala	Gly	Phe	Ala	His
			260						265					270		
50	Gln	Gly	Gly	Ile	Val	Asp	Ala	Pro	Asp	Gly	Thr	Trp	His	Tyr	Val	Ala
			275					280					285			
55	Phe	Met	Asp	Ala	Tyr	Pro	Gly	Gly	Arg	Ile	Pro	Val	Val	Ala	Pro	Leu
		290					295					300				
60	Arg	Trp	Thr	Ala	Asp	Gly	Trp	Pro	Glu	Val	Val	Thr	Asp	Ser	Gln	Gly
	305					310					315					320
65	Arg	Trp	Gly	Thr	Ser	Tyr	Pro	Ile	Pro	Val	Arg	Gly	Ala	Lys	Asn	Ala
					325					330					335	
70	Thr	Glu	Gly	Leu	Ala	Ser	Thr	Asp	Leu	Asp	Glu	Phe	Arg	Gly	Thr	Arg
				340					345					350		
75	Phe	Ser	Glu	His	Trp	Glu	Trp	Asn	His	Asn	Pro	Asp	Thr	Ser	Lys	Phe
			355					360					365			
80	Thr	Leu	Leu	Gly	Gly	Asn	Glu	Gly	Gly	Leu	Ile	Leu	Arg	Thr	Ala	Thr
		370					375					380				
85	Val	Thr	Gly	Asp	Leu	Phe	Ala	Ala	Arg	Asn	Thr	Leu	Thr	Arg	Arg	Ile

ES 2 356 614 T3

	385		390		395		400										
5	Ala	Gly	Pro	Lys	Ala	ser	Gly	Ile	Phe	Arg	Leu	Asp	Val	Arg	Gly	Met	
				405					410						415		
10	Arg	Asp	Gly	Asp	Arg	Ala	Gly	Ala	Val	Leu	Phe	Arg	Asp	Arg	Ala	Ala	
			420						425					430			
15	Tyr	Ile	Gly	Val	Trp	Lys	Gln	Gly	Asn	Glu	Ala	Arg	Ile	Val	Met	Val	
			435					440					445				
20	Asp	Asp	Leu	Arg	Leu	Asn	Glu	Asp	Gly	Trp	Arg	Thr	Ala	Ser	Thr	Gly	
		450					455					460					
25	Arg	Val	Ala	Ala	Asn	Gly	Pro	Val	Ile	Asp	Thr	Asn	Ala	Gln	Gln	Asp	
	465					470					475					480	
30	Ile	Trp	Leu	Arg	Ile	Asp	Ala	Asp	Ile	Thr	Pro	Ala	Phe	Gly	Thr	Asn	
				485						490					495		
35	Thr	Glu	Arg	Thr	Thr	Phe	Tyr	Tyr	Ser	Ile	Asp	Gly	Gly	Arg	Thr		
				500				505					510				
40	Tyr	Thr	Arg	Leu	Gly	Pro	Ala	Phe	Ala	Met	Thr	Asn	Ser	Trp	Arg	Tyr	
			515					520					525				
45	Phe	Thr	Gly	Tyr	Arg	Phe	Gly	Val	Phe	Asn	Phe	Ser	Thr	Lys	Ser	Leu	
		530					535					540					
50	Gly	Gly	Glu	Val	Lys	Val	Lys	Gly	Phe	Lys	Met	Asn	Met	Ile			
	545					550					555						
55	<210> 3																
60	<211> 2200																
	<212> ADN																
	<213> <i>Miripilus giganteus</i>																
65	<220>																
	<221> CDS																
	<222> (46)..(1974)																
70	<220>																
	<221> sig_péptido																
	<222> (46)..(94)																
75	<400> 3																
80	actagtaacg	gccgccagtg	tgctggaaag	ggctcgctcg	acacg	atg	aag	ctg	ctt							57	
									Met	Lys	Leu	Leu					
85									1								
90	ttc	ttg	ctc	ggg	gcc	ttc	gtt	gcg	caa	tgt	ctc	gcg	gtc	aca	gtg	acc	105
	Phe	Leu	Leu	Gly	Ala	Phe	Val	Ala	Gln	Cys	Leu	Ala	Val	Thr	Val	Thr	
	5					10					15				20		
95	gtt	aac	aag	aac	cct	agt	cac	acg	gta	ccg	tcc	acg	ctc	tat	ggc	ctg	153
	Val	Asn	Lys	Asn	Pro	Ser	His	Thr	Val	Pro	Ser	Thr	Leu	Tyr	Gly	Leu	

ES 2 356 614 T3

	25					30					35						
5	atg	ttt	gag	gac	atc	aac	cat	agc	ggt	gat	ggc	ggc	ctg	tac	gcg	gag	201
	Met	Phe	Glu	Asp	Ile	Asn	His	Ser	Gly	Asp	Gly	Gly	Leu	Tyr	Ala	Glu	
				40					45					50			
10	ttg	ttg	caa	aac	agg	gct	ttc	caa	cag	ggt	acc	ccg	aac	acg	gcc	gct	249
	Leu	Leu	Gln	Asn	Arg	Ala	Phe	Gln	Gln	Val	Thr	Pro	Asn	Thr	Ala	Ala	
			55					60					65				
15	gca	ctc	gct	gca	tgg	cat	ccc	atc	agc	aat	ggc	aag	ctg	gcc	gta	ata	297
	Ala	Leu	Ala	Ala	Trp	His	Pro	Ile	Ser	Asn	Ala	Lys	Leu	Ala	Val	Ile	
		70					75					80					
20	caa	gac	cca	tct	cct	gtc	tcc	aat	gca	ttg	ccg	aat	tcc	ctt	caa	ttc	345
	Gln	Asp	Pro	Ser	Pro	Val	Ser	Asn	Ala	Leu	Pro	Asn	Ser	Leu	Gln	Phe	
	85					90					95					100	
25	tcc	gtg	ccc	agt	gga	tcg	agc	ggc	agg	gtc	ggc	ttt	acc	aac	gag	ggt	393
	Ser	Val	Pro	Ser	Gly	Ser	Ser	Gly	Arg	Val	Gly	Phe	Thr	Asn	Glu	Gly	
					105					110					115		
30	ttc	tgg	gga	atc	aaa	gtc	gat	tcc	act	tgg	acg	tac	aaa	gcc	tcg	ctc	441
	Phe	Trp	Gly	Ile	Lys	Val	Asp	Ser	Thr	Trp	Thr	Tyr	Lys	Ala	Ser	Leu	
				120					125					130			
35	ttc	ttc	cgc	ttc	ccc	aca	tcg	tcg	tcc	ttc	tcg	gga	gcg	ctc	acc	gtt	489
	Phe	Phe	Arg	Phe	Pro	Thr	Ser	Ser	Ser	Phe	Ser	Gly	Ala	Leu	Thr	Val	
			135					140					145				
40	ggg	ctg	cag	acg	aac	gcc	ggg	aga	gtg	ctg	gca	cag	aac	tcc	acg	cag	537
	Gly	Leu	Gln	Thr	Asn	Ala	Gly	Arg	Val	Leu	Ala	Gln	Asn	Ser	Thr	Gln	
		150				155						160					
45	atc	cgc	ggg	acg	acc	acg	aag	tgg	acg	cag	atc	aac	ctg	gag	ctc	cac	585
	Ile	Arg	Gly	Thr	Thr	Thr	Lys	Trp	Thr	Gln	Ile	Asn	Leu	Glu	Leu	His	
	165					170					175					180	
50	cct	acc	gcc	tct	gcc	ccc	gac	gtc	agc	aac	agc	ttc	ttt	gtc	acg	att	633
	Pro	Thr	Ala	Ser	Ala	Pro	Asp	Val	Ser	Asn	Ser	Phe	Phe	Val	Thr	Ile	
					185					190					195		
55	gac	ggg	gcc	gct	ggc	gcc	ggt	cag	acg	atc	aac	ttc	gcc	atg	ttc	tcg	681
	Asp	Gly	Ala	Ala	Gly	Ala	Gly	Gln	Thr	Ile	Asn	Phe	Ala	Met	Phe	Ser	
				200					205					210			
60	ctc	ttc	cct	ccc	acg	ttc	aag	aac	agg	ccg	aat	ggg	ctg	cgc	gct	gac	729
	Leu	Phe	Pro	Pro	Thr	Phe	Lys	Asn	Arg	Pro	Asn	Gly	Leu	Arg	Ala	Asp	
			215				220						225				
65	atc	gcc	gag	act	ctg	gcc	gag	atg	ggc	ccg	tcc	ttt	ttc	cgc	ttc	cca	777
	Ile	Ala	Glu	Thr	Leu	Ala	Glu	Met	Gly	Pro	Ser	Phe	Phe	Arg	Phe	Pro	
			230				235					240					
70	ggt	ggg	aac	aac	ctg	gag	ggc	caa	acg	acg	ggc	acg	agg	tgg	cag	tgg	825
	Gly	Gly	Asn	Asn	Leu	Glu	Gly	Gln	Thr	Thr	Ala	Thr	Arg	Trp	Gln	Trp	
	245				250						255					260	
75	aat	gct	acc	gtc	ggc	tcg	ctg	ctg	gac	cgc	ccc	ggc	cgg	gtg	ggc	gac	873
	Asn	Ala	Thr	Val	Gly	Ser	Leu	Leu	Asp	Arg	Pro	Gly	Arg	Val	Gly	Asp	
				265						270					275		
80	tgg	gga	tac	gtg	aac	acg	gac	ggt	cta	ggt	ctt	ctg	gag	tat	ctc	cag	921
	Trp	Gly	Tyr	Val	Asn	Thr	Asp	Gly	Leu	Gly	Leu	Leu	Glu	Tyr	Leu	Gln	
				280				285						290			
85	ttc	ttc	gaa	gat	acg	ggc	atg	gag	ccg	atc	atg	ggc	gtc	tgg	gca	ggc	969
	Phe	Phe	Glu	Asp	Thr	Gly	Met	Glu	Pro	Ile	Met	Ala	Val	Trp	Ala	Gly	

ES 2 356 614 T3

	295					300					305						
5	tat Tyr	tct Ser 310	ctc Leu	ggc Gly	ggc Gly	aca Thr	agc Ser 315	ctt Leu	gct Ala	gag Glu	aac Asn	cag Gln 320	ctc Leu	gca Ala	ccg Pro	tac Tyr	1017
10	ata Ile 325	cag Gln	caa Gln	gcg Ala	ata Ile	gat Asp 330	cag Gln	att Ile	aac Asn	ttt Phe	gtc Val 335	atc Ile	ggg Gly	gac Asp	cct Pro	gca Ala 340	1065
15	aag Lys	agt Ser	gca Ala	cct Pro	gcg Ala 345	gcg Ala	ctc Leu	cgt Arg	gct Ala	tcc Ser 350	ctg Leu	ggc Gly	cac His	cca Pro	gag Glu 355	ccc Pro	1113
20	ttc Phe	acg Thr	ctc Leu	cgc Arg 360	ttc Phe	gtg Val	gaa Glu	gtg Val	gga Gly 365	aac Asn	gag Glu	gac Asp	ttc Phe	ttc Phe 370	gcg Ala	gcg Ala	1161
25	ggc Gly	tcg Ser	tac Tyr 375	cca Pro	tac Tyr	cgc Arg	tgg Trp	cac His 380	gac Asp	ttc Phe	gtt Val	acc Thr	gca Ala 385	ctt Leu	cag Gln	gcg Ala	1209
30	caa Gln	ttc Phe 390	ccc Pro	cag Gln	atc Ile	aga Arg	ttc Phe 395	atc Ile	gcg Ala	acc Thr	acc Thr	aac Asn 400	gcc Ala	tgg Trp	aac Asn	ccg Pro	1257
35	gtt Val 405	ctg Leu	tcc Ser	ccc Pro	gtc Val	ccg Pro 410	cag Gln	tcg Ser	tat Tyr	gat Asp	gta Val 415	cac His	gtc Val	tat Tyr	cag Gln	aca Thr 420	1305
40	ccg Pro	acc Thr	tgg Trp	ttc Phe	tac Tyr 425	caa Gln	aat Asn	gct Ala	ttc Phe	tac Tyr 430	tac Tyr	gac Asp	ggc Gly	ttc Phe	cag Gln 435	cgc Arg	1353
45	aac Asn	ggc Gly	acc Thr	aca Thr 440	tac Tyr	ttt Phe	gag Glu	ggg Gly	gag Glu 445	tac Tyr	gcc Ala	gcc Ala	atc Ile	tcc Ser 450	acc Thr	aac Asn	1401
50	ggc Ala	aac Asn	gat Asp 455	ttg Leu	ttc Phe	ggt Gly	act Thr	gtt Val 460	gcc Ala	gac Asp	ggt Gly	cgc Arg	ttg Leu 465	gcg Ala	ttc Phe	cct Pro	1449
55	aca Thr 470	gtg Val	caa Gln	agt Ser	gct Ala	acc Thr	ggg Gly 475	gag Glu	gcc Ala	gca Ala	ttc Phe	atg Met 480	acc Thr	ggc Gly	ttg Leu	gag Glu	1497
60	cgc Arg 485	aac Asn	agc Ser	gac Asp	atc Ile	gtc Val 490	ttc Phe	gcc Ala	gcg Ala	tcc Ser	tac Tyr 495	gca Ala	cct Pro	ctg Leu	ctg Leu	cag Gln 500	1545
65	cac His	gtc Val	aac Asn	tcc Ser	act Thr 505	caa Gln	tgg Trp	acc Thr	ccc Pro	gac Asp 510	ctg Leu	gtt Val	tcc Ser	tac Tyr	gac Asp 515	gcc Ala	1593
70	ggc Gly	tcc Ser	ggt Val	att Ile 520	aag Lys	tcg Ser	acg Thr	agc Ser	ttc Phe 525	ttc Phe	gcc Ala	cag Gln	aag Lys	ctg Leu 530	ttc Phe	gcc Ala	1641
75	ttg Leu	aac Asn	aag Lys 535	ggc Gly	gac Asp	caa Gln	tac Tyr	ctc Leu 540	ccg Pro	agc Ser	acg Thr	ctc Leu	ccg Pro 545	acc Thr	aac Asn	ggt Gly	1689
80	ggc Gly	acg Thr 550	ctg Leu	cac His	tgg Trp	agc Ser	atc Ile 555	act Thr	cgg Arg	gcc Ala	tct Ser	agc Ser 560	tcc Ser	ggc Gly	aag Lys	acg Thr	1737
85	ttc Phe	atc Ile	aag Lys	atc Ile	gcg Ala	aac Asn	gcc Ala	ggc Gly	agc Ser	tca Ser	gcg Ala	cag Gln	agc Ser	ctc Leu	acc Thr	ttc Phe	1785



ES 2 356 614 T3

Ala Leu Thr Val Gly Leu Gln Thr Asn Ala Gly Arg Val Leu Ala Gln  
 145 150 155 160  
 5 Asn Ser Thr Gln Ile Arg Gly Thr Thr Thr Lys Trp Thr Gln Ile Asn  
 165 170 175  
 10 Leu Glu Leu His Pro Thr Ala Ser Ala Pro Asp Val Ser Asn Ser Phe  
 180 185 190  
 15 Phe Val Thr Ile Asp Gly Ala Ala Gly Ala Gly Gln Thr Ile Asn Phe  
 195 200 205  
 20 Ala Met Phe Ser Leu Phe Pro Pro Thr Phe Lys Asn Arg Pro Asn Gly  
 210 215 220  
 25 Leu Arg Ala Asp Ile Ala Glu Thr Leu Ala Glu Met Gly Pro Ser Phe  
 225 230 235 240  
 30 Phe Arg Phe Pro Gly Gly Asn Asn Leu Glu Gly Gln Thr Thr Ala Thr  
 245 250 255  
 35 Arg Trp Gln Trp Asn Ala Thr Val Gly Ser Leu Leu Asp Arg Pro Gly  
 260 265 270  
 40 Arg Val Gly Asp Trp Gly Tyr Val Asn Thr Asp Gly Leu Gly Leu Leu  
 275 280 285  
 45 Glu Tyr Leu Gln Phe Phe Glu Asp Thr Gly Met Glu Pro Ile Met Ala  
 290 295 300  
 50 Val Trp Ala Gly Tyr Ser Leu Gly Gly Thr Ser Leu Ala Glu Asn Gln  
 305 310 315 320  
 55 Leu Ala Pro Tyr Ile Gln Gln Ala Ile Asp Gln Ile Asn Phe Val Ile  
 325 330 335  
 60 Gly Asp Pro Ala Lys Ser Ala Pro Ala Ala Leu Arg Ala Ser Leu Gly  
 340 345 350  
 65 His Pro Glu Pro Phe Thr Leu Arg Phe Val Glu Val Gly Asn Glu Asp  
 355 360 365  
 70 Phe Phe Ala Ala Gly Ser Tyr Pro Tyr Arg Trp His Asp Phe Val Thr  
 370 375 380  
 75 Ala Leu Gln Ala Gln Phe Pro Gln Ile Arg Phe Ile Ala Thr Thr Asn  
 385 390 395 400  
 80 Ala Trp Asn Pro Val Leu Ser Pro Val Pro Gln Ser Tyr Asp Val His  
 405 410 415

ES 2 356 614 T3

Val Tyr Gln Thr Pro Thr Trp Phe Tyr Gln Asn Ala Phe Tyr Tyr Asp  
 420 425 430  
 5  
 Gly Phe Gln Arg Asn Gly Thr Thr Tyr Phe Glu Gly Glu Tyr Ala Ala  
 435 440 445  
 10  
 Ile Ser Thr Asn Ala Asn Asp Leu Phe Gly Thr Val Ala Asp Gly Arg  
 450 455 460  
 15  
 Leu Ala Phe Pro Thr Val Gln Ser Ala Thr Gly Glu Ala Ala Phe Met  
 465 470 475 480  
 20  
 Thr Gly Leu Glu Arg Asn Ser Asp Ile Val Phe Ala Ala Ser Tyr Ala  
 485 490 495  
 25  
 Pro Leu Leu Gln His Val Asn Ser Thr Gln Trp Thr Pro Asp Leu Val  
 500 505 510  
 30  
 Ser Tyr Asp Ala Gly Ser Val Ile Lys Ser Thr Ser Phe Phe Ala Gln  
 515 520 525  
 35  
 Lys Leu Phe Ala Leu Asn Lys Gly Asp Gln Tyr Leu Pro Ser Thr Leu  
 530 535 540  
 40  
 Pro Thr Asn Gly Gly Thr Leu His Trp Ser Ile Thr Arg Ala Ser Ser  
 545 550 555 560  
 45  
 Ser Gly Lys Thr Phe Ile Lys Ile Ala Asn Ala Gly Ser Ser Ala Gln  
 565 570 575  
 50  
 Ser Leu Thr Phe Gln Leu Thr Gln Phe Asn Ser Val Ser Ser Thr Gly  
 580 585 590  
 55  
 Thr Leu Gln Val Leu Thr Gly Pro Glu Thr Ala Ser Asn Thr Pro Glu  
 595 600 605  
 60  
 Ala Pro Gln Ala Ile Val Pro Lys Thr Ser Thr Ile Gly Thr Gly Lys  
 610 615 620  
 65  
 Thr Phe Thr Tyr Asn Ala Pro Ala Phe Ser Val Ser Val Ile Thr Val  
 625 630 635 640  
 Thr Thr Asn