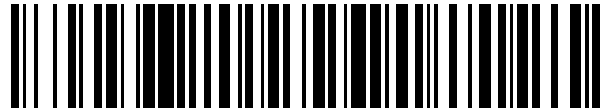


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 545 161**

21 Número de solicitud: 201430313

51 Int. Cl.:

C12N 9/42 (2006.01)

C12N 1/21 (2006.01)

C12N 15/74 (2006.01)

12

SOLICITUD DE PATENTE

A1

22 Fecha de presentación:

07.03.2014

43 Fecha de publicación de la solicitud:

08.09.2015

71 Solicitantes:

ABENGOA RESEARCH, S.L. (100.0%)
C/ Energía Solar, 1 Campus Palmas Altas
41014 Sevilla ES

72 Inventor/es:

SAADEDIN, Anas;
GÓMEZ DEL PULGAR VILLANUEVA, Eva María;
SANTERO SANTURINO, Eduardo y
TERRÓN GONZÁLEZ, Laura

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

54 Título: **Polipéptidos con actividad celulasa**

57 Resumen:

Polipéptidos con actividad celulasa.

La presente invención se refiere principalmente a un polipéptido con actividad celulasa para su uso en procesos de degradación de biomasa. La invención también se refiere a fragmentos funcionales del polipéptido, así como a los polinucleótidos que codifican dichos polipéptidos y/o fragmentos.

ES 2 545 161 A1

DESCRIPCION

Polipéptidos con actividad celulasa

5 CAMPO DE LA INVENCION

La presente invención pertenece al campo de las enzimas con actividad celulasa para la degradación de biomasa, más concretamente, se refiere a polipéptidos con actividad celulolítica o a fragmentos funcionales o variantes de los mismos, así como a los polinucleótidos que codifican dichos polipéptidos o fragmentos. La presente invención también se relaciona con una construcción génica, un vector de expresión y una célula huésped que comprenden dichos polinucleótidos, así como con un método para producir los polipéptidos, variantes o fragmentos de la invención. Por último, la invención se relaciona con el uso de los polipéptidos de la invención, fragmentos funcionales o variantes de los mismos para la degradación de biomasa, tanto para su aplicación en procedimientos para la producción de azúcares fermentables como para la producción de bioproductos por transformación química de estos azúcares, preferiblemente por fermentación microbiana de los mismos. La presente invención también incluye una composición con capacidad de degradar la biomasa que comprende al menos un polipéptido, al menos un fragmento funcional del mismo o al menos una variante del mismo de acuerdo con la presente invención y al menos algún elemento adicional que favorezca la degradación de biomasa.

ANTECEDENTES DE LA INVENCION

Las fracciones fermentables de la biomasa incluyen la celulosa y hemicelulosa. La celulosa consiste en largas cadenas insolubles de glucosa unidas covalentemente mediante enlaces beta-1,4. La hemicelulosa es una fracción heterogénea de biomasa compuesta de xilosa y azúcares menores de 5 y 6 carbonos. Aunque es un biopolímero abundante, la celulosa es altamente cristalina, insoluble en agua y muy resistente a la despolimerización. La degradación enzimática completa de celulosa a glucosa, probablemente la materia prima más deseable para procesos de fermentación, se puede conseguir mediante la acción sinérgica de 3 clases distintas de enzimas: endoglucanasas que catalizan la endohidrólisis de los enlaces (1,4)-b-D-glicosídicos de la celulosa; las exoglucanasas que llevan a cabo la hidrólisis de enlaces (1,4) en (1,4)-b-D glicanos para liberar celobiosa de los términos no reductores; y b-glucosidasas que liberan b-D-glucosa de la celobiosa (ver figura 1).

El desarrollo de un proceso económico de conversión de un producto de poco valor como la biomasa a bioetanol u otro tipo de bioproductos vía transformación química o fermentación,

requiere la optimización de varias etapas clave. En especial, una de gran importancia es la de la producción de celulasas. La utilización práctica de celulosa mediante hidrólisis con celulasas para producir glucosa requiere grandes cantidades de celulasas para despolimerizar por completo la celulosa. Además, la producción de celulasas y la hidrólisis enzimática son en proporción las etapas más costosas en este proceso.

Todas estas dificultades técnicas evidencian la necesidad de desarrollar nuevas enzimas que optimicen las tecnologías existentes para la producción de biofuel y otros bioproductos hacia sistemas más competitivos, económicos y que produzcan un menor impacto medioambiental al minimizar las emisiones de dióxido de carbono.

El hábitat del suelo es un ambiente enormemente diverso. Un gramo de suelo puede contener varios millones de bacterias pertenecientes a más de 4000 a 7000 especies diferentes. Sin embargo, solo 1% de la flora bacteriana del suelo se puede cultivar bajo condiciones estándar de laboratorio.

La metagenómica que consiste en la extracción, clonado y análisis de la totalidad del material genético de un hábitat determinado, ha emergido para explorar el potencial de la información genética de microorganismos no cultivables. Las librerías metagenómicas son una herramienta poderosa para el descubrimiento de nuevas actividades biológicas y ya han sido utilizadas con éxito para aislar nuevas enzimas hidrolíticas como amilasas, celulasas y xilanasas. Por ejemplo, Ilmberger N. ("Metagenomic cellulases highly tolerant towards the presence of ionic liquids-linking thermostability and halotolerance" Appl. Microbiol. Biotechnol. (2012)95: 135-146) describe la identificación y el aislamiento de enzimas con actividad celulasa a partir de librerías metagenómicas construidas a partir de diferentes ambientes en los que la actividad celulasa está muy presente como, por ejemplo, extractos del gusano *Teredo navalis*, muestras provenientes de una planta de biogás o heces de elefante.

El rastreo de librerías metagenómicas, que es un paso crítico en el aislamiento exitoso de actividades biológicas nuevas y mejoradas, se puede llevar a cabo mediante ensayos de actividad. Las librerías metagenómicas se construyen en vectores de expresión para permitir el aislamiento de actividades biológicas de secuencias génicas enteramente nuevas cuando estas se expresan con éxito en un determinado hospedador. Los ensayos cromogénicos de actividad son usados comúnmente para la detección de hidrolasas a partir de librerías metagenómicas. La figura 2 representa esquemáticamente el proceso de identificación y

caracterización de actividades enzimáticas mediante la generación de librerías metagenómicas.

5 Los autores de la presente invención han identificado y aislado a partir de una librería metagenómica de microorganismos del suelo, un polipéptido que muestra una alta actividad celulasa, especialmente endoglucanasa, y que posee unas características funcionales que la hacen muy apropiada para su uso en procesos industriales de degradación de biomasa y producción de bioetanol y otros bioproductos de alto valor añadido.

10 **OBJETO DE LA INVENCION**

El objeto principal de la invención es un polipéptido aislado con actividad celulolítica (en adelante el polipéptido de la invención) que comprende la secuencia SEQ ID NO1 o que posee un grado de homología o de identidad de al menos un 80% con la SEQ ID NO1 o que está codificado por un polinucleótido que hibrida bajo condiciones de baja, media, medio-
15 alta, alta o muy alta astringencia con la secuencia SEQ ID NO 6 que codifica el polipéptido maduro.

Son asimismo un objeto de la invención, los fragmentos funcionales o las variantes del polipéptido de la invención.

20 Otro objeto de la invención son las secuencias polinucleotídicas que codifican el polipéptido de la invención o los fragmentos funcionales o las variantes del mismo (en adelante polinucleótido de la invención).

25 Otro objeto de la invención viene representado por una construcción génica que comprende la secuencia polinucleotídica antes mencionada operativamente unida a una o más secuencias de control que dirigen la expresión de polipéptidos en sistemas de expresión.

Lógicamente, también es objeto de la invención un vector de expresión que comprende el
30 polinucleótido de la invención o dicha construcción génica y una célula huésped que comprende dicho vector de expresión y que es capaz de producir el polipéptido de la invención o los fragmentos funcionales o las variantes de este.

Un objeto adicional de la invención es un método para producir el polipéptido de la invención
35 que comprende:

- a) Cultivar la célula huésped antes mencionada en las condiciones necesarias para la producción del polipéptido,
- b) Recuperar el polipéptido producido.

- 5 También es objeto de la presente invención una composición celulolítica que comprende al menos un polipéptido de acuerdo con la invención o al menos un fragmento funcional o una variante del mismo y, opcionalmente, al menos un elemento que favorezca o facilite la degradación de biomasa (en adelante composición de la invención).
- 10 Otro objeto fundamental de la invención es el uso del polipéptido de la invención, de los fragmentos funcionales o las variantes del mismo o de la composición de la invención para la degradación de biomasa.

De este objeto general se derivan los siguientes más particulares.

15

El primero es un procedimiento para producir azúcares fermentables que comprende incubar biomasa con el polipéptido de la invención, con fragmentos funcionales o variantes del mismo o con la composición de la invención.

- 20 El segundo objeto que se deriva del uso antes mencionado es un procedimiento para producir un bioproducto a partir de biomasa que comprende:

- a) producir azúcares fermentables incubando biomasa con el polipéptido de la invención, con fragmentos funcionales del mismo o con la composición de la invención

25

- b) llevar a cabo una transformación química de los azúcares obtenidos en la etapa a) para obtener al menos un bioproducto,
- c) Recuperar el al menos un bioproducto obtenido en la etapa b)

BREVE DESCRIPCIÓN DE LAS FIGURAS

30

Figura 1: representa la hidrólisis de la celulosa por los tres tipos de celulasas: la endoglucanasa rompe los enlaces no covalentes en la estructura amorfa de la celulosa, la celobiohidrolasa libera celobiosa del final de la cadena y finalmente la celobiosa es hidrolizada por la beta-glucosidasa.

35

Figura 2: representa las etapas esenciales de la construcción y explotación de una librería metagenómica.

Figura 3: representación del clon 98 que contienen el gen celulasa *ce/98*. La figura muestra los diferentes dominios del polipéptido de la invención codificado por *ce/98*. En el extremo N-terminal en la posición 1-32 se encuentra la señal de translocación Twin-arginina (Tat). También se representa el dominio celulasa perteneciente a la familia 5 en la posición 41-378. Dentro del dominio celulasa hay un dominio de tipo fibronectina III. Asimismo se encuentran dos dominios de unión a carbohidratos de tipo 2 (CBM2) en las posiciones 443-570 y 664-780 y un dominio de unión a carbohidratos de tipo 6 (CBM6) en la posición 474-630.

Figura 4: representación de la actividad del polipéptido SEQ ID NO 1 a diferentes temperaturas ensayadas (30°C, 37°C, 45°C, 50°C y 56°C).

Figura 5: representación de la disminución de la densidad óptica a diferentes temperaturas ensayadas: (A) 30°C, (B) 37°C, (C) 45°C y (D) 50°C. La disminución de la densidad óptica se asocia a la degradación de la carboximetilcelulosa (CMC).

Figura 6: representación de la disminución de la densidad óptica después de la incubación de la enzima a: (A) 30°C, (B) 37°C, (C) 40°C y (D) 42°C durante diferentes periodos de tiempo. La disminución de la densidad óptica se asocia a la degradación de la CMC.

Figura 7: representación de la actividad del polipéptido SEQ ID NO 1 a diferentes pH ensayados (5, 5.5, 6, 6.5, 7, 7.5 y 8).

Figura 8: representación de la disminución de la densidad óptica a diferentes pH ensayados (5, 5.5, 6, 6.5, 7, 7.5 y 8). La disminución de la densidad óptica se asocia a la degradación de la CMC.

Figura 9: Visualización de un clon positivo para actividad celulasa después del rastreo de la librería metagenómica de bacterias del suelo usando CMC como sustrato y el método rojo congo.

Figura 10: En la imagen de la izquierda se observa resaltada la banda correspondiente al vector linearizado empleado en la clonación. En la imagen del centro se observan los dos

pocillos del gel donde se cargó el producto de PCR para obtener el inserto de interés y finalmente en la imagen de la derecha se observan tanto el vector como el inserto antes justo del paso de la ligación, a fin de poder ver la cantidad de que se disponía cada uno y poder ajustar así las proporciones de vector/inserto para la ligación.

5

DESCRIPCIÓN DETALLADA DE LA INVENCION

Con el fin de facilitar la comprensión y de aclarar el significado de determinados términos en el contexto de la presente invención se aportan las siguientes definiciones:

10

“Polipéptido”: es una secuencia de aminoácidos de una longitud de entre aproximadamente 100 y 1000 aminoácidos, preferiblemente entre aproximadamente 200 y 950 y más preferiblemente entre aproximadamente 340 y 930 aminoácidos, que puede ser producido en un sistema de expresión heterólogo o producido sintéticamente. Los aminoácidos que forman el polipéptido pueden estar modificados o no modificados y aunque los aminoácidos que forman el polipéptido son eminentemente aminoácidos proteicos (alanina, arginina, asparagina, aspartato, cisteína, fenilalanina, glicina, glutamato, glutamina, histidina, isoleucina, leucina, lisina, metionina, prolina, serina, tirosina, treonina, triptófano y valina), el polipéptido puede tener algún aminoácido proteico sustituido por alguno no proteico o sintético.

20

“Polipéptido de la invención”: es un polipéptido aislado con actividad celulolítica que comprende la secuencia SEQ ID NO1 o que posee un grado de homología o de identidad de al menos un 80% con la SEQ ID NO1 o que está codificado por un polinucleótido que hibrida bajo condiciones de baja, media, medio-alta, alta o muy alta astringencia con la secuencia SEQ ID NO 6 que codifica el polipéptido maduro.

25

“Actividad celulolítica o celulasa”: hace referencia a la capacidad que tiene el polipéptido de la invención, así como los fragmentos funcionales o las variantes del mismo de degradar por hidrólisis los enlaces (1,4)-b-D-glicosídicos de la celulosa. La actividad celulolítica o celulasa puede ser endoglucanasa, exoglucanasa, celobiohidrolasa o beta-glucosidasa dependiendo del tipo de sustrato celulósico sobre el que actúe el polipéptido. Aunque el polipéptido de la invención puede mostrar una actividad celulolítica o celulasa variada, su actividad es preferentemente endoglucanasa.

35

“**Grado de homología**”: se refiere a la correspondencia de secuencia que se da entre una determinada cadena polipeptídica con otra. El término grado de homología implica que la longitud del polipéptido frente al de referencia puede ser mayor, menor o igual. Cuando en el contexto de la invención se menciona, por ejemplo, que un polipéptido tiene un grado de homología de al menos un 80% con una secuencia determinada significa que este polipéptido puede tener una correspondencia de entre un 100% hasta un 80% con dicha secuencia y que su longitud puede ser mayor, menor o igual al de dicha secuencia.

“**Grado de identidad**”: se refiere a la correspondencia de secuencia que se da entre una determinada cadena polipeptídica con otra. El término grado de identidad implica que la longitud del polipéptido siempre es igual frente al de referencia. Cuando en el contexto de la invención se menciona por ejemplo, que un polipéptido tiene un grado de identidad de al menos un 80% con una secuencia determinada significa que este polipéptido puede tener una correspondencia de secuencia de entre un 100% hasta un 80% con dicha secuencia y la longitud del polipéptido de referencia.

“**Fragmento funcional**”: se refiere a aquellos fragmentos que se derivan del polipéptido de la invención y que conservan en mayor o menor medida la actividad celulasa del polipéptido original. El fragmento funcional puede conservar entre el 50% y el 100% de la actividad del polipéptido de la invención. También se refiere a aquellos fragmentos que conservando la actividad celulasa puedan presentar un perfil de actividad celulasa diferente al del polipéptido original.

“**Variante del polipéptido de la invención**”: se refiere a aquellas modificaciones del polipéptido de SEQ ID NO 1 que impliquen sustituciones, deleciones y/o inserciones en una o más posiciones sin afectar sustancialmente a la actividad del polipéptido. Los cambios de aminoácidos pueden ser de naturaleza menor, esto es sustituciones o inserciones conservativas que no afecten significativamente el plegado y o actividad del polipéptido; pequeñas deleciones normalmente de 1 a 30 aminoácidos; pequeñas extensiones de los extremos amino y carboxiterminal como pequeños péptidos de unión de 20-25 residuos o pequeñas extensiones que faciliten la purificación mediante el cambio de la carga neta u otra función, colas de polihistidinas, epítomos antigénicos o dominios de unión. También se contemplan variantes que puedan alterar de manera controlada y dirigida propiedades físico químicas del polipéptido. Por ejemplo, estas variantes incluyen cambios de aminoácidos que puedan mejorar la estabilidad térmica del polipéptido de la invención, cambiar su especificidad por el sustrato, su temperatura óptima o el pH óptimo.

5 “**Condiciones de muy baja astringencia**”: para sondas de al menos 100 nucleótidos de longitud implica prehibridación e hibridación a 42°C en SSPE 5X, SDS al 0.3%, 200 µg/ml de DNA fragmentado y desnaturalizado de esperma de salmón y 25% de formamida, siguiendo procedimientos estándar de Southern Blot de 12 a 24 horas. El material soporte se limpia finalmente 3 veces cada una 15 minutos usando SSC 2X y SDS al 0.2% a 45° C.

10 “**Condiciones de baja astringencia**”: para sondas de al menos 100 nucleótidos de longitud implica prehibridación e hibridación a 42°C en SSPE 5X, SDS al 0.3%, 200 µg/ml de DNA fragmentado y desnaturalizado de esperma de salmón y 25% de formamida, siguiendo procedimientos estándar de Southern Blot de 12 a 24 horas. El material soporte se limpia finalmente 3 veces cada una 15 minutos usando SSC 2X y SDS al 0.2% a 50° C.

15 “**Condiciones de media astringencia**”: para sondas de al menos 100 nucleótidos de longitud implica prehibridación e hibridación a 42°C en SSPE 5X, SDS al 0.3%, 200 µg/ml de DNA fragmentado y desnaturalizado de esperma de salmón y 35% de formamida, siguiendo procedimientos estándar de Southern Blot de 12 a 24 horas. El material soporte se limpia finalmente 3 veces cada una 15 minutos usando SSC 2X y SDS al 0.2% a 55° C.

20 “**Condiciones de media-alta astringencia**”: para sondas de al menos 100 nucleótidos de longitud implica prehibridación e hibridación a 42°C en SSPE 5X, SDS al 0.3%, 200 µg/ml de DNA fragmentado y desnaturalizado de esperma de salmón y 35% de formamida, siguiendo procedimientos estándar de Southern Blot de 12 a 24 horas. El material soporte se limpia finalmente 3 veces cada una 15 minutos usando SSC 2X y SDS al 0.2% a 60° C.

25 “**Condiciones de alta astringencia**”: para sondas de al menos 100 nucleótidos de longitud implica prehibridación e hibridación a 42°C en SSPE 5X, SDS al 0.3%, 200 µg/ml de DNA fragmentado y desnaturalizado de esperma de salmón y 50% de formamida, siguiendo procedimientos estándar de Southern Blot de 12 a 24 horas. El material soporte se limpia finalmente 3 veces cada una 15 minutos usando SSC 2X y SDS al 0.2% a 65° C.

30 “**Condiciones de muy alta astringencia**”: para sondas de al menos 100 nucleótidos de longitud significa prehibridación e hibridación a 42°C en SSPE 5X, SDS al 0.3%, 200 µg/ml de DNA fragmentado y desnaturalizado de esperma de salmón y 50% de formamida, siguiendo procedimientos estándar de Southern Blot de 12 a 24 horas. El material soporte se limpia finalmente 3 veces cada una 15 minutos usando SSC 2X y SDS al 0.2% a 70° C.

35

“**Construcción génica**”: es una sucesión de información genética codificada en una secuencia polinucleotídica generalmente de ADN de cadena simple o doble cadena, dispuesta de tal forma que potencialmente permite ejecutar el programa genético para el cual fue diseñada. La construcción génica comprende tanto elementos codificantes como no codificantes como por ejemplo secuencias de control. Normalmente, la construcción génica comprende al menos la secuencia que codifica la proteína o polipéptido de interés, así como una o más secuencias promotoras de la expresión y al menos una secuencia de terminación.

10 “**Operativamente unida**”: implica que los diferentes elementos de la construcción génica están dispuestos de tal forma que actúan de manera operativa para ejecutar correctamente el programa genético para la cual fue diseñada dicha construcción génica.

15 “**Sistemas de expresión**”: son sistemas uni o pluricelulares que permiten producir proteínas recombinantes cuando son transfectados o transformados con vectores apropiados que comprenden el gen recombinante deseado o manipulados genéticamente para incorporar dicho gen en su propio genoma. Los sistemas unicelulares de expresión pueden ser procariotas como por ejemplo bacterias o eucariotas como levaduras, células de insecto o incluso cultivos de células de mamíferos. Ejemplos de sistemas de expresión

20 pluricelulares son larvas de insectos o plantas modificadas genéticamente. En el contexto de la invención son preferidos los sistemas de expresión unicelulares y más particularmente bacterias del género *Bacillus*, *Clostridium*, *Esherichia*, *Klebsiella*, *Pseudomonas*, *Streptomyces*, *Thermoanaerobacterium* o *Zymomonas* o hongos de las especies

25 *Aspergillus awamori*, *Aspergillus foetidus*, *Aspergillus fumigatus*, *Aspergillus japonicus*, *Aspergillus nidulans*, *Aspergillus niger*, *Aspergillus oryzae*, *Bjerkandera adusta*, *Ceriporiopsis aneirina*, *Ceriporiopsis caregiea*, *Ceriporiopsis rivulosa*, *Ceriporiopsis sufrufa*, *Ceriporiopsis subvermispora*, *Coprinus cinereus*, *Coriolus hirsutus*, *Fusarium bactridioides*, *Fusarium cereales*, *Fusarium culmorum*, *Fusarium graminearum*, *Fusarium graminum*, *Fusarium heterosporum*, *Fusarium negundi*, *Fusarium oxysporum*, *Fusarium reticulatum*, *Fusarium*

30 *roseum*, *Fusarium sambucinum*, *Fusarium sarcochroum*, *Fusarium sporotrichioides*, *Fusarium sulphureum*, *Fusarium torulosum*, *Fusarium trichothecioides*, *Fusarium venenatum*, *Humicola lanuginosa*, *Mucor miehei*, *Myceliophthora thermophila*, *Neurospora crassa*, *Penicillium purpurogenum*, *Phanerochaete chrysosporium*, *Phlebia radiata*, *Pleurotus eryngii*, *Saccharomyces calenbergensis*, *Saccharomyces cerevisiae*, *Saccharomyces diastaticus*, *Saccharomyces douglasii*, *Saccharomyces kluyveri*, *Saccharomyces norvensis*,

35 *Saccharomyces oviformis*, *Thielavia terrestres*, *Trametes villosa*, *Trametes versicolor*,

Trichoderma harzianum, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei*, *Trichoderma viride* o *Yarrowia lipolytica*.

5 “**Vector de expresión**”: es un sistema de transfección o transformación de células que permite incorporar una construcción génica en una célula y cuya finalidad es la expresión por parte de dicha célula de una proteína recombinante. Los tipos de vectores de expresión más habituales son plásmidos y vectores virales aunque en el contexto de la invención incluye cualquier otro tipo de vectores comúnmente usados.

10 “**Célula huésped**”: es la célula capaz recibir por transfección o transformación material genético externo que es capaz de expresar de forma transitoria o estable. El término célula huésped también se refiere no solo a la célula directamente transfectada o transformada sino también a toda su progenie.

15 “**Biomasa**”: en el contexto de la presente invención se asimila “biomasa” con “biomasa celulósica” esto es toda materia orgánica que comprenda material lignocelulósico y que sea susceptible de ser transformada a sus componentes más elementales (azúcares fermentables) mediante un proceso industrial de degradación con enzimas celulasas. La biomasa puede ser, por ejemplo, la fracción biodegradable de los productos, residuos y
20 restos de origen biológico procedentes de la agricultura (incluyendo sustancias vegetales, tales como residuos de cultivos, y sustancias animales) industrias forestales (tales como recursos madereros) e industrias relacionadas que incluyen pesquerías y acuicultura, así como la fracción celulósica biodegradable de los residuos industriales y urbanos, tales como residuos sólidos urbanos o residuos de papel. En el contexto de la invención de manera
25 preferente la biomasa es paja o la fracción orgánica de residuos sólidos urbanos y de manera aun más preferente, la biomasa es biomasa vegetal seleccionada a partir de la lista que consiste en: biomasa rica en azúcares fermentables, tales como caña de azúcar; biomasa de almidón, por ejemplo, granos o paja de trigo; o maíz o paja de maíz o grano de maíz o fibra de maíz; o granos o paja de cebada; o granos o paja de sorgo. La biomasa
30 también puede ser, arroz, hierba, ramas, etc.

“**Degradación de biomasa**”: Es el proceso de transformación de la biomasa celulósica en sus componentes elementales (azúcares fermentables), principalmente glucosa, mediante la acción de celulasas.

35

“**Azúcares o azúcares fermentables**”: es el producto resultante del proceso de degradación de biomasa y comprende todos aquellos azúcares que son susceptibles de ser transformados químicamente o fermentados para dar lugar a bioproductos de alto valor añadido. Aunque no es el único, el principal azúcar fermentable es la glucosa.

5

“**Bioproductos**”: son los productos de alto valor añadido que pueden obtenerse por transformación química de los azúcares o por fermentación de dichos azúcares con diferentes microorganismos. Sin pretender ser limitativa, una lista de posibles microorganismos fermentantes es la siguiente: *Bacillus thermoglucosidarius*, *Clostridium butyricum*, *Corynebacterium glutamicum*, *Enterobacter aerogenes*, *Escherichia coli*, *Geobacillus thermoglucosidarius*, *Klebsiella oxytoca*, *Lactobacillus sp. Leunosoc mesenteroides*, *Thermoanaerobacter BG1L1*, *Thermoanaerobacter ethanolicus*, *Thermoanaerobacter mathranii*, *Thermoanaerobacter thermosaccharolyticum*, *Zymobacter palmae*, *Zymomonas mobilis* *Candida arabinofermentans*, *Candida boidinii*, *Candida diddensis*, *Candida fermentans*, *Chrysosporium lucknowense*, *Candida pastoris*, *Candida shehatae*, *Candida sonorensis*, *Candida tropicalis*, *Hansenula anómala*, *Hansenula polymorpha*, *Kluyveromyces fragilis*, *Kluyveromyces marxianus*, *Pichia pastoris*, *Pichia stipitis*, *Saccharomyces cerevisiae*, *Saccharomyces bulderi*, *Saccharomyces barnetti*, *Saccharomyces exiguus*, *Saccharomyces diastaticus*, *Saccharomyces uvarum* o *Schizosaccharomyces pombe* o mezclas de estos. Estos y otros microorganismos pueden proporcionar por fermentación de diferentes azúcares, bioproductos entre los cuales se pueden citar de manera no limitativa los siguientes: alcoholes como etanol, metanol, butanol, hexanol, octanol, decanol, dodecanol, 1,3-butanediol(1,3-diol), 1-alcohol; ácidos orgánicos como ácido cítrico, ácido acético, ácido itaconico, ácido láctico, ácido glutámico, ácido succínico, beta-cetoácido, beta-cetoalcohol, beta-hidroxiácido; cetonas como la acetona; gases como hidrógeno o dióxido de carbono; hidrocarburos como alcanos alquenos o alquinos; sustancias nitrogenadas como aminas, amida, nitrocompuestos o nitrilos; haluros; aminoácidos como ácido glutámico; antibióticos como penicilina o tetraciclinas; vitamina como riboflavina, vitamina B12 o betacaroteno; ácidos grasos como ácido dodecanoico, ácidos grasos trans- Δ^2 o ácido palmítico; y otros productos como etileno, glicerol, 1,3-propano-diol, betalactano, cefalosporinas, ácidos grasos trans o furano.

Polipéptidos con actividad celulasa y fragmentos funcionales de los mismos

El aspecto principal de la presente invención se refiere a un polipéptido aislado con actividad celulasa que comprende o consiste en la secuencia SEQ ID NO1 o que posee un grado de homología o de identidad de al menos un 80% con la SEQ ID NO1 o que está codificado por

35

un polinucleótido que hibrida bajo condiciones de baja, media, media-alta, alta o muy alta astringencia con la secuencia SEQ ID NO 6 que codifica el polipéptido maduro.

5 En una realización particular de la invención el polipéptido posee un grado de homología o de identidad de al menos un 85%, preferiblemente de al menos un 90%, más preferiblemente de al menos un 95% y aun más preferiblemente de al menos un 99% con la SEQ ID NO 1.

10 Por ejemplo, un polipéptido que posee un alto grado de homología (de entorno al 92%) con el polipéptido de SEQ ID NO 1 y representa una realización particular de la invención es el polipéptido con la secuencia SEQ ID NO 2. Esta secuencia posee 927 aminoácidos frente a los 928 de la SEQ ID NO 1. El polipéptido de secuencia SEQ ID NO 2 tiene una actividad celulasa comparable a la del polipéptido de SEQ ID NO 1.

15 No obstante en una realización preferida, el polipéptido de la invención se corresponde con la secuencia SEQ ID NO 1.

20 El polipéptido de SEQ ID NO 1 fue identificado y aislado a partir del screening de una metagenoteca de microorganismos del suelo. El polipéptido de SEQ ID NO 1 tiene 928 aminoácidos y posee una actividad celulasa muy marcada. Aunque el perfil de actividad indica que el polipéptido de SEQ ID NO 1 exhibe actividad exoglucanasa y beta-glucosidasa, la actividad principal del mismo es endoglucanasa.

25 Los ensayos de actividad sobre el polipéptido de la invención demuestran que éste es activo a un amplio rango de temperaturas que va desde alrededor de 30°C hasta alrededor de 65°C. Preferiblemente, el polipéptido es activo a temperaturas de entre 30°C y 60°C, más preferiblemente de entre 30°C y 50°C, más preferiblemente de entre 37°C y 50°C y más preferiblemente de entre 37°C y 45°C. La temperatura óptima de actividad del polipéptido de la invención se encuentra entre los 45 y 50°C (ver figura 4 y 5).

30 Asimismo, el polipéptido de la invención muestra una gran termorresistencia y es capaz de mantener el mismo grado de actividad después de al menos 5 días de incubación a 30°C, o después de 6 días a 37°C. El grado de actividad se mantiene alto para el polipéptido incubado a 40°C durante dos días y empieza a decrecer a partir del tercero. Finalmente,
35 con una incubación a 42°C el polipéptido pierde su actividad a los dos días. (ver figura 6).

En cuanto al pH óptimo, el polipéptido de la invención trabaja mejor a pH ligeramente ácido y es capaz de mantener la mayor parte de su actividad a pH en el rango de 5 a 7, aunque el pH óptimo de actividad se encuentra en el rango de 5-6,5 (ver figuras 7 y 8).

5 El polipéptido de SEQ ID NO 1 posee 928 aminoácidos que comprenden diferentes dominios. En su extremo N-terminal, de la posición 1 a la 32 se encuentra el péptido señal con la señal de translocación Twin-arginine. El dominio catalítico o dominio con la actividad
celulasa (perteneciente a la familia 5) se encuentra entre los aminoácidos 41 y 378. El
10 polipéptido de SEQ ID NO 1 muestra también diferentes dominios de unión a carbohidratos,
más concretamente entre las posiciones 443-570 y 664-780 se encuentran dos dominios de
unión a carbohidratos de tipo 2 (CBM2) y en la posición 474-630 un dominio de unión a
carbohidratos de tipo 6 (CBM 6). La figura 3 es una representación donde se pueden
observar los diferentes dominios y su ubicación en la secuencia del polipéptido de la
invención.

15

El polipéptido de la invención también contempla variantes del polipéptido de SEQ ID NO 1
que comprenden sustituciones, deleciones y/o inserciones en una o más posiciones. Los
cambios de aminoácidos pueden ser de naturaleza menor, esto es sustituciones o
inserciones conservativas que no afecten significativamente el plegado y o actividad del
20 polipéptido; pequeñas deleciones normalmente de 1 a 30 aminoácidos; pequeñas
extensiones de los extremos amino y carboxiterminal como pequeños péptidos de unión de
20-25 residuos o pequeñas extensiones que faciliten la purificación mediante el cambio de la
carga neta u otra función, colas de polihistidinas, epítomos antigénicos o dominios de unión.

25 También forman parte de la invención variantes con cambios del polipéptido de la invención
que puedan alterar de manera controlada y dirigida propiedades físico químicas del
polipéptido. Por ejemplo, estas variantes incluyen cambios de aminoácidos que puedan
mejorar la estabilidad térmica del polipéptido de la invención, cambiar su especificidad por el
sustrato, temperatura óptima, el pH óptimo etc.

30

Otro aspecto adicional de la invención son los fragmentos funcionales del polipéptido de la
invención. Así pues, la invención no solo se extiende al polipéptido de SEQ ID NO1 o a
polipéptidos que posean un grado de homología o de identidad de al menos un 80% con la
SEQ ID NO 1, sino también a los fragmentos funcionales de estos, es decir a aquellos
35 fragmentos que mantienen al menos entre el 50% y el 100% de la actividad celulasa del
polipéptido de la invención.

La actividad celulasa del polipéptido de la invención se encuentra en el dominio catalítico, entre los aminoácidos 41 y 378, por lo que para que un fragmento del polipéptido de la invención mantenga la actividad celulasa debe contener al menos este fragmento. Por tanto, los fragmentos funcionales de la presente invención comprenden al menos los aminoácidos entre las posiciones 41 y 378 de la SEQ ID NO 1.

En una realización particular dichos fragmentos comprenden o consisten en las secuencias SEQ ID NO 3, SEQ ID NO 4 o SEQ ID NO 5. Todos estos fragmentos mantienen la actividad celulasa del péptido aunque no tienen por qué mantener el mismo perfil de actividad.

Polinucleótidos

Otro aspecto relevante de la invención, viene representado por la secuencia polinucleotídica que codifica el polipéptido de la invención o los fragmentos funcionales del mismo.

El polinucleótido puede ser cualquier tipo de ácido nucleico de cadena doble o simple como por ejemplo ADN o ARN. Se consideran incluidos en el ámbito de la invención tanto el ADN genómico aislado a partir de clon de la librería metagenómica y que codifica el polipéptido de SEQ ID NO 1, así como el cDNA obtenido por transcripción reversa de este. También son parte de la invención los ARNm que codifican el polipéptido de la invención.

Asimismo, forman parte de la invención las modificaciones del polinucleótido que codifican las posibles variantes del polipéptido de la invención descritas en el apartado relativo al polipéptido con actividad celulasa y sus fragmentos funcionales.

En una realización particular la secuencia nucleotídica comprende o consiste en una de las siguientes secuencias: SEQ ID NO 6, SEQ ID NO 7, SEQ ID NO 8, SEQ ID NO 9 o SEQ ID NO 10.

La SEQ ID NO 6 es la secuencia polinucleotídica que codifica el polipéptido de SEQ ID NO 1. La invención también contempla las secuencias polinucleotídicas que comprendan variaciones en uno o varios nucleótidos pero que no resultan en un cambio de la secuencia polipeptídica que codifica la SEQ ID NO 6.

La SEQ ID NO 7 es la secuencia polinucleotídica que codifica el polipéptido con la secuencia SEQ ID NO2. La invención también contempla las secuencias polinucleotídicas

que comprendan variaciones en uno o varios nucleótidos pero que no resultan en un cambio de la secuencia polipeptídica que codifica la SEQ ID NO 7.

5 Las SEQ ID NO 8, 9 y 10 representan la secuencia codificadora de algunos de los fragmentos funcionales preferidos de la invención. La invención también contempla las secuencias polinucleotídicas que comprendan variaciones en uno o varios nucleótidos pero que no resultan en un cambio de la secuencia polipeptídica de estos fragmentos.

Construcción génica

10 Otro aspecto adicional de la invención es una construcción génica que comprende una secuencia polinucleotídica como las mencionadas más arriba que se encuentra operativamente unida a una o más secuencias de control que dirigen la expresión de polipéptidos en sistemas de expresión.

15 Las secuencias de control pueden ser un promotor, un polinucleótido que es reconocido por el sistema de expresión (célula huésped) para la expresión del polipéptido de la invención. Los promotores contienen secuencias de control transcripcional que median la expresión del polipéptido. El promotor puede ser cualquier polinucleótido que muestre actividad transcripcional en la célula huésped incluyendo promotores mutantes, truncados e híbridos y
 20 pueden ser obtenidos de genes que codifican polipéptidos extracelulares o intracelulares ya sean homólogos o heterólogos a la célula huésped.

Ejemplos de promotores adecuados para dirigir la transcripción de los polipéptidos de la invención en células huésped bacterianas pueden ser los promotores obtenidos del gen alfa-
 25 amylasa de *Bacillus amyloliquefaciens* (amyQ), del gen alfa-amylasa de *Bacillus licheniformis* (amyL), del gen de la penicilinas de *Bacillus licheniformis* (penP), del gen de la amilasa maltogénica de *Bacillus stearothermophilus* (amyM), del gen de la levansucrasa de *Bacillus subtilis* (sacB), de los genes xylA and xylB *Bacillus subtilis*, del gen cryIIIA de *Bacillus thuringiensis* (Agaisse and Lereclus, 1994, Molecular Microbiology 13: 97-107), del
 30 operon lac de *E. coli*, del promotor trc de *E. coli* (Egon et al., 1988, Gene 69: 301-315), del gen agarasa de *Streptomyces coelicolor* (dagA), y del gen beta-lactamase procariotico (Villa-Kamaroff et al., 1978, Proc. Natl. Acad. Sci. USA 75: 3727-3731), así como del promotor tac (DeBoer et al., 1983, Proc. Natl. Acad. Sci. USA 80: 21-25). Promotores adicionales se describen en "Useful proteins from recombinant bacteria" Gilbert et al., 1980, Scientific
 35 American 242: 74-94; y en Sambrook, E.F. Fritsch, and T. Maniatis, 1989, Molecular

Cloning, A Laboratory Manual, 2d edition, Cold Spring Harbor, New York. Ejemplos de promotores tándem se describen en WO 99/43835.

Ejemplos de promotores adecuados para dirigir la transcripción en células hospedadoras de hongos filamentosos pueden ser los promotores obtenidos de los genes de la acetamidasa de *Aspergillus nidulans* acetamidase, de la alfa amilasa neutra de *Aspergillus niger*, de la alfa amilasa acidoestable de *Aspergillus niger*, de la glucoamilasa de *Aspergillus niger* o *Aspergillus awamori* glucoamylase (glaA), de la TAKA amilasa de *Aspergillus oryzae*, de la proteasa alcalina de *Aspergillus oryzae*, de la triosa fosfato isomerasa de *Aspergillus oryzae*, de la proteasa similar a tripsina de *Fusarium oxysporum* (WO 96/00787), de la aminoglucosidasa de *Fusarium venenatum* (WO 00/56900), *Fusarium venenatum* Daria (WO 00/56900), *Fusarium venenatum* Quinn (WO 00/56900), de la lipasas de *Rhizomucor miehei*, de la proteinasa aspartica de *Rhizomucor miehei*, de la beta-glucosidasa de *Trichoderma reesei*, de la celobiohidrolasa I de *Trichoderma reesei*, de la celobiohidrolasa II de *Trichoderma reesei*, de la endoglucanasa I de *Trichoderma reesei*, de la endoglucanasa II de *Trichoderma reesei*, de la endoglucanasa III de *Trichoderma reesei*, de la endoglucanasa V de *Trichoderma reesei*, de la xilanasa I de *Trichoderma reesei*, de la xilanasa II de *Trichoderma reesei*, de la xilanasa III de *Trichoderma reesei*, de la beta xilosidasa de *Trichoderma reesei*, y del factor de elongación de la traducción de *Trichoderma reesei*, así como de promotor NA2-tpi (un promotor modificado del gen de la alfa amilasa neutra de *Aspergillus* en la que la secuencia líder no traducida se ha sustituido por una secuencia líder no traducida de un gen de la triosa fosfato isomerasa de *Aspergillus*; ejemplos no limitativos incluyen promotores modificados del gen de la alfa amilasa neutra de *Aspergillus niger* en el que la secuencia líder no traducida se ha sustituido por una secuencia líder no traducida del gen de la triosa fosfato isomerasa de *Aspergillus nidulans* o *Aspergillus oryzae*); y promotores mutantes, truncados, e híbridos de los mismos. Otros promotores se describen en US6011147. Otro promotor adecuado es el promotor Pcbh de *Myceliophthora termophila*.

En células de levaduras los promotores útiles se pueden obtener de los genes de la enolasa de *Saccharomyces cerevisiae* (ENO-1), de la galactoquinasa de *Saccharomyces cerevisiae* (GAL1), de la alcohol deshidrogenasa/gliceroaldehido-3-fosfatasa deshidrogenasa de *Saccharomyces cerevisiae* (ADH1, ADH2/GAP), de la triosa fosfato isomerasa de *Saccharomyces cerevisiae* (TPI), de la metalotioneina de *Saccharomyces cerevisiae* (CUP1), y la 3 fosfoglicerato quinasa de *Saccharomyces cerevisiae*. Otros promotores útiles para levaduras se describen en Romanos et al., 1992, Yeast 8: 423-488.

La secuencia de control puede también ser un terminador de la transcripción, que es reconocido por la célula huésped para terminar la transcripción. El terminador está operativamente unido al término 3' del polinucleótido que codifica el polipéptido. Cualquier terminador que sea funcional en la célula huésped puede usarse en el contexto de la presente invención.

Algunos terminadores apropiados para células huésped bacterianas pueden obtenerse a partir de genes de la proteasa alcalina de *Bacillus clausii* (aprH), de la alfa amilasa de *Bacillus licheniformis* (amyL), y *Escherichia coli* ribosomal RNA (rrnB).

Algunos terminadores adecuados para células huésped de hongos filamentosos se pueden obtener de los genes acetamidasa de *Aspergillus nidulans*, de la antranilato sintasa de *Aspergillus nidulans anthranilate*, de la glucoamilasa de *Aspergillus niger*, de la alfa-glucosidasa de *Aspergillus niger*, de la TAKA amilasa de *Aspergillus oryzae*, de la proteasa semejante a tripsina de *Fusarium oxysporum*, de la beta-glucosidasa de *Trichoderma reesei*, de la celobiohidrolasa I de *Trichoderma reesei*, de la celobiohidrolasa II de *Trichoderma reesei*, de la endoglucanasa I de *Trichoderma reesei*, de la endoglucanasa II de *Trichoderma reesei*, de la endoglucanasa III de *Trichoderma reesei*, de la endoglucanasa V de *Trichoderma reesei*, de la xilanasa I de *Trichoderma reesei*, de la xilanasa II de *Trichoderma reesei*, de la xilanasa III de *Trichoderma reesei*, de la beta-xilosidasa de *Trichoderma reesei* y del factor de elongación de traducción de *Trichoderma reesei*. Otra secuencia terminadora es la secuencia Tcbh.

Algunos terminadores adecuados para células huésped de levaduras se pueden obtener a partir de genes de la enolasa de *Saccharomyces cerevisiae*, del citocromo C de *Saccharomyces cerevisiae* (CYC1) y de la gliceraldehido-3-fosfato deshidrogenasa de *Saccharomyces cerevisiae*. Otros terminadores útiles en levaduras se describen en Romanos et al., 1992, *Yeast* 8: 423-488.

La secuencia control puede ser también una región estabilizadora del RNAm posterior al promotor pero anterior a la secuencia codificante de un gen que incrementa la expresión del gen. Ejemplos de secuencias estabilizadoras de RNAm se pueden obtener del gen cryIIIA de *Bacillus thuringiensis* (WO 9425612) del gen SP82 de *Bacillus subtilis* (Hue et al., 1995, *Journal of Bacteriology* 181: 3465-3471).

35

La secuencia de control puede ser también una secuencia líder, una región no traducida del RNAm que es importante en la traducción por parte de la célula huésped. La secuencia líder está operativamente unida a la región 5' terminal del polinucleótido que codifica el polipéptido. Cualquier secuencia líder funcional en la célula huésped puede usarse.

5

Secuencias líder adecuadas en hongos filamentosos se pueden obtener de los genes de la TAKA amilasa de *Aspergillus oryzae* y triosa fosfato isomerasa de *Aspergillus nidulans*.

10 Secuencias líder apropiadas para levaduras se pueden obtener de los genes de la enolasa de *Sacharomyces cerevisiae* (ENO-1), de la 3-fosfogliceratoquinasa de *Saccharomyces cerevisiae*, del de factor alfa de *Saccharomyces cerevisiae*, y del de la alcohol deshidrogenasa/Gliceraldehído-3-fosfato deshidrogenasa de *Saccharomyces cerevisiae* (ADH2/GAP).

15 La secuencia de control también puede ser una secuencia de poliadenilación una secuencia operativamente unida a término 3' del polinucleótido y que cuando es transcrito la célula huésped la reconoce como una señal para añadir residuos poliadenosina al mRNA transcrito. Cualquier señal de poliadenilación funcional en la célula huésped puede ser usada.

20

Algunas secuencias de poliadenilación adecuadas para células huésped de hongos filamentosos se pueden obtener a partir de los genes de la antranilato sintasas de *Aspergillus nidulans*, de la glucoamilasa de *Aspergillus niger*, de la alfa-glucosidasa de *Aspergillus niger*, de la TAKA amilasa de *Aspergillus oryzae* y de la proteasa similar a 25 tripsina de *Fusarium oxysporum*.

Secuencias de poliadenilación útiles en células huésped de levaduras se describen en Guo and Sherman, 1995, Mol. Cellular Biol. 15: 5983-5990.

30 La secuencia de control puede ser también un péptido señal que codifica un péptido señal unido al extremo N-terminal del polipéptido y dirige al polipéptido a las vías secretoras de señalización celular. La región 5' terminal de la secuencia codificante puede contener una secuencia codificante de un péptido señal naturalmente unido en el fragmento de traducción con el segmento de la secuencia codificante del polipéptido. Alternativamente el termino 5' 35 de la secuencia codificante puede contener una secuencia codificante de un péptido señal que es ajeno a la secuencia codificante. Una secuencia codificante de un péptido señal

ajeno puede ser necesario cuando la secuencia codificante no contiene de manera natural una secuencia codificante de un péptido señal. Por otro lado, una secuencia ajena puede estar sencillamente reemplazando una secuencia codificante del péptido secuencia natural con el fin de potenciar la secreción del polipéptido. No obstante, cualquier secuencia
 5 codificante de un péptido señal que dirija el polipéptido expresado por las vías de secreción celular puede ser usada.

Secuencias codificantes de péptidos señal adecuadas para bacterias pueden obtenerse de los genes de la amilasa maltogenica de *Bacillus* NCIB 1 1837, de la subtilisina de *Bacillus*
 10 *licheniformis*, de la beta-lactamasa de *Bacillus licheniformis*, de la amilasa de *Bacillus stearothermophilus*, de las proteasas neutras de *Bacillus stearothermophilus* (nprT, nprS, nprM) y de la proteasa A de *Bacillus subtilis*. Otros peptides señal se describen en Simonen and Palva, 1993, Microbiological Reviews 57: 109-137. Otras secuencias señal apropiadas para su uso en bacterias son las secuencias Twin arginine y las secuencia Sec.

15 Secuencias codificantes de péptidos señal adecuadas para hongos filamentosos pueden obtenerse de los genes de la amilasa neutra de *Aspergillus niger*, de la glucoamilasa de *Aspergillus niger*, de la TAKA amilasa de *Aspergillus oryzae*, de la celulasa de *Humicola insolens*, de la endoglucanasa V de *Humicola insolens*, de la lipasas de *Humicola*
 20 *lanuginosa* lipase y de la proteinasa aspartica de *Rhizomucor miehei*.

Péptidos señal para levaduras se pueden obtener de los genes del factor alfa de *Sacharomyces cerevisiae* y de la invertasa de *Saccharomyces cerevisiae*. Otras secuencias codificantes de péptidos señal se describen en Romanos et al., 1992, Yeast 8: 423-488.

25 También puede ser deseable añadir secuencias reguladoras que regulen la expresión del polipéptido en función al crecimiento de la célula huésped. Ejemplos de secuencias reguladoras son aquellas que causan que la expresión del gen se active o se inactive en respuesta a estímulos químicos o físicos, incluyendo la presencia de un compuesto
 30 regulador. Secuencias reguladoras sistemas procariotas incluyen los sistemas operadores lac, tac y trp. En levaduras se pueden usar los sistemas ADH2 o GAL1. En hongos filamentosos se pueden usar el promotor glucoamilasa de *Aspergillus niger*, el promotor TAKA alfa amilasa de *Aspergillus oryzae*, y el promotor glucoamilasa de *Aspergillus oryzae*, el promotor celobiohidrolasa I de *Trichoderma reesei* el promotor celobiohidrolasa II de
 35 *Trichoderma reesei* y el promotor Pcbh de *Myceliophthora thermophila*.

Otros ejemplos de secuencias reguladoras son aquellas que permiten la amplificación génica. En sistemas eucariotas, estas secuencias reguladoras incluyen las secuencias reguladoras del gen de la dihidrofolato reductasa que se amplifica en presencia de metotrexato y los genes de la metalotioneína que se amplifican con metales pesados. En
5 estos casos el polinucleótido que codifica el polipéptido estaría operativamente unido a la secuencia reguladora.

El vector de expresión

En otro aspecto la invención también se refiere a un vector de expresión que comprende un
10 polinucleótido o una construcción génica como los anteriormente descritos.

El vector de expresión comprende el polinucleótido de la invención, al menos un promotor y señales de terminación transcripcionales y traduccionales. Las diferentes secuencias nucleotídicas y de control pueden estar unidas para producir un vector de expresión
15 recombinante que puede incluir uno o más sitios de restricción para permitir la inserción o sustitución del polinucleótido que codifica el polipéptido de la invención en dichos sitios.

Alternativamente, el polinucleótido puede expresarse mediante la inserción del polinucleótido o la construcción génica de la invención en un vector apropiado para la
20 expresión. Al crear el vector de expresión la secuencia codificante debe estar operativamente unida con las secuencias de control apropiadas para garantizar la expresión. El vector de expresión puede ser cualquier vector (por ejemplo virus o plásmido) que pueda ser sometido a procesos de ADN recombinante y que pueda llevar a cabo la expresión del polinucleótido. La elección del vector depende esencialmente de la compatibilidad del vector
25 con la célula huésped en la cual vaya a ser introducido el vector. El vector puede ser un plásmido lineal o circular.

El vector puede ser un vector autónomamente replicante, es decir, un vector que existe como una entidad extracromosomal y cuya replicación es independiente de la replicación del
30 cromosoma como, por ejemplo, un plásmido, un elemento extracromosomal, un minicromosoma, o un cromosoma artificial. El vector puede contener cualquier medio para asegurar su autoreplicación.

Asimismo, el vector puede ser uno que al ser introducido en la célula huésped se integra en
35 el genoma y se replica junto con el o los cromosomas en los cuales se inserta. Además se puede usar un vector único o plásmido o más vectores o plásmidos que conjuntamente

contengan el ADN total que ha de ser introducido en el genoma de la célula huésped o también un transposón.

5 El vector contiene preferiblemente uno o más marcadores que permiten la selección sencilla de células transformadas o transfectadas. Un marcador de selección es un gen cuyo producto proporciona resistencia a un producto, un virus, resistencia a metales pesados, prototrofia a auxotrofos etc.

10 Ejemplos de marcadores de selección adecuados para bacterias son los genes DAL de *Bacillus licheniformis* o *Bacillus subtilis*, o marcadores que confieren resistencia a antibióticos como resistencia a ampicilina, cloranfenicol, kanamicina, neomicina, espectinomicina o tetraciclina.

15 Marcadores apropiados para levaduras incluyen pero no se limitan a ellos, ADE2, HIS3, LEU2, LYS2, MET3, TRP1 y URA3.

20 Marcadores de selección para uso en hongos filamentosos incluyen pero no se limitan a, adeA (fosforibosilaminoimidazol-succinocarboxamida sintasa), adeB (fosforibosilaminoimidazol sintasa), amdS (acetamidasa), argB (ornitina carbamoiltransferasa), bar (fosfinotricin acetiltransferasa), hph (higromicina fosfotransferasa), niaD (nitrato reductasa), pyrG (orotidina-5'-fosfato decarboxilasa), sC (sulfato adeniltransferasa) y trpC (antranilato sintasa) y equivalentes de los mismos. Los marcadores preferidos para células de *Aspergillus* son los genes amdS y Pyr G de *Aspergillus nidulans* o *Aspergillus oryzae* y el gen bar de *Streptomyces hygroscopicus*. Los marcadores preferidos para células de
25 *Trichoderma cell* son los genes adeA, adeB, amdS, hph, y pyrG. Otros marcadores preferidos son Pyr 4, Pyr 5, CysC y trp.

30 El vector de la invención debe contener preferiblemente elementos que permitan la integración del vector en el genoma de la célula huésped o la replicación autónoma del vector en la célula de manera independiente al genoma. Para la integración en el genoma el vector puede contener elementos de integración por recombinación homóloga o no homóloga. Incluso, puede incluir secuencias polinucleotídicas para dirigir la integración por recombinación homóloga en locus concretos en el(los) cromosoma(s). El elemento integracional puede ser cualquier secuencia homóloga a la secuencia diana en el genoma
35 de la célula huésped. Además los elementos integracionales pueden ser secuencias codificantes o no codificantes.

Para la replicación autónoma, el vector de la invención puede comprender un origen de replicación que le permita al vector autorreplicarse en la célula huésped. El origen de replicación puede ser cualquier replicador plasmídico que funcione en la célula huésped.

- 5 Ejemplos de orígenes de replicación bacterianos son los orígenes de replicación de los plásmidos pBR322, pUC19, pACYC177 y pACYC184 que permiten la replicación en *E. coli* y pUB1 10, pE194, pTA1060 y [rho][Alpha][Mu][beta][lota] que permiten la replicación en *Bacillus*.
- 10 Ejemplos de orígenes de replicación para usar en levaduras son los orígenes de replicación de dos micrones ARS1, ARS4, la combinación de ARS1 y CEN3, y la combinación de ARS4 y CEN6.

- Ejemplos de orígenes de replicación en hongos filamentosos son AMA1 y ANSI (Gems et al., 1991, *Gene* 98: 61 -67; Cullen et al., 1987, *Nucleic Acids Res.* 15: 9163-9175; WO 0024883). WO0024883 describe como se puede conseguir el aislamiento del gen AMA1 y la construcción de plásmidos o vectores que lo comprenden.
- 15

- Se puede insertar más de una copia del polinucleótido en la célula huésped para incrementar la producción del polipéptido de la invención. El incremento en el número de copias del polinucleótido se puede conseguir integrando al menos una copia adicional de la secuencia en el genoma de la célula huésped.
- 20

Célula huésped

- 25 Otro aspecto relevante de la invención se relaciona con una célula huésped que comprende un polinucleótido, una construcción génica o un vector de expresión como los mencionados anteriormente y que expresa el polipéptido de la invención o un fragmento funcional o una variante del mismo.
- 30 La célula huésped puede ser una célula eucariota o una célula procariota.

En una realización particular la célula huésped es una célula procariota que se selecciona preferiblemente del grupo de bacterias de los géneros *Bacillus*, *Clostridium*, *Esherichia*, *Klebsiella*, *Pseudomonas*, *Streptomyces*, *Thermoanaerobacterium* o *Zymomonas*.

- 35 En otra realización particular, la célula huésped es una célula eucariota que se selecciona preferentemente del grupo de hongos de las especies *Aspergillus awamori*, *Aspergillus*

foetidus, Aspergillus fumigatus, Aspergillus japonicus, Aspergillus nidulans, Aspergillus niger, Aspergillus oryzae, Bjerkandera adusta, Ceriporiopsis aneirina, Ceriporiopsis caregiea, Ceriporiopsis rivulosa, Ceriporiopsis sufrufa, Ceriporiopsis subvermispora, Coprinus cinereus, Coriolus hirsutus, Fusarium bactridioides, Fusarium cereales, Fusarium culmorum, Fusarium graminearum, Fusarium graminum, Fusarium heterosporum, Fusarium negundi, Fusarium oxysporum, Fusarium reticulatum, Fusarium roseum, Fusarium sambucinum, Fusarium sarcochrom, Fusarium sporotrichioides, Fusarium sulphureum, Fusarium torulosum, Fusarium trichothecioides, Fusarium venenatum, Humicola lanuginosa, Mucor miehei, Myceliophthora thermophila, Neurospora crassa, Penicillium purpurogenum, Phanerochaete chrysosporium, Phlebia radiata, Pleurotus eryngii, Saccharomyces calsbergensis, Saccharomyces cerevisiae, Saccharomyces diastaticus, Saccharomyces douglasii, Saccharomyces kluyveri, Saccharomyces norvensis, Saccharomyces oviformis, Thielavia terrestres, Trametes villosa, Trametes versicolor, Trichoderma harzianum, Trichoderma koningii, Trichoderma longibrachiatum, Trichoderma resei, Trichoderma viride o Yarrowia lipolytica.

Cualquier medio comúnmente conocido para introducir material genético en células huésped como transformación de protoplastos, transformación de células competentes, electroporación o conjugación puede ser usado para transformar o transfectar las células huésped.

Método de producción del polipéptido de la invención

Otro aspecto importante de la invención es el método para producir un polipéptido con actividad celulolítica de la invención que comprende:

- a) Cultivar una célula huésped como la mencionada anteriormente en las condiciones necesarias para la producción del polipéptido,
- b) Recuperar el polipéptido producido.

La etapa a) relativa al cultivo de las células huésped se lleva a cabo en un medio nutritivo adecuado para la célula para producir el polipéptido. Los medios nutritivos adecuados para el cultivo están disponibles comercialmente o pueden ser preparados de acuerdo con protocolos bien conocidos por un experto en la materia.

El cultivo se puede llevar a cabo a escala de laboratorio o a escala industrial en fermentadores industriales mediante métodos comúnmente conocidos por el experto en la materia. El proceso industrial puede llevarse a cabo en continuo o discontinuo.

En una realización preferida el cultivo de la etapa a) se lleva a cabo a nivel industrial en fermentadores industriales bajo condiciones controladas (temperatura, pH, nutrientes etc.). Además preferentemente el proceso se lleva a cabo en continuo.

5 El polipéptido se puede recuperar mediante cualquier método conocido y convencional. Por ejemplo, el polipéptido de la invención se puede recuperar del medio mediante colección, centrifugación, filtración, extracción, evaporación o precipitación. Asimismo, una vez recuperado el polipéptido de la invención puede ser purificado si fuera necesario por procesos comúnmente conocidos por un experto en la materia.

10

No obstante, en una realización preferida de la invención, la recuperación comprende la hidrólisis de las células huésped para liberar al medio todo el contenido celular incluido el polipéptido de la invención y el trasvase de la corriente con el sobrenadante, en el cual se encuentra el polipéptido de la invención, directamente al lugar o fermentador donde se
15 llevará a cabo el proceso de digestión o degradación de la biomasa.

Composición celulolítica

Otro aspecto de la invención es una composición celulolítica que comprende al menos un polipéptido de acuerdo con la invención o al menos un fragmento funcional o una variante
20 del mismo y, opcionalmente, al menos un elemento que favorezca o facilite la degradación de biomasa.

La composición de la invención comprende como elemento esencial de la misma al menos un polipéptido de acuerdo con la invención o al menos un fragmento funcional o una variante
25 del mismo, así como los elementos esenciales para que este ejerza correctamente su función celulolítica. Estos elementos son básicamente un tampón adecuado (preferiblemente un tampón citrato-fosfato, 50 mM, pH 5) y un medio electrolíticamente adecuado para mantener al máximo la funcionalidad del polipéptido de la invención o de sus fragmentos funcionales o variantes.

30 La composición de la invención normalmente se presenta en forma líquida aunque puede presentarse en forma sólida reconstituible.

De manera opcional aunque preferida, la composición comprende al menos un elemento adicional que favorece o facilita la degradación de biomasa. Dicho elemento que favorece o
35 facilita la degradación de biomasa puede ser cualquier agente químico, físico o biológico que ayude en la degradación completa del material lignocelulósico de la biomasa.

En una realización preferida la composición comprende como elemento que facilita la degradación de la biomasa una o más enzimas seleccionadas de entre endoglucanasas, exoglucanasas, celobiohidrolasas, beta-glucosidasas, glicosil-hidrolasas, hemicelulasas, xylanasas, enzimas lignolíticas o mezclas de las mismas.

5 Uso del polipéptido

Otro aspecto general de la invención se refiere al uso del polipéptido de la invención, de un fragmento funcional o una variante del mismo o de la composición celulolítica de la invención de para la degradación de biomasa.

10 Este aspecto general del uso del polipéptido de la invención, de un fragmento funcional o de una variante del mismo o de la composición celulolítica de la invención para la degradación de biomasa está encuadrado dentro del campo de la invención que es la producción industrial de bioproductos.

15 En este sentido, un aspecto particular de la invención se refiere a un procedimiento para producir azúcares fermentables que comprende incubar biomasa con el polipéptido de la invención, con un fragmento o variante del mismo o con la composición celulolítica de la invención.

20 Estos azúcares fermentables obtenidos se encuentran en el sobrenadante obtenido tras la incubación. Dado que la presente invención está enfocada principalmente en el campo de la producción de bioproductos normalmente la corriente con los azúcares fermentables se redirige directamente desde el tanque donde se lleva a cabo la digestión hidrolítica de la biomasa a donde se da la transformación química de este sustrato preferiblemente por
25 fermentación microbiana.

Esto no quita que en una realización particular los azúcares fermentables pueden opcionalmente ser recuperados tras la incubación y ser procesados y purificados para otros usos diferentes. La recuperación de los azúcares en este caso se podría llevar a cabo por
30 cualquier método adecuado para tal fin como por ejemplo centrifugación, filtración o deposición por gravedad.

Otro aspecto particular, relacionado con el anterior, es el procedimiento para producir un bioproducto a partir de biomasa. Este aspecto está relacionado con el anterior ya que
35 implica, en primer lugar, llevar a cabo el procedimiento antes mencionado de degradar la

biomasa hacia azúcares fermentables mediante la incubación de biomasa con el polipéptido de la invención, con un fragmento o variante del mismo o con la composición celulolítica de la invención.

- 5 Así pues, otro aspecto particular se relaciona con un procedimiento para producir un bioproducto a partir de biomasa que comprende:
- a) producir azúcares fermentables de acuerdo con el procedimiento anteriormente descrito,
 - b) llevar a cabo una transformación química de los azúcares obtenidos en la etapa
 - 10 a) para producir al menos un bioproducto,
 - c) Recuperar el al menos un bioproducto obtenido en la etapa b).

La transformación química de los azúcares puede ser un proceso termoquímico para producir bioproductos, como por ejemplo someterlos a un proceso de isomerización y posterior deshidratación e hidrogenación.

15

No obstante, en una realización particular y preferida de la invención dicha transformación comprende la fermentación microbiana de los azúcares con microorganismos fermentantes.

Tanto en el procedimiento para producir azúcares fermentables, como en el procedimiento para producir bioproductos, antes de llevar a cabo la etapa de incubación de la biomasa con el polipéptido de la invención, con un fragmento o una variante del mismo o con la composición celulolítica de la invención se puede, opcionalmente, llevar a cabo un pretratamiento de la biomasa celulósica con el fin de facilitar el proceso de digestión posterior. El pretratamiento puede consistir en la ruptura de los componentes de la pared celular de las células vegetales de la biomasa (Chandra et al., 2007, Substrate pretreatment: The key to effective enzymatic hydrolysis of lignocellulosics? Adv. Biochem. Engin. /Biotechnol. 108: 67-93; Galbe and Zacchi, 2007, Pretreatment of lignocellulosic materials for efficient bioethanol production, Adv. Biochem. Engin. /Biotechnol. 108: 41-65; Hendriks and Zeeman, 2009, Pretreatments to enhance the digestibility of lignocellulosic biomass, Bioresource Technol. 100: 10-18; Mosier et al., 2005, Features of promising technologies for pretreatment of lignocellulosic biomass, Bioresource Technol. 96: 673-686; Taherzadeh and Karimi, 2008, Pretreatment of lignocellulosic wastes to improve ethanol and biogas production: A review, Int. J. of Mol. Sci. 9: 1621-1651 ; Yang and Wyman, 2008, Pretreatment: the key to unlocking low-cost cellulosic ethanol, Biofuels Bioproducts and Biorefining-Biofpr. 2: 26-40).

20

25

30

La biomasa celulósica puede asimismo ser sometida también a una reducción del tamaño de partícula, a baños y lavados y/o condicionamiento antes del pretratamiento opcional.

Los pretratamientos convencionales de la biomasa celulósica incluyen pero no se limitan a pretratamiento con calor, pretratamiento con ácido diluido, pretratamiento con agua caliente, 5 pretratamiento alcalino, pretratamiento con cal, oxidación en mojado, explosión en mojado AFEX, pretratamiento organosolv y pretratamiento biológico. Otro pretratamientos adicionales incluyen percolación amónica, ultrasonidos, electroporación, microondas, CO₂ supercrítico, H₂O supercrítico, ozono, líquido iónico o pretratamiento por irradiación gamma.

La etapa de incubación de la biomasa celulósica (pretratada o no) con el polipéptido de la 10 invención, con un fragmento funcional o variante del mismo o con la composición celulolítica de la invención comprende la hidrólisis del material celulósico y/o hemicelulósico hacia azúcares fermentables como glucosa, celobiosa, xilosa, xilulosa, arabinosa, manosa, galactosa y/o oligosacáridos solubles.

La hidrólisis enzimática se lleva a cabo preferiblemente en un medio acuoso apropiado y en 15 condiciones que puede determinar fácilmente un experto en la materia.

En una realización preferida la hidrólisis se lleva a cabo a un pH entre 5 y 7, preferiblemente entre 5 y 6.5 y a una temperatura de entre 30°C y 55°C preferiblemente entre 37°C y 50°C.

La hidrólisis puede hacerse en un proceso discontinuo o continuo donde la biomasa se 20 añade gradualmente a la solución de hidrólisis que comprende el polipéptido de la invención, un fragmento funcional o una variante del mismo o la composición celulolítica de la invención.

La producción de azúcares fermentables se lleva a cabo generalmente en tanques o 25 reactores con agitación que controlan las condiciones de pH, temperatura y la mezcla de los componentes de reacción. Los tiempos de reacción pueden ser determinados por un experto en la materia aunque pueden durar hasta 200 horas, pero preferiblemente la etapa de producción de los azúcares fermentables se lleva a cabo entre 12 y 120 horas.

La etapa posterior, esto es la etapa b) comprende llevar a cabo una transformación química 30 de los azúcares obtenidos en la etapa anterior siendo esta transformación química mediada de una manera preferentemente con al menos un microorganismo fermentante. Para ello la corriente resultante de la etapa a) (hidrólisis enzimática) se lleva directamente al tanque donde se dará la fermentación bajo condiciones controladas.

La elección del o de los microorganismos fermentantes dependerá del tipo de bioproducto que se desee producir. En una realización particular de la invención el al menos un microorganismo fermentante se selecciona entre bacterias como *Bacillus thermoglucosidaisus*, *Clostridium butyricum*, *Corynebacterium glutamicum*, *Enterobacter*
 5 *aerogenes*, *Escherichia coli*, *Geobacillus thermoglucosidasius*, *Klebsiella oxytoca*, *Lactobacillus* sp. *Leunoscoc mesenteroides*, *Thermoanaerobacter BG1L1*, *Thermoanaerobacter ethanolicus*, *Thermoanaerobacter mathranii*, *Thermoanaerobacter thermosaccharolyticum*, *Zymobacter palmae*, *Zymomonas mobilis* u hongos como *Candida arabinofermentans*, *Candida boidinii*, *Candida diddensis*, *Candida fermentans*,
 10 *Chrysosporium lucknowense*, *Candida pastoris*, *Candida shehatae*, *Candida sonorensis*, *Candida tropicalis*, *Hansenula anómala*, *Hansenula polymorpha*, *Kluyveromyces fragilis*, *Kluyveromyces marxianus*, *Pichia pastoris*, *Pichia stipitis*, *Saccharomyces cerevisiae*, *Saccharomyces bulderi*, *Saccharomyces barnetti*, *Saccharomyces exiguus*, *Saccharomyces diastaticus*, *Saccharomyces uvarum* o *Schizosaccharomyces pombe* o una combinación de
 15 alguno de estos microorganismos.

Las condiciones para llevar a cabo la fermentación dependerán principalmente del tipo de microorganismo fermentante. Normalmente, las condiciones se adaptarán a las condiciones óptimas en las que el microorganismo en cuestión lleve a cabo el proceso de fermentación.
 20 Aunque no hay unas condiciones fijas, típicamente el tiempo de fermentación varía entre 24 y 96 horas, la temperatura entre 26°C y 60°C, preferentemente entre 32°C y 50°C y el pH entre 3 y 8, preferentemente entre 4 y 5, entre 4 y 6 o entre 4 y 7.

La cantidad de inóculo del microorganismo fermentante también dependerá igualmente del
 25 tipo de microorganismo aunque típicamente el microorganismo fermentante se puede aplicar en cantidades de entre 10^5 y 10^{12} células/ml de caldo de cultivo.

En función de los microorganismos utilizados se obtendrán diferentes bioproductos ya que cada microorganismo tiene un metabolismo característico y propio. En este sentido un
 30 experto en la materia seleccionará el microorganismo más conveniente en función del bioproducto que desee obtener. Sin pretender ser una lista limitativa algunos de los bioproductos que se pueden obtener mediante el procedimiento aquí descrito incluyen alcoholes como etanol, metanol, butanol, hexanol, octanol, decanol, dodecanol, 1,3-
 35 butanediol(1,3-diol), 1-alcohol; ácidos orgánicos como ácido cítrico, ácido acético, ácido itacónico, ácido láctico, ácido glutámico, ácido succínico, beta-cetoácido, beta-cetoalcohol,

beta-hidroxiácido; cetonas como la acetona; gases como hidrógeno o dióxido de carbono; hidrocarburos como alcanos, alquenos o alquinos; sustancias nitrogenadas como aminas, amida, nitrocompuestos o nitrilos; haluros; aminoácidos como ácido glutámico; antibióticos como penicilina o tetraciclinas; vitamina como riboflavina, vitamina B12 o betacaroteno; 5 ácidos grasos como ácido dodecanoico, ácidos grasos trans- Δ^2 o ácido palmítico; y otros productos como etileno, glicerol, 1,3-propano-diol, betalactano, cefalosporinas, ácidos grasos trans o furano.

En una realización preferida el bioproducto es un alcohol preferiblemente etanol.

10

La etapa c) del procedimiento de producción de bioproductos comprende la recuperación del mismo. El o los productos de fermentación pueden recuperarse del medio de fermentación mediante cualquier método conocido incluyendo pero sin limitarse a cromatografía, procesos electroforéticos, solubilidad diferencial, destilación o extracción. Por ejemplo, el alcohol se 15 separa del caldo de fermentación por métodos convencionales de destilación.

EJEMPLOS

Ejemplo 1: Construcción de la librería metagenómica e identificación de los clones de interés a partir de la misma

20

Construcción de la librería metagenómica

Se construyó una librería metagenómica a partir de bacterias del suelo en un fósido pMPO579 (Terron-Gonzalez *et al.*, 2013 Sci. Rep. 3: 1107). La librería metagenómica se 25 construyó de acuerdo con el protocolo del kit CopyControl™ Fosmid Library Production (Epicentre), con alguna modificación como se describe en Terron-Gonzalez *et al.* 2013. La librería comprende aproximadamente 7,4 Gigabases distribuidas entre aproximadamente 185,000 clones diferentes, con un tamaño medio de aproximadamente 40 kb, y se mantiene en la cepa EPI300-T1^R.

30

Identificación de clones que confieren actividad celulasa

Se transfirió la librería metagenómica por cruce triparental (Figurski and Helinski, 1979, Proc. Natl. Acad. Sci. USA 76:1648-1652) a la cepa MPO554, con DH5 α /pRK2013 como cepa *helper*. En presencia de salicilato, la cepa MPO554 produce NahR (el activador del 35 promotor *psal*) y la proteína N antiterminante del fago lambda, permitiendo la expresión del promotor *psal* localizado en el fósido y la antiterminación transcripcional (Terron-Gonzalez

et al., 2013 Sci. Rep. 3: 1107). Los cruces conjugativos se llevaron a cabo en LB-agar sin selección por antibiótico durante toda la noche a 37°C. Las mezclas de cruces se sembraron en placas con LB-agar suplementadas con los antibióticos necesarios para la selección de transconjugantes, arabinosa para incrementar el número de copias del fósido, salicilato para incrementar los niveles de transcripción y 1% carboximetilcelulosa (CMC) como sustrato para el rastreo. Las placas se incubaron a 30°C durante 48 h.

Se llevó a cabo un rastreo típico de celulasas usando platos de agar con CMC con tinción Rojo Congo (Teather and Wood, 1982 Appl. And Environ. Microbiol. 43:777-780). Este método tiene el problema de que las colonias se despegan debido a que las placas se inundan durante la tinción, haciendo difícil la identificación de clones positivos. Para solventar este problema, se hizo una modificación del método que consistió en la inmovilización de las colonias con una capa de cobertura de agar antes de la tinción. Después, cada plato se inundó con una solución de Rojo Congo al 1% durante 30 minutos y después NaCl 1M durante 30 minutos. Los clones positivos mostraban un halo claro alrededor de la colonia debido a la falta de tinción por hidrólisis de la CMC por las celulasas secretadas. La Figura 9 muestra una de las placas con un clon con actividad celulasa.

Se hizo el screening de alrededor de 550,000 clones, de tal forma que la probabilidad de que se analizaran todos los clones de la librería fue de 95%. Todos los clones positivos se analizaron mediante digestión con enzimas de restricción, detectándose un total de 8 fósidos diferentes con actividad celulasa (8 clones con distinto patrón de restricción).

Identificación de los genes de interés

Cada fósido se secuenció por completo y las secuencias se compararon con aquellas de bases de datos usando los programas Blastx y Blastn (Altschul *et al.*, 1997 Nucleic Acid Res. 25:3389-3402). Se identificaron ORFs potencialmente involucrados en la actividad celulasa.

Ejemplo 2: Construcción del vector y expresión de Cel98 en *E.coli* DH5 α

La secuencia del CLON98 identificado en el rastreo fue analizada mediante el programa ORF Finder para buscar pautas de lectura abiertas y posibles genes que se transcriban y se detectó así "*cel98*", una secuencia de ADN de 1787 pb (SEQ ID NO 6) que es la secuencia polinucleotídica que codifica el polipéptido de la invención que se denominó Cel98 y que posee la secuencia SEQ ID NO 1.

Se procedió al estudio y análisis del gen *cel98* mediante el programa *motifscan* (http://myhits.isb-sib.ch/cgi-bin/motif_scan). El fragmento identificado muestra en su extremo amino-terminal, de la posición 1-32, una señal de translocación Twin-arginine (tat). También se observa el dominio celulasa (perteneciente a la familia 5) en la posición 41-378. En las posiciones 443-570 y 644-780 se encuentran dos dominios de unión a carbohidratos de tipo 2 (CBM2) y en la posición 474-630 un dominio de unión a carbohidratos de tipo 6 (CBM6) (ver figura 3).

El vector elegido para subclonar el gen de *cel98* fue el pET23b. Dicho vector contiene un gen de resistencia a ampicilina, el promotor T7, el inicio de transcripción de T7, una secuencia His-tag (en el extremo C-t) y el polilinker (multicloning site). Incluye así también la secuencia de terminación T7 y los orígenes pBR322 y f1.

Para subclonar el fragmento de interés se siguió el siguiente procedimiento:

15 Diseño de los cebadores para amplificar el fragmento de interés (*cel98*)

Se procedió a diseñar cebadores empleando la secuencia de dicho gen y se decidió insertar en los cebadores señales de reconocimiento de dos enzimas de restricción diferentes para que la inserción del gen fuera dirigida (NdeI y XhoI) ya que estas no cortaban dentro de la secuencia del gen y tan sólo cortaban una vez cada una en el vector, de forma que este podía abrirse para insertar *cel98*.

Los cebadores empleados fueron los siguientes:

- Directo: CAATCATATGAGGACCCACATGAAAACG (Tm=62°C) (SEQ ID NO 11)
- Reverso: GAATCTCGAGCTTGGTGACGGTGAAAGCC (Tm=60°C) (SEQ ID NO 12)

25 Preparación de los cultivos de la cepa que contiene el vector pMPO135 y el fragmento de interés

Se prepararon cultivos en medio líquido Luria-Bertani (LB) suplementados con ampicilina (100 µg/mL). Para el cultivo del vector, la cepa DH5α de *E. coli* (se seleccionó una colonia de la placa) se inoculó en 5 ml de LB suplementado con ampicilina (el vector tiene resistencia a este antibiótico) y para el cultivo del inserto se seleccionó una colonia de UPO98 (*E. coli* MP0554 con el inserto *cel98*) en 3 ml de LB con cloranfenicol (el fósforo donde está construida la metagenoteca tiene resistencia a cloranfenicol) y arabinosa para aumentar la expresión (promotor inducible por arabinosa).

35

Extracción del DNA plasmídico (vector) y cuantificación de ADN por Nano-Drop.

Para la extracción del ADN plasmídico se empleó el kit "Nucleo Spin Plasmid (Macherey-Nagel)". Para cuantificar el ADN extraído se empleó el equipo Nano-drop ND-1000. Así las cantidades de ADN obtenidas fueron de 88,5 ng/μl para el vector y 354,5 ng/μl para el fósido que contiene el gen de interés.

Digestión del vector

Se emplearon para ello 2,5 μg de plásmido. La reacción se llevó a cabo en un volumen final de 35 μl.

10 Se necesitó, en orden de adición al tubo:

1. Agua desionizada: 1 μl (hasta completar el volumen final)
2. Tampón H (Roche) (10X): 3,5 μl
3. ADN (vector purificado): 28 μl
4. Enzimas para la digestión (NdeI (Fermentas) y XhoI (Roche)): 1,25 μl de cada enzima

15 La reacción se dejó incubando a 37°C durante dos horas y media.

Gel de agarosa (0,8%) para comprobar que las bandas del vector linearizado son las correctas.

El vector empleado posee unas 4,5 Kb. Al cortar con las enzimas de restricción se liberó un fragmento de 0,9 Kb, quedando un vector de 3,6 Kb linearizado al visualizarse en gel.

20

El gel se corrió durante 20 minutos a 135V y posteriormente se tiñó durante 20 minutos en bromuro de etidio para visualizar las bandas.

PCR del fragmento de interés

25 Empleando los cebadores previamente descritos se procedió a la amplificación por PCR del fragmento de interés "ce/98" siguiendo el siguiente programa:

1°. 5 minutos a 94°C para desnaturalizar las cadenas de ADN

2°. 35 ciclos:

- 94°C durante 30 segundos para desnaturalizar.

30

- 56°C durante 30 segundos.

- 72°C durante 5 minutos y medio para que se extienda la polimerasa. Este tiempo depende del tamaño del fragmento que se deba copiar. Se debe poner 2 minutos por Kb a amplificar, y como el inserto de interés era de 2,787 Kb, se configuró para 5' 30".

35

3°. 5 minutos a 72°C

La PCR se hizo por duplicado. Se puso también un control negativo (sin DNA para amplificar, sustituyendo su volumen por agua desionizada).

Por cada reacción se pusieron 50 µl de volumen final:

- ADN molde. Se emplearon 10-40 ng de DNA. Como teníamos a UPO98 a 354,5 ng/µl, se diluyó 20 veces y de dicha dilución se tomó 1 µl de ADN y se adicionó en 19 µl de agua desionizada, y de este último stock se puso 2 µl para la PCR.
- cebador directo (10µM) (SEQ ID NO 11): 1 µl
- cebador reverso (10µM) (SEQID NO 12): 1 µl
- dNTPs (2,5 mM): 4 µl
- polimerasa (2,5 U/µl): 0,5 µl
- tampón (10X): 5 µl
- Agua desionizada: lo que falte hasta 50 µl de volumen final

Gel de agarosa para comprobación de la PCR

Se corrió un gel pequeño al 0,8% de agarosa para visualizar el producto amplificado de PCR. Una vez corrido el gel durante 20 min a 135V se tiñó 20 minutos en bromuro de etidio.

Purificación del producto de PCR por selección de banda en gel de agarosa.

Como la PCR no salió completamente limpia, la forma de purificar el fragmento de interés fue por selección de la banda correspondiente al inserto en un gel de agarosa (0,8%). El vector cortado también se cargó en el mismo gel para coger la banda correspondiente al vector linearizado de ahí.

El gel se corrió a 100V durante una hora y posteriormente se tiñó 30 minutos en bromuro de etidio. Las bandas correspondientes al vector linearizado (3,6 Kb) y al fragmento de interés (2,7 Kb) se cortaron del gel con una cuchilla y se conservaron en tubos a -20°C.

Purificación de bandas del vector e inserto

Se empleó el kit para purificación de los fragmentos de gel (vector e inserto): ILLUSTRATION OF THE GEL BAND PURIFICATION KIT (GE HEALTHCARE).

En el último paso del protocolo a seguir del kit, se eluyó el vector en 15 µl para que estuviera lo más concentrado posible ya que este era su último paso de purificación antes de la ligación. El inserto se eluyó en 19 µl.

35

Digestión extremos del inserto

Las enzimas que se emplearon para la digestión de los extremos del inserto fueron las mismas que para digerir el vector (NdeI y XhoI). El volumen total de la reacción fueron 20 µl. El buffer que se usará será el tampón H (igual que para cortar el vector) y las enzimas serán también NdeI y XhoI. En esta ocasión se tiene menos ADN que en el caso del vector y por eso se necesita mucha menos enzima de restricción. Para 20 µl de reacción se emplearon:

- Tampón H (X10): 2 µl
- NdeI (Fermentas): 0,5 µl
- 10 -XhoI (Roche): 0,5 µl
- ADN: 17 µl de los 19 µl recuperados tras la purificación de la banda de gel.

La digestión se dejó 2 horas y media a 37°C.

15 Comprobación de la cantidad de ADN del vector y del inserto antes de proceder con la reacción de ligación

Se purificó el inserto de interés tras la reacción de digestión para liberarlo de las enzimas y tampón empleando el kit de GE HEALTHCARE de nuevo (protocolo de limpieza tras una reacción enzimática). Se preparó un gel de agarosa (0,8%) y se cargó un microlitro tanto de vector como de inserto para observar que cantidad de ADN se tenía y saber qué proporción de vector y de inserto poner en la reacción de ligación. En la figura 10 se observa el gel donde se pudo comprobar las bandas con el ADN tanto del vector como del inserto.

Ligación

- 25 El volumen final serán 10 µl:
 - Vector: 1 µl
 - Inserto: 6 µl
 - Ligasa: 1 µl
 - Tampón de ligasa (5X): 2 µl

30 Se puso un control de la religación, una reacción de ligación conteniendo sólo vector (sin inserto) para ver si el vector era capaz de religarse consigo mismo. Para ello se puso la misma cantidad de vector, pero en vez de inserto, se adicionó agua desionizada en su lugar.

35 Una vez preparadas las reacciones, se incubaron a 16°C durante 16 horas.

Transformación (con células ya competentes, por choque térmico)

Las células competentes son *E. coli* DH5 α . Los 10 μ l de reacción de ligación se introducen en un tubo conteniendo 100 μ l de células ya competentes. El proceso se llevó a cabo tanto para la ligación como para la religación:

5

1) Se mantuvo en hielo el ADN + células durante 30 minutos.

3) Se dio un choque térmico a 42°C durante exactamente 45 segundos.

4) Se retransfirieron los tubos a hielo durante 5 minutos.

5) Se adicionó 1 ml de LB y se dejó incubar a 37°C durante 1 hora y 30 minutos (sin
10 adicionar antibióticos)

6) Se sembró en placa con antibiótico (ampicilina 100 μ g/ml) y se dejaron incubando a 37°C las placas hasta el día siguiente. Las células DH5 α transformadas son resistentes a ampicilina (100) por el vector que contienen.

15 Selección de los transformantes

Para seleccionar una colonia transformante del total de colonias que crecieron en la placa tras la transformación en DH5 α se hizo una PCR de comprobación de colonias. Para ello se hizo uso de dos cebadores que amplificaran un fragmento que contuviese parte del vector y parte del inserto, para asegurarnos que dicha colonia poseía el vector conteniendo éste el
20 inserto y que no era resistente a ampicilina por contener en su interior un vector religado sin inserto. Los cebadores empleados fueron el T7 forward (amplifica parte del vector) y el UPO98intrev (amplifica dentro del inserto):

T7 fw: GTAATACGACTCACTATAGGGC (Tm=56) (SEQ ID NO 13)

UPO98intrev: TCTGGCTGACCGTGTAGG (Tm=58) (SEQ ID NO 14)

25

El fragmento esperado para amplificar era de 1,3 Kb. El programa diseñado fue el siguiente:

1°. 5 minutos a 94°C para desnaturalizar las cadenas de ADN

2°. 35 ciclos:

30

- 94°C durante 30 segundos para desnaturalizar.

- 52°C durante 30 segundos.

- 72°C durante 5 minutos y medio para que se extienda la polimerasa. Este tiempo depende del tamaño del fragmento que se deba copiar. En este caso, para amplificar 1,3 Kb, se configuró para 1' 30".

35

3°. 5 minutos a 72°C

Los tubos de reacción de PCR, para cada posible colonia transformante que se analizó por PCR se compuso de (para un volumen final de 25 μ l):

- Cebador T7 forward (10 μ M): 0,5 μ l
- Cebador UPO98intrv (10 μ M): 0,5 μ l
- 5 - dNTP's (10 mM): 0,5 μ l
- Tampón (10X): 2,5 μ l
- Mg²⁺ (25 mM): 1,5 μ l
- Polimerasa: 0,2 μ l
- Agua desionizada: 19,3 μ l

10

El ADN molde se adicionó tocando la colonia que se deseaba comprobar e inoculando con esta el tubo de PCR. De esa misma colonia se tocó e hizo una raya en una placa de agar LB con ampicilina (100 μ l/mL), de forma que en caso de resultar positivo el transformante, se tuviera biomasa suficiente para inocular caldo de cultivo y poder crecer dicho transformante.

- 15 En la PCR de candidatos se puso también un control positivo, es decir, un tubo que contenía el ADN con el que se transformó a DH5 α , y en el cual se debe ver amplificación del fragmento de 1,3 Kb.

Tras finalizar la PCR, se cargaron 5 μ l de cada tubo de reacción (28 candidatos) en un gel de agarosa al 0,8% y se corrió durante 20 minutos a 135V. Posteriormente se tiñó con bromuro de etidio durante 20 minutos y se comprobó qué candidatos tenían el inserto de interés.

Se escogió dos candidatos (se denominaron candidato 1 y 6) y de la raya que previamente se había sembrado de cada uno de ellos se puso un inóculo (medio LB suplementado con ampicilina a 100 μ l/mL) y se llevó a cabo la extracción de ADN de ambos candidatos para poder realizar un análisis por visualización de bandas tras digestión con enzimas de restricción. Así, se digirieron los dos candidatos con NdeI y XhoI (enzimas previamente utilizadas en el clonaje) y también, en una reacción a parte, con BamHI (tiene una única diana, por lo que la construcción se lineariza. El corte con esta enzima permitió ver si se tenía un inserto o más de uno dentro de un vector según la longitud de la banda linearizada visualizada en gel).

Se empleó 10 μ l de la extracción de ADN de cada candidato, para un volumen final de 20 μ l:

- Agua desionizada: 7 μ l
- 35 - Tampón H (10X): 2 μ l

- ADN: 10 µl de la extracción de ADN (66 ng/ µl y 46,4 ng/ µl para los candidatos 1 y 6 respectivamente)
- NdeI y Xho/BamHII: 0,5 µl de cada una

5 Para la digestión con NdeI y XhoI se incubó 2 horas y media a 37°C y para BamHI una hora.

Tras correr en gel y visualizar las bandas correspondientes se comprobó que ambos candidatos poseían el inserto, por lo que se escogió uno de ellos y se envió a secuenciar para asegurar que no había ninguna mutación en la secuencia. Una vez se hubo
10 comprobado esto, se congeló el subclón que contenía a *ceI98* a -80°C (se denominó pMPO1380).

Ejemplo 3: Producción de *CeI98* (SEQ ID NO 1)

15 *Transformación de la cepa de E.coli NCM631/piz227 con el subclón pMPO1380*

Se preparó un inóculo de NCM631/piz227 (cepa que se quiere transformar) el día previo a la transformación y se hizo una extracción de ADN del subclón pMPO1380 (subclón que contiene el ADN plasmídico que se quiere introducir en la NCM631/piz227). El inóculo de la NCM631/piz227 se hizo en medio LB suplementado con cloranfenicol (100 µl/mL), ya que
20 dicha cepa contiene resistencia a cloranfenicol en el plásmido piz227 (dicho plásmido permite que la bacteria *E.coli* NCM631 permanezca en niveles basales de expresión, reprimiendo la misma. Cuando es adicionado isopropil-β-D-1-tiogalactopiranosido o IPTG, se activa entonces la superproducción en la bacteria).

25 Para la transformación se diluyó el inóculo de NCM631/piz227 hasta 0,1 de densidad óptica en medio LB (sin antibióticos). Se incubó a 37°C hasta DO (600 λ) de 0,35 (fase exponencial temprana). Una vez llegó a dicha densidad óptica se repartió en alícuotas de 1 mL y se dejó enfriar en hielo 5 minutos. Se centrifugó 30 segundos y se eliminó el sobrenadante. Se resuspendió suavemente en 75 µl de LB frío y se mantuvo en hielo 5 minutos. Se añadió 75
30 µl de TSS 2X y se mezcló suavemente. Se mantuvo posteriormente en hielo 5 minutos. Tras esto, se procedió a añadir el DNA (1-5 µl, de forma que se añada a una concentración de 100 ng/µL). Se produjo entonces el choque térmico, poniendo las muestras durante 45 segundos en un baño con agua a 42°C. Las muestras fueron retransferidas luego a hielo. Se añadió a cada muestra 1 mL de LB (sin antibióticos) y se incubaron a 37°C (1 hora). Tras
35 pasar ese tiempo, se sembró cada muestra en placas con agar-LB (y antibiótico adecuado

como forma de selección, ampicilina y cloranfenicol a 100 µl/mL). Se dejó incubar las placas toda la noche a 37°C.

5 Las colonias transformantes se visualizaron al día siguiente (aquellas capaces de crecer en la placa, puesto que tienen un gen de resistencia a cloranfenicol en el vector piz227 y además un gen de resistencia a ampicilina procedente del vector del clon pMPO1380 con el que se le ha transformado).

Inducción de la expresión de Cel98 en la cepa NCM631/piz227

10 Una colonia fue seleccionada y se inoculó con ella medio LB suplementado con ampicilina y cloranfenicol (100 µl/mL para cada uno de ellos).

15 Al día siguiente se diluyó el cultivo a 0,1 (600 λ) de densidad óptica (LB con antibióticos ampicilina y cloranfenicol a 100 µl/mL para cada uno de ellos) y se dejó crecer a 37°C hasta que alcanzó 0,3 (600 λ) unidades de densidad óptica. En dicho momento, el matraz con el cultivo se puso a 26°C para atemperarse y reducir la velocidad de crecimiento del cultivo antes de añadir el inductor IPTG. Aproximadamente, a los 15 minutos, se llegó a la densidad óptica de 0,35 (600 λ) y se adicionó IPTG (0,05 mM). Se dejó incubar durante 10-12 horas a 26°C.

20 Tras la incubación se recogió el sobrenadante producido centrifugando las células a 5000 rpm durante 15 minutos y filtrando el sobrenadante con filtros de 0,22µ, aplicando vacío. El sobrenadante se distribuyó en alícuotas de 5 mL que se conservaron a -80°C.

25 **Ejemplo 4: caracterización de la actividad enzimática de Cel98 (SEQ ID NO 1)**

Se procedió a caracterizar la actividad de la enzima Cel98 con características de enzima celulasa. Para ello se empleó el ensayo de rojo congo (Haft *et al.*, Appl Microbiol Biotechnol. 2012 Apr; 94(1):223-9.). El rojo congo interacciona con la carboximetil celulosa dando un color rojo, de forma que la actividad enzimática por degradación de la CMC se puede
30 visualizar por la disminución de la intensidad de dicho color rojo, mediante medidas de densidad óptica a 580 nm.

En este ensayo se incubaron alícuotas de sobrenadante (diluido 10 veces) a las cuales se adicionó como sustrato carboximetil celulosa (CMC) a una concentración final de 0,05%. Las
35 temperaturas de incubación dependieron de la característica que se deseara evaluar de la enzima. A cada tiempo deseado se tomó una muestra de 875 µl de la reacción

(sobrenadante con CMC) y se adicionó NaOH (a una concentración final de 40 mM) 8,8 µl a cada tubo para parar la reacción enzimática. Posteriormente se adicionó 25 µl de rojo congo (a una concentración final de 0,25 mg/mL) y por último 100 µl de NaCl (a concentración final de 0,5 M). Se mezcló bien con todo el volumen invirtiendo el tubo y se incubó 30 minutos a temperatura ambiente. Se registraron las densidades ópticas a 580 nm.

Ensayos de temperatura óptima

Los ensayos fueron llevados a cabo tal y como se describe previamente, incubando en cada caso el sobrenadante con la CMC a la temperatura que se desea testar. De esta forma se ensayaron 30°C-37°C-45°C-50°C y 56°C.

Los resultados de estos ensayos se pueden observar en las figuras 4 y 5.

La figura 4 demuestra que el polipéptido de SEQ ID NO 1 es activo a temperaturas que van desde 30°C a 56°C. También se puede observar en esta figura que la temperatura óptima se encuentra entre los 45 y los 50°C. Aunque la enzima todavía mantiene actividad a los 56°C se observa que esta empieza a decrecer.

La figura 5 muestra la capacidad de degradación de Cel98 a 30°C, 37°C, 45°C y 50°C. Se puede observar que entre 37°C y 50°C la enzima es capaz de degradar prácticamente el 100% de la CMC en 35 minutos (ver figuras 5 B a D). Aunque a 30°C no se consiguen resultados comparables a los de las otras temperaturas la degradación de CMC a esta temperatura es también muy importante (ver figura 5A).

Termorresistencia

La termorresistencia del enzima fue calculada tras la incubación del sobrenadante a la temperatura de 30°C, 37°C, 40°C y 42°C durante diferentes tiempos (días en su mayoría). Se tomaron muestras cada día del sobrenadante incubado y se procedió a realizar un ensayo de rojo congo con cada muestra tomada, para observar la actividad enzimática.

Los resultados se pueden observar en las figuras 6 A a D.

En la figura 6 se puede observar la degradación de CMC llevada a cabo por el polipéptido de la invención después de su incubación a 30°C durante hasta 5 días es comparable a la de la enzima control (t=0) (ver figura 6A). También se observa claramente que la enzima deja de ser activa a partir del séptimo día de incubación a 37°C (ver figura 6B). Al incubarse

a 40°C se observa que la enzima va perdiendo actividad a los 3 días y deja de ser activa por completo al 4 día (ver figura 6C). A 42°C la enzima deja de ser activa a partir de los dos días (ver figura 6D).

5 pH óptimo

Los ensayos para detectar el pH óptimo de actividad del enzima fueron llevados a cabo diluyendo en estos casos 1/10 el sobrenadante en diferentes tampones (desde pH 5 a pH 8). Se realizaron ensayos de rojo congo con cada tampón y se registraron los datos de DO a 580 nm.

10

pH 5: tampón citrato

pH 5,5-6: tampón 2-(N-morfolino) ácido etanosulfónico (MES)

pH 6,5-7: tampón piperazin-N,N'-bis(2-ácido etanosulfónico) (PIPES)

pH 7,5-8: tampón 2-Amino-2-(hidroximetil)-1,3-propanodiol (TRIZMA)

15

Los resultados se pueden observar en las figuras 7y 8.

20

La figura 7 demuestra que Cel98 trabaja mejor a un pH ligeramente ácido (pH 5) y que es capaz de mantener la mayor parte de su actividad a pH en el rango de 5 a 7. A pH 7 Cel 98 mantiene aproximadamente el 50% de su actividad, mientras que entre 5.5 y 6.5 mantiene aproximadamente el 75% de su actividad.

25

La figura 8 muestra claramente el mejor efecto de degradación de CMC de la enzima pH 5. Cel98 trabaja también bien entre 5.5 y 7, preferiblemente entre 5.5 y 6.5. A los pH de 7.5 y 8 Cel98 claramente produce un efecto degradación de la CMC muy inferior al de pH inferiores.

Ejemplo 5: producción de bioetanol a partir biomasa celulósica pretratada mediante el uso de Cel98 o una composición celulolítica que comprende Cel98

30

Se utilizó paja de maíz como sustrato para la hidrólisis enzimática. Pueden llevarse a cabo pre-tratamientos tales como hidrotermólisis controlada por pH, irradiación, tratamientos ácidos-alcalinos, solventes orgánicos, pretratamientos de filtrado con amonio, oxidación húmeda, etc. En este caso se llevó a cabo el pretratamiento con ácido caliente diluido.

35

Se llevó a cabo una reacción de hidrólisis con un cóctel enzimático control durante 48 horas. Se partió de una concentración de polisacárido preferiblemente superior al 50% del peso seco del material que lo contiene (preferiblemente, pero no limitado a 25-150 mg/mL) y un

cóctel enzimático control con un contenido de Cel98 que puede variar entre 0,001-1,0 mg/mL, 0,01-0,1 mg/mL ó 0,05-0,5 mg/mL testada junto con otras enzimas en un cóctel enzimático (Bgl, Cbh1, Cbh2, Eg).

- 5 Esta fase de hidrólisis inicial se llevó a cabo en matraces de 2 L, en medio buffer fosfato-citrato (50 mM) a pH 5 durante 48 horas a 40°C (preferiblemente entre 37-42°C).

Se estudió de este modo el efecto de Cel 98 y se pudo apreciar que la adición al cóctel mínimo de la enzima Cel 98 analizada supuso un incremento en la capacidad de sacarificación con respecto al control de un 15-30%.

10

La glucosa producida se determinó empleando el kit "D-Glucose (GOPOD Format) Assay Kit" de Megazyme.

15

20

25

REIVINDICACIONES

1. Un polipéptido aislado con actividad celulasa que comprende la secuencia SEQ ID NO1 o que posee un grado de homología o identidad de al menos un 80% con la SEQ ID NO1 o que está codificado por un polinucleótido que hibrida bajo condiciones de baja, 5 media, media-alta, alta o muy alta astringencia con la secuencia SEQ ID NO 6 que codifica el polipéptido maduro.
2. Un polipéptido de acuerdo con la reivindicación 1 que posee un grado de homología o identidad de al menos un 85%, preferiblemente de al menos un 90%, más preferiblemente 10 de al menos un 95% y aún más preferiblemente de al menos un 99%.
3. Un polipéptido de acuerdo con cualquiera de las reivindicaciones anteriores que es activo a temperaturas de entre 30 y 65°C.
- 15 4. Un fragmento funcional o una variante de un polipéptido de acuerdo con cualquiera de las reivindicaciones anteriores.
5. Un fragmento de acuerdo con la reivindicación 4 que comprende al menos los aminoácidos 41 a 378 de la SEQ ID NO 1.
- 20 6. Un fragmento de acuerdo con la reivindicación 4 o 5 seleccionado entre SEQ ID NO 3, SEQ ID NO 4 o SEQ ID NO 5.
7. Una secuencia polinucleotídica que codifica un polipéptido de acuerdo con cualquiera 25 de las reivindicaciones 1 a 3 o un fragmento o una variante de acuerdo con cualquiera de las reivindicaciones 4 a 6.
8. Una secuencia polinucleotídica de acuerdo con la reivindicación 7 cuya secuencia se selecciona de entre SEQ ID NO 6, SEQ ID NO 8, SEQ ID NO 9 o SEQ ID NO 10.
- 30 9. Una construcción génica que comprende una secuencia polinucleotídica de acuerdo con la reivindicación 7 u 8 operativamente unida a una o más secuencias de control que dirigen la expresión de polipéptidos en sistemas de expresión.
- 35 10. Un vector de expresión que comprende una construcción génica de acuerdo con la reivindicación 9.

11. Una célula huésped que comprende una construcción génica de acuerdo con la reivindicación 9 o un vector de acuerdo con la reivindicación 10 y que expresa el polipéptido de acuerdo con cualquiera de las reivindicaciones 1 a 3 o un fragmento o una variante de acuerdo con cualquiera de las reivindicaciones 4-6.

5

12. Una célula huésped de acuerdo con la reivindicación 11 donde la célula es un organismo unicelular procariota o eucariota.

13. Una célula huésped de acuerdo con cualquiera de las reivindicaciones 11 o 12 donde
10 la célula procariota se selecciona del grupo de bacterias de los géneros *Bacillus*,
Clostridium, *Esherichia*, *Klebsiella*, *Pseudomonas*, *Streptomyces*, *Thermoanaerobacterium* o
Zymomonas y la célula eucariota se selecciona del grupo de hongos de las especies
Aspergillus awamori, *Aspergillus foetidus*, *Aspergillus fumigatus*, *Aspergillus japonicus*,
Aspergillus nidulans, *Aspergillus niger*, *Aspergillus oryzae*, *Bjerkandera adusta*, *Ceriporiopsis*
15 *aneirina*, *Ceriporiopsis caregiea*, *Ceriporiopsis rivulosa*, *Ceriporiopsis sufrufa*, *Ceriporiopsis*
subvermispora, *Coprinus cinereus*, *Coriolus hirsutus*, *Fusarium bactridioides*, *Fusarium*
cereales, *Fusarium culmorum*, *Fusarium graminearum*, *Fusarium graminum*, *Fusarium*
heterosporum, *Fusarium negundi*, *Fusarium oxysporum*, *Fusarium reticulatum*, *Fusarium*
roseum, *Fusarium sambucinum*, *Fusarium sarcochroum*, *Fusarium sporotrichioides*,
20 *Fusarium sulphureum*, *Fusarium torulosum*, *Fusarium trichothecioides*, *Fusarium*
venenatum, *Humicola lanuginosa*, *Mucor miehei*, *Myceliophthora thermophila*, *Neurospora*
crassa, *Penicillium purpurogenum*, *Phanerochaete chrysosporium*, *Phlebia radiata*, *Pleurotus*
eryngii, *Saccharomyces calenbergensis*, *Saccharomyces cerevisiae*, *Saccharomyces*
diastaticus, *Saccharomyces douglasii*, *Saccharomyces kluyveri*, *Saccharomyces norvensis*,
25 *Saccharomyces oviformis*, *Thielavia terrestres*, *Trametes villosa*, *Trametes versicolor*,
Trichoderma harzianum, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma*
reseii, *Trichoderma viride* o *Yarrowia lipolytica*.

14. Un método para producir un polipéptido con actividad celulolítica que comprende:
30 a) Cultivar una célula huésped de acuerdo con cualquiera de las reivindicaciones 11
a 13 en las condiciones necesarias para la producción del polipéptido,
b) Recuperar el polipéptido producido.

15. Una composición celulolítica que comprende al menos un polipéptido de acuerdo con
35 cualquiera de las reivindicaciones 1 a 3 y/o al menos un fragmento o una variante de

acuerdo con cualquiera de las reivindicaciones 4 a 6 y, opcionalmente, al menos un elemento que favorezca o facilite la degradación de biomasa.

5 16. Una composición de acuerdo con la reivindicación 15 que comprende al menos un elemento adicional que se selecciona de entre una o más endoglucanasas, exoglucanasas, celobiohidrolasas, beta-glucosidasas, glicosil-hidrolasas, hemicelulasas, xylanases, enzimas ligninolíticas o mezclas de las mismas.

10 17. Uso del polipéptido de acuerdo con cualquiera de las reivindicaciones 1 a 3, del fragmento o de la variante de acuerdo con cualquiera de las reivindicaciones 4 a 6 o de la composición de acuerdo con cualquiera de las reivindicaciones 15 a 16 para la degradación de biomasa.

15 18. Uso de acuerdo con la reivindicación 17 donde la degradación de biomasa se lleva a cabo en un procedimiento para la producción de bioproductos.

20 19. Un procedimiento para producir azúcares fermentables que comprende incubar biomasa con el polipéptido de acuerdo con cualquiera de las reivindicaciones 1 a 3, con el fragmento o una variante de acuerdo con cualquiera de las reivindicaciones 4 a 6 o con la composición de acuerdo con cualquiera de las reivindicaciones 15 a 16.

25 20. Un procedimiento de acuerdo con la reivindicación 19 que comprende recuperar los azúcares obtenidos después de la incubación con el polipéptido de acuerdo con cualquiera de las reivindicaciones 1 a 3, con el fragmento o una variante de acuerdo con cualquiera de las reivindicaciones 4 a 6 o con la composición de acuerdo con cualquiera de las reivindicaciones 15 a 16.

30 21. Un procedimiento para producir un bioproducto a partir de biomasa que comprende:
a) producir azúcares fermentables de acuerdo con el procedimiento de la reivindicación 19,
b) llevar a cabo una transformación química de los azúcares obtenidos en la etapa a) para producir al menos un bioproducto,
c) Recuperar el al menos un bioproducto obtenido en la etapa b).

35 22. Un procedimiento de acuerdo con la reivindicación 21 donde la transformación de los azúcares comprende la fermentación con al menos un microorganismo fermentante.

23. Un procedimiento de acuerdo con la reivindicación 22 donde el al menos un microorganismo fermentante se selecciona entre bacterias como *Bacillus thermoglucosidaisus*, *Clostridium butyricum*, *Corynebacterium glutamicum*, *Enterobacter aerogenes*, *Escherichia coli*, *Geobacillus themoglucosidasius*, *Klebsiella oxytoca*,
 5 *Lactobacillus sp.* *Leunoscoc mesenteroides*, *Thermoanaerobacter BG1L1*, *Thermoanaerobacter ethanolicus*, *Thermoanaerobacter mathranii*, *Thermoanaerobacter thermosaccharolyticum*, *Zymobacter palmae*, *Zymomonas mobilis* u hongos como *Candida arabinofementans*, *Candida boidinii*, *Candida diddensis*, *Candida fermentans*, *Chrysosporium lucknowense*, *Candida pastoris*, *Candida shehatae*, *Candida sonorensis*,
 10 *Candida tropicalis*, *Hansenula anómala*, *Hansenula polymorpha*, *Kluyveromyces fragilis*, *Kluyveromyces marxianus*, *Pichia pastoris*, *Pichia stipitis*, *Saccharomyces cerevisiae*, *Saccharomyces bulderi*, *Saccharomyces barnetti*, *Saccharomyces exiguus*, *Saccharomyces diastaticus*, *Saccharomyces uvarum* o *Schizosaccharomyces pombe* o una combinación de alguno de estos microorganismos.

15

24. Un procedimiento de acuerdo con cualquiera de las reivindicaciones 21 a 23 donde el bioproducto obtenido se selecciona entre alcoholes como etanol, metanol, butanol, hexanol, octanol, decanol, dodecanol, 1,3-butanediol(1,3-diol), 1-alcohol; ácidos orgánicos como ácido cítrico, ácido acético, ácido itacónico, ácido láctico, ácido glutámico, ácido succínico,
 20 beta-cetoácido, beta-cetoalcohol, beta-hidroxiácido; cetonas como la acetona; gases como hidrógeno o dióxido de carbono; hidrocarburos como alcanos, alquenos o alquinos; sustancias nitrogenadas como aminas, amida, nitrocompuestos o nitrilos; haluros; aminoácidos como ácido glutámico; antibióticos como penicilina o tetraciclinas; vitamina como riboflavina, vitamina B12 o betacaroteno; ácidos grasos como ácido dodecanoico,
 25 ácidos grasos trans- Δ^2 o ácido palmítico; y otros productos como etileno, glicerol, 1,3-propano-diol, betalactano, cefalosporinas, ácidos grasos trans o furano.

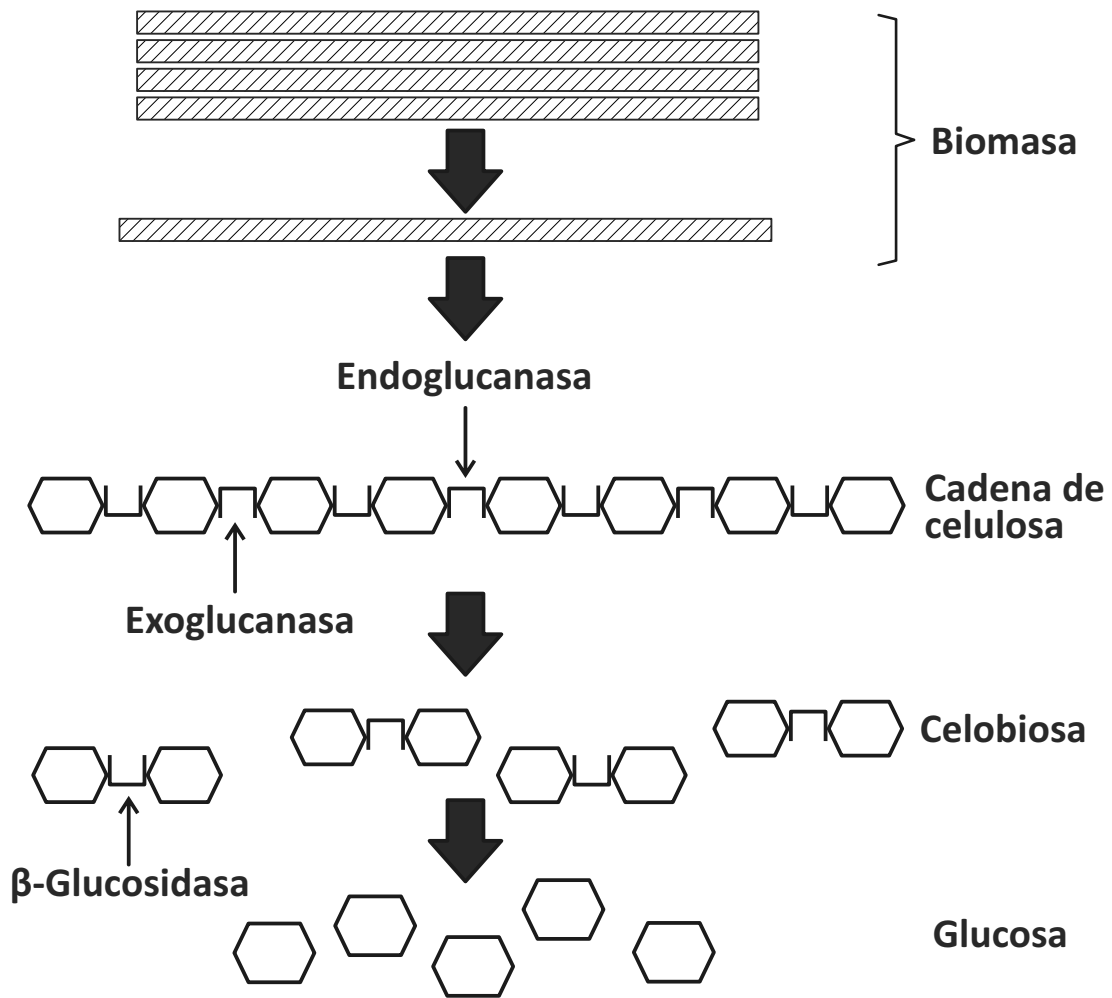


FIG. 1

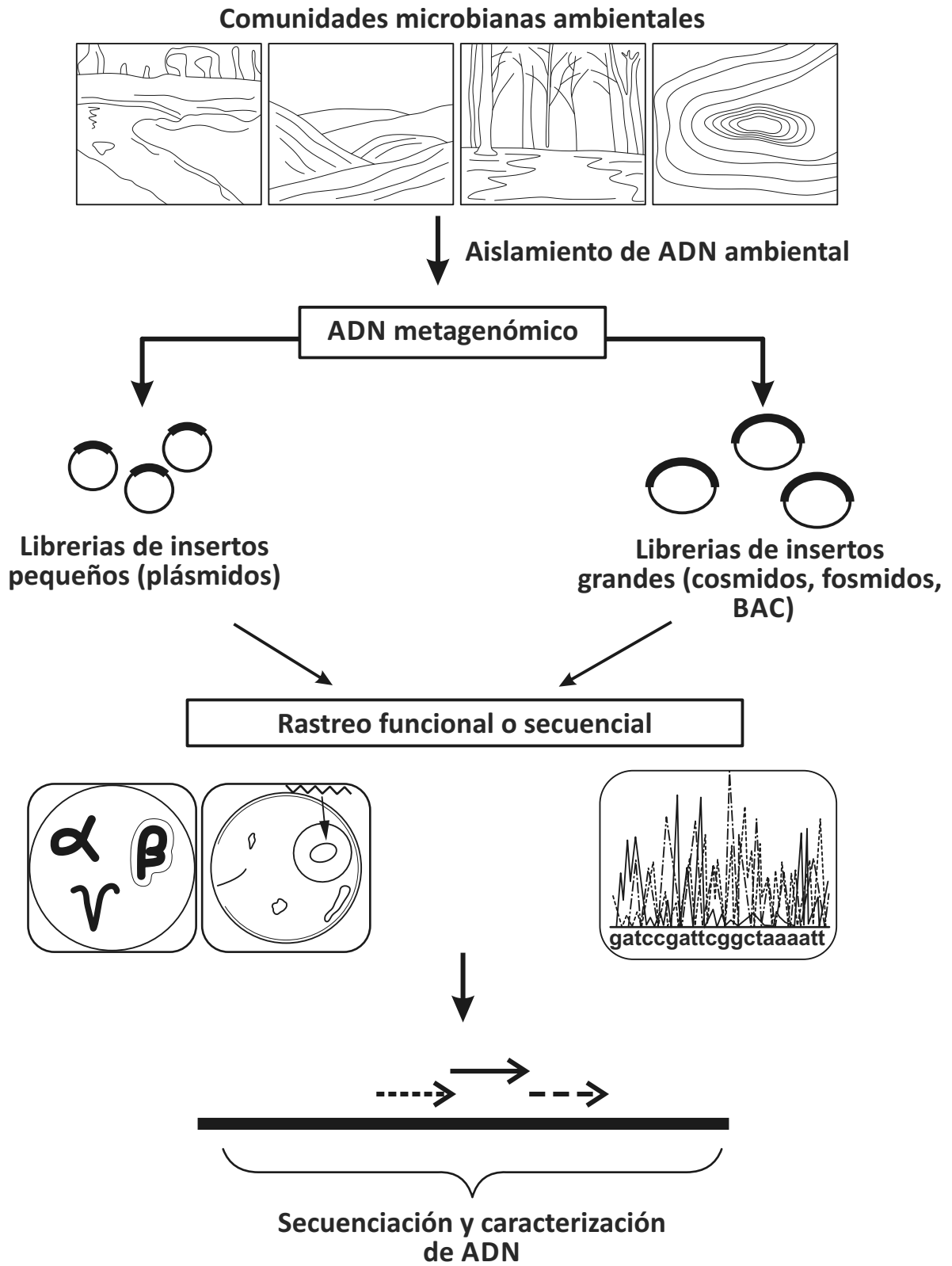


FIG. 2

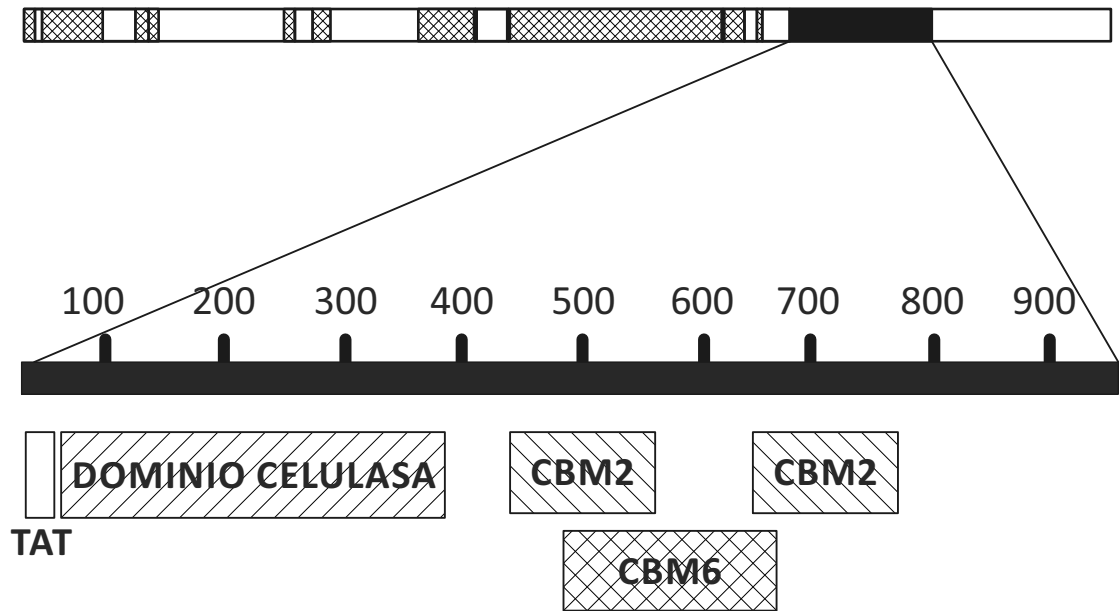


FIG. 3

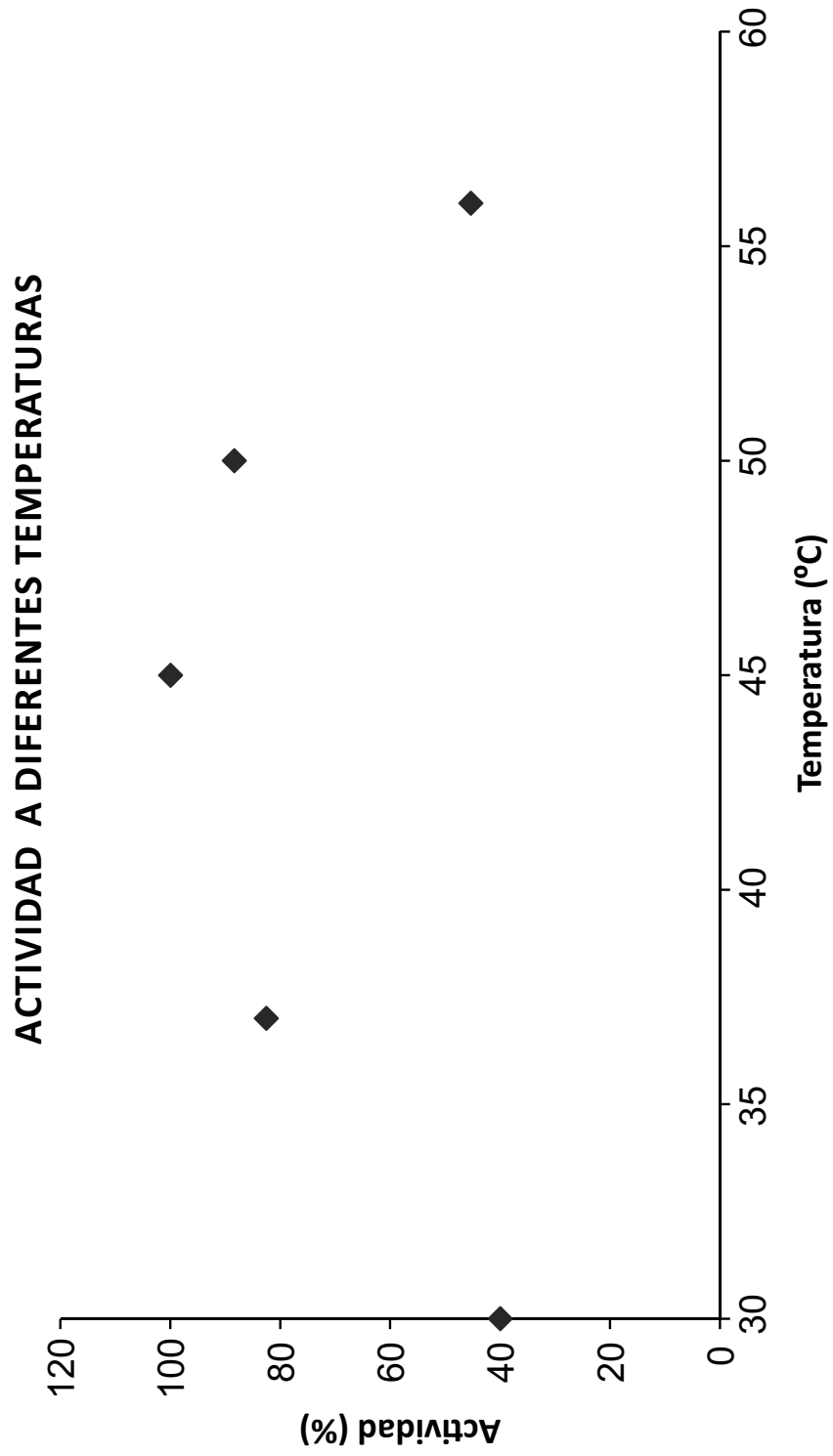


FIG. 4

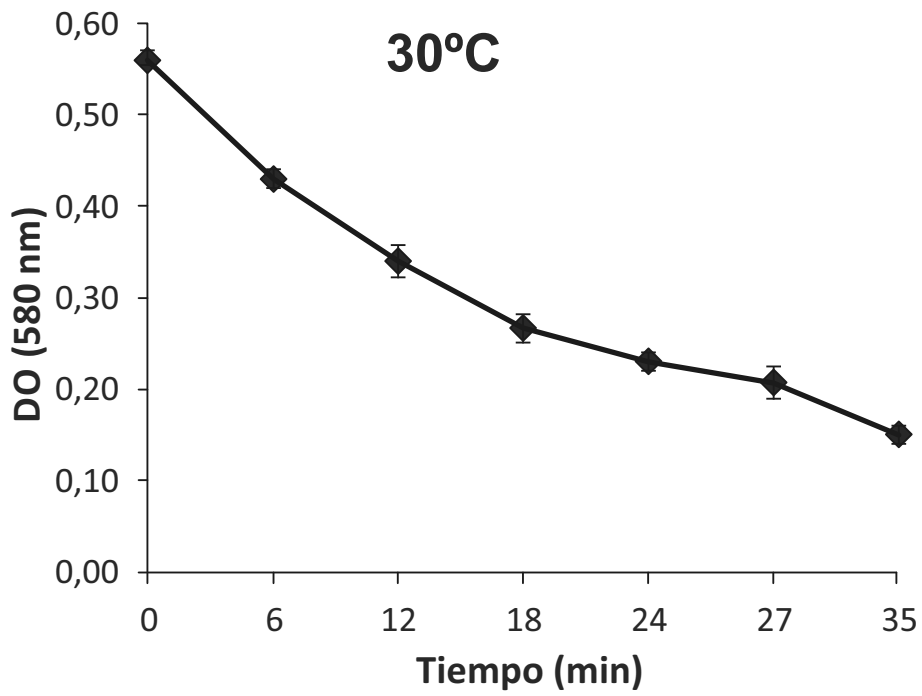


FIG. 5a

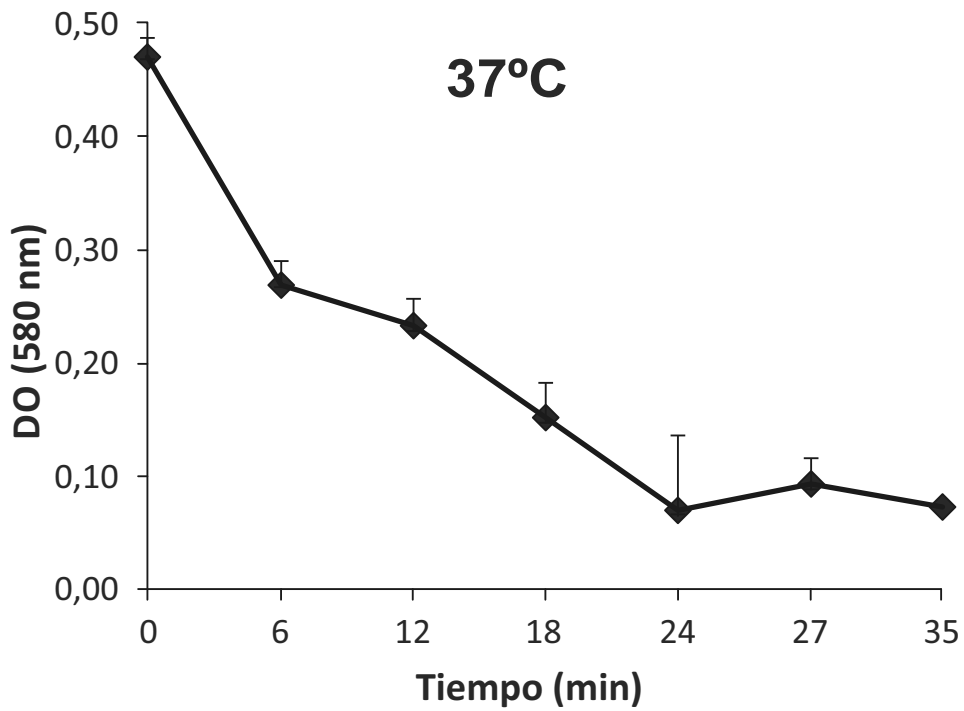


FIG. 5b

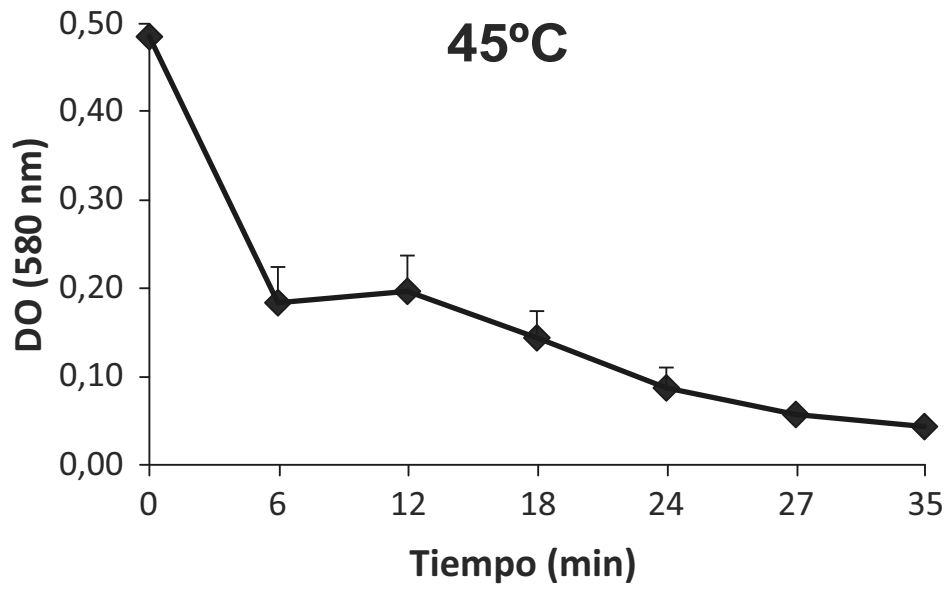


FIG. 5c

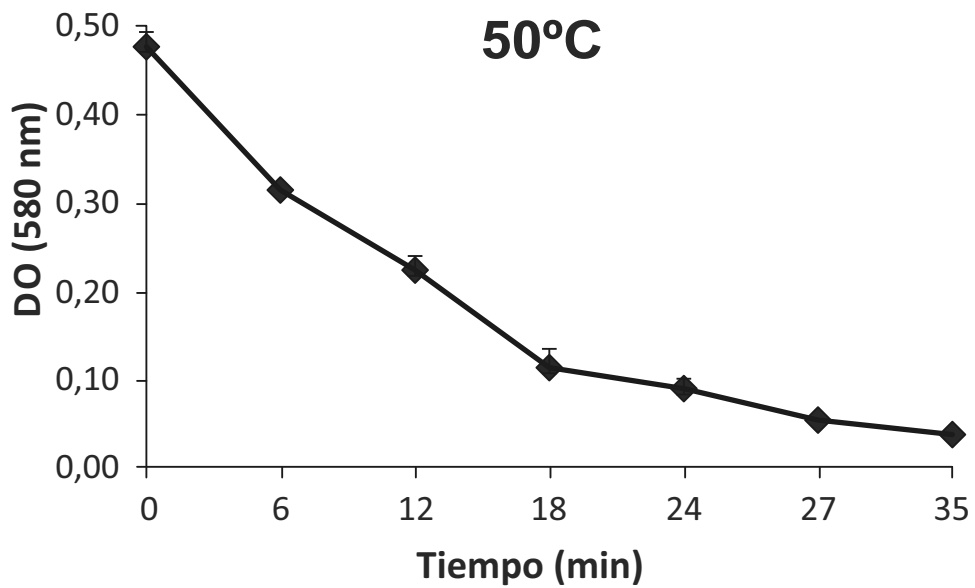


FIG. 5d

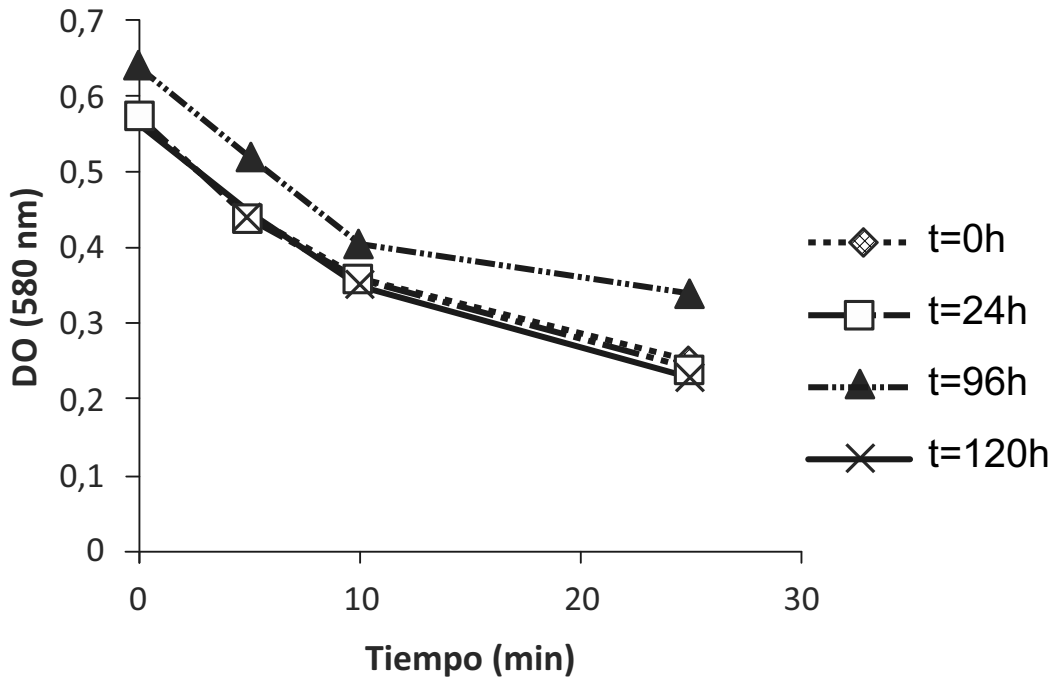


FIG. 6a

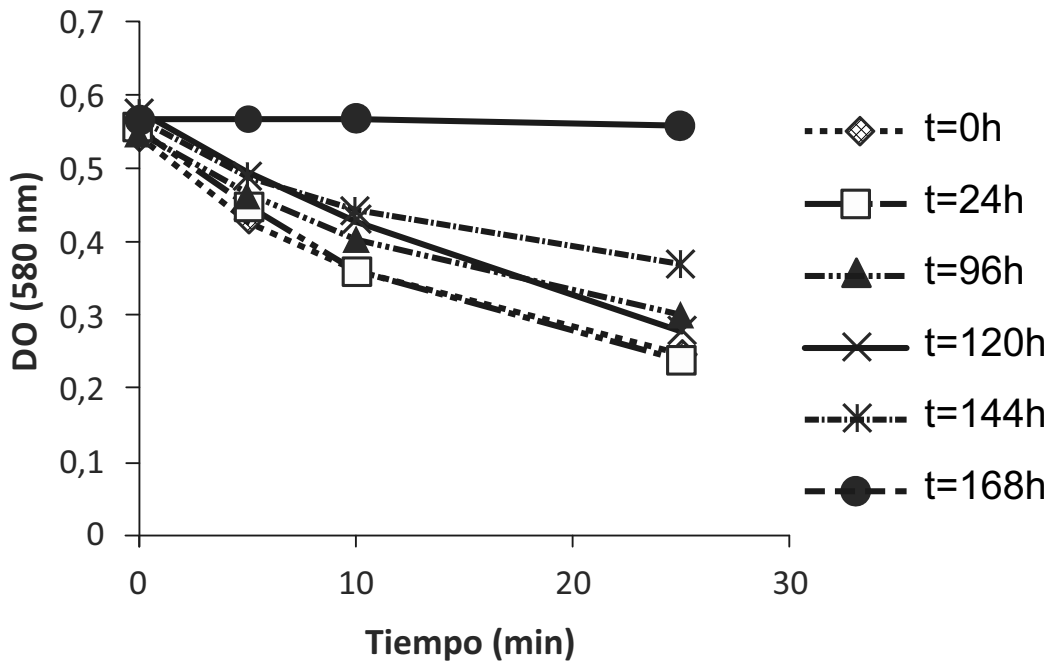


FIG. 6b

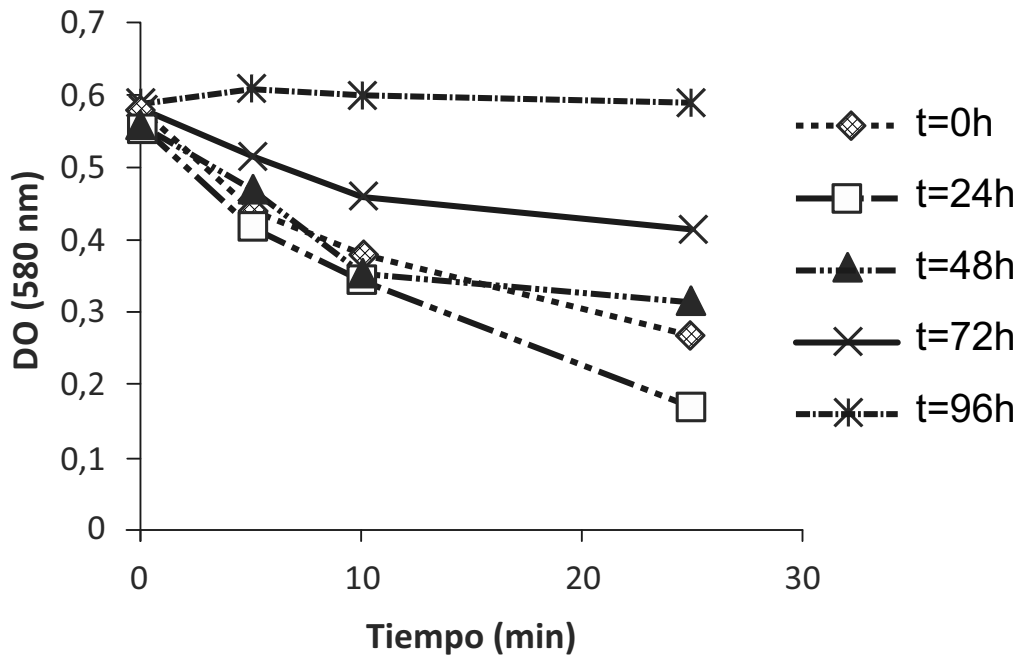


FIG. 6c

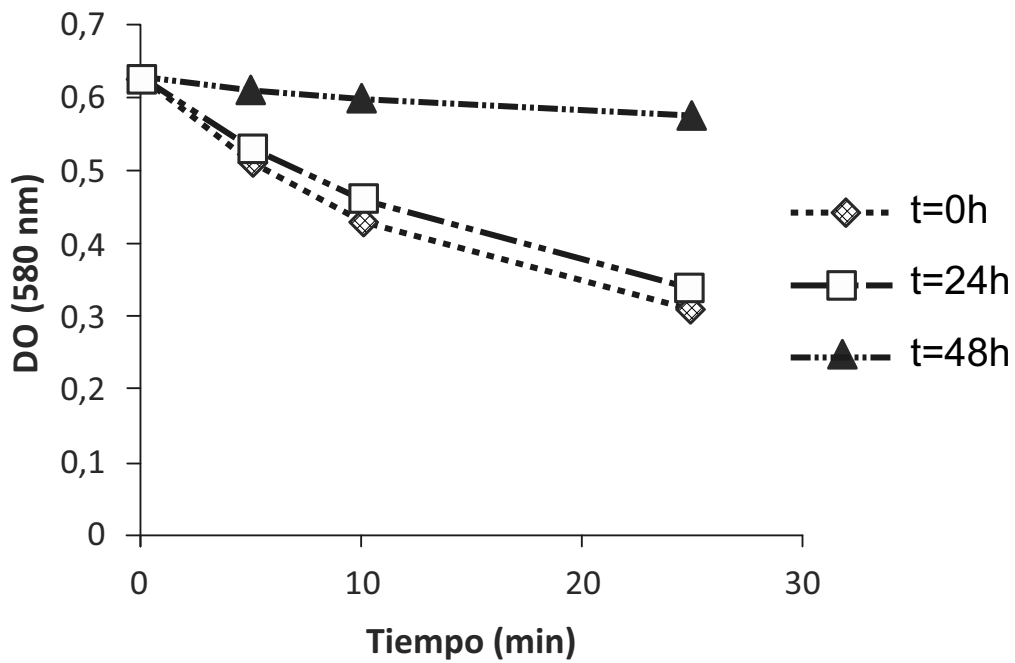


FIG. 6d

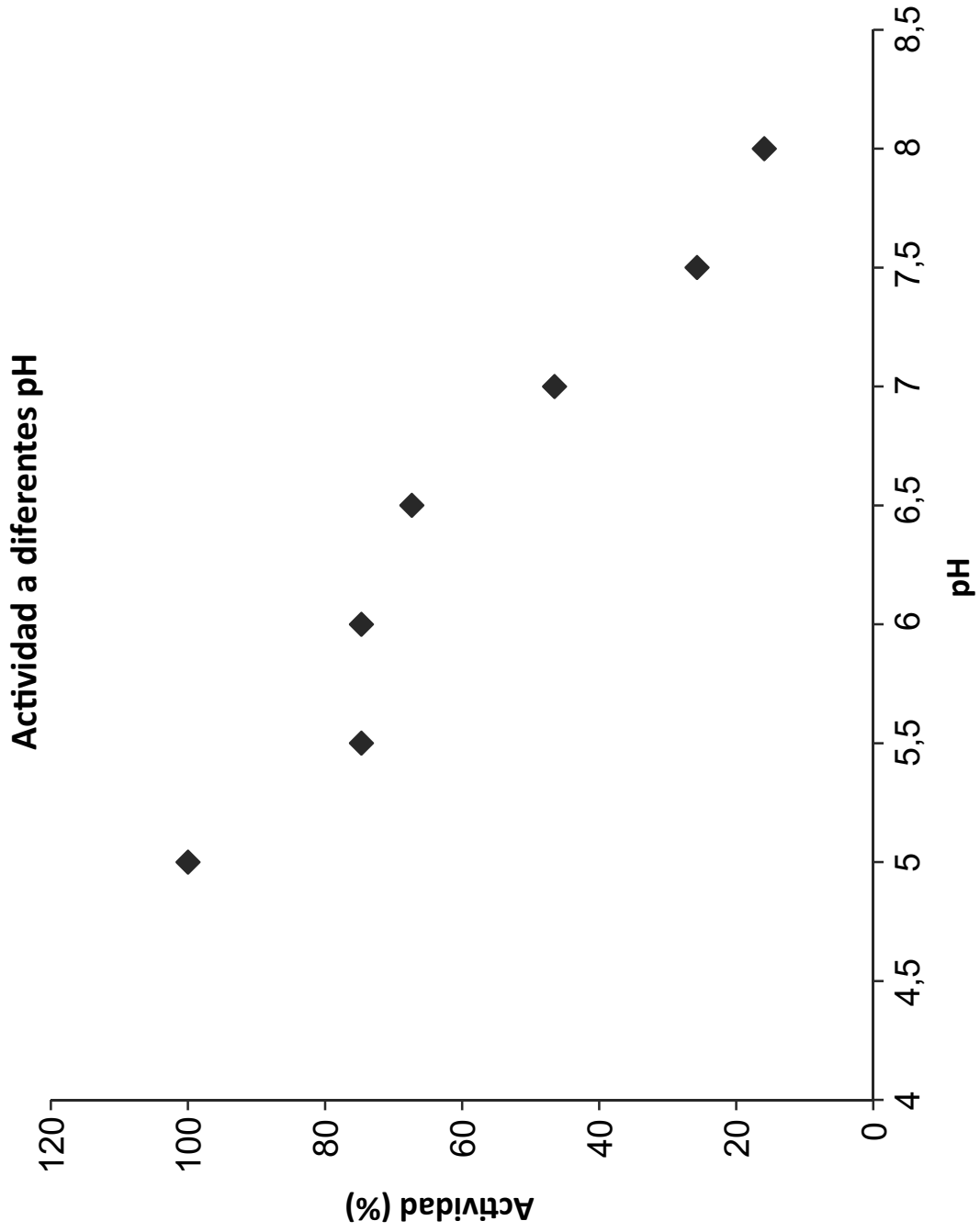


FIG. 7

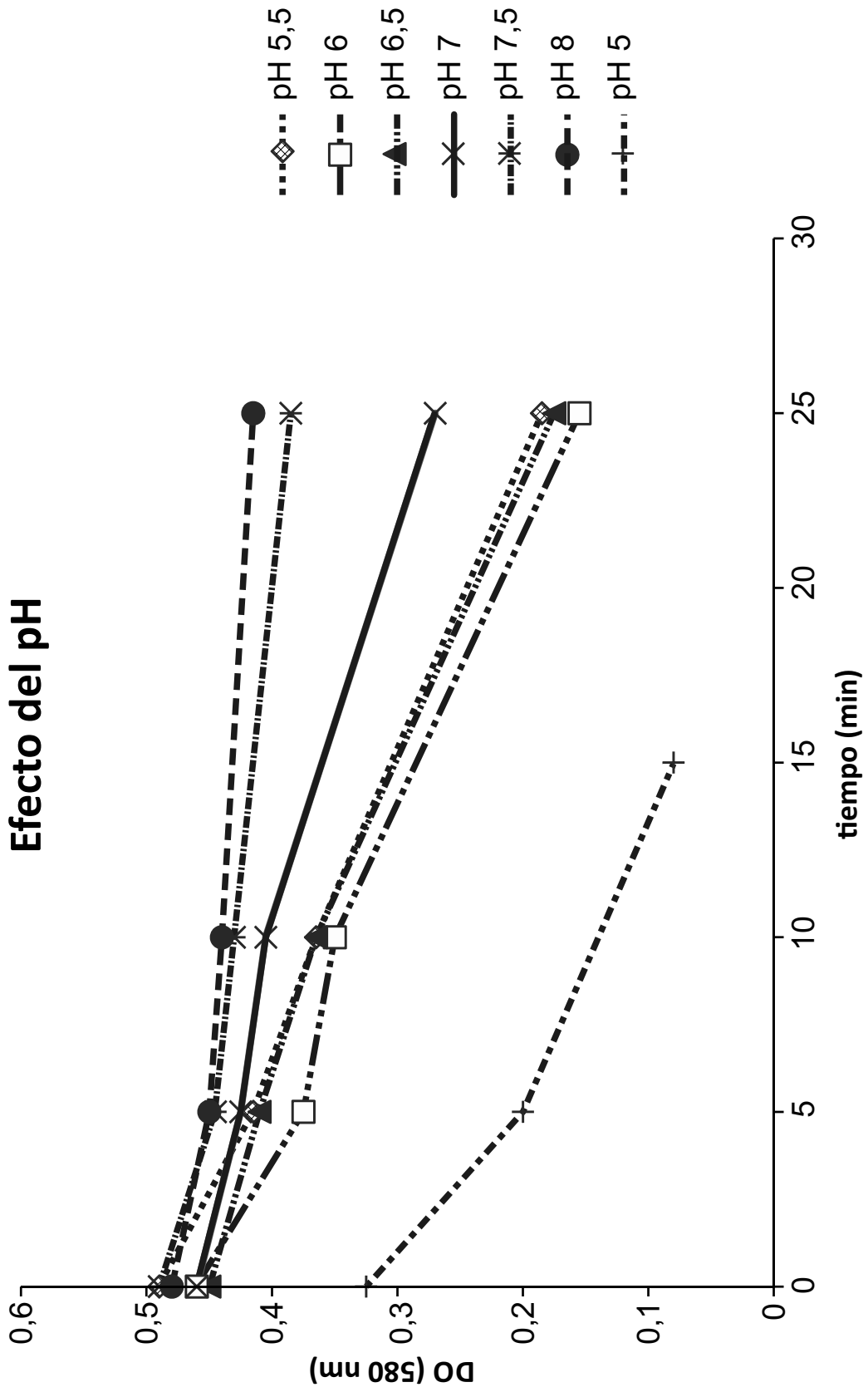


FIG. 8

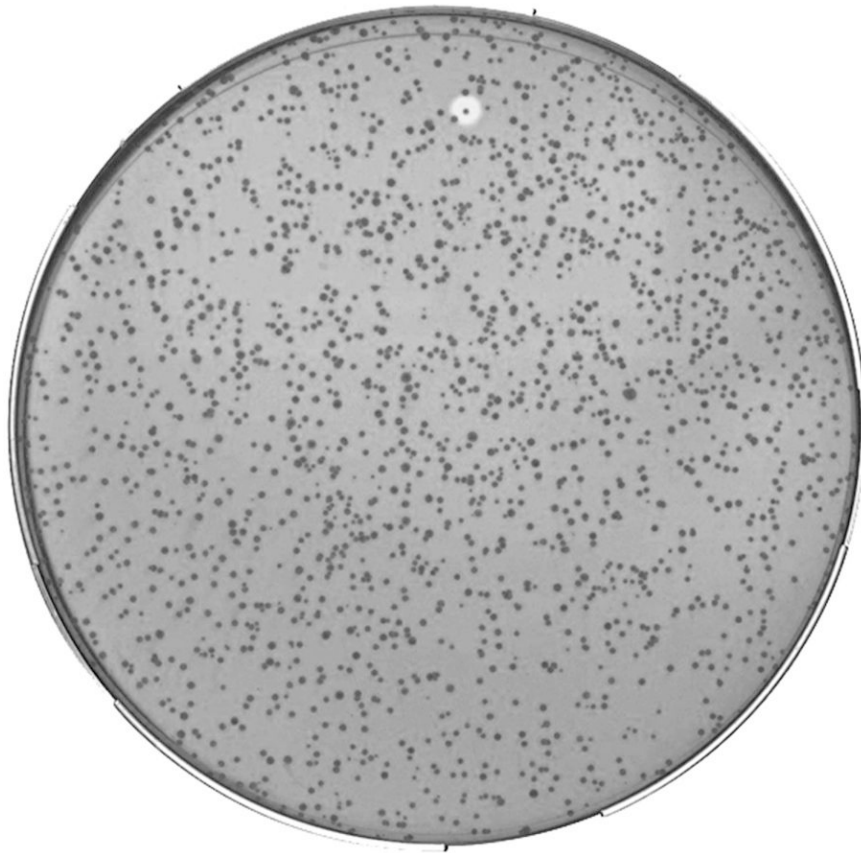


FIG. 9

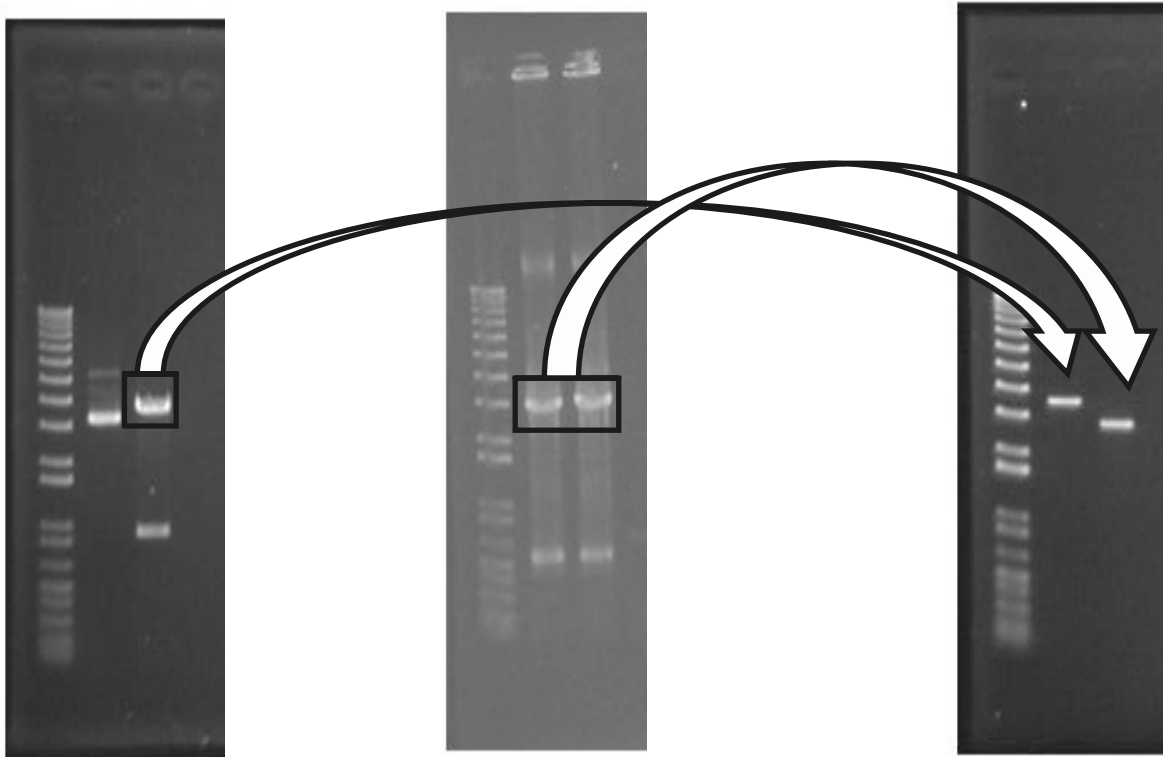


FIG. 10

Lista de Secuencias

<110> ABENGOA RESEARCH, S.L.

<120> POLIPÉPTIDOS CON ACTIVIDAD CELULASA

<130> 194/13

<160> 14

<170> PatentIn version 3.3

<210> 1

<211> 928

<212> PRT

<213> Organismo del suelo no identificado

<400> 1

Met Arg Thr His Met Lys Thr Ser Leu Lys Leu Arg Ala Leu Gly Ala
1 5 10 15

Leu Phe Ala Ala Leu Val Gly Ala Val Ala Ala Pro Ala Ala Ser Ala
20 25 30

Ala Gly Tyr Tyr Ala Gln Gln Gly Lys Ile Tyr Asp Ala Ser Gly Gln
35 40 45

Gln Val Pro Ile Arg Gly Ile Ser His Tyr Gly Phe Asn Ala Asn Ile
50 55 60

Leu Gln Pro Gln Tyr Leu Trp Ser Met Gly Trp Lys Glu Gln Ile Ala
65 70 75 80

Gln Ile Lys Ala Leu Gly Phe Asn Ala Val Arg Leu Pro Tyr Val Pro
85 90 95

Asp Thr Leu Tyr Val Thr Thr Pro Val Asp Gln Leu Ser Tyr Val Asp
100 105 110

Pro Gly Lys Asn Ala Asp Leu Ile Gly Lys Thr Pro Leu Gln Ala Met
115 120 125

Asp Met Trp Met Ala Glu Ala Asp Arg Gln Gly Leu Tyr Val Met Leu
130 135 140

Asp Phe His Ser Val Ser Asn Asp Arg Gln Tyr Pro Thr Trp Phe Val
145 150 155 160

Asp Asn Ser Ala Asp Phe Gly Leu Ile Tyr Asn Ser Ser Ala Tyr Thr
165 170 175

Val Asp Asn Trp Ile Arg Asp Leu Lys Phe Val Ala Ala Arg Tyr Ala
180 185 190

Asn Leu Pro His Phe Phe Ala Ile Asp Ile Tyr Asn Glu Pro Asn Gly
195 200 205

ES 2 545 161 A1

Ile Val Arg Trp Ser Thr Gly Asp Ala Asn Met Thr Gln Ser Lys Asn
 210 215 220

Tyr Trp Lys Asp Ala Ala Glu Lys Ala Ala Ala Gly Val Leu Ala Ala
 225 230 235 240

Asn Pro Asn Leu Met Ile Phe Val Gln Gly Ile Asn Gly Asn Trp Asp
 245 250 255

Gly Val Glu Lys Ser Asn Ile Pro Met Asn Tyr Gly Glu Asp Phe Gln
 260 265 270

Pro Gln Arg Tyr Gln Pro Leu Asn Ile Pro Ala Asp Lys Leu Val Leu
 275 280 285

Ser Pro His Thr Tyr Gly Pro Asp Val Tyr Val Lys Ser Thr Phe Asn
 290 295 300

Ala Ala Asn Phe Pro Ala Asn Leu Ala Ala Asp Trp Glu Thr Leu Phe
 305 310 315 320

Gly Gln Phe Tyr Pro Thr His Pro Val Val Pro Gly Glu Trp Gly Gly
 325 330 335

Arg Tyr Gly Ile Gly Gly Asp Pro Arg Asp Val Ala Trp Gln Asn Ala
 340 345 350

Phe Val Asp Tyr Met Leu Gly Lys Gly Met Arg Asp Ser Phe Tyr Trp
 355 360 365

Cys Tyr Thr Pro Asn Ser Gly Asp Thr Gly Gly Ile Leu Asp Asp Asn
 370 375 380

Leu Gln Val Arg Thr Asp Lys Met Ala Leu Leu Gln Lys Leu Trp Gly
 385 390 395 400

Thr Ser Gly Thr Ser Ala Ala Ser Ile Ala Leu Ser Ser Ser Ala Tyr
 405 410 415

Thr Val Ser Gln Ser Ala Gly Ser Val Ser Leu Ser Val Asn Arg Ser
 420 425 430

Gly Gly Thr Gly Ala Val Ser Val Ser Tyr Ala Thr Ala Gly Gly Ser
 435 440 445

Ala Val Ser Gly Thr Asp Phe Thr Ala Lys Ser Gly Thr Leu Asn Trp
 450 455 460

Ala Ala Gly Asp Thr Ala Ala Lys Thr Ile Thr Ile Pro Val Ser Asn
 465 470 475 480

ES 2 545 161 A1

Ala Thr Pro Phe Ser Gly Ser Lys Thr Phe Thr Val Ser Leu Ser Gly
 485 490 495

Ala Ser Ser Gly Ala Ala Leu Ala Thr Pro Asn Ile Ala Thr Val Thr
 500 505 510

Ile Asn Gly Ser Ala Val Thr Ala Ser Ala Gly Thr Leu Ala Leu Gly
 515 520 525

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

Val Ala Arg Ser Gly Gly Ser Ser Gly Ala Val Ser Val Lys Tyr Ala
 545 550 555 560

Thr Ala Asn Gly Thr Ala Val Ala Gly Thr Asp Tyr Thr Ala Ala Ser
 565 570 575

Gly Thr Leu Ser Trp Ala Ser Gly Asp Thr Ala Ser Lys Ser Phe Ser
 580 585 590

Val Ala Ile Ser Asn Ala Thr Pro Phe Thr Gly Ser Lys Ser Phe Ser
 595 600 605

Ile Thr Leu Ser Gly Thr Ser Gly Gly Ala Thr Leu Gly Ala Ala Ser
 610 615 620

Ala Ser Val Thr Ile Asn Gly Ser Gly Ser Ser Ser Gly Ser Thr
 625 630 635 640

Thr Ser Tyr Thr Gln Pro Tyr Ile Ala Gly Phe Thr Pro Ser Ser Gly
 645 650 655

Pro Ala Gly Thr Thr Val Thr Ile Ser Gly Lys Gly Phe Thr Gly Leu
 660 665 670

Ser Ser Val Thr Ile Ala Gly Thr Ala Ala Thr Phe Lys Val Gly Ser
 675 680 685

Asp Thr Ser Leu Thr Leu Thr Val Pro Ser Asn Ala Ser Ser Gly Gln
 690 695 700

Ile Ala Leu Ile Asn Pro Lys Tyr Ala Ala Trp Ser Gly Ser Pro Phe
 705 710 715 720

Ser Val Thr Ser Ser Ser Ser Thr Thr Thr Ala Pro Gly Thr Leu Ala
 725 730 735

Leu Gly Ala Ser Ser Tyr Ser Val Ala Gln Ser Thr Gly Ile Ala Thr
 740 745 750

ES 2 545 161 A1

Leu Ser Val Gly Arg Ala Ser Gly Ser Ser Gly Ala Val Ser Val Gln
755 760 765

Tyr Ala Thr Val Asn Gly Ser Ala Val Ala Gly Thr Asp Tyr Thr Ala
770 775 780

Thr Ser Gly Thr Leu Gln Trp Thr Ser Gly Asp Thr Ala Ala Lys Thr
785 790 795 800

Ile Ser Ile Pro Leu Ser Thr Gln Lys Ala Phe Thr Gly Ser Lys Ser
805 810 815

Phe Ser Leu Lys Leu Ser Thr Thr Ser Gly Gly Ala Ala Leu Gly Ser
820 825 830

Pro Ser Ser Ala Thr Val Thr Ile Asn Gly Gly Ser Thr Ala Tyr Ala
835 840 845

Gln Pro Tyr Ile Gln Thr Phe Ser Pro Gln Ser Gly Ala Val Gly Thr
850 855 860

Val Val Thr Ile Thr Gly Ser Gly Phe Thr Gly Leu Ser Arg Val Thr
865 870 875 880

Leu Gly Ser Ser Ser Ala Thr Phe Gln Val Val Ser Asp Thr Gln Leu
885 890 895

Lys Val Thr Ile Pro Ser Gly Ala Ser Thr Ala Gln Ile Ala Leu Phe
900 905 910

Asn Ser Lys Tyr Ala Ala Trp Thr Pro Ser Ala Phe Thr Val Thr Lys
915 920 925

<210> 2
<211> 927
<212> PRT
<213> Organismo del suelo no identificado
<400> 2

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

Leu Leu Ala Thr Leu Val Gly Ala Ala Ala Ala Pro Ala Ala Ser Ala
20 25 30

Ala Gly Tyr Tyr Ala Gln Gln Gly Lys Ile Tyr Asp Ala Ser Gly Gln
35 40 45

Gln Val Pro Ile Arg Gly Ile Ser His Tyr Gly Phe Asn Ala Asn Ile
50 55 60

Leu Gln Pro Gln Tyr Leu Trp Ser Met Gly Trp Lys Glu Gln Ile Ala
65 70 75 80

ES 2 545 161 A1

Gln Ile Lys Ala Leu₈₅ Gly Phe Asn Ala Val₉₀ Arg Leu Pro Tyr Val₉₅ Pro
 Asp Thr Leu Tyr₁₀₀ Val Thr Thr Pro Val₁₀₅ Asp Gln Leu Ser Tyr₁₁₀ Val Asp
 Pro Gly Lys₁₁₅ Asn Pro Glu Leu Leu₁₂₀ Gly Lys Thr Pro Leu₁₂₅ Glu Val Leu
 Asp Leu₁₃₀ Trp Met Ala Glu Ala₁₃₅ Asp Arg Gln Gly Leu Tyr Val Met Leu
 Asp Phe His Ser Val Ser₁₅₀ Asn Asp Arg Gln Tyr₁₅₅ Gln Thr Trp Phe Val₁₆₀
 Asp Asn Ser Ala Asp₁₆₅ Phe Gly Leu Ile Tyr₁₇₀ Asn Ser Gln Ala Tyr₁₇₅ Thr
 Val Asp Asn Trp₁₈₀ Ile Arg Asp Leu Lys₁₈₅ Phe Val Ala Ala Arg Tyr Ala
 Asn Leu Pro₁₉₅ His Phe Phe Ala Ile₂₀₀ Asp Ile Tyr Asn Glu Pro Asn Gly
 Val Val₂₁₀ Arg Trp Ser Thr Gly₂₁₅ Asp Ala Asn Met Thr₂₂₀ Gln Ser Lys Asn
 Tyr Trp Lys Asp Ala Ala₂₃₀ Glu Lys Gly Ala Ala₂₃₅ Gly Val Leu Ala Ala₂₄₀
 Asn Pro Asn Leu Met₂₄₅ Ile Phe Val Gln Gly₂₅₀ Ile Asn Gly Asn Trp Asp
 Gly Ile Glu Asn₂₆₀ Thr Asn Ile Pro Met₂₆₅ Asn Tyr Gly Glu Asp Phe Gln
 Pro Gln Ala Tyr Gln Pro Leu Asn₂₈₀ Ile Pro Asn Asp Lys₂₈₅ Leu Val Leu
 Ser Pro₂₉₀ His Thr Tyr Gly Pro₂₉₅ Asp Val Tyr Val Lys₃₀₀ Ser Thr Phe Asn
 Ala Ala Asn Phe Pro Ala Asn Leu Ala Ala Asp₃₁₅ Trp Glu Thr Leu Phe₃₂₀
 Gly Gln Phe Tyr Pro₃₂₅ Ser His Pro Val Val₃₃₀ Pro Gly Glu Trp Gly₃₃₅ Gly
 Arg Tyr Gly Val₃₄₀ Gly Gly Asp Ser Arg₃₄₅ Asp Val Ala Trp Gln Asn Ala₃₅₀

ES 2 545 161 A1

Phe Val Asp Tyr Met Leu Gly Lys Gly Met Arg Asp Ser Phe Tyr Trp
 355 360 365
 Cys Tyr Thr Pro Asn Ser Gly Asp Thr Gly Gly Ile Leu Asp Asp Ser
 370 375 380
 Leu Asn Val Arg Thr Asp Lys Met Ala Leu Leu Gln Lys Leu Trp Gly
 385 390 395 400
 Thr Ser Gly Ser Thr Ala Ser Ile Ala Leu Ser Ser Ser Ala Tyr Thr
 405 410 415
 Val Ser Gln Ser Ala Gly Ser Val Ser Leu Ser Val Ser Arg Ser Gly
 420 425 430
 Gly Thr Gly Ala Val Ser Val Ser Tyr Ala Thr Ala Gly Gly Thr Ala
 435 440 445
 Val Ser Gly Thr Asp Phe Thr Pro Lys Ser Gly Thr Leu Ser Trp Ala
 450 455 460
 Ala Gly Asp Thr Ala Ala Lys Thr Ile Ala Ile Pro Val Ser Asn Ala
 465 470 475 480
 Thr Pro Phe Thr Gly Ser Lys Thr Phe Thr Val Ser Leu Ser Gly Ala
 485 490 495
 Ser Ser Gly Ala Ser Leu Ala Thr Pro Asn Ile Ala Thr Val Thr Ile
 500 505 510
 Asn Gly Ser Gly Val Val Ala Ser Ala Gly Ser Leu Ala Leu Gly Ala
 515 520 525
 Ser Ser Tyr Ser Val Ala Gln Ser Ala Gly Ser Leu Thr Val Thr Val
 530 535 540
 Ala Arg Ser Gly Gly Ser Ser Gly Ala Val Ser Val Lys Tyr Ala Thr
 545 550 555 560
 Ala Asn Gly Thr Ala Val Ala Gly Ser Asp Tyr Thr Ala Ala Ser Gly
 565 570 575
 Thr Leu Ser Trp Ala Ser Gly Glu Thr Ala Ser Lys Ser Phe Ser Val
 580 585 590
 Ala Ile Ser Asn Ala Thr Pro Tyr Thr Gly Ser Lys Ser Phe Ser Ile
 595 600 605
 Ala Leu Ser Gly Ala Ser Gly Gly Ala Thr Leu Gly Ala Ala Ser Ala
 610 615 620

ES 2 545 161 A1

Ser Val Thr Ile Asn Gly Ser Gly Ser Ser Ser Gly Ser Gly Thr Thr
 625 630 635 640
 Ser Tyr Thr Gln Pro Tyr Ile Ala Gly Phe Thr Pro Ser Ser Gly Pro
 645 650 655
 Gly Gly Thr Thr Val Thr Ile Ser Gly Lys Gly Phe Thr Gly Leu Ser
 660 665 670
 Ser Val Ser Ile Ala Gly Thr Ala Ala Thr Phe Lys Val Ser Ser Asp
 675 680 685
 Thr Ser Leu Thr Leu Thr Val Pro Gly Ser Ala Gly Ser Gly Gln Ile
 690 695 700
 Ala Leu Ile Asn Pro Lys Tyr Ala Ala Trp Ser Gly Ser Pro Phe Ser
 705 710 715 720
 Val Thr Ser Ser Ser Thr Thr Thr Ala Pro Gly Thr Leu Ala Leu
 725 730 735
 Gly Ala Ser Ser Tyr Ser Val Ala Gln Ser Ser Gly Ile Ala Thr Leu
 740 745 750
 Ser Val Ala Arg Ala Ser Gly Ser Ser Gly Ala Val Ser Val Lys Tyr
 755 760 765
 Ala Thr Val Asp Gly Ser Ala Val Ala Gly Thr Asp Tyr Thr Ala Val
 770 775 780
 Ser Gly Thr Leu Ser Trp Ala Ser Gly Asp Ala Ala Ala Lys Thr Val
 785 790 795 800
 Ser Ile Pro Ile Ser Thr Leu Lys Ala Phe Thr Gly Ser Lys Ser Phe
 805 810 815
 Ser Leu Lys Leu Ser Thr Thr Ser Gly Gly Ala Thr Leu Gly Ser Pro
 820 825 830
 Ser Trp Ala Thr Val Thr Ile Asn Gly Gly Ser Thr Ala Tyr Thr Gln
 835 840 845
 Pro Tyr Ile Gln Thr Phe Ser Pro Arg Ser Gly Ala Ala Gly Thr Val
 850 855 860
 Val Thr Ile Thr Gly Ser Gly Phe Thr Gly Leu Ser Arg Val Thr Val
 865 870 875 880
 Gly Ser Ser Ser Ala Thr Phe Leu Val Val Ser Asp Thr Gln Leu Gln
 885 890 895
 Val Thr Ile Pro Ser Gly Ala Ser Thr Ala Gln Ile Ala Leu Phe Asn

ES 2 545 161 A1

900

905

910

Pro Lys Tyr Ala Ala Trp Thr Pro Ser Ala Phe Thr Val Thr Pro
 915 920 925

<210> 3
 <211> 338
 <212> PRT
 <213> Organismo del suelo no identificado
 <400> 3

Lys Ile Tyr Asp Ala Ser Gly Gln Gln Val Pro Ile Arg Gly Ile Ser
 1 5 10 15

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

Met Gly Trp Lys Glu Gln Ile Ala Gln Ile Lys Ala Leu Gly Phe Asn
 35 40 45

Ala Val Arg Leu Pro Tyr Val Pro Asp Thr Leu Tyr Val Thr Thr Pro
 50 55 60

Val Asp Gln Leu Ser Tyr Val Asp Pro Gly Lys Asn Ala Asp Leu Ile
 65 70 75 80

Gly Lys Thr Pro Leu Gln Ala Met Asp Met Trp Met Ala Glu Ala Asp
 85 90 95

Arg Gln Gly Leu Tyr Val Met Leu Asp Phe His Ser Val Ser Asn Asp
 100 105 110

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

Ile Tyr Asn Ser Ser Ala Tyr Thr Val Asp Asn Trp Ile Arg Asp Leu
 130 135 140

Lys Phe Val Ala Ala Arg Tyr Ala Asn Leu Pro His Phe Phe Ala Ile
 145 150 155 160

Asp Ile Tyr Asn Glu Pro Asn Gly Ile Val Arg Trp Ser Thr Gly Asp
 165 170 175

Ala Asn Met Thr Gln Ser Lys Asn Tyr Trp Lys Asp Ala Ala Glu Lys
 180 185 190

Ala Ala Ala Gly Val Leu Ala Ala Asn Pro Asn Leu Met Ile Phe Val
 195 200 205

Gln Gly Ile Asn Gly Asn Trp Asp Gly Val Glu Lys Ser Asn Ile Pro
 210 215 220

ES 2 545 161 A1

Met Asn Tyr Gly Glu Asp Phe Gln Pro Gln Arg Tyr Gln Pro Leu Asn
225 230 235 240

Ile Pro Ala Asp Lys Leu Val Leu Ser Pro His Thr Tyr Gly Pro Asp
245 250 255

Val Tyr Val Lys Ser Thr Phe Asn Ala Ala Asn Phe Pro Ala Asn Leu
260 265 270

Ala Ala Asp Trp Glu Thr Leu Phe Gly Gln Phe Tyr Pro Thr His Pro
275 280 285

Val Val Pro Gly Glu Trp Gly Gly Arg Tyr Gly Ile Gly Gly Asp Pro
290 295 300

Arg Asp Val Ala Trp Gln Asn Ala Phe Val Asp Tyr Met Leu Gly Lys
305 310 315 320

Gly Met Arg Asp Ser Phe Tyr Trp Cys Tyr Thr Pro Asn Ser Gly Asp
325 330 335

Thr Gly

- <210> 4
- <211> 378
- <212> PRT
- <213> Organismo del suelo no identificado
- <400> 4

Met Arg Thr His Met Lys Thr Ser Leu Lys Leu Arg Ala Leu Gly Ala
1 5 10 15

Leu Phe Ala Ala Leu Val Gly Ala Val Ala Ala Pro Ala Ala Ser Ala
20 25 30

Ala Gly Tyr Tyr Ala Gln Gln Gly Lys Ile Tyr Asp Ala Ser Gly Gln
35 40 45

Gln Val Pro Ile Arg Gly Ile Ser His Tyr Gly Phe Asn Ala Asn Ile
50 55 60

Leu Gln Pro Gln Tyr Leu Trp Ser Met Gly Trp Lys Glu Gln Ile Ala
65 70 75 80

Gln Ile Lys Ala Leu Gly Phe Asn Ala Val Arg Leu Pro Tyr Val Pro
85 90 95

Asp Thr Leu Tyr Val Thr Thr Pro Val Asp Gln Leu Ser Tyr Val Asp
100 105 110

Pro Gly Lys Asn Ala Asp Leu Ile Gly Lys Thr Pro Leu Gln Ala Met

ES 2 545 161 A1

115 120 125
 Asp Met Trp Met Ala Glu Ala Asp Arg Gln Gly Leu Tyr Val Met Leu
 130 135 140
 Asp Phe His Ser Val Ser Asn Asp Arg Gln Tyr Pro Thr Trp Phe Val
 145 150 155
 Asp Asn Ser Ala Asp Phe Gly Leu Ile Tyr Asn Ser Ser Ala Tyr Thr
 165 170 175
 Val Asp Asn Trp Ile Arg Asp Leu Lys Phe Val Ala Ala Arg Tyr Ala
 180 185 190
 Asn Leu Pro His Phe Phe Ala Ile Asp Ile Tyr Asn Glu Pro Asn Gly
 195 200 205
 Ile Val Arg Trp Ser Thr Gly Asp Ala Asn Met Thr Gln Ser Lys Asn
 210 215 220
 Tyr Trp Lys Asp Ala Ala Glu Lys Ala Ala Ala Gly Val Leu Ala Ala
 225 230 235 240
 Asn Pro Asn Leu Met Ile Phe Val Gln Gly Ile Asn Gly Asn Trp Asp
 245 250 255
 Gly Val Glu Lys Ser Asn Ile Pro Met Asn Tyr Gly Glu Asp Phe Gln
 260 265 270
 Pro Gln Arg Tyr Gln Pro Leu Asn Ile Pro Ala Asp Lys Leu Val Leu
 275 280 285
 Ser Pro His Thr Tyr Gly Pro Asp Val Tyr Val Lys Ser Thr Phe Asn
 290 295 300
 Ala Ala Asn Phe Pro Ala Asn Leu Ala Ala Asp Trp Glu Thr Leu Phe
 305 310 315 320
 Gly Gln Phe Tyr Pro Thr His Pro Val Val Pro Gly Glu Trp Gly Gly
 325 330
 Arg Tyr Gly Ile Gly Gly Asp Pro Arg Asp Val Ala Trp Gln Asn Ala
 340 345 350
 Phe Val Asp Tyr Met Leu Gly Lys Gly Met Arg Asp Ser Phe Tyr Trp
 355 360 365
 Cys Tyr Thr Pro Asn Ser Gly Asp Thr Gly
 370 375

<210> 5
 <211> 570

ES 2 545 161 A1

<212> PRT

<213> Organismo del suelo no identificado

<400> 5

Met Arg Thr His Met Lys Thr Ser Leu Lys Leu Arg Ala Leu Gly Ala
 1 5 10 15

Leu Phe Ala Ala Leu Val Gly Ala Val Ala Ala Pro Ala Ala Ser Ala
 20 25 30

Ala Gly Tyr Tyr Ala Gln Gln Gly Lys Ile Tyr Asp Ala Ser Gly Gln
 35 40 45

Gln Val Pro Ile Arg Gly Ile Ser His Tyr Gly Phe Asn Ala Asn Ile
 50 55 60

Leu Gln Pro Gln Tyr Leu Trp Ser Met Gly Trp Lys Glu Gln Ile Ala
 65 70 75 80

Gln Ile Lys Ala Leu Gly Phe Asn Ala Val Arg Leu Pro Tyr Val Pro
 85 90 95

Asp Thr Leu Tyr Val Thr Thr Pro Val Asp Gln Leu Ser Tyr Val Asp
 100 105 110

Pro Gly Lys Asn Ala Asp Leu Ile Gly Lys Thr Pro Leu Gln Ala Met
 115 120 125

Asp Met Trp Met Ala Glu Ala Asp Arg Gln Gly Leu Tyr Val Met Leu
 130 135 140

Asp Phe His Ser Val Ser Asn Asp Arg Gln Tyr Pro Thr Trp Phe Val
 145 150 155 160

Asp Asn Ser Ala Asp Phe Gly Leu Ile Tyr Asn Ser Ser Ala Tyr Thr
 165 170 175

Val Asp Asn Trp Ile Arg Asp Leu Lys Phe Val Ala Ala Arg Tyr Ala
 180 185 190

Asn Leu Pro His Phe Phe Ala Ile Asp Ile Tyr Asn Glu Pro Asn Gly
 195 200 205

Ile Val Arg Trp Ser Thr Gly Asp Ala Asn Met Thr Gln Ser Lys Asn
 210 215 220

Tyr Trp Lys Asp Ala Ala Glu Lys Ala Ala Ala Gly Val Leu Ala Ala
 225 230 235 240

Asn Pro Asn Leu Met Ile Phe Val Gln Gly Ile Asn Gly Asn Trp Asp
 245 250 255

ES 2 545 161 A1

Gly Val Glu Lys Ser Asn Ile Pro Met Asn Tyr Gly Glu Asp Phe Gln
 260 265 270

Pro Gln Arg Tyr Gln Pro Leu Asn Ile Pro Ala Asp Lys Leu Val Leu
 275 280 285

Ser Pro His Thr Tyr Gly Pro Asp Val Tyr Val Lys Ser Thr Phe Asn
 290 295 300

Ala Ala Asn Phe Pro Ala Asn Leu Ala Ala Asp Trp Glu Thr Leu Phe
 305 310 315 320

Gly Gln Phe Tyr Pro Thr His Pro Val Val Pro Gly Glu Trp Gly Gly
 325 330 335

Arg Tyr Gly Ile Gly Gly Asp Pro Arg Asp Val Ala Trp Gln Asn Ala
 340 345 350

Phe Val Asp Tyr Met Leu Gly Lys Gly Met Arg Asp Ser Phe Tyr Trp
 355 360 365

Cys Tyr Thr Pro Asn Ser Gly Asp Thr Gly Gly Ile Leu Asp Asp Asn
 370 375 380

Leu Gln Val Arg Thr Asp Lys Met Ala Leu Leu Gln Lys Leu Trp Gly
 385 390 395 400

Thr Ser Gly Thr Ser Ala Ala Ser Ile Ala Leu Ser Ser Ser Ala Tyr
 405 410 415

Thr Val Ser Gln Ser Ala Gly Ser Val Ser Leu Ser Val Asn Arg Ser
 420 425 430

Gly Gly Thr Gly Ala Val Ser Val Ser Tyr Ala Thr Ala Gly Gly Ser
 435 440 445

Ala Val Ser Gly Thr Asp Phe Thr Ala Lys Ser Gly Thr Leu Asn Trp
 450 455 460

Ala Ala Gly Asp Thr Ala Ala Lys Thr Ile Thr Ile Pro Val Ser Asn
 465 470 475 480

Ala Thr Pro Phe Ser Gly Ser Lys Thr Phe Thr Val Ser Leu Ser Gly
 485 490 495

Ala Ser Ser Gly Ala Ala Leu Ala Thr Pro Asn Ile Ala Thr Val Thr
 500 505 510

Ile Asn Gly Ser Ala Val Thr Ala Ser Ala Gly Thr Leu Ala Leu Gly
 515 520 525

Ala Ser Ser Tyr Ser Val Ala Gln Asn Ala Gly Thr Leu Thr Val Thr

ES 2 545 161 A1

530

535

540

Val Ala Arg Ser Gly Gly Ser Ser Gly Ala Val Ser Val Lys Tyr Ala
 545 550 555 560

Thr Ala Asn Gly Thr Ala Val Ala Gly Thr
 565 570

<210> 6
 <211> 2787
 <212> DNA
 <213> Organismo del suelo no identificado

<400> 6
 atgaggacc acatgaaaac gaggcctgaaa ctccgggcat tgggcgcgct ctttgcggcg 60
 ctggtcggcg ccgtcgcggc gccggccgcc agcgccgcgg gctattacgc gcagcagggc 120
 aagatctacg acgccagcgg ccagcaggta ccgatccgcg gcatcagcca ctacggcttc 180
 aatgccaaca tcctgcagcc gcagtatctg tggtcgatgg gctggaagga gcagatcgcg 240
 cagatcaagg cgctcggctt caacgccgtg cgtctgccgt acgtgccgga cacgctgtac 300
 gtgacgacgc cggtcgacca gctcagctac gtcgatccgg gcaagaacgc ggacctgatc 360
 ggcaagacgc cgctgcaggc gatggacatg tggatggccg aggccgatcg ccagggcctc 420
 tacgtcatgc tcgacttcca cagcgtctcg aatgatcgcc agtatccgac ctggttcgctc 480
 gacaacagcg ccgacttcgg cctgatctac aacagctcgg cctacaccgt cgacaactgg 540
 atccgcgatc tgaagtctgt cgccgcgcgc tacgccaacc tgccgcactt cttcgccatc 600
 gacatctaca acgagccgaa cggcatcgtg cgctggagca ccggcgacgc caacatgacg 660
 cagtcgaaga actactggaa ggatgccgcc gagaaggctg ccgccggcgt gcttgcggcc 720
 aatccgaacc tgatgatctt cgtgcagggc atcaacggca actgggatgg cgtcgagaag 780
 agcaacatcc cgatgaacta cggcgaggac ttccagccgc agcgctacca gccgctcaac 840
 atcccggccg acaagctggt gctgagcccg cacacctacg gcccgacgt ctacgtgaag 900
 tcgacgttca acgccgcaa cttcccggcc aatctggcgg ccgactggga gacgctgttc 960
 ggccagttct atccgacgca cccggtcgtg ccgggcgagt ggggcggccg ttacggcatc 1020
 ggcggtgatc cgcgcgacgt cgcctggcag aacgccttcg tcgactacat gctcggcaag 1080
 ggcatgcgcg acagcttcta ctggtgctac acgccgaaca gggcgacac cggcggcatc 1140
 ctcgacgaca acctgcaggt ggcaccgac aagatggcgc tgctgcagaa gctgtggggc 1200
 accagcggca ccagcgcggc atcgatcgcg ctgtcgtcga ggcctacac ggtcagccag 1260
 agcgcgggtt cggtgagcct gtcggtcaat cgcagcggcg gcacgggcgc ggtctcggtc 1320
 agctatgcga ccgccggcgg cagcgcggtt tccggcaccg acttcacggc caagagcggc 1380
 acgctgaact gggccgcggg cgacaccgcg gcgaagacca tcacgattcc ggtcagcaac 1440
 gcgacgccgt tcagcggcag caagacctc accgtctcgc tgtcgggcgc cagcagcggc 1500
 gccgcgctcg ccacgccgaa catcgccacg gtgacgatca acggttcggc ggtcacggcc 1560

ES 2 545 161 A1

tcggccggca cgctcgcgct cggcgcgctc agctacagcg tcgcgcagaa tgccggcagc 1620
ctgaccgtca cggtcgcgcg cagcggcggc tcgagcggcg cggtcagcgt caagtacgcc 1680
accgccaacg gcacggcggc cgccggcacc gactacacgg ctgccagcgg tacgctgagc 1740
tgggcttcgg gcgacaccgc gagcaagagc ttcagcgtcg cgatcagcaa cgccacgccg 1800
ttcaccggca gcaagagctt cagcatcacg ctgtccggca catccggcgg cgccacgctc 1860
ggcgcggcca gcgccagcgt cacgatcaac ggctccggca gcagcagcag cggcagcacg 1920
accagctaca cgcagccgta catcgccggc ttcacgccga gctcgggtcc ggccggcagc 1980
acggtgacga tcagcggcaa gggcttcacc ggcctgagca gcgtgacct cgctggcacc 2040
gccgcgacct tcaaggtcgg cagcgacacc tcgctgacgc tgaccgtgcc gagcaatgcc 2100
agcagcggcc agatcgact gatcaaccg aagtacgcgg cctggtcggg cagcccgttc 2160
agcgtgacct cgtcgtcgtc gacgacgacc gcgccgggca cgcttgact cggcgcgctc 2220
agctacagcg tcgcgcagag cacgggcatt gccacgctca gcgtcggctc tgccagcggc 2280
tccagcggcg ccgtcagcgt gcagtacgcg accgtcaacg gcagcgcctg cgccggcacc 2340
gactacacgg cgaccagcgg cacgctgacg tggacctcgg gcgacaccgc ggccaagacc 2400
atcagcattc cgctcagcac gcagaaggcc ttcaccggca gcaagtcggt ctcgctgaaa 2460
ctgagcacga cctccggcgg cgccgcgctc ggctcgccgt cgctcggccac ggtgacgatc 2520
aacggcggct cgaccgctta tgcgcagccg tacatccaga cttctcgcg gcagagcggc 2580
gcggtcggga cgggtggtgac gatcacgggc tcgggcttca ctggcctgtc acgggtcacg 2640
ctcggttcgt cgtcggcgac tttccagggt gtcagcgata cgcagctgaa ggtcacgatt 2700
ccgtccggtg ccagcaccgc gcagatcgcg ctgttcaatt cgaagtacgc ggcattggacg 2760
ccgtcggctt tcaccgtcac caagtga 2787

<210> 7
<211> 2784
<212> DNA
<213> Organismo del suelo no identificado

<400> 7
atgaggacc ctatgaagat gagcctgaaa ctccgggcgc tgggcgcgct ccttgcgacg 60
ctggtcggcg ccgccggc gccggccgag agcggccg cgtattacgc gcagcagggc 120
aagatctacg acgccagcgg ccagcaggtg ccgatccgcg gcacagcca ctacggtttc 180
aatgccaaca tcctgcagcc gcagtacctc tggtcgatgg gctggaagga gcagatcgcg 240
cagatcaagg cgctcggctt caacgccgtg cgtctgccgt acgtgccgga cacgctgtac 300
gtgacgacgc cggtcgacca gctgagctat gtcgaccgga gcaagaacc ggaactgctc 360
ggcaagacgc cgctggaggt actggacctg tggatggcgg aggccgaccg tcagggcctc 420
tacgtcatgc tcgatttcca cagcgtctcg aacgaccgcc aatatcagac ctggttcgctc 480
gacaacagcg ccgacttcgg cctgatctac aacagccagg cctataccgt cgacaactgg 540
atccgcgatc tcaagttcgt cgccgcgagc tacgcgaacc tgccgcactt cttcgcgatc 600

ES 2 545 161 A1

gacatctaca acgagccgaa cggcgtcgtg cgctggagca ccggcgacgc caacatgacg 660
cagtcgaaga actactggaa ggacgccgcc gagaagggcg ccgccggcgt gctcgcggcc 720
aaccccaacc tgatgatctt cgtgcagggc atcaacggca actgggacgg catcgagaac 780
accaacatcc cgatgaacta cggcgaggac ttccagccgc aggcctacca gccgctcaac 840
atcccgaacg acaagctggt gctgtcgccg cacacctatg gcccggacgt ctatgtgaag 900
tcgacgttca acgccgcaa cttcccggcc aacctcgccg ccgactggga gacgctgttc 960
ggccagttct atccgtcgca tccggtcgtg cccggcgagt ggggcggtcg ctacggcgtc 1020
ggcggcgact cgcgcgacgt cgcctggcag aacgcgttcg tcgactacat gctcggcaag 1080
ggcatgcgtg acagcttcta ctggtgctac acgccgaaca gcggcgacac cggcggcatc 1140
ctcgacgaca gcctcaacgt gcgcaccgac aagatggcgc tgctgcagaa gctgtggggc 1200
acgagcggca gcacggcgtc gatcgcgctg tcatcgagcg cctacaccgt cagccagagc 1260
gccggttcgg tgagcctgtc agtcagccgc agcggcggca cgggcgcggt gtcggtcagc 1320
tatgcgacgg ccggcggcac cgcggtgtcc ggaccgact tcacgccgaa gagcggcacg 1380
ctgagctggg ccgccggcga cacggcggct aagaccatcg ccattccggt cagcaaccg 1440
acgccattca ccggcagcaa gacctcacc gtttcgctgt ccggcgccag cagcggcgc 1500
tcgctcgcca cgccgaacat cgccacggtg acgatcaacg gttcgggctg cgtcgcctcg 1560
gccggcagcc tcgcgctcgg ggcatcgagc tacagcgtcg cgcagagcgc cggctcgtg 1620
accgtcacgg tcgcgcgag cggcggttcg agcggtgcgg tcagcgtcaa gtacgccacc 1680
gccaacggca ccgccgttg cggcagcgac tacacggccg ccagcggcac gctgagctgg 1740
gttcgggcg agacggcgag caagtcgttc agcgtcgcga tcagcaacgc cacgccgtac 1800
accggcagca agagcttcag catcgcgctg tccggcgcgt ccggcggcgc cacgctcggc 1860
gcggccagcg ccagcgtcac gatcaacggc tcgggcagca gcagcggcag cggcacgacg 1920
agctacacgc agccgtacat cgccggcttc acgccgagct cgggtccggg cggcacgacg 1980
gtgacgatca gcggcaaggg cttaccggc ctgagcagcg tgagcatcgc cggcaccgcc 2040
gcgacattca aggtcagcag cgacacctg ctgacgtga ccgtgccggg cagcgcggc 2100
agcggccaga tcgcgctgat caaccgaag tacgccgctt ggtcgggag cccgttcagc 2160
gtgacctcgt cgtcgtcgac cacgaccgcg ccgggcacgc tcgcgctcgg tgctcagac 2220
tacagcgtcg cgcagagctc gggcatcgcc acgctcagcg tcgcgcgcgc cagcggctcc 2280
agcggcgccg tcagcgtgaa gtacgcgacg gtcgacggca gtgccgtcgc cggcaccgac 2340
tacacggcgg tcagcggcac gctgagctgg gcttcgggcg atgctgcagc caagaccgtc 2400
agcatcccga tcagcagcgt gaaggccttc accggcagca agtcgttctc gctgaagctg 2460
agcacgacct ccggcggcgc cacgctcggc tcgccgtcgt gggccacggt gacgatcaac 2520
ggcggctcga ccgcgtacac gcagccgtac atccagacct tctcgccacg cagcggcgcg 2580
gccggcacgg tggtgacgat caccggctcg ggcttcaccg gcctgtcgcg ggtcacggtc 2640
ggctcgtcgt cggcgacttt cctggtggtc agcgacacgc agctgcaggt cacgattccg 2700

ES 2 545 161 A1

tccggtgcca gcaccgcca gatcgcgctg ttcaaccga agtacgcggc atggacgcc 2760
 tcggctttca ccgtcacccc gtga 2784

<210> 8
 <211> 1014
 <212> DNA
 <213> Organismo del suelo no identificado

<400> 8
 aagatctacg acgccagcgg ccagcaggta ccgatccgcg gcatcagcca ctacggcttc 60
 aatgccaaca tcctgcagcc gcagtatctg tggtcgatgg gctggaagga gcagatcgcg 120
 cagatcaagg cgctcggctt caacgccgtg cgtctgccgt acgtgccgga cacgctgtac 180
 gtgacgacgc cggtcgacca gctcagctac gtcgatccgg gcaagaacgc ggacctgatc 240
 ggcaagacgc cgctgcaggc gatggacatg tggatggccg aggccgatcg ccagggcctc 300
 tacgtcatgc tcgacttcca cagcgtctcg aatgatcgcc agtatccgac ctggttcgtc 360
 gacaacagcg ccgacttcgg cctgatctac aacagctcgg cctacaccgt cgacaactgg 420
 atccgcgatc tgaagtctgt cgccgcgcgc tacgccaacc tgccgcactt cttcgccatc 480
 gacatctaca acgagccgaa cggcatcgtg cgctggagca ccggcgacgc caacatgacg 540
 cagtcgaaga actactggaa ggatgccgcc gagaaggctg ccgccggcgt gcttgccggcc 600
 aatccgaacc tgatgatctt cgtgcagggc atcaacggca actgggatgg cgtcgagaag 660
 agcaacatcc cgatgaacta cggcgaggac ttccagccgc agcgtacca gccgctcaac 720
 atcccggccg acaagctggt gctgagcccg cacacctag gcccgacgt ctacgtgaag 780
 tcgacgttca acgccgcaa cttcccggcc aatctggcgg ccgactggga gacgctgttc 840
 ggccagttct atccgacga cccggtcgtg ccgggcgagt ggggcggccg ttacggcatc 900
 ggcggtgatc cgcgcgacgt cgcctggcag aacgccttcg tcgactacat gctcggcaag 960
 ggcatgcgcg acagcttcta ctggtgctac acgccgaaca gcggcgacac cggc 1014

<210> 9
 <211> 1134
 <212> DNA
 <213> Organismo del suelo no identificado

<400> 9
 atgaggacc acatgaaaac gagcctgaaa ctccgggcat tgggcgcgct ctttgccggcg 60
 ctggtcggcg ccgtcgcggc gccggccgcc agcgcgcgg gctattacgc gcagcagggc 120
 aagatctacg acgccagcgg ccagcaggta ccgatccgcg gcatcagcca ctacggcttc 180
 aatgccaaca tcctgcagcc gcagtatctg tggtcgatgg gctggaagga gcagatcgcg 240
 cagatcaagg cgctcggctt caacgccgtg cgtctgccgt acgtgccgga cacgctgtac 300
 gtgacgacgc cggtcgacca gctcagctac gtcgatccgg gcaagaacgc ggacctgatc 360
 ggcaagacgc cgctgcaggc gatggacatg tggatggccg aggccgatcg ccagggcctc 420
 tacgtcatgc tcgacttcca cagcgtctcg aatgatcgcc agtatccgac ctggttcgtc 480

ES 2 545 161 A1

gacaacagcg ccgacttcgg cctgatctac aacagctcgg cctacaccgt cgacaactgg 540
atccgcgatc tgaagtctgt cgccgcgcgc tacgccaacc tgccgcactt cttcgccatc 600
gacatctaca acgagccgaa cggcatcgtg cgctggagca ccggcgacgc caacatgacg 660
cagtcgaaga actactggaa ggatgccgcc gagaaggctg ccgccggcgt gcttgcggcc 720
aatccgaacc tgatgatctt cgtgcagggc atcaacggca actgggatgg cgtcgagaag 780
agcaacatcc cgatgaacta cggcgaggac ttccagccgc agcgctacca gccgctcaac 840
atcccggccg acaagctggt gctgagcccc cacacctacg gcccgacgt ctacgtgaag 900
tcgacgttca acgccgcaa cttcccggcc aatctggcgg ccgactggga gacgctgttc 960
ggccagttct atccgacgca cccggtcgtg ccgggcgagt ggggcggccg ttacggcatc 1020
ggcggtgatc cgcgcgacgt cgcctggcag aacgccttcg tcgactacat gctcggcaag 1080
ggcatgcgcg acagcttcta ctggtgctac acgccgaaca gcggcgacac cggc 1134

<210> 10
<211> 1710
<212> DNA
<213> Organismo del suelo no identificado

<400> 10
atgaggacc acatgaaaac gagcctgaaa ctccgggcat tgggcgcgct ctttgcggcg 60
ctggtcggcg ccgtcgcggc gccggccgcc agcgccgcgg gctattacgc gcagcagggc 120
aagatctacg acgccagcgg ccagcaggta ccgatccgcg gcatcagcca ctacggcttc 180
aatgccaaca tcctgcagcc gcagtatctg tggtcgatgg gctggaagga gcagatcgcg 240
cagatcaagg cgctcggctt caacgccgtg cgtctgccgt acgtgccgga cacgctgtac 300
gtgacgacgc cggtcgacca gctcagctac gtcgatccgg gcaagaacgc ggacctgac 360
ggcaagacgc cgctgcaggg gatggacatg tggatggccg aggccgatcg ccagggcctc 420
tacgtcatgc tcgacttcca cagcgtctcg aatgatcgcc agtatccgac ctggttcgctc 480
gacaacagcg ccgacttcgg cctgatctac aacagctcgg cctacaccgt cgacaactgg 540
atccgcgatc tgaagtctgt cgccgcgcgc tacgccaacc tgccgcactt cttcgccatc 600
gacatctaca acgagccgaa cggcatcgtg cgctggagca ccggcgacgc caacatgacg 660
cagtcgaaga actactggaa ggatgccgcc gagaaggctg ccgccggcgt gcttgcggcc 720
aatccgaacc tgatgatctt cgtgcagggc atcaacggca actgggatgg cgtcgagaag 780
agcaacatcc cgatgaacta cggcgaggac ttccagccgc agcgctacca gccgctcaac 840
atcccggccg acaagctggt gctgagcccc cacacctacg gcccgacgt ctacgtgaag 900
tcgacgttca acgccgcaa cttcccggcc aatctggcgg ccgactggga gacgctgttc 960
ggccagttct atccgacgca cccggtcgtg ccgggcgagt ggggcggccg ttacggcatc 1020
ggcggtgatc cgcgcgacgt cgcctggcag aacgccttcg tcgactacat gctcggcaag 1080
ggcatgcgcg acagcttcta ctggtgctac acgccgaaca gcggcgacac cggcggcatc 1140
ctcgacgaca acctgcaggt gcgcaccgac aagatggcgc tgctgcagaa gctgtggggc 1200

ES 2 545 161 A1

accagcggca ccagcgggc atcgatcgcg ctgtcgtcga ggcctacac ggtcagccag 1260
 agcgcgggtt cggtgagcct gtcggtaaat cgcagcggcg gcacgggccc ggtctcggtc 1320
 agctatgca cgcggcggg cagcggcgtt tccggcaccg acttcacggc caagagcggg 1380
 acgctgaact gggccgccc cgacaccgcg gcgaagacca tcacgattcc ggtcagcaac 1440
 gcgacgccgt tcagcggcag caagacctc accgtctcgc tgtcgggccc cagcagcggc 1500
 gccgcgctcg ccacgccaa catcgccacg gtgacgatca acggttcggc ggtcacggcc 1560
 tcggccggca cgctcgcgt cggcgcgctg agctacagcg tcgcgagaa tgccggcagc 1620
 ctgaccgtca cggtcgcgcg cagcggcggg tcgagcggcg cggtcagcgt caagtacgcc 1680
 accgccaacg gcacggcggg cgccggcacc 1710

<210> 11
 <211> 28
 <212> DNA
 <213> Artificial

<220>
 <223> cebador directo para la amplificación de Cel98

<400> 11
 caatcatatg aggaccaca tgaaaacg 28

<210> 12
 <211> 29
 <212> DNA
 <213> Artificial

<220>
 <223> cebador reverso para la amplificación de Cel 98

<400> 12
 gaatctcgag cttggtgac gtgaaagcc 29

<210> 13
 <211> 22
 <212> DNA
 <213> Artificial

<220>
 <223> cebador T7 fw

<400> 13
 gtaatacgac tcactatagg gc 22

<210> 14
 <211> 18
 <212> DNA
 <213> Artificial

<220>
 <223> cebador UP098intrev

<400> 14
 tctggctgac cgtgtagg 18



- ②① N.º solicitud: 201430313
 ②② Fecha de presentación de la solicitud: 07.03.2014
 ③② Fecha de prioridad:

INFORME SOBRE EL ESTADO DE LA TECNICA

⑤① Int. Cl.: Ver Hoja Adicional

DOCUMENTOS RELEVANTES

Categoría	⑤⑥ Documentos citados	Reivindicaciones afectadas
A	US 20030096342 A1 (ADNEY et al.) 22.05.2003, todo el documento.	1-24
A	ES 2262197 T3 (NOVOZYMES A/S) 16.11.2006, página 2, líneas 5-11; página 3, línea 26 – página 4, línea 13.	1-24
A	WO 2012142171 A2 (ALLOPARTIS BIOTECHNOLOGIES, INC. [US/US]) 18.10.2012, página 3, párrafo [0011] – página 5 párrafo [0017].	1-24
A	LYKIDIS A. et al. Genome Sequence and Analysis of the Soil Cellulolytic Actinomycete Thermobifida fusca. Journal of Bacteriology. 2007, Vol. 189(6), páginas: 2477-2486.	1-24

Categoría de los documentos citados

X: de particular relevancia
 Y: de particular relevancia combinado con otro/s de la misma categoría
 A: refleja el estado de la técnica

O: referido a divulgación no escrita
 P: publicado entre la fecha de prioridad y la de presentación de la solicitud
 E: documento anterior, pero publicado después de la fecha de presentación de la solicitud

El presente informe ha sido realizado

para todas las reivindicaciones

para las reivindicaciones nº:

Fecha de realización del informe
09.06.2014

Examinador
M. D. García Grávalos

Página
1/4

CLASIFICACIÓN OBJETO DE LA SOLICITUD

C12N9/42 (2006.01)

C12N1/21 (2006.01)

C12N15/74 (2006.01)

Documentación mínima buscada (sistema de clasificación seguido de los símbolos de clasificación)

C07K, C12N

Bases de datos electrónicas consultadas durante la búsqueda (nombre de la base de datos y, si es posible, términos de búsqueda utilizados)

INVENES, EPODOC, WPI, NPL, BIOSIS, MEDLINE, EMBASE, USPTO PATENT DATABASE, PUBMED, GOOGLE PATENTS, GOOGLE SCHOLAR, EBI, STN.

Fecha de Realización de la Opinión Escrita: 09.06.2014

Declaración

Novedad (Art. 6.1 LP 11/1986)	Reivindicaciones 1-24	SI
	Reivindicaciones	NO
Actividad inventiva (Art. 8.1 LP11/1986)	Reivindicaciones 1-24	SI
	Reivindicaciones	NO

Se considera que la solicitud cumple con el requisito de aplicación industrial. Este requisito fue evaluado durante la fase de examen formal y técnico de la solicitud (Artículo 31.2 Ley 11/1986).

Base de la Opinión.-

La presente opinión se ha realizado sobre la base de la solicitud de patente tal y como se publica.

1. Documentos considerados.-

A continuación se relacionan los documentos pertenecientes al estado de la técnica tomados en consideración para la realización de esta opinión.

Documento	Número Publicación o Identificación	Fecha Publicación
D01	US 20030096342 A1	22.05.2003
D02	ES 2262197 T3	16.11.2006
D03	WO 2012142171 A2	18.10.2012
D04	LYKIDIS A. et al. Journal of Bacteriology. 2007, Vol. 189(6), páginas: 2477-2486.	2007

2. Declaración motivada según los artículos 29.6 y 29.7 del Reglamento de ejecución de la Ley 11/1986, de 20 de marzo, de Patentes sobre la novedad y la actividad inventiva; citas y explicaciones en apoyo de esta declaración

El objeto técnico de la presente invención es un péptido, definido por la SEQ ID NO: 1, que presenta actividad celulasa (reivindicaciones 1-18), así como su utilidad en procesos de degradación de la biomasa y en procedimientos de producción de azúcares fermentables para obtención de bioproductos (reivindicaciones 19-24).

El documento D01 divulga una enzima glucosidasa, *Gux1*, aislada de la bacteria *Acidothermus cellulolyticus*; formas recombinantes de esta enzima, incluyendo polipéptidos, fusiones, variaciones y derivados y su uso para degradación de carbohidratos (ver todo el documento).

El documento D02 divulga una enzima, una endoglucanasa con actividad celulolítica a alta temperatura, aislada de la bacteria *Dictyoglomus* sp. Se refiere también a una secuencia de ADN que codifica dicha enzima y a un método para obtenerla; a una composición que la comprende y al uso de la enzima y de la composición en aplicaciones industriales diversas (ver página 2, líneas 5-11; página 3, línea 26 - página 4, línea 13).

El documento D03 divulga nuevas enzimas con actividad celulolítica aisladas de *Thermobifida fusca* y su uso para degradación de la biomasa (ver página 3, párrafo [0011] - página 5, párrafo [0017]).

El documento D04 divulga la identificación de enzimas con actividad celulasa en el genoma de *Thermobifida fusca* (todo el documento).

1. NOVEDAD Y ACTIVIDAD INVENTIVA (Art. 6.1 y Art. 8.1 LP 11/1986)

El objeto técnico de la presente invención es un péptido, definido por la SEQ ID NO: 1, que presenta actividad celulasa, así como su utilidad en procesos de degradación de la biomasa y en procedimientos de producción de azúcares fermentables para obtención de bioproductos.

1.1. REIVINDICACIONES 1-24

El uso de péptidos con actividad celulasa para degradación de carbohidratos o biomasa es conocido en el estado de estado de la técnica. De hecho, los documentos D01-D04 anticipan enzimas glucosidasas y su uso para degradación de carbohidratos y aplicaciones industriales diversa. Sin embargo, las secuencias correspondientes a estas enzimas no coinciden con las reivindicadas en la presente solicitud.

La presente invención reivindica un péptido, definido por la SEQ ID NO: 1, que presenta actividad celulasa. Dicho péptido no ha sido encontrado en el estado de la técnica por lo que se considera que la invención proporciona una enzima nueva con actividad celulasa para uso en procesos de degradación de la biomasa y en procedimientos de producción de azúcares fermentables para obtención de bioproductos.

En consecuencia, las reivindicaciones 1-24 cumplen el requisito de novedad y actividad inventiva (Art. 6.1 y Art. 8.1 LP 11/1986).