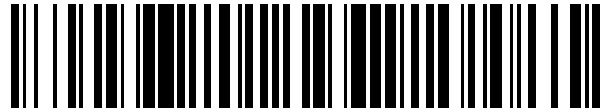


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 530 518**

51 Int. Cl.:

C12P 21/06 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **28.04.2004 E 04750928 (6)**

97 Fecha y número de publicación de la concesión europea: **19.11.2014 EP 1618183**

54 Título: **Nueva celulasa 029cel de Bacillus**

30 Prioridad:

29.04.2003 US 466831 P

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

03.03.2015

73 Titular/es:

**DANISCO US INC. (100.0%)
925 Page Mill Road
Palo Alto, California 94304 , US**

72 Inventor/es:

**JONES, BRIAN, E.;
GRANT, WILLIAM , D.;
HEAPHY, SHAUN y
GRANT, SUSAN**

74 Agente/Representante:

DE ELZABURU MÁRQUEZ, Alberto

ES 2 530 518 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín europeo de patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre concesión de Patentes Europeas).

DESCRIPCIÓN

Nueva celulasa 029cel de *Bacillus*

Campo de la invención

5 Esta invención se refiere a una nueva celulasa a la que se hace referencia en esta memoria como 029cel. También se describen ácidos nucleicos que codifican la celulasa, composiciones que comprenden dicha celulasa, métodos para identificar nuevas celulasas y métodos para utilizar dichas composiciones. Preferiblemente, la(s) celulasa(s) se aísla(n) de especies del género *Bacillus*, preferiblemente de *B. agaradhaerens*. La presente invención se refiere además al uso de la nueva celulasa en composiciones reconocidas en la técnica como ventajosas por tener una
10 celulasa añadida a ellas, incluyendo como un aditivo en una composición detergente, en el tratamiento de tejidos que contienen celulosa, en el tratamiento de pulpa y papel y en el tratamiento de almidón para la producción de etanol o jarabe de maíz con alto contenido de fructosa.

Antecedentes de la invención

15 La celulosa y la hemicelulosa son los más abundantes materiales vegetales producidos por fotosíntesis. Pueden ser degradadas y empleadas como fuente de energía por numerosos microorganismos, incluyendo bacterias, levaduras y hongos, que producen enzimas extracelulares capaces de la hidrólisis de los sustratos poliméricos hasta azúcares monoméricos (aro et al., 2001). Puesto que nos aproximamos a los límites de los recursos no renovables, el potencial de la celulosa para llegar a ser un importante recurso energético renovable es enorme (Krishna et al., 2001). La utilización eficaz de la celulosa a través de procesos biológicos es un planteamiento para superar la escasez de alimentos, piensos y combustibles (Ohmiya et al. 1997).

20 Las celulasas son enzimas que hidrolizan celulosa (enlaces beta-1,4-glucano o beta-D-glucosídicos) para dar lugar a la formación de glucosa, celobiosa, celo-oligosacáridos y similares. Las celulasas han sido tradicionalmente divididas en tres clases principales: endoglucanasas (EC 3.2.1.4) ("EG"), exoglucanasas o celobiohidrolasas (EC 3.2.1.91) ("CBH") y beta-glucosidasas ([beta]-D-glucósido glucohidrolasa; EC 3.2.1.21) ("BG") (Knowles et al., 1987; Shulein, 1988). Las endoglucanasas actúan principalmente sobre las partes amorfas de la fibra de celulosa, mientras que las celobiohidrolasas también son capaces de degradar la celulosa cristalina (Nevalainen y Penttila, 1995). Por lo tanto, para la solubilización eficaz de la celulosa cristalina se necesita la presencia de una celobiohidrolasa en un sistema de celulasas (Suurnakki et al., 2000). La beta-glucosidasa actúa para liberar unidades de D-glucosa de la celobiosa, los celo-oligosacáridos y otros glucósidos (Freer, 1993).

30 Con objeto de convertir eficazmente la celulosa cristalina en glucosa, se requiere el sistema de celulasas completo que comprende componentes de cada una de las clasificaciones CBH, EG y BG, con componentes aislados menos eficaces en cuanto a hidrolizar la celulosa cristalina (Filho et al., 1996). Se ha observado una relación sinérgica entre los componentes de celulasa de las diferentes clasificaciones. En particular, las celulasas de tipo EG y las celulasas de tipo CBH interactúan sinérgicamente para degradar más eficazmente la celulosa. Véase, por ejemplo, Wood, 1985.

35 Aunque se han descrito previamente composiciones de celulasa, persiste la necesidad de nuevas y mejoradas composiciones de celulasa para uso en detergentes domésticos, composiciones para lavado a la piedra o detergentes para la colada, etcétera. Son particularmente interesantes las celulasas que presentan unas cualidades mejoradas.

Breve sumario de la invención

40 Un objeto de la presente invención es proporcionar una nueva celulasa que tenga propiedades beneficiosas para uso en detergentes, el tratamiento de textiles, la conversión de biomasa y la fabricación de pulpa y papel.

Un objeto de la presente invención es proporcionar polipéptidos que tengan actividad celulolítica y polinucleótidos que codifiquen los polipéptidos. Los polipéptidos pueden mejorar la degradación del material de la pared celular, por ejemplo, la celulosa y/o la hemicelulosa. Los polipéptidos pueden también mejorar la estabilidad o actividad de otras
45 enzimas implicadas en la degradación del material de la pared celular de plantas, por ejemplo, la biomasa.

Un objeto de la presente invención es proporcionar una nueva celulasa y derivados de la misma, métodos para producir dichas celulasas, y composiciones que comprendan dichas nuevas celulasas. La presente invención se refiere además al uso de la nueva celulasa y derivados de la misma en composiciones reconocidas en la técnica como ventajosas por tener una celulasa añadida a ellas, incluyendo como un aditivo en una composición detergente,
50 en el tratamiento de textiles tales como tejidos que contienen celulosa y fibras útiles para los mismos, como un aditivo para pienso para animales, en la conversión de biomasa, en el tratamiento de pulpa y papel y en el tratamiento de almidón para la producción de etanol o jarabe de maíz con alto contenido de fructosa.

Otro objeto de la presente invención es proporcionar un método para producir una nueva celulasa a través de la expresión heteróloga desde células huésped recombinantes.

Aún otro objeto de la presente invención es proporcionar una secuencia de ácido nucleico que codifique la celulasa de la invención. En un aspecto, el ácido nucleico y la secuencia de aminoácidos facilitan la producción comercial de la nueva celulasa y las composiciones de celulasa de la invención.

5 Aún otro objeto de la presente invención es proporcionar una nueva celulasa que tenga unas propiedades excelentes para uso en detergentes, el tratamiento de textiles, como un suplemento para pienso y en la fabricación de pulpa y papel. En un aspecto más, la celulasa resulta útil en la conversión de biomasa.

10 En un primer aspecto, la invención incluye un polinucleótido aislado que tiene una secuencia que codifica 029cel, una secuencia complementaria de la secuencia de codificación del *gen 029cel*, y una composición que comprende el polinucleótido. El polinucleótido puede ser mRNA, DNA, cDNA, DNA genómico, o un análogo antisentido de los mismos.

En una realización, un polinucleótido *029cel* puede comprender una molécula de ácido nucleico aislada que se hibrida con el complemento del ácido nucleico presentado como ID. SEC. nº 2 bajo condiciones de rigor moderado a alto, donde la molécula de ácido nucleico codifica un polipéptido 029cel que presenta actividad ligante de celulosa.

15 El polinucleótido que tiene una identidad secuencial de al menos 80%, 85%, 90%, 95%, 98% o más con respecto a la secuencia presentada como ID. SEC. nº 2 puede codificar una proteína 029cel. En una realización específica, el polinucleótido comprende una secuencia sustancialmente idéntica a la ID. SEC. nº 2.

20 En un segundo aspecto, se proporciona una nueva celulasa o un derivado que son obtenibles de una especie de *Bacillus*. Preferiblemente, la celulasa de la invención comprende una secuencia de aminoácidos de acuerdo con la Figura 3 (ID. SEC. nº 3), o un derivado de la misma, que tiene una identidad secuencial superior al 90%, preferiblemente una identidad secuencial superior a 95% y más preferiblemente una identidad secuencial superior al 97% con respecto a una porción activa de la misma.

En un tercer aspecto, la presente invención se refiere a una construcción de ácido nucleico que comprende la secuencia de nucleótidos, que codifica el polipéptido de la invención, operativamente unida a una o más secuencias de control que dirigen la producción del polipéptido en un huésped adecuado.

25 La invención proporciona además vectores de expresión recombinantes que contienen una secuencia de ácido nucleico que codifica 029cel o una variante de la misma por ajuste, operativamente unida a elementos reguladores eficaces para la expresión de la proteína en un huésped seleccionado. En un aspecto relacionado, la invención incluye una célula huésped que contiene el vector.

30 En un cuarto aspecto, la presente invención se refiere a un vector de expresión recombinante que comprende la construcción de ácido nucleico de la invención.

En un quinto aspecto, la presente invención se refiere a una célula huésped recombinante que comprende la construcción de ácido nucleico de la invención.

35 La invención incluye además un método para producir 029cel mediante técnicas recombinantes, mediante el cultivo de células huésped procarióticas o eucarióticas recombinantes que comprenden una secuencia de ácido nucleico que codifica 029cel bajo unas condiciones eficaces para promover la expresión de la proteína, y la recuperación posterior de la proteína de la célula huésped o del medio de cultivo celular.

En un sexto aspecto, la presente invención se refiere a un método para producir un polipéptido de la invención, método que comprende: (a) cultivar un microorganismo, que en su forma de tipo silvestre es capaz de producir el polipéptido, para producir el polipéptido; y (b) recuperar el polipéptido.

40 En un séptimo aspecto, la invención proporciona una composición enzimática útil en la conversión de celulosa en etanol. En una realización preferida, la composición enzimática comprende 029cel. La composición puede comprender además enzimas celulasas adicionales tales como endoglucanasas y/o celobiohidrolasas. La composición puede estar enriquecida en 029cel.

45 En una realización, la invención proporciona un método para identificar nuevas enzimas aislando el DNA total de la comunidad microbiana de un entorno, construyendo un banco de DNA genómico en *E. coli*, explorando el banco en cuanto a la expresión de actividad celulasa, identificando el gen de celulasa en el clon positivo para celulasa y caracterizando la nueva enzima celulasa.

Se proporcionan además en esta memoria métodos analíticos para detectar ácidos nucleicos *029cel* y proteínas 029cel, que también forman parte de la invención.

50 Según aún otra realización de la invención, se proporciona un método para transformar un microorganismo adecuado con una secuencia de ácido nucleico que codifica una celulasa de acuerdo con la invención. Se proporciona un método para producir la celulasa de acuerdo con la invención a partir de dicho microorganismo transformado.

Un objeto más de la invención es proporcionar un vector de expresión particularmente eficaz en *Streptomyces*. Las bacterias del género *Streptomyces* sirven como células huésped alternas para la producción de diversas proteínas y, con respecto a la expresión y producción de celulasas, pueden ofrecer diversas ventajas con respecto a las células huésped del género *Bacillus*, particularmente cuando las células se cultivan con una densidad celular elevada. Un vector de expresión preferido comprende una secuencia polinucleotídica reguladora que incluye una secuencia promotora derivada de un gen de glucosa isomerasa de *Actinoplanes*, una secuencia señal derivada de un gen de celulasa de *Streptomyces*, y una secuencia de DNA que codifica una celulasa, particularmente una celulasa de acuerdo con la invención.

En una realización preferida de la presente invención, una celulasa de longitud completa es obtenible de bacterias del género *Bacillus*.

Otros objetos, características y ventajas de la presente invención se harán evidentes a partir de la siguiente descripción detallada. Sin embargo, se debería entender que la descripción detallada y los ejemplos específicos, aunque indican realizaciones preferidas de la invención, sólo se proporcionan a modo de ilustración, ya que diversos cambios y modificaciones dentro del alcance y el espíritu de la invención se harán evidentes a un experto en la técnica a partir de esta descripción detallada.

Breve descripción de los dibujos

La Figura 1 ilustra la secuencia nucleotídica ambiental (ID. SEC. nº 1).

La Figura 2 ilustra una secuencia de ácido nucleico que codifica la nueva celulasa (ID. SEC. nº 2).

La Figura 3 ilustra la deducida secuencia de aminoácidos de la celulasa de la invención (ID. SEC. nº 3).

Descripción detallada

La invención será ahora descrita en detalle sólo a modo de referencia usando las definiciones y ejemplos siguientes.

A menos que se defina otra cosa en esta memoria, todos los términos técnicos y científicos empleados en esta memoria tienen un significado igual al comúnmente entendido por quien tiene una experiencia normal en la técnica a la cual pertenece esta invención. Singleton et al., *DICTIONARY OF MICROBIOLOGY AND MOLECULAR BIOLOGY*, 2ª edición, John Wiley and Sons, New York (1994), y Hale y Marham, *THE HARPER COLLINS DICTIONARY OF BIOLOGY*, Harper Perennial, New York (1991), proporcionan a alguien con experiencia un diccionario general de muchos de los términos empleados en esta invención. Aunque en la práctica o el examen de la presente invención se pueden emplear métodos y materiales similares o equivalentes a los descritos en esta memoria, se describen los métodos y materiales preferidos. Los intervalos numéricos incluyen los números que definen el intervalo. A menos que se indique otra cosa, los ácidos nucleicos se escriben de izquierda a derecha en orientación 5' a 3'; las secuencias de aminoácidos se escriben de izquierda a derecha en orientación amino a carboxilo, respectivamente. Los profesionales son particularmente dirigidos a Sambrook et al., 1989, y F. M. Ausubel et al., 1993, para las definiciones y términos de la técnica. Se ha de entender que esta invención no se limita a la metodología, los protocolos y los reactivos particulares descritos, ya que estos pueden variar.

Los encabezamientos proporcionados en esta memoria no son limitaciones de los distintos aspectos o realizaciones de la invención que se pueden tener por referencia a la memoria descriptiva en su totalidad. En consecuencia, los términos definidos inmediatamente a continuación se definen más detalladamente por referencia a la memoria descriptiva en su totalidad.

Todas las publicaciones citadas en esta memoria se incorporan expresamente a esta memoria por referencia con la finalidad de describir y revelar composiciones y metodologías que podrían ser empleadas en relación con la invención.

I. Definiciones

"Celulasa", "enzimas celulolíticas" o "enzimas celulasas" significa la endoglucanasa bacteriana del invento descrita en esta memoria. Tres tipos diferentes de enzimas celulasas actúan sinérgicamente para convertir la celulosa y sus derivados en glucosa.

El término "celulasa" se refiere a una categoría de enzimas capaces de hidrolizar polímeros de celulosa en oligómeros celo-oligosacáridos más cortos, celobiosa y/o glucosa. Se han obtenido numerosos ejemplos de celulasas, tales como exoglucanasas, exocelobiohidrolasas, endoglucanasas y glucosidasas, de organismos celulolíticos, incluyendo particularmente hongos y bacterias. Las enzimas producidas por estos microbios son mezclas de proteínas con tres tipos de acciones útiles en la conversión de celulosa en glucosa: endoglucanasas (EG), celobiohidrolasas (CBH) y beta-glucosidasa. Estos tres tipos diferentes de enzimas celulasas actúan sinérgicamente para convertir la celulosa y sus derivados en glucosa.

Muchos microbios producen enzimas que hidrolizan la celulosa, incluyendo el hongo del género *Trichoderma* que

causa podredumbre de la madera, las bacterias del compost de los géneros *Thermomonospora*, *Bacillus* y *Cellulomonas*; las bacterias del género *Streptomyces*; y los hongos de los géneros *Humicola*, *Aspergillus* y *Fusarium*.

Mediante la expresión "célula huésped" se quiere significar una célula que contiene un vector y sustenta la replicación, y/o la transcripción o la transcripción y la traducción (expresión), de la construcción de expresión. Las células huésped para uso en la presente invención pueden ser células procarióticas, tales como células de *E. coli*, o células eucarióticas tales como células de levadura, vegetal, insecto, anfibio o mamífero. En una realización de acuerdo con la presente invención, "célula huésped" significa las células del género *Bacillus*. En otra realización preferida de acuerdo con la invención, "célula huésped" significa las células del género *Streptomyces*. Un *Streptomyces* significa cualquier cepa bacteriana que es un miembro del género *Streptomyces*, según se clasifica en Buchanan et al., *The Shorter Bergey's Manual for Determinative Bacteriology* (Williams & Wilkins, 1982). Las cepas particularmente preferidas del género *Streptomyces* incluyen *S. lividans*, *S. rubiginosus* y *S. coelicolor*. Se describe *S. lividans* en Lomovskaya et al., *J. Virology* 9: 258 (1972). Sin embargo, quien tiene experiencia se dará cuenta de que se puede emplear cualquier célula huésped apropiada, tal como, por ejemplo, una célula bacteriana, fúngica, eucariótica o vegetal.

El término "recombinante", cuando se utiliza con referencia, por ejemplo, a una célula o a un ácido nucleico, una proteína o un vector, indica que la célula, el ácido nucleico, la proteína o el vector ha sido modificado mediante la introducción de un ácido nucleico o una proteína heterólogos o mediante la alteración de un ácido nucleico o una proteína nativos, o que la célula procede de una célula así modificada. De este modo, por ejemplo, las células recombinantes expresan genes que no se encuentran en la forma nativa (no recombinante) de la célula o expresan genes nativos que, en cualquier caso, se expresan anormalmente, se infraexpresan o no se expresan en absoluto.

La expresión "secuencias señal secretora" significa una secuencia de DNA que codifica un polipéptido (un "péptido secretor" o "péptido señal secretor") que, como un componente de un polipéptido más grande, dirige al polipéptido más grande a través de una ruta secretora de una célula en que se sintetiza. El péptido más grande es comúnmente escindido para que se separe el péptido secretor durante el tránsito a través de la ruta secretora, produciéndose el péptido señal secretor y un péptido más pequeño al que comúnmente se hace referencia como "polipéptido maduro".

Como se emplean en esta memoria, las frases "preparación completa de celulasas" y "composición completa de celulasas" se utilizan indistintamente y se refieren a composiciones tanto de origen natural como de origen no natural. Una composición "de origen natural" es una producida por una fuente de origen natural y que comprende, por ejemplo, uno o más componentes de tipo celobiohidrolasa, uno o más componentes de tipo endoglucanasa y uno o más componentes de β -glucosidasa, en donde cada uno de estos componentes se encuentra en la proporción producida por la fuente. Ciertos hongos producen sistemas completos de celulasas que incluyen exocelobiohidrolasas o celulasas de tipo CBH, endoglucanasas o celulasas de tipo EG y beta- glucosidasas o celulasas de tipo BG (Schulein, 1988). Sin embargo, a veces estos sistemas carecen de celulasas de tipo CBH, y las celulasas bacterianas también incluyen típicamente pocas celulasas de tipo CBH o ninguna. Una composición de origen natural es una que es producida por un organismo no modificado con respecto a las enzimas celulolíticas de modo que la proporción de las enzimas componentes es indistinta de la producida por el organismo nativo.

Una composición de "origen no natural" abarca las composiciones producidas al: (1) combinar enzimas celulolíticas componentes en una proporción de origen natural o en una proporción de origen no natural, es decir, alterada; o (2) modificar un organismo para que sobreexpresen o infraexpresen una o más enzimas celulolíticas; o (3) modificar un organismo para que se suprima al menos una enzima celulolítica; o (4) modificar un organismo para que exprese una enzima celulolítica componente heteróloga.

Como se emplea en esta memoria, el término "promotor" se refiere a una secuencia de ácido nucleico que actúa para dirigir la transcripción de un gen cadena abajo. El promotor será generalmente apropiado para la célula huésped en la que se está expresando el gen diana. El promotor junto con otras secuencias de ácido nucleico reguladoras de la transcripción y la traducción (también denominadas "secuencias de control") son necesarios para que se exprese un gen dado. En general, las secuencias reguladoras de la transcripción y la traducción incluyen, pero no se limitan a, secuencias promotoras, sitios de unión al ribosoma, secuencias de inicio y parada de la transcripción, secuencias de inicio y parada de la traducción, y secuencias potenciadoras o activadoras. El promotor puede ser el promotor normalmente asociado con el gen cadena abajo o puede ser heterólogo, es decir, de otro gen u otro microorganismo con tal de que actúe para dirigir el gen. Cuando la célula huésped de transformación es del género *Bacillus*, un promotor preferido es el promotor *aprE*. En un aspecto, el promotor es un promotor inducible. En un aspecto, cuando la célula huésped es un hongo filamentoso, el promotor es el promotor *cbh1* de *T. reesei*, que está depositado en GenBank bajo el Número de Acceso D86235. En otro aspecto, el promotor es un promotor *cbh II* o de xilanasas de *T. reesei*.

Un ácido nucleico está "operativamente unido" cuando está situado en una relación funcional con otra secuencia de ácido nucleico. Por ejemplo, un DNA que codifica un líder secretor, es decir, un péptido señal, está operativamente unido al DNA para un polipéptido si se expresa como una preproteína que participa en la secreción del polipéptido; un promotor o un potenciador está operativamente unido a una secuencia de codificación si afecta a la transcripción de la secuencia; o un sitio de unión al ribosoma está operativamente unido a una secuencia de codificación si está

colocado de modo que facilita la traducción. En general, "operativamente unido" significa que las secuencias de DNA que están unidas son contiguas, y, en el caso de un líder secretor, contiguas y en fase de lectura. Sin embargo, los potenciadores no tienen que ser contiguos. La unión se lleva a cabo por ligación en sitios de restricción convenientes. Si dichos sitios no existen, se usan los adaptadores o conectores oligonucleotídicos sintéticos de acuerdo con la práctica convencional.

"Construcción de DNA" o "vector de DNA" significa una secuencia de nucleótidos que comprende uno o más fragmentos de DNA que codifican la nueva celulasa. Los "vectores de expresión" están incluidos en los "vectores de DNA". Los vectores de expresión típicos contienen secuencias reguladoras tales como terminadores de la transcripción y la traducción, secuencias de iniciación de la transcripción y la traducción, secuencias señal, y promotores útiles para la regulación de la expresión del ácido nucleico particular. El término "promotor" se emplea en su sentido ordinario para referirse a una secuencia polinucleotídica implicada en el control de la iniciación de la transcripción de una secuencia polinucleotídica que codifica una proteína. Una "secuencia señal" se refiere a un péptido señal o una porción de una proteína que es capaz de dirigir el transporte de una proteína deseada en forma bioactiva desde un huésped. La forma madura de una proteína extracelular carece de la secuencia señal, que es escindida durante el proceso de secreción. Aunque sin pretender limitar la invención, el número de restos de aminoácido en un péptido señal puede ser entre aproximadamente 5 y aproximadamente 100 restos de aminoácido. La secuencia señal puede ser modificada para obtener sitios de clonación que permitan la ligación de DNA o la inserción de DNA que codifica una celulasa. Los vectores comprenden opcionalmente casetes de expresión genéricos que contienen al menos una secuencia terminadora independiente, secuencias que permiten la replicación del casete en procariontes, eucariontes o ambos (por ejemplo, vectores lanzadera), y marcadores de selección para sistemas tanto procarióticos como eucarióticos. Los vectores son adecuados para la replicación e integración en procariontes, eucariontes o ambos. Véanse Gilman y Smith, *Gene* 8: 81-97 (1979); Roberts *et al.*, *Nature* 328: 731-734 (1987); Berger y Kimmel, *GUIDE TO MOLECULAR CLONING TECHNIQUES, METHODS IN ENZYMOLOGY*, Volumen 152, Academic Press, Inc., San Diego, California, EE.UU. ("Berger"); B. Scheider *et al.*, *Protein Expr. Purif.* 6435: 10 (1995); Sambrook *et al.*, *MOLECULAR CLONING - A LABORATORY MANUAL* (2ª edición), Volúmenes 1-3, Cold Springs Harbor Publishing (1989) ("Sambrook"); y *CURRENT PROTOCOLS IN MOLECULAR BIOLOGY*, Ausubel *et al.* (redactores), Current Protocols, una empresa mixta entre Greene Publishing Associates, Inc., y John Wiley & Sons, Inc. (1997, Suplemento) ("Ausubel"). Se conocen vectores de clonación útiles en *Streptomyces*, y se hace referencia a ellos en las Patentes de EE.UU. números 4.338.397, 4.411.994, 4.513.085, 4.513.086, 4.745.056, 5.514.590 y 5.622.866 y el Documento WO88/07079.

Como se emplea en esta memoria, el término "gen" significa el segmento de DNA implicado en producir una cadena polipeptídica, que puede incluir o no regiones que precedan y sigan a la región de codificación, por ejemplo, secuencias 5' no traducidas (5' UTR; del inglés, 5' *untranslated region*) o "líder" y secuencias 3' UTR o "tráiler", así como secuencias intermedias (intrones) entre segmentos de codificación individuales (exones).

El término "heterólogo", cuando se utiliza con referencia a porciones de un ácido nucleico, indica que el ácido nucleico comprende dos o más subsecuencias que no se encuentran normalmente en la misma relación mutua en la naturaleza. Por ejemplo, el ácido nucleico se produce típicamente de forma recombinante, teniendo dos o más secuencias, por ejemplo, de genes no relacionados colocadas para crear un nuevo ácido nucleico funcional, tal como, por ejemplo, un promotor de una fuente y una región de codificación de otra fuente. Similarmente, una proteína heteróloga se referirá a menudo a dos o más subsecuencias que no se hallan en la misma relación mutua en la naturaleza (por ejemplo, una proteína de fusión).

La expresión "% de homología" se emplea indistintamente en esta memoria con la expresión "% de identidad" y se refiere al nivel de identidad de secuencias de ácido nucleico o de aminoácidos entre la secuencia de ácido nucleico que codifica O29cel o la secuencia de aminoácidos de O29cel cuando son alineadas utilizando un programa para alineación de secuencias.

Por ejemplo, como se emplea en esta memoria, una homología de 80% significa la misma cosa que una identidad secuencial de 80% determinada por un algoritmo definido, y, en consecuencia, un compuesto homólogo de una secuencia dada tiene una identidad secuencial superior al 80% con respecto a un tramo de la secuencia dada. Los niveles ejemplares de identidad secuencial incluyen, pero no se limitan a, una identidad secuencial de 80, 85, 90, 95, 98% o más con respecto a una secuencia dada, por ejemplo, la secuencia de codificación para O29cel, como se describe en esta memoria.

Los programas informáticos ejemplares que se pueden utilizar para determinar la identidad entre dos secuencias incluyen, pero no se limitan a, el juego de programas BLAST, por ejemplo, BLASTN, BLASTX y TBLASTX, BLASTP y TBLASTN, públicamente asequibles en Internet en www.ncbi.nih.gov/BLAST. Véanse también Altschul *et al.*, 1990, y Altschul *et al.*, 1997.

Las búsquedas de secuencias se llevan típicamente a cabo usando el programa BLASTN cuando se evalúa una secuencia de ácido nucleico dada en relación con secuencias de ácido nucleico en GenBank DNA Sequences y otras bases de datos públicas. Se prefiere el programa BLASTX para la búsqueda de secuencias de ácido nucleico que han sido traducidas en todos los marcos de lectura frente a secuencias de aminoácidos en GenBank Protein Sequences y otras bases de datos públicas. Tanto BLASTN como BLASTX se ejecutan usando los parámetros por

omisión de una penalización por hueco abierto de 11,0 y una penalización por hueco extendido de 1,0, y se utiliza la matriz BLOSUM-62 (véanse, por ejemplo, Altschul *et al.*, 1997).

5 Con objeto de determinar el "% de identidad" entre dos o más secuencias se lleva a cabo una alineación preferida de las secuencias seleccionadas utilizando, por ejemplo, el programa CLUSTAL-W en MacVector versión 6.5, ejecutado con los parámetros por omisión, que incluyen una penalización por hueco abierto de 10,0 y una penalización por hueco extendido de 0,1 y una matriz de similitudes BLOSUM 30.

Los términos "aislado" y "purificado", como se emplean en esta memoria, se refieren a un ácido nucleico o aminoácido que es separado de al menos un componente con el que se asocia naturalmente.

10 En el presente contexto, la expresión "polipéptido sustancialmente puro" significa una preparación polipeptídica que contiene a lo sumo 10% en peso de otro material polipeptídico con el que se asocia nativamente (se prefieren porcentajes más pequeños del otro material polipeptídico, por ejemplo, a lo sumo 8% en peso, a lo sumo 6% en peso, a lo sumo 5% en peso, a lo sumo 4% en peso, a lo sumo 3% en peso, a lo sumo 2% en peso, a lo sumo 1% en peso, y a lo sumo 0,5% en peso). Por lo tanto, se prefiere que el polipéptido sustancialmente puro tenga una pureza de al menos 92%, es decir, que el polipéptido constituya al menos el 92% en peso del material polipeptídico total presente en la preparación, y se prefieren porcentajes mayores tales como una pureza de al menos 94%, una pureza de al menos 95%, una pureza de al menos 96%, una pureza de al menos 97%, una pureza de al menos 98%, una pureza de al menos 99% y una pureza de a lo sumo 99,5%. Los polipéptidos descritos en esta memoria están preferiblemente en una forma sustancialmente pura. En particular, se prefiere que los polipéptidos descritos en esta memoria estén en "forma esencialmente pura", es decir, que la preparación polipeptídica esté esencialmente exenta de otro material polipeptídico con el que se asocia nativamente. Esto puede ser llevado a cabo, por ejemplo, preparando el polipéptido por medio de métodos recombinantes bien conocidos. En esta memoria, la expresión "polipéptido sustancialmente puro" es sinónima de las expresiones "polipéptido aislado" y "polipéptido en forma aislada".

25 En general, las moléculas de ácido nucleico que codifican la 029cel se hibridarán con la secuencia proporcionada en esta memoria como ID. SEC. n° 2 (la 029cel) bajo condiciones de rigor moderado a elevado. Sin embargo, en ciertos casos se emplea una secuencia nucleotídica que codifica 029cel y posee una utilización de codones sustancialmente diferente, aunque la proteína codificada por la secuencia nucleotídica que codifica 029cel tiene la misma, o sustancialmente la misma, secuencia de aminoácidos que la proteína nativa. Por ejemplo, la secuencia de codificación puede ser modificada para facilitar una expresión más rápida de 029cel en un sistema de expresión procariontico o eucariótico particular, de acuerdo con la frecuencia con que un codón particular es utilizado por el huésped. Por ejemplo, Te'o *et al.* (2000) describen la optimización de genes para expresión en hongos filamentosos.

35 Se considera que una secuencia de ácido nucleico es "selectivamente hibridable" con una secuencia de ácido nucleico de referencia si las dos secuencias se hibridan específicamente entre sí bajo unas condiciones de hibridación y lavado de rigor moderado a elevado. Las condiciones de hibridación se basan en la temperatura de fusión (T_m) de la sonda o complejo ligante de ácido nucleico. Por ejemplo, el "rigor máximo" tiene típicamente lugar a aproximadamente $T_m - 5$ °C (5 °C por debajo de la T_m de la sonda); el "rigor elevado" a aproximadamente 5-10 °C por debajo de la T_m ; el "rigor moderado o intermedio" a aproximadamente 10-20 °C por debajo de la T_m de la sonda; y el "rigor bajo" a aproximadamente 20-25 °C por debajo de la T_m . Funcionalmente, se pueden emplear las condiciones de rigor máximo para identificar secuencias que tienen una identidad absoluta o una identidad casi absoluta con la sonda de hibridación, mientras que las condiciones de rigor elevado se utilizan para identificar secuencias que tienen una identidad secuencial de aproximadamente 80% o más con la sonda.

45 Las condiciones de hibridación de rigores moderado y elevado son bien conocidas en la técnica (véanse, por ejemplo, Sambrook *et al.*, 1989, capítulos 9 y 11, y F. M. Ausubel *et al.*, 1993). Un ejemplo de condiciones de rigor elevado incluye hibridación a aproximadamente 42 °C en formamida al 50%, SSC 5X, disolución de Denhardt 5X, SDS al 0,5% y 100 µg/ml de DNA portador desnaturalizado, seguida de lavado dos veces en SSC 2X y SDS al 0,5% a temperatura ambiental y dos veces más en SSC 0,1X y SDS al 0,5% a 42 °C.

Como se emplean en esta memoria, las expresiones "transformada", "establemente transformada" y "transgénica" con referencia a una célula significan que la célula tiene una secuencia de ácido nucleico no nativa (heteróloga) integrada en su genoma o como un plásmido episómico que se mantiene a lo largo de múltiples generaciones.

50 Como se emplea en esta memoria, el término "expresión" se refiere al proceso mediante el cual se produce un polipéptido basándose en la secuencia de ácido nucleico de un gen. El proceso incluye tanto transcripción como traducción.

55 En el contexto de insertar una secuencia de ácido nucleico en una célula, el término "introducida" significa "transfección" o "transformación" o "transducción" e incluye la referencia a la incorporación de una secuencia de ácido nucleico a una célula eucariótica o procariótica, donde la secuencia de ácido nucleico puede ser incorporada al genoma de la célula (por ejemplo, un cromosoma, plásmido, plástido o DNA mitocondrial), convertida en un replicón autónomo, o expresada transitoriamente (por ejemplo, mRNA transfectado).

De esto resulta que la frase "expresión de 029cel" se refiere a la transcripción y traducción del gen de la celulasa 029cel, cuyos productos incluyen RNA precursor, mRNA, polipéptido, y polipéptidos postraduccionalmente procesados. A modo de ejemplo, los ensayos para la expresión de 029cel incluyen transferencia Western para la proteína 029cel, análisis por transferencia Northern y ensayos por reacción en cadena de la polimerasa con transcriptasa inversa (RT-PCR; del inglés, reverse transcriptase polymerase chain reaction) para mRNA de 029cel, y ensayos de actividad endoglucanasa como los descritos en S. P. Shoemaker y R. D. Jr. Brown (Biochim. Biophys. Acta, 1978, 523: 133-146) y Schulein (1988).

Como se emplea en esta memoria, la expresión "agente tensioactivo" se refiere a cualquier compuesto generalmente reconocido en la técnica por tener cualidades superficialmente activas. De este modo, por ejemplo, los agentes tensioactivos comprenden agentes tensioactivos aniónicos, catiónicos y no iónicos tales como los hallados comúnmente en detergentes. Los agentes tensioactivos aniónicos incluyen alquilbencenosulfonatos lineales o ramificados; alquil- o alquenil-éter-sulfatos que tienen grupos alquilo o grupos alquenilo lineales o ramificados; alquil- o alquenil-sulfatos; olefinasulfonatos; y alcanosulfonatos. Los agentes tensioactivos anfóteros incluyen sulfonatos de sales de amonio cuaternario y agentes tensioactivos anfóteros de tipo betaína. Dichos agentes tensioactivos anfóteros tienen tanto los grupos cargados positivos como los negativos en la misma molécula. Los agentes tensioactivos no iónicos pueden comprender éteres de polioxialquileo así como alcanolamidas de ácidos grasos superiores o aductos de las mismas con óxido de alquileo, monoésteres de ácido graso y glicerol, y similares.

Como se emplea en esta memoria, la expresión "tejido que contiene celulosa" se refiere a tejidos cosidos o no cosidos, hilos o fibras hechos de celulosa que contiene o no algodón o de mezclas de celulosas que contienen o no algodón, incluyendo productos celulósicos naturales y productos celulósicos artificiales (tales como yute, lino, ramio, rayón y Lyocell).

Como se emplea en esta memoria, la expresión "tejido que contiene algodón" se refiere a tejidos cosidos o no cosidos, hilos o fibras hechos de algodón puro o de mezclas de algodón, incluyendo telas de algodón tejidas en telar, tejidos de punto de algodón, mahones de algodón, hilos de algodón, algodón en rama y similares.

Como se emplea en esta memoria, la expresión "composición para lavado a la piedra" se refiere a una formulación para uso en el lavado a la piedra de tejidos que contienen celulosa. Las composiciones para lavado a la piedra se emplean para modificar tejidos que contienen celulosa antes de su venta, es decir, durante el proceso de fabricación. Por contraste, las composiciones detergentes están destinadas a la limpieza de prendas sucias y no se utilizan durante el proceso de fabricación.

Como se emplea en esta memoria, la expresión "composición detergente" se refiere a una mezcla que está destinada al uso en un medio de lavado para la colada de tejidos que contienen celulosa sucios. En el contexto de la presente invención, dichas composiciones pueden incluir, además de celulosas y agentes tensioactivos, enzimas hidrolíticas adicionales, agentes mejoradores de la limpieza, agentes blanqueadores, activadores del blanqueo, azuletes y colorantes fluorescentes, inhibidores del apelmazamiento, agentes enmascaradores, activadores de celulosas, antioxidantes y agentes solubilizantes.

Como se emplean en esta memoria, las expresiones "activa" y "biológicamente activa" se refieren a una actividad biológica asociada con una proteína particular y se usan indistintamente en esta memoria. Por ejemplo, la actividad enzimática asociada con una proteasa es la proteólisis y, por lo tanto, una proteasa activa tiene actividad proteolítica. De esto resulta que la actividad biológica de una proteína dada se refiere a toda actividad biológica típicamente atribuida a esa proteína por quienes tienen experiencia en la técnica.

Cuando se emplea en disoluciones enzimáticas, el componente 029cel se añade generalmente en una cantidad suficiente para permitir la máxima velocidad de liberación de azúcares solubles de la biomasa. La cantidad de componente 029cel añadida depende del tipo de biomasa que se va a hidrolizar, la cual puede ser fácilmente determinada por el técnico experto. Sin embargo, cuando se emplea, el porcentaje ponderal del componente 029cel con respecto a cualesquier otros componentes de tipo celulasa presentes en la composición de celulasa es de preferiblemente alrededor de 1, preferiblemente alrededor de 5, preferiblemente alrededor de 10, preferiblemente alrededor de 15 o preferiblemente alrededor de 20 por ciento en peso, a preferiblemente alrededor de 25, preferiblemente alrededor de 30, preferiblemente alrededor de 35, preferiblemente alrededor de 40, preferiblemente alrededor de 45 o preferiblemente alrededor de 50 por ciento en peso. Además, los intervalos preferidos pueden ser de aproximadamente 0,5 a aproximadamente 15 por ciento en peso, de aproximadamente 0,5 a aproximadamente 20 por ciento en peso, de aproximadamente 1 a aproximadamente 10 por ciento en peso, de aproximadamente 1 a aproximadamente 15 por ciento en peso, de aproximadamente 1 a aproximadamente 20 por ciento en peso, de aproximadamente 1 a aproximadamente 25 por ciento en peso, de aproximadamente 5 a aproximadamente 20 por ciento en peso, de aproximadamente 5 a aproximadamente 25 por ciento en peso, de aproximadamente 5 a aproximadamente 30 por ciento en peso, de aproximadamente 5 a aproximadamente 35 por ciento en peso, de aproximadamente 5 a aproximadamente 40 por ciento en peso, de aproximadamente 5 a aproximadamente 45 por ciento en peso, de aproximadamente 5 a aproximadamente 50 por ciento en peso, de aproximadamente 10 a aproximadamente 20 por ciento en peso, de aproximadamente 10 a aproximadamente 25 por ciento en peso, de aproximadamente 10 a aproximadamente 30 por ciento en peso, de aproximadamente 10 a aproximadamente 35 por ciento en peso, de aproximadamente 10 a aproximadamente 40 por ciento en peso, de aproximadamente 10 a

aproximadamente 45 por ciento en peso, de aproximadamente 10 a aproximadamente 50 por ciento en peso, de aproximadamente 15 a aproximadamente 20 por ciento en peso, de aproximadamente 15 a aproximadamente 25 por ciento en peso, de aproximadamente 15 a aproximadamente 30 por ciento en peso, de aproximadamente 15 a aproximadamente 35 por ciento en peso, de aproximadamente 15 a aproximadamente 40 por ciento en peso, de aproximadamente 15 a aproximadamente 45 por ciento en peso, y de aproximadamente 15 a aproximadamente 50 por ciento en peso.

II. Biología molecular

Esta invención depende de técnicas rutinarias en el campo de la genética recombinante. Unos textos básicos en que se describen los métodos generales para uso en esta invención incluyen Sambrook *et al.*, *Molecular Cloning, A Laboratory Manual* (2ª edición, 1989); Kriegler, *Gene Transfer and Expression: A Laboratory Manual* (1990); y Ausubel *et al.* (redactores), *Current Protocols in Molecular Biology* (1994).

Para obtener una expresión de alto nivel de un gen clonado, el gen heterólogo es preferiblemente situado a aproximadamente la misma distancia del promotor a la que está el gen de celulasa de origen natural. Sin embargo, como es sabido en la técnica, se puede admitir cierta variación en esta distancia sin pérdida de la función del promotor.

Los expertos en la técnica son conscientes de que se puede modificar un promotor natural por recambio, sustitución, adición o eliminación de uno o más nucleótidos sin que cambie su función. La práctica de la invención abarca, y no está limitada por, dichas alteraciones en el promotor.

El vector/construcción de expresión contiene típicamente una unidad de transcripción o casete de expresión que contiene todos los elementos adicionales necesarios para la expresión de la secuencia heteróloga. De este modo, un casete de expresión típico contiene un promotor operativamente unido a la secuencia de ácido nucleico heteróloga y señales necesarias para una eficaz poliadenilación del transcrito, sitios de unión al ribosoma, y una terminación de la traducción. Los elementos adicionales del casete pueden incluir potenciadores y, si se emplea DNA genómico como el gen estructural, intrones con sitios funcionales dadores y aceptores de ajustes.

La práctica de la invención no está restringida por la elección del promotor en la construcción genética. La única restricción a la elección del promotor es que sea funcional en la célula huésped empleada. Un promotor preferido cuando la célula huésped de transformación es del género *Bacillus* es el promotor *aprE*.

Además de una secuencia promotora, el casete de expresión debería contener también una región de terminación de la transcripción, cadena abajo del gen estructural, para proporcionar una terminación eficaz. La región de terminación se puede obtener del mismo gen que la secuencia promotora o se puede obtener de genes diferentes.

El particular vector de expresión utilizado para transportar la información genética a la célula no es particularmente crítico. Se puede emplear cualquiera de los vectores convencionales utilizados para expresión en células eucarióticas o procarióticas. Los vectores de expresión bacterianos estándares incluyen los bacteriófagos λ y M13, así como plásmidos tales como plásmidos basados en pBR322, pSKF, pET23D, y sistemas de expresión de fusión tales como MBP, GST y LacZ. También se pueden añadir etiquetas epitópicas a las proteínas recombinantes para proporcionar métodos de aislamiento convenientes, por ejemplo, c-myc.

Los elementos que se incluyen típicamente en vectores de expresión también incluyen un replicón, un gen que codifica resistencia a antibióticos para permitir la selección de bacterias que contienen plásmidos recombinantes, y sitios de restricción únicos en regiones no esenciales del plásmido para permitir la inserción de secuencias heterólogas. El particular gen de resistencia a antibióticos elegido no es crítico; son adecuados cualesquier de los muchos genes de resistencia conocidos en la técnica.

Los métodos de transformación de la presente invención pueden dar lugar a la integración estable de todo, o parte de, el vector de transformación en el genoma del hongo filamentoso. Sin embargo, también se contempla la transformación que da lugar al mantenimiento de un vector de transformación extracromosómico autorreplicante.

El gen que codifica la celulasa de la presente invención puede ser clonado usando vectores de fago λ (expresión) y células huésped de *E. coli*. (Alternativamente, se puede utilizar una clonación por PCR usando cebadores de consenso creados sobre dominios conservados). Los solicitantes han descubierto que la transformación del gen que codifica la celulasa de la presente invención y la expresión en *E. coli* dan lugar a una proteína activa. Después de una primera operación de clonación en *E. coli*, se puede transferir un gen de celulasa de acuerdo con la presente invención a un huésped de expresión industrial más preferido, tal como una especie del género *Bacillus* o *Streptomyces*, un hongo filamentoso tal como uno del género *Aspergillus* o *Trichoderma*, o una levadura tal como una del género *Saccharomyces*. La expresión y secreción de alto nivel obtenibles en estos organismos huésped permiten la acumulación de la celulasa en el medio de fermentación, del cual puede ser posteriormente recuperada.

En Ferrari *et al.*, Patente de EE.UU. nº 5.264.366, se proporciona un protocolo de transformación y expresión general preferido para cepas de *Bacillus* con proteasas suprimidas. En, por ejemplo, Berka *et al.*, Patente de EE.UU. nº 5.364.770, se describen la transformación y expresión en *Aspergillus*.

Se pueden utilizar muchos métodos de transfección estándares para producir líneas celulares de *Trichoderma reesei* que expresen grandes cantidades de la proteína heteróloga. Algunos de los métodos publicados para la introducción de construcciones de DNA en cepas de *Trichoderma* productoras de celulasas incluyen Lorito, Hayes, DiPietro y Harman, 1993, Curr. Genet. 24: 349-356; Goldman, VanMontagu y Herrera-Estrella, 1990, Curr. Genet. 17: 169-174; Penttila, Nevalainen, Ratto, Salminen y Knowles, 1987, Gene 6: 155-164; para *Aspergillus*, Yelton, Hamer y Timberlake, 1984, Proc. Natl. Acad. Sci. USA 81: 1470-1474; para *Fusarium*, Bajar, Podila y Kolattukudy, 1991, Proc. Natl. Acad. Sci. USA 88: 8202-8212; para *Streptomyces*, Hopwood et al., 1985, The John Innes Foundation, Norwich, Reino Unido; y para *Bacillus*, Brigidi, DeRossi, Bertarini, Riccardi y Matteuzzi, 1990, FEMS Microbiol. Lett. 55: 135-138.

5 Sin embargo, se puede utilizar cualquiera de los bien conocidos procedimientos para introducir secuencias nucleotídicas extrañas en células huésped. Estos incluyen el uso de transfección con fosfato cálcico, Polybrene, fusión de protoplastos, electroporación, biolística, liposomas, microinyección, vectores plasmídicos, vectores virales y cualquiera de los otros métodos bien conocidos para introducir DNA genómico clonado, cDNA, DNA sintético u otro material genético extraño en una célula huésped (véase, por ejemplo, Sambrook et al., *supra*). También tiene
10 utilidad el método de transfección mediada por *Agrobacterium*, descrito en la Patente de EE.UU. n° 6.255.115. Sólo es necesario que el procedimiento particular de ingeniería genética empleado permita introducir exitosamente al
15 menos un gen en la célula huésped capaz de expresar el gen heterólogo.

Una vez que se ha introducido el vector de expresión en las células, las células transfectadas son cultivadas bajo unas condiciones que favorecen la expresión de genes bajo el control de secuencias promotoras de genes de
20 celulasa. Del modo descrito más adelante se pueden cultivar grandes partidas de células transformadas. Finalmente, el producto se recupera del cultivo usando técnicas estándares.

Por lo tanto, la presente invención proporciona la expresión y la secreción potenciada de las celulasas de la invención cuya expresión está bajo el control de secuencias promotoras de genes de celulasa, incluyendo genes de
25 celulasa de origen natural, secuencias de DNA de fusión y diversas construcciones heterólogas. La invención también proporciona procesos para que se expresen y secreten niveles elevados de las celulasas de la invención.

III. Identificación de ácidos nucleicos y secuencias proteicas codificadas

Se preparó un banco genómico de *Bacillus agaradhaerans* (DSM 8721) usando técnicas estándares conocidas en este campo técnico. Este organismo produce una celulasa alcalina, (endo-1,4-beta-glucanasa), que pertenece a la familia 5 de celulasas de las glicosil hidrolasas; endoglucanasa 5A, EC 3.2.1.4, Swiss-Prot: 085465, nombre de
30 entrada: GUN5_BACAG; número de acceso del EBI: AF067428, cuyo gen tiene una longitud de 1203 pares de bases (Davies et al., 1998). Se detectaron clones positivos para celulasa con una incidencia de 1/3000 en el ensayo en placa. En el proceso para aislar un gen de acuerdo con un aspecto de la presente invención, se emplearon cebadores degenerados basados en la secuencia de codificación para esta enzima. Sin embargo, inesperadamente,
35 no se obtuvo producto alguno por PCR al usar cebadores conocidos por multiplicar la conocida celulasa de *B. agaradhaerans*. Por lo tanto, la secuencia completa del inserto que codifica la celulasa fue determinada por "paseo con cebadores" (*primer walking*).

En el proceso para aislar un gen de acuerdo con el segundo aspecto de la presente invención se utiliza su homología con una secuencia de nucleótidos que comprende toda, o parte de, la secuencia de nucleótidos de ID. SEC. n° 1 o ID. SEC. n° 2, como se muestran en la lista de secuencias. Los ejemplos de dichos procesos incluyen:

- 40 a) explorar un banco génico que supuestamente contiene un gen 029cel usando la secuencia de nucleótidos como una sonda.
- b) preparar un cebador basado en la información sobre la secuencia de nucleótidos y llevar luego a cabo una PCR usando una muestra que supuestamente contiene un gen 029cel como un molde.

Más específicamente, el proceso a) anterior comprende:

- 45 a) preparar un banco génico que supuestamente contiene un gen de celulasa, explorar el banco génico usando una secuencia de nucleótidos que comprende toda, o parte de, la secuencia de nucleótidos de ID. SEC. n° 2, como se muestra en la lista de secuencias, para seleccionar secuencias que se hibriden con la secuencia de nucleótidos que comprende toda, o parte de, la secuencia de nucleótidos de ID. SEC. n° 2, como se muestra en la lista de secuencias del banco génico, aislar luego las secuencias seleccionadas, y aislar un gen 029cel de las
50 secuencias que han sido seleccionadas y aisladas del banco génico.

El banco génico puede ser un banco de DNA genómico o un banco de cDNA, y se puede preparar de acuerdo con un procedimiento conocido.

IV. Expresión proteica

Las proteínas de la presente invención se producen al cultivar células transformadas con un vector de expresión que contiene el gen de celulasa de la invención cuya expresión está bajo el control de secuencias promotoras. La
55

presente invención es particularmente útil para potenciar la producción intracelular y/o extracelular de proteínas. La proteína puede ser homóloga o heteróloga.

Las proteínas de la presente invención pueden ser también modificadas de modo que se formen moléculas quiméricas que comprendan una proteína de interés fusionada con otro polipéptido heterólogo u otra secuencia de aminoácidos heteróloga. En una realización, dicha molécula quimérica comprende una fusión de la proteína de interés con una etiqueta polipeptídica que proporciona un epítopo al que se puede unir selectivamente un anticuerpo anti-etiqueta. La etiqueta epitópica se coloca generalmente en el extremo amino o carboxilo de la proteína de interés.

En la técnica se conocen bien diversas etiquetas polipeptídicas y sus respectivos anticuerpos. Los ejemplos incluyen etiquetas de poli-histidina (poli-his) o poli-histidina-glicocola (poli-his-gly); HIS6 y etiquetas para quelación de metales, la etiqueta polipeptídica flu HA y su anticuerpo 12CA5 [Field *et al.*, *Mol. Cell. Biol.* 8: 2159-2165 (1988)]; la etiqueta c-myc y los anticuerpos 8F9, 3C7, 6E10, G4, B7 y 9E10 hacia ella [Evan *et al.*, *Molecular and Cellular Biology* 5: 3610-3616 (1985)]; y la etiqueta glicoproteína D (gD) del virus del herpes simple y su anticuerpo [Paborsky *et al.*, *Protein Engineering* 3 (6): 547-553 (1990)]. Otras etiquetas polipeptídicas incluyen el péptido FLAG [Hopp *et al.*, *BioTechnology* 6: 1204-1210 (1988)]; el péptido epitópico KT3 [Martin *et al.*, *Science* 255: 192-194 (1992)]; el péptido epitópico de tubulina [Skinner *et al.*, *J. Biol. Chem.* 266: 15.163-15.166 (1991)]; y la etiqueta peptídica de la proteína del gen 10 de T7 [Lutz-Freyermuth *et al.*, *Proc. Natl. Acad. Sci. USA* 87: 6393-6397 (1990)].

Las condiciones apropiadas para la expresión de dicho gen 029cel comprenden proporcionar al cultivo los componentes necesarios para el crecimiento y/o la expresión de la celulasa de la invención. Las condiciones óptimas para la producción de las proteínas variarán con la elección de la célula huésped y con la elección de la proteína que se va a expresar. Dichas condiciones serán fácilmente determinadas por un experto en la técnica a través de una experimentación u optimización rutinaria.

La proteína de interés es típicamente purificada o aislada después de la expresión. La proteína de interés puede ser aislada o purificada de diversas maneras conocidas por los expertos en la técnica dependiendo de qué otros componentes están presentes en la muestra. Los métodos de purificación estándares incluyen técnicas electroforéticas, moleculares, inmunológicas y cromatográficas, incluyendo las cromatografías HPLC de intercambio iónico, hidrófoba, de afinidad y de fase inversa y el cromatografía. Por ejemplo, la proteína de interés puede ser purificada usando una columna estándar con anticuerpo anti-proteína de interés. También son útiles las técnicas de ultrafiltración y diafiltración, junto con la concentración de proteínas. Para una guía general sobre técnicas de purificación adecuadas, véase Scopes, *Protein Purification* (1982). El grado de purificación necesario variará dependiendo del uso de la proteína de interés. En ciertos casos no será necesaria la purificación.

V. Utilidad de la celulasa

El tratamiento de textiles de acuerdo con la presente invención contempla el procesamiento o la limpieza de textiles con una composición que comprende la celulasa de esta invención. Dicho tratamiento incluye, pero no se limita a, el lavado a la piedra, la modificación de la textura, el tacto y/o el aspecto de tejidos que contienen celulosa, y otras técnicas empleadas durante la fabricación o el lavado/reacondicionamiento de tejidos que contienen celulosa. Además, en el contexto de esta invención, el tratamiento contempla la separación de algodón "inmaduro" o "muerto" de tejidos o fibras celulósicas. El algodón inmaduro es significativamente más amorfo que el algodón maduro y, a causa de ello, por ejemplo, se tiñe desigualmente. La composición contemplada en la presente invención incluye además un componente de celulasa para uso en el lavado de un tejido manufacturado sucio que contiene celulosa. Por ejemplo, se puede emplear una celulasa de esta invención en una composición detergente para el lavado de ropa sucia. Las composiciones detergentes útiles de acuerdo con la presente invención incluyen formulaciones especiales tales como composiciones para prelavado, remojo previo y restauración de colores de uso doméstico. Dichas composiciones de tratamiento, como se describen en esta memoria, pueden estar en forma de un concentrado que requiere dilución o en forma de una disolución diluida o en una forma que puede ser aplicada directamente al tejido que contiene celulosa. En, por ejemplo, la Publicación EP n° 220.016 y las Solicitudes GB números 1.368.599 y 2.095.275 se describen técnicas de tratamiento generales para el tratamiento de textiles con celulasa.

El tratamiento de un material celulósico de acuerdo con la presente invención contempla además el tratamiento de pienso para animales, pulpa y/o papel, alimento y grano con fines conocidos en la técnica. Por ejemplo, se sabe que las celulasas aumentan el valor del pienso para animales, mejoran la capacidad de drenaje de la pulpa de madera, mejoran los productos alimenticios y reducen la fibra del grano durante el proceso de molienda de granos en estado húmedo o en estado seco.

El tratamiento de acuerdo con la presente invención comprende preparar una disolución acuosa que contiene una cantidad eficaz de una celulasa o una combinación de celulasas junto con otros ingredientes opcionales que incluyen, por ejemplo, un tampón, un agente tensioactivo y/o un agente desengrasante. Una cantidad eficaz de una composición de enzima celulasa es una concentración de enzima celulasa suficiente para su finalidad prevista. De esta manera, por ejemplo, una "cantidad eficaz" de celulasa en una composición para lavado a la piedra de acuerdo con la presente invención es la cantidad que proporcionará el efecto deseado, por ejemplo, producir un aspecto

desgastado y deslucido en costuras y en piezas de tejido. Similarmente, una "cantidad eficaz" de celulasa en una composición destinada a mejorar el tacto y/o aspecto de un tejido que contiene celulosa es la cantidad que produce mejoras mensurables en el tacto, por ejemplo, mejorar la suavidad del tejido, o el aspecto, por ejemplo, eliminar las bolitas y fibrillas que tienden a reducir la nitidez en el aspecto de un tejido. La cantidad de celulasa empleada también depende del equipo empleado, los parámetros de procesamiento empleados (la temperatura de la disolución de tratamiento de celulasa, el tiempo de exposición a la disolución de celulasa, y similares) y la actividad celulasa (por ejemplo, una disolución particular requerirá una concentración menor de celulasa cuando se utilice una composición de celulasa más activa en comparación con una composición de celulasa menos activa). La concentración exacta de celulasa en la disolución acuosa de tratamiento a la que se añade el tejido que se va a tratar puede ser fácilmente determinada por el técnico experto basándose en los factores anteriores así como en el resultado deseado. En los procesos de lavado a la piedra se ha preferido generalmente que la celulasa esté presente en la disolución acuosa de tratamiento en una concentración de aproximadamente 0,5 a 5000 ppm, y lo más preferiblemente de aproximadamente 10 a 200 ppm, de proteína total. En las composiciones para la mejora del tacto y/o el aspecto de un tejido que contiene celulosa se ha preferido generalmente que la celulasa esté presente en la disolución acuosa de tratamiento en una concentración de aproximadamente 0,1 a 2000 ppm, y lo más preferiblemente de aproximadamente 0,5 a 200 ppm, de proteína total.

En una realización de tratamiento preferida, se emplea un tampón en la composición de tratamiento de modo que la concentración de tampón sea suficiente para mantener el pH de la disolución dentro del intervalo en que la celulasa empleada presenta actividad. El pH en que la celulasa presenta actividad depende de la naturaleza de la celulasa empleada. La concentración exacta de tampón empleada dependerá de diversos factores que el técnico experto puede tener en cuenta fácilmente. Por ejemplo, en una realización preferida, el tampón así como la concentración de tampón se seleccionan para mantener el pH de la disolución final de celulasa dentro del intervalo de pHs requerido para una actividad celulasa óptima. La determinación del intervalo de pHs óptimo de las celulosas de la invención puede ser llevada a cabo de acuerdo con técnicas bien conocidas. Los tampones adecuados para el pH dentro del intervalo de actividad de la celulasa son también bien conocidos por los expertos en este campo técnico.

Además de la celulasa y un tampón, la composición de tratamiento puede contener opcionalmente un agente tensioactivo. Los agentes tensioactivos adecuados incluyen cualquier agente tensioactivo compatible con la celulasa que se utiliza y con el tejido, incluyendo, por ejemplo, agentes tensioactivos aniónicos, no iónicos y anfóteros. Los agentes tensioactivos aniónicos adecuados incluyen, pero no se limitan a, alquilbencenosulfonatos lineales o ramificados; alquil- o alquenil-éter-sulfatos que tienen grupos alquilo o grupos alquenilo lineales o ramificados; alquil- o alquenil-sulfatos; olefinasulfonatos; alcanosulfonatos y similares. Los contraiones adecuados para los agentes tensioactivos aniónicos incluyen, pero no se limitan a, iones de metales alcalinos tales como sodio y potasio; iones de metales alcalinotérreos tales como calcio y magnesio; ion amonio; y alcanolaminas que tienen de 1 a 3 grupos alcohol de número de carbonos 2 o 3. Los agentes tensioactivos anfóteros incluyen, por ejemplo, sulfonatos de sales de amonio cuaternario y agentes tensioactivos anfóteros de tipo betaína. Dichos agentes tensioactivos anfóteros tienen tanto los grupos cargados positivos como los negativos en la misma molécula. Los agentes tensioactivos no iónicos comprenden generalmente éteres de polioxialquilenos así como alcanolamidas de ácidos grasos superiores o aductos de las mismas con óxido de alquilenos, y monoésteres de ácido graso y glicerol. También se pueden emplear mezclas de agentes tensioactivos de las maneras conocidas por los expertos en la técnica.

Se puede preparar una composición concentrada de celulasa para uso en los métodos descritos en esta memoria. Dichos concentrados contienen cantidades concentradas de la composición de celulasa anteriormente descrita, un tampón y un agente tensioactivo, preferiblemente en una disolución acuosa. Cuando se formula así, el concentrado de celulasa puede ser fácilmente diluido con agua para preparar rápida y exactamente preparaciones de celulasa que tengan la concentración requerida de cada componente. Cuando se formulan concentrados acuosos, estos concentrados pueden ser diluidos para alcanzar la concentración requerida de los componentes en la disolución de celulasa, como se indicó anteriormente. Como es bien evidente, dichos concentrados de celulasa permiten la fácil formulación de las disoluciones de celulasa y permiten también un transporte factible de la composición al lugar donde se utilizará. El concentrado de tratamiento puede estar en cualquier forma reconocida en la técnica, por ejemplo, en forma de líquido, emulsión, gel o pasta. Dichas formas son bien conocidas por los expertos en la técnica.

Cuando se emplea un concentrado sólido de celulasa, la composición de celulasa puede ser un gránulo, un polvo, un aglomerado o un disco sólido. Los gránulos pueden ser formulados para que contengan materiales que reduzcan la velocidad de disolución de los gránulos en el medio de lavado. Dichos materiales y gránulos se describen en la Patente de EE.UU. nº 5.254.283.

También se pueden utilizar o disponer otros materiales en la composición de celulasa de la presente invención, según se desee, incluyendo piedras, pumita, cargas, disolventes, activadores de enzimas y agentes anti-redeposición, dependiendo del uso final de la composición.

A modo de ejemplo se describirán métodos de lavado a la piedra en detalle; sin embargo, los parámetros descritos son fácilmente modificados por el técnico experto para otras aplicaciones, es decir, mejorar el tacto y/o el aspecto de un tejido. Se pone el tejido que contiene celulosa en contacto con la composición para lavado a la piedra que contiene celulasa, que contiene una cantidad eficaz de la celulasa, entremezclando la composición de tratamiento

5 con la composición para lavado a la piedra y poniendo así la enzima celulasa próxima al tejido. Posteriormente se agita la disolución acuosa que contiene la celulasa y el tejido. Si la composición de tratamiento es una disolución acuosa, el tejido puede ser directamente remojado en la disolución. Similarmente, cuando la composición para lavado a la piedra es un concentrado, se diluye el concentrado en un baño de agua con el tejido que contiene celulosa. Cuando la composición para lavado a la piedra está en forma sólida, por ejemplo, un gel o una barrita sólida para prelavado, se puede poner en contacto la composición para lavado a la piedra aplicando directamente la composición al tejido o al líquido de lavado.

10 El tejido que contiene celulosa es incubado con la disolución para lavado a la piedra bajo unas condiciones eficaces para permitir que la acción enzimática confiera un aspecto de lavado a la piedra al tejido que contiene celulosa. Por ejemplo, durante el lavado a la piedra se pueden ajustar el pH, la proporción de líquido, la temperatura y el tiempo de reacción para optimizar las condiciones bajo las cuales actúa la composición para lavado a la piedra. "Condiciones eficaces" se refiere necesariamente al pH, la proporción de líquido y la temperatura que permiten que la enzima celulasa reaccione eficazmente con el tejido que contiene celulosa, para producir en este caso el efecto de lavado a la piedra. Dentro de la experiencia en la técnica está el maximizar las condiciones para utilizar las composiciones para lavado a la piedra de acuerdo con la presente invención.

15 La proporción de líquido durante el lavado a la piedra, es decir, la proporción entre el peso de la disolución de la composición para lavado a la piedra (es decir, el líquido de lavado) y el peso del tejido, empleada en esta memoria, es generalmente una cantidad suficiente para alcanzar el deseado efecto de lavado a la piedra en el tejido de mahón y depende del proceso utilizado. Preferiblemente, las proporciones de líquido son de aproximadamente 4:1 a aproximadamente 50:1, más preferiblemente de aproximadamente 5:1 a aproximadamente 20:1, y lo más preferiblemente de aproximadamente 10:1 a aproximadamente 15:1.

20 Las temperaturas de reacción durante el lavado a la piedra con las presentes composiciones para lavado a la piedra vienen reguladas por dos factores competitivos. En primer lugar, mayores temperaturas corresponden generalmente a cinéticas de reacción potenciadas, es decir, reacciones más rápidas, que permiten tiempos de reacción reducidos en comparación con los tiempos de reacción requeridos a temperaturas menores. En consecuencia, las temperaturas de reacción son generalmente al menos aproximadamente 10 °C y superiores. En segundo lugar, la celulasa es una proteína que pierde actividad por encima de una temperatura de reacción dada, temperatura que depende de la naturaleza de la celulasa utilizada. De este modo, si se permite que la temperatura de reacción sea demasiado elevada, se pierde la actividad celulolítica como resultado de la desnaturalización de la celulasa. Aunque las temperaturas estándares para la utilización de celulosas en la técnica están generalmente en el intervalo de 35 °C a 65 °C, y se esperaría que estas condiciones fueran también adecuadas para la celulasa de la invención, deberían determinarse las condiciones de temperatura óptimas de acuerdo con técnicas bien conocidas con respecto a la celulasa específica utilizada.

25 Los tiempos de reacción dependen de las condiciones específicas bajo las cuales tiene lugar el lavado a la piedra. Por ejemplo, el pH, la temperatura y la concentración de celulasa afectarán al tiempo de reacción óptimo. En general, los tiempos de reacción son de aproximadamente 5 minutos a aproximadamente 5 horas, y preferiblemente de aproximadamente 10 minutos a aproximadamente 3 horas, y más preferiblemente de aproximadamente 20 minutos a aproximadamente 1 hora.

30 De acuerdo con aún otra realización preferida de la presente invención, se puede emplear la celulasa de la invención en una composición detergente. Las composiciones detergentes de acuerdo con la presente invención son útiles como composiciones para prelavado o composiciones para remojo previo, o para la limpieza durante el ciclo regular de lavado o enjuague. Preferiblemente, la composición detergente de la presente invención comprende una cantidad eficaz de celulasa, y un agente tensioactivo, e incluye opcionalmente otros ingredientes descritos más adelante.

35 Una cantidad eficaz de celulasa empleada en las composiciones detergentes de esta invención es una cantidad suficiente para impartir los efectos deseables que se sabe que produce la celulasa sobre los tejidos que contienen celulosa, por ejemplo, eliminación de bolitas, suavizado, anti-formación de bolitas, eliminación de fibras superficiales, anti-matiz grisáceo y limpieza. Preferiblemente, en la composición detergente se emplea la celulasa en una concentración de aproximadamente 10 ppm a aproximadamente 20.000 ppm de detergente.

40 La concentración de enzima celulasa empleada en la composición detergente es preferiblemente seleccionada de modo que, tras la dilución en un medio de lavado, la concentración de enzima celulasa esté en el intervalo de aproximadamente 0,01 a aproximadamente 1000 ppm, preferiblemente de aproximadamente 0,02 ppm a aproximadamente 500 ppm, y lo más preferiblemente de aproximadamente 0,5 ppm a aproximadamente 250 ppm, de proteína total. La cantidad de enzima celulasa empleada en la composición detergente dependerá del grado en que se diluirá el detergente tras la adición de agua para formar una disolución de lavado.

45 Las composiciones detergentes de la presente invención pueden estar en cualquier forma reconocida en la técnica, tal como, por ejemplo, como un líquido, en gránulos, en emulsiones, en geles o en pastas. Tales formas son bien conocidas por el técnico experto. Cuando se emplea una composición detergente sólida, la celulasa se formula preferiblemente como gránulos. Preferiblemente, los gránulos pueden ser formulados para que contengan además un agente protector de celulosas. El gránulo puede ser formulado para que contenga materiales que reduzcan la

velocidad de disolución del gránulo en el medio de lavado. En la Patente de EE.UU. nº 5.254.283 se describen dichos materiales y gránulos.

En las composiciones detergentes de esta invención se emplea un agente superficialmente activo, es decir, un agente tensioactivo, incluyendo agentes tensioactivos aniónicos, no iónicos y anfóteros bien conocidos por su uso en composiciones detergentes.

Los agentes tensioactivos aniónicos adecuados para uso en la composición detergente de esta invención incluyen alquilbencenosulfonatos lineales o ramificados; alquil- o alquenil-éter-sulfatos que tienen grupos alquilo o grupos alquenilo lineales o ramificados; alquil- o alquenil-sulfatos; olefinasulfonatos; y alcanosulfonatos. Los contraiones adecuados para los agentes tensioactivos aniónicos incluyen iones de metales alcalinos tales como sodio y potasio; iones de metales alcalinotérreos tales como calcio y magnesio; ion amonio; y alcanolaminas que tienen de 1 a 3 grupos alcano de número de carbonos 2 o 3. Los agentes tensioactivos anfóteros incluyen sulfonatos de sales de amonio cuaternario y agentes tensioactivos anfóteros de tipo betaína. Dichos agentes tensioactivos anfóteros tienen tanto los grupos cargados positivos como los negativos en la misma molécula. Los agentes tensioactivos no iónicos comprenden generalmente éteres de polioalquilenos así como alcanolamidas de ácidos grasos superiores o aductos de las mismas con óxido de alquilenos, monoésteres de ácido graso y glicerol, y similares. En la Solicitud de Patente Británica nº 2.094.826 A se describen agentes tensioactivos adecuados para uso en esta invención. También se pueden utilizar mezclas de dichos agentes tensioactivos. El agente tensioactivo o la mezcla de agentes tensioactivos se emplea generalmente en las composiciones detergentes de esta invención en una cantidad de aproximadamente 1 por ciento en peso a aproximadamente 95 por ciento en peso de la composición detergente total, y preferiblemente de aproximadamente 5 por ciento en peso a aproximadamente 45 por ciento en peso de la composición detergente total. Además de la composición de celulosa y el (los) agente(s) tensioactivo(s), las composiciones detergentes de esta invención pueden contener opcionalmente uno o más de los componentes siguientes:

Hidrolasas salvo celulasa

Las hidrolasas adecuadas incluyen éster carboxilato hidrolasa, tioéster hidrolasa, monoéster fosfato hidrolasa y diéster fosfato hidrolasa, que actúan sobre el enlace éster; glicósido hidrolasa, que actúa sobre compuestos glicosílicos; una enzima que hidroliza compuestos N-glicosílicos; tioéter hidrolasa, que actúa sobre el enlace éter; y α -amino-acil-peptido hidrolasa, peptidil-aminoácido hidrolasa, acil-aminoácido hidrolasa, dipéptido hidrolasa y peptidil-peptido hidrolasa, que actúan sobre el enlace peptídico. Entre ellas, son preferibles la éster carboxilato hidrolasa, la glicósido hidrolasa y la peptidil-peptido hidrolasa. Las hidrolasas adecuadas incluyen (1) proteasas pertenecientes a las peptidil-peptido hidrolasas, tales como pepsina, pepsina B, quimosina, tripsina, quimotripsina A, quimotripsina B, elastasa, enterocinasa, catepsina C, papaína, quimopapaína, ficina, trombina, fibrinolisisina, renina, subtilisina, aspergillopeptidasa A, colagenasa, clostridiopeptidasa B, calicreína, gastrisina, catepsina D, bromelina, queratinasa, quimotripsina C, pepsina C, aspergillopeptidasa B, urocinasa, carboxipeptidasas A y B, y aminopeptidasa; (2) glicósido hidrolasas (la celulasa que es un ingrediente esencial está excluida de este grupo) tales como α -amilasa, β -amilasa, glucoamilasa, invertasa, lisozima, pectinasa, quitinasa y dextranasa. Entre ellas son preferibles la α -amilasa y la β -amilasa. Actúan en sistemas que van de ácidos a neutros, pero una que se obtiene de bacterias presenta una actividad elevada en un sistema alcalino; y (3) éster carboxilato hidrolasas, incluyendo carboxil esterasa, lipasa, pectina esterasa y clorofilasa. Entre ellas es particularmente eficaz la lipasa.

La hidrolasa distinta de la celulasa se incorpora a la composición detergente en tanto como se requiera de acuerdo con la finalidad. Se debería incorporar preferiblemente en una cantidad de 0,001 a 5 por ciento en peso, y más preferiblemente de 0,02 a 3 por ciento en peso, en términos de proteína purificada. Esta enzima se debería emplear en forma de gránulos hechos de enzima cruda sola o en combinación con otros componentes de la composición detergente. Los gránulos de enzima cruda se utilizan en una cantidad tal que la enzima purificada es de 0,001 a 50 por ciento en peso en los gránulos. Los gránulos se utilizan en una cantidad de 0,002 a 20 y preferiblemente de 0,1 a 10 por ciento en peso. Como con las celulasas, estos gránulos pueden ser formulados para que contengan un agente protector de enzimas y un material retardador de la disolución.

Agentes tensioactivos catiónicos y sales de ácidos grasos de cadena larga

Dichos agentes tensioactivos catiónicos y sales de ácidos grasos de cadena larga incluyen sales de ácidos grasos saturados o insaturados, sales de alquil- o alquenil-éter-ácido carboxílico, sales o ésteres de ácido α -sulfograso, agentes tensioactivos de tipo aminoácido, agentes tensioactivos de éster fosfato, y sales de amonio cuaternario que incluyen aquéllas que tienen de 3 a 4 sustituyentes alquílicos y hasta 1 sustituyente alquílico sustituido con fenilo. En la Solicitud de Patente Británica nº 2.094.826 A se describen adecuados agentes tensioactivos catiónicos y sales de ácidos grasos de cadena larga. La composición puede contener de aproximadamente 1 a aproximadamente 20 por ciento en peso de dichos agentes tensioactivos catiónicos y sales de ácidos grasos de cadena larga.

Agentes mejoradores de la limpieza

A. Agentes secuestradores divalentes

La composición puede contener de aproximadamente 0 a aproximadamente 50 por ciento en peso de uno o más

5 componentes mejoradores de la limpieza seleccionados del grupo que consiste en sales de metal alcalino y sales de alanolamina de los compuestos siguientes: fosfatos, fosfonatos, fosfonocarboxilatos, sales de aminoácidos, aminopoliacetatos, electrolitos de alto peso molecular, polímeros que no se disocian, sales de ácidos dicarboxílicos y sales de aluminosilicato. En la Solicitud de Patente Británica nº 2.094.826 A se describen agentes secuestradores divalentes adecuados.

B. Alcalis o electrolitos inorgánicos

10 La composición puede contener de aproximadamente 1 a aproximadamente 50 por ciento en peso, preferiblemente de aproximadamente 5 a aproximadamente 30 por ciento en peso, con respecto a la composición, de una o más sales de metal alcalino de los compuestos siguientes como alcalis o electrolitos inorgánicos: silicatos, carbonatos y sulfatos así como alcalis orgánicos tales como trietanolamina, dietanolamina, monoetanolamina y trisopropanolamina.

Agentes anti-redepósito

15 La composición puede contener de aproximadamente 0,1 a aproximadamente 5 por ciento en peso de uno o más de los compuestos siguientes como agentes anti-redepósito: polietilenglicol, poli (alcohol vinílico), polivinilpirrolidona y carboximetilcelulosa.

Entre ellos, una combinación de carboximetilcelulosa y/o polietilenglicol con la composición de celulasa de la presente invención proporciona una composición especialmente útil para eliminar la suciedad.

Agentes blanqueadores

20 El uso de la celulasa de la presente invención en combinación con un agente blanqueador tal como monopersulfato potásico, percarbonato sódico, perborato sódico, aducto de sulfato sódico/peróxido de hidrógeno y aducto de cloruro sódico/peróxido de hidrógeno o/y un colorante blanqueador fotosensible tal como una sal cincica o aluminica de ftalocianina sulfonada mejora adicionalmente los efectos detergentes. Similarmente, se pueden utilizar agentes blanqueadores y catalizadores del blanqueo como los descritos en el Documento EP 684.304.

Azuletes y colorantes fluorescentes

25 Si es necesario, se pueden incorporar diversos azuletes y colorantes fluorescentes a la composición. En la Solicitud de Patente Británica nº 2.094.826 A se describen azuletes y colorantes fluorescentes adecuados.

Inhibidores del apelmazamiento

30 Se pueden incorporar los siguientes inhibidores del apelmazamiento al detergente en polvo: sales del ácido p-toluenosulfónico, sales del ácido xilenosulfónico, sales del ácido acético, sales del ácido sulfosuccínico, talco, sílice finamente pulverizada, sílices amorfas, arcilla, silicato cálcico (tal como Micro-Cell de Johns Manville Co.), carbonato de calcio y óxido de magnesio.

Antioxidantes

35 Los antioxidantes incluyen, por ejemplo, terc-butil-hidroxitolueno, 4,4'-butiliden-bis(6-terc-butil-3-metilfenol), 2,2'-butiliden-bis(6-terc-butil-4-metilfenol), cresol monoestirenado, cresol diestirenado, fenol monoestirenado, fenol diestirenado y 1,1-bis(4-hidroxi-fenil)ciclohexano.

Agentes solubilizantes

Los agentes solubilizantes incluyen, por ejemplo, alcoholes inferiores tales como etanol, sales de bencenosulfonato, sales de alquil inferior-bencenosulfonato tales como sales de p-toluenosulfonato, glicoles tales como propilenglicol, sales de acetilbencenosulfonato, acetamidas, amidas del ácido piridinadicarboxílico, sales de benzoato, y urea.

40 La composición detergente de la presente invención puede ser utilizada en un amplio intervalo de pHs, de pHs ácidos a alcalinos. En una realización preferida, la composición detergente de la presente invención se puede utilizar en medios de lavado detergentes ligeramente ácidos, neutros o alcalinos que tengan un pH de más de 5 a no más de aproximadamente 12.

45 Con las composiciones detergentes de esta invención se pueden utilizar, si se desea, aparte de los ingredientes anteriores, perfumes, tampones, conservantes, colorantes y similares. Tales componentes se emplean convencionalmente en las cantidades hasta ahora utilizadas en la técnica.

50 Cuando una base detergente usada en la presente invención está en forma de polvo, puede ser una que se prepare mediante cualesquier métodos de preparación conocidos, incluyendo un método de secado por pulverización y un método de granulación. Se prefiere la base detergente obtenida particularmente mediante el método de secado por pulverización, el método de aglomeración, el método de mezclamiento en estado seco o los métodos sin torre de secado por pulverización. La base detergente obtenida mediante el método de secado por pulverización no está

restringida con respecto a las condiciones de preparación. La base detergente obtenida mediante el método de secado por pulverización es de gránulos huecos que se obtienen al pulverizar una suspensión acuosa de ingredientes térmicamente resistentes, tales como agentes tensioactivos y agentes mejoradores de la limpieza, en un espacio caliente. Después del secado por pulverización se pueden añadir perfumes, enzimas, agentes blanqueadores y agentes alcalinos inorgánicos mejoradores de la limpieza. Con una base detergente granular y muy densa obtenida de modo tal como mediante el método de secado por pulverización-granulación o aglomeración, también se pueden añadir varios ingredientes después de la preparación de la base.

Cuando la base detergente es un líquido, puede ser una disolución homogénea o una dispersión no homogénea. Para suprimir la descomposición de la carboximetilcelulosa por la celulasa en el detergente, es deseable que la carboximetilcelulosa sea granulada o revestida antes de su incorporación a la composición.

Las composiciones detergentes de esta invención pueden ser incubadas con tejidos que contienen celulosa, por ejemplo, tejidos sucios, en usos industriales y domésticos con las temperaturas, tiempos de reacción y proporciones de líquido convencionalmente empleados en estos entornos.

Los detergentes de acuerdo con la presente invención pueden ser adicionalmente formulados como un detergente para prelavado en la disolución apropiada en un pH intermedio en que existe actividad suficiente para proporcionar las mejoras deseadas: suavizado, eliminación de bolitas, prevención de la formación de bolitas, eliminación de fibras superficiales o limpieza. Cuando la composición detergente es una composición para remojo previo (por ejemplo, prelavado o pretratamiento), sea una composición líquida, pulverizable, en gel o en pasta, la enzima celulasa se emplea generalmente en una cantidad de aproximadamente 0,0001 a aproximadamente 1 por ciento en peso con respecto al peso total de la composición para remojo previo o pretratamiento. En dichas composiciones se puede emplear opcionalmente un agente tensioactivo, y, cuando se emplea, está generalmente presente en una concentración de aproximadamente 0,005 a aproximadamente 20 por ciento en peso con respecto al peso total de la composición para remojo previo. El resto de la composición comprende componentes convencionales usados en la composición para remojo previo, es decir, diluyente, tampones, otras enzimas (proteasas) y similares en sus concentraciones convencionales.

Se contempla que las composiciones que comprenden enzimas celulasas descritas en esta memoria se puedan emplear en uso doméstico como una composición independiente adecuada para restaurar el color de tejidos deslucidos (véase, por ejemplo, la Patente de EE.UU. n° 4.738.682), y se puedan también emplear en un quitamanchas y para la eliminación de bolitas y la anti-formación de bolitas (prevención de la formación de bolitas).

El uso de la celulasa de acuerdo con la invención puede ser particularmente eficaz en aditivos para pienso y en el procesamiento de pulpa y papel. Estas aplicaciones industriales adicionales se describen, por ejemplo, en la Publicación PCT n° 95/16360 y en la Concedida Patente Finlandesa n° 87372, respectivamente.

Con objeto de ilustrar adicionalmente la presente invención y sus ventajas, se proporcionan los ejemplos específicos siguientes con el conocimiento de que se presentan para ilustrar la presente invención y no se deben considerar en modo alguno restrictivos de su alcance.

Ejemplos

Se presentan los ejemplos siguientes para ilustrar, pero no limitar, la invención reivindicada.

Ejemplo 1

Recogida y procesamiento de muestras

En este ejemplo se ilustra cómo recoger muestras y procesarlas para obtener DNA suficiente para crear un banco de cDNA.

Se recogieron muestras de agua (250 ml) de la zona litoral de Sonachi (Crater) Lake, Kenia, usando un vaso de precipitados de acero inoxidable, de 250 ml de capacidad, montado en el extremo de una pértiga flexible y extensible de 1 m de longitud, y se colocaron en recipientes de plástico de cierre hermético (Whirlpak) para su transporte al laboratorio a temperatura ambiental. La temperatura de las aguas superficiales era 28 °C, con un pH de 10 y una conductividad de 7,23 mS·cm⁻¹ (a 27 °C).

Para recoger la flora microbiana, se filtró agua (750 ml) de Sonachi (Crater) Lake, Kenia, *in situ* (usando una bomba de vacío de funcionamiento manual) a través de una serie de filtros de membrana estériles (47 mm de diámetro) compuestos de nitrato de celulosa o acetato de celulosa, de tamaño de poro decreciente, hasta que se detuvo todo el flujo de agua. La serie de filtros tenía 8 µm, 3 µm y 0,22 µm. Los filtros de membrana individuales se pusieron inmediatamente en 10 ml de tampón para estabilización de células (TES), frío y estéril, que contenía Tris·HCl 10 mM, pH de 8,0, EDTA 1 mM y NaCl al 5% en peso/volumen, en tubos universales de plástico estériles de 30 ml de capacidad, y se mantuvieron sobre hielo en una nevera portátil refrigerada hasta que pudieron ser procesados más tarde, normalmente a las 4 horas del muestreo. Se dispersó el material microbiano de los filtros mediante mezclado energético y formación de remolinos con glóbulos de vidrio estériles (5 ml) y se recogieron las células

como sedimento en tubos de microcentrífuga por centrifugación a 13.000g durante 5 minutos. Se repartió el material microbiano en partes alícuotas en tubos de microcentrífuga, en volúmenes que se estimó que contenían el equivalente de 10^8 a 10^9 células bacterianas, obteniéndose un total de 12 tubos. Se extrajo el DNA usando el kit GenomicPrep™ para aislamiento de DNA de Células y Tejidos (Amersham Pharmacia Biotech, Piscataway, New Jersey, EE.UU.) siguiendo las instrucciones del fabricante. Se resuspendieron las células de cada tubo en 600 µl de la Disolución para Lisis Celular proporcionada y se incubó a 80 °C durante 5 minutos para lisar las células. Las muestras preparadas mediante este método son estables a temperatura ambiental durante al menos 18 meses, y se transportaron de nuevo al laboratorio en esta forma. La extracción de DNA se completó mediante tratamiento con RNasa A, precipitación de proteínas y precipitación del DNA con isopropanol siguiendo el protocolo del fabricante. Cada sedimento de DNA se disolvió en 100 µl de tampón de Tris 10 mM estéril, pH de 8,5.

Se estimó la producción de DNA desarrollando muestras de 5 µl en un gel de agarosa al 0,5% en peso/volumen y comparando con cantidades conocidas de DNA genómico bacteriano. Se reunieron las muestras, obteniéndose un total de aproximadamente 20 µg de DNA. Puesto que las producciones eran pequeñas, el material fue complementado con aproximadamente un 30% de material extra extraído de las muestras de agua que fueron recogidas al mismo tiempo como material *in situ* y almacenadas a 4 °C en el laboratorio hasta que fueran necesarias. Esta cantidad de DNA, aproximadamente 30 µg, fue la cantidad de material de partida que experimentos preliminares habían mostrado que era necesaria para llevar a cabo la prueba y la digestión de restricción en masa y el fraccionamiento por tamaños para obtener material suficiente para la construcción del banco.

Ejemplo 2

Construcción del banco

En el ejemplo siguiente se detalla cómo preparar un banco de DNA para uso en la exploración y detección de nuevas secuencias en *E. coli*.

Preparación de DNA

El DNA reunido fue utilizado para la construcción del banco de DNA genómico. El DNA purificado fue sometido a digestión parcial con *Sau3A1* para obtener un tamaño medio de fragmento de aproximadamente 5 kb. El DNA restringido fue sometido a fraccionamiento por tamaños mediante electroforesis en agarosa al 0,5% en TAE (Tris-acetato 0,04 M, EDTA 0,001 M, pH de 8,0). El material en el intervalo de 1,5 a 10 kb fue escindido y fue repuesto en un pocillo del mismo corte de tamaños en una parte no usada del gel de agarosa, y fue concentrado hasta una banda estrecha mediante una corriente eléctrica invertida. Se escindió la banda de DNA y se extrajo el DNA usando el kit QIAEXII de QIAGEN (Crawley, Reino Unido) para extracción de gel, siguiendo las directrices del fabricante. El DNA eluido fue precipitado con etanol y resuspendido en tampón de Tris·HCl 10 mM, pH de 8,5.

Preparación de bancos Lambda

El DNA restringido fue clonado en un vector Lambda usando el kit de vector ZAP-Express™ (predigerido con *Bam*H1 y tratado con fosfatasa alcalina) y el extracto de empaquetamiento Gigapak® III Gold (Stratagene, Amsterdam, Holanda) siguiendo el protocolo del fabricante. Los bancos primarios fueron multiplicados según el protocolo sembrando partes alícuotas que contenían $\sim 5 \times 10^4$ pfu con la cepa huésped XL1-Blue MRF' de *E. coli* en placas Petri de 150 mm y eluyendo el fago en tampón. Los bancos multiplicados se guardaron en dimetilsulfóxido al 7% en volumen/volumen a -80 °C tras su congelación en nitrógeno líquido. El título primario total fue $1,8 \times 10^6$ pfu y, después de la multiplicación, $6,8 \times 10^9$ pfu·ml⁻¹.

Evaluación de la calidad del banco

Se escindió el vector fagémido pBK-CMV del banco Lambda ZAP usando el fago auxiliar ExAssist (Stratagene) del modo descrito por el fabricante y se utilizó para infectar la cepa XL0LR de *E. coli*. Los clones que contenían el plásmido fueron aislados mediante siembra en placas de agar Luria – Bertani (LB) que contenía 50 µg·ml⁻¹ de kanamicina. Se utilizó la exploración azul:blanco en presencia de Xgal (5-bromo-4-cloro-3-indol-β-D-galactósido) e IPTG (isopropiltio-β-D-galactósido) para determinar la eficacia de la clonación. Si no se ha clonado DNA en el vector Lambda, el gen de β-galactosidasa se expresa en presencia del inductor IPTG, lo que da lugar a la escisión del sustrato Xgal análogo para producir un pigmento azul en la colonia. Sin embargo, si se ha clonado exitosamente un fragmento del DNA genómico en el vector Lambda, altera el gen de modo que no se produce la enzima y la colonia permanece blanca. Por lo tanto, la proporción de colonias azules a blancas puede ser utilizada para calcular el porcentaje de clones que contienen un inserto. Para este banco, la exploración azul:blanco dio una proporción de 7 colonias azules a 286 colonias blancas, lo que indica que el 97% de los clones contenía un inserto del DNA genómico. Se seleccionaron al azar veinticuatro colonias y se preparó DNA plasmídico utilizando el sistema de purificación de DNA Wizard® Plus SV Miniprep (Promega, Southampton, Reino Unido). Para determinar tamaños de inserto se utilizó un análisis de restricción usando *Pst*I y *Hind*III que flanquean el sitio de clonación de *Bam*H1, seguido de electroforesis en gel de agarosa. Se halló que un clon de los 24 no tenía inserto detectable. El resto tenía insertos que variaban de 1,5 kb a 8,0 kb.

Ejemplo 3

Exploración del banco en cuanto a celulasas

Bancos de DNA en el fagémido pBK-CMV fueron explorados en cuanto a actividad celulasa en un ensayo de los clones de *E. coli* en placa. Para detectar la actividad celulasa, los bancos genómicos fueron sembrados en agar LB que contenía kanamicina, carboximetilcelulosa (sal sódica de baja viscosidad; Sigma, Poole, Reino Unido) al 0,5% en peso/volumen e IPTG (15 µl de una disolución 0,5 M extendidos sobre la superficie del agar en una placa Petri de 7 cm de diámetro). Después de un cultivo durante la noche a 37 °C, las colonias fueron cubiertas con 3 ml de agarosa fundida al 0,7% en peso/volumen disuelta en agua, que había sido enfriada a 50 °C. Una vez que se hubo hecho esto, las placas fueron inundadas con una disolución de rojo Congo al 0,1% en peso/volumen durante 30 minutos, lo que fue seguido de 2 lavados con NaCl 1 M. Los clones positivos que presentaban actividad celulasa extracelular estaban rodeados por un halo amarillo contra un fondo rojo (R. Teather y P. J. Wood, Applied & Environmental Microbiology, 43: 770-780, 1982).

La exploración de 110.000 clones de *E. coli* pBK-CMV produjo 4 zonas de aclaramiento que indicaban posibles colonias productoras de celulasa. Tres de éstas fueron exitosamente recuperadas como clones productores de celulasa después de homogeneizar el tapón de agar separado de la zona aclarada, sembrar con asa en estrías para obtener colonias individuales, y confirmar el fenotipo mediante la prueba de rojo Congo.

Ejemplo 4

Caracterización de un clon positivo para celulasa

Se aisló DNA plasmídico de los tres clones positivos para celulasa y se determinó el tamaño de los insertos mediante digestión de restricción del modo anteriormente descrito. Los tres tenían el mismo tamaño (aproximadamente 3,5 kb) y los fragmentos del mismo tamaño después de la digestión, según se determinó mediante electroforesis en gel. Esto indicaba que los tres productos de aislamiento eran idénticos, obtenidos por multiplicación de un solo clon. Esto fue confirmado por el primer ciclo de secuenciación del DNA plasmídico (usando sitios de cebador en el plásmido pBK-CMV). Esto fue llevado a cabo por el Protein and Nucleic Acid Chemistry Laboratory de la Universidad de Leicester usando la química de terminadores "BigDye" de Perkin Elmer y el secuenciador de DNA automatizado ABI modelo 377. Se obtuvo la cobertura completa de la secuencia mediante "paseo con cebadores" desde ambos extremos, el 5' y el 3', del inserto. Se editó la secuencia usando el editor multiseccional Seqed™ de Applied Biosystems, versión 1.0.3. Se ensambló la secuencia con programas en el GCG Wisconsin Package, versión 10.2-UNIX, disponible en la Universidad de Leicester. Esto permitió identificar un inserto de DNA ambiental de 3410 bases nucleotídicas (Figura 1).

Ejemplo 5

Identificación del gen de celulasa

Se identificaron posibles marcos de lectura abiertos (ORF; del inglés, open reading frames) en la secuencia de nucleótidos del DNA ambiental insertado del clon 029cel usando la función ORF Find del programa MapDraw (DNASTAR, Brighton, Massachusetts, EE.UU.) u ORF Search del Vector NTI Suite de programas (InforMax®, North Bethesda, Maryland, EE.UU.).

Esto permitió identificar un ORF compuesto de 1746 nucleótidos correspondientes a una proteína de 581 aminoácidos, comenzando en la posición 3004 de la secuencia del inserto y acabando en la posición 1259. La secuencia de este ORF fue escindida usando EditSeq (DNASTAR) y fue examinada mediante programas BLAST.

En la Figura 2 se muestra la secuencia de nucleótidos de este ORF.

Un examen de la secuencia de nucleótidos usando el programa BLASTn, que compara una secuencia de nucleótidos en cuestión con una base de datos de secuencias de nucleótidos no redundantes, indicó una falta de similitud significativa con cualquier secuencia conocida.

Un examen de la secuencia de nucleótidos usando el programa BLASTx, que compara los productos de traducción conceptuales de seis marcos de una secuencia de nucleótidos en cuestión (ambas cadenas) con una base de datos de secuencias proteicas, reveló sorprendentemente una similitud muy pequeña (no más del 29%) con diversas endocelulasas bacterianas. La máxima puntuación de alineación reveló una identidad del 25% (132 aminoácidos) con una región de 527 aminoácidos de CelJ, una enzima que comprende 1601 aminoácidos, el componente catalítico más grande del celulosoma de *Clostridium thermocellum* (identificación de proteína: BAA1207070.1; acceso: D83704.1). Es muy probable que una enzima con esta pequeña homología no hubiera sido detectada usando métodos convencionales en que se emplean sondas de DNA basadas en secuencias de genes de celulasa conocidas, especialmente dada la altísima diversidad de celulasas ya caracterizadas.

En la Figura 3 se muestra la proteína traducida, compuesta de 581 aminoácidos.

Ejemplo 6

Caracterización de la enzima

Influencia de la sal

5 Células de *E. coli* pBK-CMV que contienen el gen 029cel son suspendidas en 5 ml de tampón (Tris-HCl 20 mM, pH de 8,0, NaCl 500 mM, EDTA 0,1 mM, Triton X-100 al 0,1%) y son deshechas por sonicación sobre hielo. Los extractos de sonicación son examinados mediante un ensayo de difusión en agar sobre carboximetilcelulosa (CMC) en diferentes concentraciones de NaCl. Se ponen los extractos de sonicación (100 µl) y diluciones 1 en 10 en pocillos excavados en placas de CMC-agar que contienen cantidades variables de NaCl. Se incuban las placas a 37 °C durante 16 horas y se miden en milímetros las zonas de aclaramiento resultantes, que indican hidrólisis de
10 celulosa. La celulasa 029cel es activa a lo largo del intervalo de NaCl al 0 – 25% en peso/volumen, aunque la actividad en NaCl al 25% en peso/volumen es menor que la actividad en NaCl al 0%.

Influencia del pH

15 Se investiga la influencia del pH sobre la actividad celulasa usando el método del gradiente de pHs en placa descrito por Grant y Tindall ("Isolation of alkaliphilic bacteria" en *Microbial Growth and Survival in Extreme Environments*, Academic Press, Londres, 1980, páginas 27-36). Se vierte un medio de agar que contiene CMC hasta una profundidad de 1 cm en placas Petri cuadradas y se deja endurecer. Se corta una cavidad uniforme de 1 cm de anchura desde un borde de la placa y se vierte agar que contiene Na₂CO₃·10H₂O al 20% en peso/volumen y NaOH 0,2 M (preparado mezclando volúmenes iguales de NaOH 0,4 M/Na₂CO₃·10H₂O al 40% en peso/volumen y agar al 4% en peso/volumen estériles a 60 °C) en la cavidad. Se desarrollan las placas a 37 °C durante la noche para
20 permitir que se forme un gradiente uniforme de pH 12 a pH 7. Para examinar la tolerancia de la celulasa 029cel al pH, se corta una cavidad estrecha a través del gradiente (agar) en ángulo recto con respecto a la cavidad original y se llena con 1 ml de extracto de células sometidas a sonicación. Se deja que se desarrollen las placas durante la noche a 37 °C. Las placas son tratadas con rojo Congo durante 30 minutos para visualizar la zona de hidrólisis de celulosa. Los resultados indican que la celulasa 029cel es activa en un pH de aproximadamente 11,5.

25

REIVINDICACIONES

1. Un polinucleótido aislado, seleccionado del grupo que consiste en:
- (a) una secuencia de ácido nucleico que tiene una identidad secuencial de al menos 85% con respecto a la ID. SEC. nº 1, o el complemento de la misma;
- 5 (b) una secuencia de ácido nucleico que codifica, o es complementaria de una secuencia que codifica, un polipéptido 029cel que tiene una identidad secuencial de al menos 85% con respecto a la secuencia de aminoácidos ID. SEC. nº 3;
- (c) una secuencia de ácido nucleico que codifica, o es complementaria de una secuencia que codifica, un polipéptido 029cel que tiene una identidad secuencial de al menos 90% con respecto a la secuencia de aminoácidos ID. SEC. nº 3;
- 10 (d) una secuencia de ácido nucleico que codifica, o es complementaria de una secuencia que codifica, un polipéptido 029cel que tiene una identidad secuencial de al menos 95% con respecto a la secuencia de aminoácidos ID. SEC. nº 3; y
- (e) una secuencia de ácido nucleico que codifica, o es complementaria de una secuencia que codifica, un polipéptido 029cel que tiene la secuencia de aminoácidos ID. SEC. nº 3;
- 15 en donde dicho polinucleótido aislado codifica un polipéptido que tiene la actividad biológica de una celulasa y en donde la identidad se determina mediante el programa CLUSTAL-W en MacVector versión 6.5, ejecutado con los parámetros por omisión que incluyen una penalización por hueco abierto de 10,0, una penalización por hueco extendido de 0,1 y una matriz de similitudes BLOSUM 30.
- 20 2. Un polinucleótido aislado, seleccionado del grupo que consiste en:
- (a) una secuencia de ácido nucleico ID. SEC. nº 1, o el complemento de la misma;
- (b) una secuencia de ácido nucleico que se hibrida, bajo condiciones de rigor elevado, con la secuencia ID. SEC. nº 1, o el complemento de la misma;
- (c) una secuencia de ácido nucleico ID. SEC. nº 2, o el complemento de la misma; y
- 25 (d) una secuencia de ácido nucleico que se hibrida, bajo condiciones de rigor elevado, con la secuencia ID. SEC. nº 2, o el complemento de la misma,
- en donde dicho polinucleótido aislado codifica un polipéptido que tiene la actividad biológica de una celulasa y en donde la hibridación se lleva a cabo a 42 °C en formamida al 50%, SSC 6X, disolución de Denhardt 5X, SDS al 0,5% y 100 µg/ml de DNA portador desnaturalizado, seguida de lavado dos veces en SSPE 2X y SDS al 0,5% a temperatura ambiental y dos veces más en SSPE 0,1X y SDS al 0,5% a 42 °C.
- 30 3. El polinucleótido aislado de la Reivindicación 1, en donde el polinucleótido se selecciona del grupo: mRNA, DNA, cDNA, DNA genómico, y un análogo antisentido de los mismos.
4. El polinucleótido aislado de la Reivindicación 3, en donde dicho polinucleótido es una molécula de RNA.
5. El polinucleótido aislado de la Reivindicación 1 que codifica una enzima que tiene actividad celulasa, en donde la enzima es aislada de una fuente del género *Trichoderma*.
- 35 6. El polinucleótido aislado de la Reivindicación 5, en donde la enzima es aislada de *Trichoderma reesei*.
7. Una construcción de expresión que comprende una secuencia polinucleotídica que codifica una secuencia de aminoácidos que tiene actividad celulasa y
- (i) que tiene una identidad secuencial de al menos 85% con respecto a la secuencia de aminoácidos ID. SEC. nº 3, o
- 40 (ii) que es complementaria de una secuencia de nucleótidos que tiene una identidad secuencial de al menos 85% con respecto a una secuencia de nucleótidos que codifica la secuencia de aminoácidos ID. SEC. nº 3,
- en donde la identidad se determina mediante el programa CLUSTAL-W en MacVector versión 6.5, ejecutado con los parámetros por omisión que incluyen una penalización por hueco abierto de 10,0, una penalización por hueco extendido de 0,1 y una matriz de similitudes BLOSUM 30.
- 45 8. Un vector de expresión que comprende el polinucleótido de la Reivindicación 1.
9. Un vector de expresión que comprende un polinucleótido aislado de la Reivindicación 1, operativamente unido a secuencias de control reconocidas por una célula huésped transformada con el vector.

10. Un vector de expresión según la Reivindicación 9, que comprende una secuencia polinucleotídica reguladora que incluye una secuencia promotora derivada de un gen de glucosa isomerasa de *Actinoplanes*, una secuencia señal derivada de un gen de celulasa de *Streptomyces*, y una secuencia polinucleotídica que codifica una celulasa 029cel.
- 5 11. Un vector que comprende la construcción de expresión de la Reivindicación 7.
12. Una célula huésped transformada con el vector de la Reivindicación 8.
13. La célula huésped de la Reivindicación 12, que es una célula procariótica.
14. La célula huésped de la Reivindicación 12, que es una célula eucariótica.
- 10 15. Un polipéptido 029cel sustancialmente purificado con la actividad biológica de una celulasa, que comprende una secuencia seleccionada del grupo que consiste en:
- (a) una secuencia de aminoácidos que tiene una identidad secuencial de al menos 85% con respecto a la secuencia de aminoácidos ID. SEC. nº 3;
- (b) una secuencia de aminoácidos que tiene una identidad secuencial de al menos 90% con respecto a la secuencia de aminoácidos ID. SEC. nº 3;
- 15 (c) una secuencia de aminoácidos que tiene una identidad secuencial de al menos 95% con respecto a la secuencia de aminoácidos ID. SEC. nº 3; y
- (d) una secuencia de aminoácidos ID. SEC. nº 3;
- en donde la identidad se determina mediante el programa CLUSTAL-W en MacVector versión 6.5, ejecutado con los parámetros por omisión que incluyen una penalización por hueco abierto de 10,0, una penalización por hueco extendido de 0,1 y una matriz de similitudes BLOSUM 30.
- 20 16. El polipéptido celulasa sustancialmente purificado según la Reivindicación 15 o un derivado que es obtenible de un microorganismo del género *Bacillus*.
17. Un método para producir una celulasa, que comprende las operaciones de:
- (a) cultivar la célula huésped según la Reivindicación 12 en un medio de cultivo adecuado bajo unas condiciones adecuadas para producir la celulasa; y
- 25 (b) obtener dicha celulasa producida.
18. El método de la Reivindicación 17, en donde la célula huésped es una célula de hongo filamentoso o levadura.
19. El método de la Reivindicación 17, en donde la célula huésped es una bacteria.
20. El método de la Reivindicación 19, en donde la bacteria es del género *Streptomyces*.
- 30 21. Una enzima purificada que tiene actividad celulasa, preparada mediante el método de la Reivindicación 17.
22. Un oligonucleótido antisentido complementario de un RNA mensajero que codifica un polipéptido que tiene la secuencia ID. SEC. nº 3, en donde, tras la exposición a una célula huésped que produce celulasa, dicho oligonucleótido disminuye o inhibe la producción de celulasa por dicha célula huésped.
23. El oligonucleótido antisentido de la Reivindicación 22, en donde la célula huésped es un hongo filamentoso.
- 35 24. Una composición detergente, composición que comprende un agente tensioactivo y un polipéptido seleccionado del grupo que consiste en:
- (a) una secuencia de aminoácidos que tiene una identidad secuencial de al menos 85% con respecto a la secuencia de aminoácidos ID. SEC. nº 3;
- (b) una secuencia de aminoácidos que tiene una identidad secuencial de al menos 90% con respecto a la secuencia de aminoácidos ID. SEC. nº 3;
- 40 (c) una secuencia de aminoácidos que tiene una identidad secuencial de al menos 95% con respecto a la secuencia de aminoácidos ID. SEC. nº 3; y
- (d) una secuencia de aminoácidos ID. SEC. nº 3;
- en donde la identidad se determina mediante el programa CLUSTAL-W en MacVector versión 6.5, ejecutado con los

parámetros por omisión que incluyen una penalización por hueco abierto de 10,0, una penalización por hueco extendido de 0,1 y una matriz de similitudes BLOSUM 30.

25. El detergente según la Reivindicación 24, en donde dicho detergente es un detergente para la colada.
26. El detergente según la reivindicación 24, en donde dicho detergente es un detergente para la vajilla.
- 5 27. Un aditivo para pienso, que comprende una celulasa según la Reivindicación 15.
28. Un método para tratar la pulpa de madera, que comprende poner dicha pulpa de madera en contacto con una celulasa según la Reivindicación 15.
29. Un método para convertir biomasa en azúcares, que comprende poner dicha biomasa en contacto con una celulasa según la Reivindicación 15.
- 10 30. El método de la Reivindicación 29, que comprende además la generación de jarabe de maíz con alto contenido de fructosa.
31. Un método para producir etanol, método que comprende las operaciones de:
 - (a) poner una composición de biomasa en contacto con una composición enzimática que comprende 029cel según la Reivindicación 15 para producir una disolución de azúcar;
 - 15 (b) añadir un microorganismo fermentador a la disolución de azúcar; y
 - (c) cultivar el microorganismo fermentador bajo unas condiciones suficientes para producir etanol.

Figura 1A

Secuencia de nucleótidos del DNA ambiental insertado (029cel)

ID. SEC. nº 1

| | | | | | |
|--------------|-------------|-------------|-------------|-------------|------|
| ATCAACACGC | TGAAAGTAA | TTTCAAGGGT | AAGGCCATCG | GTTGCCGCCG | 50 |
| GGGTAGAAAT | GTGCGTTGG | ATTTTCGTTGA | GCGGCGTCGC | CGGCGTTCCA | 100 |
| CCGAGGGCAT | AGCGCAGCAG | GTTGGCGATG | CCACCGGTGA | GGCCTTCGGG | 150 |
| GCCGCCTACG | ATGTTGTGCT | CAGCCGCCCA | TGCGATGTAG | CCGTCCGGCT | 200 |
| CGGGTTCGCT | CGCGGGGGTG | AAGAAGACAA | TGTCGTGCGAG | ATAAAGGTTG | 250 |
| CCGCTTCCGC | TCTCAACGCC | GCCGAGGTTG | AATTGGATTT | CGCAAATTCT | 300 |
| CGTTAGGTCC | AGCACGGAAT | CGCCGACGAG | GTCGGCTATG | GGAATCTGAA | 350 |
| TGCGCCCATATA | GGTTGGGTA | CGCGGAAGGG | ACACGTAGGG | ACCCACTTTG | 400 |
| TCATTGGGCG | AGACGAGCCG | GACAAAGATT | TGGTGCGCCG | CCTGCGAGGG | 450 |
| GCCTTGGAGG | GCGAGAGAAA | GGTACGTGAG | GGCGCTGATG | TCGTGCCTGG | 500 |
| GACCGTCTCC | CCAGTTGTCG | AGATTGAGCC | CAAATCCGGC | CCACCATCCG | 550 |
| GCGATAGTGT | AGCTCCAATG | GTAGTGACGC | TCACCCTCGA | AGCCGCCGCT | 600 |
| GGAGAGTTCC | TGCAAGCCGT | CGCCCCAAAT | GCCCGTGATG | AGCGTTGCCT | 650 |
| CGTCACGGTA | GATCACAAGT | TCGGCGGCGG | GTGCCGGGGG | AAGATCGCCT | 700 |
| TGAGTGATCA | CGAGAGTGGC | GGTGGCGCTG | CCTTCGTGAT | TAGGGTCCGGT | 750 |
| AATGGTGGCG | ACGACCGTGT | AGCTACCGGG | CCCCACTGGC | GCATGGGTGG | 800 |
| AACCGTTGTA | GGTAAAGGAG | ACGTCAAGCC | CCACGGGATG | GGTCTCGGCA | 850 |
| AGAGCGGCCCT | TGGGGGTGCC | GTCGAAAACG | TGTTCCAAAT | TGGAGAGCGT | 900 |
| GATGGTGGCG | GGTGCCTTGA | GCACAGTCAC | AGAAACAGTG | GATTGCACGG | 950 |
| GATCGTGCGC | TGCCGTGTCT | GCAGGTGTGA | AGACCACGCT | GTAAAAACGG | 1000 |
| GTTCCGGCGG | ACGGTGCAAG | GCCGGACAGG | ACAAAGGCAA | AGTCCGCCGGG | 1050 |
| GACGGCGGCT | ACTCCGCCGC | TCAGGCCGGC | CTCCGCAAGG | GTTTGCCCGA | 1100 |
| AGGTGATGGG | TGCGGCTGTG | GGCCACATCT | CCACAAGGCC | GGTGTCCCCC | 1150 |
| TCGTACGCA | CCGGCATGAG | GGCGGAGAGG | AGATGAATGT | AACTGGCTTG | 1200 |
| GTAATTGATG | TCGGGCTCGG | TGATTTCCCA | TGAGTTCTCC | GGCCAAAAC | 1250 |
| CATTCCAATC | AAGGTAGGCT | TTTTGCACGG | GTTGGTCTCG | GATCGCCTGA | 1300 |
| ATGCTTCCGC | TGTATTTGGG | CATTGGGACC | CGCCCCGAAAG | AAAACCAGGA | 1350 |
| GCGGGACCGT | AGAGTGAAGT | GAGGGCATTG | TCCCAGTCCG | GCCATCGCGG | 1400 |
| AACCAATGGT | GGTAGATTTT | ATTGGCTGCA | CGGTGAGCGC | CGCTGGCATA | 1450 |
| CATGTTGCTA | AGATAGACCA | TGCCATTGG | GTTCACTCCG | TGGAGATAGT | 1500 |
| GCAGGTAGCC | CATCGCGGCA | TCGCGATGCG | CGGCCGCGTC | GGCGGGGTTG | 1550 |
| AGCCCAAGCC | TCCGTACCCC | CTCGAAGAAA | AAGCCAGCCT | GAGACTTTGT | 1600 |
| TTTGTTTCGAG | CCCCACGTGT | AATCCTGATC | C TTCAGGTAG | GCGCGGTAGG | 1650 |
| CGTCGGTCTG | GTTATTCCAT | GCACCGAGAA | ACTCCCCACC | GTTTATAGAA | 1700 |
| GCCGCCATCC | GGTTGCGGAT | GTCGGCAGAG | ACGCTAGGCG | TCGCTCCCGG | 1750 |
| GAGGGTTCGTG | TAGTGGGCGA | GAGCTTTTTG | TAGCTCACCT | TGAAAGGGGA | 1800 |
| AGAAATACCA | CCACTGCACG | GGCTCCATAT | CGAGATAGCG | CACATCGAAG | 1850 |
| AAATCGCGAT | AGACCGCACC | GCCCGTGCGC | TCGAAGAGCA | TGGCGGCGGC | 1900 |
| CATCACACGG | TTGGCTAGCG | TATCGTGGGC | ATTGCGCGAG | GGGCTCACGG | 1950 |
| AAGCAAATCC | GGTGTGTCG | AAAGGCACAT | GAGGATGGAC | CATGGTCCAA | 2000 |
| TTCCATGCGG | CGATGGCAGC | GGATTCGAGG | GTGACGGCAT | AATCGCTCAT | 2050 |
| GCCTACGCTC | TCAAAGACAG | TCGCCCCGAG | GGCGAAAGCG | GCGGCAGCCA | 2100 |
| TGGCAGTGGC | CTCGGTGCGAG | ACGGGGCCGT | AGTAACGCGG | ATGGGTGTCG | 2150 |
| GTGCTCGGCG | GGCTGGCGCT | CTGGTGCCCC | GTCACGGAAA | CTTTCCCGAG | 2200 |
| AATAGCCCCG | CTCGGCTCCT | GCATGCGTAA | GAGCCAGTCC | ATTCCCCATT | 2250 |
| TGACTTCGTC | AAGCAGGTGCG | GGGACACCGT | TGCCGGATTC | CGGGATGCCA | 2300 |
| AAATCATCGG | TAAAGACGTC | AGGCCGCCCT | TGATAGGCAA | GGAGCAGCTC | 2350 |
| CAGGATGACG | CGCCCCGTCC | ACTCGCTGTA | CTTGTTGAAA | TCGCCCGCAT | 2400 |

Figura 1B

| | | | | | |
|------------|------------|------------|------------|------------|------|
| CGAACCAACC | GCCGCTGAGA | TCGCGCTCCA | AGGAGGCATT | CCCCATATCC | 2450 |
| CAGATGGGGC | GGCTGGCGAC | GTCCTGCGGG | TGAGAAGCGG | CATCGGCCCA | 2500 |
| GTTCGCGTGG | GCGTAGGGCA | CCTCCTTGGC | AAACCCGGAG | CGCTGATAGA | 2550 |
| AGAACATGCG | CACGGCCTCG | CGCAGGACAA | CATCGTAAAC | ATCCGCGCCA | 2600 |
| ATGGCGAAAC | TATCGGAATG | AGTGTTGTTG | GCAGGATCGT | GGATGCGGTA | 2650 |
| GTGGCCGGGC | TCGGCAACTA | CCGTAAAATC | AAACCACCAC | ACGCGGTCTC | 2700 |
| CCGATTGAAT | ATGGATGGCG | CCGCCGTTCC | ACGGGACCGG | TGAGCCGGAG | 2750 |
| AAAACCACGA | CGCCATCGTT | CACGCGACGG | ACCTCCAGCG | TTGCGCCGGG | 2800 |
| GCTGTAGCTC | TCGGCGCTGT | TCCAGCCAAT | CTGCGGGTCG | GCGATCACCG | 2850 |
| CCACCTTGGT | GGCATCGGCG | GGGTAACCGA | ATTGGTCGAT | GCGGATTTTA | 2900 |
| TCGGTGTGGG | TGGAGGCGAC | GAGGGCGGAG | CTGCCCATGA | GCAGCAAGAA | 2950 |
| AAAGCCCCT | GTCGGCCCGA | TACCAAAAAA | ACGAATAGGG | AGAGAAAAAT | 3000 |
| TCATAGCAGG | ATGTGGATAC | GGAAAGGGGG | AAAACGGTGC | AAAGACCCAA | 3050 |
| GCCCAACGCT | TGGCGAAAAC | TGGATGGTTG | GTTTATCAAG | AAAAGCGCTT | 3100 |
| TTGAGCCAAA | AGCTGCGGGC | AATCCTTATT | GCGTTTCACA | ATATTTTCAC | 3150 |
| ATCGTCGGCG | GCACGACTTT | TCGATGGGCG | ACTTGACAGC | GTATTCTCTC | 3200 |
| AGGCGCGAGG | CTGCAAACCT | TATGAAAAAA | GGCCCGCGCA | GCGATCTGTC | 3250 |
| CCCGGTCAAA | ATCCAGTCAA | GGTTTGTTCA | AGGGTTTGAG | GTCTGATAGA | 3300 |
| GGCACAGTCG | AGCCATCAGC | AGTCGCATTG | AGTAGGGTTG | TTGGAGAAAG | 3350 |
| TGTGCAAATG | ACCGCTGCCG | AAGGAACTGT | GGAGACAAAA | AGCATATTTT | 3400 |
| CCTCGCCAAG | | | | | 3410 |

Figura 2

La secuencia de nucleótidos del ORF de 029cel

| | | | | | |
|-------------|------------|-------------|------------|------------|------|
| ATGAATTTTT | CTCTCCCTAT | TCGTTTTTTTT | GGTATCGGGC | CGACAGCGGG | 50 |
| CTTTTTCTTG | CTGCTCATGG | GCAGCTCCGC | CCTCGTCGCC | TCCACCCACA | 100 |
| CCGATAAAAT | CCGCATCGAC | CAATTCGGTT | ACCCCGCCGA | TGCCACCAAG | 150 |
| GTGGCGGTGA | TCGCCGACCC | GCAGATTGGC | TGGAACAGCG | CCGAGAGCTA | 200 |
| CAGCCCCGGC | GCAACGCTGG | AGGTCCGTCG | CGTGAACGAT | GGCGTCGTGG | 250 |
| TTTTCTCCGG | CTCACCGGTC | CCGTGGAACG | GCGGCGCCAT | CCATATTCAA | 300 |
| TCGGGAGACC | GCGTGTGGTG | GTTTGATTTT | ACGGTAGTTG | CCGAGCCCGG | 350 |
| CCACTACCGC | ATCCACGATC | CTGCCAACAA | CACTCATTCC | GATAGTTTCG | 400 |
| CCATTGGCGC | GGATGTTTAC | GATGTTGTCC | TGCGCGAGGC | CGTGCGCATG | 450 |
| TTCTTCTATC | AGCGCTCCGG | GTTTGCCAAG | GAGGTGCCCT | ACGCCACGC | 500 |
| GAACTGGGGC | GATGCCGCTT | CTCACCCGCA | GGACGTCGCC | AGCCGCCCCA | 550 |
| TCTGGGATAT | GGGGAATGCC | TCCTTGGAGC | GCGATCTCAG | CGGCGGTGG | 600 |
| TTCGATGCGG | GCGATTTCAA | CAAGTACAGC | GAGTGGACGG | GGCGCGTCAT | 650 |
| CCTGGAGCTG | CTCCTTGCCT | ATCAAGGGCG | GCCTGACGTC | TTTACCGATG | 700 |
| ATTTTGGCAT | CCCGGAATCC | GGCAACGGTG | TCCCGACCT | GCTTGACGAA | 750 |
| GTCAAATGGG | GAATGGACTG | GCTCTTACGC | ATGACGAGC | CGAGCGGGGC | 800 |
| TATTCTCGGG | AAAGTTTCCG | TGACGGGGCA | CCAGAGCGCC | AGCCCGCCGA | 850 |
| GCACCGACAC | CCATCCGCGT | TACTACGGCC | CCGTCTCGAC | CGAGGCCACT | 900 |
| GCCATGGGCTG | CCGCCGCTTT | CGCCCTCGGG | GCGACTGTCT | TTGAGAGCGT | 950 |
| AGGCATGAGC | GATTATGCCG | TCACCCTCGA | ATCCGCTGCC | ATCGCCGCAT | 1000 |
| GGAATGGAC | CATGGTCCAT | CCTCATGTGC | CTTTCGACAA | CACCGGATTT | 1050 |
| GCTTCCGTGA | GCCCCTCGCG | CAATGCCAC | GATACGCTAG | CCAACCGTGT | 1100 |
| GATGGCCGCC | GCCATGCTCT | TCGAGCGCAC | GGGCGGTGCG | GTCTATCGCG | 1150 |
| ATTTCTTCGA | TGTGCGCTAT | CTCGATATGG | AGCCCGTGCA | GTGGTGGTAT | 1200 |
| TTCTTCCCCT | TTCAAGGTGA | GCTACAAAAA | GCTCTCGCCC | ACTACACGAC | 1250 |
| CCTCCCGGGA | GCGACGCCTA | GCGTCTCTGC | CGACATCCGC | AACCGGATGG | 1300 |
| CGGCTTCTAT | AAACGGTGGG | GAGTTTCTCG | GTGCATGGAA | TAACCAGACC | 1350 |
| GACGCCTACC | GCGCCTACCT | GAAGGATCAG | GATTACACGT | GGGGCTCGAA | 1400 |
| CAAAACAAAG | TCTCAGGCTG | GCTTTTTCTT | CGAGGGGGTA | CGGAGGCTTG | 1450 |
| GGCTCAACCC | CGCCGACGCG | GCCGCGCATC | GCGATGCCGC | GATGGGCTAC | 1500 |
| CTGCACTATC | TCCACGGAGT | GAACCCAATG | GGCATGGTCT | ATCTTAGCAA | 1550 |
| CATGTATGCC | AGCGGCGCTG | ACCGTGCAGC | CAATGAAATC | TACCACCATT | 1600 |
| GGTTCGCGA | TGGCCGGACT | GGGACAATGC | CCTCACTTCA | CTCTACGGTC | 1650 |
| CCGCTCCTGG | TTTTCTTTCG | GGCGGGTCCC | AATGCCCAA | TACAGCGGAA | 1700 |
| GCATTCAGGC | GATCCGAGAC | CAACCCGTGC | AAAAGCCTA | CCTTGA | 1746 |

ES 2 530 518 T3

Figura 3

La proteína 029cel traducida, compuesta de 581 aminoácidos

| | | | | | |
|-------------|------------|------------|------------|------------|-----|
| MNFSLPPIRFF | GIGPTAGFFL | LLMGSSALVA | STHTDKIRID | QFGYPADATK | 50 |
| VAVIADPQIG | WNSAESYSPG | ATLEVRRVND | GVVVFSGSPV | PWNGGAIHIQ | 100 |
| SGDRVWWFDF | TVVAEPGHYR | IHDPANNTHS | DSFAIGADVY | DVVLREAVRM | 150 |
| FFYQRSGFAK | EVPYAHANWA | DAASHPQDVA | SRPIWDMGNA | SLERDLSGGW | 200 |
| FDAGDFNKYS | EWVGRVILEL | LLAYQGRPDV | FTDDFGIPES | GNGVPDLLDE | 250 |
| VKWGMDWLLR | MQEPSGAILG | KVSVTGHQSA | SPPSTDTHPR | YYGPVSTEAT | 300 |
| AMAAAFAALG | ATVFESVGMS | DYAVTLESAA | IAAWNWTMVH | PHVFPDNTGF | 350 |
| ASVSPSRNAH | DTLANRVMAA | AMLFERTGGA | VYRDFFDVRY | LDMEPVQWWY | 400 |
| FFPFQGELQK | ALAHYTTLPG | ATPSVSADIR | NRMAASINGG | EFLGAWNNQT | 450 |
| DAYRAYLKDQ | DYTWGSNKTK | SQAGFFFEGV | RRLGLNPADA | AAHRDAAMGY | 500 |
| LHYLHGVPNPM | GMVYLSNMYA | SGADRAANEI | YHHWFRDGRT | GTMPSLHSTV | 550 |
| PLLVFFRAGP | NAQIQRKHSG | DPRPTRAKSL | P | | 581 |