

República Federativa do Brasil
Ministério do Desenvolvimento, Indústria
e do Comércio Exterior
Instituto Nacional da Propriedade Industrial

(21) **PI0610031-7 A2**

(22) Data de Depósito: 27/04/2006

(43) Data da Publicação: 18/05/2010
(RPI 2054)



(51) *Int.Cl.:*

C12N 9/42 (2010.01)

C12N 9/24 (2010.01)

C12N 15/00 (2010.01)

C11D 3/386 (2010.01)

(54) Título: **POLIPEPTÍDEO ISOLADO, POLINUCLEOTÍDEO ISOLADO, CONSTRUÇÃO DE ÁCIDO NUCLEICO, CÉLULA HOSPEDEIRA RECOMBINANTE, MÉTODOS PARA PRODUZIR O POLIPEPTÍDEO E PARA PRODUZIR UM MUTANTE DE UMA CÉLULA PRECURSORA, CÉLULA MUTANTE, MÉTODOS PARA PRODUZIR UMA PROTEÍNA, E PARA PRODUZIR UM POLINUCLEOTÍDEO, POLINUCLEOTÍDEO, MUTANTE, COMPOSIÇÃO DETERGENTE, E, MÉTODO PARA DEGRADAR BIOMASSA CONTENDO CELULOSE E HEMI-CELULOSE**

(57) Resumo: POLIPEPTÍDEO ISOLADO, POLINUCLEOTÍDEO ISOLADO, CONSTRUÇÃO DE ÁCIDO NUCLEICO, CÉLULA HOSPEDEIRA RECOMBINANTE, MÉTODOS PARA PRODUZIR O POLIPEPTÍDEO E PARA PRODUZIR UM MUTANTE DE UMA CÉLULA PRECURSORA, CÉLULA MUTANTE, MÉTODOS PARA PRODUZIR UMA PROTEÍNA, E PARA PRODUZIR UM POLINUCLEOTÍDEO, POLINUCLEOTÍDEO MUTANTE, COMPOSIÇÃO DETERGENTE, E, MÉTODO PARA DEGRADAR BIOMASSA CONTENDO CELULOSE E HEMI-CELULOSE. A presente invenção diz respeito a polipeptídeos isolados tendo atividade de endoglicanase e polinucleotídeos isolados que codificam os polipeptídeos. A invenção também diz respeito a construções de ácido nucleico, vetores e células hospedeiras que compreendem os polinucleotídeos assim como a métodos para produzir e usar os polipeptídeos.

(30) Prioridade Unionista: 27/04/2005 US 60/675,601

(73) Titular(es): Novozymes, Inc.

(72) Inventor(es): ALFREDO LOPEZ DE LEON , ELENA VLASENKO, HANSHU DING, MICHEAL REY , PAUL HARRIS

(74) Procurador(es): Momsen, Leonardos & CIA.

(86) Pedido Internacional: PCT US2006016244 de 27/04/2006

(87) Publicação Internacional: WO 2006/116682 de 02/11/2006

“POLIPEPTÍDEO ISOLADO, POLINUCLEOTÍDEO ISOLADO, CONSTRUÇÃO DE ÁCIDO NUCLEICO, VETOR DE EXPRESSÃO RECOMBINANTE, CÉLULA HOSPEDEIRA RECOMBINANTE, MÉTODOS PARA PRODUZIR O POLIPEPTÍDEO E PARA PRODUZIR
5 UM MUTANTE DE UMA CÉLULA PRECURSORA, CÉLULA MUTANTE, MÉTODOS PARA PRODUZIR UMA PROTEÍNA E PARA PRODUZIR UM POLINUCLEOTÍDEO, POLINUCLEOTÍDEO MUTANTE, PLANTA TRANSGÊNICA, PARTE DE PLANTA OU CÉLULA VEGETAL, COMPOSIÇÃO DETERGENTE, E, MÉTODO
10 PARA DEGRADAR BIOMASSA CONTENDO CELULOSE E HEMI-CELULOSE”

Declaração como para os Direitos às Invenções Feitas Sob Pesquisa e Desenvolvimento Patrocinados pelo Governo Federal

Esta invenção foi feita com apoio do Governo sob o
15 Subcontrato NREL Nº ZCO-30017-02, Prime Contract DE-AC36-98G010337 outorgado pelo Departamento de Energia. O governo tem certos direitos nesta invenção.

Fundamentos da invenção

Campo da invenção

20 A presente invenção diz respeito a polipeptídeos isolados tendo atividade de endoglicanase e polinucleotídeos isolados que codificam os polipeptídeos. A invenção também diz respeito às construções de ácido nucleico, vetores e células hospedeiras que compreendem os polinucleotídeos assim como aos métodos para produzir e usar os polipeptídeos.

25 **Descrição da Técnica Relacionada**

A celulose é um polímero da glicose do açúcar simples covalentemente ligada pelas ligações beta-1,4. Muitos microorganismos produzem enzimas que hidrolisam glicanos ligados em beta. Estas enzimas incluem endoglicanases, celobioidrolases e beta-glicosidases. As

endoglicanases digerem o polímero de celulose em locais aleatórios, abrindo-o ao ataque pelas celobioidrolases. As celobioidrolases seqüencialmente liberam moléculas de celobiose das extremidades do polímero de celulose. A celobioidrolase I é uma 1,4-D-glicano celobioidrolase (E.C. 3.2.1.91) a
5 atividade da qual catalisa a hidrólise de ligações 1,4-beta-D-glicosídicas na celulose, celotetrioise ou qualquer polímero contendo glicose ligado em beta-1,4, liberando celobiose das extremidades redutoras da cadeia. A celobioidrolase II é uma 1,4-D-glicano celobioidrolase (E.C. 3.2.1.91) a atividade da qual catalisa a hidrólise de ligações 1,4-beta-D-glicosídicas na
10 celulose, celotetrioise ou qualquer polímero contendo glicose ligado em beta-1,4, liberando celobiose das extremidades não redutoras da cadeia. A celobiose é um dímero ligado em beta-1,4 solúvel em água de glicose. As beta-glicosidases hidrolisam celobiose a glicose.

A conversão de matérias primas celulósicas em etanol tem as
15 vantagens da pronta disponibilidade de grandes quantidades de matéria prima, o desejo de evitar queimar ou depositar em aterro sanitário os materiais e a limpeza do combustível de etanol. Madeira, resíduos agrícolas, safras herbáceas e resíduos sólidos municipais têm sido considerados como matérias primas para a produção de etanol. Estes materiais primariamente consistem de
20 celulose, hemi-celulose e lignina. Uma vez que a celulose é convertida para glicose, a glicose é facilmente fermentada pela levedura em etanol.

Kvesitadaze *et al.*, 1995, *Applied Biochemistry e Biotechnology* 50: 137-143, descrevem a isolamento e as propriedades de uma endoglicanase termoestável a partir de uma cepa mutante termofílica de
25 *Thielavia terrestris*. Gilbert *et al.*, 1992, *Bioresource Technology* 39: 147-154, descrevem a caracterização das enzimas presentes no sistema da celulose de *Thielavia terrestris* 255B. Breuil *et al.*, 1986, *Biotechnology Letters* 8: 673-676, descrevem a produção e localização de celulasas e beta-glicosidases das cepas de *Thielavia terrestris* 0464 e NRRL 8126.

Seria uma vantagem na técnica identificar novas endoglicanases tendo propriedades melhoradas, tais como taxas de hidrólise melhoradas, melhor estabilidade térmica, absorção reduzida para lignina e a capacidade para hidrolisar componentes não celulósicos da biomassa, tais como hemi-celulose, além de hidrolisar a celulose. As endoglicanases com uma ampla faixa de atividades colaterais sobre a hemi-celulose podem ser especialmente benéficas para melhorar o rendimento da hidrólise global de substratos de biomassa ricos em hemi-celulose, complexos.

É um objetivo da presente invenção fornecer polipeptídeos melhorados tendo atividade de endoglicanase e polinucleotídeos que codificam os polipeptídeos.

Sumário da invenção

A presente invenção diz respeito a polipeptídeos isolados tendo atividade de endoglicanase selecionados do grupo que consiste de:

(a) um polipeptídeo que compreende uma seqüência de aminoácido que tenha pelo menos 60% de identidade com o polipeptídeo maduro que codifica a seqüência da SEQ ID NO: 2;

(b) um polipeptídeo que é codificado por uma seqüência de nucleotídeo que hibridiza sob pelo menos condições de estringência média com (i) o polipeptídeo maduro que codifica a seqüência da SEQ ID NO: 1, (ii) a seqüência de DNA genômico que compreende o polipeptídeo maduro que codifica a seqüência da SEQ ID NO: 1 ou (iii) um filamento complementar de (i) ou (ii); e

(c) uma variante que compreende uma substituição, deleção e/ou inserção conservativa de um ou mais aminoácidos do polipeptídeo maduro da SEQ ID NO: 2.

A presente invenção também diz respeito a polinucleotídeos isolados que codificam polipeptídeos tendo atividade de endoglicanase, selecionado do grupo que consiste de:

(a) um polinucleotídeo que codifica um polipeptídeo que compreende uma sequência de aminoácido que tenha pelo menos 60% de identidade com o polipeptídeo maduro da SEQ ID NO: 2;

5 (b) um polinucleotídeo tendo pelo menos 60% de identidade com o polipeptídeo maduro que codifica a sequência da SEQ ID NO: 1; e

(c) um polinucleotídeo que hibridiza sob pelo menos condições de estringência baixa com (i) o polipeptídeo maduro que codifica a sequência da SEQ ID NO: 1, (ii) a sequência de DNA genômico que compreende o polipeptídeo maduro que codifica a sequência da SEQ ID NO: 10 1 ou (iii) um filamento complementar de (i) ou (ii).

Em um aspecto preferido, o polipeptídeo maduro tem os aminoácidos de 18 a 336 da SEQ ID NO: 2. Em um outro aspecto preferido, o polipeptídeo maduro que codifica a sequência tem os nucleotídeos de 52 a 1008 da SEQ ID NO: 1.

15 A presente invenção também diz respeito a construções de ácido nucleico, vetores de expressão recombinantes e células hospedeiras recombinantes que compreendem os polinucleotídeos.

A presente invenção também diz respeito a métodos para produzir um tal polipeptídeo tendo atividade de endoglicanase que 20 compreendem: (a) cultivar uma célula hospedeira recombinante compreendendo uma construção de ácido nucleico que compreende um polinucleotídeo que codifica o polipeptídeo sob condições condutivas para a produção do polipeptídeo; e (b) recuperar o polipeptídeo.

A presente invenção também diz respeito a métodos de usar os 25 polipeptídeos tendo atividade de endoglicanase em detergentes e na conversão de celulose para glicose.

A presente invenção diz respeito ainda às construções de ácido nucleico que compreendem um gene que codifica uma proteína, em que o gene é operavelmente ligado a uma sequência de nucleotídeo que codifica um

peptídeo de sinal que compreende ou que consiste dos aminoácidos de 1 a 17 da SEQ ID NO: 2, em que o gene é estranho para a sequência de nucleotídeo.

Descrição Resumida das Figuras

As Figuras 1A e 1B mostram a sequência de cDNA e a sequência de aminoácido deduzida de uma endoglicanase de *Thielavia terrestris* NRRL 8126 (CEL7F) (SEQ ID NOs: 1 e 2, respectivamente).

A Figura 2 mostra um mapa de restrição de pTter7F.

A Figura 3 mostra um mapa de restrição de pAlLo1.

A Figura 4 mostra um mapa de restrição de pBANe10.

A Figura 5 mostra um mapa de restrição de pAlLo2.

A Figura 6 mostra um mapa de restrição de pAlLo22.

A Figura 7 mostra a conversão relativa de beta-glicano (1% p/v) depois de 2 horas de hidrólise no pH 5,5 e 60°C.

A Figura 8 mostra a conversão relativa de beta-glicano (1% p/v) depois de 24 horas de hidrólise no pH 5,5 e 60°C.

Definições

Atividade de endoglicanase: O termo “atividade de endoglicanase” é aqui definido como uma endo-1,4-beta-D-glicano 4-glicanoidrolase (E.C. No. 3.2.1.4) que catalisa a endoidrólise de ligações 1,4-beta-D-glicosídicas em celulose, derivados de celulose (tais como carboximetil celulose e hidroxietil celulose), liquenina, ligações beta-1,4 em beta-1,3 glicanos mistos tais como beta-D-glicanos ou xiloglicanos de cereal e outro material vegetal contendo componentes celulósicos. Para os propósitos da presente invenção, a atividade de endoglicanase é determinada usando a hidrólise da carboximetil celulose (CMC) de acordo com o procedimento de Ghose, 1987, Pure and Appl. Chem. 59: 257-268. Uma unidade de atividade de endoglicanase é definida como 1,0 mol de açúcares redutores produzido por minuto a 50°C, pH 4,8.

Em um aspecto preferido, os polipeptídeos da presente

invenção tendo atividade de endoglicanase têm ainda atividade enzimática contra um ou mais substratos selecionados do grupo que consiste de xilano, xiloglicano, arabinosilano, galactana, galactomanana, dextrano e quitina. A atividade dos polipeptídeos tendo atividade de endoglicanase nestes substratos
5 de polissacarídeo é determinada como a quantidade relativa de corante liberada de diferentes substratos tingidos com AZCL depois incubando os substratos (5 g por litro) com um polipeptídeo tendo atividade de endoglicanase da presente invenção (1 mg de proteína por g de substrato) por 1 e 92 horas sem agitação no pH 5,0 (50 mM de acetato de sódio) e 50°C. A
10 liberação de corante foi determinada pela medição da absorbância a 590 nm.

Em um aspecto mais preferido, os polipeptídeos da presente invenção tendo atividade de endoglicanase têm ainda atividade enzimática contra xilano. Em um outro aspecto mais preferido, os polipeptídeos da presente invenção tendo atividade de endoglicanase têm ainda atividade
15 enzimática contra xiloglicano. Em um outro aspecto mais preferido, os polipeptídeos da presente invenção tendo atividade de endoglicanase têm ainda atividade enzimática contra arabinosilano. Em um outro aspecto mais preferido, os polipeptídeos da presente invenção tendo atividade de endoglicanase têm ainda atividade enzimática contra galactana. Em um outro
20 aspecto mais preferido, os polipeptídeos da presente invenção tendo atividade de endoglicanase têm ainda atividade enzimática contra galactomanana. Em um outro aspecto mais preferido, os polipeptídeos da presente invenção tendo atividade de endoglicanase têm ainda atividade enzimática contra dextrano. Em um outro aspecto mais preferido, os polipeptídeos da presente invenção
25 tendo atividade de endoglicanase têm ainda atividade enzimática contra quitina. Em um outro aspecto mais preferido, os polipeptídeos da presente invenção tendo atividade de endoglicanase têm ainda atividade enzimática contra xilano, xiloglicano, arabinosilano, galactana, galactomanana, dextrano e quitina.

Os polipeptídeos da presente invenção têm pelo menos 20%, preferivelmente pelo menos 40%, mais preferivelmente pelo menos 50%, mais preferivelmente pelo menos 60%, mais preferivelmente pelo menos 70%, mais preferivelmente pelo menos 80%, ainda mais preferivelmente pelo menos 90%, o mais preferivelmente pelo menos 95% e ainda mais preferivelmente pelo menos 100% da atividade de endoglicanase do polipeptídeo que consiste da seqüência de aminoácido mostrada como os aminoácidos de 18 a 336 da SEQ ID NO: 2.

Família 7 glicosídeo hidrolase ou Família GH7: Os termos “Família 7 glicosídeo hidrolase” ou “Família GH7” ou “CEL7F” são aqui definidos como um polipeptídeo que cai dentro da Família 7 da glicosídeo hidrolase de acordo com Henrissat B., 1991, A classification of glycosyl hydrolases based on amino-acid sequence similarities, *Biochem. J.* 280: 309-316 e Henrissat B. e Bairoch A., 1996, Updating the seqüência-based classification of glycosyl hydrolases, *Biochem. J.* 316: 695-696.

Polipeptídeo isolado: O termo “polipeptídeo isolado” como aqui usado refere-se a um polipeptídeo que é pelo menos 20% puro, preferivelmente pelo menos 40% puro, mais preferivelmente pelo menos 60% puro, ainda mais preferivelmente pelo menos 80% puro, o mais preferivelmente pelo menos 90% puro e ainda mais preferivelmente pelo menos 95% puro, como determinado pelo SDS-PAGE.

Polipeptídeo substancialmente puro: O termo “polipeptídeo substancialmente puro” indica aqui uma preparação de polipeptídeo que contenha no máximo 10%, preferivelmente no máximo 8%, mais preferivelmente no máximo 6%, mais preferivelmente no máximo 5%, mais preferivelmente no máximo 4%, mais preferivelmente no máximo 3%, ainda mais preferivelmente no máximo 2%, o mais preferivelmente no máximo 1% e ainda mais preferivelmente no máximo 0,5% em peso de outro material de polipeptídeo com o qual o mesmo esteja naturalmente associado. É, portanto,

preferido que o polipeptídeo substancialmente puro seja pelo menos 92% puro, preferivelmente pelo menos 94% puro, mais preferivelmente pelo menos 95% puro, mais preferivelmente pelo menos 96% puro, mais preferivelmente pelo menos 96% puro, mais preferivelmente pelo menos 97% puro, mais preferivelmente pelo menos 98% puro, ainda mais preferivelmente pelo menos 99%, o mais preferivelmente pelo menos 99,5% puro e ainda mais preferivelmente 100% puro em peso do material de polipeptídeo total presente na preparação.

Os polipeptídeos da presente invenção estão preferivelmente em uma forma substancialmente pura. Em particular, é preferido que os polipeptídeos estejam na “forma essencialmente pura”, isto é, que a preparação de polipeptídeo seja essencialmente livre de outro material de polipeptídeo com o qual o mesmo esteja naturalmente associado. Isto pode ser realizado, por exemplo, pela preparação do polipeptídeo por meio de métodos recombinantes bem conhecidos ou pelos métodos de purificação clássicos.

Aqui, o termo “polipeptídeo substancialmente puro” é sinônimo com os termos “polipeptídeo isolado” e “polipeptídeo na forma isolada.”

Polipeptídeo maduro: O termo “polipeptídeo maduro” é aqui definido como um polipeptídeo tendo atividade de endoglicanase que está na sua forma final a seguir da tradução e quaisquer modificações pós-traducionais, tais como processamento de terminal N, truncagem de terminal C, glicosilação, etc.

Identidade: A relacionabilidade entre duas seqüências de aminoácido ou entre duas seqüências de nucleotídeo é descrita pelo parâmetro “identidade”.

Para os propósitos da presente invenção, o grau de identidade entre duas seqüências de aminoácido é determinado pelo método Clustal (Higgins, 1989, CABIOS 5: 151-153) usando o software LASERGENE®

MEGALIGN[®] (DNASTAR, Inc., Madison, WI) com uma tabela de identidade e os seguintes parâmetros de alinhamento múltiplo: Penalidade de intervalo de 10 e penalidade de comprimento de intervalo de 10. Os Parâmetros de alinhamento aos pares são Ktuple = 1, penalidade de intervalo = 3, janela = 5 e diagonais = 5.

Para os propósitos da presente invenção, o grau de identidade entre duas seqüências de nucleotídeo é determinada pelo método de Wilbur-Lipman (Wilbur e Lipman, 1983, Proceedings of the National Academy of Science USA 80: 726-730) usando o software LASERGENE[®] MEGALIGN[®] (DNASTAR, Inc., Madison, WI) com uma tabela de identidade e os seguintes parâmetros de alinhamento múltiplo: Penalidade de intervalo de 10 e penalidade de comprimento de intervalo de 10. Parâmetros de alinhamento aos pares são Ktuple = 3, penalidade de intervalo = 3 e janela = 20.

Fragmento de polipeptídeo: O termo “fragmento de polipeptídeo” é aqui definido como um polipeptídeo tendo um ou mais aminoácidos deletados do término amino e/ou carboxila da SEQ ID NO: 2 ou uma seqüência homóloga desta, em que o fragmento tem atividade de endoglicanase. Preferivelmente, um fragmento contém pelo menos 270 resíduos de aminoácido, mais preferivelmente pelo menos 285 resíduos de aminoácido e o mais preferivelmente pelo menos 300 resíduos de aminoácido, por exemplo, aminoácidos de 18 a 336 da SEQ ID NO: 2.

Subseqüência: O termo “subseqüência” é aqui definido como uma seqüência de nucleotídeo tendo um ou mais nucleotídeos deletados da extremidade 5' e/ou 3' da SEQ ID NO: 1 ou uma seqüência homóloga desta, em que a subseqüência codifica um fragmento de polipeptídeo tendo atividade de endoglicanase. Preferivelmente, uma subseqüência contém pelo menos 810 nucleotídeos, mais preferivelmente pelo menos 855 nucleotídeos e o mais preferivelmente pelo menos 900 nucleotídeos.

Variante alélica: O termo “variante alélica” indica aqui

qualquer de duas ou mais formas alternativas de um gene que ocupa o mesmo local cromossômico. A variação alélica surge naturalmente através da mutação e pode resultar em polimorfismo dentro das populações. As mutações de gene podem ser silenciosas (nenhuma mudança no polipeptídeo codificado) ou pode codificar polipeptídeos tendo seqüências de aminoácido alteradas. Uma variante alélica de um polipeptídeo é um polipeptídeo codificado por uma variante alélica de um gene.

Polinucleotídeo isolado: O termo “polinucleotídeo isolado” como aqui usado refere-se a um polinucleotídeo que é pelo menos 20% puro, preferivelmente pelo menos 40% puro, mais preferivelmente pelo menos 60% puro, ainda mais preferivelmente pelo menos 80% puro, o mais preferivelmente pelo menos 90% puro e ainda mais preferivelmente pelo menos 95% puro, como determinado pela eletroforese em agarose.

Polinucleotídeo substancialmente puro: O termo “polinucleotídeo substancialmente puro” como aqui usado refere-se a uma preparação de polinucleotídeo livre de outros nucleotídeos estranhos ou não desejado e em um forma adequada para o uso dentro de sistemas de produção de proteína geneticamente engendrada. Assim, um polinucleotídeo substancialmente puro contém no máximo 10%, preferivelmente no máximo 8%, mais preferivelmente no máximo 6%, mais preferivelmente no máximo 5%, mais preferivelmente no máximo 4%, mais preferivelmente no máximo 3%, ainda mais preferivelmente no máximo 2%, o mais preferivelmente no máximo 1% e ainda mais preferivelmente no máximo 0,5% em peso de outro material de polinucleotídeo com o qual o mesmo esteja naturalmente associado. Um polinucleotídeo substancialmente puro, entretanto, pode incluir que ocorrem naturalmente regiões 5' e 3' não traduzidas, tais como promotores e terminadores. É preferido que o polinucleotídeo substancialmente puro seja pelo menos 90% puro, preferivelmente pelo menos 92% puro, mais preferivelmente pelo menos 94% puro, mais

preferivelmente pelo menos 95% puro, mais preferivelmente pelo menos 96% puro, mais preferivelmente pelo menos 97% puro, ainda mais preferivelmente pelo menos 98% puro, o mais preferivelmente pelo menos 99% e ainda mais preferivelmente pelo menos 99,5% puro em peso. Os polinucleotídeos da presente invenção estão preferivelmente em uma forma substancialmente pura. Em particular, é preferido que os polinucleotídeos aqui divulgados estejam na “forma essencialmente pura”, isto é, que a preparação de polinucleotídeo esteja essencialmente livre de outro material de polinucleotídeo com o qual o mesmo esteja naturalmente associado. Aqui, o termo “polinucleotídeo substancialmente puro” é sinônimo com o termo “polinucleotídeo isolado” e “polinucleotídeo na forma isolada.” Os polinucleotídeos podem ser de origem genômica, cDNA, RNA, semi-sintética, sintética ou quaisquer combinações destes.

Poli-peptídeo maduro que codifica a seqüência: O termo “poli-peptídeo maduro que codifica a seqüência” é aqui definido como uma seqüência de nucleotídeo que codifica um poli-peptídeo maduro tendo atividade de endoglicanase.

cDNA: O termo “cDNA” é aqui definido como uma molécula de DNA que pode ser preparada pela transcrição reversa a partir de uma molécula de mRNA madura, combinada obtida de uma célula eucariótica. O cDNA carece de seqüências de intron que estão usualmente presentes no DNA genômico correspondente. O transcrito de RNA inicial, primário é um precursor para o mRNA que é processado através de uma série de etapas antes de aparecer como mRNA maduro combinado. Estas etapas incluem a remoção de seqüências de intron por um processo chamado combinação. O cDNA derivado de mRNA carece, portanto, de quaisquer seqüências de intron.

Construção de ácido nucleico: O termo “construção de ácido nucleico” como aqui usado refere-se a uma molécula de ácido nucleico, de

filamento único ou duplo, que é isolado de um gene que ocorre naturalmente ou que é modificado para conter segmentos de ácidos nucleicos em uma maneira que de outro modo não existiria na natureza. O termo construção de ácido nucleico é sinônimo com o termo “cassete de expressão” quando a
5 construção de ácido nucleico contém as seqüências de controle requeridas para a expressão de uma seqüência codificadora da presente invenção.

Seqüência de controle: O termo “seqüências de controle” é aqui definido como incluindo todos os componentes, que são necessários ou vantajosos para a expressão de um polinucleotídeo que codifica um
10 polipeptídeo da presente invenção. Cada seqüência de controle pode ser nativa ou estranha para a seqüência de nucleotídeo que codifica o polipeptídeo ou nativa ou estranha entre si. Tais seqüências de controle incluem, mas não são limitadas a, uma líder, seqüência de poliadenilação, seqüência propeptídica, promotor, seqüência de peptídeo de sinal e
15 terminador de transcrição. No mínimo, as seqüências de controle incluem um promotor e sinais de parada transcrpcionais e traducionais. As seqüências de controle podem ser fornecidas com ligadores com o propósito de introduzir sítios de restrição específicos que facilitem a ligação das seqüências de controle com a região codificadora da seqüência de nucleotídeo que codifica
20 um polipeptídeo.

Operavelmente ligado: O termo “operavelmente ligado” indica aqui uma configuração em que uma seqüência de controle é colocada em uma posição apropriada em relação à seqüência codificadora da seqüência de polinucleotídeo tal que a seqüência de controle direcione a expressão da
25 seqüência codificadora de um polipeptídeo.

Seqüência codificadora: Quando aqui usado o termo “seqüência codificadora” significa uma seqüência de nucleotídeo, que diretamente especifica a seqüência de aminoácido do seu produto de proteína. Os limites da seqüência codificadora são no geral determinados por uma

matriz de leitura aberta, que usualmente começa com o códon de partida ATG ou códons de partida alternativos tais como GTG e TTG e termina com um códon de parada tal como TAA, TAG e TGA. A sequência codificadora pode ser um DNA, cDNA ou sequência de nucleotídeo recombinante.

5 Expressão: O termo “expressão” inclui qualquer etapa envolvida na produção do polipeptídeo incluindo, mas não limitado a, transcrição, modificação pós transcricional, tradução, modificação pós traducional e secreção.

10 Vetor de expressão: O termo “vetor de expressão” é aqui definido como uma molécula de DNA linear ou circular que compreende um polinucleotídeo que codifica um polipeptídeo da invenção e que é operavelmente ligado aos nucleotídeos adicionais que fornecem a sua expressão.

15 Célula hospedeira: O termo “célula hospedeira”, como aqui usado, inclui qualquer tipo de célula que é suscetível à transformação, transfecção, transdução e outros com uma construção de ácido nucleico ou vetor de expressão que compreende um polinucleotídeo da presente invenção.

20 Modificação: O termo “modificação” significa aqui qualquer modificação química do polipeptídeo que consiste do polipeptídeo maduro da SEQ ID NO: 2 ou uma sequência homóloga desta assim como manipulação genética do DNA que codifica este polipeptídeo. A modificação pode ser substituições, deleções e/ou inserções de um ou mais aminoácidos assim como substituições de uma ou mais cadeias laterais de aminoácido.

25 Variante artificial: Quando aqui usado, o termo “variante artificial” significa um polipeptídeo tendo atividade de endoglicanase produzida por um organismo que expressa uma sequência de nucleotídeo modificada da SEQ ID NO: 1 ou uma sequência homóloga desta ou a região codificadora madura desta. A sequência de nucleotídeo modificada é obtida através da intervenção humana pela modificação da sequência de nucleotídeo

divulgada na SEQ ID NO: 1 ou uma sequência homóloga desta ou a região codificadora madura desta.

Descrição Detalhada da Invenção

Polipectídeos Tendo Atividade de Endoglicanase

5 Em um primeiro aspecto, a presente invenção diz respeito a polipeptídeos isolados que compreendem uma sequência de aminoácido que tem um grau de identidade com o polipeptídeo maduro da SEQ ID NO: 2 de pelo menos 60%, preferivelmente pelo menos 65%, mais preferivelmente pelo menos 70%, mais preferivelmente pelo menos 75%, mais preferivelmente pelo menos 80%, mais preferivelmente pelo menos 85%, ainda mais preferivelmente pelo menos 90%, o mais preferivelmente pelo menos 95% e
10 ainda mais preferivelmente pelo menos 97%, 98% ou 99%, tendo atividade de endoglicanase (daqui em diante “polipeptídeos homólogos”). Em um aspecto preferido, os polipeptídeos homólogos têm uma sequência de aminoácido que difere em dez aminoácidos, preferivelmente em cinco aminoácidos, mais preferivelmente em quatro aminoácidos, ainda mais preferivelmente em três aminoácidos, o mais preferivelmente em dois aminoácidos e ainda mais preferivelmente em um aminoácido do polipeptídeo maduro da SEQ ID NO:
15 2.

20 Um polipeptídeo da presente invenção preferivelmente compreende a sequência de aminoácido da SEQ ID NO: 2 ou uma variante alélica desta; ou um fragmento desta que tem atividade de endoglicanase. Em um aspecto preferido, um polipeptídeo compreende a sequência de aminoácido da SEQ ID NO: 2. Em um outro aspecto preferido, um
25 polipeptídeo compreende o polipeptídeo maduro da SEQ ID NO: 2. Em um outro aspecto preferido, um polipeptídeo compreende os aminoácidos de 18 a 336 da SEQ ID NO: 2 ou uma variante alélica desta; ou um fragmento desta que tem atividade de endoglicanase. Em um outro aspecto preferido, um polipeptídeo compreende os aminoácidos de 18 a 336 da SEQ ID NO: 2. Em

um outro aspecto preferido, um polipeptídeo consiste da seqüência de aminoácido da SEQ ID NO: 2 ou uma variante alélica desta; ou um fragmento desta que tem atividade de endoglicanase. Em um outro aspecto preferido, um polipeptídeo consiste da seqüência de aminoácido da SEQ ID NO: 2. Em um
5 outro aspecto preferido, um polipeptídeo consiste do polipeptídeo maduro da SEQ ID NO: 2. Em um outro aspecto preferido, um polipeptídeo consiste dos aminoácidos 18 a 336 da SEQ ID NO: 2 ou uma variante alélica desta; ou um fragmento desta que tem atividade de endoglicanase. Em um outro aspecto preferido, um polipeptídeo consiste dos aminoácidos de 18 a 336 da SEQ ID
10 NO: 2.

Em um segundo aspecto, a presente invenção diz respeito a polipeptídeos isolados tendo atividade de endoglicanase que são codificados pelos polinucleotídeos que hibridizam sob condições de estringência muito baixa, preferivelmente condições de estringência baixa, mais preferivelmente
15 condições de estringência média, mais preferivelmente condições de estringência média-alta, ainda mais preferivelmente condições de estringência alta e o mais preferivelmente condições de estringência muito altas com (i) a seqüência codificadora do polipeptídeo maduro da SEQ ID NO: 1, (ii) a seqüência de DNA genômico que compreende a seqüência codificadora do
20 polipeptídeo maduro da SEQ ID NO: 1, (iii) uma subsequência de (i) ou (ii) ou (iv) um filamento complementar de (i), (ii) ou (iii) (J. Sambrook, E. F. Fritsch e T. Maniatis, 1989, Molecular Cloning, A Laboratory Manual, 2^a edição, Cold Spring Harbor, Nova Iorque). Uma subsequência da SEQ ID NO: 1 contém pelo menos 100 nucleotídeos contíguos ou preferivelmente
25 pelo menos 200 nucleotídeos contíguos. Além disso, a subsequência pode codificar um fragmento de polipeptídeo que tem atividade de endoglicanase. Em um aspecto preferido, a seqüência codificadora do polipeptídeo maduro tem os nucleotídeos de 52 a 1008 da SEQ ID NO: 1.

A seqüência de nucleotídeo da SEQ ID NO: 1 ou uma

subseqüência desta, assim como a seqüência de aminoácido da SEQ ID NO: 2 ou um fragmento desta, podem ser usados para planejar uma sonda de ácido nucleico para identificar e clonar os polipeptídeos que codificam o DNA tendo atividade de endoglicanase a partir de cepas de gêneros ou espécies diferentes de acordo com métodos bem conhecidos na técnica. Em particular, tais sondas podem ser usadas para a hibridização com o genoma ou cDNA do gênero ou espécie de interesse, seguindo os procedimentos de Southern blotting padrão, de modo a identificar e isolar o gene correspondente nesse lugar. Tais sondas podem ser consideravelmente mais curtas do que a seqüência inteira, mas deve ter pelo menos 14, preferivelmente pelo menos 25, mais preferivelmente pelo menos 35 e o mais preferivelmente pelo menos 70 nucleotídeos no comprimento. É, entretanto, preferido que a sonda de ácido nucleico tenha pelo menos 100 nucleotídeos no comprimento. Por exemplo, a sonda de ácido nucleico podem ter pelo menos 200 nucleotídeos, preferivelmente pelo menos 300 nucleotídeos, mais preferivelmente pelo menos 400 nucleotídeos ou o mais preferivelmente pelo menos 500 nucleotídeos no comprimento. sondas ainda mais longas podem ser usadas, por exemplo, sondas de ácido nucleico que têm pelo menos 600 nucleotídeos, preferivelmente pelo menos 700 nucleotídeos, mais preferivelmente pelo menos 800 nucleotídeos ou o mais preferivelmente pelo menos 900 nucleotídeos no comprimento. Sondas tanto de DNA quanto de RNA podem ser usadas. As sondas são tipicamente rotuladas para detectar o gene correspondente (por exemplo, com ^{32}P , ^3H , ^{35}S , biotina ou avidina). Tais sondas são abrangidas pela presente invenção.

Uma biblioteca de DNA ou cDNA genômica preparada a partir de tais outros organismos, portanto, pode ser triada quanto ao DNA que hibridiza com as sondas descritas acima e que codifica um polipeptídeo tendo atividade de endoglicanase. O genômico ou outro DNA a partir do qual tais outros organismos podem ser separados pela eletroforese em gel de agarose

ou poliacrilamida ou outras técnicas de separação. O DNA das bibliotecas ou o DNA separado podem ser transferidos e imobilizados em nitrocelulose ou outro material carregador adequado. De modo a identificar um clone ou DNA que seja homólogo com a SEQ ID NO: 1 ou uma subsequência desta, o material carregador é usado em um Southern blot.

Para os propósitos da presente invenção, a hibridização indica que a sequência de nucleotídeo hibridiza com uma sonda de ácido nucleico rotulada que corresponde à sequência de nucleotídeo mostrada na SEQ ID NO: 1, a sequência de DNA genômico que compreende a SEQ ID NO: 1, seu filamento complementar ou uma subsequência desta, sob condições de estringência de muito baixas a muito altas. As moléculas às quais a sonda de ácido nucleico hibridiza sob estas condições podem ser detectadas usando, por exemplo, películas de raio X.

Em um aspecto preferido, a sonda de ácido nucleico é a sequência codificadora do polipeptídeo maduro da SEQ ID NO: 1. Em um outro aspecto preferido, a sonda de ácido nucleico tem os nucleotídeos de 52 a 1008 da SEQ ID NO: 1. Em um outro aspecto preferido, a sonda de ácido nucleico é uma sequência de polinucleotídeo que codifica o polipeptídeo da SEQ ID NO: 2 ou uma subsequência desta. Em um outro aspecto preferido, a sonda de ácido nucleico é a SEQ ID NO: 1. Em um outro aspecto preferido, a sonda de ácido nucleico é a sequência codificadora do polipeptídeo maduro da SEQ ID NO: 1. Em um outro aspecto preferido, a sonda de ácido nucleico é a sequência de polinucleotídeo contida no plasmídeo pTter7F que está contida na *E. coli* NRRL B-30837, em que a sequência de polinucleotídeo desta codifica um polipeptídeo tendo atividade de lipase. Em um outro aspecto preferido, a sonda de ácido nucleico é a sequência codificadora do polipeptídeo maduro contida no plasmídeo pTter7F que está contida na *E. coli* NRRL B-30837.

Para sondas longas de pelo menos 100 nucleotídeos no

comprimento, as condições de estringência de muito baixas a muito altas são definidas como pré-hibridização e hibridização a 42°C em 5X SSPE, 0,3% de SDS, 200 µg/ml de DNA de espermatozoides de salmão cisalhado e desnaturado e 25% de formamida para as estringências muito baixas e baixas, 35% de formamida para as estringências médias e média-altas ou 50% de formamida para as estringências altas e muito altas, seguindo os procedimentos de Southern blotting padrão idealmente por 12 a 24 horas.

Para sondas longas de pelo menos 100 nucleotídeos no comprimento, o material carregador é finalmente lavado três vezes cada uma por 15 minutos usando 2X SSC, 0,2% de SDS preferivelmente pelo menos a 45°C (estringência muito baixa), mais preferivelmente pelo menos a 50°C (estringência baixa), mais preferivelmente pelo menos a 55°C (estringência média), mais preferivelmente pelo menos a 60°C (estringência média-alta), ainda mais preferivelmente pelo menos a 65°C (estringência alta) e o mais preferivelmente pelo menos a 70°C (estringência muito alta).

Para sondas curtas que têm cerca de 15 nucleotídeos a cerca de 70 nucleotídeos no comprimento, as condições de estringência são definidas como lavagem de pré-hibridização, hibridização e pós-hibridização a cerca de 5°C a cerca de 10°C abaixo da T_m calculada, usando o cálculo de acordo com Bolton e McCarthy (1962, Proceedings of the National Academy of Sciences USA 48: 1390) em 0,9 M de NaCl, 0,09 M de Tris-HCl pH 7,6, 6 mM de EDTA, 0,5% de NP-40, 1X solução de Denhardt, 1 mM de pirofosfato de sódio, 1 mM de fosfato monobásico de sódio, 0,1 mM de ATP e 0,2 mg de RNA de levedura por ml seguindo os procedimentos de Southern blotting padrão idealmente por 12 a 24 horas.

Para sondas curtas que têm cerca de 15 nucleotídeos a cerca de 70 nucleotídeos no comprimento, o material carregador é lavado uma vez em 6X SSC mais 0,1% de SDS por 15 minutos e duas vezes cada uma por 15 minutos usando 6X SSC de 5°C a 10°C abaixo da T_m calculada.

Em um terceiro aspecto, a presente invenção diz respeito a variantes artificiais que compreendem uma substituição, deleção e/ou inserção conservativa de um ou mais aminoácidos da SEQ ID NO: 2 ou uma sequência homóloga desta; ou o polipeptídeo maduro desta. Preferivelmente, as mudanças de aminoácido são de uma natureza menor, isto é substituições ou inserções de aminoácido conservativas que não afetam significativamente a dobra e/ou atividade da proteína; deleções pequenas, tipicamente de um a cerca de 30 aminoácidos; extensões de terminal amino ou carboxila pequenas, tais como um resíduo de metionina de terminal amino; um peptídeo ligador pequeno de até cerca de 20 a 25 resíduos; ou uma extensão pequena que facilite a purificação pela mudança da carga líquida ou uma outra função, tal como um trato de poli-histidina, um epítipo antigênico ou um domínio de ligação.

Os exemplos de substituições conservativas estão dentro do grupo de aminoácidos básicos (arginina, lisina e histidina), aminoácidos ácidos (ácido glutâmico e ácido aspártico), aminoácidos polares (glutamina e asparagina), aminoácidos hidrofóbicos (leucina, isoleucina e valina), aminoácidos aromáticos (fenilalanina, triptofano e tirosina) e aminoácidos pequenos (glicina, alanina, serina, treonina e metionina). As substituições de aminoácido que no geral não alteram a atividade específica são conhecidas na técnica e são descritas, por exemplo, por H. Neurath e R. L. Hill, 1979, Em, The proteins, Academic Press, Nova Iorque. As mudanças que mais habitualmente ocorre são Ala/Ser, Val/Ile, Asp/Glu, Thr/Ser, Ala/Gly, Ala/Thr, Ser/Asn, Ala/Val, Ser/Gly, Tyr/Phe, Ala/Pro, Lys/Arg, Asp/Asn, Leu/Ile, Leu/Val, Ala/Glu e Asp/Gly.

Além dos 20 aminoácidos padrão, os aminoácidos não padrão (tais como 4-hidroxiprolina, 6-N-metil lisina, ácido 2-aminoisobutírico, isovalina e alfa-metil serina) podem ser substituídos no lugar de resíduos de aminoácido de um polipeptídeo do tipo selvagem. Um número limitado de

aminoácidos não conservativos, aminoácidos que não são codificados pelo código genético e aminoácidos não naturais podem ser substituídos no lugar de resíduos de aminoácido. Os “aminoácidos não naturais” têm sido modificados depois da síntese da proteína e/ou têm uma estrutura química na(s) sua(s) cadeia(s) lateral(is) diferentes daquela dos aminoácidos padrão. Os aminoácidos não naturais podem ser quimicamente sintetizados e preferivelmente, são comercialmente disponíveis e incluem ácido pipecólico, ácido tiazolidino carboxílico, desidroprolina, 3- e 4-metilprolina e 3,3-dimetilprolina.

10 Alternativamente, as mudanças de aminoácido são de uma natureza tal que as propriedades psicoquímica dos polipeptídeos sejam alteradas. Por exemplo, as mudanças de aminoácido podem melhorar a estabilidade térmica do polipeptídeo, alterar a especificidade de substrato, mudar o pH ideal e outros.

15 Os aminoácidos essenciais no polipeptídeo precursor podem ser identificados de acordo com procedimentos conhecidos na técnica, tais como mutagênese direcionada ao sítio ou mutagênese de varredura de alanina (Cunningham e Wells, 1989, Science 244: 1081-1085). Na técnica mencionado em segundo lugar, as mutações de alanina individuais são
20 introduzidas em cada resíduo na molécula e as moléculas mutantes resultantes são testadas quanto a atividade biológica (isto é, atividade de endoglicanase) para identificar resíduos de aminoácido que são críticos para a atividade da molécula. Ver também, Hilton *et al.*, 1996, J. Biol. Chem. 271: 4699-4708. O sítio ativo da enzima ou outra interação biológica também pode ser
25 determinada pela análise química da estrutura, como determinado por técnicas tais como ressonância magnética nuclear, cristalografia, difração de elétron ou rotulação de fotoafinidade, em conjunção com a mutação de aminoácidos de sítio de contato putativo. Ver, por exemplo, de Vos *et al.*, 1992, Science 255: 306-312; Smith *et al.*, 1992, J. Mol. Biol. 224: 899-904; Wlodaver *et al.*,

1992, FEBS Lett. 309: 59-64. As identidades de aminoácidos essenciais também podem ser deduzidas da análise de identidades com polipeptídeos que estejam relacionados a um polipeptídeo de acordo com a invenção.

As substituições de aminoácido individuais ou múltiplas
5 podem ser feitas e testadas usando métodos conhecidos de mutagênese, recombinação e/ou embaralhamento, seguidos por um procedimento de triagem relevante, tal como aqueles divulgados por Reidhaar-Olson e Sauer, 1988, Science 241: 5357; Bowie e Sauer, 1989, Proc. Natl. Acad. Sci. USA 86: 2152-2156; WO 95/17413; ou WO 95/22625. Outros métodos que podem
10 ser usados incluem PCR propensa a erro, demonstração de fago (por exemplo, Lowman *et al.*, 1991, Biochem. 30: 10832-10837; Patente U.S. Nº 5.223.409; WO 92/06204) e mutagênese direcionada à região (Derbyshire *et al.*, 1986, Gene 46: 145; Ner *et al.*, 1988, DNA 7: 127).

Os métodos de mutagênese/embaralhamento podem ser
15 combinados com métodos de triagem de alto rendimento, automatizados para detectar a atividade de polipeptídeos clonados, mutagenizados expressados pelas células hospedeiras (Ness *et al.*, 1999, Nature Biotechnology 17: 893-896). As moléculas de DNA mutagenizadas que codificam polipeptídeos ativos podem ser recuperados das células hospedeiras e rapidamente
20 seqüenciadas usando métodos padrão na técnica. Estes métodos permitem a determinação rápida da importância de resíduos de aminoácido individuais em um polipeptídeo de interesse e podem ser aplicados aos polipeptídeos de estrutura desconhecida.

O número total de substituições, deleções e/ou inserções de
25 aminoácido do polipeptídeo maduro da SEQ ID NO: 2, tal como dos aminoácidos de 18 a 336 da SEQ ID NO: 2, é 10, preferivelmente 9, mais preferivelmente 8, mais preferivelmente 7, mais preferivelmente no máximo 6, mais preferivelmente 5, mais preferivelmente 4, ainda mais preferivelmente 3, o mais preferivelmente 2 e ainda mais preferivelmente 1.

Fontes de Polipeptídeos Tendo Atividade de Endoglicanase

Um polipeptídeo da presente invenção pode ser obtido a partir de microorganismos de qualquer gênero. Para os propósitos da presente invenção, o termo “obtido a partir de” como aqui usado em conexão com uma fonte dada deve significar que o polipeptídeo codificado por uma seqüência de nucleotídeo é produzido pela fonte ou por uma cepa em que a seqüência de nucleotídeo da fonte foi inserida. Em um aspecto preferido, o polipeptídeo obtido a partir de uma dada fonte é secretado extracelularmente.

Um polipeptídeo da presente invenção pode ser um polipeptídeo bacteriano. Por exemplo, o polipeptídeo pode ser um polipeptídeo bacteriano gram positivo tal como um polipeptídeo de *Bacillus*, por exemplo, um polipeptídeo de *Bacillus alkalophilus*, *Bacillus amiloliquefaciens*, *Bacillus brevis*, *Bacillus circulans*, *Bacillus coagulans*, *Bacillus lautus*, *Bacillus lentus*, *Bacillus licheniformis*, *Bacillus megaterium*, *Bacillus stearothermophilus*, *Bacillus subtilis* ou *Bacillus thuringiensis* tendo atividade de endoglicanase; ou um polipeptídeo de *Streptomyces* tendo atividade de endoglicanase, por exemplo, um polipeptídeo de *Streptomyces lividans* ou *Streptomyces murinus* tendo atividade de endoglicanase; ou um polipeptídeo bacteriano gram negativo, por exemplo, um polipeptídeo de *E. coli* ou um *Pseudomonas sp.* tendo atividade de endoglicanase.

Um polipeptídeo da presente invenção também pode ser um polipeptídeo fúngico e mais preferivelmente um polipeptídeo de levedura tal como um polipeptídeo de *Candida*, *Kluyveromyces*, *Pichia*, *Saccharomyces*, *Schizosaccharomyces* ou *Yarrowia* tendo atividade de endoglicanase; ou mais preferivelmente um polipeptídeo fúngico filamentoso tal como um polipeptídeo de *Acremonium*, *Aspergillus*, *Aureobasidium*, *Cryptococcus*, *Filibasidium*, *Fusarium*, *Humicola*, *Magnaporthe*, *Mucor*, *Myceliophthora*, *Neocafflmastix*, *Neurospora*, *Paecilomyces*, *Penicillium*, *Piromyces*, *Schizophilum*, *Talaromyces*, *Thermoascus*, *Thielavia*, *Tolypocladium* ou

Trichoderma tendo atividade de endoglicanase.

Em um aspecto preferido, o polipeptídeo é um polipeptídeo de *Saccharomyces carlsbergensis*, *Saccharomyces cerevisiae*, *Saccharomyces diastaticus*, *Saccharomyces douglasfi*, *Saccharomyces kluyveri*,
 5 *Saccharomyces norbensis* ou *Saccharomyces oviformis* tendo atividade de endoglicanase.

Em um outro aspecto preferido, o polipeptídeo é um polipeptídeo de *Aspergillus aculeatus*, *Aspergillus awamori*, *Aspergillus fumigatus*, *Aspergillus foetidus*, *Aspergillus japonicus*, *Aspergillus nidulans*,
 10 *Aspergillus niger*, *Aspergillus oryzae*, *Fusarium bactridioides*, *Fusarium cerealis*, *Fusarium crookwellense*, *Fusarium culmorum*, *Fusarium graminearum*, *Fusarium graminum*, *Fusarium heterosporum*, *Fusarium negundi*, *Fusarium oxysporum*, *Fusarium reticulatum*, *Fusarium roseum*, *Fusarium sambucinum*, *Fusarium sarcochroum*, *Fusarium sporotrichioides*,
 15 *Fusarium sulphureum*, *Fusarium torulosum*, *Fusarium trichothecioides*, *Fusarium venenatum*, *Humicola insolens*, *Humicola lanuginosa*, *Mucor miehei*, *Myceliophthora thermophila*, *Neurospora crassa*, *Penicillium purpurogenum*, *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei* ou *Trichoderma viride* tendo atividade
 20 de endoglicanase.

Em um outro aspecto preferido, o polipeptídeo é um polipeptídeo de *Thielavia achromatica*, *Thielavia albomyces*, *Thielavia albopilosa*, *Thielavia australeinsis*, *Thielavia fimeti*, *Thielavia microspora*, *Thielavia ovispora*, *Thielavia peruviana*, *Thielavia spededonium*, *Thielavia*
 25 *setosa*, *Thielavia subthermophila*, *Thielavia terrestris*, *Thielavia terricola*, *Thielavia thermophila*, *Thielavia variospora* ou *Thielavia wareingfi* tendo atividade de endoglicanase.

Em um aspecto mais preferido, o polipeptídeo é um polipeptídeo de *Thielavia terrestris* tendo atividade de endoglicanase e o mais

preferivelmente um *Thielavia terrestris* NRRL 8126 tendo atividade de endoglicanase, por exemplo, o polipeptídeo da SEQ ID NO: 2 ou o polipeptídeo maduro desta.

5 Será entendido que para as espécies anteriormente mencionadas a invenção abrange os estados tanto perfeito quanto imperfeito e outros equivalentes taxonômicos, por exemplo, anamorfos, independente do nome da espécie pelo qual eles são conhecidos. Aqueles habilitados na técnica facilmente reconhecerão a identidade de equivalentes apropriados.

10 A cepas destas espécies são facilmente acessíveis ao público em várias coleções de cultura, tais como a American Type Culture Collection (ATCC), Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSM), Centraalbureau Voor Schimmelcultures (CBS) e Agricultural Research Service Patent Culture Collection, Northern Regional Research Center (NRRL).

15 Além disso, tais polipeptídeos podem ser identificados e obtidos a partir de outras fontes incluindo microorganismos isolados da natureza (por exemplo, solo, compostos, água, etc.) usando as sondas mencionadas acima. As técnicas para isolar microorganismos de habitats naturais são bem conhecidos na técnica. O polinucleotídeo pode ser depois
20 obtido pela triagem de similaridade de uma biblioteca genômica ou de cDNA de um tal microorganismo. Uma vez que uma seqüência de polinucleotídeo que codifica um polipeptídeo foi detectada com a(s) sonda(s), o polinucleotídeo pode ser isolado ou clonado utilizando-se técnicas que são bem conhecidas por aqueles de habilidade comum na técnica (ver, por
25 exemplo, Sambrook *et al.*, 1989, *supra*).

Os polipeptídeos da presente invenção também incluem polipeptídeos fundidos ou polipeptídeos de fusão cliváveis em que um outro polipeptídeo é fundido no terminal N ou no terminal C do polipeptídeo ou fragmento desta. Um polipeptídeo fundido é produzido fundindo-se uma

seqüência de nucleotídeo (ou uma porção desta) que codifica um outro polipeptídeo a uma seqüência de nucleotídeo (ou uma porção desta) da presente invenção. As técnicas para produzir polipeptídeos de fusão são conhecidas no ramo e incluem ligar as seqüências codificadoras que codificam os polipeptídeos de modo que elas estejam na matriz e que a expressão do polipeptídeo fundido esteja sob o controle do(s) mesmo(s) promotor(es) e terminador(es).

Polinucleotídeos

A presente invenção também diz respeito aos polinucleotídeos isolados que compreendem ou que consistem de uma seqüência de nucleotídeo que codifica um polipeptídeo da presente invenção tendo atividade de endoglicanase.

Em um aspecto preferido, a seqüência de nucleotídeo compreende ou consiste da SEQ ID NO: 1. Em um outro aspecto mais preferido, a seqüência de nucleotídeo compreende ou consiste da seqüência contida no plasmídeo pTter7F que está contido na *E. coli* NRRL B-30837. Em um outro aspecto preferido, a seqüência de nucleotídeo compreende ou consiste da região codificadora do polipeptídeo maduro da SEQ ID NO: 1. Em um outro aspecto preferido, a seqüência de nucleotídeo compreende ou consiste dos nucleotídeos de 52 a 1008 da SEQ ID NO: 1. Em um outro aspecto mais preferido, a seqüência de nucleotídeo compreende ou consiste da região codificadora do polipeptídeo maduro contida no plasmídeo pTter7F que está contido na *E. coli* NRRL B-30837. A presente invenção também abrange seqüências de nucleotídeo que codificam um polipeptídeo que compreende ou que consiste da seqüência de aminoácido da SEQ ID NO: 2 ou o polipeptídeo maduro desta, que diferem da SEQ ID NO: 1 ou a seqüência codificadora do polipeptídeo maduro desta em virtude da degenerescência do código genético. A presente invenção também diz respeito às subseqüências da SEQ ID NO: 1 que codificam fragmentos da

SEQ ID NO: 2 que têm atividade de endoglicanase.

A presente invenção também diz respeito a mutantes de polinucleotídeo que compreendem pelo menos uma mutação na seqüência codificadora do polipeptídeo maduro da SEQ ID NO: 1, em que a seqüência de nucleotídeo mutante codifica o polipeptídeo maduro da SEQ ID NO: 2. Em um aspecto preferido, o polipeptídeo maduro tem os aminoácidos de 18 a 336 da SEQ ID NO: 2.

As técnicas usadas para isolar ou clonar um polinucleotídeo que codifica um polipeptídeo são conhecidas no ramo e incluem a isolamento de DNA genômico, preparação de cDNA ou uma combinação destas. A clonagem dos polinucleotídeos da presente invenção de tal DNA genômico pode ser efetuada, por exemplo, usando-se a reação da cadeia da polimerase (PCR) bem conhecida ou triagem de anticorpo de bibliotecas de expressão para detectar fragmentos de DNA clonados com características estruturais compartilhadas. Ver, por exemplo, Innis *et al.*, 1990, PCR: A Guide to Methods and Application, Academic Press, Nova Iorque. Outros procedimentos de amplificação de ácido nucleico tais como a reação de cadeia da ligase (LCR), transcrição ativada ligada (LAT) e amplificação com base em seqüência de nucleotídeo (NASBA) podem ser usados. Os polinucleotídeos podem ser clonados a partir de uma cepa de *Thielavia* ou um outro organismo ou organismo relacionado e assim, por exemplo, podem ser uma espécie alélica ou variante do polipeptídeo que codifica a região da seqüência de nucleotídeo.

A presente invenção também diz respeito a polinucleotídeos que compreendem seqüências de nucleotídeo que têm um grau de identidade com a seqüência codificadora do polipeptídeo maduro da SEQ ID NO: 1 de pelo menos 60%, preferivelmente pelo menos 65%, mais preferivelmente pelo menos 70%, mais preferivelmente pelo menos 75%, mais preferivelmente pelo menos 80%, mais preferivelmente pelo menos 85%, mais

preferivelmente pelo menos 90%, ainda mais preferivelmente pelo menos 95% e o mais preferivelmente pelo menos 97% de identidade, que codifica um polipeptídeo ativo. Em um aspecto preferido, a seqüência codificadora do polipeptídeo maduro tem os nucleotídeos de 52 a 1008 da SEQ ID NO: 1.

5 A modificação de uma seqüência de nucleotídeo que codifica um polipeptídeo da presente invenção pode ser necessária para a síntese de polipeptídeos substancialmente similares ao polipeptídeo. O termo “substancialmente similar” ao polipeptídeo refere-se às formas que não ocorrem naturalmente do polipeptídeo. Estes polipeptídeos podem diferir em
10 algum modo engendrado do polipeptídeo isolado da sua fonte nativa, por exemplo, variantes artificiais que diferem na atividade específica, termoestabilidade, pH ideal ou semelhante. A seqüência variante pode ser construída com base na seqüência de nucleotídeo apresentada como a região que codifica o polipeptídeo da SEQ ID NO: 1, por exemplo, uma
15 subseqüência desta e/ou pela introdução de substituições de nucleotídeo que não dão origem a uma outra seqüência de aminoácido do polipeptídeo codificado pela seqüência de nucleotídeo, mas que corresponde ao uso de códon do organismo hospedeiro pretendido para a produção da enzima ou pela introdução de substituições de nucleotídeo que podem dar origem a uma
20 seqüência de aminoácido diferente. Para uma descrição geral de substituição de nucleotídeo, ver, por exemplo, Ford *et al.*, 1991, Protein Expression and Purification 2: 95-107.

 Estará evidente àqueles habilitados na técnica que tais substituições podem ser feitas fora das regiões críticas para a função da
25 molécula e ainda resultar em um polipeptídeo ativo. Os resíduos de aminoácido essenciais para a atividade do polipeptídeo codificado por um polinucleotídeo isolado da invenção e portanto preferivelmente não submetido à substituição, podem ser identificados de acordo com procedimentos conhecidos na técnica, tais como mutagênese direcionada ao sítio ou

mutagênese de varredura de alanina (ver, por exemplo, Cunningham e Wells, 1989, Science 244: 1081-1085). Na técnica mencionada em segundo lugar, mutações são introduzidas em cada resíduo positivamente carregado na molécula e as moléculas mutantes resultantes são testadas quanto a atividade de endoglicanase para identificar resíduos de aminoácido que são críticos para a atividade da molécula. Os sítios de interação de enzima substrato também podem ser determinados pela análise da estrutura tridimensional como determinada por tais técnicas como análise de ressonância magnética nuclear, cristalografia ou rotulação por fotoafinidade (ver, por exemplo, de Vos *et al.*, 1992, Science 255: 306-312; Smith *et al.*, 1992, Journal of Molecular Biology 224: 899-904; Wlodaver *et al.*, 1992, FEBS Letters 309: 59-64).

A presente invenção também diz respeito a polinucleotídeos isolados que codificam um polipeptídeo da presente invenção, que hibridizam sob condições de estringência muito baixa, preferivelmente condições de estringência baixa, mais preferivelmente condições de estringência média, mais preferivelmente condições de estringência média-alta, ainda mais preferivelmente condições de estringência alta e o mais preferivelmente condições de estringência muito altas com (i) a seqüência codificadora do polipeptídeo maduro da SEQ ID NO: 1, (ii) a seqüência de DNA genômico que compreende a seqüência codificadora do polipeptídeo maduro da SEQ ID NO: 1 ou (iii) um filamento complementar de (i) ou (ii); ou variante alélicas e subseqüências deste (Sambrook *et al.*, 1989, *supra*), como aqui definido. Em um aspecto preferido, a seqüência codificadora do polipeptídeo maduro tem os nucleotídeos de 52 a 1008 da SEQ ID NO: 1.

A presente invenção também diz respeito a polinucleotídeos isolados obtidos (a) pela hibridização de uma população de DNA sob condições de estringência muito baixas, baixas, médias, média-altas, altas ou muito altas com (i) a seqüência codificadora do polipeptídeo maduro da SEQ ID NO: 1, (ii) a seqüência de DNA genômico que compreende a seqüência

codificadora do polipeptídeo maduro da SEQ ID NO: 1 ou (iii) um filamento complementar de (i) ou (ii); e (b) isolar o polinucleotídeo de hibridização, que codifica um polipeptídeo tendo atividade de endoglicanase. Em um aspecto preferido, a seqüência codificadora do polipeptídeo maduro tem os nucleotídeos de 52 a 1008 da SEQ ID NO: 1.

Construções de Ácido Nucleico

A presente invenção também diz respeito às construções de ácido nucleico que compreendem um polinucleotídeo isolado da presente invenção operavelmente ligado a uma ou mais seqüências de controle que direcionam a expressão da seqüência codificadora em uma célula hospedeira adequada sob condições compatíveis com as seqüências de controle.

Um polinucleotídeo isolado que codifica um polipeptídeo da presente invenção pode ser manipulado em uma variedade de modos para fornecer a expressão do polipeptídeo. A manipulação da seqüência de polinucleotídeo antes da sua inserção em um vetor pode ser desejável ou necessária dependendo do vetor de expressão. As técnicas para modificar as seqüências de polinucleotídeo que utilizam os métodos de DNA recombinante são bem conhecidas no ramo.

A seqüência de controle pode ter uma seqüência promotora apropriada, uma seqüência de nucleotídeo que é reconhecida por uma célula hospedeira para a expressão de um polinucleotídeo que codifica um polipeptídeo da presente invenção. A seqüência promotora contém seqüências de controle transcricional que medeiam a expressão do polipeptídeo. O promotor pode ser qualquer seqüência de nucleotídeo que mostre atividade transcricional na célula hospedeira de escolha incluindo promotores mutantes, truncados e híbridos e podem ser obtidos a partir dos genes que codificam os polipeptídeos extracelulares ou intracelulares homólogos ou heterólogos para a célula hospedeira.

Os exemplos de promotores adequados para direcionar a

transcrição das construções de ácido nucleico da presente invenção, especialmente em uma célula hospedeira bacteriana, são os promotores obtidos do operon lac de *E. coli*, gene da agarase (dagA) de *Streptomyces coelicolor*, gene da levansucrase (sacB) de *Bacillus subtilis*, gene da alfa-amilase (amil) de *Bacillus licheniformis*, gene da amilase maltogênica (amyM) de *Bacillus stearothermophilus*, gene da alfa-amilase (amyQ) de *Bacillus amiloliquefaciens*, gene da penicilinase (penP) de *Bacillus licheniformis*, genes xilA e xilB de *Bacillus subtilis* e gene da beta-lactamase procariótica (VillaKamaroff *et al.*, 1978, Proceedings of the National Academy of Sciences USA 75: 3727-3731), assim como o promotor *tac* (DeBoer *et al.*, 1983, Proceedings of the National Academy of Sciences USA 80: 21-25). Outros promotores são descritos em "Useful proteins from recombinant bacteria" em Scientific American, 1980, 242: 74-94; e em Sambrook *et al.*, 1989, *supra*.

Os exemplos de promotores adequados para direcionar a transcrição das construções de ácido nucleico da presente invenção em uma célula hospedeira fúngica filamentosa são promotores obtidos a partir dos genes para a TAKA amilase de *Aspergillus oryzae*, proteinase aspártica de *Rhizomucor miehei*, alfa-amilase neutra de *Aspergillus niger*, alfa-amilase estável a ácido de *Aspergillus niger*, glicoamilase (glaA) de *Aspergillus niger* ou *Aspergillus awamori*, lipase de *Rhizomucor miehei*, protease alcalina de *Aspergillus oryzae*, triose fosfato isomerase de *Aspergillus oryzae*, acetamidase de *Aspergillus nidulans*, amiloglicosidase de *Fusarium venenatum* (WO 00/56900), *Fusarium venenatum* Dada (WO 00/56900), *Fusarium venenatum* Quinn (WO 00/56900), protease equivalente a tripsina de *Fusarium oxysporum* (WO 96/00787), beta-glicosidase de *Trichoderma reesei*, celobioidrolase I de *Trichoderma reesei*, celobioidrolase II de *Trichoderma reesei*, endoglicanase I de *Trichoderma reesei*, endoglicanase II de *Trichoderma reesei*, endoglicanase III de *Trichoderma reesei*,

endoglicanase IV de *Trichoderma reesei*, endoglicanase V de *Trichoderma reesei*, xilanase I de *Trichoderma reesei*, xilanase II de *Trichoderma reesei*, beta-xilosidase de *Trichoderma reesei*, assim como o promotor NA2-tpi (um híbrido dos promotores dos genes para a alfa-amilase neutra de *Aspergillus niger* e triose fosfato isomerase de *Aspergillus oryzae*); e promotores destas mutantes, truncados e híbridos.

Em um hospedeiro de levedura, os promotores úteis são obtidos a partir dos genes para a enolase (ENO-1) de *Saccharomyces cerevisiae*, galactoquinase (GAL1) de *Saccharomyces cerevisiae*, álcool desidrogenase/gliceraldeído-3-fosfato desidrogenase (ADH1, ADH2/ GAP) de *Saccharomyces cerevisiae*, triose fosfato isomerase (TPI) de *Saccharomyces cerevisiae*, metalotionina (CUP1) de *Saccharomyces cerevisiae* e 3-fosfoglicerato quinase de *Saccharomyces cerevisiae*. Outros promotores úteis para as células hospedeiras de levedura são descritas por Romanos *et al.*, 1992, *Yeast* 8: 423-488.

A seqüência de controle também pode ser uma seqüência terminadora de transcrição adequada, uma seqüência reconhecida por uma célula hospedeira para terminar a transcrição. A seqüência terminadora é operavelmente ligada ao terminal 3' da seqüência de nucleotídeo que codifica o polipeptídeo. Qualquer terminador que seja funcional na célula hospedeira de escolha pode ser usada na presente invenção.

Os terminadores preferidos para as células hospedeiras filamentosas fúngicas são obtidos a partir dos genes para a TAKA amilase de *Aspergillus oryzae*, glicoamilase de *Aspergillus niger*, antranilato sintase de *Aspergillus nidulans*, alfa-glicosidase de *Aspergillus niger* e protease equivalente a tripsina de *Fusarium oxysporum*.

Os terminadores preferidos para as células hospedeiras de levedura são obtidos a partir dos genes para a enolase de *Saccharomyces cerevisiae*, citocromo C (CYC1) de *Saccharomyces cerevisiae* e

gliceraldeído-3-fosfato desidrogenase de *Saccharomyces cerevisiae*. Outros terminadores úteis para as células hospedeiras de levedura são descritas por Romanos *et al.*, 1992, *supra*.

A seqüência de controle também pode ser uma seqüência líder adequada, uma região não traduzida de um mRNA que é importante para a tradução pela célula hospedeira. A seqüência líder é operavelmente ligada ao terminal 5' da seqüência de nucleotídeo que codifica o polipeptídeo. Qualquer seqüência líder que seja funcional na célula hospedeira de escolha pode ser usada na presente invenção.

Os líderes preferidos para as células hospedeiras fúngica filamentosa são obtidos a partir dos genes para a TAKA amilase de triose fosfato isomerase de *Aspergillus oryzae* e *Aspergillus nidulans*.

Os líderes adequados para as células hospedeiras de levedura são obtidos a partir dos genes para a enolase (ENO-1) de *Saccharomyces cerevisiae*, 3-fosfoglicerato quinase de *Saccharomyces cerevisiae*, fator alfa de *Saccharomyces cerevisiae* e álcool desidrogenase/gliceraldeído-3-fosfato desidrogenase (ADH2/GAP) de *Saccharomyces cerevisiae*.

A seqüência de controle também pode ser uma seqüência de poliadenilação, uma seqüência operavelmente ligada ao terminal 3' da seqüência de nucleotídeo e que, quando transcrita, é reconhecida pela célula hospedeira como um sinal para adicionar resíduos de poliadenosina ao mRNA transcrito. Qualquer seqüência de poliadenilação que seja funcional na célula hospedeira de escolha pode ser usada na presente invenção.

As seqüências de poliadenilação preferidas para as células hospedeiras fúngica filamentosa são obtidas a partir dos genes para a TAKA amilase de *Aspergillus oryzae*, glicoamilase de *Aspergillus niger*, antranilato sintase de *Aspergillus nidulans*, protease equivalente a tripsina de *Fusarium oxysporum* e alfa-glicosidase de *Aspergillus niger*.

As seqüências de poliadenilação úteis para as células

hospedeiras de levedura são descritas por Guo e Shermano, 1995, *Molecular Cellular Biology* 15: 5983-5990.

A seqüência de controle também pode ser uma região codificadora de peptídeo de sinal que codifica uma seqüência de aminoácido ligada ao terminal amino de um polipeptídeo e direciona o polipeptídeo codificado no caminho secretor da célula. A extremidade 5' da seqüência codificadora da seqüência de nucleotídeo pode inerentemente conter uma região codificadora de peptídeo de sinal naturalmente ligada na matriz de leitura de tradução com o segmento da região codificadora que codifica o polipeptídeo secretado. Alternativamente, a extremidade 5' da seqüência codificadora pode conter uma região codificadora de peptídeo de sinal que seja estranha para a seqüência codificadora. A região codificadora de peptídeo de sinal estranha pode ser requerida onde a seqüência codificadora não contém naturalmente uma região codificadora de peptídeo de sinal. Alternativamente, a região codificadora de peptídeo de sinal estranha pode simplesmente substituir a região codificadora de peptídeo de sinal natural de modo a realçar a secreção do polipeptídeo. Entretanto, qualquer região codificadora de peptídeo de sinal que direcione o polipeptídeo expressado no caminho secretor de uma célula hospedeira de escolha, isto é, secretado em um meio de cultura, pode ser usado na presente invenção.

As regiões codificadoras de peptídeo de sinal eficazes para as células hospedeiras bacteriana são as regiões codificadoras de peptídeo de sinal obtidas a partir dos genes para a amilase maltogênica de *Bacillus* NCIB 11837, alfa-amilase de *Bacillus stearothermophilus*, subtilisina de *Bacillus licheniformis*, beta-lactamase de *Bacillus licheniformis*, proteases neutras de *Bacillus stearothermophilus* (nprT, nprS, nprM) e *Bacillus subtilis* prsA. Outros peptídeos de sinal são descritos por Simonen e Palva, 1993, *Microbiological Reviews* 57: 109-137.

As regiões codificadoras de peptídeo de sinal eficazes para as

células hospedeiras fúngicas filamentosas são as regiões codificadoras de peptídeo de sinal obtidas a partir dos genes para TAKA amilase de *Aspergillus oryzae*, amilase neutra de *Aspergillus niger*, glicoamilase de *Aspergillus niger*, proteinase de aspártica de *Rhizomucor miehe*, celulase de *Humicola insolens*, endoglicanase V de *Humicola insolens* e lipase de *Humicola lanuginosa*.

Em um aspecto preferido, o peptídeo de sinal tem os aminoácidos de 1 a 17 da SEQ ID NO: 2. Em um outro aspecto preferido, a região codificadora de peptídeo de sinal tem os nucleotídeos de 1 a 51 da SEQ ID NO: 1 que codifica os aminoácidos de 1 a 17 da SEQ ID NO: 2.

Os peptídeos de sinal úteis para as células hospedeiras de levedura são obtidos a partir dos genes para o fator alfa de *Saccharomyces cerevisiae* e invertase de *Saccharomyces cerevisiae*. Outras regiões codificadoras de peptídeo de sinal úteis são descritas por Romanos *et al.*, 1992, *supra*.

A seqüência de controle também pode ser uma região codificadora de propeptídeo que codifica uma seqüência de aminoácido posicionada no terminal amino de um polipeptídeo. O polipeptídeo resultante é conhecido como uma proenzima ou propolipeptídeo (ou um zimogênio em alguns casos). Um propolipeptídeo é no geral inativo e pode ser convertido a um polipeptídeo ativo maduro pela clivagem catalítica ou autocatalítica do propeptídeo a partir do propolipeptídeo. A região codificadora de propeptídeo pode ser obtida a partir dos genes para a protease alcalina de *Bacillus subtilis* (aprE), protease neutra de *Bacillus subtilis* (nprT), fator alfa de *Saccharomyces cerevisiae*, proteinase aspártica de *Rhizomucor miehei* e lacase de *Myceliophthora thermophila* (WO 95/33836).

Onde as regiões tanto de peptídeo de sinal quanto de propeptídeo estão presentes no terminal amino de um polipeptídeo, a região de propeptídeo é posicionada contígua ao terminal amino de um polipeptídeo

e a região de peptídeo de sinal é posicionada contígua ao terminal amino da região de propeptídeo.

5 Também pode ser desejável adicionar seqüências regulatórias que possibilitem a regulação da expressão do polipeptídeo em relação ao crescimento da célula hospedeira. Os exemplos de sistemas regulatórios são aqueles que causam a expressão do gene a ser ligado ou desligado em resposta a um estímulo químico ou físico, incluindo a presença de um composto regulatório. Os sistemas regulatórios em sistemas procarióticos incluem os sistemas operadores lac, tac e trp. Em levedura, o sistema ADH2
10 ou sistema GAL1 pode ser usado. Em fungos filamentosos, o promotor de TAKA alfa-amilase, o promotor da glicoamilase de *Aspergillus niger* e o promotor da glicoamilase de *Aspergillus oryzae* podem ser usados como seqüências reguladoras. Outros exemplos de seqüências reguladoras são aquelas que possibilitam a amplificação do gene. Em sistemas eucarióticos,
15 estes incluem o gene da diidrofoliato redutase que é amplificado na presença de metotrexato e os genes da metalotioneína que são amplificados com metais pesados. Nestes casos, a seqüência de nucleotídeo que codifica o polipeptídeo seria operavelmente ligada com a seqüência reguladora.

Vetores de Expressão

20 A presente invenção também diz respeito a vetores de expressão recombinantes que compreendem um polinucleotídeo da presente invenção, um promotor e sinais de parada transcricionais e traducionais. Os vários ácidos nucleicos e seqüências de controle descritos aqui podem ser unidos entre si para produzir um vetor de expressão recombinante que pode
25 incluir um ou mais sítios de restrição convenientes para permitir a inserção ou substituição da seqüência de nucleotídeo que codifica o polipeptídeo em tais sítios. Alternativamente, uma seqüência de nucleotídeo da presente invenção pode ser expressada pela inserção da seqüência de nucleotídeo ou uma construção de ácido nucleico que compreenda a seqüência em um vetor

apropriado para a expressão. Na criação do vetor de expressão, a sequência codificadora está localizada no vetor de modo que a sequência codificadora esteja operavelmente ligada com as sequências de controle apropriadas para a expressão.

5 O vetor de expressão recombinante pode ser qualquer vetor (por exemplo, um plasmídeo ou vírus) que pode ser convenientemente submetido aos procedimentos de DNA recombinantes e podem realizar a expressão da sequência de nucleotídeo. A escolha do vetor tipicamente dependerá da compatibilidade do vetor com a célula hospedeira na qual o
10 vetor deva ser introduzido. Os vetores podem ser plasmídeos lineares ou circulares fechados.

O vetor pode ser um vetor que replica autonomamente, isto é, um vetor que exista como uma entidade extracromossômica, a replicação do qual é independente da replicação cromossômica, por exemplo, um
15 plasmídeo, um elemento extracromossômico, um minicromossoma ou um cromossoma artificial. O vetor pode conter qualquer meio para garantir a auto-replicação. Alternativamente, o vetor pode ser um que, quando introduzido na célula hospedeira, é integrado no genoma e replicado junto com o(s) cromossoma(s) no qual ele foi integrado. Além disso, um vetor ou
20 plasmídeo individuais ou dois ou mais vetores ou plasmídeos que juntos contêm o DNA total a ser introduzido no genoma da célula hospedeira ou um transposon pode ser usado.

Os vetores da presente invenção preferivelmente contêm um ou mais marcadores selecionáveis que permitem a seleção fácil de células
25 transformadas, transfectadas, transduzidas ou células semelhantes. Um marcador selecionável é um gene, o produto do qual fornece resistência a biocida ou viral, resistência aos metais pesados, prototrofia aos auxótrofos e outros.

Os exemplos de marcadores selecionáveis bacterianos são os

genes *dal* de *Bacillus subtilis* ou *Bacillus licheniformis* ou marcadores que conferem resistência a antibiótico tal como resistência a ampicilina, canamicina, cloranfenicol ou tetraciclina. Os marcadores adequados para as células hospedeiras de levedura são ADE2, HIS3, LEU2, LYS2, MET3, TRP1 e URA3. Os marcadores selecionáveis para o uso em uma célula hospedeira fúngica filamentosa incluem, mas não são limitados a, *amdS* (acetamidase), *argB* (ornitina carbamoiltransferase), *bar* (fosfinotricina acetiltransferase), *hph* (higromicina fosfotransferase), *niaD* (nitrato redutase), *pyrG* (orotidina-5'-fosfato descarboxilase), *sC* (sulfate adeniltransferase) e *trpC* (antranilato sintase), assim como equivalentes destes. Preferidos para o uso em uma célula de *Aspergillus* são os genes *amdS* e *pyrG* de *Aspergillus nidulans* ou *Aspergillus oryzae* e o gene *bar* de *Streptomyces hygroscopicus*.

Os vetores da presente invenção preferivelmente contêm um elemento ou elementos que permitem a integração do vetor no genoma da célula hospedeira ou replicação autônoma do vetor na célula independente do genoma.

Para a integração no genoma da célula hospedeira, o vetor pode contar com a sequência do polinucleotídeo que codifica o polipeptídeo ou qualquer outro elemento do vetor para a integração no genoma pela recombinação homóloga ou não homólogo. Alternativamente, o vetor pode conter sequências de nucleotídeo adicionais para direcionar a integração pela recombinação homóloga no genoma da célula hospedeira em uma localização ou localizações exatas no(s) cromossoma(s). Para aumentar a probabilidade de integração em uma localização exata, os elementos integracionais devem preferivelmente conter um número suficiente de ácidos nucleicos, tal como 100 a 10.000 pares de base, preferivelmente 400 a 10.000 pares de base e o mais preferivelmente 800 a 10.000 pares de base, que têm um alto grau de identidade com a sequência alvo correspondente para realçar a probabilidade de recombinação homóloga. Os elementos integracionais podem ser qualquer

seqüência que seja homóloga com a seqüência alvo no genoma da célula hospedeira. Além disso, os elementos integracionais podem ser seqüências de nucleotídeo não codificadoras ou codificadoras. Por outro lado, o vetor pode ser integrado no genoma da célula hospedeira pela recombinação não homóloga.

Para a replicação autônoma, o vetor pode compreender ainda uma origem de replicação que possibilite que o vetor replique autonomamente na célula hospedeira em questão. A origem de replicação pode ser qualquer replicador plasmídico que medie a replicação autônoma que funcione em uma célula. O termo “origem de replicação” ou “replicador plasmídico” é aqui definido como uma seqüência de nucleotídeo que possibilite que um plasmídeo ou vetor replique *in vivo*.

Os exemplos de origens bacterianas de replicação são as origens de replicação de plasmídeos pBR322, pUC19, pACYC177 e pACYC184 que permitem a replicação na *E. coli* e pUB110, pE194, pTA1060 e pAIV1111 permitindo a replicação em *Bacillus*.

Os exemplos de origens de replicação para o uso em uma célula hospedeira de levedura são a origem de 2 micron de replicação, ARS1, ARS4, a combinação de ARS1 e CEN3 e a combinação de ARS4 e CEN6.

Os exemplos de origens de replicação úteis em uma célula fúngica filamentosa são AMA1 e ANSI (Gems *et al.*, 1991, Gene 98: 61-67; Cullen *et al.*, 1987, Nucleic Acids Research 15: 9163-9175; WO 00/24883). A isolamento do gene AMA1 e a construção de plasmídeos ou vetores que compreendem o gene podem ser realizadas de acordo com os métodos divulgados na WO 00/24883.

Mais do que uma cópia de um polinucleotídeo da presente invenção pode ser inserida na célula hospedeira para aumentar a produção do produto de gene. Um aumento no número de cópia do polinucleotídeo pode ser obtido integrando-se pelo menos uma cópia adicional da seqüência no

genoma da célula hospedeira ou pela inclusão de um gene marcador selecionável amplificável com o polinucleotídeo onde as células contendo cópias amplificadas do gene marcador selecionável e por meio deste cópias adicionais do polinucleotídeo, podem ser selecionadas cultivando-se as células na presença do agente selecionável apropriado.

Os procedimentos usados para ligar os elementos descritos acima para construir os vetores de expressão recombinantes da presente invenção são bem conhecidos por uma pessoa habilitada na técnica (ver, por exemplo, Sambrook *et al.*, 1989, *supra*).

10 **Células hospedeiras**

A presente invenção também diz respeito a células hospedeiras recombinantes, que compreendem um polinucleotídeo da presente invenção, que são vantajosamente usadas na produção recombinante dos polipeptídeos. Um vetor que compreende um polinucleotídeo da presente invenção é introduzido em uma célula hospedeira de modo que o vetor seja mantido como um integrante cromossômico ou como um vetor extra-cromossômico auto-replicante como descrito mais no princípio. O termo “célula hospedeira” abrange qualquer progênie de uma célula precursora que não é idêntica à célula precursora devido às mutações que ocorrem durante a replicação. A escolha de uma célula hospedeira em um grau maior dependerá do gene que codifica o polipeptídeo e da sua fonte.

A célula hospedeira pode ser um microorganismo unicelular, por exemplo, um procariota ou um microorganismo não unicelular, por exemplo, um eucariota.

Os microorganismos unicelulares úteis são células bacterianas tais como bactérias gram positivas incluindo, mas não limitadas a, uma célula de *Bacillus*, por exemplo, *Bacillus alkalophilus*, *Bacillus amiloliquefaciens*, *Bacillus brevis*, *Bacillus circulans*, *Bacillus clausii*, *Bacillus coagulans*, *Bacillus lautus*, *Bacillus lentus*, *Bacillus licheniformis*, *Bacillus megaterium*,

Bacillus stearothermophilus, *Bacillus subtilis* e *Bacillus thuringiensis*; ou um *Streptomyces* célula, por exemplo, *Streptomyces lividans* e *Streptomyces murinus* ou bactérias gram negativas tais como *E. coli* e *Pseudomonas sp.* Em um aspecto preferido, a célula hospedeira bacteriana é uma célula de *Bacillus* 5 *lentus*, *Bacillus licheniformis*, *Bacillus stearothermophilus* ou *Bacillus subtilis*. Em um outro aspecto preferido, a célula de *Bacillus* é um *Bacillus* alcalofílico.

A introdução de um vetor em uma célula hospedeira bacteriana, por exemplo, pode ser efetuada pela transformação de protoplasto 10 (ver, por exemplo, Chang e Cohen, 1979, Molecular General Genetics 168: 111-115), usando células competentes (ver, por exemplo, Young e Spizizen, 1961, Journal of Bacteriology 81: 823-829 ou Dubnau e Davidoff-Abelson, 1971, Journal of Molecular Biology 56: 209-221), eletroporação (ver, por exemplo, Shigekawa e Dower, 1988, Biotechniques 6: 742-751) ou 15 conjugação (ver, por exemplo, Koehler e Thorne, 1987, Journal of Bacteriology 169: 5771-5278).

A célula hospedeira também pode ser um eucariota, tal como uma célula de mamífero, inseto, planta ou fungo.

Em um aspecto preferido, a célula hospedeira é uma célula 20 fúngica. “Fungos” como aqui usado inclui o filo *Ascomycota*, *Basidiomycota*, *Chytridiomycota* e *Zygomycota* (como definido por Hawksworth *et al.*, Em, Ainsworth e Bisby’s Dictionary of The Fungi, 8ª edição, 1995, CAB International, University Press, Cambridge, UK) assim como o *Oomycota* (como citado em Hawksworth *et al.*, 1995, *supra*, página 171) e todos os 25 fungos mitospóricos (Hawksworth *et al.*, 1995, *supra*).

Em um aspecto mais preferido, a célula hospedeira fúngica é uma célula de levedura. “Levedura” como aqui usado inclui a levedura ascosporogênica (*Endomycetales*), levedura basidiosporogênica e levedura pertencente aos Fungos Imperfeitos (*Blastomycetes*). Visto que a classificação

de levedura pode mudar no futuro, para o propósito desta invenção, levedura deve ser definida como descrito em *Biology and Activities of Yeast* (Skinner, F. A., Passmore, S. M. e Davenport, R. R., eds, Soc. App. Bacteriol. Symposium Series Nº 9, 1980).

5 Em um aspecto ainda mais preferido, a célula hospedeira de levedura é uma célula de *Candida*, *Hansenula*, *Kluyveromyces*, *Pichia*, *Saccharomyces*, *Schizosaccharomyces* ou *Yarrowia*.

 Em um aspecto mais preferido, a célula de levedura hospedeira é uma célula de *Saccharomyces carlsbergensis*, *Saccharomyces cerevisiae*, *Saccharomyces diastaticus*, *Saccharomyces douglasii*,
10 *Saccharomyces kluyveri*, *Saccharomyces norbensis* ou *Saccharomyces oviformis*. Em um outro aspecto mais preferido, a célula hospedeira de levedura é uma célula *Kluyveromyces lactis*. Em um outro aspecto mais preferido, a célula hospedeira de levedura é uma célula *Yarrowia lipolitica*.

15 Em um outro aspecto mais preferido, a célula hospedeira fúngica é uma célula fúngica filamentosa. “Fungo filamentoso” inclui todas as formas filamentosas da subdivisão *Eumycota* e *Oomycota* (como definido por Hawksworth *et al.*, 1995, *supra*). Os fungos filamentosos são no geral caracterizados por uma parede micelial composta de quitina, celulose,
20 glicano, quitosano, manana e outros polissacarídeos complexos. O crescimento vegetativo é pelo alongamento hifal e o catabolismo de carbono é obrigatoriamente aeróbico. Ao contrário, o crescimento vegetativo pelas leveduras tais como *Saccharomyces cerevisiae* é pelo brotamento de um talo unicelular e o catabolismo do carbono pode ser fermentativo.

25 Em um aspecto ainda mais preferido, a célula hospedeira fúngica filamentosa é uma célula de *Acremonium*, *Aspergillus*, *Aureobasidium*, *Bjerkandera*, *Ceriporiopsis*, *Coprinus*, *Coriolus*, *Cryptococcus*, *Filibasidium*, *Fusarium*, *Humicola*, *Magnaporthe*, *Mucor*, *Myceliophthora*, *Neocallimastix*, *Neurospora*, *Paecilomyces*, *Penicillium*,

Phanerochaete, *Phlebia*, *Piromyces*, *Pleurotus*, *Schizophyllum*, *Talaromyces*, *Thermoascus*, *Thielavia*, *Tolypocladium*, *Trametes* ou *Trichoderma*.

Em um aspecto mais preferido, a célula hospedeira fúngica filamentosa é uma célula de *Aspergillus awamori*, *Aspergillus fumigatus*,
 5 *Aspergillus foetidus*, *Aspergillus japonicus*, *Aspergillus nidulans*, *Aspergillus niger* ou *Aspergillus oryzae*. Em um outro aspecto mais preferido, a célula hospedeira fúngica filamentosa é uma célula de *Fusarium bactridioides*, *Fusarium cerealis*, *Fusarium crookwellense*, *Fusarium culmorum*, *Fusarium graminearum*, *Fusarium graminum*, *Fusarium heterosporum*, *Fusarium negundi*,
 10 *Fusarium oxysporum*, *Fusarium reticulatum*, *Fusarium roseum*, *Fusarium sambucinum*, *Fusarium sarcochroum*, *Fusarium sporotrichioides*, *Fusarium sulphureum*, *Fusarium torulosum*, *Fusarium trichothecioides* ou *Fusarium venenatum*. Em um outro aspecto mais preferido, a célula hospedeira fúngica filamentosa é uma célula *Bjerkandera adusta*,
 15 *Ceriporiopsis aneirina*, *Ceriporiopsis aneirina*, *Ceriporiopsis caregiea*, *Ceriporiopsis gilvescens*, *Ceriporiopsis pannocinta*, *Ceriporiopsis rivulosa*, *Ceriporiopsis subrufa*, *Ceriporiopsis subvermispora*, *Coprinus cinereus*, *Coriolus hirsutus*, *Humicola insolens*, *Humicola lanuginosa*, *Mucor miehei*, *Myceliophthora thermophila*, *Neurospora crassa*, *Penicillium purpurogenum*,
 20 *Phanerochaete chrysosporium*, *Phlebia radiata*, *Pleurotus eryngii*, *Thielavia achromatica*, *Thielavia albomyces*, *Thielavia albopilosa*, *Thielavia australeinsis*, *Thielavia fimeti*, *Thielavia microspora*, *Thielavia ovispora*, *Thielavia peruviana*, *Thielavia spededonium*, *Thielavia setosa*, *Thielavia subthermophila*, *Thielavia terrestris*, *Thielavia terricola*, *Thielavia thermophila*, *Thielavia variospora*, *Thielavia wareingii*,
 25 *Trametes villosa*, *Trametes versicolor*, *Trichoderma harzianum*, *Trichoderma longibrachiatum*, *Trichoderma reesei* ou *Trichoderma viride*.

As células fúngicas podem ser transformadas por um processo que envolve a formação de protoplasto, transformação dos protoplastos e

regeneração da parede celular em uma maneira conhecida por si. Os procedimentos adequados para a transformação de células hospedeiras de *Aspergillus* e *Trichoderma* são descritos na EP 238 023 e Yelton *et al.*, 1984, Proceedings of the National Academy of Sciences USA 81: 1470-1474. Os métodos adequados para transformar espécies de *Fusarium* são descritos por Malardier *et al.*, 1989, Gene 78: 147-156 e WO 96/00787. As leveduras podem ser transformadas usando os procedimentos descritos por Becker e Guarente, Em Abelson, J. N. e Simon, M. I., editores, Guide to Yeast Genetics and Molecular Biology, Methods in Enzymology, Volume 194, pp 182-187, Academic Press, Inc., Nova Iorque; Ito *et al.*, 1983, Journal of Bacteriology 153: 163; e Hinnen *et al.*, 1978, Proceedings of the National Academy of Sciences USA 75: 1920.

Métodos de Produção

A presente invenção também diz respeito a métodos para produzir um polipeptídeo da presente invenção, que compreendem: (a) cultivar uma célula, que na sua forma do tipo selvagem é capaz de produzir o polipeptídeo, sob condições condutivas para a produção do polipeptídeo; e (b) recuperar o polipeptídeo. Em um aspecto preferido, a célula é do gênero *Thielavia*. Em um aspecto mais preferido, a célula é *Thielavia terrestris*. Em um aspecto mais preferido, a célula é *Thielavia terrestris* NRRL 8126.

A presente invenção também diz respeito a métodos para produzir um polipeptídeo da presente invenção, que compreendem: (a) cultivar uma célula hospedeira sob condições condutivas para a produção do polipeptídeo; e (b) recuperar o polipeptídeo.

A presente invenção também diz respeito a métodos para produzir um polipeptídeo da presente invenção, que compreendem: (a) cultivar uma célula hospedeira sob condições condutivas para a produção do polipeptídeo, em que a célula hospedeira compreende uma sequência de nucleotídeo mutante que compreende pelo menos uma mutação na sequência

codificadora do polipeptídeo maduro da SEQ ID NO: 1, em que a sequência de nucleotídeo mutante codifica um polipeptídeo que consiste do polipeptídeo maduro da SEQ ID NO: 2 e (b) recuperar o polipeptídeo. Em um aspecto preferido, o polipeptídeo maduro tem os aminoácidos de 18 a 336 da SEQ ID NO: 2.

Nos métodos de produção da presente invenção, as células são cultivadas em um meio nutriente adequado para a produção do polipeptídeo usando métodos bem conhecidos na técnica. Por exemplo, a célula pode ser cultivada pelo cultivo em frasco agitado e fermentação em escala pequena ou escala grande (incluindo fermentações contínuas, por batelada, retro alimentadas ou de estado sólido) em fermentadores de laboratório ou industriais realizada em um meio adequado e sob condições que permitam que o polipeptídeo seja expressado e/ou isolado. O cultivo ocorre em um meio nutriente adequado que compreende fontes de carbono e nitrogênio e sais inorgânicos, usando procedimentos conhecidos na técnica. Os meios adequados são disponíveis de fornecedores comerciais ou podem ser preparados de acordo com composições publicadas (por exemplo, em catálogos da American Type Culture Collection). Se o polipeptídeo é secretado no meio nutriente, o polipeptídeo pode ser recuperado diretamente do meio. Se o polipeptídeo não é secretado no meio, o mesmo pode ser recuperado de lisados celulares.

Os polipeptídeos podem ser detectados usando métodos conhecidos na técnica que são específicos para os polipeptídeos. Estes métodos de detecção podem incluir o uso de anticorpos específicos, a formação de um produto enzimático ou o desaparecimento de um substrato enzimático. Por exemplo, um ensaio de enzima pode ser usado para determinar a atividade do polipeptídeo como aqui descrita.

O polipeptídeo resultante pode ser recuperado usando métodos conhecidas na técnica. Por exemplo, o polipeptídeo pode ser recuperado do

meio nutriente pelos procedimentos convencionais incluindo, mas não limitado a, centrifugação, filtração, extração, secagem por pulverização, evaporação ou precipitação.

Os polipeptídeos da presente invenção podem ser purificados por uma variedade de procedimentos conhecidos na técnica incluindo, mas não limitado a, cromatografia (por exemplo, troca iônica, de afinidade, hidrofóbica, cromatofocalização e por exclusão de tamanho), procedimentos eletroforético (por exemplo, a focalização isoeletrica preparativa), solubilidade diferencial (por exemplo, precipitação com sulfato de amônio), SDS-PAGE ou extração (ver, por exemplo, Protein Purification, J. C. Janson e Lars Ryden, editores, VCH, Publishers, Nova Iorque, 1989) para obter os polipeptídeos substancialmente puros.

Plantas

A presente invenção também diz respeito a uma planta transgênica, parte de planta ou célula vegetal que foram transformadas com uma sequência de nucleotídeo que codifica um polipeptídeo tendo atividade de endoglicanase da presente invenção de modo a expressar e produzir o polipeptídeo em quantidades recuperáveis. O polipeptídeo podem ser recuperado da planta ou parte de planta. Alternativamente, a planta ou parte de planta contendo o polipeptídeo recombinante podem ser usadas como tais para melhorar a qualidade de um alimento ou ração, por exemplo, melhorar o valor nutricional, palatabilidade e propriedades reológicas ou para destruir um fator antinutritivo.

A planta transgênica pode ser dicotiledônea (uma dicot) ou monocotiledônea (uma monocot). Os exemplos de plantas monocot são gramas, tais como a grama dos prados (grama azul, *Poa*), gramas forrageiras tais como *Festuca*, *Lolium*, grama temperada, tal como *Agrostis* e cereais, por exemplo, trigo, aveias, centeio, cevada, arroz, sorgo e milho.

Os exemplos de plantas dicot são tabaco, legumes, tais como

tremoços, batata, beterraba açucareira, ervilha, feijão e soja e plantas crucíferas (família *Brassicaceae*), tais como couve-flor, semente de colza e o organismo modelo intimamente relacionado *Arabidopsis thaliana*.

Os exemplos de partes de plantas são haste, calo, folhas, raiz, frutas, sementes e tubérculos assim como os tecidos individuais que compreendem estas partes, por exemplo, epiderme, mesófilo, parênquima, tecidos vasculares, meristemas. Os compartimentos de célula vegetal específicos, tais como cloroplastos, apoplastos, mitocôndria, vacúolos, peroxissomas e citoplasma também são considerados ser uma parte da planta. Além disso, qualquer célula vegetal, qualquer que seja a origem do tecido, é considerada ser uma parte da planta. Do mesmo modo, partes de planta tais como tecidos específicos e células isoladas para facilitar a utilização da invenção são também considerados partes de planta, por exemplo, embriões, endospermas, aleurona e revestimentos de semente.

Também incluídos dentro do escopo da presente invenção estão a progênie de tais plantas, partes de planta e células vegetais.

A planta ou célula vegetal transgênicas que expressam um polipeptídeo da presente invenção podem ser construídos de acordo com métodos conhecidos na técnica. Em resumo, a planta ou célula vegetal é construída pela incorporação de uma ou mais construções de expressão que codificam um polipeptídeo da presente invenção no genoma hospedeiro vegetal ou genoma de cloroplasto e propagando a planta ou célula vegetal modificadas resultantes em uma planta ou célula vegetal transgênica.

A construção de expressão é convenientemente uma construção de ácido nucleico que compreende um polinucleotídeo que codifica um polipeptídeo da presente invenção operavelmente ligado com seqüências reguladoras apropriadas requeridas para a expressão da seqüência de nucleotídeo na planta ou parte da planta de escolha. Além disso, a construção de expressão pode compreender um marcador selecionável útil

para identificar células hospedeiras nas quais a construção de expressão foi integrada e as seqüências de DNA necessárias para a introdução da construção na planta em questão (o último depende do método de introdução de DNA a ser usado).

5 A escolha de seqüências reguladoras, tais como as seqüências promotoras e terminadoras e opcionalmente as seqüências de sinal ou trânsito é determinada, por exemplo, com base em quando, onde e como o polipeptídeo é desejado ser expressado. Por exemplo, a expressão do gene que codifica um polipeptídeo da presente invenção pode ser constitutiva ou
10 indutível ou pode ser específico de desenvolvimento, estágio ou tecido e o produto de gene pode ser alvejado a um tecido específico ou parte de planta tal como sementes ou folhas. As seqüências reguladoras são, por exemplo, descritas por Tague *et al.*, 1988, Plant Physiology 86: 506.

Para a expressão constitutiva, o promotor de 35S-CaMV, da
15 ubiquitina 1 do milho e da actina 1 do arroz podem ser usada (Franck *et al.*, 1980, Cell 21: 285-294, Christensen *et al.*, 1992, Plant Mol. Biol. 18: 675-689; Zhang *et al.*, 1991, Plant Cell 3: 1155-1165). Os promotores específicos de órgão podem ser, por exemplo, um promotor de tecidos de depósito de armazenagem tal como sementes, tubérculos de batata e frutas (Edwards &
20 Coruzzi, 1990, Ann. Rev. Genet. 24: 275-303) ou de tecidos de depósito metabólico tal como meristemas (Ito *et al.*, 1994, Plant Mol. Biol. 24: 863-878), um promotor específico de semente tal como os promotores da glutelina, prolamina, globulina ou albumina do arroz (Wu *et al.*, 1998, Plant and Cell Physiology 39: 885-889), um promotor da *Vicia faba* da legumina
25 B4 e o gene da proteína de semente desconhecida de *Vicia faba* (Conrad *et al.*, 1998, Journal of Plant Physiology 152: 708-711), um promotor de uma proteína de corpo oleoso de semente (Chen *et al.*, 1998, Plant and Cell Physiology 39: 935-941), o promotor napA da proteína de armazenagem da *Brassica napus* ou qualquer outro promotor específico de semente conhecido

na técnica, por exemplo, como descrito na WO 91/14772. Além disso, o promotor pode ser um promotor específico de folha tal como o promotor *rbcs* do arroz ou tomate (Kyoizuka *et al.*, 1993, *Plant Physiology* 102: 991-1000, o promotor do gene da adenina metiltransferase do vírus *chlorella* (Mitra e Higgins, 1994, *Plant Molecular Biology* 26: 85-93) ou o promotor do gene *aldP* do arroz (Kagaya *et al.*, 1995, *Molecular and General Genetics* 248: 668-674) ou um promotor indutível por fermento tal como o promotor *pin2* da batata (Xu *et al.*, 1993, *Plant Molecular Biology* 22: 573-588). Do mesmo modo, o promotor pode ser indutível pelos tratamentos abióticos tais como temperatura, seca ou alterações na salinidade ou induzido pelas substâncias exogenamente aplicadas que ativam o promotor, por exemplo, etanol, estrogênios, hormônios vegetais tais como etileno, ácido abscísico e ácido giberélico e metais pesados.

Um elemento realçador de promotor também pode ser usado para se obter a expressão mais alta de um polipeptídeo da presente invenção na planta. Por exemplo, o elemento realçador de promotor pode ser um intron que é colocado entre o promotor e a sequência de nucleotídeo que codifica um polipeptídeo da presente invenção. Por exemplo, Xu *et al.*, 1993, *supra*, divulga o uso do primeiro intron do gene da actina 1 do arroz para realçar a expressão.

O gene marcador selecionável e qualquer outras partes da construção de expressão podem ser escolhidos daqueles disponíveis na técnica.

A construção de ácido nucleico é incorporada no genoma vegetal de acordo com técnicas convencionais conhecidas na técnica, incluindo a transformação mediada por *Agrobacterium*, transformação mediada por vírus, microinjeção, bombardeamento de partícula, transformação biolística e eletroporação (Gasser *et al.*, 1990, *Science* 244: 1293; Potrykus, 1990, *Bio/Technology* 8: 535; Shimamoto *et al.*, 1989,

Nature 338: 274).

Presentemente, a transferência de gene mediada por *Agrobacterium tumefaciens* é o método de escolha para gerar dicots transgênicos (para uma revisão, ver Hooykas e Schilperoort, 1992, Plant Molecular Biology 19: 15-38) e também pode ser usada para transformar monocots, embora outros métodos de transformação sejam freqüentemente usados para estas plantas. Presentemente, o método de escolha para gerar monocots transgênicos é o bombardeamento de partícula (partículas de ouro ou tungstênio microscópicas revestidas com o DNA transformador) de calos embrionários ou embriões em desenvolvimento (Christou, 1992, Plant Journal 2: 275-281; Shimamoto, 1994, Current Opinion Biotechnology 5: 158-162; Vasil *et al.*, 1992, Bio/Technology 10: 667-674). Um método alternativo para a transformação de monocots está fundamentado na transformação de protoplasto como descrito por Omirulleh *et al.*, 1993, Plant Molecular Biology 21: 415-428.

Seguindo a transformação, os transformantes tendo incorporado a construção de expressão são selecionados e regenerados em plantas inteiras de acordo com métodos bem conhecidos na técnica. Freqüentemente o procedimento de transformação é planejado para a eliminação seletiva de genes de seleção durante a regeneração ou nas gerações seguintes usando-se, por exemplo, a co-transformação com duas construções de T-DNA separadas ou excisão específica de sítio do gene de seleção por uma recombinase específica.

A presente invenção também diz respeito a métodos para produzir um polipeptídeo da presente invenção que compreende: (a) cultivar uma planta transgênica ou uma célula vegetal que compreenda um polinucleotídeo que codifique um polipeptídeo tendo atividade de endoglicanase da presente invenção sob condições condutivas para a produção do polipeptídeo; e (b) recuperar o polipeptídeo.

Remoção ou Redução da Atividade de Endoglicanase

A presente invenção também diz respeito a métodos para produzir um mutante de uma célula precursora, que compreende romper ou deletar uma seqüência de polinucleotídeo ou uma porção desta, que codifica um polipeptídeo da presente invenção, que resulta na célula mutante produzir menos do polipeptídeo do que a célula quando cultivadas sob as mesmas condições.

A célula mutante pode ser construída reduzindo-se ou eliminando-se a expressão de uma seqüência de nucleotídeo que codifica um polipeptídeo da presente invenção usando métodos bem conhecidos na técnica, por exemplo, inserções, rompimentos, substituições ou deleções. Em um aspecto preferido, a seqüência de nucleotídeo é inativada. A seqüência de nucleotídeo a ser modificada ou inativada pode ser, por exemplo, a região codificadora ou uma parte desta essencial para a atividade de ou um elemento regulador requerido para a expressão da região codificadora. Um exemplo de uma tal seqüência reguladora ou de controle pode ser uma seqüência promotora ou uma parte funcional desta, isto é, uma parte que é suficiente para afetar a expressão da seqüência de nucleotídeo. Outras seqüências de controle para a modificação possível incluem, mas não são limitadas a, uma seqüência líder, de poliadenilação, seqüência de propeptídeo, seqüência de peptídeo de sinal, terminador de transcrição e ativador transcricional.

A modificação ou inativação da seqüência de nucleotídeo pode ser realizada submetendo-se a célula precursora à mutagênese e selecionando quanto as células mutantes em que a expressão da seqüência de nucleotídeo foi reduzida ou eliminada. A mutagênese, que pode ser específica ou aleatória, pode ser realizada, por exemplo, pelo uso de um agente mutanizador físico ou químico adequado, pelo uso de um oligonucleotídeo adequado ou submetendo-se a seqüência de DNA à mutagênese gerada pela PCR. Além disso, a mutagênese pode ser realizada pelo uso de qualquer combinação

destes agentes mutagenizadores.

Os exemplos de um agente mutagenizador físico ou químico adequados para o presente propósito incluem irradiação de ultravioleta (UV), hidroxilamina, N-metil-N'-nitro-N-nitrosoguanidina (MNNG), O-metil
5 hidroxilamina, ácido nitroso, sulfonato de etil metano (EMS), bissulfito de sódio, ácido fórmico e análogos de nucleotídeo.

Quando tais agentes são usados, a mutagênese é tipicamente realizada incubando-se a célula precursora a ser mutagenizada na presença do agente mutagenizante de escolha sob condições adequadas e triagem e/ou
10 seleção para mutantes celulares que exibem expressão reduzida ou nenhuma do gene.

A modificação ou inativação da seqüência de nucleotídeo pode ser realizada pela introdução, substituição ou remoção de um ou mais nucleotídeos no gene ou um elemento regulador requerido para a sua
15 transcrição ou tradução. Por exemplo, os nucleotídeos podem ser inseridos ou removidos de modo a resultar na introdução de um códon de parada, a remoção do códon de início ou uma mudança na matriz de leitura aberta. Tal modificação ou inativação podem ser realizadas pela mutagênese direcionada ao sítio ou mutagênese gerada pela PCR de acordo com métodos conhecidos
20 na técnica. Embora, em princípio, a modificação possa ser realizada *in vivo*, isto é, diretamente na célula que expressa a seqüência de nucleotídeo a ser modificada, é preferido que a modificação seja realizada *in vitro* como exemplificado abaixo.

Um exemplo de um modo conveniente para eliminar ou
25 reduzir a expressão de uma seqüência de nucleotídeo por uma célula é fundamentado em técnicas de substituição de gene, deleção de gene ou rompimento de gene. Por exemplo, no método do rompimento de gene, uma seqüência de ácido nucleico que corresponda à seqüência de nucleotídeo endógena é mutagenizada *in vitro* para produzir uma seqüência de ácido

nucleico defeituosa que é depois transformada na célula precursora para produzir um gene defeituoso. Pela recombinação homóloga, a sequência de ácido nucleico defeituosa substitui a sequência de nucleotídeo endógena. Pode ser desejável que a sequência de nucleotídeo defeituosa também codifica um

5 marcador que possa ser usado para a seleção de transformantes em que a sequência de nucleotídeo foi modificada ou destruída. Em um aspecto particularmente preferido, a sequência de nucleotídeo é rompida com um marcador selecionável tal como aqueles aqui descritos.

Alternativamente, a modificação ou inativação da sequência de

10 nucleotídeo pode ser realizada pelas técnicas de anti-sentido ou RNAi estabelecidas usando uma sequência complementar à sequência de nucleotídeo. Mais especificamente, a expressão da sequência de nucleotídeo por uma célula pode ser reduzida ou eliminada pela introdução de uma sequência complementar à sequência de nucleotídeo do gene que pode ser

15 transcrita na célula e é capaz de hibridização para o mRNA produzido na célula. Sob condições que permitam que a sequência de nucleotídeo de anti-sentido complementar hibridize ao mRNA, a quantidade de proteína traduzível é assim reduzida ou eliminada.

A presente invenção diz respeito ainda a uma célula mutante

20 de uma célula precursora que compreende um rompimento ou deleção de uma sequência de nucleotídeo que codifica o polipeptídeo ou uma sequência de controle desta, que resulta na célula mutante produzir menos do polipeptídeo ou no polipeptídeo comparado com a célula precursora.

As células mutantes deficientes em polipeptídeo assim criadas

25 são particularmente úteis como células hospedeiras para a expressão de polipeptídeos homólogos e/ou heterólogos. Portanto, a presente invenção diz respeito ainda a métodos para produzir um polipeptídeo homólogo ou heterólogo que compreendem: (a) cultivar a célula mutante sob condições condutivas para a produção do polipeptídeo; e (b) recuperar o polipeptídeo. O

termo “polipeptídeos heterólogos” é aqui definido como polipeptídeos que não são nativos para a célula hospedeira, uma proteína nativa em que as modificações foram feitas para alterar a sequência nativa ou uma proteína nativa cuja expressão é quantitativamente alterada como um resultado de uma manipulação da célula hospedeira pelas técnicas de DNA recombinante.

Em um outro aspecto, a presente invenção diz respeito a um método para produzir um produto de proteína essencialmente livre de atividade de endoglicanase pela fermentação de uma célula que produz tanto um polipeptídeo da presente invenção assim como o produto de proteína de interesse pela adição de uma quantidade eficaz de um agente capaz de inibir a atividade de endoglicanase para o caldo de fermentação antes, durante ou depois que a fermentação tenha sido completada, recuperar o produto de interesse do caldo de fermentação e opcionalmente submetendo-se o produto recuperado para outra purificação.

Em um outro aspecto, a presente invenção diz respeito a um método para produzir um produto de proteína essencialmente livre de atividade de endoglicanase pelo cultivo da célula sob condições permitindo a expressão do produto, submetendo o caldo de cultura resultante a um tratamento de pH e temperatura combinados de modo a reduzir a atividade de endoglicanase substancialmente e recuperar o produto do caldo de cultura. Alternativamente, o tratamento de pH e temperatura combinados pode ser realizado em uma preparação de enzima recuperada do caldo de cultura. O tratamento de pH e temperatura combinados pode opcionalmente ser usado em combinação com um tratamento com um inibidor de endoglicanase.

De acordo com este aspecto da invenção, é possível remover pelo menos 60%, preferivelmente pelo menos 75%, mais preferivelmente pelo menos 85%, ainda mais preferivelmente pelo menos 95% e o mais preferivelmente pelo menos 99% da atividade de endoglicanase. A remoção completa da atividade de endoglicanase pode ser obtida pelo uso deste

método.

O tratamento de pH e temperatura combinados é preferivelmente realizado a um pH na faixa de 2 a 3 ou 10 a 11 e uma temperatura na faixa de pelo menos 75 a 85°C por um período suficiente de tempo para atingir o efeito desejado, onde tipicamente, de 1 a 3 horas é suficiente.

Os métodos usados para o cultivo e purificação do produto de interesse podem ser realizados pelos métodos conhecidas na técnica.

Os métodos da presente invenção para produzir um produto essencialmente livre de endoglicanase é de interesse particular na produção de polipeptídeos eucarióticos, em particular proteínas fúngicas tais como enzimas. A enzima pode ser selecionada, por exemplo, de uma enzima amilolítica, enzima lipolítica, enzima proteolítica, enzima celulítica, oxidorreductase ou enzima que degrada a parede da célula vegetal. Os exemplos de tais enzimas incluem uma aminopeptidase, amilase, amiloglicosidase, carboidrase, carboxipeptidase, catalase, celulase, quitinase, cutinase, ciclodextrina glicosiltransferase, desoxirribonuclease, esterase, galactosidase, beta-galactosidase, glicoamilase, glicose oxidase, glicosidase, haloperoxidase, hemicelulase, invertase, isomerase, lacase, ligase, lipase, liase, manosidase, oxidase, enzima pectinolítica, peroxidase, fitase, fenoloxidase, polifenoloxidase, enzima proteolítica, ribonuclease, transferase, transglutaminase ou xilanase. As células deficientes em endoglicanase também podem ser usadas para expressar proteínas heterólogas de interesse farmacêutico tais como hormônios, fatores de crescimento, receptores e outros.

Será entendido que o termo “polipeptídeos eucarióticos” inclui não apenas polipeptídeos nativos, mas também aqueles polipeptídeos, por exemplo, enzimas, que foram modificadas pelas substituições, deleções ou adições de aminoácido ou outras de tais modificações para realçar a atividade,

termoestabilidade, tolerância ao pH e outros.

Em um outro aspecto, a presente invenção diz respeito a um produto de proteína essencialmente livre de atividade de endoglicanase que é produzido por um método da presente invenção.

5 Composições

A presente invenção também diz respeito às composições que compreendem um polipeptídeo da presente invenção. Preferivelmente, as composições são enriquecidas em um tal polipeptídeo. O termo “enriquecidas” indica que a atividade de endoglicanase da composição foi
10 aumentada, por exemplo, com um fator de enriquecimento de pelo menos 1,1.

A composição pode compreender um polipeptídeo da presente invenção como o componente enzimático principal, por exemplo, uma composição de mono-componente. Alternativamente, a composição pode compreender atividades enzimáticas múltiplas, tais como uma
15 aminopeptidase, amilase, carboidrase, carboxipeptidase, catalase, celulase, quitinase, cutinase, ciclodextrina glicosiltransferase, desoxirribonuclease, esterase, alfa-galactosidase, beta-galactosidase, glicoamilase, alfa-glicosidase, beta-glicosidase, haloperoxidase, invertase, lacase, lipase, manosidase, oxidase, enzima pectinolítica, peptidoglutaminase, peroxidase, fitase,
20 polifenoloxidase, enzima proteolítica, ribonuclease, transglutaminase ou xilanase. A(s) enzima(s) adicionais pode(m) ser produzida(s), por exemplo, por um microorganismo pertencente ao gênero *Aspergillus*, preferivelmente *Aspergillus aculeatus*, *Aspergillus awamori*, *Aspergillus fumigatus*, *Aspergillus foetidus*, *Aspergillus japonicus*, *Aspergillus nidulans*, *Aspergillus niger* ou *Aspergillus oryzae*; *Fusarium*, preferivelmente *Fusarium bac*
25 *bactridioides*, *Fusarium cerealis*, *Fusarium crookwellense*, *Fusarium culmorum*, *Fusarium graminearum*, *Fusarium graminum*, *Fusarium heterosporum*, *Fusarium negundi*, *Fusarium oxysporum*, *Fusarium reticulatum*, *Fusarium roseum*, *Fusarium sambucinum*, *Fusarium*

sarcochroum, *Fusarium sulphureum*, *Fusarium torulozeum*, *Fusarium trichothecioides* ou *Fusarium venenatum*; *Humicola*, preferivelmente *Humicola insolens* ou *Humicola lanuginosa*; ou *Trichoderma*, preferivelmente *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei* ou *Trichoderma viride*.

As composições de polipeptídeo podem ser preparadas de acordo com métodos conhecidas na técnica e podem estar na forma de uma composição líquida ou uma composição seca. Por exemplo, a composição de polipeptídeo pode estar na forma de um granulado ou um microgranulado. O polipeptídeo a ser incluído na composição pode ser estabilizado de acordo com métodos conhecidos na técnica.

Os exemplos são dados abaixo de usos preferidos das composições de polipeptídeo da invenção. A dosagem da composição de polipeptídeo da invenção e outras condições sob as quais a composição é usada podem ser determinadas com base nos métodos conhecidos na técnica.

Usos

A presente invenção também é direcionada aos métodos para usar os polipeptídeos tendo atividade de endoglicanase ou composições destas.

20 Degradação de Biomassa para Monossacarídeos, Dissacarídeos e Polissacarídeos

Os polipeptídeos tendo atividade de endoglicanase e células hospedeiras da presente invenção, podem ser usadas na produção de monossacarídeos, dissacarídeos e polissacarídeos como matérias primas químicas ou de fermentação da biomassa para a produção de etanol, plásticos ou outros produtos ou intermediários. Os polipeptídeos tendo atividade de endoglicanase podem estar na forma de um caldo de fermentação bruto com ou sem as células removidas ou na forma de uma preparação de enzima semi-purificada ou purificada. Alternativamente, uma célula hospedeira da presente

invenção pode ser usada como uma fonte do polipeptídeo tendo atividade de endoglicanase em um processo de fermentação com a biomassa.

A biomassa pode incluir, mas não é limitada a, fontes de madeira, resíduo sólido municipal, papel residual e resíduos de safra (ver, por exemplo, Wiseloge *et al*, 1995, em Handbook on Bioethanol (Charles E. Wyman, editor), pp. 105-118, Taylor & Francis, Washington D.C.; Wyman, 1994, Bioresource Technology 50: 3-16; Lynd, 1990, Applied Biochemistry and Biotechnology 24/25: 695-719; Mosier *et al*, 1999, Recent Progress in Bioconversion of Lignocellulosics, em Advances in Biochemical Engineering/ Biotechnology, T. Scheper, editor geral, Volume 65, pp. 23-40, Springer-Verlag, Nova Iorque).

O polissacarídeo predominante na parede da célula primária de biomassa é a celulose, a segunda mais abundante é a hemi-celulose e a terceira é a pectina. A parede celular secundária, produzida depois que a célula parou de crescer, também contém polissacarídeos e é fortalecida pela lignina polimérica covalentemente reticulada para hemi-celulose. A celulose é um homopolímero de anidroclobiose e assim um beta-(1-4)-D-glicano linear, enquanto que as hemi-celuloses incluem uma variedade de compostos, tais como xilanos, xiloglicanos, arabinoxilanos e mananas em estruturas ramificadas complexas com um espectro de substituintes. Embora no geral polimorfa, a celulose é encontrada em tecido vegetal primariamente como uma matriz cristalina insolúvel de cadeias de glicano paralelas. Nas hemi-celuloses usualmente o hidrogênio liga-se à celulose, assim como a outras hemi-celuloses, que ajuda a estabilizar a matriz de parede celular.

Três classes principais de glicoidrolases são usadas para romper a biomassa celulósica:

(1) As “endo-1,4-beta-glicanases” ou 1,4-beta-D-glicano-4-glicanoidrolases (EC 3.2.1.4), que atuam aleatoriamente sobre substratos de 1,4-beta-glicano solúveis e insolúveis.

(2) As “exo-1,4-beta-D-glicanases” incluindo tanto as 1,4-beta-D-glicano glicoidrolases (EC 3.2.1.74), que liberam D-glicose a partir de 1,4-beta-D-glicanos e hidrolisam D-celobiose lentamente e celobioidrolases (1,4-beta-D-glicano celobioidrolases, EC 3.2.1.91), que liberam D-celobiose dos 1,4-beta-glicanos.

(3) As “beta-D-glicosidases” ou beta-D-glucosídeo glicoidrolases (EC 3.2.1.21), que atuam para liberar unidades de D-glicose a partir de celobiose e celodextrinas solúveis, assim como uma série de glicosídeos.

Estas três classes de enzimas trabalham juntas sinergisticamente resultando na descristalização e hidrólise eficientes da celulose nativa a partir da biomassa para produzir açúcares redutores.

Os polipeptídeos tendo atividade de endoglicanase da presente invenção podem ser usados em conjunção com as enzimas mencionadas acima para degradar ainda mais o componente de celulose do substrato de biomassa, (ver, por exemplo, Brigham *et al.*, 1995, em Handbook on Bioethanol (Charles E. Wymano, editor), pp. 119-141, Tailor & Francis, Washington D.C.; Lee, 1997, Journal of Biotechnology 56: 1-24).

O etanol pode ser produzido pela degradação enzimática de biomassa e conversão dos sacarídeos liberados ao etanol. Este tipo de etanol é freqüentemente aludido como bioetanol ou biocombustível. Este pode ser usado como um aditivo de combustível ou diluente em misturas a partir de menos do que 1% e até 100% (um substituto de combustível).

Composições Detergentes

Os polipeptídeos tendo atividade de endoglicanase da presente invenção podem ser adicionados a uma composição detergente e assim se tornar um componente desta.

A composição detergente da presente invenção pode ser, por exemplo, formulada como uma composição detergente de lavagem de roupas

a mão ou em máquina incluindo uma composição de aditivo de lavar roupas adequada para o pré tratamento de tecidos tingidos e uma composição amaciante de tecido adicionada ao enxágüe ou formulada como uma composição detergente para o uso nas operações de limpeza de superfície dura doméstica geral ou formuladas para operações de lavagem de louça manual ou em máquina.

Em um aspecto específico, a presente invenção fornece um aditivo de detergente que compreende os polipeptídeos tendo atividade de endoglicanase da presente invenção. O aditivo de detergente assim como a composição detergente podem compreender uma ou mais outras enzimas tais como uma protease, lipase, cutinase, uma amilase, carboidrase, celulase, pectinase, mananase, arabinase, galactanase, xilanase, oxidase, por exemplo, uma lacase e/ou peroxidase.

No geral as propriedades dos componentes enzimáticos devem ser compatíveis com o detergente selecionado, (isto é, pH ideal, compatibilidade com outros ingredientes enzimáticos e não enzimáticos, etc.) e os componentes enzimáticos devem estar presentes em quantidades eficazes.

Proteases: As proteases adequadas incluem aquelas de origem animal, vegetal ou microbiana. A origem microbiana é preferida. Mutantes quimicamente modificados ou engendrados de proteína são incluídos. A protease pode ser uma serina protease ou uma metaloprotease, preferivelmente uma protease microbiana alcalina ou uma protease equivalente à tripsina. Os exemplos de proteases alcalinas são subtilisinas, especialmente aquelas derivadas de *Bacillus*, por exemplo, subtilisina Novo, subtilisina Carlsberg, subtilisina 309, subtilisina 147 e subtilisina 168 (descrita na WO 89/06279). Os exemplos de protease equivalentes a tripsinas são tripsina (por exemplo, de origem porcina ou bovina) e a protease *Fusarium* descrita na WO 89/06270 e WO 94/25583.

Os exemplos de proteases úteis são as variantes descritas na

WO 92/19729, WO 98/20115, WO 98/20116 e WO 98/34946, especialmente as variantes com substituições em uma ou mais das seguintes posições: 27, 36, 57, 76, 87, 97, 101, 104, 120, 123, 167, 170, 194, 206, 218, 222, 224, 235 e 274.

5 As enzimas de protease comercialmente disponíveis preferidas incluem Alcalase[®], Savinase[®], Primase[®], Duralase[®], Esperase[®] e Kannase[®] (Novozymes NS), Maxatase[®], Maxacal[®], Maxapem[®], Properase[®], Purafect[®], Purafect OXP[®], FN2[®] e FN3[®] (Genencor International Inc.).

10 Lipases: As lipases adequadas incluem aquelas de origem bacteriana ou fúngica. Os mutantes quimicamente modificados ou engendrados de proteína são incluídos. Os exemplos de lipases úteis incluem lipases de *Humicola* (sinônimo *Thermomyces*), por exemplo, de *H. lanuginosa* (*T. lanuginosus*) como descrito na EP 258 068 e EP 305 216 ou de *H. insolens* como descrito na WO 96/13580, uma lipase de *Pseudomonas*, por exemplo, de *P. alcaligenes* ou *P. pseudoalcaligenes* (EP 218 272), *P. cepacia* (EP 331 376), *P. stutzeri* (GB 1.372.034), *P. fluorescens*, *Pseudomonas* sp. cepa SD 705 (WO 95/06720 e WO 96/27002), *P. wisconsinensis* (WO 96/12012), uma lipase de *Bacillus*, por exemplo, de *B. subtilis* (Dartois *et al*, 1993, Biochemica et Biophysica Acta, 1131, 253-360), *B. stearothermophilus* 20 (JP 64/744992) ou *B. pumilus* (WO 91/16422).

Outros exemplos são variantes de lipase tais como aquelas descritas na WO 92/05249, WO 94/01541, EP 407 225, EP 260 105, WO 95/35381, WO 96/00292, WO 95/30744, WO 94/25578, WO 95/14783, WO 95/22615, WO 97/04079 e WO 97/07202.

25 As lipases comercialmente disponíveis preferidas incluem Lipolase[®], Lipex[®] e Lipolase Ultra[®] (Novozymes NS).

Amilases: As amilases adequadas (α e/ou β) incluem aquelas de origem bacteriana ou fúngica. Mutantes quimicamente modificados ou engendrados de proteína são incluídos. As amilases incluem, por exemplo, α -

amilases obtidas de *Bacillus*, por exemplo, uma cepa especial de *Bacillus licheniformis*, descrita em mais detalhes na GB 1.296.839.

Os exemplos de amilases úteis são as variantes descritas na WO 94/02597, WO 94/18314, WO 96/23873 e WO 97/43424, especialmente
 5 as variantes com substituições em uma ou mais das seguintes posições: 15, 23, 105, 106, 124, 128, 133, 154, 156, 181, 188, 190, 197, 202, 208, 209, 243, 264, 304, 305, 391, 408 e 444.

As amilases comercialmente disponíveis são Duramil[®], Termamil[®], Fungamil[®] e BAN[®] (Novozymes A/S), Rapidase[®] e Purastar[®] (da
 10 Genencor International Inc.).

Celulases: As celulases adequadas incluem aquelas de origem bacteriana ou fúngica. Os mutantes quimicamente modificados ou engendrados de proteína são incluídos. As celulases adequadas incluem celulases dos gêneros *Bacillus*, *Pseudomonas*, *Humicola*, *Fusarium*,
 15 *Thielavia*, *Acremonium* ou *Trichoderma* por exemplo, as celulases fúngicas produzidas a partir de *Humicola insolens*, *Myceliophthora thermophila* e *Fusarium oxysporum* divulgadas na Patente U.S. Nº 4.435.307, Patente U.S. Nº 5.648.263, Patente U.S. Nº 5.691.178, Patente U.S. Nº 5.776.757 e WO 89/09259.

20 As celulases especialmente adequadas são as celulases alcalinas e neutras tendo benefícios de cuidado com as cores. Os exemplos de tais celulases são celulases descritas na EP 0 495 257, EP 0 531 372, WO 96/11262, WO 96/29397, WO 98/08940. Outros exemplos são variantes de celulase tais como aquelas descritas na WO 94/07998, EP 0 531 315, Patente
 25 U.S. Nº 5.457.046, Patente U.S. Nº 5.686.593, Patente U.S. Nº 5.763.254, WO 95/24471, WO 98/12307 e PCT/DK98/00299.

As celulases comercialmente disponíveis incluem Celluclast[®], Celluzyme[®] e Carezyme[®] (Novozymes A/S), Clazinase[®] e Puradax HA[®] (Genencor International Inc.) e KAC-500(B)[®] (Kao Corporation).

Peroxidasas/Oxidases: As peroxidases/oxidases adequadas incluem aquelas de origem vegetal, bacteriana ou fúngica. Os mutantes quimicamente modificados ou engendrados de proteína são incluídos. Os exemplos de peroxidases úteis incluem peroxidases de *Coprinus*, por exemplo, de *C. cinereus* e variantes destes como aquelas descritas na WO 93/24618, WO 95/10602 e WO 98/15257.

As peroxidases comercialmente disponíveis incluem Guardzyme® (Novozymes A/S).

O(s) componente(s) enzimático(s) pode(m) ser incluído(s) em uma composição detergente pela adição de aditivos separados contendo uma ou mais enzimas ou pela adição de um aditivo combinado que compreende todas estas enzimas. Um aditivo de detergente da presente invenção, isto é, um aditivo separado ou um aditivo combinado, podem ser formulados, por exemplo, como um granulado, líquido, pasta fluida, etc. As formulações de aditivo de detergente preferidas são granulados, em particular granulados não empoáveis, líquidos, em particular líquidos estabilizados ou pastas fluidas.

Os granulados não empoáveis podem ser produzidos, por exemplo, como divulgado nas Patentes U.S. Nº 4.106.991 e 4.661.452 e pode opcionalmente ser revestido pelos métodos conhecidas na técnica. Os exemplos de materiais de revestimento cerosos são produtos de poli(óxido de etileno) (polietileno glicol, PEG) com pesos moleculares médios de 1000 a 20000; nonilfenóis etoxilados tendo de 16 a 50 unidades de óxido de etileno; álcoois graxos etoxilados em que o álcool contém de 12 a 20 átomos de carbono e em que existem de 15 a 80 unidades de óxido de etileno; álcoois graxos; ácidos graxos; e mono- e di- e triglicerídeos de ácidos graxos. Os exemplos de materiais de revestimento formadores de película adequados para a aplicação pelas técnicas de leito de fluido são dadas na GB 1483591. As preparações de enzima líquidas, por exemplo, podem ser estabilizadas pela adição de um poliol tal como propileno glicol, um açúcar ou álcool de açúcar,

ácido láctico ou ácido bórico de acordo com métodos estabelecidos. As enzimas protegidas podem ser preparadas de acordo com o método divulgado na EP 238,216.

5 A composição detergente da presente invenção pode estar em qualquer forma conveniente, por exemplo, uma barra, um tablete, um pó, um grânulo, uma pasta ou um líquido. Um detergente líquido pode ser aquoso, tipicamente contendo até 70% de água e de 0 a 30% de solvente orgânico ou não aquoso.

10 A composição detergente compreende um ou mais tensoativos, que podem ser não iônicos incluindo semi-polares e/ou aniônicos e/ou catiônicos e/ou zwitteriônicos. Os tensoativos estão tipicamente presentes em um nível de 0,1% a 60% em peso.

15 Quando nesse ponto incluído o detergente usualmente conterá de cerca de 1% a cerca de 40% de um tensoativo aniônico tal como alquilbenzenossulfonato linear, alfa-olefinsulfonato, alquil sulfato (sulfato de álcool graxo), etoxissulfato álcool, alcanossulfonato secundário, éster metílico de ácido alfa-sulfo graxo, ácido alquil- ou alquenilsuccínico ou sabão.

20 Quando nesse ponto incluído o detergente usualmente conterá de cerca de 0,2% a cerca de 40% de um tensoativo não iônico tal como etoxilato álcool, etoxilato de nonilfenol, alquilpoliglicosídeo, óxido de alquildimetilamino, monoetanolamida de ácido graxo etoxilado, monoetanolamida de ácido graxo, amida de ácido graxo de poliidroxi alquila ou derivados de N-acil N-alquila de glicosamina ("glicamidas").

25 O detergente pode conter de 0 a 65% de um formador de detergente ou agente de complexação tal como zeólito, difosfato, trifosfato, fosfonato, carbonato, citrato, ácido nitrilotriacético, ácido etilenodiaminotetraacético, ácido dietilenotriaminopentaacético, ácido alquil- ou alquenilsuccínico, silicatos solúveis ou silicatos dispostos em camadas (por exemplo, SKS-6 da Hoechst).

O detergente pode compreender um ou mais polímeros. Os exemplos são carboximetilcelulose, poli(vinilpirrolidona), poli(etileno glicol), álcool polivinílico, poli(vinilpiridina-N-óxido), poli(vinil-imidazol), policarboxilatos tais como poliacrilatos, copolímeros do ácido maleico/acrílico e metacrilato de laurila/copolímeros de ácido acrílico.

O detergente pode conter um sistema de branqueamento que pode compreender uma fonte de H_2O_2 tal como perborato ou percarbonato que podem ser combinados com um ativador de branqueamento que forme perácido tal como tetraacetililenodiamina ou nonanoiloxibenzenossulfonato. Alternativamente, o sistema de branqueamento pode compreender peroxiácidos, por exemplo, do tipo amida, imida ou sulfona.

O(s) componente(s) enzimático(s) da composição detergente da presente invenção pode(m) ser estabilizado(s) usando agentes de estabilização convencionais, por exemplo, um poliol tal como propileno glicol ou glicerol, um açúcar ou álcool de açúcar, ácido láctico, ácido bórico ou um derivado de ácido bórico, por exemplo, um éster de borato aromático ou um derivado de ácido fenil borônico tal como o ácido 4-formilfenil borônico e a composição pode ser formulada como descrito, por exemplo, na WO 92/19709 e WO 92/19708.

O detergente também pode conter outros ingredientes detergentes convencionais tais como condicionadores de tecido incluindo argilas, reforçadores de espuma, supressores de espuma, agentes anti corrosão, agentes de suspensão de sujeira, agentes anti-redeposição de sujeira, corantes, bactericidas, abrillantadores óticos, hidrótropos, inibidores de manchamento ou perfumes.

Nas composições detergentes qualquer componente enzimático, em particular os polipeptídeos tendo atividade de endoglicanase da presente invenção, podem ser adicionados em uma quantidade que corresponde de 0,01 a 100 mg de proteína enzimática por litro de líquido de

lavagem, preferivelmente de 0,05 a 5 mg de proteína enzimática por litro de líquido de lavagem, em particular de 0,1 a 1 mg de proteína enzimática por litro de líquido de lavagem.

Os polipeptídeos tendo atividade de endoglicanase da presente invenção podem ser adicionalmente incorporados nas formulações detergentes divulgadas na WO 97/07202 que é aqui incorporada como referência.

Peptídeo de sinal

A presente invenção também diz respeito às construções de ácido nucleico que compreendem um gene que codificam uma proteína operavelmente ligada a uma seqüência de nucleotídeo que codifica um peptídeo de sinal que compreende ou que consiste dos aminoácidos de 1 a 17 da SEQ ID NO: 2, que permitem a secreção da proteína em um meio de cultura, em que o gene é estranho à seqüência de nucleotídeo.

Em um aspecto preferido, a seqüência de nucleotídeo compreende os nucleotídeos de 1 a 51 da SEQ ID NO: 1. Em um outro aspecto preferido, a seqüência de nucleotídeo consiste dos nucleotídeos de 1 a 51 da SEQ ID NO: 1.

A presente invenção também diz respeito a vetores de expressão recombinantes e células hospedeiras recombinantes que compreendem tais construções de ácido nucleico.

A presente invenção também diz respeito a métodos para produzir uma proteína que compreende: (a) cultivar uma tal célula hospedeira recombinante sob condições adequadas para a produção da proteína; e (b) recuperar a proteína.

A proteína pode ser nativa ou heteróloga a uma célula hospedeira. O termo “proteína” não é aqui intencionado a se referir a um comprimento específico do produto codificado e, portanto, abrange peptídeos, oligopeptídeos e proteínas. O termo “proteína” também abrange dois ou mais

polipeptídeos combinados para formar o produto codificado. As proteínas também incluem polipeptídeos híbridos que compreendem uma combinação de seqüências de polipeptídeo parciais ou completas obtidas de pelo menos duas proteínas diferentes em que uma ou mais podem ser heterólogas ou
5 nativas para a célula hospedeira. As proteínas incluem ainda variações alélicas e engendradas que ocorrem naturalmente das proteínas mencionadas acima e proteínas híbridas.

Preferivelmente, a proteína é um hormônio ou variante deste, enzima, receptor ou porção deste, anticorpo ou porção deste ou repórter. Em
10 um aspecto mais preferido, a proteína é uma oxidoredutase, transferase, hidrolase, liase, isomerase ou ligase. Em um aspecto ainda mais preferido, a proteína é uma aminopeptidase, amilase, carboidrase, carboxipeptidase, catalase, celulase, quitinase, cutinase, ciclodextrina glicosiltransferase, desoxirribonuclease, esterase, alfa-galactosidase, beta-galactosidase,
15 glicoamilase, alfa-glicosidase, beta-glicosidase, invertase, lacase, lipase, manosidase, mutanase, oxidase, enzima pectinolítica, peroxidase, fitase, polifenoloxidase, enzima proteolítica, ribonuclease, transglutaminase ou xilanase.

O gene pode ser obtido de qualquer procariótico, eucariótico
20 ou outra fonte.

A presente invenção é ainda descrita pelos seguintes exemplos que não devem ser interpretados como limitantes do escopo da invenção.

Exemplos

Materiais

25 Os produtos químicos usados como tampões e substratos foram produtos comerciais de grau pelo menos reagente.

Os géis de SDS-PAGE, tampão de carga e tampão de condução foram obtidos da Invitrogen/Novex (Carlsbad, CA). A tripsina modificada grau de seqüenciamento foi da Princeton Separations (Aldelphia,

NJ). As manchas de proteína de BioSafe Commassie Blue G250 foram obtidas da BioRad Laboratories (Hercules, CA).

Cepas

5 A cepa *Aspergillus oryzae* Ja1250 (WO 99/61651) foi usada para a expressão de um polipeptídeo de *Thielavia terrestris* tendo atividade de endoglicanase. *Thielavia terrestris* NRRL cepa 8126 foi usada como a fonte de um gene para um polipeptídeo da Família 7F tendo atividade de endoglicanase.

Meios

10 As placas de PDA foram compostas por litro de 39 gramas de ágar de dextrose de batata.

O meio NNCYP foi composto por litro de 5,0 g de NH_4NO_3 , 0,5 g de $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 0,3 g de CaCl_2 , 2,5 g de ácido cítrico, 1,0 g de Bacto Peptona, 5,0 g de extrato de levedura, 1 ml de metais traço COVE e K_2HPO_4 15 suficiente para se obter um pH final de aproximadamente 5,4.

O meio NNCYPmod foi composto por litro de 1,0 g de NaCl , 5,0 g de NH_4NO_3 , 0,2 g de $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 0,2 g de CaCl_2 , 2,0 g de ácido cítrico, 1,0 g de Bacto Peptona, 5,0 g de extrato de levedura, 1 ml de solução de metais traço COVE e K_2HPO_4 suficiente para se obter o pH final de 20 aproximadamente 5,4.

A solução de metais traço COVE foi composta por litro de 0,04 g de $\text{Na}_2\text{B}_4\text{O}_7 \cdot 10\text{H}_2\text{O}$, 0,4 g de $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$, 1,2 g de $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$, 0,7 g de $\text{MnSO}_4 \cdot \text{H}_2\text{O}$, 0,8 g de $\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$ e 10 g de $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$.

25 As placas LB foram compostas por litro de 10 g de triptona, 5 g de extrato de levedura, 5 g de cloreto de sódio e 15 g de Bacto Ágar.

O meio MDU2BP foi composto por litro de 45 g de maltose, 1 g de $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 1 g de NaCl , 2 g de K_2HSO_4 , 12 g de KH_2PO_4 , 2 g de uréia e 500 µl de metais traço AMG, o pH foi ajustado para 5,0 e depois esterilizado em filtro com uma unidade de filtração de 0,22 µm.

Os metais traço AMG foram compostos por litro de 14,3 g de $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$, 2,5 g de $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$, 0,5 g de $\text{NiCl}_2 \cdot 6\text{H}_2\text{O}$, 13,8 g de $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$, 8,5 g de $\text{MnSO}_4 \cdot 7\text{H}_2\text{O}$ e 3 g de ácido cítrico.

5 O meio SOC foi composto de 2% de triptona, 0,5% de extrato de levedura, 10 mM de NaCl, 2,5 mM de KCl, 10 mM de MgCl_2 e 10 mM de MgSO_4 , esterilizado pela autoclavagem e depois adicionados à glicose esterilizada em filtro a 20 mM.

O meio de congelamento foi composto de 60% de SOC e 40% de glicerol.

10 O meio 2X YT foi composto por litro de 16 g de triptona, 10 g de extrato de levedura, 5 g de NaCl e 15 g de Bacto ágar, esterilizado pela autoclavagem.

Exemplo 1: Construção de biblioteca de cDNA de rótulos de seqüência expressada (EST)

15 *Thielavia terrestris* NRRL 8126 foi cultivada em 50 ml de meio NNCYPmod suplementado com 1% de glicose em um frasco de 250 ml a 45°C, 200 rpm por 24 horas. Uma alíquota de dois ml da cultura líquida de 24 horas foi usada para semear um frasco de 500 ml contendo 100 ml de meio NNCYPmod suplementado com 2% de Sigmacell-20 (Sigma Chemical Co.,
20 St. Louis, MO). A cultura foi incubada a 45°C, 200 rpm por 3 dias. Os micélios foram colhidos por filtração através de um funil de Buchner com um pré filtro de fibra de vidro (Nalgene, Rochester NY), lavados duas vezes com 10 mM de Tris-HCl-1 mM de EDTA pH 8 (TE) e rapidamente congelados em nitrogênio líquido.

25 O RNA total foi isolado usando os seguintes métodos. Os micélios congelados de *Thielavia terrestris* NRRL 8126 foram cultivados em um triturador de café elétrico. O material triturado foi misturado 1:1 v/v com 20 ml de Fenazol (Ambion, Inc., Austin, TX) em um tubo Falcon de 50 ml. Uma vez que os micélios foram colocados em suspensão, eles foram extraídos

com clorofórmio e três vezes com uma mistura de fenol-clorofórmio-álcool isoamílico 25:24:1 v/v/v. A partir da fase aquosa resultante, o RNA foi precipitado pela adição de 1/10 volume de acetato de sódio 3 M pH 5,2 e 1,25 volume de isopropanol. O RNA precipitado foi recuperado pela centrifugação a 12.000 x g por 30 minutos a 4°C. A pelota final foi lavada com etanol a 70% frio, secada ao ar e recolocada em suspensão em 500 ml de água tratada com pirocarbonato de dietila (DEPC-água).

A qualidade e quantidade do RNA purificado foi avaliada com um Bioanalisador Agilent 2100 (Agilent Technologies, Inc., Palo Alto, CA). O mRNA poliadenilado foi isolado de 360 µg de RNA total com a ajuda de um Kit Poli (a) Purist Magnetic (Ambion, Inc., Austin, TX) de acordo com as instruções do fabricante.

Para criar a biblioteca de cDNA, um Kit CloneMiner[®] (Invitrogen, Carlsbad, CA) foi utilizado para construir uma biblioteca direcional que não requer o uso da clonagem de enzima de restrição, por meio deste reduzindo o número de clones quiméricos e tendência de tamanho.

Para garantir o sucesso da síntese do primeiro filamento do cDNA, duas reações foram realizadas em paralelo com duas concentrações diferentes de mRNA (2,2 e 4,4 µg de poli(a)+ mRNA). As amostras de mRNA foram misturadas com um iniciador Biotina-attB2-Oligo(dt) (Kit CloneMiner[®], Invitrogen, Carlsbad, CA), 1X tampão do primeiro filamento (Invitrogen, Carlsbad, CA), 2 µl de ditiontreitol 0,1 M (DTT), 10 mM de cada dNTP e água a um volume final de 18 e 16 µl, respectivamente. As misturas de reação foram cuidadosamente misturadas e depois 2 e 4 µl de transcriptase reversa SuperScript[®] (Invitrogen, Carlsbad, CA) foram adicionados e incubados a 45°C por 60 minutos para sintetizar o primeiro filamento complementar.

Para a síntese do segundo filamento, para cada reação do primeiro filamento foram adicionados 30 µl de 5X tampão do segundo

filamento (Invitrogen, Carlsbad, CA), 3 µl de cada dNTP 10 mM, 10 unidades de DNA ligase de *E. coli* (Invitrogen, Carlsbad, CA), 40 unidades de DNA polimerase I de *E. coli* (Invitrogen, Carlsbad, CA) e 2 unidades de RNase H *E. coli* (Invitrogen, Carlsbad, CA) em um volume total de 150 µl. As misturas
 5 foram depois incubadas a 16°C por duas horas. Depois da incubação de duas horas, 2 µl de T4 DNA polimerase (Invitrogen, Carlsbad, CA) foram adicionados a cada reação e incubados a 16°C por 5 minutos para criar um cDNA abruptamente terminado. As reações de cDNA foram extraídas com uma mistura de fenol-clorofórmio-álcool isoamílico 25:24:1 v/v/v e
 10 precipitada na presença de 20 µg de glicogênio, 120 µl de acetato de amônio 5 M e 660 µl de etanol. Depois da centrifugação a 12.000 x g por 30 minutos a 4°C, as pelotas de cDNA foram lavadas com etanol a 70% frio, secadas sob vácuo por 2 a 3 minutos e recolocadas em suspensão em 18 µl de DEPC-água. Para cada amostra de cDNA recolocada em suspensão foram adicionados 10
 15 µl de 5X tampão adaptado (Invitrogen, Carlsbad, CA), 10 µg de adaptador attB1 (Invitrogen, Carlsbad, CA) mostrado abaixo, 7 µl de DTT 0,1 M e 5 unidades de T4 DNA ligase (Invitrogen, Carlsbad, CA).

Filamento de topo adaptador attB1:

5'-TCGTCGGGGACAACCTTTGTACAAAAAAGTTGG-3' (SEQ ID NO: 3)

20 Filamento de fundo adaptador attB1:

3'-CCCCTGTTGAAACATGTTTTTTCAACCP-5' (SEQ ID NO: 4)

As reações de ligação foram incubadas durante a noite a 16°C. Os adaptadores em excesso foram removidos pela cromatografia de exclusão de tamanho em 1 ml de resina Sephacril® S-500 HR (Amersham Biosciences,
 25 Piscataway, NJ). As frações de coluna foram coletadas de acordo com as instruções do Kit CloneMiner® e as frações 3 a 14 foram analisadas com um Bioanalisador Agilent para determinar a fração na qual os adaptadores attB1 começaram a eluir. Esta análise mostrou que os adaptadores começaram a eluir em torno da fração 10 ou 11. Para a primeira biblioteca as frações de 6 a

11 foram reunidas e para a segunda biblioteca as frações 4 a 11 foram reunidas.

A clonagem do cDNA foi realizada pela recombinação de DNA homólogo de acordo com o protocolo Gateway System (Invitrogen, Carlsbad, CA) usando a BP Clonase[®] (Invitrogen, Carlsbad, CA) como a recombinase. Cada reação de recombinação BP Clonase[®] conteve aproximadamente 70 ng de cDNA flanqueado com attB, 250 ng de pDONR[®]222, 2 µl de tampão 5X BP Clonase[®], 2 µl de tampão TE e 3 µl de BP Clonase[®]. Todos os reagentes foram obtidos da Invitrogen, Carlsbad, CA.

As reações de recombinação foram incubadas a 25°C durante a noite.

As reações de recombinação BP inativadas por calor foram depois divididas em 6 alíquotas e eletroporadas em células eletrocompetentes ElectroMax[®] DH10B (Invitrogen, Carlsbad, CA) usando um BioRad Gene Pulser II (BioRad, Hercules, CA) com os seguintes parâmetros: Voltagem: 2,0 kV; Resistência: 200 Ω; e Capacidade: 25 µF. As células eletroporadas foram recolocadas em suspensão em 1 ml de meio SOC e incubadas a 37°C por 60 minutos com agitação constante a 200 rpm. Depois do período de incubação, as células transformadas foram reunidas e misturadas 1:1 com meio de congelamento. Uma alíquota de 200 µl foi removida para a titulação da biblioteca e depois o resto de cada biblioteca foi aliquotada em criofrascos de 1,8 ml (Wheaton Science Products, Millville, NJ) e armazenados congelados a -80°C.

As diluições em quatro séries de cada biblioteca foram preparadas: 1/100, 1/1000, 1/10⁴, 1/10⁵. De cada diluição 100 µl foram plaqueados em placas LB de 150 mm suplementadas com 50 µg de canamicina por ml e incubadas a 37°C durante a noite. O número de colônias em cada placa de diluição foram contadas e usadas para calcular o número total de transformantes em cada biblioteca.

A primeira biblioteca mostrou ter 5,4 milhões de clones

independentes e a segunda biblioteca mostrou ter 9 milhões de clones independentes.

Exemplo 2: Preparação de direcionador e seqüenciamento de nucleotídeo de clones de cDNA

5 As alíquotas de ambas as bibliotecas foram misturadas e plaqueadas em placas LB de 25 x 25 cm suplementadas com 50 µg de canamicina por ml. as colônias individuais foram dispostas em série sobre placas de 96 reservatórios contendo 100 µl de LB suplementado com 50 µg de canamicina por ml com a ajuda de um Genetix QPix Robot (Genetix Inc.,
10 Boston, MA). Quarenta e cinco placas de 96 reservatórios foram obtidas para um total de 4320 clones individuais. As placas foram incubadas durante a noite a 37°C com agitação a 200 rpm. Depois da incubação, 100 µl de glicerol a 50% estéril foram adicionados a cada reservatório. Os transformantes foram replicados com a ajuda de uma ferramenta de 96 pinos (Boekel, Feasterville,
15 PA) em placas de microcultura secundária, de 96 reservatórios de cavidade profunda (Advanced Genetic Technologies Corporation, Gaithersburg, MD) contendo 1 ml de Magnificent Broth® (MacConnell Research, San Diego, CA) suplementado com 50 µg de canamicina por ml em cada reservatório. As placas de microtítulo primárias foram armazenadas congeladas a -80°C. As
20 placas de cavidade profunda secundárias foram incubadas a 37°C durante a noite com agitação vigorosa a 300 rpm em um agitador rotativo. Para impedir derramar e contaminação cruzada e para permitir aeração suficiente, cada placa de cultura secundária foi coberta com uma almofada de polipropileno (Advanced Genetic Technologies Corporations, Gaithersburg, MD) e uma
25 cobertura de cavidade de microtituladora plástica. O DNA plasmídico foi preparado com um MWG Robot-Smart 384 (MWG Biotech Inc., High Point, NC) e Montage Plasmide Miniprep Kits (Millipore, Billerica, MA).

As reações de seqüenciamento foram realizadas usando a química terminadora de Big-Dye® (Applied Biosystems, Inc., Foster City,

CA) (Giesecke *et al*, 1992, Journal of Virology Methods 38: 47-60) e um iniciador de seqüenciamento M13 Avançado (-20) mostrado abaixo. 5'-GTAAAACGACGGCCAG-3' (SEQ ID NO: 5)

As reações de seqüenciamento foram realizadas em um formato de 384 reservatórios com um Robot-Smart 384 (MWG Biotech Inc., High Point, NC) e remoção de terminador com Kits Millipore MultiScreen Seq384 Seqüencing Clean-up (Millipore, Billerica, MA). As reações contiveram 6 µl de DNA plasmídico e 4 µl de mistura master de seqüenciamento contendo 2 µl de tampão de seqüenciamento 5x (Millipore, Billerica, MA), 1 µl de terminador Big-Dye® (Applied Biosystems, Inc., Foster City, CA), 1,6 pmoles de iniciador M13 Avançado e 1 µl de água. O seqüenciamento do DNA de passagem única foi realizado com um Seqüenciador de DNA Automatizado ABI PRISM Modelo 3700 (Applied Biosystems, Foster City, CA).

Exemplo 3: Análise de dados de seqüência de DNA de clones de cDNA

Chamada de base, designação de valor de qualidade e corte de vetor foram realizados com a ajuda do software PHRED/PHRAP (University of Washington, Seattle, WA). A análise em grupos dos ESTs foi realizada com um Parcel Transcript Assembler v. 2.6.2. (Paracel, Inc., Pasadena, CA). A análise da formação de grupo de EST indicou 395 grupos independentes.

A análise de homologia de seqüência das seqüências EST montadas contra a base de dados PIR foi realizada com o programa Blastx (Altschul *et al.*, 1990, J. Mol. Biol. 215: 403-410) em um grupo Linux de 32 nodos (Paracel, Inc., Pasadena, CA) usando a matriz BLOSUM 62 (Henikoff, 1992, Proc. Natl. Acad. Sci. USA 89: 10915-10919). Dos 395 grupos, 246 tiveram acertos blast para genes conhecidos nas bases de dados de proteína publicas e 149 não tiveram nenhum acerto significativo contra estas bases de dados. Entre estes 246 genes, 13 tiveram acertos contra homólogos bem caracterizados de genes de glicosil hidrolase.

Exemplo 4: Identificação de clones de cDNA que codificam uma endoglicanase da Família 7 (CEL7F)

Um clone de cDNA que codifica uma endoglicanase da Família 7 (CEL7F) foi inicialmente identificado pela sua identidade com a proteína EG-1 da endoglicanase da Família 7 de *Trichoderma longibrachiatum* (NREF NF00756647). Esta análise indicou que as duas proteínas foram 44% idênticas ao nível da proteína em um trecho de 113 aminoácidos (339 pares de base). Depois desta identificação inicial o clone Tter08C4 foi recuperado da placa de estoque congelada original e riscado sobre uma placa LB suplementada com 50 µg de canamicina por ml. A placa foi incubada durante a noite a 37°C e no dia contíguo uma colônia única da placa foi usada para inocular 3 ml de LB suplementado com 50 µg de canamicina por ml. A cultura líquida foi incubada durante a noite a 37°C e o DNA plasmídico foi preparado com um BioRobot 9600 (QIAGEN, Inc., Valencia, CA). O DNA plasmídico do clone Tter08C4 foi seqüenciado mais uma vez com a química do terminador Big-Dye® como descrita acima, usando o M13 avançado e um iniciador Poli-T mostrado abaixo para seqüenciar a extremidade 3' do clone.

5'-TTTTTTTTTTTTTTTTTTTTTTTTTTTVN-3' (SEQ ID NO: 6)

onde V = G, A, C e N = G, A, C, T

A análise de homologia Blastx da informação de seqüência indicou que a proteína codificada pelo clone Tter08C4 foi similar à proteína EG1 de *Trichoderma reesei* (NREF NF00494331). Estas proteínas foram 46% idênticas em um trecho de 365 aminoácidos.

A análise da seqüência de proteína deduzida do clone Tter08C4 com o programa Interproscan (Zdobnov e Apweiler, 2001, Bioinformatics 17: 847-8) mostrou que o gene contido na assinatura de seqüência das proteínas da Família 7. Esta assinatura de seqüência conhecida como o padrão Pfam PF00840 (Bateman *et al.*, 2002, Nucleic Acids Research

30: 276-280) foram encontrados 18 aminoácidos do aminoácido de partida metionina confirmando que o clone Tter08C4 codifica uma endoglicanase da Família 7.

A sequência de cDNA (SEQ ID NO: 1) e a sequência de aminoácido deduzida (SEQ ID NO: 2) da endoglicanase de *Thielavia terrestris* são mostradas nas Figuras 1A e 1B. O clone de cDNA codifica um polipeptídeo de 336 aminoácidos. O teor de %G+C do clone de cDNA do gene é 67,5% e da região codificadora da proteína madura (nucleotídeos de 55 a 1011 da SEQ ID NO: 1) é 67,5%. Usando o programa de software SignalP (Nielsen *et al.*, 1997, Protein Engineering 10: 1-6), um peptídeo de sinal de 17 resíduos foi prognosticado. A proteína madura prognosticada contém 319 aminoácidos com uma massa molecular de 33,3 kDa.

Um alinhamento comparativo de sequências da Família 7 de endoglicanase foi determinada pelo método Clustal W (Higgins, 1989, *supra*) usando o software LASERGENE® MEGALIGN® (DNASTAR, Inc., Madison, WI) com uma tabela de identidade e os seguintes parâmetros de alinhamento múltiplo: Penalidade de intervalo de 10 e penalidade de comprimento de intervalo de 10. Parâmetros de alinhamento aos pares foram Ktuple = 1, penalidade de intervalo = 3, janela = 5 e diagonais = 5. O alinhamento mostrou que a sequência de aminoácido deduzida do gene CEL7F de *Thielavia terrestris* maduro compartilha 46,7% de identidade com a sequência de aminoácido deduzida da região catalítica do gene da endoglicanase I de *Trichoderma reesei* (NREF NF00494331, Uniprot Q5BMS5) e 44,9% de identidade com a sequência de aminoácido deduzida do gene da endoglicanase I de *Trichoderma reesei* de tamanho natural (NREF NF00494331, Uniprot Q5BMS5). A análise do alinhamento das regiões catalíticas destas proteínas mostrou que a CEL7F endoglicanase de *Thielavia terrestris* perdeu pelo menos três motivos de sequência separados que são conservados em todos os outros membros conhecidos da Família 7 de glicosil

hidrolase a partir de fungos. Dois destes consistem de mais do que dez resíduos de aminoácido e contêm resíduos altamente conservados que não estão presentes na CEL7F endoglicanase de *Thielavia terrestris*.

Uma vez que a identidade do clone Tter08C4 foi confirmada, uma alíquota de 0,5 µl de DNA plasmídico deste clone, designado pTter7F (Figura 2), foi transferido em um frasco de células TOP10 da *E. coli* (Invitrogen, Carlsbad, CA), suavemente misturadas e incubadas em gelo por 10 minutos. As células foram depois submetidas a choque térmico a 42°C por 30 segundos e incubadas mais uma vez em gelo por 2 minutos. As células foram recolocadas em suspensão em 250 µl de meio SOC e incubadas a 37°C por 60 minutos com agitação constante (200 rpm). Depois do período de incubação, duas alíquotas de 30 µl foram plaqueadas sobre placas de LB suplementadas com 50 µg de canamicina por ml e incubadas durante a noite a 37°C. No dia contíguo uma colônia única foi escolhida e riscada sobre três criofrascos de 1,8 ml contendo cerca de 1,5 ml de LB agarose suplementada com 50 µg de canamicina por ml. Os frascos foram selados com PetriSeal® (Diversified Biotech, Boston MA) e depositados com o Agricultural Research Service Patent Culture Collection, Northern Regional Research Center, 1815 University Street, Peoria, Illinois, 61604, como NRRL B-30802, com uma data de depósito de 11 de abril de 2005.

Exemplo 5: Construção de vetor de expressão pAlLo2

O vetor de expressão pAlLo1 foi construído pela modificação de pBANe6 (Patente U.S. Nº 6.461.837), que compreende um híbrido dos promotores a partir do genes para a alfa-amilase neutra de *Aspergillus niger* e triose fosfato isomerase *Aspergillus oryzae* (promotor NA2-tpi), sequência terminadora da amiloglicosidase de *Aspergillus niger* (terminador AMG) e gene da acetamidase de *Aspergillus nidulans* (amdS). Todas as etapas de mutagênese foram verificadas pelo seqüenciamento usando a química do terminador Big-Dye® como descrita. A modificação de pBANe6 foi realizada

eliminando-se primeiro três sítios de restrição Nco I nas posições de 2051, 2722 e 3397 pares de base do marcador de seleção *amdS* pela mutagênese direcionada ao sítio. Todas as mudanças foram planejadas para serem “silenciosas” deixando a seqüência de proteína real do produto de gene *amdS* inalterada. A remoção destes três sítios foi realizada simultaneamente com um Kit de Mutagênese Direcionada ao Sítio *in vitro* GeneEditor® (Promega, Madison, WI) de acordo com as instruções do fabricante usando os seguintes iniciadores (o nucleotídeo sublinhado representa a base mudada):

AMDS3NcoMut (2050): 5'-GTGCCCCATGATACGCCTCCGG-3' (SEQ ID NO: 7)

AMDS2NcoMut (2721): 5'-GAGTCGTATTTCCAAGGCTCCTGACC-3' (SEQ ID NO: 8)

AMDS1NcoMut (3396): 5'-GGAGGCCATGAAGTGGACCAACGG-3' (SEQ ID NO: 9)

Um plasmídeo que compreende todas as três mudanças de seqüência esperadas foi depois submetido à mutagênese direcionada ao sítio, usando um Kit de Mutagênese Direcionada ao Sítio QuickChange® (Stratagene, La Jolla, CA), para eliminar o sítio de restrição Nco I na extremidade do terminador AMG na posição 1643. Os seguintes iniciadores (o nucleotídeo sublinhado representa a base mudada) foram usados para a mutagênese:

Iniciador superior para mutagenizar a seqüência terminadora AMG: 5'-CACCGTGAAAGCCATGCTCTTTCCTTCGTGTAGAAGACCAGACAG-3' (SEQ ID NO: 10)

Iniciador inferior para mutagenizar a seqüência terminadora AMG: 5'-CTGGTCTTCTACACGAAGGAAAGAGCATGGCTTTCACGGTGTCTG-3' (SEQ ID NO: 11)

A última etapa na modificação de pBANE6 foi a adição de um novo sítio de restrição Nco I no começo do poliligador usando um Kit de

Mutagênese Direcionada ao Sítio QuickChange® e os seguintes iniciadores (os nucleotídeos sublinhados representam as bases mudadas) para produzir pAlLo1 (Figura 3).

Iniciador Superior para mutagenizar o promotor NA2-tpi: 5'-
 5 CTATATACACA ACTGGATTTACCATGGGCCCCGCGGCCGCAGATC-3'
 (SEQ ID NO: 12)

Iniciador Inferior para mutagenizar o promotor NA2-tpi: 5'-
 GATCTGCGGCCGCGGGCCCATGGTAAATCCAGTTGTGTATATAG-3'
 (SEQ ID NO: 13)

10 O gene amdS de pAlLo1 foi trocado com o gene pirG de *Aspergillus nidulans*. O plasmídeo pBANE10 (Figura 4) foi usado como uma fonte para o gene pirG como um marcador de seleção. A análise da sequência de pBANE10 mostrou que o marcador pirG foi contido dentro de um fragmento de restrição Nsi I e não contém os sítios de restrição Nco I ou Pac
 15 I. Visto que o amdS também é flanqueado pelos sítios de restrição Nsi I a estratégia para comutar o marcador de seleção foi uma simples troca de fragmentos de restrição Nsi I. O DNA plasmídico de pAlLo1 e pBANE10 foi digerido com a enzima de restrição Nsi I e os produtos purificados pela eletroforese em gel de agarose. O fragmento Nsi I de pBANE10 contendo o
 20 gene pirG foi ligado à cadeia principal de pAlLo1 para substituir o fragmento de DNA de Nsi I original contendo o gene amdS. Os clones recombinantes foram analisados pela digestão de restrição para determinar que eles tiveram o inserto correto e também a sua orientação. Um clone com o gene pirG transcrito na direção anti-horária foi selecionado. O novo plasmídeo foi
 25 designado pAlLo2 (Figura 5).

Exemplo 6: Clonagem do gene da endoglicanase da Família CEL7F em um vetor de expressão de *Aspergillus oryzae*

Dois iniciadores de oligonucleotídeo sintético, mostrados abaixo, foram planejados para amplificar pela PCR a matriz de leitura aberta

de tamanho natural de EST Tter08C4 de *Thielavia terrestris* que codifica uma endoglicanase da Família CEL7F. Um Kit de Clonagem na Fusão (BD Biosciences, Palo Alto, CA) foi usado para clonar o fragmento diretamente em pAlLo2.

5 Iniciador Avançado Em Fusão:

5'-**ACTGGATTACCATGACCCTACGGCTCCCTGTCATCA**-3' (SEQ ID NO: 14)

Iniciador Reverso Em Fusão:

10 5'-**TCACCTCTAGTTAATTAAGTAGTTCTTCGTGGTAGACC**-3' (SEQ ID NO: 15)

As letras em negrito representam a sequência codificadora. A sequência remanescente contém a identidade de sequência comparada com os sítios de inserção de pAlLo2.

15 Cinquenta picomoles de cada um dos iniciadores acima foram usados em uma reação de PCR contendo 50 ng de pTter11C9 DNA, 1X tampão de Amplificação Pfx (Invitrogen, Carlsbad, CA), 6 µl de 10 mM de mistura of dATP, dTTP, dGTP e dCTP, 2,5 unidades de Pfx DNA Polimerase Platinum (Invitrogen, Carlsbad, CA), 1 µl de MgSO₄ 50 mM e 5 µl de 10X
20 solução realçadora pCRx (Invitrogen, Carlsbad, CA) em um volume final de 50 µl. Um Eppendorf Mastercycler 5333 (Eppendorf Scientific, Inc., Westbury, NY) foi usado para amplificar o fragmento programado para um ciclo a 98°C por 2 minutos; e 35 ciclos cada a 94°C por 30 segundos, 65°C por 30 segundos e 68°C por 1,5 minuto. Depois dos 35 ciclos, a reação foi incubada a 68°C por 10 minutos e depois esfriada a 10°C até processada
25 adicionalmente. Um produto de reação de PCR de 1,4 kb foi isolado em um gel de agarose GTG a 0,8% (Cambrex Bioproducts One Meadowlands Plaza East Rutherford, New Jersey 07073) usando 40 mM de tampão base de Tris-acetato de sódio 20 mM-EDTA dissódico 1 mM (TAE) e 0,1 µg de brometo de etídio por ml. A faixa de DNA foi visualizada com a ajuda de um Dark

Reader[®] (Clare Chemical Research, Dolores, CO) para evitar mutações induzidas por UV. A faixa de DNA de 1,4 kb foi excisada com uma lâmina de barbear descartável e purificada com um copo giratório Ultrafree-DA (Millipore, Billerica, MA) de acordo com as instruções do fabricante.

5 O vetor pAlLo2 foi linearizado pela digestão com Nco I e Pac I. O fragmento foi purificado pela eletroforese em gel e ultrafiltração como descrito acima. A clonagem do fragmento de PCR purificado no vetor pAlLo2 linearizado e purificado foi realizado com um Kit de Clonagem In-Fusion (BD Biosciences, Palo Alto, CA). A reação (20 µl) conteve 1X Tampão In-
10 Fusion (BD Biosciences, Palo Alto, CA), 1X BSA (BD Biosciences, Palo Alto, CA), 1 µl de enzima In-Fusion (diluída 1:10) (BD Biosciences, Palo Alto, CA), 100 ng de pAlLo2 digerida com Nco I e Pac I e 50 ng do produto de PCR CEL7F purificado de *Thielavia terrestris*. A reação foi incubada na temperatura ambiente por 30 minutos. Uma amostra de 2 µl da reação foi
15 usada para transformar células XL10 SoloPac[®] Gold da *E. coli* (Stratagene, La Jolla, CA) de acordo com as instruções do fabricante. Depois do período de recuperação, duas alíquotas de 100 µl da reação de transformação foram plaqueadas em placas 2X YT de 150 mm suplementado com 100 µg de ampicilina por ml. As placas foram incubadas durante a noite a 37°C. Um
20 conjunto de oito clones putativos recombinantes foi selecionado aleatoriamente das placas de seleção e o DNA plasmídico foi preparado de cada um usando um BioRobot 9600. Os clones foram analisados pela digestão de restrição com Xho I. Dois clones que tiveram o padrão de digestão de restrição esperado foram depois seqüenciados para confirmar que não houve
25 mutações no inserto clonado. O Clone #1 foi selecionado e designado pAlLo22 (Figura 6).

Exemplo 7: Expressão do gene da endoglicanase da Família CEL7F *Thielavia terrestris* em *Aspergillus oryzae* JAL250

Protoplastos de *Aspergillus oryzae* Jal250 (WO 99/61651)

foram preparados de acordo com o método de Christensen *et al*, 1988, Bio/Technology 6: 1419-1422. Cinco microgramas de pAlLo22 (assim como pAlLo2 como um controle de vetor) foram usados para transformar protoplastos de *Aspergillus oryzae* JAL250.

5 A transformação de *Aspergillus oryzae* Ja1250 com pAlLo2 produziu cerca de 50 transformantes. Oito transformantes foram isolados em placas de PDA individuais e incubados por cinco dias a 34°C.

10 As placas de esporo confluentes foram lavadas com 5 ml de 0,01% de Tween 80 e a suspensão de esporo foi usada para inocular 25 ml de meio MDU2BP em 125 ml frascos de agitação de vidro. As culturas de transformante foram incubadas a 34°C com agitação constante a 200 rpm. No quinto dia após a inoculação, as culturas foram centrifugadas a 6000 x g e seus sobrenadantes coletados. Cinco µl de cada sobrenadante foram misturados com um volume igual de 2X tampão de carga (10% de R-
15 mercaptoetanol) e carregados em um gel de SDS-PAGE de 1,5 mm de 8% a 16% de Tris-Glicina e tingido com Simply Blue SafeStain (Invitrogen, Carlsbad, CA). Os perfis de SDS-PAGE dos caldos de cultura mostraram que seis dos oito transformantes tiveram uma nova faixa de proteína de aproximadamente 40 kDa. O transformante número 7 foi selecionado para
20 estudos adicionais e designado Ja1250AlLo22 de *Aspergillus oryzae*.

Exemplo 8: Culturas em frasco agitado grandes de *Aspergillus oryzae* Ja1250AlLo22

25 Esporos de *Aspergillus oryzae* Ja1250AlLo22 foram espalhados sobre uma placa de PDA e incubadas por cinco dias a 34°C. A placa de esporo confluyente foi lavada duas vezes com 5 ml de 0,01% de Tween 80 para maximizar o número de esporos coletados. A suspensão de esporo foi depois usada para inocular 500 ml de meio MDU2BP em um frasco de Fernbach de dois litros. A cultura de transformante foi incubada a 34°C com agitação constante (200 rpm). No quinto dia após a inoculação, o

caldo de cultura foi coletado pela filtração em uma unidade de filtração de Náilon de 75 mm, 500 ml com um tamanho de poro de 0,45 µm com um pré filtro de fibra de vidro. Uma amostra de 5 µl do caldo foi analisada pela SDS-PAGE como descrito acima para confirmar que o padrão de proteína foi o mesmo como aquele obtido antes. Uma vez que o caldo mostrou conter a faixa de proteína de 40 kDa, o caldo foi submetido à caracterização enzimática.

Exemplo 9: Caracterização da CEL7F endoglicanase de *Thielavia terrestris*

O caldo de Ja1250AlLo22 de *Aspergillus oryzae* descrito no Exemplo 8 foi filtrado através de um filtro de 0,22 µm de tamanho de poro (Millipore, Billerica, MA), concentrado usando uma célula agitada Amicon equipada com uma membrana PM10 (Millipore, Billerica, MA) e dessalinizada usando uma coluna Econo-Pac 10DG (BioRad Laboratories, Hercules, CA).

O caldo de *Aspergillus oryzae* Ja1250 (apenas o vetor), como um controle negativo, foi tratado do mesmo modo como acima.

Os substratos tingidos usados para avaliar a especificidade de substrato da CEL7F endoglicanase de *Thielavia terrestris* incluíram: AZCL-arabinoxilano (trigo), AZCL-β-glicano, AZCL-dextrano, AZCL-HE-celulose, AZCL-galactana da batata, AZCL-galactomanana (Alfarroba), AZCL-xilano (vidoeiro), AZCL-xiloglicano (Megazyme, Bray, Irlanda) e Quitina Azure (Sigma, St Louis, MO).

Os ensaios de atividade foram realizados em placas de 96 reservatórios profundos (Axygen Scientific, Union City, CA) seladas por um selador de placa (ALPS-300, Abgene, Epsom, UK). Oitocentos µl dos substratos acima (6,25 g por litro de acetato de sódio 50 mM pH 5,0) foram transferidos em cada reservatório da placa de 96 reservatórios profundos, seguido por 180 µl de acetato de sódio 50 mM pH 5,0 e 20 µl de solução de

CEL7F endoglicanase *Thielavia terrestris* (0,25 g/L) para iniciar as reações. A concentração de substrato e a carga de enzima na mistura de reação final foram 5 g por litro e 1 mg de enzima por g de substrato, respectivamente. O caldo de J1250 de *Aspergillus oryzae* foi testado junto com o caldo de CEL7

5 endoglicanase de *Thielavia terrestris* sob as mesmas condições, servindo como um controle negativo. As reações foram incubadas a 50°C sem mistura. Antes da amostragem, as placas de reservatório profundo foram centrifugadas em uma centrífuga de placa (Sorvall RT7, Global Medical Instrumentation, Ramsey, MN) a 3000 rpm por 5 minutos. Uma amostra de 150 µl do

10 sobrenadante foi transferida em uma placa de filtração de 96 reservatórios (tamanho de poro de 0,45 µm, Millipore, Billerica, MA), submetida à vácuo e o filtrado foi coletado. Uma amostra de 100 µl do filtrado foi transferida em uma outra placa de 96 reservatórios e a absorbância a 590 nm foi medida usando um Spectra MAX340 (Molecular Devices, Sunnyvale, CA).

15 Depois de incubações em 1 hora e 92 horas, o corante liberado dos substratos tingidos diferentes pela CEL7F endoglicanase de *Thielavia terrestris* (depois de subtrair o corante liberado pelo J1250 de *Aspergillus oryzae*) é mostrado na Tabela 1 como valores de 590 nm relativos.

Tabela 1. $A_{590 \text{ nm}}$ relativa depois de incubar endoglicanase de *Thielavia terrestris* com substratos tingidos diferentes por 1 hora e 92 horas a 50°C, pH 5,0.

Tempo em horas	AX	βG	Dex	HEC	Gal	GM	Xly	XG	Quitina Azure
1	0,00	0,22	0,00	0,16	0,00	0,00	0,01	0,01	0,00
92	0,58	0,78	0,00	1,00	0,00	0,08	0,32	0,75	0,00

AX: AZCL-arabinosilano

βG: AZCL-β-glicano

Dex: AZCL-dextrano

HEC: AZCL-HE-celulose

Gal: AZCL-galactana de batata

GM: AZCL-galactomanana

Xly: AZCL-xilano

XG: AZCL-xiloglicano

A CEL7F endoglicanase de *Thielavia terrestris* possuiu atividade contra AZCL-β-glicano e AZCL-HE-celulose depois de 1 hora. A CEL7F endoglicanase de *Thielavia terrestris* também mostrou atividade sobre

substratos tingidos com arabinoxilano, xilano e xiloglicano e baixa atividade sobre o substrato tingido com galactomanana depois de uma incubação de 92 horas.

Exemplo 10: Hidrólise de forragem de milho pré tratada com a CEL7F endoglicanase de *Thielavia terrestris*

5 A forragem de milho foi pré tratada no U.S. Department of Energy National Renewable Energy Laboratory (NREL) usando ácido sulfúrico diluído. As seguintes condições foram usadas para o pré tratamento: 0,048 g de ácido sulfúrico / g de biomassa seca a 190°C e 25% p/p de sólidos
10 secos por volta de 1 minuto. Os sólidos insolúveis em água na forragem de milho pré tratada (PCS) conteve 52% de celulose, 3,6% de hemi-celulose e 29,8% de lignina. Celulose e hemi-celulose foram determinadas por uma hidrólise de ácido sulfúrico de dois estágios com análise subsequente de açúcares pela cromatografia líquida de alto desempenho usando o
15 Procedimento Analítico Padrão NREL #002. A lignina foi determinada gravimetricamente depois de hidrolisar as frações de celulose e hemi-celulose com ácido sulfúrico usando o Procedimento Analítico Padrão NREL #003. Antes da hidrólise enzimática, o PCS foi lavado com um grande volume de água duplamente destilada até que o pH fosse mais alto do que 4,0 e foi
20 depois peneirado através de uma peneira de malha 100 e autoclavado a 121°C por 30 minutos.

A hidrólise de PCS foi conduzida em placas de 96 reservatórios profundos, (Axygen Scientific, Union City, CA) selado por um selador de placa (ALPS-300, Abgene, Epsom, UK), com um volume de
25 reação total de 1,0 ml. A hidrólise de PCS (10 mg/ml em tampão de acetato de sódio 50 mM pH 5,0) foi realizado usando 1,25 mg de CEL7F endoglicanase de *Thielavia terrestris* (preparada como descrita no Exemplo 9), por grama de PCS. O caldo de Ja1250 de *Aspergillus oryzae* (preparado como descrito no Exemplo 9) foi conduzido como um controle. A hidrólise de PCS

foi realizada a 50°C, pH 5,0. As reações foram conduzidas em duplicatas e as alíquotas foram tiradas durante o curso da hidrólise. As reações de hidrólise de PCS foram interrompidas misturando-se uma alíquota de 20 µl de cada hidrolisado com 180 µl de NaOH 0,11 M (reagente de interrupção). As diluições em série apropriadas foram geradas para cada amostra e o teor de açúcar redutor determinado usando um ensaio da hidrazida do ácido para-hidroxibenzóico (PHBAH, Sigma, St. Louis, MO) adaptado para um formato de microplaca de 96 reservatórios como descrito abaixo. Em resumo, uma alíquota de 90 µl de uma amostra apropriadamente diluída foi colocada em uma microplaca de fundo cônico de 96 reservatórios. As reações foram iniciadas pela adição de 60 µl de 1,5% (p/v) de PHBAH em NaOH a 2% para cada reservatório. As placas foram aquecidas descobertas a 95°C por 10 minutos. As placas foram deixadas esfriar até a temperatura ambiente (RT) e 50 µl de H₂O destilada adicionados a cada reservatório. Uma alíquota de 100 µl de cada reservatório foi transferida a uma placa de 96 reservatórios de fundo chato e a absorbância a A_{410 nm} medida usando uma SpectraMax Microplate Reader (Molecular Devices, Sunnyvale, CA). Os padrões de glicose (0,1 a 0,0125 mg/ml diluídos com 0,4% de hidróxido de sódio) foram usados para preparar uma curva padrão para traduzir os valores de A_{410nm} obtidos em equivalentes de glicose. Os equivalentes resultantes foram usados para calcular a porcentagem de conversão de PCS celulose para cada reação. O grau de conversão da celulose para açúcar redutor (conversão,%) foi calculado usando a seguinte equação:

$$\begin{aligned}\text{Conversão (\%)} &= \text{RS}_{(\text{mg/ml})} * 100 * 162 / (\text{Celulose}_{(\text{mg/ml})} * 180) \\ &= \text{RS}_{(\text{mg/ml})} * 100 / (\text{Celulose}_{(\text{mg/ml})} * 1,111)\end{aligned}$$

Nesta equação, RS é a concentração de açúcar redutor em solução medida em equivalentes de glicose (mg/ml) e o fator 1,111 reflete o peso ganho na conversão da celulose para glicose.

A hidrólise de PCS pela CEL7F endoglicanase de *Thielavia*

terrestris (1,25 mg/g de PCS) produziu uma conversão de celulose de 2,1% depois de 120 horas. Jal250 *Aspergillus oryzae* (1,25 mg/g de PCS) produziu menos do que 1% de conversão depois de 120 horas.

5 **Exemplo 11: Hidrólise de beta-glicano solúvel de cevada pela endoglicanase de *Thielavia terrestris***

10 A Cel7F endoglicanase de *Thielavia terrestris* foi testada na forma do caldo de Jal250A1Lo22 de *Aspergillus oryzae* descrita no Exemplo 8. O caldo foi concentrado e trocado para acetato de sódio 50 mM pH 5,0 usando o filtro centrífugo Centricon Plus-20 com membrana de poliétersulfona Biomax-5 (5000 NMWL) da Millipore (Bedford, MA). O
caldo de Jal250 de *Aspergillus oryzae* (vetor sozinho) foi tratado da mesma maneira como acima.

15 A concentração de proteína nas soluções de enzima foi determinada usando o ensaio de microplaca do Ácido Bicinconínico (BCA) de acordo com as instruções do fabricante para um Kit de Reagente de Ensaio de Proteína de BCA (Pierce Chemical Co., Rockford, IL).

As diluições de enzima foram preparadas frescas antes de cada experimento a partir das soluções de enzima de estoque, que foram armazenadas a -20°C.

20 A atividade da Cel7F endoglicanase de *Thielavia terrestris* sobre beta-glicano solúvel da cevada (viscosidade do meio, 230 kDa, Megazyme International Ireland Ltd., Bray, Irlanda) foi determinada no pH 5,5 (acetato de sódio 50 mM com 0,02% de azida de sódio) e 60°C. Os resultados foram comparados com aqueles para Cel7B (EGI) endoglicanase
25 de *Trichoderma reesei*. A Cel7B (EGI) endoglicanase de *Trichoderma reesei* recombinante pode ser preparada de acordo com Takashima *et al*, 1998, Journal of Biotechnology 65: 163-171.

A concentração inicial de beta-glicano nas reações de hidrólise foi de 1,0% (p/v). Reações de um ml foram conduzidas sem agitar em Placas

de 96 reservatórios profundos de Eppendorf (1,2 ml, VWR Scientific, West Chester, PA). As enzimas foram usadas em três cargas de proteína, 0,05, 0,1 e 0,2 mg por g de glicano. Nas reações de controle, as endoglicanases foram substituídas com 50 mM de acetato de sódio pH 5,5 contendo 0,02% de azida de sódio (controle de tampão) ou com caldo de Jal250 de *Aspergillus oryzae* concentrado e trocado no tampão não contendo nenhuma enzima recombinantemente expressada (controle de Jal250).

As alíquotas foram removidas das reações de hidrólise em 2 horas e 24 horas, diluídas com água deionizada e analisadas quanto aos açúcares redutores usando o ensaio da hidrazida do ácido p-hidróxi-benzóico (PHBAH) como descrito no Exemplo 10. A conversão relativa de beta-glicano como uma função da carga de proteína em dois tempos de incubação, 2 horas e 24 horas, é mostrada nas Figuras 7 e 8, respectivamente. A conversão relativa é mostrada como uma porcentagem de conversão obtida depois da hidrólise de 24 horas de beta-glicano pela Cel7F endoglicanase de *Thielavia terrestris* (0,2 mg de proteína por g de glicano).

A Cel7F endoglicanase de *Thielavia terrestris* mostrou conversão mais alta de beta-glicano do que a Cel7B endoglicanase de *Trichoderma reesei* e continuou a produzir novos grupos finais redutores além do tempo de 2 horas de incubação. Ao contrário, a Cel7B endoglicanase de *Trichoderma reesei* quase não mostrou nenhum aumento adicional na concentração de açúcar redutor depois de 2 horas de hidrólise.

Depósito de Material Biológico

O seguinte material biológico foi depositado sob os termos do Tratado de Budapest com a Agricultural Research Service Patent Culture Collection, Northern Regional Research Center, 1815 University Street, Peoria, Illinois, 61604 e dado o seguinte número de acesso:

Depósito	Número de Acesso	Data de Depósito
pTter7F de <i>E. coli</i>	NRRL B-30837	11 de abril de 2005

A cepa foi depositada sob as condições que garante que o acesso à cultura estará disponível durante a pendência deste pedido de patente a uma pessoa determinada pelo Representante de Patentes e Marcas a ser designado para isto sob 37 C.F.R. §1,14 e 35 U.S.C. §122. O depósito
5 representa uma cultura substancialmente pura da cepa depositada. O depósito está disponível como requerido pelas leis de patente estrangeiras em países em que as contrapartes do pedido objeto ou a sua progênie são depositados. Entretanto, deve ser entendido que a disponibilidade de um depósito não constitui uma licença para praticar a invenção objeto em detrimento dos
10 direitos de patente outorgados pela ação governamental.

A invenção descrita e aqui reivindicada não deve ser limitada no escopo pelos aspectos específicos aqui divulgados, visto que estes aspectos são pretendidos como ilustrações de diversos aspectos da invenção. Pretende-se que quaisquer aspectos equivalentes estejam dentro do escopo desta
15 invenção. De fato, várias modificações da invenção além daquelas mostradas e descritas aqui tornar-se-ão evidentes àqueles habilitados na técnica a partir da descrição precedente. Pretende-se também que tais modificações caiam dentro do escopo das reivindicações anexas. No caso de conflito, a presente divulgação incluindo as definições controlarão.

20 Várias referências são aqui citadas, as divulgações das quais são incorporadas por referência em suas totalidades.

Número de referência do pedido do agente ou requerente 10802.204-WO	Pedido Internacional no. A ser designado
---	---

INDICAÇÕES RELACIONADAS A UM MICROORGANISMO DEPOSITADO

(Regra PCT 13bis)

A. As indicações feitas abaixo se referem ao microorganismo referido na descrição na página 61, linha 18.	
B. IDENTIFICAÇÃO DE	Outros depósitos estão identificados em uma folha adicional <input type="checkbox"/>
Nome da instituição depositária: Agricultural Research Service Patent Culture Collection (NRRL)	
Endereço da instituição depositária (<i>incluindo código postal e país</i>) Northern Regional Research Center 1815 University Street Peoria, IL 61604 US	
Data do depósito 11 de abril de 2005	Número de Acesso B-30837
C. INDICAÇÕES ADICIONAIS (<i>deixar em branco se não aplicável</i>)	<i>Esta informação continua em uma folha adicional</i> <input type="checkbox"/>
Com relação às designações em que uma patente europeia e/ou australiana é procurada, durante a pendência do pedido de patente, uma amostra do microorganismo depositado deve somente ser fornecida a um perito independente nomeado pela pessoa solicitando a amostra (Regra 28(4) EPC/Regulation 3.25 da Regra Estatutária Australiana 1991 no. 71)	
D. ESTADOS DESIGNADOS PARA OS QUAIS FORAM FEITAS INDICAÇÕES (<i>se as indicações não são para todos os Estados designados</i>)	
E. FORNECIMENTO SEPARADO DE INFORMAÇÕES (<i>deixar em branco se não aplicável</i>)	
A indicação relacionada abaixo será submetida ao Escritório Internacional posteriormente (especificar a natureza geral das indicações, por exemplo, "Número de Acesso do Depósito")	

Somente para uso do Escritório receptor <input type="checkbox"/> Esta folha foi recebida com o pedido internacional	Somente para uso do Escritório Internacional <input type="checkbox"/> Esta folha foi recebida pelo Escritório internacional
Funcionário autorizado:	Funcionário autorizado:

LISTAGEM DAS SEQUÊNCIAS

<110> Novozymes, Inc.

<120> **Poli-peptídeos tendo atividade de endoglicanase e polinucleotídeos que codificam os mesmos.**

<130> 10802.204-WO

<150> 60/675,601

<151> 2005-04-27

<160> 15

<170> PatentIn version 3.2

<210> 1

<211> 1011

<212> DNA

<213> Thielavia terrestris

<400> 1

```

atgaccctac ggctccctgt catcagcctg ctggcctcgc tggcagcagg cgcctgctgc      60
gtcccacggg cggagtttca ccccccctct cgcacttgga aatgcacgac ctccggggggc      120
tgctgacagc agaacaccag cgtcgtcctg gaccgtgact cgaagtacgc cgcacacagc      180
gccggctcgc ggacggaatc ggattacgag gcaatgggag tgtccacttc gggcaatgcc      240
gtgacgtgtt accactacgt caagaccaac ggcacctcgc tcccgccttc gccgcgcctc      300
taoctcctgg gcgcggacgg caagtacgtg cttatggacc tcctcaacca ggagctgtcg      360
gtggacgtcg acttctcggc gctgccgtgc ggcgagaacg gggccttcta cctgtccgag      420
atggcggcgg acgggcgggg cgacgcgggg gcggggcgac ggtactgoga cgcgcagtcg      480
cagggctact gctgcaaoga gatggacatc ctcgaggcca actogatggc gacggccatg      540
acgccgcacc cgtgcaaggg caacaactgc gaccgcagcg gctgcggcta caaccogtac      600
gccagcggcc agcgcggctt ctacggggcc ggcaagacgg tcgacacgag caagcccttc      660
accgtctga cgcagttcgc cgcagcgggc ggcaagctga ccagatcac ccgcaagtac      720
atccagaacg gccgggagat cggcggcggc ggcaccatct ccagctgagg ctccgagttt      780
tcgacggggc gcctgaccgg catggggcag gcctgggggc gcggaatggg gctggccatg      840
agcatotgga acgacgcggc ccaggagatg gcatggctcg atgccggcaa caacggccct      900
tgccccagtg gccagggcag cccgtccgtc attcagtcgc agcatccga caccacgtc      960
gtcttctcca acatcaggtg gggcgacatc ggggtctacca cgaagaacta g      1011

```

<210> 2

<211> 336

<212> PRT

<213> Thielavia terrestris

<400> 2

Met Thr Leu Arg Leu Pro Val Ile Ser Leu Leu Ala Ser Leu Ala Ala
 1 5 10 15

Gly Ala Val Val Val Pro Arg Ala Glu Phe His Pro Pro Leu Pro Thr
 20 25 30

Trp Lys Cys Thr Thr Ser Gly Gly Cys Val Gln Gln Asn Thr Ser Val
 35 40 45

Val Leu Asp Arg Asp Ser Lys Tyr Ala Ala His Ser Ala Gly Ser Arg
 50 55 60

Thr Glu Ser Asp Tyr Ala Ala Met Gly Val Ser Thr Ser Gly Asn Ala
 65 70 75 80

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

Ser Pro Arg Ile Tyr Leu Leu Gly Ala Asp Gly Lys Tyr Val Leu Met
 100 105 110

Asp Leu Leu Asn Gln Glu Leu Ser Val Asp Val Asp Phe Ser Ala Leu
 115 120 125

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

Gly Arg Gly Asp Ala Gly Ala Gly Asp Gly Tyr Cys Asp Ala Gln Cys
 145 150 155 160

Gln Gly Tyr Cys Cys Asn Glu Met Asp Ile Leu Glu Ala Asn Ser Met
 165 170 175

Ala Thr Ala Met Thr Pro His Pro Cys Lys Gly Asn Asn Cys Asp Arg
 180 185 190

Ser Gly Cys Gly Tyr Asn Pro Tyr Ala Ser Gly Gln Arg Gly Phe Tyr
 195 200 205

Gly Pro Gly Lys Thr Val Asp Thr Ser Lys Pro Phe Thr Val Val Thr
 210 215 220

Gln Phe Ala Ala Ser Gly Gly Lys Leu Thr Gln Ile Thr Arg Lys Tyr
 225 230 235 240

Ile Gln Asn Gly Arg Glu Ile Gly Gly Gly Thr Ile Ser Ser Cys
 245 250 255

Gly Ser Glu Ser Ser Thr Gly Gly Leu Thr Gly Met Gly Glu Ala Leu
 260 265 270

Gly Arg Gly Met Val Leu Ala Met Ser Ile Trp Asn Asp Ala Ala Gln
 275 280 285

Glu Met Ala Trp Leu Asp Ala Gly Asn Asn Gly Pro Cys Ala Ser Gly
 290 295 300

Gln Gly Ser Pro Ser Val Ile Gln Ser Gln His Pro Asp Thr His Val
 305 310 315 320

Val Phe Ser Asn Ile Arg Trp Gly Asp Ile Gly Ser Thr Thr Lys Asn
 325 330 335

<210> 3
 <211> 32
 <212> DNA
 <213> Escherichia coli

<400> 3
 tcgtcgggga caactttgta caaaaaagtt gg 32

<210> 4
 <211> 27
 <212> DNA
 <213> Escherichia coli

<400> 4
 cccctgttga aacatgtttt ttcaacc 27

<210> 5
 <211> 16
 <212> DNA
 <213> Escherichia coli

<400> 5
 gtaaaacgac ggccag 16

<210> 6
 <211> 25
 <212> DNA
 <213> Escherichia coli

<220>
 <221> misc_feature
 <222> (25)..(25)
 <223> N = A, C, G, OR T

<400> 6
 tttttttttt tttttttttt tttvn

25

<210> 7
 <211> 22
 <212> DNA
 <213> Aspergillus nidulans

<400> 7
 gtgccccatg ataagcctcc gg

22

<210> 8
 <211> 26
 <212> DNA
 <213> Aspergillus nidulans

<400> 8
 gagtcgtatt tccaaggctc ctgacc

26

<210> 9
 <211> 24
 <212> DNA
 <213> Aspergillus nidulans

<400> 9
 ggaggccatg aagtggacca acgg

24

<210> 10
 <211> 45
 <212> DNA
 <213> Aspergillus niger

<400> 10
 caccgtgaaa gccatgctct ttcttcgtg tagaagacca gacag

45

<210> 11
 <211> 45
 <212> DNA
 <213> Aspergillus niger

<400> 11
 ctggtcttct acacgaagga aagagcatgg ctttcacggt gtctg

45

<210> 12
 <211> 44
 <212> DNA
 <213> Aspergillus niger

<400> 12

ctatatacac aactggattt accatgggcc cgcggcgca gatc

44

<210> 13

<211> 44

<212> DNA

<213> *Aspergillus niger*

<400> 13

gatctgcggc cgcggggcca tggtaaatac agttgtgtat atag

44

<210> 14

<211> 36

<212> DNA

<213> *Thielavia terrestris*

<400> 14

actggattac catgacccta cggctccctg tcatca

36

<210> 15

<211> 38

<212> DNA

<213> *Thielavia terrestris*

<400> 15

tcacctctag ttaattaact agttcttcgt ggtagacc

38

REIVINDICAÇÕES

1. Polipeptídeo isolado, caracterizado pelo fato de que tem atividade de endoglicanase, selecionado do grupo que consiste de:

5 (a) um polipeptídeo que compreenda uma seqüência de aminoácido que tenha pelo menos 60 % de identidade com o polipeptídeo maduro da SEQ ID NO: 2;

10 (b) um polipeptídeo que é codificado por um polinucleotídeo que hibridiza sob pelo menos condições de estringência média com (i) o polipeptídeo maduro que codifica a seqüência da SEQ ID NO: 1, (ii) a seqüência de DNA genômico que compreende o polipeptídeo maduro que codifica a seqüência da SEQ ID NO: 1, ou (iii) um filamento complementar de (i) ou (ii); e

15 (c) uma variante que compreenda uma substituição, deleção e/ou inserção conservativa de um ou mais aminoácidos do polipeptídeo maduro da SEQ ID NO: 2.

2. Polipeptídeo de acordo com a reivindicação 1, caracterizado pelo fato de que compreende uma seqüência de aminoácido que tenha pelo menos 60 % de identidade com o polipeptídeo maduro da SEQ ID NO: 2.

20 3. Polipeptídeo de acordo com a reivindicação 2, caracterizado pelo fato de que compreende uma seqüência de aminoácido que tenha pelo menos 65 % de identidade com o polipeptídeo maduro da SEQ ID NO: 2.

4. Polipeptídeo de acordo com a reivindicação 3, caracterizado pelo fato de que compreende uma seqüência de aminoácido que tenha pelo menos 70 % de identidade com o polipeptídeo maduro da SEQ ID NO: 2.

25 5. Polipeptídeo de acordo com a reivindicação 4, caracterizado pelo fato de que compreende uma seqüência de aminoácido que tenha pelo menos 75 % de identidade com o polipeptídeo maduro da SEQ ID NO: 2.

6. Polipeptídeo de acordo com a reivindicação 5, caracterizado pelo fato de que compreende uma seqüência de aminoácido que tenha pelo

menos 80 % de identidade com o polipeptídeo maduro da SEQ ID NO: 2.

7. Polipeptídeo de acordo com a reivindicação 6, caracterizado pelo fato de que compreende uma seqüência de aminoácido que tenha pelo menos 85 % de identidade com o polipeptídeo maduro da SEQ ID NO: 2.

5 8. Polipeptídeo de acordo com a reivindicação 7, caracterizado pelo fato de que compreende uma seqüência de aminoácido que tenha pelo menos 90 % de identidade com o polipeptídeo maduro da SEQ ID NO: 2.

9. Polipeptídeo de acordo com a reivindicação 8, caracterizado pelo fato de que compreende uma seqüência de aminoácido que tenha pelo
10 menos 95 % de identidade com o polipeptídeo maduro da SEQ ID NO: 2.

10. Polipeptídeo de acordo com a reivindicação 1, caracterizado pelo fato de que compreende a seqüência de aminoácido da SEQ ID NO: 2, ou um fragmento desta tendo atividade de endoglicanase.

11. Polipeptídeo de acordo com a reivindicação 10, caracterizado pelo fato de que compreende a seqüência de aminoácido da
15 SEQ ID NO: 2.

12. Polipeptídeo de acordo com a reivindicação 11, caracterizado pelo fato de que compreende o polipeptídeo maduro da SEQ ID NO: 2.

13. Polipeptídeo de acordo com a reivindicação 1, caracterizado pelo fato de que consiste da SEQ ID NO: 2, ou um fragmento desta tendo atividade de endoglicanase.

14. Polipeptídeo de acordo com a reivindicação 13, caracterizado pelo fato de que consiste da SEQ ID NO: 2.

15. Polipeptídeo de acordo com a reivindicação 14, caracterizado pelo fato de que consiste do polipeptídeo maduro da SEQ ID NO: 2.

16. Polipeptídeo de acordo com a reivindicação 1, caracterizado pelo fato de que é codificado por um polinucleotídeo que

hibridiza sob pelo menos condições de estringência média com (i) o polipeptídeo maduro que codifica a sequência da SEQ ID NO: 1, (ii) a sequência de DNA genômico que compreende o polipeptídeo maduro que codifica a sequência da SEQ ID NO: 1, ou (iii) um filamento complementar de (i) ou (ii).

17. Polipeptídeo de acordo com a reivindicação 16, caracterizado pelo fato de que é codificado por um polinucleotídeo que hibridiza sob pelo menos condições de estringência média-alta com (i) o polipeptídeo maduro que codifica a sequência da SEQ ID NO: 1, (ii) a sequência de DNA genômico que compreende o polipeptídeo maduro que codifica a sequência da SEQ ID NO: 1, ou (iii) um filamento complementar de (i) ou (ii).

18. Polipeptídeo de acordo com a reivindicação 17, caracterizado pelo fato de que é codificado por um polinucleotídeo que hibridiza sob pelo menos condições de estringência alta com (i) o polipeptídeo maduro que codifica a sequência da SEQ ID NO: 1, (ii) a sequência de DNA genômico que compreende o polipeptídeo maduro que codifica a sequência da SEQ ID NO: 1, ou (iii) um filamento complementar de (i) ou (ii).

19. Polipeptídeo de acordo com a reivindicação 1, caracterizado pelo fato de que é codificado pelo polinucleotídeo contido no plasmídeo pTter7F que está contido na *E. coli* NRRL B-30837.

20. Polipeptídeo de acordo com a reivindicação 1, caracterizado pelo fato de que o polipeptídeo é uma variante que compreende uma substituição, deleção e/ou inserção conservativa de um ou mais aminoácidos dos aminoácidos 18 a 336 da SEQ ID NO: 2.

21. Polipeptídeo de acordo com qualquer uma das reivindicações de 1 a 20, caracterizado pelo fato de que o polipeptídeo maduro é os aminoácidos 18 a 336 da SEQ ID NO: 2.

22. Polipeptídeo de acordo com qualquer uma das reivindicações de 1 a 20, caracterizado pelo fato de que o polipeptídeo maduro que codifica a sequência tem os nucleotídeos de 52 a 1008 da SEQ ID NO: 1.

5 23. Polipeptídeo de acordo com qualquer uma das reivindicações de 1 a 22, caracterizado pelo fato de que ainda têm atividade enzimática contra um ou mais substratos selecionados do grupo que consiste de xilano, xiloglicano, arabinoxilano, 1,4-galactano, galactomanano, dextrano e quitina.

10 24. Polinucleotídeo isolado, caracterizado pelo fato de que compreende uma sequência de nucleotídeo que codifica o polipeptídeo como definido em qualquer uma das reivindicações de 1 a 23.

15 25. Polinucleotídeo isolado de acordo com a reivindicação 24, caracterizado pelo fato de que compreende pelo menos uma mutação no polipeptídeo maduro que codifica a sequência da SEQ ID NO: 1, em que a sequência de nucleotídeo mutante codifica o polipeptídeo maduro da SEQ ID NO: 2.

20 26. Construção de ácido nucleico, caracterizada pelo fato de que compreende o polinucleotídeo como definido nas reivindicações 24 ou 25 operavelmente ligado a uma ou mais sequências de controle que direcionam a produção do polipeptídeo em um hospedeiro de expressão.

25 27. Vetor de expressão recombinante, caracterizado pelo fato de que compreende a construção de ácido nucleico como definida na reivindicação 26.

28. Célula hospedeira recombinante, caracterizada pelo fato de que compreende a construção de ácido nucleico como definida na reivindicação 26.

29. Método para produzir o polipeptídeo como definido em qualquer uma das reivindicações de 1 a 23, caracterizado pelo fato de que

compreende: (a) cultivar uma célula, que na sua forma do tipo selvagem é capaz de produzir o polipeptídeo, sob condições condutivas para a produção do polipeptídeo; e (b) recuperar o polipeptídeo.

30. Método para produzir o polipeptídeo como definido em qualquer uma das reivindicações de 1 a 23, caracterizado pelo fato de que compreende: (a) cultivar uma célula hospedeira que compreenda uma construção de ácido nucleico que compreende uma seqüência de nucleotídeo que codifica o polipeptídeo sob condições condutivas para a produção do polipeptídeo; e (b) recuperar o polipeptídeo.

31. Método para produzir um mutante de uma célula precursora, caracterizado pelo fato de que compreende romper ou deletar uma seqüência de nucleotídeo que codifica o polipeptídeo como definido em qualquer uma das reivindicações de 1 a 23, que resulta na produção do mutante menos do polipeptídeo do que da célula precursora.

32. Célula mutante, caracterizada pelo fato de ser produzida pelo método como definido na reivindicação 31.

33. Célula mutante de acordo com a reivindicação 32, caracterizado pelo fato de que compreende ainda um gene que codifica uma proteína nativa ou heteróloga.

34. Método para produzir uma proteína, caracterizado pelo fato de que compreende: (a) cultivar a célula mutante como definida em reivindicação 33 sob condições condutivas para a produção da proteína; e (b) recuperar a proteína.

35. Polinucleotídeo isolado, caracterizado pelo fato de que é obtido pela (a) hibridização de uma população de DNA sob condições de estringência média com (i) o polipeptídeo maduro que codifica a seqüência da SEQ ID NO: 1, (ii) a seqüência de DNA genômico que compreende o polipeptídeo maduro que codifica a seqüência da SEQ ID NO: 1, ou (iii) um filamento complementar de (i) ou (ii); e (b) isolar o polinucleotídeo de

hibridização, que codifica um polipeptídeo tendo atividade de endoglicanase.

36. Polinucleotídeo isolado de acordo com a reivindicação 35, caracterizado pelo fato de que é obtido pela (a) hibridização de uma população de DNA sob condições de estringência média-alta com (i) o polipeptídeo maduro que codifica a sequência da SEQ ID NO: 1, (ii) a sequência de DNA genômico que compreende o polipeptídeo maduro que codifica a sequência da SEQ ID NO: 1, ou (iii) um filamento complementar de (i) ou (ii); e (b) isolar o polinucleotídeo de hibridização, que codifica um polipeptídeo tendo atividade de endoglicanase.

37. Polinucleotídeo isolado de acordo com a reivindicação 36, caracterizado pelo fato de que é obtido pela (a) hibridização de uma população de DNA sob condições de estringência alta com (i) o polipeptídeo maduro que codifica a sequência da SEQ ID NO: 1, (ii) a sequência de DNA genômico que compreende o polipeptídeo maduro que codifica a sequência da SEQ ID NO: 1, ou (iii) um filamento complementar de (i) ou (ii); e (b) isolar o polinucleotídeo de hibridização, que codifica um polipeptídeo tendo atividade de endoglicanase.

38. Polinucleotídeo isolado de acordo com qualquer uma das reivindicações de 35 a 37, caracterizado pelo fato de que o polipeptídeo maduro que codifica a sequência tem os nucleotídeos de 52 a 1008 da SEQ ID NO: 1.

39. Método para produzir um polinucleotídeo compreendendo uma sequência de nucleotídeo mutante, caracterizado pelo fato de que compreende: (a) introduzir pelo menos uma mutação no polipeptídeo maduro que codifica a sequência da SEQ ID NO: 1, em que a sequência de nucleotídeo mutante codifica um polipeptídeo que consiste do polipeptídeo maduro da SEQ ID NO: 2; e (b) recuperar o polinucleotídeo que compreende a sequência de nucleotídeo mutante.

40. Polinucleotídeo mutante, caracterizado pelo fato de que é

produzido pelo método como definido na reivindicação 39.

41. Método para produzir um polipeptídeo, caracterizado pelo fato de que compreende: (a) cultivar uma célula que compreende o polinucleotídeo mutante como definido na reivindicação 40 que codifica o polipeptídeo sob condições condutivas para a produção do polipeptídeo; e (b) recuperar o polipeptídeo.

42. Construção de ácido nucleico, caracterizada pelo fato de que compreende um gene que codifica uma proteína operavelmente ligada a uma sequência de nucleotídeo que codifica um peptídeo de sinal que compreende ou que consiste dos aminoácidos 1 a 17 da SEQ ID NO: 2, em que o gene é estranho à sequência de nucleotídeo.

43. Vetor de expressão recombinante, caracterizado pelo fato de que compreende a construção de ácido nucleico como definida na reivindicação 42.

44. Célula hospedeira recombinante, caracterizada pelo fato de que compreende a construção de ácido nucleico como definida na reivindicação 42.

45. Método para produzir uma proteína, caracterizado pelo fato de que compreende: (a) cultivar a célula hospedeira recombinante como definida na reivindicação 44 sob condições condutivas para a produção da proteína; e (b) recuperar a proteína.

46. Método para produzir o polipeptídeo como definido em qualquer uma das reivindicações de 1 a 23, caracterizado pelo fato de que compreende: (a) cultivar uma planta transgênica ou uma célula vegetal que compreendam um polinucleotídeo que codifique um polipeptídeo tendo atividade de endoglicanase sob condições condutivas para a produção do polipeptídeo; e (b) recuperar o polipeptídeo.

47. Planta transgênica, parte de planta ou célula vegetal, caracterizada pelo fato de que foi transformada com um polinucleotídeo que

codifica o polipeptídeo como definido em qualquer uma das reivindicações de 1 a 23.

5 48. Composição detergente, caracterizada pelo fato de que compreende o polipeptídeo como definido em qualquer uma das reivindicações de 1 a 23 e um tensoativo.

49. Método para degradar biomassa contendo celulose e hemi-celulose, caracterizado pelo fato de que compreende tratar a biomassa com uma quantidade eficaz do polipeptídeo como definido em qualquer uma das reivindicações de 1 a 23 e recuperar a biomassa degradada.

10 50. Método de acordo com a reivindicação 49, caracterizado pelo fato de que ainda compreende tratar a biomassa com uma quantidade eficaz de endo-1,4-beta-glicanase, exo-1,4-beta-D-glicanase e/ou beta-D-glicosidase.

15 51. Método para degradar biomassa contendo celulose e hemi-celulose, caracterizado pelo fato de que compreende tratar a biomassa com uma célula hospedeira como definida na reivindicação 28 e recuperar a biomassa degradada.

20 52. Método de acordo com a reivindicação 51, caracterizado pelo fato de que ainda compreende tratar a biomassa com uma quantidade eficaz de endo-1,4-beta-glicanase, exo-1,4-beta-D-glicanase e/ou beta-D-glicosidase.

M T L R L P V I S L L A S L A A G
 1 ATGACCCTACGGCTCCCTGTCATCAGCCTGCTGGCCTCGCTGGCAGCAGGC
 A V V V P R A E F H P P L P T W K
 52 GCCGTCGTCGTCCCACGGGCGGAGTTTCACCCCCCTCTCCCGACTTGGAAA
 C T T S G G C V Q Q N T S V V L D
 103 TGCACGACCTCCGGGGGCTGCGTGCAGCAGAACACCAGCGTCGTCCTGGAC
 R D S K Y A A H S A G S R T E S D
 154 CGTGA CT CGAAGTACGCCGCACACAGCGCCGGCTCGCGGACGGAATCGGAT
 Y A A M G V S T S G N A V T L Y H
 205 TACGCGGCAATGGGAGTGTCCACTTCGGGCAATGCCGTGACGCTGTACCAC
 Y V K T N G T L V P A S P R I Y L
 256 TACGTCAAGACCAACGGCACCCCTCGTCCCCGCTTCGCCGCGCATCTACCTC
 L G A D G K Y V L M D L L N Q E L
 307 CTGGGCGCGGACGGCAAGTACGTGCTTATGGACCTCCTCAACCAGGAGCTG
 S V D V D F S A L P C G E N G A F
 358 TCGGTGGACGTCGACTTCTCGGCGCTGCCGTGCGGCGAGAACGGGGCCTTC
 Y L S E M A A D G R G D A G A G D
 409 TACCTGTCCGAGATGGCGGCGGACGGGCGGGGCGACGCGGGGGCGGGCGAC
 G Y C D A Q C Q G Y C C N E M D I
 460 GGGTACTGCGACGCGCAGTGCCAGGGCTACTGCTGCAACGAGATGGACATC
 L E A N S M A T A M T P H P C K G
 511 CTCGAGGCCAACTCGATGGCGACGGCCATGACGCCGCACCCGTGCAAGGGC
 N N C D R S G C G Y N P Y A S G Q
 562 AACAACTGCGACCGCAGCGGCTGCGGCTACAACCCGTACGCCAGCGGCCAG
 R G F Y G P G K T V D T S K P F T
 613 CGCGGCTTCTACGGGCCCCGGCAAGACGGTCGACACGAGCAAGCCCTTCACC
 V V T Q F A A S G G K L T Q I T R
 664 GTCGTCACGCAGTTCGCCGCCAGCGGCGGCAAGCTGACCCAGATCACCCGC
 K Y I Q N G R E I G G G G T I S S
 715 AAGTACATCCAGAACGGCCGGGAGATCGGCGGCGGCGGCACCATCTCCAGC
 C G S E S S T G G L T G M G E A L
 766 TGGCGCTCCGAGTCTTCGACGGGCGGCCTGACCGGCATGGGCGAGGCGCTG
 G R G M V L A M S I W N D A A Q E
 817 GGGCGCGGAATGGTGCTGGCCATGAGCATCTGGAACGACGCGGCCCAGGAG
 M A W L D A G N N G P C A S G Q G

Fig. 1A

868 ATGGCATGGCTCGATGCCGGCAACAACGGCCCTTGCGCCAGTGGCCAGGGC
S P S V I Q S Q H P D T H V V F S
919 AGCCCGTCCGTCATTCAGTCGCAGCATCCCGACCCACGTCGTCTTCTCC
N I R W G D I G S T T K N *
970 AACATCAGGTGGGGCGACATCGGGTCTACCACGAAGAACTAG

Fig. 1B

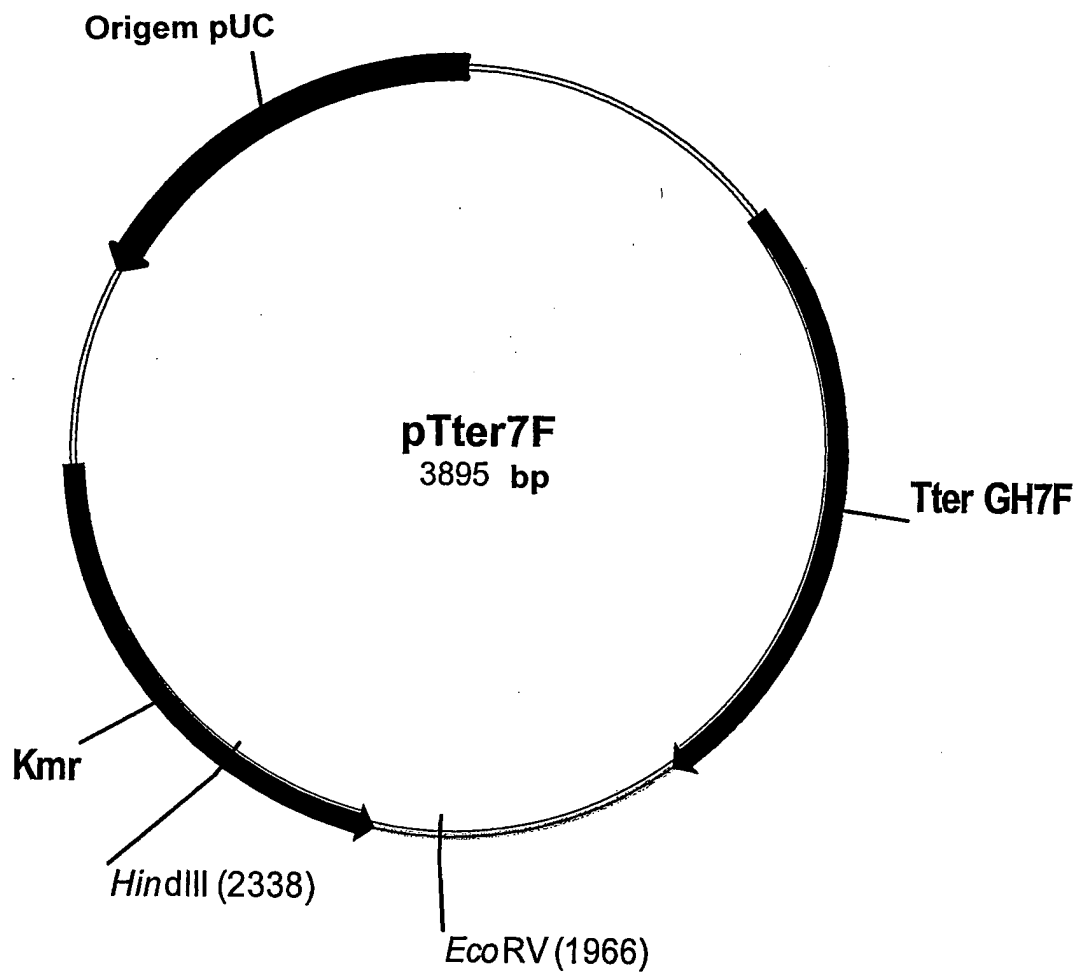


Fig. 2

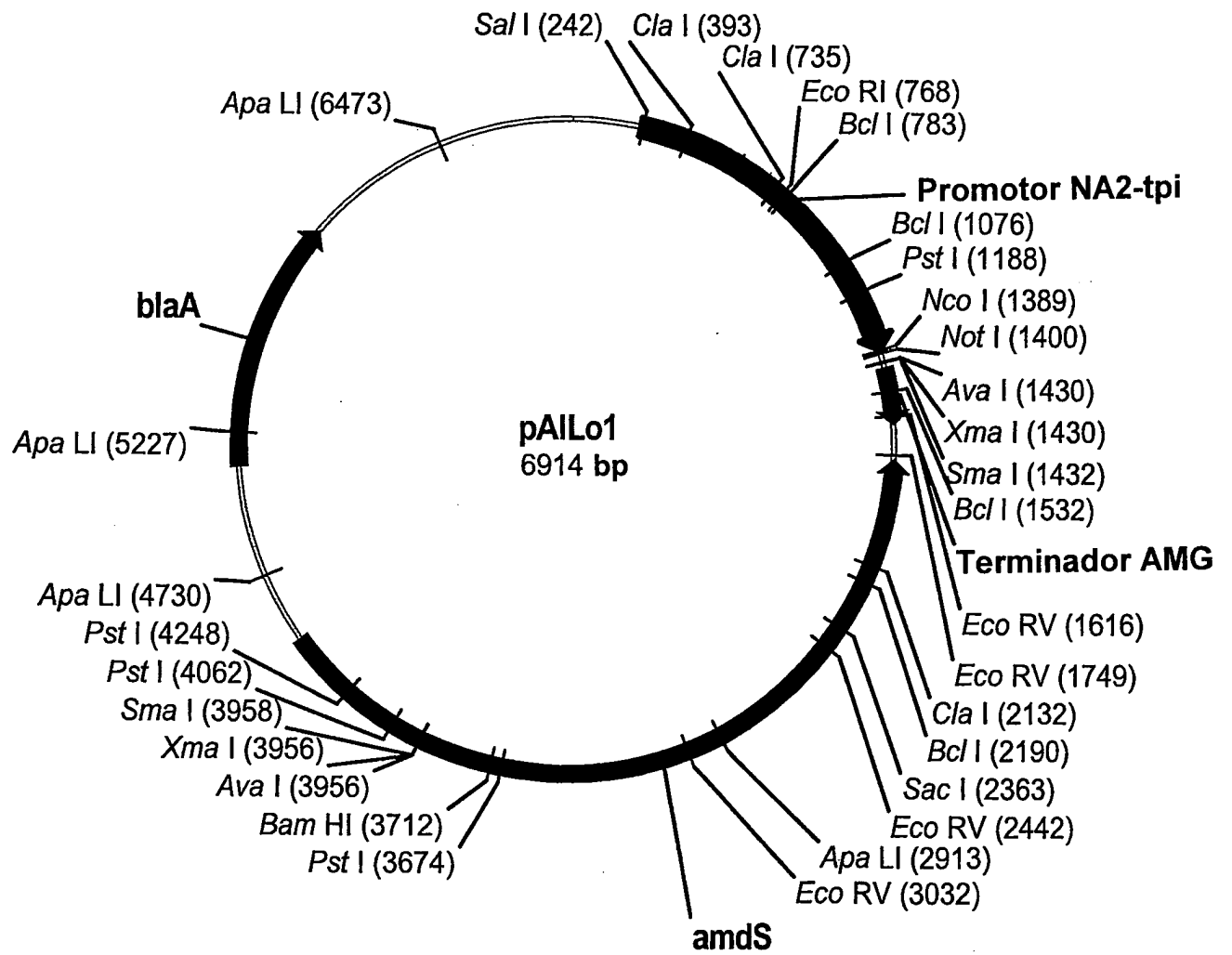


Fig. 3

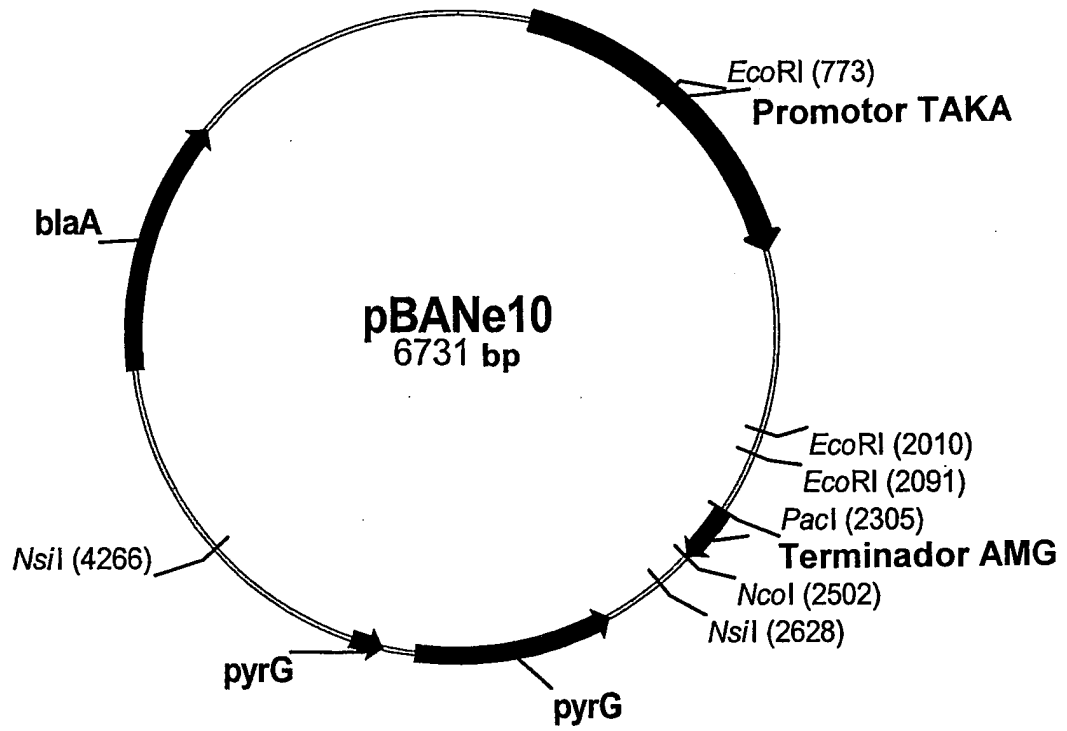


Fig. 4

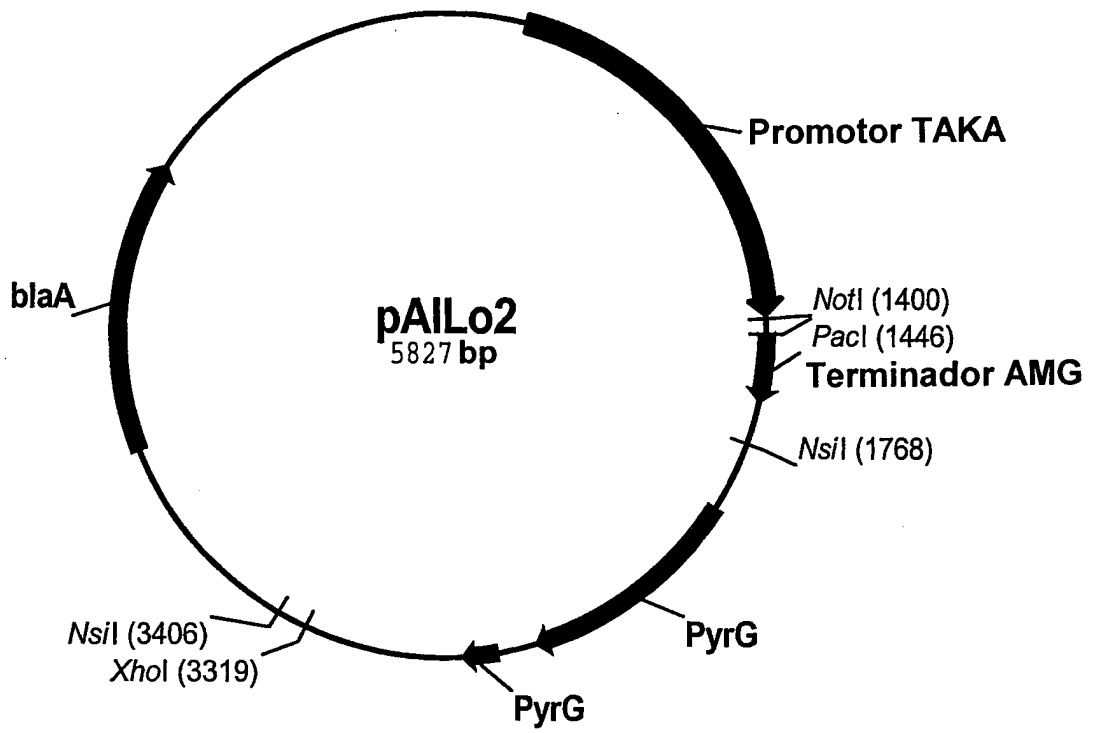


Fig. 5

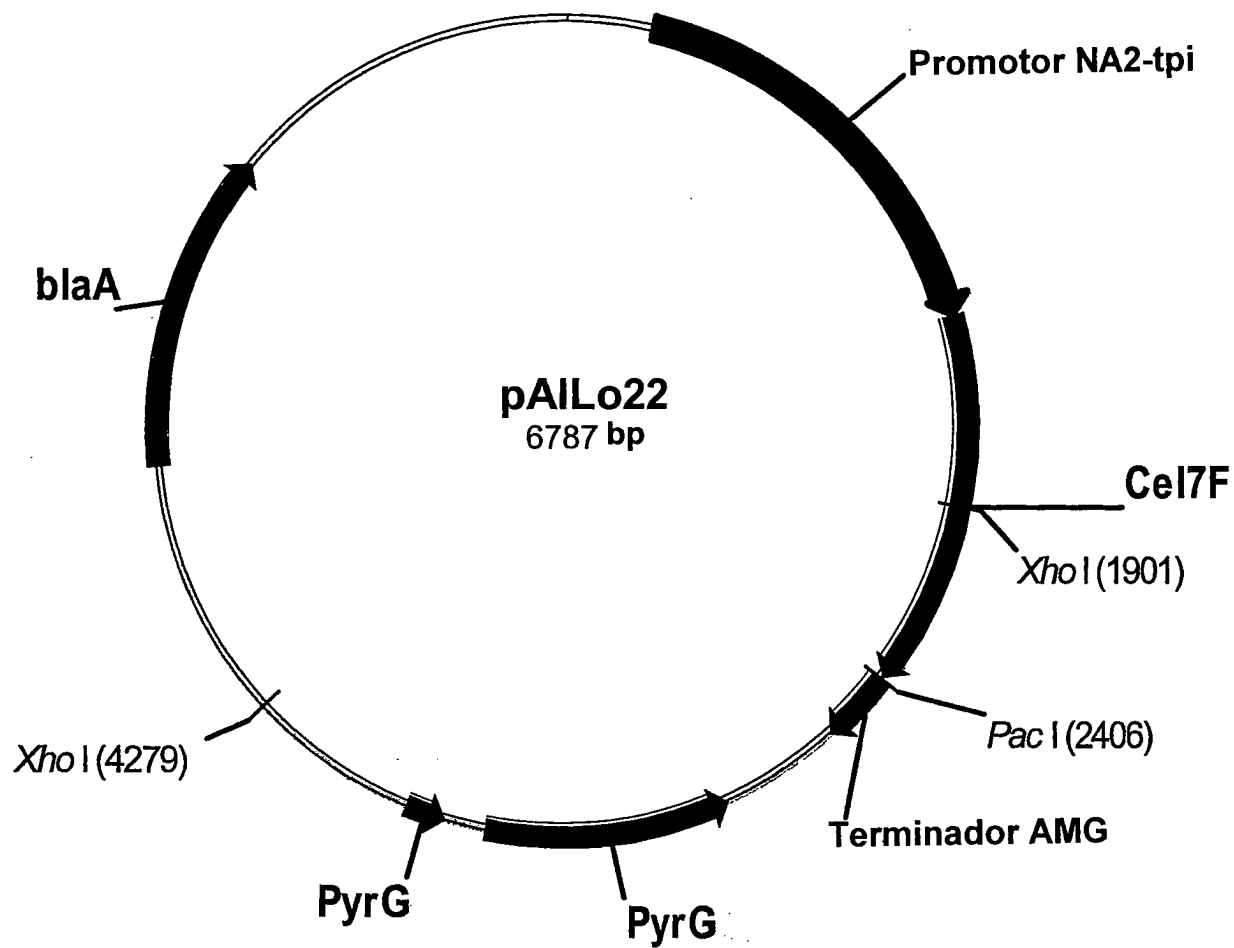


Fig. 6

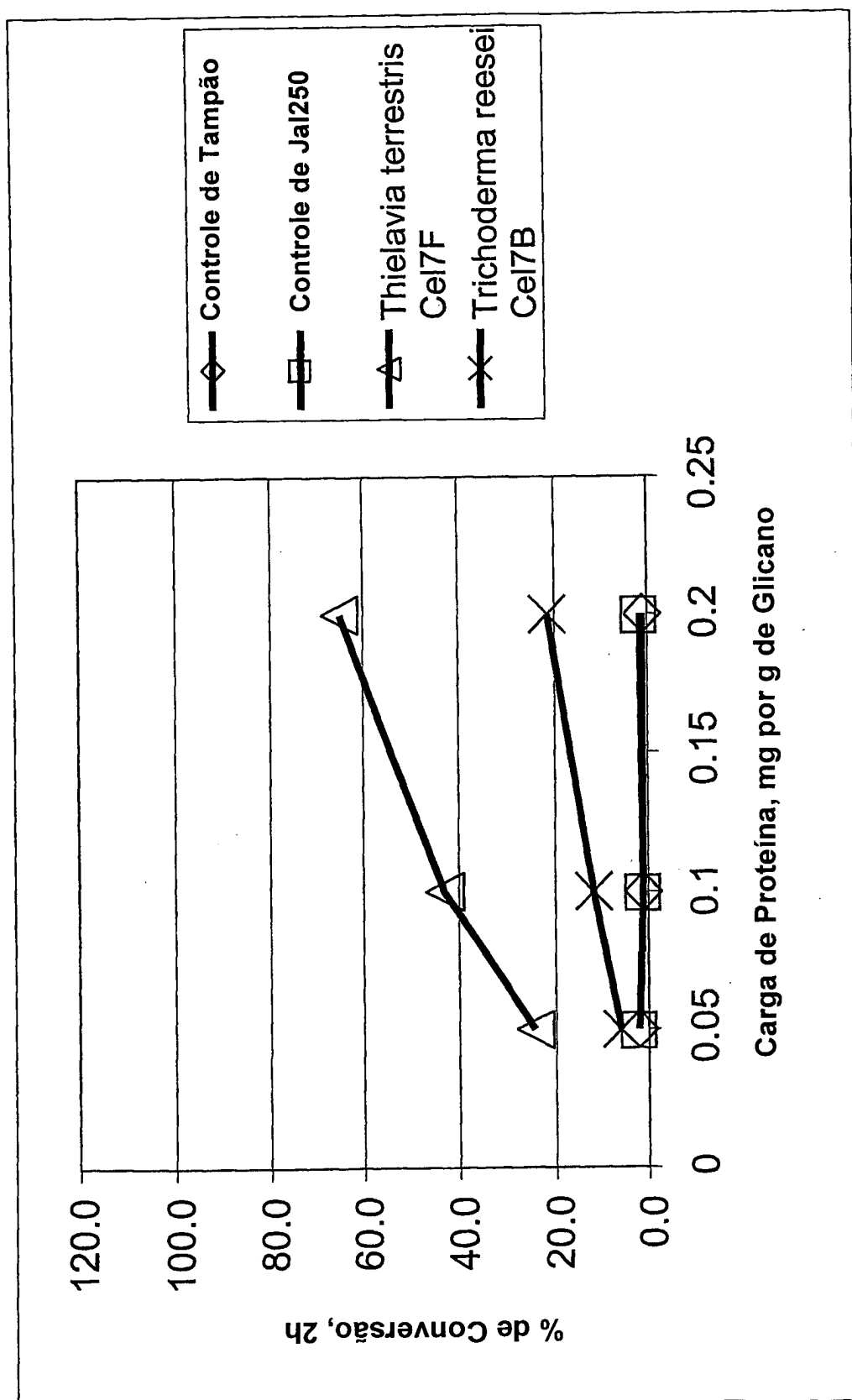


Fig. 7

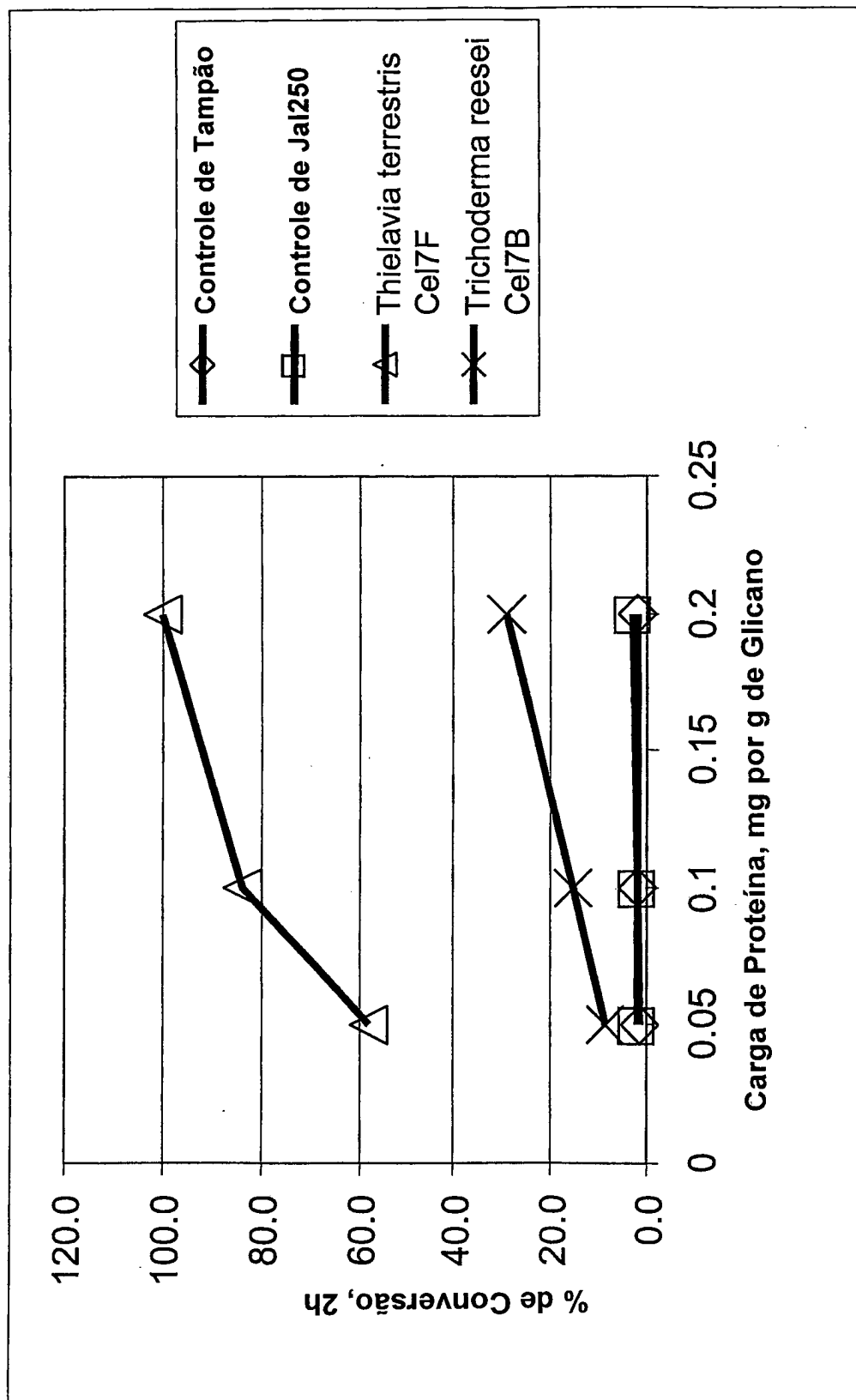


Fig. 8

RESUMO

“POLIPEPTÍDEO ISOLADO, POLINUCLEOTÍDEO ISOLADO, CONSTRUÇÃO DE ÁCIDO NUCLEICO, VETOR DE EXPRESSÃO RECOMBINANTE, CÉLULA HOSPEDEIRA RECOMBINANTE, 5 MÉTODOS PARA PRODUZIR O POLIPEPTÍDEO E PARA PRODUZIR UM MUTANTE DE UMA CÉLULA PRECURSORA, CÉLULA MUTANTE, MÉTODOS PARA PRODUZIR UMA PROTEÍNA E PARA PRODUZIR UM POLINUCLEOTÍDEO, POLINUCLEOTÍDEO MUTANTE, PLANTA TRANSGÊNICA, PARTE DE PLANTA OU 10 CÉLULA VEGETAL, COMPOSIÇÃO DETERGENTE, E, MÉTODO PARA DEGRADAR BIOMASSA CONTENDO CELULOSE E HEMI-CELULOSE”

A presente invenção diz respeito a polipeptídeos isolados tendo atividade de endoglicanase e polinucleotídeos isolados que codificam 15 os polipeptídeos. A invenção também diz respeito a construções de ácido nucleico, vetores e células hospedeiras que compreendem os polinucleotídeos assim como a métodos para produzir e usar os polipeptídeos.

“POLIPEPTÍDEO ISOLADO, POLINUCLEOTÍDEO ISOLADO, CONSTRUÇÃO DE ÁCIDO NUCLEICO, CÉLULA HOSPEDEIRA RECOMBINANTE, MÉTODOS PARA PRODUZIR O POLIPEPTÍDEO E PARA PRODUZIR UM MUTANTE DE UMA CÉLULA PRECURSORA, 5 CÉLULA MUTANTE, MÉTODOS PARA PRODUZIR UMA PROTEÍNA, E PARA PRODUZIR UM POLINUCLEOTÍDEO, POLINUCLEOTÍDEO MUTANTE, PLANTA TRANSGÊNICA, PARTE DE PLANTA OU CÉLULA VEGETAL, COMPOSIÇÃO DETERGENTE, E, MÉTODO PARA DEGRADAR BIOMASSA CONTENDO CELULOSE E HEMI- 10 CELULOSE”

Declaração como para os Direitos às Invenções Feitas Sob Pesquisa e Desenvolvimento Patrocinados pelo Governo Federal

Esta invenção foi feita com apoio do Governo sob o Subcontrato NREL Nº ZCO-30017-02, Prime Contract DE-AC36-98G010337 15 outorgado pelo Departamento de Energia. O governo tem certos direitos nesta invenção.

Fundamentos da invenção

Campo da invenção

A presente invenção diz respeito a polipeptídeos isolados 20 tendo atividade de endoglicanase e polinucleotídeos isolados que codificam os polipeptídeos. A invenção também diz respeito às construções de ácido nucleico, vetores e células hospedeiras que compreendem os polinucleotídeos assim como aos métodos para produzir e usar os polipeptídeos.

25 **Descrição da Técnica Relacionada**

A celulose é um polímero da glicose do açúcar simples covalentemente ligada pelas ligações beta-1,4. Muitos microorganismos produzem enzimas que hidrolisam glicanos ligados em beta. Estas enzimas incluem endoglicanases, celobiohidrolases e beta-glicosidases. As

REIVINDICAÇÕES

1. Polipeptídeo isolado, caracterizado pelo fato de que tem atividade de endoglicanase, selecionado do grupo que consiste de:

5 (a) um polipeptídeo que compreenda uma seqüência de aminoácido que tenha preferivelmente pelo menos 60 % de identidade, mais preferivelmente pelo menos 65 % de identidade, mais preferivelmente pelo menos 70 % de identidade, mais preferivelmente pelo menos 80 % de identidade, mais ainda preferivelmente pelo menos 85 % de identidade, mais preferivelmente pelo menos 90 % de identidade, e ainda mais preferivelmente
10 pelo menos 95 % de identidade com o polipeptídeo maduro da SEQ ID NO: 2;

(b) um polipeptídeo que é codificado por um polinucleotídeo que hibridiza sob preferivelmente pelo menos condições de estringência média, mais preferivelmente pelo menos condições de estringência média-alta, e mais preferivelmente pelo menos condições de estringência alta com (i)
15 o polipeptídeo maduro que codifica a seqüência da SEQ ID NO: 1, (ii) a seqüência de DNA genômico que compreende o polipeptídeo maduro que codifica a seqüência da SEQ ID NO: 1, ou (iii) um filamento complementar de (i) ou (ii); e

20 (c) uma variante que compreenda uma substituição, deleção e/ou inserção de um ou mais aminoácidos do polipeptídeo maduro da SEQ ID NO: 2.

2. Polipeptídeo de acordo com a reivindicação 1, caracterizado pelo fato de que compreende ou consiste do polipeptídeo maduro da SEQ ID
25 NO: 2, ou um fragmento desta tendo atividade de endoglicanase.

3. Polipeptídeo de acordo com a reivindicação 1, caracterizado pelo fato de que é codificado pelo polinucleotídeo contido no plasmídeo pTter7F que está contido na *E. coli* NRRL B-30837.

4. Polipeptídeo de acordo com qualquer uma das

reivindicações de 1 a 3, caracterizado pelo fato de que o polipeptídeo maduro é os aminoácidos 18 a 336 da SEQ ID NO: 2 e o polipeptídeo maduro que codifica a sequência é os nucleotídeos de 52 a 1008 da SEQ ID NO: 1.

5 5. Polipeptídeo de acordo com qualquer uma das reivindicações de 1 a 4, caracterizado pelo fato de que ainda têm atividade enzimática contra um ou mais substratos selecionados do grupo que consiste de xilano, xiloglicano, arabinoxilano, 1,4-galactano, galactomanano, dextrano e quitina.

10 6. Polinucleotídeo isolado, caracterizado pelo fato de que compreende uma sequência de nucleotídeo que codifica o polipeptídeo como definido em qualquer uma das reivindicações de 1 a 5.

15 7. Construção de ácido nucleico, caracterizada pelo fato de que compreende o polinucleotídeo como definido na reivindicação 6 operavelmente ligado a uma ou mais sequências de controle que direcionam a produção do polipeptídeo em um hospedeiro de expressão.

8. Célula hospedeira recombinante, caracterizada pelo fato de que compreende a construção de ácido nucleico como definida na reivindicação 7.

20 9. Método para produzir o polipeptídeo como definido em qualquer uma das reivindicações de 1 a 5, caracterizado pelo fato de que compreende: (a) cultivar uma célula, que na sua forma do tipo selvagem é capaz de produzir o polipeptídeo, sob condições condutivas para a produção do polipeptídeo; e (b) recuperar o polipeptídeo.

25 10. Método para produzir o polipeptídeo como definido em qualquer uma das reivindicações de 1 a 5, caracterizado pelo fato de que compreende: (a) cultivar uma célula hospedeira que compreenda uma construção de ácido nucleico que compreende uma sequência de nucleotídeo que codifica o polipeptídeo sob condições condutivas para a produção do polipeptídeo; e (b) recuperar o polipeptídeo.

11. Método para produzir um mutante de uma célula precursora, caracterizado pelo fato de que compreende romper ou deletar uma sequência de nucleotídeo que codifica o polipeptídeo como definido em qualquer uma das reivindicações de 1 a 5, que resulta na produção do mutante menos do polipeptídeo do que da célula precursora.

12. Célula mutante, caracterizada pelo fato de ser produzida pelo método como definido na reivindicação 11.

13. Célula mutante de acordo com a reivindicação 12, caracterizado pelo fato de que compreende ainda um gene que codifica uma proteína nativa ou heteróloga.

14. Método para produzir uma proteína, caracterizado pelo fato de que compreende: (a) cultivar a célula mutante como definida na reivindicação 13 sob condições condutivas para a produção da proteína; e (b) recuperar a proteína.

15. Polinucleotídeo isolado, caracterizado pelo fato de que é obtido pela (a) hibridização de uma população de DNA sob pelo menos preferivelmente condições de estringência média com (i) o polipeptídeo maduro que codifica a sequência da SEQ ID NO: 1, (ii) a sequência de DNA genômico que compreende o polipeptídeo maduro que codifica a sequência da SEQ ID NO: 1, ou (iii) um filamento complementar de (i) ou (ii); e (b) isolar o polinucleotídeo de hibridização, que codifica um polipeptídeo tendo atividade de endoglicanase.

16. Método para produzir um polinucleotídeo compreendendo uma sequência de nucleotídeo mutante, caracterizado pelo fato de que compreende: (a) introduzir pelo menos uma mutação no polipeptídeo maduro que codifica a sequência da SEQ ID NO: 1, em que a sequência de nucleotídeo mutante codifica um polipeptídeo que consiste do polipeptídeo maduro da SEQ ID NO: 2; e (b) recuperar o polinucleotídeo que compreende a sequência de nucleotídeo mutante.

17. Polinucleotídeo mutante, caracterizado pelo fato de que é produzido pelo método como definido na reivindicação 16.

18. Método para produzir um polipeptídeo, caracterizado pelo fato de que compreende: (a) cultivar uma célula que compreende o polinucleotídeo mutante como definido na reivindicação 17 que codifica o polipeptídeo sob condições condutivas para a produção do polipeptídeo; e (b) recuperar o polipeptídeo.

19. Construção de ácido nucleico, caracterizada pelo fato de que compreende um gene que codifica uma proteína operavelmente ligada a uma sequência de nucleotídeo que codifica um peptídeo de sinal que compreende ou que consiste dos aminoácidos 1 a 17 da SEQ ID NO: 2, em que o gene é estranho à sequência de nucleotídeo.

20. Célula hospedeira recombinante, caracterizada pelo fato de que compreende a construção de ácido nucleico como definida na reivindicação 19.

21. Método para produzir uma proteína, caracterizado pelo fato de que compreende: (a) cultivar a célula hospedeira recombinante como definida na reivindicação 20 sob condições condutivas para a produção da proteína; e (b) recuperar a proteína.

22. Método para produzir o polipeptídeo como definido em qualquer uma das reivindicações de 1 a 5, caracterizado pelo fato de que compreende: (a) cultivar uma planta transgênica ou uma célula vegetal que compreendam um polinucleotídeo que codifique um polipeptídeo tendo atividade de endoglicanase sob condições condutivas para a produção do polipeptídeo; e (b) recuperar o polipeptídeo.

23. Planta transgênica, parte de planta ou célula vegetal, caracterizada pelo fato de que foi transformada com um polinucleotídeo que codifica o polipeptídeo como definido em qualquer uma das reivindicações de 1 a 5.

24. Composição detergente, caracterizada pelo fato de que compreende o polipeptídeo como definido em qualquer uma das reivindicações de 1 a 5 e um tensoativo.

5 25. Método para degradar biomassa contendo celulose e hemi-celulose, caracterizado pelo fato de que compreende tratar a biomassa com uma quantidade eficaz do polipeptídeo como definido em qualquer uma das reivindicações de 1 a 5 ou a célula hospedeira como definida na reivindicação 8 e recuperar a biomassa degradada.

10 26. Método de acordo com a reivindicação 25, caracterizado pelo fato de que ainda compreende tratar a biomassa com uma quantidade eficaz de endo-1,4-beta-glicanase, exo-1,4-beta-D-glicanase e/ou beta-D-glicosidase.

RESUMO

“POLIPEPTÍDEO ISOLADO, POLINUCLEOTÍDEO ISOLADO,
CONSTRUÇÃO DE ÁCIDO NUCLEICO, CÉLULA HOSPEDEIRA
RECOMBINANTE, MÉTODOS PARA PRODUZIR O POLIPEPTÍDEO E
5 PARA PRODUZIR UM MUTANTE DE UMA CÉLULA PRECURSORA,
CÉLULA MUTANTE, MÉTODOS PARA PRODUZIR UMA PROTEÍNA,
E PARA PRODUZIR UM POLINUCLEOTÍDEO, POLINUCLEOTÍDEO
MUTANTE, PLANTA TRANSGÊNICA, PARTE DE PLANTA OU
CÉLULA VEGETAL, COMPOSIÇÃO DETERGENTE, E, MÉTODO
10 PARA DEGRADAR BIOMASSA CONTENDO CELULOSE E HEMI-
CELULOSE”

A presente invenção diz respeito a polipeptídeos isolados
tendo atividade de endoglicanase e polinucleotídeos isolados que codificam
os polipeptídeos. A invenção também diz respeito a construções de ácido
15 nucleico, vetores e células hospedeiras que compreendem os polinucleotídeos
assim como a métodos para produzir e usar os polipeptídeos.

“POLIPEPTÍDEO ISOLADO, POLINUCLEOTÍDEO ISOLADO, CONSTRUÇÃO DE ÁCIDO NUCLEICO, CÉLULA HOSPEDEIRA RECOMBINANTE, MÉTODOS PARA PRODUZIR O POLIPEPTÍDEO E PARA PRODUZIR UM MUTANTE DE UMA CÉLULA PRECURSORA, CÉLULA MUTANTE, MÉTODOS PARA PRODUZIR UMA PROTEÍNA, E PARA PRODUZIR UM POLINUCLEOTÍDEO, POLINUCLEOTÍDEO MUTANTE, COMPOSIÇÃO DETERGENTE, E, MÉTODO PARA DEGRADAR BIOMASSA CONTENDO CELULOSE E HEMI-CELULOSE”

10 **Declaração como para os Direitos às Invenções Feitas Sob Pesquisa e Desenvolvimento Patrocinados pelo Governo Federal**

Esta invenção foi feita com apoio do Governo sob o Subcontrato NREL Nº ZCO-30017-02, Prime Contract DE-AC36-98G010337 outorgado pelo Departamento de Energia. O governo tem certos direitos nesta invenção.

Fundamentos da invenção

Campo da invenção

A presente invenção diz respeito a polipeptídeos isolados tendo atividade de endoglicanase e polinucleotídeos isolados que codificam os polipeptídeos. A invenção também diz respeito às construções de ácido nucleico, vetores e células hospedeiras que compreendem os polinucleotídeos assim como aos métodos para produzir e usar os polipeptídeos.

Descrição da Técnica Relacionada

25 A celulose é um polímero da glicose do açúcar simples covalentemente ligada pelas ligações beta-1,4. Muitos microorganismos produzem enzimas que hidrolisam glicanos ligados em beta. Estas enzimas incluem endoglicanases, celobioidrolases e beta-glicosidases. As

REIVINDICAÇÕES

1. Polipeptídeo isolado, caracterizado pelo fato de que tem atividade de endoglicanase, selecionado do grupo que consiste de:

5 (a) um polipeptídeo que compreenda uma sequência de aminoácido que tenha preferivelmente pelo menos 60 % de identidade, mais preferivelmente pelo menos 65 % de identidade, mais preferivelmente pelo menos 70 % de identidade, mais preferivelmente pelo menos 80 % de identidade, mais ainda preferivelmente pelo menos 85 % de identidade, mais preferivelmente pelo menos 90 % de identidade, e ainda mais preferivelmente
10 pelo menos 95 % de identidade com o polipeptídeo maduro da SEQ ID NO: 2;

(b) um polipeptídeo que é codificado por um polinucleotídeo que hibridiza sob preferivelmente pelo menos condições de estringência média, mais preferivelmente pelo menos condições de estringência média-alta, e mais preferivelmente pelo menos condições de estringência alta com (i)
15 o polipeptídeo maduro que codifica a sequência da SEQ ID NO: 1, (ii) a sequência de DNA genômico que compreende o polipeptídeo maduro que codifica a sequência da SEQ ID NO: 1, ou (iii) um filamento complementar de (i) ou (ii); e

20 (c) uma variante que compreenda uma substituição, deleção e/ou inserção de um ou mais aminoácidos do polipeptídeo maduro da SEQ ID NO: 2.

2. Polipeptídeo de acordo com a reivindicação 1, caracterizado pelo fato de que compreende ou consiste do polipeptídeo maduro da SEQ ID
25 NO: 2, ou um fragmento desta tendo atividade de endoglicanase.

3. Polipeptídeo de acordo com a reivindicação 1, caracterizado pelo fato de que é codificado pelo polinucleotídeo contido no plasmídeo pTter7F que está contido na *E. coli* NRRL B-30837.

4. Polipeptídeo de acordo com qualquer uma das

reivindicações de 1 a 3, caracterizado pelo fato de que o polipeptídeo maduro é os aminoácidos 18 a 336 da SEQ ID NO: 2 e o polipeptídeo maduro que codifica a sequência é os nucleotídeos de 52 a 1008 da SEQ ID NO: 1.

5 5. Polipeptídeo de acordo com qualquer uma das reivindicações de 1 a 4, caracterizado pelo fato de que ainda têm atividade enzimática contra um ou mais substratos selecionados do grupo que consiste de xilano, xiloglicano, arabinoxilano, 1,4-galactano, galactomanano, dextrano e quitina.

10 6. Polinucleotídeo isolado, caracterizado pelo fato de que compreende uma sequência de nucleotídeo que codifica o polipeptídeo como definido em qualquer uma das reivindicações de 1 a 5.

15 7. Construção de ácido nucleico, caracterizada pelo fato de que compreende o polinucleotídeo como definido na reivindicação 6 operavelmente ligado a uma ou mais sequências de controle que direcionam a produção do polipeptídeo em um hospedeiro de expressão.

8. Célula hospedeira recombinante, caracterizada pelo fato de que compreende a construção de ácido nucleico como definida na reivindicação 7.

20 9. Método para produzir o polipeptídeo como definido em qualquer uma das reivindicações de 1 a 5, caracterizado pelo fato de que compreende: (a) cultivar uma célula, que na sua forma do tipo selvagem é capaz de produzir o polipeptídeo, sob condições condutivas para a produção do polipeptídeo; e (b) recuperar o polipeptídeo.

25 10. Método para produzir o polipeptídeo como definido em qualquer uma das reivindicações de 1 a 5, caracterizado pelo fato de que compreende: (a) cultivar uma célula hospedeira que compreenda uma construção de ácido nucleico que compreende uma sequência de nucleotídeo que codifica o polipeptídeo sob condições condutivas para a produção do polipeptídeo; e (b) recuperar o polipeptídeo.

11. Método para produzir um mutante de uma célula precursora, caracterizado pelo fato de que compreende romper ou deletar uma sequência de nucleotídeo que codifica o polipeptídeo como definido em qualquer uma das reivindicações de 1 a 5, que resulta na produção do mutante menos do polipeptídeo do que da célula precursora.

12. Célula mutante, caracterizada pelo fato de ser produzida pelo método como definido na reivindicação 11.

13. Célula mutante de acordo com a reivindicação 12, caracterizado pelo fato de que compreende ainda um gene que codifica uma proteína nativa ou heteróloga.

14. Método para produzir uma proteína, caracterizado pelo fato de que compreende: (a) cultivar a célula mutante como definida na reivindicação 13 sob condições condutivas para a produção da proteína; e (b) recuperar a proteína.

15. Polinucleotídeo isolado, caracterizado pelo fato de que é obtido pela (a) hibridização de uma população de DNA sob pelo menos preferivelmente condições de estringência média com (i) o polipeptídeo maduro que codifica a sequência da SEQ ID NO: 1, (ii) a sequência de DNA genômico que compreende o polipeptídeo maduro que codifica a sequência da SEQ ID NO: 1, ou (iii) um filamento complementar de (i) ou (ii); e (b) isolar o polinucleotídeo de hibridização, que codifica um polipeptídeo tendo atividade de endoglicanase.

16. Método para produzir um polinucleotídeo compreendendo uma sequência de nucleotídeo mutante, caracterizado pelo fato de que compreende: (a) introduzir pelo menos uma mutação no polipeptídeo maduro que codifica a sequência da SEQ ID NO: 1, em que a sequência de nucleotídeo mutante codifica um polipeptídeo que consiste do polipeptídeo maduro da SEQ ID NO: 2; e (b) recuperar o polinucleotídeo que compreende a sequência de nucleotídeo mutante.

17. Polinucleotídeo mutante, caracterizado pelo fato de que é produzido pelo método como definido na reivindicação 16.

18. Método para produzir um polipeptídeo, caracterizado pelo fato de que compreende: (a) cultivar uma célula que compreende o polinucleotídeo mutante como definido na reivindicação 17 que codifica o polipeptídeo sob condições condutivas para a produção do polipeptídeo; e (b) recuperar o polipeptídeo.

19. Construção de ácido nucleico, caracterizada pelo fato de que compreende um gene que codifica uma proteína operavelmente ligada a uma sequência de nucleotídeo que codifica um peptídeo de sinal que compreende ou que consiste dos aminoácidos 1 a 17 da SEQ ID NO: 2, em que o gene é estranho à sequência de nucleotídeo.

20. Célula hospedeira recombinante, caracterizada pelo fato de que compreende a construção de ácido nucleico como definida na reivindicação 19.

21. Método para produzir uma proteína, caracterizado pelo fato de que compreende: (a) cultivar a célula hospedeira recombinante como definida na reivindicação 20 sob condições condutivas para a produção da proteína; e (b) recuperar a proteína.

22. Método para produzir o polipeptídeo como definido em qualquer uma das reivindicações de 1 a 5, caracterizado pelo fato de que compreende: (a) cultivar uma planta transgênica ou uma célula vegetal que compreendam um polinucleotídeo que codifique um polipeptídeo tendo atividade de endoglicanase sob condições condutivas para a produção do polipeptídeo; e (b) recuperar o polipeptídeo.

23. Composição detergente, caracterizada pelo fato de que compreende o polipeptídeo como definido em qualquer uma das reivindicações de 1 a 5 e um tensoativo.

24. Método para degradar biomassa contendo celulose e hemi-

celulose, caracterizado pelo fato de que compreende tratar a biomassa com uma quantidade eficaz do polipeptídeo como definido em qualquer uma das reivindicações de 1 a 5 ou a célula hospedeira como definida na reivindicação 8 e recuperar a biomassa degradada.

- 5 25. Método de acordo com a reivindicação 24, caracterizado pelo fato de que ainda compreende tratar a biomassa com uma quantidade eficaz de endo-1,4-beta-glicanase, exo-1,4-beta-D-glicanase e/ou beta-D-glicosidase.

RESUMO

5 “POLIPEPTÍDEO ISOLADO, POLINUCLEOTÍDEO ISOLADO,
CONSTRUÇÃO DE ÁCIDO NUCLEICO, CÉLULA HOSPEDEIRA
RECOMBINANTE, MÉTODOS PARA PRODUZIR O POLIPEPTÍDEO E
PARA PRODUZIR UM MUTANTE DE UMA CÉLULA PRECURSORA,
CÉLULA MUTANTE, MÉTODOS PARA PRODUZIR UMA PROTEÍNA,
E PARA PRODUZIR UM POLINUCLEOTÍDEO, POLINUCLEOTÍDEO
MUTANTE, COMPOSIÇÃO DETERGENTE, E, MÉTODO PARA
10 DEGRADAR BIOMASSA CONTENDO CELULOSE E HEMI-
CELULOSE”

A presente invenção diz respeito a polipeptídeos isolados
tendo atividade de endoglicanase e polinucleotídeos isolados que codificam
os polipeptídeos. A invenção também diz respeito a construções de ácido
nucleico, vetores e células hospedeiras que compreendem os polinucleotídeos
15 assim como a métodos para produzir e usar os polipeptídeos.