



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

(21) **PI0710218-6 A2**

(22) Data de Depósito: 30/03/2007
(43) Data da Publicação: 02/08/2011
(RPI 2117)



(51) *Int.Cl.:*
C12N 9/42 2006.01

(54) Título: **POLIPEPTÍDEO ISOLADO, POLINUCLEOTÍDEO, CONSTRUÇÃO DE ÁCIDO NUCLEICO, VETOR DE EXPRESSÃO RECOMBINANTE, CÉLULA HOSPEDEIRA RECOMBINANTE, MÉTODOS PARA PRODUZIR O POLIPEPTÍDEO, PARA PRODUZIR UM MUTANTE A PARTIR DE UMA CÉLULA PRECURSORA, PARA PRODUZIR UMA PROTEÍNA, PARA PRODUZIR UM POLINUCLEOTÍDEO, PARA DEGRADAR OU CONVERTER UM MATERIAL CELULÓSICO, PARA PRODUZIR UMA SUBSTÂNCIA E PARA INIBIR A EXPRESSÃO DE UM POLIPEPTÍDEO EM UMA CÉLULA, CÉLULA MUTANTE, PLANTA TRANSGÊNICA, PARTE DE PLANTA OU CÉLULA DE PLANTA, E, MOLÉCULA DE RNA DE FILAMENTO DUPLO**

(57) Resumo: POLIPEPTÍDEO ISOLADO, POLINUCLEOTÍDEO, CONSTRUÇÃO DE ÁCIDO NUCLEICO, VETOR DE EXPRESSÃO RECOMBINANTE, CÉLULA HOSPEDEIRA RECOMBINANTE, MÉTODOS PARA PRODUZIR O POLIPEPTÍDEO, PARA PRODUZIR UM MUTANTE DE UMA CÉLULA PRECURSORA, PARA PRODUZIR UMA PROTEÍNA, PARA PRODUZIR UM POLINUCLEOTÍDEO, PARA DEGRADAR OU CONVERTER UM MATERIAL CELULÓSICO, PARA PRODUZIR UMA SUBSTÂNCIA, E PARA IMBIR A EXPRESSÃO DE UM POLIPEPTÍDEO EM UMA CÉLULA, CÉLULA MUTANTE, PLANTA TRANSGÊNICA, PARTE DE PLANTA OU CÉLULA DE PLANTA, E, MOLÉCULA DE RNA DE FILAMENTO DUPLO A presente invenção diz respeito a polipeptídeos isolados tendo atividade de endoglucanase 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 métodos para produzir e usar os polipeptídeos.

(30) Prioridade Unionista: 30/03/2006 US 60/788523

(73) Titular(es): Novozymes A/S, Novozymes, Inc.

(72) Inventor(es): Elena Vlasenko, Elizabeth Zaretsky, Marcus Sakari Kauppinen, Paul Harris

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

(86) Pedido Internacional: PCT US2007065639 de 30/03/2007

(87) Publicação Internacional: WO 2007/057621 de 15/05/2008

“POLIPEPTÍDEO ISOLADO, POLINUCLEOTÍDEO, CONSTRUÇÃO DE
ÁCIDO NUCLEICO, VETOR DE EXPRESSÃO RECOMBINANTE,
CÉLULA HOSPEDEIRA RECOMBINANTE, MÉTODOS PARA
PRODUIR O POLIPEPTÍDEO, PARA PRODUIR UM MUTANTE DE
5 UMA CÉLULA PRECURSORA, PARA PRODUIR UMA PROTEÍNA,
PARA PRODUIR UM POLINUCLEOTÍDEO, PARA DEGRADAR OU
CONVERTER UM MATERIAL CELULÓSICO, PARA PRODUIR UMA
SUBSTÂNCIA, E PARA INIBIR A EXPRESSÃO DE UM POLIPEPTÍDEO
EM UMA CÉLULA, CÉLULA MUTANTE, PLANTA TRANSGÊNICA,
10 PARTE DE PLANTA OU CÉLULA DE PLANTA, E, MOLÉCULA DE
RNA DE FILAMENTO DUPLO”

Referência a uma Listagem de Seqüência

Este pedido contém uma Listagem de Seqüência na forma
legível em computador. A forma legível em computador é aqui incorporadas
15 por referência.

Referência a um Depósito de Material Biológico

Este pedido contém uma referência aos depósitos de material
biológico que foi fabricado no Northern Regional Research Center (NRRL)
sob o Tratado de Budapeste e designados os números de acesso NRRL B-
20 30898 e NRRL B-30901, depósitos microbianos estes que são aqui
incorporados por referência.

Fundamentos da invenção

Campo da invenção

A presente invenção diz respeito a polipeptídeos isolados
25 tendo atividade de endoglucanase 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 métodos para produzir e usar os polipeptídeos.

Descrição da Técnica Relacionada

A celulose é um polímero do açúcar simples glicose covalentemente ligado pelas ligações beta-1,4. Muitos microorganismos produzem enzimas que hidrolisam glicanos ligados em beta. Estas enzimas incluem endoglucanases, celobioidrolases e beta-glucosidases. As endoglucanases digerem o polímero de celulose em locais aleatórios, abrindo-o ao ataque pelas celobioidrolases. As celobioidrolases sequencialmente liberam moléculas de celobiose das extremidades do polímero de celulose. A celobioidrolase I é uma atividade de 1,4-beta-D-glicano celobioidrolase (E.C. 3.2.1.91) que catalisa a hidrólise de ligações 1,4-beta-D-glicosídicas em celulose, celotetrioise ou qualquer polímero contendo celulose ligada em beta-1,4, liberando celobiose das extremidades redutoras da cadeia. A celobioidrolase II é uma atividade de 1,4-beta-D-glicano celobioidrolase (E.C. 3.2.1.91) que catalisa a hidrólise de ligações 1,4-beta-D-glicosídicas em celulose, celotetrioise ou qualquer polímero contendo celulose ligada 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-glucosidases hidrolisam celobiose em glicose.

A conversão de estoques de alimentação celulósicos em etanol tem a vantagem da disponibilidade fácil de grandes quantidades de estoque de alimentação, o desejo de evitar queimar ou aterrar os materiais e a limpeza do combustível de etanol. Madeira, resíduos agrícolas, safras herbáceas e resíduos sólidos municipais foram considerados como estoques de alimentação para a produção de etanol. Estes materiais primariamente consistem de celulose, hemicelulose e lignina. Uma vez que a celulose é convertida para glicose, a glicose é facilmente fermentada pela levedura em etanol.

Kvesitadze *et al*, 1995, Applied Biochemistry and Biotechnology 50: 137-143, descreve a isolamento e propriedades de uma endoglucanase termoestável de uma cepa mutante termofílica de *Thielavia*

terrestris. Gilbert *et al.*, 1992. Bioresource Technology 39: 147-154, descrevem a caracterização das enzimas presentes no sistema de celulose de *Thielavia terrestris* 255B Breuil *et al.*, 1986, Biotechnology Letters 8: 673-676, descreve a produção e localização de celulasas e beta-glucosidases das
5 cepas C464 e NRRL 8126 de *Thielavia terrestris*.

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

É um objetivo da presente invenção fornecer polipeptídeos tendo atividade de endoglucanase melhorada e polinucleotídeos que
15 codifiquem os polipeptídeos.

Sumário da invenção

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

20 (a) um polipeptídeo que compreende uma seqüência de aminoácido tendo pelo menos 90% de identidade com o polipeptídeo maduro da SEQ ID NO: 2 ou pelo menos 60% de identidade com o polipeptídeo maduro da SEQ ID NO: 4;

25 (b) um polipeptídeo que é codificado por um polinucleotídeo que hibridiza sob condições de pelo menos alta estringência com (i) a seqüência que codifica o polipeptídeo maduro da SEQ ID NO: 1, (ii) a seqüência de cDNA contida na seqüência que codifica o polipeptídeo maduro da SEQ ID NO: 1 ou (iii) um filamento complementar de (i) ou (ii) ou sob condições de pelo menos estringência média com (i) a seqüência que codifica

o polipeptídeo maduro da SEQ ID NO: 3, (ii) a sequência de DNA genômico que compreende a sequência que codifica o polipeptídeo maduro da SEQ ID NO: 3 ou (iii) um filamento complementar de (i) ou (ii);

5 (c) um polipeptídeo que é codificado por um polinucleotídeo que compreende uma sequência de nucleotídeo tendo pelo menos 90% de identidade com a sequência que codifica o polipeptídeo maduro da SEQ ID NO: 1 ou pelo menos 60% de identidade com a sequência que codifica o polipeptídeo maduro da SEQ ID NO: 3; e

10 (d) uma variante que compreende uma substituição, deleção e/ou inserção de um ou mais aminoácidos do polipeptídeo maduro da SEQ ID NO: 2 ou SEQ ID NO: 4.

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

15 (a) um polinucleotídeo que codifica um polipeptídeo que compreende uma sequência de aminoácido tendo pelo menos 90% de identidade com o polipeptídeo maduro da SEQ ID NO: 2 ou pelo menos 60% de identidade com o polipeptídeo maduro da SEQ ID NO: 4;

20 (b) um polinucleotídeo que hibridiza sob condições de pelo menos alta estringência com (i) a sequência que codifica o polipeptídeo maduro da SEQ ID NO: 1, (ii) a sequência de cDNA contida na sequência que codifica o polipeptídeo maduro da SEQ. ID NO: 1. ou (iii) um filamento complementar de (i) ou (ii) ou sob condições de pelo menos estringência média com (i) a sequência que codifica o polipeptídeo maduro da SEQ ID
25 NO: 3, (ii) a sequência de DNA genômico que compreende a sequência que codifica o polipeptídeo maduro da SEQ ID NO: 3 ou (iii) um filamento complementar de (i) ou (ii);

(c) um polinucleotídeo que compreende uma sequência de nucleotídeo tendo pelo menos 90% de identidade com a sequência que

codifica o polipeptídeo maduro da SEQ ID NO: 1 ou pelo menos 60% de identidade com a sequência que codifica o polipeptídeo maduro da SEQ ID NO: 3; e

(d) um polinucleotídeo que codifica uma variante que compreende uma substituição, deleção e/ou inserção de um ou mais aminoácidos do polipeptídeo maduro da SEQ ID NO: 2 ou SEQ ID NO: 4,

Em um aspecto preferido, o polipeptídeo maduro tem os aminoácidos de 19 a 396 da SEQ ID NO: 2, Em um outro aspecto preferido, o polipeptídeo maduro tem os aminoácidos de 20 a 400 da SEQ ID NO: 4, Em um outro aspecto preferido, a sequência que codifica o polipeptídeo maduro tem os nucleotídeos de 110 a 1557 da SEQ ID NO: 1. Em um outro aspecto preferido, a sequência que codifica o polipeptídeo maduro tem os nucleotídeos de 58 a 1200 da SEQ ID NO: 3.

A presente invenção também diz respeito a construções de ácido nucleico, vetores de expressão recombinantes, células hospedeiras recombinantes que compreendem os polinucleotídeos e métodos de produzir um polipeptídeo tendo atividade de endoglucanase.

A presente invenção também diz respeito a métodos de inibir a expressão de um polipeptídeo em uma célula, que compreendem administrar à célula ou expressar na célula uma molécula de RNA de filamento duplo (dsRNA), em que a dsRNA compreende uma subsequência de um polinucleotídeo da presente invenção. A presente também diz respeito a uma tal molécula de RNA inibitória de filamento duplo (dsRNA), em que opcionalmente a dsRNA é uma siRNA ou uma molécula miRNA.

A presente invenção também diz respeito a métodos de usar os polipeptídeos tendo atividade de endoglucanase na conversão de celulose para glicose e várias substâncias.

A presente invenção também diz respeito a plantas que compreendem um polinucleotídeo isolado que codifica um tal polipeptídeo

tendo atividade de endoglucanase.

A presente invenção também diz respeito a métodos para produzir um tal polipeptídeo tendo atividade de endoglucanase, que compreendem: (a) cultivar uma planta transgênica ou uma célula de planta
5 que compreendam um polinucleotídeo que codifica um tal polipeptídeo tendo atividade de endoglucanase sob condições condutivas para a produção do polipeptídeo; e (b) recuperar o polipeptídeo.

A presente invenção diz respeito ainda às construções de ácido nucleico que compreendem um gene que codifica uma proteína, em que o
10 gene está operavelmente ligado a uma seqüência de nucleotídeo que codifica um peptídeo de sinal que compreende ou que consiste dos aminoácidos de 1 a 18 da SEQ ID NO: 2 ou aminoácidos de 1 a 19 da SEQ ID NO: 4, em que o gene é estranho para a primeira e segunda seqüências de nucleotídeo.

Descrição Resumida das Figuras

15 A Figura 1 mostra um mapa de restrição de pPH23.

As Figuras 2A e 2B mostram a seqüência de DNA genômico e a seqüência de aminoácido deduzida de uma endoglucanase NRRL 8126 CEL6B de *Thielavia terrestris* (SEQ ID NOs: 1 e 2, respectivamente).

A Figura 3 mostra um mapa de restrição de pCIBG146.

20 As Figuras 4A e 4B mostram a seqüência de cDNA e a seqüência de aminoácido deduzida de uma endoglucanase NRRL 8126 CEL6C de *Thielavia terrestris* (SEQ ID NOs: 3 e 4, respectivamente).

A Figura 5 mostra um mapa de restrição de pAlLo1.

A Figura 6 mostra um mapa de restrição de pBANE10,

25 A Figura 7 mostra um mapa de restrição de pAlLo2,

A Figura 8 mostra um mapa de restrição de pEJG105,

A Figura 9 mostra um mapa de restrição de pA2I3G146.

Definições

Atividade de endoglucanase: O termo “atividade de

endoglucanase” é aqui definido como uma endo-1,4-beta-D-glicano 4-glicanoidrolase (E.C. No. 32.1.4) que catalisa a endoidrólise de ligações 1,4-beta-D-glicosídicas em celulose, derivados de celulose (tais como carboximetil celulose e hidroxietil celulose), lignocelulose, derivados de lignocelulose, lichenina, ligações beta-1,4 em beta-1,3 glicanos misturados 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 endoglucanase 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 endoglucanase é definida como 1,0 μ mol de açúcares redutores produzidos por minuto a 50°C, pH 5,0.

Em um aspecto preferido, os polipeptídeos da presente invenção tendo atividade de endoglucanase têm ainda atividade de enzima voltada para um ou mais substratos selecionados do grupo que consiste de xilano, xiloglicano, arabinoxilano, 1,4-beta-D-manana e galactomanana. A atividade dos polipeptídeos tendo atividade de endoglucanase sobre estes substratos de polissacarídeo é determinada como porcentagem do substrato hidrolisado para reduzir açúcares depois de incubar o substrato (5 mg por ml) com um polipeptídeo tendo atividade de endoglucanase da presente invenção (5 mg de proteína por g de substrato) por 24 horas com agitação intermitente no pH 5,0 (50 mM de acetato de sódio) e 50°C. Os açúcares redutores em misturas de hidrólise são determinados pelo ensaio da hidrazida do ácido p-hidroxibenzóico (PHBAH).

Em um aspecto mais preferido, os polipeptídeos da presente invenção tendo atividade de endoglucanase têm ainda atividade de enzima voltada para o xilano. Em um outro aspecto mais preferido, os polipeptídeos da presente invenção tendo atividade de endoglucanase têm ainda atividade de enzima voltada para o xiloglicano. Em um outro aspecto mais preferido, os

polipeptídeos da presente invenção tendo atividade de endoglucanase têm ainda atividade de enzima voltada para o arabinoxilano. Em um outro aspecto mais preferido, os polipeptídeos da presente invenção tendo atividade de endoglucanase têm ainda atividade de enzima voltada para a 1,4-beta-D-manana. Em um outro aspecto mais preferido, os polipeptídeos da presente invenção tendo atividade de endoglucanase têm ainda atividade de enzima voltada para a galactomanana. Em um outro aspecto mais preferido, os polipeptídeos da presente invenção tendo atividade de endoglucanase têm ainda atividade de enzima voltada para o xilano, xiloglicano, arabinoxilano, 1,4-beta-D-manana e/ou galactomanana.

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 endoglucanase do polipeptídeo maduro da SEQ ID NO: 2 ou SEQ ID NO: 4.

Família da glicosídeo hidrolase 6 ou Família GH6: Os termos “Família da glicosídeo hidrolase 6” ou “Família GH6” são aqui definidos como um polipeptídeo que cai na Família 6 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 sequence-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 pela SDS-PAGE.

Polipectídeo substancialmente puro: O termo “polipectídeo substancialmente puro” denota aqui uma preparação de polipectídeo que 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 polipectídeo com que o mesmo está nativa ou recombinantemente associado.

É, por tanto, preferido que o polipectídeo 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 polipectídeo total presente na preparação.

Os polipectídeos da presente invenção estão preferivelmente em uma forma substancialmente pura. Em particular, é preferido que os polipectídeos estejam na “forma essencialmente pura”, isto é, que a preparação de polipectídeo esteja essencialmente livre de outro material de polipectídeo com que esteja nativa ou recombinantemente associado. Isto pode ser realizado, por exemplo, preparando-se o polipectídeo por meio de métodos recombinantes bem conhecidos ou pelos métodos de purificação clássicos.

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

Polipectídeo maduro: O termo “polipectídeo maduro” é aqui

definido como um polipeptídeo tendo atividade de endoglucanase que está na sua forma final a seguir da tradução e quaisquer modificações pós traducionais, tais como o processamento de terminal N, truncagem de terminal C, glicosilação, etc.

5 **Seqüência codificadora de polipeptídeo maduro:** O termo “seqüência codificadora de polipeptídeo maduro” é aqui definido como uma seqüência de nucleotídeo que codifica um polipeptídeo maduro tendo atividade de endoglucanase.

10 **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 usando o algoritmo de Needleman-Wunsch (Needleman e Wunsch, 1970, J. Mol, Biol. 48: 443-453) como implementado no programa Needle da EMBOSS com penalidade de abertura de intervalo de 10, penalidade de extensão de intervalo de 0,5 e a matriz EBLOSUM62. A saída da “identidade mais longa” rotulada por Needle é usada como a identidade percentual e é calculada como segue:

(Resíduos Idênticos x 100)/(Comprimento de Alinhamento - Número de Intervalos no Alinhamento)

20 Para os propósitos da presente invenção, o grau de identidade entre duas seqüências de nucleotídeo é determinado usando o algoritmo de Needleman-Wunsch (Needleman e Wunsch, 1970, J. Mol. Biol. 48: 443-453) como implementado no programa Needle de EMBOSS com penalidade de abertura de intervalo de 10, penalidade de extensão de intervalo de 0,5 e a matriz EDNAFULL. A saída de “identidade mais longa” rotulada por Needle

25 é usada como a identidade percentual e é calculada como segue:

(Resíduos Idênticos x 100)/(Comprimento de Alinhamento - Número de Intervalos no Alinhamento)

Seqüência homóloga: O termo “seqüência homóloga” é aqui definido como uma proteína prognosticada que dá um valor E (ou contagem de probabilidade) de menos do que 0,001 em uma pesquisa fasta (Pearson, W. R., 1999, em *Bioinformatics Methods and Protocols*, S. Misener e S. A. Krawetz, ed., pp. 185-219) com a SEQ ID NO: 2 ou SEQ ID NO: 4.

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 do polipeptídeo maduro da SEQ ID NO: 2 ou SEQ ID NO: 4 ou uma seqüência homóloga destas, em que o fragmento tem atividade de endoglucanase. Em um aspecto preferido, um fragmento contém pelo menos 320 resíduos de aminoácido, mais preferivelmente pelo menos 340 resíduos de aminoácido e o mais preferivelmente pelo menos 360 resíduos de aminoácido do polipeptídeo maduro da SEQ ID NO: 2 ou uma seqüência homóloga desta. Em um outro aspecto preferido, um fragmento contém pelo menos 320 resíduos de aminoácido, mais preferivelmente pelo menos 340 resíduos de aminoácido e o mais preferivelmente pelo menos 360 resíduos de aminoácido do polipeptídeo maduro da SEQ ID NO: 4 ou uma seqüência homóloga desta.

Subseqüência: O termo “subseqüência” é aqui definido como uma seqüência de nucleotídeo tendo um ou mais nucleotídeos deletados das extremidades 5’ e/ou 3’ da seqüência que codifica o polipeptídeo maduro da SEQ ID NO: 1 ou SEQ ID NO: 3: ou uma seqüência homóloga desta; em que a subseqüência codifica um fragmento de polipeptídeo tendo atividade de endoglucanase. Em um aspecto preferido, uma subseqüência contém pelo menos 960 nucleotídeos, mais preferivelmente pelo menos 1020 nucleotídeos e o mais preferivelmente pelo menos 1080 nucleotídeos da seqüência que codifica o polipeptídeo maduro da SEQ ID NO: 1 ou uma seqüência homóloga desta, Em um outro aspecto preferido, um subseqüência contém pelo menos 960 nucleotídeos, mais preferivelmente pelo menos 1020 nucleotídeos e o

mais preferivelmente pelo menos 1080 nucleotídeos da sequência que codifica o polipeptídeo maduro da SEQ ID NO: 3 ou uma sequência homóloga desta.

Variante alélica: O termo “variante alélica” denota aqui qualquer uma 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 podem 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 desejados e em uma forma adequada para o uso dentro de sistemas de produção de proteína geneticamente engendradas. 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 está nativa ou recombinantemente associado. Um polinucleotídeo substancialmente puro,

entretanto, pode incluir regiões não traduzidas 5' e 3' que ocorrem naturalmente, 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 seja essencialmente livre de outro material de polinucleotídeo com o qual o mesmo esteja nativa ou recombinantemente associado. Aqui, o termo “polinucleotídeo substancialmente puro” é sinônimo com os termos “polinucleotídeo isolado” e “polinucleotídeo na forma isolada.” Os polinucleotídeos podem ser de origem genômica, cDNA, RNA, semi-sintética, sintética ou quaisquer combinações destes.

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 determinada 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 extremidades com um códon de parada tal como TAA, TAG e TGA. A seqüência codificadora pode ser um DNA, cDNA ou seqüência de nucleotídeo recombinante.

Seqüência codificadora de polipeptídeo maduro: O termo “seqüência codificadora de polipeptídeo maduro” é aqui definido como uma seqüência de nucleotídeo que codifica um polipeptídeo maduro tendo atividade de endoglucanase.

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 qualquer seqüência 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 isolada 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 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 para incluir todos os componentes, que são necessários ou vantajosos para a expressão de um polinucleotídeo que codifica um polipeptídeo da presente invenção. Cada seqüência de controle pode ser nativa ou estranha à seqüência de nucleotídeo que codifica o polipeptídeo ou nativa ou estranha entre si. Tais seqüências de controle incluem, mas não são limitadas a, um líder, seqüência de poliadenilação, seqüência de propeptídeo, promotor, seqüência de peptídeo de sinal e terminador de transcrição. Em um mínimo, as seqüências de controle incluem um promotor e sinais de parada transcricional e traducional. 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 um polipeptídeo.

Operavelmente ligado: O termo “operavelmente ligado” denota 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 seqüência codificadora de um polipeptídeo.

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.

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 presente invenção e que é operavelmente ligado aos nucleotídeos adicionais que fornecem a sua expressão.

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.

Modificação: O termo “modificação” significa aqui qualquer modificação química do polipeptídeo que consiste do polipeptídeo maduro da SEQ ID NO: 2 ou SEQ ID NO: 4; ou uma seqüência homóloga destas; assim como manipulação genética do DNA que codifica um tal 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.

Variante artificial: Quando aqui usado, o termo “variante artificial” significa um polipeptídeo tendo atividade de endoglucanase produzido por um organismo que expressa uma seqüência de nucleotídeo

modificada do polipeptídeo maduro que codifica a seqüência da SEQ ID NO: 1 ou SEQ ID NO: 3; ou uma seqüência homóloga destas. A seqüência de nucleotídeo modificada é obtida através da intervenção humana pela modificação da seqüência de nucleotídeo divulgada na SEQ ID NO: 1 ou SEQ ID NO: 3; ou uma seqüência homóloga destas.

Descrição Detalhada da Invenção

Polipeptídeos Tendo Atividade de Endoglucanase

Em um primeiro aspecto, a presente invenção diz respeito a polipeptídeos isolados compreendendo uma seqüência de aminoácido que tenha um grau de identidade com o polipeptídeo maduro da SEQ ID NO: 2 ou SEQ ID NO: 4 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 ainda mais preferivelmente pelo menos 96%, 97%, 98% ou 99% que tenha atividade de endoglucanase (a seguir “polipeptídeos homólogos”). Em um aspecto preferido, os polipeptídeos homólogos têm uma seqüência 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: 2 ou SEQ ID NO: 4.

Um polipeptídeo da presente invenção preferivelmente compreende a seqüência de aminoácido da SEQ ID NO: 2 ou uma variante alélica desta; ou um fragmento desta que tenha atividade de endoglucanase. Em um aspecto preferido, um polipeptídeo compreende a seqüência de aminoácido da SEQ ID NO: 2. Em um outro aspecto preferido, um polipeptídeo compreende o polipeptídeo maduro da SEQ ID NO: 2. Em um outro aspecto preferido, um polipeptídeo compreende aminoácidos 1 a 42 da

SEQ ID NO: 2. Em um outro aspecto preferido, um polipeptídeo compreende os aminoácidos de 19 a 396 da SEQ ID NO: 2 ou uma variante alélica desta; ou um fragmento desta que tenha atividade de endoglucanase. Em um outro aspecto preferido, um polipeptídeo compreende os aminoácidos de 19 a 396 da SEQ ID NO: 2. Em um outro aspecto preferido, um polipeptídeo consiste da seqüência do aminoácido da SEQ ID NO: 2 ou uma variante alélica desta; ou um fragmento desta que tem atividade de endoglucanase. Em um outro aspecto preferido, um polipeptídeo consiste da seqüência de aminoácido da SEQ ID NO: 2. Em um outro aspecto preferido, um polipeptídeo consiste do polipeptídeo maduro da SEQ ID NO: 2. Em um outro aspecto preferido, um polipeptídeo consiste dos aminoácidos de 19 a 396 da SEQ ID NO: 2 ou uma variante alélica desta; ou um fragmento desta que tenha atividade de endoglucanase. Em um outro aspecto preferido, um polipeptídeo consiste dos aminoácidos de 19 a 396 da SEQ ID NO: 2.

Um polipeptídeo da presente invenção preferivelmente compreende a seqüência de aminoácido da SEQ ID NO: 4 ou uma variante alélica desta; ou um fragmento desta que tenham atividade de endoglucanase. Em um aspecto preferido, um polipeptídeo compreende a seqüência de aminoácido da SEQ ID NO: 4. Em um outro aspecto preferido, um polipeptídeo compreende o polipeptídeo maduro da SEQ ID NO: 4. Em um outro aspecto preferido, um polipeptídeo compreende os aminoácidos de 20 a 400 da SEQ ID NO: 4 ou uma variante alélica desta; ou um fragmento desta que tem atividade de endoglucanase. Em um outro aspecto preferido, um polipeptídeo compreende aminoácidos de 20 a 400 da SEQ ID NO: 4. Em um outro aspecto preferido, um polipeptídeo consiste da seqüência de aminoácido da SEQ ID NO: 4 ou uma variante alélica desta; ou um fragmento desta que tem atividade de endoglucanase. Em um outro aspecto preferido, um polipeptídeo consiste da seqüência de aminoácido da SEQ ID NO: 4. Em um outro aspecto preferido, um polipeptídeo consiste do polipeptídeo maduro da

SEQ ID NO: 4. Em um outro aspecto preferido, um polipeptídeo consiste dos aminoácidos de 20 a 400 da SEQ ID NO: 4 ou uma variante alélica desta; ou um fragmento desta que tenham atividade de endoglucanase. Em um outro aspecto preferido, um polipeptídeo consiste de aminoácidos de 20 a 400 da
5 SEQ ID NO: 4.

Em um segundo aspecto, a presente invenção diz respeito aos polipeptídeos isolados tendo atividade de endoglucanase que são codificados pelos polinucleotídeos que hibridizam sob condições de estringência muito baixa, preferivelmente condições de estringência baixa, mais preferivelmente
10 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 alta com (i) a sequência codificadora do polipeptídeo maduro da SEQ ID NO: 1 ou SEQ ID NO: 3, (ii) a sequência de cDNA contida na sequência codificadora do
15 polipeptídeo maduro da SEQ ID NO: 1 ou a sequência de DNA genômico compreende a sequência codificadora do polipeptídeo maduro da SEQ ID NO: 3, (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
20 Harbor, Nova Iorque). Uma subsequência da sequência codificadora do polipeptídeo maduro da SEQ ID NO: 1 ou SEQ ID NO: 3 contém pelo menos 100 nucleotídeos contíguos ou preferivelmente pelo menos 200 nucleotídeos contíguos. Além disso, a subsequência pode codificar um fragmento de polipeptídeo que tenha atividade de endoglucanase. Em um aspecto preferido
25 a sequência codificadora do polipeptídeo maduro tem os nucleotídeos de 110 a 1557 da SEQ ID NO: 1. Em um outro aspecto preferido, a sequência codificadora do polipeptídeo maduro tem os nucleotídeos de 58 a 1200 da SEQ ID NO: 3. Em um outro aspecto preferido, o filamento complementar é o filamento complementar de tamanho natural da sequência codificadora do

polipeptídeo maduro da SEQ ID NO: 1 ou SEQ ID NO: 3.

A sequência de nucleotídeo da SEQ ID NO: 1 ou SEQ ID NO: 3 ou uma subsequência destas, assim como a sequência de aminoácido da SEQ ID NO: 2 ou SEQ ID NO: 4 ou um fragmento destas, podem ser usadas para planejar uma sonda de ácido nucleico para identificar e clonar polipeptídeos que codificam DNA tendo atividade de endoglucanase a partir de cepas de gêneros ou espécies diferentes de acordo com os métodos bem conhecidos na técnica. Em particular, tais sondas podem ser usadas para a hibridização com o DNA genômico ou cDNA do gênero ou espécie de interesse, seguindo procedimentos de Southern blotting padrão, de modo a identificar e isolar o gene correspondente nisto. Tais sondas podem ser consideravelmente mais curtas do que a sequência inteira, mas devem ter pelo menos 14, preferivelmente pelo menos 25, mais preferivelmente pelo menos 35 e o mais preferivelmente pelo menos 70 nucleotídeos no comprimento. Entretanto, é preferido que a sonda de ácido nucleico tenha pelo menos 100 nucleotídeos no comprimento. Por exemplo, a sonda de ácido nucleico pode ter pelo menos 200 nucleotídeos, preferivelmente pelo menos 300 nucleotídeos, mais preferivelmente pelo menos 400 nucleotídeos ou mais preferivelmente pelo menos 500 nucleotídeos. Sondas ainda mais longas podem ser usadas, por exemplo, sondas de ácido nucleico que tenham pelo menos 600 nucleotídeos, pelo menos preferivelmente pelo menos 700 nucleotídeos, mais preferivelmente pelo menos 800 nucleotídeos ou o mais preferivelmente pelo menos 900 nucleotídeos no comprimento. As 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 genômico ou cDNA preparada a partir de tais outras cepas, portanto, pode ser triada quanto ao DNA que hibridiza com as sondas descritas acima e que codifica um polipeptídeo tendo

atividade de endoglucanase. O DNA genômico ou outro DNA de tais outras cepas pode ser separado pela eletroforese em gel de agarose ou poliacrilamida ou outras técnicas de separação. O DNA das bibliotecas ou o DNA separado podem ser transferido para 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 SEQ ID NO: 3 ou uma subsequência destas, o material carregador é preferivelmente 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 correspondente à sequência codificadora do polipeptídeo maduro da SEQ ID NO: 1 ou SEQ ID NO: 3, a sequência de cDNA contida na sequência codificadora do polipeptídeo maduro da SEQ ID NO: 1 ou a sequência de DNA genômico que compreende a sequência codificadora de polipeptídeo maduro da SEQ ID NO: 3, seu filamento complementar ou uma subsequência destas, sob condições de stringência muito baixa a muito alta. As moléculas com as quais a sonda de ácido nucleico hibridiza sob estas condições podem ser detectadas usando por exemplo, película de raio X.

Em um aspecto preferido, a sonda de ácido nucleico é a sequência codificadora de polipeptídeo maduro da SEQ ID NO: 1. Em um outro aspecto preferido, a sonda de ácido nucleico tem os nucleotídeos de 110 a 1557 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 de polinucleotídeo contida no plasmídeo pPF147 que está contido na *E. coli* NRRL B-30898, em que a sequência de polinucleotídeo desta codifica um polipeptídeo tendo atividade de endoglucanase. Em um outro aspecto preferido, a sonda de ácido nucleico é a região que codifica o polipeptídeo maduro contida no plasmídeo pPH47

que está contido na *E. coli* NRRL B-30898.

Em um aspecto preferido, a sonda de ácido nucleico é a sequência que codifica o polipeptídeo maduro da SEQ ID NO: 3. Em um outro aspecto preferido, a sonda de ácido nucleico tem os nucleotídeos de 58 a 1200 da SEQ ID NO: 3. Em um outro aspecto preferido, a sonda de ácido nucleico é uma sequência de polinucleotídeo que codifica o polipeptídeo da SEQ ID NO: 4 ou uma subsequência desta. Em um outro aspecto preferido, a sonda de ácido nucleico é a SEQ ID NO: 3. Em um outro aspecto preferido, a sonda de ácido nucleico é a sequência de polinucleotídeo contida no plasmídeo pCIBG146 que está contida na *E. coli* NRRL B-30901, em que a sequência de polinucleotídeo desta codifica um polipeptídeo tendo atividade de endoglucanase. Em um outro aspecto preferido, a sonda de ácido nucleico é a região que codifica o polipeptídeo maduro contido no plasmídeo pCIBG146 que está contido na *E. coli* NRRL B-30901.

Para sondas longas de pelo menos 100 nucleotídeos no comprimento, condições de estringência muito baixa a muito alta são definidas como pré hibridização e hibridização a 42°C em 5X SSPE, 0,3% de SDS, 200 µg/ml de DNA de esperma de salmão cisalhado e desnaturado e 25% de formamida para as estringências muito baixas e baixas, 35% de formamida para as estringências médias e médias-altas ou 50% de formamida para as estringências altas e muito altas, seguindo os procedimentos de Southern blotting padrão por 12 a 24 horas de modo ideal.

Para as sondas longas de pelo menos 100 nucleotídeos no comprimento, o material carregador é finalmente lavado três vezes cada 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 são de cerca de 15 nucleotídeos a cerca de 70 nucleotídeos no comprimento, as condições de estringência são definidas como pré hibridização, hibridização e lavagem pós hibridização de 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 por 12 a 24 horas de modo ideal.

Para sondas curtas que são de cerca de 15 nucleotídeos a cerca de 70 nucleotídeos no comprimento, o material carregador é lavado uma vez em 6X SCC mais 0,1% de SDS por 15 minutos e duas vezes cada 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 polipeptídeos isolados codificados pelos polinucleotídeos que compreendem ou que consistem de seqüências de nucleotídeo que tenham um grau de identidade com a seqüência que codifica o polipeptídeo maduro da SEQ ID NO: 1 ou SEQ ID NO: 3 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 que codifica o polipeptídeo maduro tem os nucleotídeos de 110 a 1557 da SEQ ID NO: 1. Em um outro aspecto preferido, a seqüência que codifica o polipeptídeo maduro tem os nucleotídeos de 58 a 1200 da SEQ ID NO: 3. Ver a seção polinucleotídeo aqui.

Em um quarto aspecto, a presente invenção diz respeito a variantes artificiais que compreendem uma substituição, deleção e/ou inserção de um ou mais (ou de vários) aminoácidos do polipeptídeo maduro da SEQ ID NO: 2 ou SEQ ID NO: 4; ou uma sequência homóloga desta.

5 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;
 10 um peptídeo ligador pequeno de até cerca de 20 a 25 resíduos; ou uma extensão pequena que facilita a purificação pela mudança da carga líquida ou uma outra função, tal como um trecho de poli-histidina, um epítopo antigênico ou um domínio de ligação.

Os exemplos de substituições conservativas estão dentro do
 15 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
 20 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. A mudanças que ocorrem mais habitualmente 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,
 25 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 dos 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 dos resíduos de aminoácido. Os “aminoácidos não naturais” foram modificados depois da síntese da proteína e/ou têm uma estrutura química na(s) sua(s) cadeia(s) lateral(is) diferente 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 físico químicas dos polipeptídeos são 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 a mutagênese direcionada ao sítio ou mutagênese de varredura de alanina (Cunningham e Wells, 1989, Science 244: 1081-1085). Na última técnica, as mutações de alanina únicas são introduzidas em cada resíduo na
20 molécula e as moléculas mutantes resultantes são testadas quanto a atividade biológica (isto é, a atividade de endoglucanase) 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 podem ser determinados pela análise física da
25 estrutura, como determinada 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 com um polipeptídeo de acordo com a invenção.

As substituições, deleções e/ou inserções de aminoácido únicas ou múltiplas podem ser feitas e testadas usando métodos conhecidos de mutagênese, recombinação e/ou embaralhamento, seguido por um procedimento de triagem relevante, tal como aqueles divulgados por Reidhaar-Olson e Sauer, 1988, *Science* 241: 53-57; Bowie e Sauer, 1989, *Proc. Natl. Acad. Sci. USA* 86: 2152-2156; WO 95/17413; ou WO 95/22625.

Outros métodos que podem ser usados incluem a 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 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 recuperadas das células hospedeiras e rapidamente 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 aminoácido do polipeptídeo maduro da SEQ ID NO: 2 ou SEQ ID NO: 4, tais como os aminoácidos de 19 a 396 da SEQ ID NO: 2 ou os aminoácidos de 20 a 400 da SEQ ID NO: 4, 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 Endoglucanase

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 dada fonte deve significar que o polipeptídeo codificado por uma seqüência de nucleotídeo é produzida 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 extracelularmente secretado.

Um polipeptídeo tendo atividade de endoglucanase 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*, *Streptococcus*, *Streptomyces*, *Staphylococcus*, *Enterococcus*, *Lactobacillus*, *Lactococcus*, *Clostridium*, *Geobacillus* ou *Oceanobacillus* tendo atividade de endoglucanase ou um polipeptídeo bacteriano Gram negativo tal como um polipeptídeo de *E. coli*, *Pseudomonas*, *Salmonella*, *Campylobacter*, *Helicobacter*, *Flavobacterium*, *Fusobacterium*, *Ilyobacter*, *Neisseria* ou *Ureaplasma* tendo atividade de endoglucanase.

Em um aspecto preferido, o polipeptídeo é um polipeptídeo de *Bacillus alkalophilus*, *Bacillus amyloliquefaciens*, *Bacillus brevis*, *Bacillus circularis*, *Bacillus clausii*, *Bacillus coagulans*, *Bacillus finnis*, *Bacillus lautus*, *Bacillus lentus*, *Bacillus licheniformis*, *Bacillus megaterium*, *Bacillus pumilus*, *Bacillus stearothermophilus*, *Bacillus subtilis* ou *Bacillus thuringiensis* tendo atividade de endoglucanase.

Em um outro aspecto preferido, o polipeptídeo é um polipeptídeo de *Streptococcus equisimilis*, *Streptococcus pyogenes*, *Streptococcus uberis* ou *Streptococcus equi* subsp. *Zooepidemicus* tendo atividade de endoglucanase.

Em um outro aspecto preferido, o polipeptídeo é um polipeptídeo de *Streptomyces achromagenes*, *Streptomyces avermitilis*, *Streptomyces coelicolor*, *Streptomyces griseus* ou *Streptomyces lividans* tendo atividade de endoglucanase.

5 Um polipeptídeo tendo atividade de endoglucanase 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 endoglucanase; ou mais preferivelmente um polipeptídeo
10 fúngico filamentoso tal como um *Acremanium*, *Aspergillus*, *Aureobasidium*, *Chrysosporium*, *Cryptococcus*, *Filibasidium*, *Fusarium*, *Humicola*, *Magnaporthe*, *Mucor*, *Myceliophthora*, *Neocallimastix*, *Neurospora*, *Paecilomyces*, *Penicillium*, *Piromyces*, *Schizophyllum*, *Talaromyces*, *Thermoascus*, *Thielavia*, *Tolypocladium* ou *Trichoderma* tendo atividade de
15 endoglucanase.

Em um aspecto preferido, o polipeptídeo é um polipeptídeo de *Saccharomyces carlsbergensis*, *Saccharomyces cerevisiae*, *Saccharomyces diastaticus*, *Saccharomyces douglasii*, *Saccharomyces kluyveri*, *Saccharomyces norbensis* ou *Saccharomyces oviformis* tendo atividade de
20 endoglucanase.

Em um outro aspecto preferido, o polipeptídeo é um polipeptídeo de *Aspergillus aculeatus*, *Aspergillus awamori*, *Aspergillus fumigatus*, *Aspergillus foetidus*, *Aspergillus japonicus*, *Aspergillus nidulans*, *Aspergillus niger*, *Aspergillus oryzae*, *Chrysosporium keratinophilum*,
25 *Chrysosporium lucknowense*, *Chrysosporium tropicum*, *Chrysosporium merdarium*, *Chrysosporium mops*, *Chrysosporium pannicola*, *Chrysosporium queenslandicum*, *Chrysosporium zonatum*, *Fusarium bactridioides*, *Fusarium cerealis*, *Fusarium crookwellense*, *Fusarium culmorum*, *Fusarium graminearum*, *Fusarium graminum*, *Fusarium heterosporum*, *Fusarium*

negundi, *Fusarium oxisporum*, *Fusarium reticulatum*, *Fusarium roseum*,
Fusarium sambucinum, *Fusarium sarcochroum*. *Fusarium sporotrichioides*,
Fusarium sulphureum, *Fusarium torulosum*, *Fusarium trichothecioides*,
Fusarium venenatum, *Humicola insolens*, *Humicola lanuginosa*, *Mucor*
5 *miehei*, *Myceliophthora thermophila*, *Neurospora crassa*, *Penicillium*
brasilianum, *Penicillium camembertii*, *Penicillium capsulatum*, *Penicillium*
chrysogenum, *Penicillium citreonigrum*, *Penicillium citrinum*, *Penicillium*
claviforme, *Penicillium corylophitum*, *Penicillium crustosum*, *Penicillium*
digitatum, *Penicillium expansum*, *Penicillium funiculosum*. *Penicillium*
10 *giabrum*, *Penicillium granulatum*, *griseofulvum*. *Penicillium islandicum*,
Penicillium italicum, *Penicillium janthinellum*, *Penicillium lividum*,
Penicillium megasporum, *Penicillium melinii*, *Penicillium notatum*,
Penicillium oxalicum, *Penicillium puberulum*, *Penicillium purpurescens*,
Penicillium purpurogenum, *Penicillium roquefortii*, *Penicillium rugulosum*,
15 *Penicillium spinulosum*, *Penicillium waksmanii*, *Trichoderma harzianum*,
Trichoderma koningii, *Trichoderma longibrachiatum*. *Trichoderma reesei* ou
Trichoderma viride tendo atividade de endoglucanase.

Em um outro aspecto preferido, o polipeptídeo é um
 polipeptídeo de *Thielavia achromatica*, *Thielavia albomyces*, *Thielavia*
 20 *albopilosa*, *Thielavia australeinsis*, *Thielavia fimeti*, *Thielavia microspota*,
Thielavia ovispora, *Thielavia peniviana*, *Thielavia spededonium*, *Thielavia*
setosa, *Thielavia subthermophila*, *Thielavia terrestris*, *Thielavia terricola*,
Thielavia thermophila, *Thielavia variospora* ou *Thielavia wareingii* tendo
 atividade de endoglucanase.

25 Em um aspecto mais preferido, o polipeptídeo é um
 polipeptídeo de *Thielavia terrestris* e o mais preferivelmente um polipeptídeo
 de *Thielavia terrestris* NRRL 8126, por exemplo, o polipeptídeo da SEQ ID
 NO: 2 ou SEQ ID NO: 4 ou o polipeptídeo maduro destas.

Será entendido que para as espécies anteriormente

mencionadas, a invenção abrange os estados tanto perfeitos quanto imperfeitos 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.

As 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 and Zellkulturen GmbH (DSM), Centraalbureau Voor Schimmelcultures (CBS) e Agricultural Research Service Patent Culture Collection, Northern Regional Research Center (NRRL).

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 conhecidas na técnica. O polinucleotídeo pode ser depois obtido triando-se similarmente uma biblioteca genômica ou de cDNA de um tal microorganismo. Uma vez que uma seqüência de polinucleotídeo que codifica um polipeptídeo tenha sido 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 exemplo, Sambrook *et al.*, 1989, *supra*).

Os polipeptídeos da presente invenção também incluem polipeptídeos fundidos ou polipeptídeos de fusão clivados em que um outro polipeptídeo é fundido no término N ou no término C do polipeptídeo ou fragmento deste. 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 na técnica e incluem ligar as seqüências codificadoras que codificam os polipeptídeos de modo que eles estejam na matriz e que a expressão do polipeptídeo fundido esteja sob o controle dos mesmos promotor(es) e terminador.

5 Um polipeptídeo de fusão pode compreender ainda um sítio de clivagem. Na secreção da proteína de fusão, o sítio é clivado liberando o polipeptídeo tendo atividade de endoglucanase da proteína de fusão.

Os exemplos de sítios de clivagem incluem, mas não são limitados a, um sítio Kex2 que codifica o dipeptídeo Lys-Arg (Martin *et al* 10 2003, J. Ind. Microbiol. Biotechnol 3: 568-76; Svetina *et al*, 2000, J. Biotechnol 76: 245-251; Rasmussen-Wilson *et al*, 1997, App. Environ, Microbiol. 63: 3488-3493; Ward *et al.*, 1995, Biotechnology 13: 498-503; e Contreras *et al*, 1991, Biotechnology 9: 378-381), um sítio Ile-(Glu ou Asp)-Gly-Arg, que é clivado por uma protease de Fator Xa depois do resíduo de arginina (Eaton *et al*, 1966, Biochem. 25: 505-512); um sítio Asp-Asp-Asp-Asp-Lys, que é clivado por uma enterocinase depois da lisina (Collins-Racie *et al.*, 1995, Biotechnology 13: 982-987); um sítio His-Tyr-Glu ou sítio His-Tyr-Asp, que é clivado pela Genenase I (Carter *et al.*, 1989, Proteins: Structure, Function and Genetics 6: 240-248); um sítio Leu-Val-Pro-Arg-Gly-Ser, que é clivado pela trombina depois do Arg (Stevens, 2003, Drug 20 Discovery World 4: 3548); um sítio Glu-Asn-Leu-Tyr-Phe-Gln-Gly, que é clivado pela TEV protease depois do Gln (Stevens, 2003, supra); e um sítio Leu-Glu-Val-Leu-Phe-Gln-Gly-Pro, que é clivado por uma forma geneticamente engendrada de rinovírus 3C humano protease depois do Gln 25 (Stevens, 2003, supra).

Polinucleotídeos

A presente invenção também diz respeito a um polinucleotídeo isolado que compreende ou consiste de uma seqüência de nucleotídeo que codifica um polipeptídeo da presente invenção tendo atividade de

endoglucanase.

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
5 contida no plasmídeo pPH47 que está contido na *E. coli* NRRL 8-30898. Em um outro aspecto preferido, a seqüência de nucleotídeo compreende ou consiste da região que codifica o polipeptídeo maduro da SEQ ID NO: 1. Em um outro aspecto preferido, a seqüência de nucleotídeo compreende ou consiste dos nucleotídeos de 110 a 1557 da SEQ ID NO: 1. Em um outro
10 aspecto mais preferido, a seqüência de nucleotídeo compreende ou consiste da região que codifica o polipeptídeo maduro contido no plasmídeo pPH47 que está contido na *E. coli* NRRL 8-30898. A presente invenção também abrange as seqüências de nucleotídeo que codificam os polipeptídeos que compreendem ou que consistem da seqüência de aminoácido da SEQ ID NO:
15 2 ou o polipeptídeo maduro desta, que difere da SEQ ID NO: 1 ou a seqüência que codifica o polipeptídeo maduro desta em virtude da degenerescência do código genético. A presente invenção também diz respeito às subsequências da SEQ ID NO: 1 que codificam fragmentos da SEQ ID NO: 2 que têm atividade de endoglucanase.

20 Em um outro aspecto preferido, a seqüência de nucleotídeo compreende ou consiste da SEQ ID NO: 3. Em um outro aspecto mais preferido, a seqüência de nucleotídeo compreende ou consiste da seqüência contida no plasmídeo pClBG146 que está contido na *E. coli* NRRL B-30901. Em um outro aspecto preferido, a seqüência de nucleotídeo compreende ou
25 consiste do polipeptídeo maduro que codifica a região da SEQ ID NO: 3. Em um outro aspecto preferido, a seqüência de nucleotídeo compreende ou consiste dos nucleotídeos de 58 a 1200 da SEQ ID NO: 3. Em um outro aspecto mais preferido, a seqüência de nucleotídeo compreende ou consiste da região codificadora do polipeptídeo maduro contida no plasmídeo pClBG146

que está contido na *E. coli* NRRL B-30901. A presente invenção também abrange seqüências de nucleotídeo que codificam polipeptídeos que compreendem ou que consistem da seqüência de aminoácido da SEQ ID NO: 4 ou o polipeptídeo maduro desta, que difere da SEQ ID NO: 3 ou a seqüência
5 que codifica o polipeptídeo maduro desta em virtude da degenerescência do código genético. A presente invenção também diz respeito às subsequências da SEQ ID NO: 3 que codificam fragmentos da SEQ ID NO: 4 que têm atividade de endoglucanase.

A presente invenção também diz respeito a polinucleotídeos
10 mutantes que compreendem ou consistem de pelo menos uma mutação na seqüência que codifica o polipeptídeo maduro da SEQ ID NO: 1 ou SEQ ID NO: 3, em que a seqüência de nucleotídeo mutante codifica o polipeptídeo maduro da SEQ ID NO: 2 ou SEQ ID NO: 4. Em um aspecto preferido, o polipeptídeo maduro tem os aminoácidos de 19 a 396 da SEQ ID NO: 2. Em
15 um outro aspecto preferido, o polipeptídeo maduro tem os aminoácidos de 20 a 400 da SEQ ID NO: 4.

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 destes. A
20 clonagem dos polinucleotídeos da presente invenção a partir 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 clonado com características estruturais compartilhadas. Ver, por exemplo, Innis *et al.*, 1990, PCR: A
25 Guide to Methods and Applications, Academic Press, Nova Iorque. Outros procedimentos de amplificação de ácido nucleico tais como a reação da cadeia de ligase (LCR), transcrição ativada ligada (LAT) e amplificação com base na seqüência de nucleotídeo (NASBA) podem ser usados. Os polinucleotídeos podem ser clonados a partir de uma cepa de *Myceliophthora*

termomophila CBS 117.65, basidiomiceto CBS 494.95 ou basidiomiceto CBS 495.95 ou um outro organismo ou organismo relacionado e assim, por exemplo, pode ser uma variante alélica ou de espécie variante da região que codifica o polipeptídeo da sequência de nucleotídeo.

5 A presente invenção também diz respeito aos polinucleotídeos que compreendem ou consistem das sequências de nucleotídeo que têm um grau de identidade com a sequência que codifica o polipeptídeo maduro da SEQ ID NO: 1 ou SEQ ID NO: 3 de pelo menos 60%, preferivelmente pelo menos 65%, mais preferivelmente pelo menos 70%, mais preferivelmente
10 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 sequência que codifica o polipeptídeo maduro tem os
15 nucleotídeos de 110 a 1557 da SEQ ID NO: 1. Em um outro aspecto preferido a sequência que codifica o polipeptídeo maduro tem os nucleotídeos de 58 a 1200 da SEQ ID NO: 3.

 A modificação de uma sequência de nucleotídeo que codifica um polipeptídeo da presente invenção pode ser necessária para a síntese de
20 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 algum modo engendrado a partir do polipeptídeo isolado da sua fonte nativa, por exemplo, variantes artificiais que diferem em atividade específica,
25 termoestabilidade, pH ótimo ou semelhante. A sequência variante pode ser construída com base na sequência de nucleotídeo apresentada como a região que codifica o polipeptídeo da SEQ ID NO: 1 ou SEQ ID NO: 3, por exemplo, uma subsequência destas e/ou pela introdução de substituições de nucleotídeo que não dão origem a uma outra sequência de aminoácido do

polipeptídeo codificado pela seqüência de nucleotídeo, mas que corresponde ao uso de códon do organismo hospedeiro intencionado para a produção da enzima ou pela introdução de substituições de nucleotídeo que podem dar origem a uma seqüência de aminoácido diferentes. 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 molécula e ainda resulta 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, pode ser identificado 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, *supra*). Na última técnica, as mutações são introduzidas em cada resíduo positivamente carregado na molécula e as moléculas mutantes resultantes são testadas quanto a atividade de endoglucanase para identificar resíduos de aminoácido que são críticos para a atividade da molécula. Os sítios de interação de substrato-enzima também podem ser determinados pela análise da estrutura tridimensional como determinado por técnicas tais como a análise de ressonância magnética nuclear, cristalografia ou rotulação de fotoafinidade (ver, por exemplo, de Vos *et al.*, 1992, *supra*; Smith *et al.*, 1992, *supra*; Wlodaver *et al.*, 1992, *supra*).

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 alta com sob pelo menos condições de estringência alta com (i) a seqüência que codifica o polipeptídeo maduro da SEQ ID NO: 1 ou SEQ ID NO: 3, (ii) a seqüência de cDNA contida na seqüência que codifica o polipeptídeo maduro da SEQ ID NO: 1 ou a seqüência de DNA genômico que compreende a seqüência que codifica o polipeptídeo maduro da SEQ ID NO: 3, (iii) uma subsequência de (i) ou (ii) ou (iv) um filamento complementar de (i) ou (ii); ou variantes alélicas destas (Sambrook *et al.*, 1989, *supra*), como aqui definido. Em um aspecto preferido, a seqüência que codifica o polipeptídeo maduro tem os nucleotídeos de 110 a 1557 da SEQ ID NO: 1. Em um outro aspecto preferido, a seqüência que codifica o polipeptídeo maduro tem os nucleotídeos de 58 a 1200 da SEQ ID NO: 3. Em um outro aspecto preferido, o filamento complementar é o filamento complementar de tamanho natural da seqüência que codifica o polipeptídeo maduro da SEQ ID NO: 1 ou SEQ ID NO: 3.

A presente invenção também diz respeito a polinucleotídeos isolados obtidos pela (a) hibridização de uma população de DNA sob condições de estringência muito baixa, baixa, média, média-alta, alta ou muito alta com (i) a seqüência que codifica o polipeptídeo maduro da SEQ ID NO: 1 ou SEQ ID NO: 3, (ii) a seqüência de cDNA contida na seqüência que codifica o polipeptídeo maduro da SEQ ID NO: 1 ou o a seqüência de DNA genômico que compreende a seqüência codificadora do polipeptídeo maduro da SEQ ID NO: 3 ou (iii) um filamento complementar de (i) ou (ii); e (b) isolar o polinucleotídeo de hibridização, que codifica um polipeptídeo tendo atividade de endoglucanase. Em um aspecto preferido, a seqüência que codifica o polipeptídeo maduro tem os nucleotídeos de 110 a 1557 da SEQ ID NO: 1. Em um outro aspecto preferido, a seqüência que codifica o polipeptídeo maduro tem os nucleotídeos de 58 a 1200 da SEQ ID NO: 3. Em um outro aspecto preferido, o filamento complementar é o filamento

complementar de tamanho natural da sequência que codifica o polipeptídeo maduro da SEQ ID NO: 1 ou SEQ ID NO: 3.

Construções de ácido nucleico

5 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 sequências de controle que direcionam a expressão da sequência codificadora em uma célula hospedeira adequada sob condições compatíveis com as sequências de controle.

10 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 sequência do polinucleotídeo antes da sua inserção em um vetor pode ser desejável ou necessária dependendo do vetor de expressão. As técnicas para modificar as sequências de polinucleotídeo utilizando métodos de DNA recombinante são bem conhecidos na técnica.

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

25 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*, do gene da agarase de *Streptomyces*

coelicolor (dagA), do gene da levansacarase de *Bacillus subtilis* (sacB), do gene da alfa-amilase de *Bacillus licheniformis* (amil), do gene da amilase maltogênica de *Bacillus stearothermophilus* (amyM), do gene da alfa-amilase de *Bacillus amiloliquefaciens* (amyQ), do gene da penicilinase de *Bacillus licheniformis* (penP), genes da *xylA* e *xylB* de *Bacillus subtilis* e do gene da beta-lactamase procariótica (Villa-Kamaroff *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: 2125). Outros promotores são descritos em “Useful proteins from recombinant bacteria” no 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 amilase TAKA de *Aspergillus oryzae*, proteinase aspártica de *Rhizomucor miehei*, alfa-amilase neutra de *Aspergillus niger*, alfa-amilase estável em ácido de *Aspergillus niger*, glicoamilase de *Aspergillus niger* ou *Aspergillus awamori* (glaA), lipase de *Rhizomucor miehei*, protease alcalina de *Aspergillus oryzae*, triose fosfato isomerase de *Aspergillus oryzae*, acetamidase de *Aspergillus nidulans*, amiloglucosidase de *Fusarium venenatum* (WO 00/56900), Daria de *Fusarium venenatum* (WO 00/56900), Quinn de *Fusarium venenatum* (WO 00/56900), protease equivalente a tripsina de *Fusarium oxysporum* (WO 96/00787), beta-glucosidase de *Trichoderma reesei*, celobioidrolase I de *Trichoderma reesei*, celobioidrolase II de *Trichoderma reesei*, endoglucanase I de *Trichoderma reesei*, endoglucanase II de *Trichoderma reesei*, endoglucanase III de *Trichoderma reesei*, endoglucanase IV de *Trichoderma reesei*, endoglucanase 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 mutantes, truncados e híbridos destes.

Em um hospedeiro de levedura, os promotores úteis são
5 obtidos dos genes para a enolase de *Saccharomyces cerevisiae* (ENO-1), galactocinase de *Saccharomyces cerevisiae* (GAL1), álcool desidrogenase/gliceraldeído-3-fosfato desidrogenase de *Saccharomyces cerevisiae* (ADH1,ADH2/GAP), triose fosfato isomerase de *Saccharomyces cerevisiae* (TPI), metalotionina de *Saccharomyces cerevisiae* (CUP1) e 3-
10 fosfoglicerato cinase 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
15 célula hospedeira para terminar a transcrição. A seqüência terminadora está operavelmente ligada ao término 3' da seqüência de nucleotídeo que codifica o polipeptídeo. Qualquer terminado que seja funcional na célula hospedeira de escolha pode ser usado na presente invenção.

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

Os terminadores preferidos para as células hospedeiras de
25 levedura são obtidos a partir dos genes para a enolase de *Saccharomyces cerevisiae*, citocromo C de *Saccharomyces cerevisiae* (CYC1) e gliceraldeído-3-fosfato desidrogenase de *Saccharomyces cerevisiae*. Outros terminadores úteis para as células hospedeiras de levedura são descritos 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 seja importante para a tradução pela célula hospedeira. A seqüência líder é operavelmente ligada ao término 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 de fungo filamentoso são obtidos a partir dos genes para a amilase TAKA de *Aspergillus oryzae* e triose fosfato isomerase de *Aspergillus nidulans*.

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

A seqüência de controle também pode ser uma seqüência de poliadenilação, uma seqüência operavelmente ligada ao término 3' da seqüência de nucleotídeo e que, quando transcrita, é reconhecida pela célula hospedeira como um sinal para a adição de 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 fúngicas filamentosas são obtidas a partir dos genes para a amilase TAKA de *Aspergillus oryzae*, glicoamilase de *Aspergillus niger*, antranilato sintase de *Aspergillus nidulans*, protease equivalente a tripsina de *Fusarium oxysporum* e alfa-glucosidase de *Aspergillus niger*.

As seqüências de poliadenilação úteis para as células hospedeiras de levedura são descritas por Guo e Sherman, 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 término amino de um polipeptídeo e direciona o polipeptídeo codificado dentro do caminho secretor da célula. A extremidade 5' da seqüência codificadora da seqüência de nucleotídeo pode inerentemente
5 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
10 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
15 região codificadora de peptídeo de sinal que direcione o polipeptídeo expressado dentro do caminho secretor de uma célula hospedeira de escolha, isto é secretado em um meio de cultura, pode ser usada na presente invenção.

As regiões codificadoras de peptídeo de sinal eficazes para as células hospedeiras bacterianas são as regiões codificadoras de peptídeo de
20 sinal obtidas a partir dos genes para a amilase maltogênica NCIB 11837 de *Bacillus*, alfa-amilase de *Bacillus stearothermophilus*, subtilisina de *Bacillus licheniformis*, beta-lactamase de *Bacillus licheniformis*, proteases neutras de *Bacillus stearothermophilus* (nprT, nprS, nprM) e prsA de *Bacillus subtilis*. Outros peptídeos de sinal são descritos por Simonen e Palva, 1993,
25 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 amilase TAKA de *Aspergillus oryzae*, amilase neutra de *Aspergillus niger*, glicoamilase de

Aspergillus niger, proteinase aspártica de *Rhizomucor miehei*, celulase de *Humicola insolens*, endoglucanase V de *Humicola insolens* e lipase de *Humicola lanuginosa*.

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 descritos por Romanos *et al.*, 1992, *supra*.

Em um aspecto preferido, o peptídeo de sinal compreende ou consiste dos aminoácidos de 1 a 18 da SEQ ID NO: 2. Em um outro aspecto preferido, a região codificadora de polipeptídeo de sinal tem os nucleotídeos de 56 a 109 da SEQ ID NO: 1.

Em um outro aspecto preferido, o peptídeo de sinal compreende ou consiste dos aminoácidos de 1 a 19 da SEQ ID NO: 4. Em um outro aspecto preferido, a região codificadora de polipeptídeo de sinal compreende ou consiste dos nucleotídeos de 1 a 57 da SEQ ID NO: 3.

A seqüência de controle também pode ser uma região codificadora de pró-peptídeo que codifica uma seqüência de aminoácido posicionada no término amino de um polipeptídeo. O polipeptídeo resultante é conhecido como uma pró enzima ou pró-polipeptídeo (ou um zimogeno em alguns casos). Um pró-polipeptídeo é no geral inativo e pode ser convertido a um polipeptídeo ativo maduro pela clivagem catalítica ou auto-catalítica do pró-peptídeo a partir do pró-polipeptídeo. A região codificadora de pró-peptídeo 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 pró-peptídeo estão presentes no término amino de um polipeptídeo, a região de

pró-peptídeo é posicionada a seguir do terminal amino de um polipeptídeo e a região de peptídeo de sinal é posicionada a seguir do terminal amino da região de pró-peptídeo.

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

Vetores de Expressão

A presente invenção também diz respeito a vetores de expressão recombinantes que compreendem um polinucleotídeo da presente invenção, um promotor e sinais de parada transcricional e traducional. Os vários ácidos nucleicos e seqüências de controle aqui descritos podem ser unidos entre si para produzir um vetor de expressão recombinante que pode 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 polinucleotídeo da presente

invenção pode ser expressada inserindo-se a 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 seqüência codificadora está localizada no vetor de modo que a seqüência codificadora seja operavelmente ligada com as seqüências de controle apropriadas para a expressão.

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

O vetor pode ser um vetor que replique autonomamente, isto é, um vetor que existe como uma entidade extracromossômica, a replicação do qual é independente da replicação cromossômica, por exemplo, um 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(s) qual(is) o mesmo foi integrado. Além disso, um único vetor ou plasmídeo ou dois ou mais vetores ou plasmídeos que juntos contêm o DNA total a ser introduzido no genoma da célula hospedeira ou um transposon podem ser usados.

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

outras.

Os exemplos de marcadores bacterianos selecionáveis são os genes *dal* de *Bacillus subtilis* ou *Bacillus licheniformis* ou marcadores que conferem resistência a antibiótico tal como resistência à 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 filamentosa fúngica incluem, mas não são limitados a, *amdS* (acetamidase), *argB* (ornitina carbamoil-transferase), *bar* (fosfinotricina acetiltransferase), *hph* (higromicina fosfotransferase), *niaD* (nitrato reductase), *pyrG* (orotidina-5-fosfato descarboxilase), *sC* (sulfato 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 a replicação autônoma do vetor na célula independente do genoma.

Para a integração no genoma da célula hospedeira, o vetor pode contar com a seqüência do polinucleotídeo que codifica o polipeptídeo ou qualquer outro elemento do vetor para a integração no genoma pela recombinação homóloga ou não homóloga. Alternativamente, o vetor pode conter seqüências de nucleotídeo adicionais para direcionar a integração pela recombinação homóloga no genoma da célula hospedeira em um local ou locais exatos no(s) cromossoma(s). Para aumentar a probabilidade de integração em um local exato, os elementos integracionais devem preferivelmente conter um número suficiente de ácidos nucleicos, tais como 100 a 10.000 pares de base, preferivelmente de 400 a 10.000 pares de base e o

mais preferivelmente de 800 a 10.000 pares de base, que tenham um alto grau de identidade com a seqüência alvo correspondente para realçar a probabilidade de recombinação homóloga. Os elementos integracionais podem ser qualquer seqüência que seja homóloga com a seqüência alvo no
5 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
10 uma origem de replicação que permita que o vetor replique autonomamente na célula hospedeira em questão. A origem de replicação pode ser qualquer plasmídeo replicador que medeie a replicação autônoma que funcione em uma célula. O termo “origem de replicação” ou “plasmídeo replicador” é aqui definido como uma seqüência de nucleotídeo que permite que um plasmídeo
15 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 pAM β 1 que permitem a replicação em *Bacillus*.

20 Os exemplos de origens de replicação para o uso em uma célula hospedeira de levedura são a origem de 2 microns 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 ANS1 (Gems *et al.*, 1991, Gene 98: 61-67;
25 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 realizados de acordo com os métodos divulgados na WO 00/24883.

Mais do que uma cópia de um polinucleotídeo da presente

invenção pode ser inserida em uma 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 incluindo-se um gene marcador selecionável amplificável com o polinucleotídeo onde as células contendo cópias amplificadas do gene marcador selecionável e deste modo 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*).

Células hospedeiras

A célula hospedeira pode ser qualquer célula útil na produção recombinante de um polipeptídeo da presente invenção, por exemplo, um procariota ou um eucariota.

A célula hospedeira procariótica pode ser qualquer bactéria Gram positiva ou uma bactéria Gram negativa. As bactérias Gram positivas incluem, mas não são limitadas a, *Bacillus*, *Streptococcus*, *Streptomyces*, *Staphylococcus*, *Enterococcus*, *Lactobacillus*, *Lactococcus*, *Clostridium*, *Geobacillus* e *Oceanobacillus*. As bactérias Gram negativas incluem, mas não são limitadas a, *E. coli*, *Pseudomonas*, *Salmonella*, *Campylobacter*, *Helicobacter*, *Flavobacterium*, *Fusobacterium*, *Llyobacter*, *Neisseria* e *Ureaplasma*.

A célula hospedeira bacteriana pode ser qualquer célula de *Bacillus*. As células de *Bacillus* úteis na prática da presente invenção incluem, mas não são limitadas a, células de *Bacillus alkalophilus*, *Bacillus amiloliquefaciens*, *Bacillus brevis*, *Bacillus circulans*, *Bacillus clausii*, *Bacillus coagulans*, *Bacillus firmus*, *Bacillus lautus*, *Bacillus lentus*, *Bacillus*

licheniformis, *Bacillus megaterium*, *Bacillus pumilus*, *Bacillus stearothermophilus*, *Bacillus subtilis* e *Bacillus thuringiensis*.

Em um aspecto preferido, a célula hospedeira bacteriana é uma célula de *Bacillus amiloliquefaciens*, *Bacillus lentus*, *Bacillus licheniformis*,
 5 *Bacillus stearothermophilus* ou *Bacillus subtilis*. Em um aspecto mais preferido, a célula hospedeira bacteriana é uma célula de *Bacillus amiloliquefaciens*. em um outro aspecto mais preferido, a célula hospedeira bacteriana é uma célula de *Bacillus clausii*. Em um outro aspecto mais preferido, a célula hospedeira bacteriana é uma célula de *Bacillus*
 10 *licheniformis*. Em um outro aspecto mais preferido, a célula hospedeira bacteriana é uma célula de *Bacillus subtilis*.

A célula hospedeira bacteriana também pode ser qualquer célula de *Streptococcus*. As células de *Streptococcus* úteis na prática da presente invenção incluem, mas não são limitadas a, *Streptococcus*
 15 *equisimilis*, *Streptococcus pyogenes*, *Streptococcus uberis* e *Streptococcus equi* subsp. *Zooepidemicus*.

Em um aspecto preferido, a célula hospedeira bacteriana é uma célula de *Streptococcus equisimilis*. Em um outro aspecto preferido, a célula hospedeira bacteriana é uma célula de *Streptococcus pyogenes*. Em um outro
 20 aspecto preferido, a célula hospedeira bacteriana é uma célula de *Streptococcus uberis*. Em um outro aspecto preferido, a célula hospedeira bacteriana é uma célula de *Streptococcus equi* subsp. *Zooepidemicus*.

A célula hospedeira bacteriana também pode ser qualquer célula de *Streptomyces*. As células de *Streptomyces* útil na prática da presente invenção incluem, mas não são limitadas a, *Streptomyces achromogenes*,
 25 *Streptomyces avermitilis*, *Streptomyces coelicolor*, *Streptomyces griseus* e *Streptomyces lividans*.

Em um aspecto preferido, a célula hospedeira bacteriana é uma célula de *Streptomyces achromogenes*. Em um outro aspecto preferido a

célula hospedeira bacteriana é uma célula de *Streptomyces avermitilis*. Em um outro aspecto preferido, a célula hospedeira bacteriana é uma célula de *Streptomyces coelicolor*. Em um outro aspecto preferido, a célula hospedeira bacteriana é uma célula de *Streptomyces griseus*. Em um outro aspecto preferido, a célula hospedeira bacteriana é uma célula de *Streptomyces lividans*.

A introdução de DNA em uma célula de *Bacillus*, por exemplo, pode ser efetuada pela transformação de protoplasto (ver, por exemplo, Chang e Cohen, 1979, Molecular General Genetics 168: 111-115), pelo uso de 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), por eletroporação (ver, por exemplo, Shigekawa e Dower, 1988, Biotechniques 6, 742-751) ou pela conjugação (ver, por exemplo, Koehler e Theme, 1987, Journal of Bacteriology 169: 5271-5278). A introdução de DNA em uma célula de *E. coli*, por exemplo, pode ser efetuada pela transformação de protoplasto (ver, por exemplo, Hanahan, 1983, J. Mol. Biol. 166: 557-580) ou eletroporação (ver, por exemplo, Dower *et al.*, 1988, Nucleic Acids Res. 16: 6127-6145). A introdução de DNA em uma célula de *Streptomyces*, por exemplo, pode ser efetuada pela transformação de protoplasto e eletroporação (ver, por exemplo, Gong *et al.*, 2004, Folic Microbial. (Praha) 49: 399-405), pela conjugação (ver, por exemplo, Mazodier *et al.*, 1989, J. Bacteriol. 171: 3583-3585) ou pela transdução (ver, por exemplo, Burke *et al.*, 2001, Proc. Natl. Acad. Sci. USA 98: 6289-6294). A introdução de DNA em uma célula de *Pseudomonas*, por exemplo, pode ser efetuada pela eletroporação (ver, por exemplo, Choi *et al.*, 2006, J. Microbial. Methods 64: 391-397) ou pela conjugação (ver, por exemplo, Pinedo e Smets, 2005, Appl. Environ. Microbiol. 71: 51-57). A introdução de DNA em uma célula de *Streptococcus*, por exemplo, pode ser efetuada pela competência natural (ver, por exemplo, Perry e Kuramitsu,

1981, *Infect. Immun*, 32: 1295-1297), pela transformação de protoplasto (ver, por exemplo, Catt e Jollick, 1991, *Microbios*. 68: 189-2070, pela eletroporação (ver, por exemplo, Buckley *et al.*, 1999, *Appl. Environ. Microbiol.* 65: 3800-3804) ou pela conjugação (ver, por exemplo, Clewell, 5 1981, *Microbial. Rev.* 45: 409-436), Entretanto, qualquer método conhecido na para a introdução de DNA em uma célula hospedeira pode ser usado.

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 10 fúngica. “Fungos” como aqui usado inclui os filos *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 15 171) e todos os 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 levedura ascosporogênea (*Endomycetales*), levedura basidiosporogênea e levedura pertencente ao *Fungi Imperfecti* (*Blastomycetes*). Visto que a classificação de 20 levedura pode mudar no futuro, para os propósitos desta invenção, levedura deve ser definida como descrito em *Biology and Activities of Yeast* (Skinner, EA., Passmore, S. M. e Davenport, R. R., eds, Soc. App. Bacterial. Symposium Series No. 9, 1980).

Em um aspecto ainda mais preferido, a célula hospedeira de 25 levedura é uma célula de *Candida*, *Hansenula*, *Kluyveromyces*, *Pichia*, *Saccharomyces*, *Schizosaccharomyces* ou *Yarrowia*,

Em um aspecto mais preferido, a célula hospedeira de levedura é uma célula de *Saccharomyces carlsbergensis*, *Saccharomyces cerevisiae*, *Saccharomyces diastaticus*, *Saccharomyces douglasii*, *Saccharomyces*

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

5 Em um outro aspecto mais preferido, a célula hospedeira fúngica é uma célula fúngica filamentosa. “Fungos Filamentosos” incluem 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 de micélios composta de quitina, 10 celulose, 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 de carbono pode ser fermentativo.

15 Em um aspecto ainda mais preferido, a célula hospedeira fúngica filamentosa é uma célula de *Acremonium*, *Aspergillus*, *Aureobasidium*, *Bjerkandera*, *Ceriporiopsis*, *Chrysosporium*, *Coprinus*, *Coriolus*, *Cryptococcus*, *Filibasidium*, *Fusarium*, *Humicola*, *Magnaporthe*, *Mucor*, *Myceliophthora*, *Neocallimastix*, *Neurospora*, *Paecilomyces*, 20 *Penicillium*, *Phanerochaete*, *Phlebia*, *Piromyces*, *Pieurotus*, *Schizophyllum*, *Talammyces*, *Thermoascus*, *Thielavia*, *Tolypocladium*, *Trametes* ou *Trichoderma*.

Em um aspecto mais preferido, a célula hospedeira fúngica filamentosa é uma célula de *Aspergillus awatroti*, *Aspergillus fumigatus*, 25 *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, *Fusarium oxisporum*, *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
5 hospedeira fúngica filamentosa é uma célula de *Bjerkandera adusta*,
Ceriporiopsis aneirina, *Ceriporiopsis aneirina*, *Ceriporiopsis caregiea*,
Ceriporiopsis gilvescens, *Ceriporiopsis pannocinta*, *Ceriporiopsis rivulosa*,
Ceriporiopsis subrufa, *Ceriporiopsis subvermispora*, *Chrysosporium*
keratinophilum, *Chrysosporium lucknowense*, *Chrysosporium tropicum*,
10 *Chrysosporium merdarium*, *Chrysosporium inops*, *Chrysosporium pannicola*,
Chrysosporium queensiandicum, *Chrysosporium zonatum*, *Coprinus cinereus*,
Coriolus hirsutus, *Humicola insolens*, *Humicola lanuginosa*, *Mucor miehei*,
Mycelophthora thermophila, *Neurospora Grasse*, *Penicillium brasilianum*,
Penicillium purpurogenum, *Phanerochaete chrysosporium*, *Phlebia radiata*,
15 *Pleurotus eryngii*, *Thielavia terrestris*, *Trametes villosa*, *Trametes versicolor*,
Trichoderma harzianum, *Trichoderma koningii*, *Trichoderma*
longibrachiatum, *Trichoderma taeseli* ou *Trichoderma viride*.

As células fúngicas podem ser transformadas por um processo
 envolvendo a formação de protoplasto, transformação dos protoplastos e
 20 regeneração da parede celular em uma maneira conhecida por si. Os
 procedimentos adequados para a transformação de células hospedeiras de
Aspergillus e *Trichoderma* são descritas 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 descritas por
 25 Malardier *et al.*, 1989, Gene 78: 147-156 e WO 96/00787. A levedura pode
 ser transformada 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

Métodos de Produção

5 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 produz 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
10 um aspecto mais preferido a célula é de *Thielavia terrestris*. Em um aspecto mais preferido a célula é de *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
15 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 seqüência de
20 nucleotídeo mutante tendo pelo menos uma mutação na seqüência codificadora do polipeptídeo maduro da SEQ ID NO: 1 ou SEQ ID NO: 3, em que a seqüência de nucleotídeo mutante codifica um polipeptídeo que compreende ou consiste do polipeptídeo maduro da SEQ ID NO: 2 ou SEQ ID NO: 4 e (b) recuperar o polipeptídeo.

25 Em um aspecto preferido, o polipeptídeo maduro tem os aminoácidos de 19 a 396 da SEQ ID NO: 2. Em um outro aspecto preferido, o polipeptídeo maduro tem os aminoácidos de 20 a 400 da SEQ ID NO: 4. Em um outro aspecto preferido, a seqüência que codifica o polipeptídeo maduro tem os nucleotídeos de 110 a 1557 da SEQ ID NO: 1. Em um outro aspecto

preferido, a sequência que codifica o polipeptídeo maduro tem os nucleotídeos de 58 a 1200 da SEQ ID NO: 3.

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 pequena escala ou grande escala (incluindo as fermentações contínuas, batelada, lote alimentado ou estado sólido) em fermentadores de laboratório ou industriais realizadas em um meio adequado e sob condições que possibilitem 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 de célula.

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

O polipeptídeo resultante pode ser recuperado usando métodos conhecidos 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 de íon, afinidade, hidrofóbica, cromatofocalização e exclusão de tamanho), procedimentos eletroforéticos (por exemplo, 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 se obter polipeptídeos substancialmente puros.

Plantas

A presente invenção também diz respeito a plantas, por exemplo, uma planta transgênica, parte de planta ou célula de planta, compreendendo um polinucleotídeo isolado que codifica um polipeptídeo tendo atividade de endoglucanase da presente invenção de modo a expressar e produzir o polipeptídeo em quantidades recuperáveis. O polipeptídeo pode ser recuperado da planta ou parte da planta. Alternativamente, a planta ou parte da 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*), grama forrageira tal como *Festuca*, *Lolium*, grama temperada, tal como *Agrostis* e cereais, por exemplo, trigo, aveia, centeio, cevada, arroz, sorgo e milho (milho).

Os exemplos de plantas dicot são tabaco, legumes, tais como tremoço, batata, beterraba açucareira, ervilha, feijão e soja e plantas crucíferas (família *Brassicaceae*), tais como couve-flor, colza e o organismo modelo intimamente relacionado *Arabidopsis thaliana*.

Os exemplos de partes de planta 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. Compartimentos de célula de planta específicos, tais como cloroplastos, apoplastos, mitocôndria, vacúolos, peroxissomas e citoplasma também são considerados ser uma parte da planta. Além disso, qualquer célula de planta, qualquer que seja a origem do tecido, é considerada ser uma parte da planta. Do mesmo modo, partes da planta tais como tecidos e células específicos isolados para facilitar a utilização da invenção também são consideradas partes da planta, por exemplo, embriões, endospermas, aleurona e revestimentos de sementes.

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

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

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

A escolha de seqüências reguladoras, tais como seqüências promotoras e terminadoras e opcionalmente 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 indutível ou pode ser específica de desenvolvimento, estágio ou tecido e o produto de gene pode ser alvejado a um tecido ou parte da planta específicos tais como sementes ou folhas. As seqüências regulatórias são, por exemplo, descritas por Tague *et al.*, 1988, Plant Physiology 86: 506.

Para a expressão constitutiva, os promotores 35S-CaMV, da ubiquitina 1 do milho e da actina 1 do arroz podem ser usados (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 dos tecidos de depósito de armazenagem tais como sementes, tubérculos da batata e frutas (Edwards & Coruzzi, 1990, Ann. Rev. Genet. 24: 275-303) ou de tecidos de depósito metabólico tais como meristemas (Ito *et al.*, 1994, Plant Mol. Biol. 24: 863-878), um promotor específico de semente tal como o promotor 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 B4 e o gene da proteína de semente desconhecido da *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 de *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 seqüência de nucleotídeo que codifica um polipeptídeo da presente invenção. Por exemplo, Xu *et al.*, 1993, supra, divulgam o uso do primeiro intron do gene da actina 1 do arroz para realçar a expressão.

O gene marcador selecionável e quaisquer 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 no ramo, 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 pela *Agrobacterium tumefaciens* é o método de escolha para gerar dicots transgênicas (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 frequentemente usados para estas plantas. Presentemente, o método de escolha para gerar monocots transgênicas 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. Frequentemente 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 a 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 compreendem (a) cultivar um planta transgênica ou uma célula de planta que compreendem um polinucleotídeo que codifica um polipeptídeo tendo atividade de endoglucanase da presente invenção sob condições condutivas para a produção do polipeptídeo; e (b) recuperar o polipeptídeo.

Remoção ou Redução da Atividade de Endoglucanase

A presente invenção também diz respeito a métodos para produzir um mutante de uma célula precursora, que compreende romper ou deletar uma sequência de polinucleotídeo ou uma porção desta, que codifiquem um polipeptídeo da presente invenção, que resulte na célula

mutante produzindo menos do polipeptídeo do que a célula precursora quando cultivada sob as mesmas condições.

A célula mutante pode ser construída pela redução ou eliminação da expressão de uma sequência 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 sequência de nucleotídeo é inativada. A sequência de nucleotídeo a ser modificada ou inativada pode ser, por exemplo, a região codificadora ou uma parte desta essencial para a atividade ou um elemento regulador requerido para a expressão da região codificadora. Um exemplo de uma tal sequência reguladora ou de controle pode ser uma sequência promotora ou uma parte funcional desta, isto é, uma parte que é suficiente para afetar a expressão da sequência de nucleotídeo. Outras sequências de controle para a modificação possível incluem, mas não são limitadas a, uma líder, sequência de poliadenilação, sequência de propeptídeo, sequência de peptídeo de sinal, terminador de transcrição e ativador transcricional.

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

Os exemplos de um agente mutagenizante físico ou químico adequado para o presente propósito incluem irradiação ultravioleta (UV), hidroxilamina, N-metil-N'-nitro-N-nitrosoguanidina (MNNG), O-metil hidroxilamina, ácido nitroso, metano sulfonato de etila (EMS), bissulfito de

sódio, ácido fórmico e análogos de nucleotídeo.

Quando tais agentes são usados, a mutagênese é tipicamente realizada pela incubação da célula precursora a ser mutagenizada na presença do agente mutagenizante de escolha sob condições e triagem adequadas e/ou selecionando quanto a células mutantes 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 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 parada ou uma mudança na matriz de leitura aberta. Tal modificação ou inativação pode ser realizada pela mutagênese direcionada ao sítio ou mutagênese gerada pela PCR de acordo com métodos conhecidos na técnica. Embora, em princípio, a modificação possa ser realizada *in vivo*, i.e., diretamente na célula para expressar 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 reduzir a expressão de uma seqüência de nucleotídeo por uma célula está fundamentado em técnicas de substituição de gene, deleção de gene ou rompimento de gene. Por exemplo, no método de rompimento de gene, uma seqüência de ácido nucleico que corresponde 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. Por recombinação homóloga, a seqüência de ácido nucleico defeituosa substitui a seqüência de nucleotídeo endógena. Pode ser desejável que a seqüência de nucleotídeo defeituosa também codifique um marcador que pode ser usado para a seleção de transformantes em que a

seqüência de nucleotídeo foi modificada ou destruída. Em um aspecto particularmente preferido, a seqüência de nucleotídeo é rompida com um marcador selecionável tal como aqueles aqui descritos.

Alternativamente, a modificação ou inativação da seqüência de nucleotídeo pode ser realizada pelas técnicas de anti-sentido ou RNAi estabelecidas usando uma seqüência complementar à seqüência de nucleotídeo. Mais especificamente, a expressão da seqüência de nucleotídeo por uma célula pode ser reduzida ou eliminada pela introdução de uma seqüência complementar à seqüência de nucleotídeo do gene que pode ser transcrita na célula e é capaz de hibridizar ao mRNA produzido na célula. Sob condições que permitam que a seqüência de nucleotídeo de anti-sentido complementar hibridize com o mRNA, a quantidade de proteína traduzida é assim reduzida ou eliminada.

A presente invenção diz respeito ainda a uma célula mutante de uma célula precursora que compreende um rompimento ou deleção de uma seqüência de nucleotídeo que codifica o polipeptídeo ou uma seqüência de controle desta, que resulta na célula mutante produzindo menos do polipeptídeo ou nenhum polipeptídeo comparada com a célula precursora.

As células mutantes deficientes em polipeptídeo assim criadas 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 modificações foram feitas para alterar a seqüência nativa ou uma proteína nativa cuja expressão é quantitativamente alterada como um resultado de uma manipulação da célula hospedeira pelas técnicas de DNA recombinantes.

Em um outro aspecto, a presente invenção diz respeito a um método para produzir um produto de proteína essencialmente livre de atividade de endoglucanase 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 endoglucanase para o caldo de fermentação antes, durante ou depois que a fermentação tem sido completada, recuperar o produto de interesse do caldo de fermentação e opcionalmente submeter o produto recuperado a 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 endoglucanase cultivando-se a célula sob condições que permitam 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 endoglucanase 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 ser opcionalmente usado em combinação com um tratamento com um inibidor de endoglucanase.

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 endoglucanase. A remoção completa da atividade de endoglucanase pode ser obtida pelo uso deste método.

O tratamento de pH e temperatura combinados é preferivelmente realizado em 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 conhecidos na técnica.

Os métodos da presente invenção para produzir um produto essencialmente isento de endoglucanase 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 de célula de planta. Os exemplos de tais enzimas incluem uma aminopeptidase, amilase, amiloglucosidase, carboidrase, carboxipeptidase, catalase, celulase, quitinase, cutinase, ciclodextrina glicosiltransferase, desoxirribonuclease, esterase, galactosidase, beta-galactosidase, glicoamilase, glicose oxidase, glucosidase, haloperoxidase, hemicelulase, invertase, isomerase, lactase, ligase, lipase, liase, manosidase, oxidase, enzima pectinolítica, peroxidase, fitase, fenoloxidase, polifenoloxidase, enzima proteolítica, ribonuclease, transferase, transglutaminase ou xilanase. As células deficientes em endoglucanase também podem ser usadas para expressar proteínas heterólogas de interesse farmacêutico tais como hormônios, fatores de crescimento, receptores e outros,

Deve ser entendido que o termo “polipeptídeos eucarióticos” inclui não apenas polipeptídeos nativos, mas também aqueles polipeptídeos, por exemplo, enzimas, que foram modificados pelas substituições, deleções ou adições de aminoácido ou outra de tais modificações para realçar a atividade, termoestabilidade, tolerância ao pH e outras.

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

Métodos de Inibir a Expressão de um Polipeptídeo

A presente invenção também diz respeito a métodos de inibir a expressão de um polipeptídeo em uma célula, que compreendem administrar à célula ou expressar na célula uma molécula de RNA de filamento duplo (dsRNA), em que o dsRNA compreende uma subsequência ou porção de um polinucleotídeo da presente invenção. Em um aspecto preferido, o dsRNA tem cerca de 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 ou mais nucleotídeos duplex no comprimento. Em um outro aspecto preferido, o polipeptídeo tem atividade de endoglucanase.

O dsRNA é preferivelmente um RNA interferente pequeno (siRNA) ou um micro RNA (miRNA). Em um aspecto preferido, o dsRNA é RNA interferente pequeno (siRNAs) para inibir a transcrição. Em um outro aspecto preferido, o dsRNA é micro RNA (miRNAs) para inibir a tradução.

A presente invenção também diz respeito a tais moléculas de RNA de filamento duplo (dsRNA) para inibir a expressão de um polipeptídeo em uma célula, em que o dsRNA compreende uma subsequência ou porção de um polinucleotídeo que codifica o polipeptídeo maduro da SEQ ID NO: 2 ou SEQ ID NO: 4. Embora a presente invenção não seja limitada por qualquer mecanismo particular de ação, o dsRNA pode entrar em uma célula e causar a degradação de um RNA de filamento simples (ssRNA) de seqüências similares ou idênticas, incluindo mRNAs endógenos. Quando uma célula é exposta ao dsRNA, o mRNA do gene homólogo é seletivamente degradado por um processo chamado de interferência de RNA (RNAi).

Os dsRNAs da presente invenção podem ser usados em produtos terapêuticos de silenciamento de gene. Em um aspecto, a invenção fornece métodos para degradar seletivamente RNA usando os dsRNAs da presente invenção. O processo pode ser praticado *in vitro*, *ex vivo* ou *in vivo*. Em um aspecto, as moléculas de dsRNA podem ser usadas para gerar uma mutação de perda de função em uma célula, um órgão ou um animal. Os métodos para fabricar e usar moléculas de dsRNA para degradar

seletivamente RNA são bem conhecidos na técnica, ver, por exemplo, a Patente U.S. Nº 6.506.559, Patente U.S. Nº 6.511.824; Patente U.S. Nº 6.515.109; e Patente U.S. Nº 6.489.127.

Composições

5 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 “enriquecido” indica que a atividade de endoglucanase da composição foi aumentada, por exemplo, com um fator de enriquecimento de pelo menos 1,1.

10 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 aminopeptidase, amilase, carboidrase, carboxipeptidase, catalase, celulase, quitinase, cutinase, ciclodextrina glicosiltransferase, desoxirribonuclease, esterase, alfa-galactosidase, beta-galactosidase, glicoamilase, alfa-glucosidase, beta-glucosidase, haloperoxidase, invertase, lacase, lipase, manosidase, oxidase, enzima pectinolítica, peptidoglutaminase, peroxidase, fitase, polifenoloxidase, enzima proteolítica, ribonuclease, transglutaminase
 15 ou xilanase. A(s) enzima(s) adicional(is) 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*
 20 *bactridioides*, *Fusarium cerealis*, *Fusarium crookwellense*, *Fusarium culmorum*, *Fusarium graminearum*, *Fusarium graminum*, *Fusarium heterosporum*, *Fusarium negundi*, *Fusarium oxisporum*, *Fusarium reticulatum*, *Fusarium toseum*, *Fusarium sambucinum*, *Fusarium sarcochroum*, *Fusarium sulphureum*. *Fusarium toruloseum*, *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*.

5 As composições de polipeptídeo podem ser preparadas de acordo com métodos conhecidos na técnica e podem estar na forma de uma composição líquida ou uma 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
10 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 pode ser determinada com base em métodos conhecidos na técnica.

15 **Usos**

A presente invenção também diz respeito a métodos para degradar ou converter um material celulósico, que compreendem tratar o material celulósico com uma composição que compreenda uma quantidade eficaz de um polipeptídeo tendo atividade de endoglucanase da presente
20 invenção. Em um aspecto preferido, o método compreende ainda recuperar o material celulósico degradado ou convertido.

Os polipeptídeos e células hospedeiras da presente invenção podem ser usados na produção de monossacarídeos, dissacarídeos e polissacarídeos como estoques de alimentação química ou fermentação de
25 biomassa celulósica para a produção de etanol, plásticos, outros produtos ou intermediários. A composição que compreende o polipeptídeo tendo atividade de endoglucanase pode 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, a composição pode compreender

uma célula hospedeira da presente invenção como uma fonte do polipeptídeo tendo atividade de endoglucanase em um processo de fermentação com a biomassa. A célula hospedeira também pode conter genes nativo ou heterólogos que codificam outras proteínas e enzimas, mencionadas acima, úteis no processamento de biomassa. Em particular, os polipeptídeos e células hospedeiras da presente invenção podem ser usadas para aumentar o valor de resíduos de processamento (grãos de destiladores secos, grãos gastos da fabricação de cerveja, bagaço de cana de açúcar, etc.) pela degradação parcial ou completa de celulose ou hemicelulose. As composições também podem compreender outras proteínas e enzimas úteis no processamento de biomassa, por exemplo, celobioidrolase, beta-glucosidase, enzimas hemicelulolítica, realçadores (WO 2005/074647, WO 2005/074656), etc.

Nos métodos da presente invenção, qualquer material celulósico, tal como biomassa, pode ser usado. É aqui entendido que o termo “material celulósico” abrange lignocelulose. A biomassa pode incluir, mas não é limitado a, recursos lenhoso, resíduos sólidos municipais, papel usado, safras 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 de célula primária de biomassa é a celulose, o segundo mais abundante é a hemi-celulose e o terceiro é a pectina. A parede de célula secundária, produzida depois que a célula parou de crescer, também contém polissacarídeos e é reforçada pela lignina polimérica covalentemente reticulada a hemicelulose. A celulose é um homopolímero de anidrocetobiose e assim um beta-(1-4)-D-glicano linear,

enquanto que as hemiceluloses 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. As hemiceluloses usualmente ligam-se por hidrogênio à celulose, assim como a outras hemiceluloses, o que ajuda a estabilizar a matriz de parede celular.

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

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

(2) As “exo-1,4-beta-D-glucanases” 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 de 1,4-beta-glicanos.

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

Os polipeptídeos tendo atividade de endoglucanase da presente invenção são preferivelmente usados em conjunção com outras proteínas celulolíticas, por exemplo, exo-1,4-beta-D-glucanases e beta-D-glucosidases, para degradar o componente de celulose do substrato de biomassa, (ver, por exemplo. Brigham *et al.*, 1995, em Handbook on Bioethanol (Charles E. Wyman, editor), pp, 119-141 Taylor & Francis, Washington DC.: Lee, 1997, Journal of Biotechnology 56: 1-24). O termo “proteínas celulolíticas” é aqui definido como aquelas proteínas ou misturas de proteínas mostradas como sendo capazes de hidrolisar ou converter ou degradar a celulose sob as

condições testadas.

As *exo-1,4-beta-D-glucanases* e *beta-D-glucosidases* podem ser produzidas por qualquer método conhecidos na técnica (ver, por exemplo, Bennett, J. W. e LaSure, L. (eds.), *More Gene Manipulations in Fungi*, Academic Press, CA, 1991).

As quantidades ótimas de um polipeptídeo tendo atividade de endoglucanase e outras proteínas celulolíticas depende de diversos fatores incluindo, mas não são limitados à mistura de proteínas celulolíticas componentes, o substrato celulósico, a concentração de substrato celulósico, o(s) pré tratamento(s) do substrato celulósico, temperatura, tempo, pH e inclusão de organismo de fermentação (por exemplo, levedura para a Sacarificação e Fermentação Simultâneas).

Em um aspecto preferido, a quantidade de polipeptídeo tendo atividade de endoglucanase por g de material celulósico é de cerca de 0,5 a cerca de 50 mg, preferivelmente de cerca de 0,5 a cerca de 40 mg, mais preferivelmente de cerca de 0,5 a cerca de 25 mg, mais preferivelmente de cerca de 0,75 a cerca de 20 mg, mais preferivelmente de cerca de 0,75 a cerca de 15 mg, ainda mais preferivelmente de cerca de 0,5 a cerca de 10 mg e o mais preferivelmente de cerca de 2,5 a cerca de 10 mg por g de material celulósico.

Em um outro aspecto preferido, a quantidade de proteínas celulolíticas por g de material celulósico é de cerca de 0,5 a cerca de 50 mg, preferivelmente de cerca de 0,5 a cerca de 40 mg, mais preferivelmente de cerca de 0,5 a cerca de 25 mg, mais preferivelmente de cerca de 0,75 a cerca de 20 mg, mais preferivelmente de cerca de 0,75 a cerca de 15 mg, ainda mais preferivelmente de cerca de 0,5 a cerca de 10 mg e o mais preferivelmente de cerca de 2,5 a cerca de 10 mg por g de material celulósico.

Nos métodos da presente invenção, a composição pode ser suplementada por uma ou mais atividades de enzima adicionais para melhorar

a degradação do material celulósico. As enzimas adicionais preferidas são hemicelulases, esterases (por exemplo, lipases, fosfolipases e/ou cutinases), proteases, lacases, peroxidases ou misturas destes.

Nos métodos da presente invenção, a(s) enzima(s) adicional(is) pode(m) ser adicionadas antes ou durante a fermentação, incluindo durante ou depois da propagação do(s) microorganismo(s) de fermentação.

As enzimas podem ser derivadas ou obtidas a partir de qualquer origem adequada, incluindo, origem bacteriana, fúngica, de levedura ou mamífera. O termo “obtida” significa aqui que a enzima pode ter sido isolada de um organismo que naturalmente produz a enzima como uma enzima nativa. O termo “obtida” também significa aqui que a enzima pode ter sido produzida recombinantemente em um organismo hospedeiro, em que a enzima recombinantemente produzida é nativa ou estranha para o organismo hospedeiro ou tem uma sequência de aminoácido modificada, por exemplo, tendo um ou mais aminoácidos que são deletados, inseridos e/ou substituídos, isto é, uma enzima recombinantemente produzida que é um mutante e/ou um fragmento de uma sequência de aminoácido nativa ou uma enzima produzida pelos processos de embaralhamento de ácido nucleico conhecidos na técnica. Abrangidos dentro do significado de uma enzima nativa estão as variantes naturais e dentro do significado de uma enzima estranha estão variantes obtidas recombinantemente, tais como pela mutagênese direcionada ao sítio ou embaralhamento.

As enzimas também podem ser purificadas. O termo “purificada” como aqui usado abrange enzimas livres de outros componentes do organismo a partir do qual a mesma é derivada. O termo “purificada” também abrange enzimas livres de componentes do organismo nativo a partir do qual a mesma é obtida. As enzimas podem ser purificadas, apenas com quantidades menores de outras proteínas estando presentes. A expressão

“outras proteínas” diz respeito em particular a outras enzimas. O termo “purificada” como aqui usado também refere-se à remoção de outros componentes, particularmente outras proteínas e mais particularmente outras enzimas presentes na célula de origem da enzima da invenção. A enzima pode ser “substancialmente pura,” isto é, livre de outros componentes do organismo no qual a mesma é produzida, isto é, por exemplo, um organismo hospedeiro para as enzimas recombinantemente produzidas. Em um aspecto preferido, as enzimas são pelo menos 75% (p/p), preferivelmente pelo menos 80%, mais preferivelmente pelo menos 85%, mais preferivelmente pelo menos 90%, mais preferivelmente pelo menos 95%, mais preferivelmente pelo menos 96%, mais preferivelmente pelo menos 97%, ainda mais preferivelmente pelo menos 98% ou o mais preferivelmente pelo menos 99% puras. Em um outro aspecto preferido, a enzima é 100% pura.

As enzimas usadas na presente invenção pode estar em qualquer forma adequada para o uso nos processos aqui descritos, tais como, por exemplo, um caldo de fermentação bruto com ou sem células, um pó ou granulado secos, um granulado não pulverulento, um líquido, um líquido estabilizado ou uma enzima protegida. Os granulados podem ser produzidos, por exemplo, como divulgado nas Patentes U.S. Nºs 4.106.991 e 4.661.452 e podem ser opcionalmente revestidos por processos conhecidos na técnica. As preparações de enzima líquidas, por exemplo, podem ser estabilizadas pela adição de estabilizadores tais como um açúcar, um álcool de açúcar ou um outro poliol e/ou ácido láctico ou um outro ácido orgânico de acordo com processo estabelecido. As enzimas protegidas podem ser preparadas de acordo com o processo divulgado na EP 238.216.

Os métodos da presente invenção podem ser usados para processar um material celulósico em muitos produtos orgânicos, produtos químicos e combustíveis úteis, Além de etanol, algumas mercadorias e produtos químicos especiais que podem ser produzidos a partir de celulose

incluem xilose, acetona, acetato, glicina, lisina, ácidos orgânicos (por exemplo, ácido láctico), 1,3-propanodiol, butanodiol, glicerol, etileno glicol, furfural, poliidroxiclcanoatos, cis, ácido cis-mucônico e ração de animal (Lynd, L. R., Wyman, C. E e Gerngross, T. U., 1999. Biocommodity Engineering, *Biotechnol. Prog.*, 15: 777-793; Philippidis, G. P.. 1996, Cellulose bioconversion technology, no *Handbook on Bioethanol Production and Utilization*, Wyman, C. E., ed., Taylor & Francis, Washington, DC, 179-212; e Ryu. D. D. Y. e Mandeis, M., 1980, Cellulases: biosynthesis and applications, *Enz. Microb. Technol.*, 2: 91-102), os benefícios de co-produção potenciais estendem-se além da síntese de produtos orgânicos múltiplos a partir de carboidrato fermentáveis. Os resíduos ricos em lignina que permanecem depois do processamento biológico podem ser convertidos para produtos químicos derivados de lignina ou usados para energizar a produção.

Os métodos convencionais usados para processar o material celulósico de acordo com os métodos da presente invenção são bem entendidos por aqueles habilitados na técnica. Os métodos da presente invenção podem ser implementados usando qualquer aparelho que processe biomassa convencional configurado para operar de acordo com a invenção.

Um tal aparelho pode incluir um reator de lote agitado, um reator agitado de fluxo contínuo com ultrafiltração, um reator de coluna de tampão de fluxo contínuo (Gusakov, A. V. e Sinitsyn, A. P., 1985, Kinetics of the enzymatic hydrolysis of cellulose: 1. A mathematical model for a batch reactor process, *Enz. Microb. Technol*, 7: 346-352), um reator de atrito (Ryu, S. K. e Lee, J. M., 1983, Bioconversion of waste cellulose by using a attrition bioreactor, *Biotechnol. Bioeng*, 25: 53-65) ou um reator com agitação intensiva induzida por um campo eletromagnético (Gusakov, A. V., Sinitsyn, A. P., Davydkin, I. Y., Davydkin, V. Y., Protas, O. V., 1996, Enhancement of enzymatic cellulose hydrolysis using a novel type of bioreactor with intensive stirring induced by electromagnetic field, *Appl. Biochem. Biotechnol.* 56:

141-153).

Os métodos convencionais incluem, mas não são limitados a, sacarificação, fermentação, hidrólise e fermentação separadas (SHF), sacarificação e fermentação simultâneas (SSF), sacarificação e cofermentação simultâneas (SSCF), hidrólise e fermentação híbridas (HHF) e conversão microbiana direta (DMC).

A SHF usa etapas de processo separadas para primeiro hidrolisar enzimaticamente a celulose para glicose e depois fermentar a glicose para etanol. Na SSF, a hidrólise enzimática de celulose e a fermentação de glicose para etanol são combinadas em uma etapa (Philippidis, G. P, 1996, Cellulose bioconversion technology, no Handbook on Bioethanol: Production and Utilization, Wyman, C. E., ed., Taylor & Francis, Washington, DC, 179-212). A SSCF inclui a cofermentação de açúcares múltiplos (Sheehan, J. e Himmel, M., 1999, Enzymes, energy and the environment: A strategic perspective on the U.S. Department of Energy's research and development activities for bioethanol, Biotechnol. Prog. 15: 817-827). A HHF inclui duas etapas separadas realizadas no mesmo reator mas em temperaturas diferentes, i.e., sacarificação enzimática de alta temperatura seguida por SSF em uma temperatura mais baixa que a cepa de fermentação possa tolerar. DMC combina todos os três processos (produção de celulose, hidrólise da celulose e fermentação) em uma etapa (Lynd, L R., Weimer. P. J., van Zil, W. H. e Pretorius, S., 2002, Microbial cellulose utilization: Fundamentals and biotechnology, Microbiol. Mol. Biol. Reviews 66: 506-577).

“Fermentação” ou “processo de fermentação” referem-se a qualquer processo de fermentação ou qualquer processo que compreenda uma etapa de fermentação. Um processo de fermentação inclui, sem limitação, processos de fermentação usados para produzir produtos de fermentação incluindo álcoois (por exemplo, arabinitol, butanol, etanol, glicerol, metanol,

1,3-propanodiol, sorbitol e xilitol); ácidos orgânicos (por exemplo, ácido acético, ácido acetônico, ácido adípico, ácido ascórbico, ácido cítrico, ácido 2,5-diceto-D-glicônico, ácido fórmico, ácido fumárico, ácido glucárico, ácido glicônico, ácido glicurônico, ácido glutárico, ácido 3-hidroxipropiônico, ácido itacônico, ácido láctico, ácido málico, ácido malônico, ácido oxálico, ácido propiônico, ácido succínico e ácido xilônico); cetonas (por exemplo, acetona); aminoácidos (por exemplo, ácido aspártico, ácido glutâmico, glicina, lisina, serina e treonina); gases (por exemplo, metano, hidrogênio (H₂), dióxido de carbono (CO₂) e monóxido de carbono (CO)). Os processos de fermentação também incluem processos de fermentação usados na indústria de álcool consumível (por exemplo, cerveja e vinho), indústria de laticínio (por exemplo, produtos de laticínio fermentado), indústria do couro e indústria do tabaco.

A presente invenção diz respeito ainda a métodos de produzir uma substância, que compreendem: (a) sacarificar um material celulósico com uma composição que compreenda uma quantidade eficaz de um polipeptídeo tendo atividade de endoglucanase; (b) fermentar o material celulósico sacarificado da etapa (a) com um ou mais microorganismos de fermentação; e (c) recuperar a substância da fermentação. A composição que compreende o polipeptídeo tendo atividade de endoglucanase pode 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 ou a composição pode compreender uma célula hospedeira da presente invenção como uma fonte do polipeptídeo tendo atividade de endoglucanase em um processo de fermentação com a biomassa.

A substância pode ser qualquer substância derivada da fermentação. Em uma forma de realização preferida, a substância é um álcool. Deve ser entendido que o termo “álcool” abrange uma substância que contém uma ou mais porções de hidroxila. Em uma forma de realização mais

preferida, o álcool é arabinitol. Em uma outra forma de realização mais preferida, o álcool é butanol. Em uma outra forma de realização mais preferida, o álcool é etanol. Em uma outra forma de realização mais preferida, o álcool é glicerol. Em uma outra forma de realização mais preferida, o álcool é metanol. Em uma outra forma de realização mais preferida, o álcool é 1,3-propanodiol. Em uma outra forma de realização mais preferida, o álcool é sorbitol. Em uma outra forma de realização mais preferida, o álcool é xilitol. Ver, por exemplo, Gong, C. S., Cao, N. J., Du, J. e Tsao, G. T., 1999, Ethanol Production from renewable resources. em *Advances in Biochemical Engineering/ Biotechnology*, Scheper, T., ed., Springer-Verlag Berlin Heidelberg, Alemanha, 65: 207-241; Silveira, M. M. e Jonas, R., 2002, The biotechnological production of sorbitol, *App. Microbiol, Biotechnol.* 59: 400-408; Nigam, P. e Singh, D., 1995, Processes for fermentative production of xilitol – a sugar substitute, *Process Biochemistry* 30 (2): 117-124; Ezeji, T. C., Qureshi, N. e Blaschek. H. P., 2003, Production of acetone, butanal and ethanol by *Clostridium beijerinckii* BA101 and in situ recovery by gas stripping, *World Journal of Microbiology and Biotechnology* 19 (6): 595-603.

Em uma outra forma de realização preferida, a substância é um ácido orgânico. Em uma outra forma de realização mais preferida, o ácido orgânico é ácido acético. Em uma outra forma de realização mais preferida, o ácido orgânico é o ácido acetônico. Em uma outra forma de realização mais preferida, o ácido orgânico é o ácido adípico. Em uma outra forma de realização mais preferida, o ácido orgânico é o ácido ascórbico. Em uma outra forma de realização mais preferida, o ácido orgânico é o ácido cítrico. Em uma outra forma de realização mais preferida, o ácido orgânico é o ácido 2,5-diceto-D-glicônico. Em uma outra forma de realização mais preferida, o ácido orgânico é o ácido fórmico, Em uma outra forma de realização mais preferida, o ácido orgânico é o ácido fumárico. Em uma outra forma de realização mais preferida, o ácido orgânico é o ácido glucárico. Em uma outra forma de

realização mais preferida, o ácido orgânico é o ácido glicônico. Em uma outra forma de realização mais preferida, o ácido orgânico é o ácido glicurônico. Em uma outra forma de realização mais preferida, o ácido orgânico é o ácido glutárico. Em uma outra forma de realização preferida, o ácido orgânico é o ácido 3-hidroxipropiônico. Em uma outra forma de realização mais preferida, o ácido orgânico é o ácido itacônico. Em uma outra forma de realização mais preferida, o ácido orgânico é o ácido láctico. Em uma outra forma de realização mais preferida, o ácido orgânico é o ácido málico, Em uma outra forma de realização mais preferida, o ácido orgânico é o ácido malônico. Em uma outra forma de realização mais preferida, o ácido orgânico é o ácido oxálico. Em uma outra forma de realização mais preferida, o ácido orgânico é o ácido propiônico. Em uma outra forma de realização mais preferida, o ácido orgânico é o ácido succínico. Em uma outra forma de realização mais preferida, o ácido orgânico é o ácido xilônico. Ver, por exemplo, Chen, R. e Lee, Y. Y., 1997, Membrane-mediated extractive fermentation for lactic acid production from celulosic biomass, Appl. Biochem. Biotechnol. 63-65: 435-448.

Em uma outra forma de realização preferida, a substância é uma cetona. Deve ser entendido que o termo “cetona” abrange uma substância que contém uma ou mais porções de cetona. Em uma outra forma de realização mais preferida, a cetona é acetona. Ver, por exemplo, Qureshi e Blaschek, 2003, supra.

Em uma outra forma de realização preferida, a substância é um aminoácido. Em uma outra forma de realização mais preferida, o ácido orgânico é o ácido aspártico. Em uma outra forma de realização mais preferida, o aminoácido é o ácido glutâmico. Em uma outra forma de realização mais preferida, o aminoácido é glicina. Em uma outra forma de realização mais preferida, o aminoácido é lisina. Em uma outra forma de realização mais preferida, o aminoácido é serina. Em uma outra forma de

realização mais preferida, o aminoácido é treonina. Ver, por exemplo, Richard, A. e Margaritis, A., 2004, Empirical modeling of batch fermentation kinetics for poli(glutamic acid) production and other microbial biopolymers, *Biotechnology and Bioengineering* 87 (4): 501-515.

5 Em uma outra forma de realização preferida, a substância é um gás. Em uma outra forma de realização mais preferida, o gás é metano. Em uma outra forma de realização mais preferida, o gás é H₂. Em uma outra forma de realização mais preferida, o gás é CO₂. Em uma outra forma de realização mais preferida, o gás é CO. Ver, por exemplo, Kataoka, N., A. Miya e K. Kiriya, 1997, Studies on hydrogen production by continuous culture system of hydrogen-producing anaerobic bacteria, *Water Science and Technology* 36 (6-7): 41-47; e Gunaseelan V. N. em *Biomass and Bioenergy*, Vol. 13 (1-2), pp. 83-114, 1997, Anaerobic digestion of biomass for methane production: A review.

15 A produção de uma substância a partir de material celulósico tipicamente requer quatro etapas principais. Estas quatro etapas são pré-tratamento, hidrólise enzimática, fermentação e recuperação. É exemplificado abaixo um processo para produzir etanol, mas deve ser entendido que processos similares podem ser usados para produzir outras substâncias, por exemplo, as substâncias descritas acima.

20 Pré tratamento. Na etapa de pré-tratamento ou pré-hidrólise, o material celulósico é aquecido para quebrar a lignina e a estrutura de carboidrato, solubilizar a maior parte da hemicelulose e tornar a fração de celulose acessível às enzimas celulolíticas. O aquecimento é realizado diretamente com vapor ou em pasta fluída onde um catalisador também pode ser adicionado ao material para acelerar as reações. Os catalisadores incluem ácidos fortes, tais como ácido sulfúrico e SO₂, ou álcali, tal como hidróxido de sódio. O propósito do estágio de pré-tratamento é facilitar a penetração das enzimas e microorganismos. A biomassa celulósica também pode ser

submetida a um pré tratamento de explosão de vapor hidrotérmica (Ver a Patente U.S. Pedido Nº 20020154730).

Sacarificação. Na etapa de hidrólise enzimática, também conhecida como sacarificação, as enzimas como aqui descritas são adicionadas ao material pré tratado para converter a fração de celulose para glicose e/ou outros açúcares. A sacarificação é no geral realizada em reatores ou fermentadores de tanque agitado sob condições de pH, temperatura e mistura controladas. Uma etapa de sacarificação pode durar até 200 horas, A sacarificação pode ser realizada em temperaturas de cerca de 30°C a cerca de 65°C, em particular em torno de 50°C e em um pH na faixa entre cerca de 4 e cerca de 5, especialmente em torno do pH 4,5. Para produzir glicose que pode ser metabolizada pela levedura, a hidrólise é tipicamente realizada na presença de uma beta-glucosidase,

Fermentação. Na etapa de fermentação, os açúcares, liberados do material celulósico como um resultado das etapas de pré tratamento e hidrólise enzimática, são fermentados até etanol por um organismo de fermentação, tal como levedura. A fermentação também pode ser realizada simultaneamente com a hidrólise enzimática no mesmo vaso, mais uma vez sob condições de pH, temperatura e mistura controladas. Quando a sacarificação e fermentação são realizadas simultaneamente no mesmo vaso, o processo é no geral chamado de sacarificação e fermentação simultâneas ou SSF.

Qualquer substrato celulósico ou matéria prima adequados pode ser usado em um processo de fermentação da presente invenção. O substrato é no geral selecionado com base no produto de fermentação desejado, isto é, a substância a ser obtida a partir da fermentação e no processo utilizado, como é bem conhecido na técnica. Os exemplos de substratos adequados para o uso nos métodos da presente invenção incluem materiais contendo celulose, tais como madeira ou resíduos vegetais ou

açúcares de peso molecular baixo DP1-3 obtido a partir de material celulósico processado que podem ser metabolizados pelo microorganismo de fermentação e que podem ser fornecidos pela adição direta ao meio de fermentação.

5 O termo “meio de fermentação” será entendido referir-se a um meio antes que o(s) microorganismo(s) de fermentação seja(sejam) adicionado(s), tal como, um meio que resulta de um processo de sacarificação, assim como um meio usado em um processo de sacarificação e fermentação simultâneas (SSF).

10 “Microorganismo de fermentação” refere-se a qualquer microorganismo adequado para o uso em um processo de fermentação desejado. Os microorganismos de fermentação adequados de acordo com a invenção são capazes de fermentar isto é, converter, açúcares, tais como glicose, xilose, arabinose, manose, galactose ou oligossacarídeos direta ou
15 indiretamente no produto de fermentação desejado. Os exemplos de microorganismos de fermentação incluem organismos fúngicos, tais como levedura. A levedura preferida inclui cepas da *Saccharomyces* spp. e em particular, *Saccharomyces cerevisiae*. A levedura comercialmente disponível inclui, por exemplo, Red Star[®]/Lesaffre Ethanol Red (disponível da Red
20 Star/Lesaffre, USA) FALI (disponível da Fleischmann's Yeast, uma divisão da Burns Philp Food Inc., USA), SUPERSTART (disponível da Alltech), GERT STRAND (disponível da Gert. Strand AB, Suécia) e FERMIOL (disponível da DSM Specialties).

25 Em uma forma de realização preferida, a levedura é uma *Saccharomyces* spp. Em uma forma de realização mais preferida, a levedura é *Saccharomyces cerevisiae*. Em uma outra forma de realização mais preferida, a levedura é *Saccharomyces distaticus*. Em uma outra forma de realização mais preferida, a levedura é *Saccharomyces uvarum*. Em uma outra forma de realização preferida, a levedura é uma *Kluyveromyces*. Em uma outra forma

de realização mais preferida, a levedura é *Kluyveromyces marxianus*. Em uma outra forma de realização mais preferida, a levedura é *Kluyveromyces fragilis*. Em uma outra forma de realização preferida, a levedura é uma *Candida*, Em uma outra forma de realização mais preferida, a levedura é *Candida pseudotropicalis*. Em uma outra forma de realização mais preferida, a levedura é *Candida brassicae*. Em uma outra forma de realização preferida, a levedura é uma *Clavispora*. Em uma outra forma de realização mais preferida, a levedura é *Clavispora lusitaniae*. Em uma outra forma de realização mais preferida, a levedura é *Clavispora opuntiae*. Em uma outra forma de realização preferida, a levedura é uma *Pachysolen*. Em uma outra forma de realização mais preferida, a levedura é *Pachysolen tannophilus*. Em uma outra forma de realização preferida, a levedura é uma *Bretannomyces*. Em uma outra forma de realização mais preferida, a levedura é *Bretannomyces clausenii* (Philippidis, G. P., 1996, Cellulose bioconversion technology, em Handbook on Bioethanol: Production and Utilization, Wyman, C. E., ed., Taylor & Francis, Washington, DC, 179-212).

As bactérias que podem fermentar eficientemente glicose para etanol incluem, por exemplo, *Zymomonas mobilis* e *Clostridium thermocellum* (Philippidis, 1996, supra).

É bem conhecido na técnica que os organismos descritos acima também podem ser usados para produzir outras substâncias, como aqui descrito.

A clonagem de genes heterólogos em *Saccharomyces cerevisiae* (Chen, Z., Ho, N. W. Y., 1993, Cloning and improving the expressão of *Pichia stipitis* xilose reductase gene in *Saccharomyces cerevisiae*, Appl. Biochem. Biotechnol. 39-40: 135-147; Ho, N. W. Y., Chen, Z. Brainard, A. P., 1998, Genetically engineered *Saccharomyces* yeast capable of effectively cofermenting glucose and xylose, Appl. Environ. Microbiol. 64: 1852-1859) ou em bactérias tais como *Escherichia coli* (Beall,

D. S., Ohta, K., Ingram, L. O., 1991, Parametric studies of ethanol production from xylose and other sugars by recombinant *Escherichia coli*, *Biotech. Bioeng.* 38: 296-303); *Klebsiella oxitoca* (Ingram, L. O., Comes, P. F., Lai, X., Moniruzzaman, M., Wood, B. E., Yomano, L. P., York, S. W., 1998, 5 Metabolic engineering of bacteria for ethanol production, *Biotechnol. Bioeng.* 58: 204-214) e *Zymomonas mobilis* (Zhang, M. Eddy, C., Deanda, K., Finkelstein, M. e Picataggio, S., 1995, Metabolic engineering of a pentose metabolism pathway in etanologenic *Zymomonas mobilis*, *Science* 267: 240-243; Deanda, K., Zhang, M., Eddy, C. e Picataggio, S., 1996, Development 10 of a arabinose-fermenting *Zymomonas mobilis* strain by metabolic pathway engineering, *Appl. Environ. Microbial.* 62: 4465-4470) tem levado à construção de organismos capazes de converter hexoses e pentoses para etanol (co-fermentação).

Levedura ou um outro microorganismo tipicamente é 15 adicionado à celulose degradada ou hidrolisado e a fermentação é continuada por cerca de 24 a cerca de 96 horas, tal como de cerca de 35 a cerca de 60 horas. A temperatura está tipicamente entre cerca de 26°C a cerca de 40°C, em particular em cerca de 32°C e de cerca do pH 3 a cerca do pH 6, em particular em torno do pH 4-5.

20 Em um forma de realização preferida, levedura ou um outro microorganismo é aplicado à celulose degradada ou hidrolisado e a fermentação é continuada por cerca de 24 a cerca de 96 horas, tal como tipicamente 35 a 60 horas, Em uma forma de realização preferida, a temperatura está no geral entre cerca de 26 a cerca de 40°C, em particular de 25 cerca de 32°C e o pH é no geral de cerca de pH 3 a cerca de pH 6, preferivelmente em torno do pH 4-5. Levedura ou um outro microorganismo são preferivelmente aplicados em quantidades de aproximadamente 10^5 a 10^2 , preferivelmente de aproximadamente 10^7 a 10^{10} , de modo especial de aproximadamente 5×10^7 contagens viáveis por ml de caldo de fermentação.

Durante uma fase de produção de etanol a contagem de célula de levedura deve estar preferivelmente na faixa de aproximadamente 10^7 a 10^{10} , especialmente em torno de aproximadamente 2×10^8 . Orientação adicional com respeito ao uso de levedura para fermentação pode ser encontrado, por exemplo, em “The Alcohol Textbook (Editores K. Jacques, T. P. Lyons e D. R. Kelsall, Nottingham University Press, Reino Unido 1999), que é por meio deste incorporada por referência.

O processo mais amplamente usado na técnica é o processo de sacarificação e fermentação simultâneas (SSF) onde não há estágio de contenção para a sacarificação, significado que a levedura e a enzima são adicionadas juntas.

Para a produção de etanol, a seguir da fermentação a massa é destilada para extrair o etanol. O etanol obtido de acordo com o processo da invenção pode ser usado como, por exemplo, etanol de combustível; etanol de beber, isto é, bebida alcoólica neutra potável ou etanol industrial.

Um estimulador de fermentação pode ser usado em combinação com qualquer um dos processos enzimáticos aqui descritos para melhorar ainda mais o processo de fermentação e em particular, o desempenho do microorganismo de fermentação, tal como, realce da taxa e rendimento de etanol. Um “estimulador de fermentação” refere-se aos estimuladores para o cultivo do microorganismos de fermentação, em particular, levedura. Os estimuladores de fermentação para cultivo incluem vitaminas e minerais. Os exemplos de vitaminas incluem multivitaminas, biotina, pantotenato, ácido nicotínico, meso-inositol, tiamina, piridoxina, ácido para-aminobenzóico, ácido fólico, riboflavina e Vitaminas A, B, C, D e E. Ver, por exemplo, Alfenore *et al.*, improving ethanol production and viability of *Saccharomyces cerevisiae* by a vitamin feeding strategy during fed-batch process, Springer-Verlag (2002), que é por meio deste incorporada por referência. Os exemplos de minerais incluem minerais e sais minerais que

possam fornecer nutrientes que compreendam P, K, Mg, S, Ca, Fe, Zn, Mn e Cu

Recuperação. O álcool é separado do material celulósico fermentado e purificado pelos métodos convencionais de destilação. Etanol com uma pureza de até cerca de 96% em vol. de etanol pode ser obtido, que pode ser usado como, por exemplo, etanol de combustível, etanol de beber, isto é, bebidas alcoólicas neutras potáveis ou etanol industrial.

Para outras substâncias, qualquer método conhecido na técnica pode ser usado incluindo, mas não são limitados a, cromatografia (por exemplo, de troca iônica, de afinidade, hidrofóbica, cromatofocalização e exclusão de tamanho), procedimentos eletroforéticos (por exemplo, focalização isoeletrica preparativa), solubilidade diferencial (por exemplo, precipitação com sulfato de amônio), SDS-PAGE, destilação ou extração.

Nos métodos da presente invenção, o polipeptídeo tendo atividade de endoglucanase e outra(s) proteína(s) celulolítica(s) pode(m) ser suplementada(s) por uma ou mais atividades de enzima adicionais para melhorar a degradação do material celulósico. As enzimas adicionais preferidas são hemicelulases, esterases (por exemplo, lipases, fosfolipases e/ou cutinases), proteases, lacases, peroxidases ou misturas destas.

Nos métodos da presente invenção, a(s) enzima(s) adicional(is) pode(m) ser adicionada(s) antes ou durante a fermentação, incluindo durante ou depois da propagação do(s) microorganismo(s) de fermentação.

Peptídeos de Sinal

A presente invenção também diz respeito 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 18 da SEQ ID NO: 2 ou aminoácidos de 1 a 19 da SEQ ID NO: 4, em que o

gene é estranho para a primeira e segunda seqüências de nucleotídeo.

Em um aspecto preferido, a seqüência de nucleotídeo compreende ou consiste dos nucleotídeos de 56 a 109 da SEQ ID NO: 1. Em um outro aspecto preferido, a seqüência de nucleotídeo compreende ou
5 consiste dos nucleotídeos de 1 a 57 da SEQ ID NO: 3.

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
10 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
15 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 a partir de pelo
20 menos duas proteínas diferentes em que uma ou mais podem ser heterólogas ou nativas para a célula hospedeira. As proteínas incluem ainda variações alélicas que ocorrem naturalmente e variações engendradas das proteínas mencionadas acima e proteínas híbridas.

Preferivelmente, a proteína é um hormônio ou variante deste,
25 enzima, receptor ou porção deste, anticorpo ou porção deste ou repórter. Em 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, glicoamilase, alfa-glucosidase, beta-glucosidase, invertase, lacase, uma outra lipase, manosidase, mutanase, oxidase, enzima pectinolítica, peroxidase, fitase, polifenoloxidase, enzima proteolítica, ribonuclease, transglutaminase ou xilanase.

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

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

10 **Exemplos**

Materiais

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

Cepas

15 *Thielavia terrestris* NRRL 8126 foi usada como a fonte dos polipeptídeos da Família 6 tendo atividade de endoglucanase. A *Saccharomyces cerevisiae* cepa W3124 (MATa: ura 3-52; leu 2-3, 112; his 3-D200: pep 4-1137: prc1::HIS3: prb1; LEU2; cir⁺) foi usada para a triagem de bibliotecas de expressão de *Thielavia terrestris* NRRL 8126 quanto a
20 atividade de endoglucanase. *Aspergillus oryzae* cepa HowB104 (negativa em alfa-amilase) foi usada para a expressão do polipeptídeo CEL6B de *Thielavia terrestris* NRRL 8126 tendo atividade de endoglucanase. *Aspergillus oryzae* cepa JaL250 (WO 99/61651) foi usada para a expressão do polipeptídeo CEL6C de *Thielavia terrestris* NRRL 8126 tendo atividade de endoglucanase.

25 **Meios e Soluções**

O meio LB foi composto por litro de 10 g de triptona, 5 g de extrato de levedura e 5 g de cloreto de sódio,

O meio LB ampicilina foi composto por litro de 10 g de triptona, 5 g de extrato de levedura, 5 g de cloreto de sódio e 50 µg de

ampicilina por ml (esterilizado por filtração, adicionado depois da autoclavagem).

As placas de LB ampicilina foram compostas por litro de meio LB ampicilina e 15 g de bacto ágar.

5 O meio YEG foi composto de 0,5% de extrato de levedura e 2% de glicose.

O meio YPD foi composto de 1% de extrato de levedura, 2% de peptona e esterilizado por filtração 2% de glicose adicionados depois da autoclavagem.

10 O meio YPM foi composto de 1% de extrato de levedura, 2% de peptona e esterilizado por filtração 2% de maltodextrina adicionados depois da autoclavagem.

O meio SC-URA com galactose foi composto por litro de 100 ml de 10X sais Basais, 28 ml de casaminoácidos a 20% sem vitaminas, 10 ml de triptofano a 1%, 3,6 ml de treonina a 5% (esterilizada por filtração, adicionada depois da autoclavagem) e 100 ml de galactose a 20% (esterilizada por filtração, adicionada depois da autoclavagem).

O meio SC-URA com glicose foi composto por litro de 100 ml de 10X solução de sais Basais, 28 ml de casaminoácidos a 20% sem
20 vitaminas, 10 ml de triptofano a 1%, 3,6 ml de treonina a 5% (esterilizada por filtração, adicionada depois da autoclavagem) e 100 ml de glicose a 20% (esterilizada por filtração, adicionada depois da autoclavagem).

A solução 10X sais Basais foi composta por litro de 75 g de base de nitrogênio de levedura, 113 g de ácido succínico e 68 g de NaOH.

25 SC-ágar foi composto por litro de meio SC-URA (com glicose ou galactose como indicado) e 20 g de ágar.

As placas com 0,1% de AZCL HE celulose SC ágar com galactose foram compostas por litro de meio SC-URA com galactose, 20 g de ágar e 0,1% de AZCL HE celulose (Megazyme, Wcklow, Irlanda).

O meio PD com celulose foi composto por litro de 24 gramas de dextrose de batata (Difco) e 30 gramas de Solcafloc (Diacel disponível da Dicalie-Europe-Nord, Gent, Bélgica)

5 O meio de dextrose de batata foi composto por litro de 39 gramas de dextrose de batata (Difco).

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

10 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_2SO_4 , 12 g de KH_2PO_4 , 7 g de extrato de levedura, 2 g de uréia e 0,5 ml de solução de metais traço AMG, pH ajustado para 5,0.

A solução de metais traço AMG foi composta 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 $\text{MnSO}_4 \cdot \text{H}_2\text{O}$ e 3 g de ácido cítrico.

15 As placas COVE foram compostas por litro de 342,3 g de sacarose, 20 ml de solução salina de COVE, 10 ml de acetamida 1 M, 10 ml de CsCl_2 1,5 M e 25 g de Noble ágar.

20 A solução salina de COVE foi composta por litro de 26 g de KCl , 26 g de $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 76 g de KH_2PO_4 e 50 ml de solução de metais traço COVE.

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 A solução de metais traço foi composta por litro de 41,2 mg de $\text{FeCl}_3 \cdot 6\text{H}_2\text{O}$, 11,6 mg de $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$, 5,4 mg de $\text{MnSO}_4 \cdot \text{H}_2\text{O}$, 2,0 mg de $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$, 0,48 mg de H_3BO_3 e 67,2 mg de ácido cítrico.

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 , 10 mM de MgSO_4 e esterilizado por filtração glicose a 20 mM, adicionados depois da

autoclavagem.

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

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.

O TE foi composto de 10 mM de Tris pH 7,4 e 0,1 mM de EDTA.

Exemplo 1: Extração de DNA genômico de *Thielavia terrestris* NRRL 8126

Thielavia terrestris NRRL 8126 foi cultivada em 25 ml de meio YEG a 37°C e 250 rpm por 24 horas. Micélios foram depois coletados pela filtração através Miracloth® (CalBiochem, La Jolla, CA, USA) e lavado uma vez com 25 ml de tampão de 10 mM de Tris-1 mM de EDTA (TE). O tampão em excesso foi drenado da preparação de micélios, que foi subsequentelemente congelada em nitrogênio líquido. A preparação de micélios congelados foi triturada a um pó fino em um moedor de café elétrico e o pó foi adicionado a um tubo de centrifuga plástico descartável contendo 20 ml de tampão de TE e 5 ml de 20% p/v de dodecil sulfato de sódio (SDS). A mistura foi suavemente invertida várias vezes para garantir a mistura e extraída duas vezes com um volume igual de fenol:clorofórmio:álcool isoamílico (25:24:1 v/v/v). Acetato de sódio (solução 3 M) foi adicionada à amostra extraída a uma concentração final de 0,3 M seguida por 2,5 volumes de etanol gelado para precipitar o DNA. O tubo foi centrifugado a 15.000 x g por 30 minutos para pelotizar o DNA. A pelota de DNA foi deixada secar ao ar por 30 minutos antes da recolocação em suspensão em 0,5 ml de tampão TE. A ribonuclease A isenta de Dnase foi adicionada à pelota de DNA recolocada em suspensão a uma concentração de 100 µg por ml e a mistura foi depois incubada a 37°C por 30 minutos. A proteinase K (200 µg/ml) foi adicionada e o tubo foi incubado um adicional de uma hora a 37°C. Finalmente, a amostra

foi centrifugada por 15 minutos a 12.000 x g e o sobrenadante foi aplicado a um tubo de distribuição Qiaprep® 8 (QIAGEN Inc., Valencia, CA, USA). As colunas foram lavadas duas vezes com 1 ml de PB (QIAGEN Inc., Valencia, CA, USA) e 1 ml de PE (QIAGEN Inc., Valencia, CA, USA) sob vácuo. O DNA isolado foi eluído com 100 µl de TE, precipitado com etanol, lavado com 70% de etanol, secado sob vácuo, recolocado em suspensão em tampão TE e armazenado a 4°C.

Para gerar o DNA genômico para a amplificação pela PCR, *Thielavia terrestris* NRRL 8126 foi cultivada em 50 ml de meio NNCYP suplementado com 1% de glicose em um frasco agitado com defletor a 42°C e 200 rpm por 24 horas. Os micélios foram colhidos pela filtração, lavado duas vezes em TE (10 mM de Tris-1 mM de EDTA) e congelado sob nitrogênio líquido. Um pedaço do tamanho de uma ervilha de micélios congelados foi colocado em suspensão em 0,7 ml de dodecil sulfato de lítio a 1% em TE e rompidos pela agitação com um volume igual de pérolas de zirconia/sílica de 0,1 mm (Biospec Products, Inc., Bartlesville, OK, USA) por 45 segundos em um FastPrep FP120 (ThermoSavant, Holbrook, NY, USA). Os escombros foram removidos pela centrifugação a 13.000 x g por 10 minutos e o sobrenadante depurado foi levado a 2,5 M de acetato de amônio e incubados em gelo por 20 minutos. Depois do período de incubação, os ácidos nucleicos foram precipitados pela adição de 2 volumes de etanol. Depois da centrifugação por 15 minutos em uma microcentrífuga a 4°C, a pelota foi lavada em 70% de etanol e secada ao ar. O DNA foi recolocado em suspensão em 120 µl de 0,1X TE e incubados com 1 µl de RNase A isento de DNase a 37°C por 20 minutos. Acetato de amônio foi adicionado a 2,5 M e o DNA foi precipitado com 2 volumes de etanol. A pelota foi lavada em etanol a 70%, secada ao ar e recolocada em suspensão em tampão TE.

Exemplo 2: A amplificação pela PCR de um fragmento de gene cel6b de DNA genômico de *Thielavia terrestris* NRRL 8126

Os iniciadores foram planejados com base nos motivos conservados encontrados em outras glicosil hidrolases da Família 6. As seqüências de peptídeo específicas usadas para planejar o iniciador foram:

5 EPDSLANLVT (correspondente aos aminoácidos de 167 a 176 da SEQ ID NO: 2)
W[I,V]KPGGE[C,S] (aminoácidos de 343 a 350 da SEQ ID NO: 2)

A estratégia CODEHOP foi utilizada (Rose *et al.*, 1998, Nucleic Acids Res. 26: 1628-1635) para planejar os seguintes iniciadores:

10 Iniciador de Sentido:

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

Iniciador de Anti-sentido:

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

15 A amplificação foi realizada em um volume de 30 µl contendo tampão 1X AmpliTaq (Applied Biosystems, Foster City, CA, USA), 2,5 unidades de AmpliTaq® DNA polimerase (Applied Biosystems, Foster City, CA, USA), 1 µM de cada iniciador de sentido e anti-sentido e aproximadamente 1 µg de DNA genômico de *Thielavia terrestris*. A amplificação foi realizada em um Robocycler® (Stratagene, La Jolla, CA, 20 USA) programado por 1 ciclo a 96°C por 3 minutos e 72°C por 3 minutos (durante o que a DNA polimerase foi adicionada); e 35 ciclos cada um a 94°C por 45 segundos, 58°C por 45 segundos e 72°C por 1 minuto; seguidos por uma extensão final a 72°C por 7 minutos.

25 Os produtos da reação foram fracionados em um gel de agarose a 1,6% usando 40 mM de tampão Tris base-20 mM de acetato de sódio-1 mM de EDTA dissódico (TAE). e uma faixa de aproximadamente 800 pares de base foi excisada, purificada usando um Kit de Extração em Gel QIAEX® II (QIAGEN Inc., Valencia, CA, USA) e subclonada usando um Kit TOPO TA (Invitrogen, Carlsbad, CA, USA). O plasmídeo de um

transformante da *E. coli* foi sequenciado e descoberto conter um inserto de 789 pares de base que codifica uma proteína da Família 6 (CEL6B). Este plasmídeo foi designado pPH23 (Figura 1).

Exemplo 3: Construção e triagem de biblioteca de DNA genômico de

5 *Thielavia terrestris* NRRL 8126

Uma biblioteca de DNA genômico de *Thielavia terrestris* NRRL 8126 foi construída usando o vetor de clonagem de bacteriófago hZipLox (Life Technologies, Gaithersburg, MD, USA) com células Y1090ZL da *E. coli* (Life Technologies, Gaithersburg, MD, USA) como um hospedeiro para plaqueamento e purificação de bacteriófago recombinante e *E. coli* DH10Bzip (Life Technologies, Gaithersburg, MD, USA) para a excisão de clones pZL1 individuais contendo o gene cel6b.

DNA genômico de *Thielavia terrestris* NRRL 8126 preparado como descrito no Exemplo 1 foi parcialmente digerido com Tsp 5091 e fracionados por tamanho em géis de agarose a 1% usando tampão TAE. Os fragmentos de DNA que migram na faixa de tamanho de 3 a 7 kb foram excisados e eluídos a partir do gel usando reagentes Prep-a-Gene (BioRad Laboratories, Hercules, CA, USA). Os fragmentos de DNA eluídos foram ligados com ramos do vetor AZipLox clivados e desfosforilados com *Eco* RI (Life Technologies, Gaithersburg, MD, USA) e as misturas de ligação foram embaladas usando extratos de embalagem comerciais (Stratagene, La Jolla, CA, USA). As bibliotecas de DNA embaladas foram plaqueadas e amplificadas em células Y1090ZL da *E. coli*. A biblioteca de DNA genômico não amplificada conteve $3,1 \times 10^6$ pfu/ml (títulos de fundo sem nenhum DNA foram $2,0 \times 10^4$ pfu/ml).

Um fragmento de sonda cel6b de *Thielavia terrestris* foi amplificado a partir de pPH23 usando iniciadores homólogos ao vetor TOPO e Herculase® DNA Polymerase (Stratagene, La Jolla, CA, USA) como mostrado abaixo.

5'-CTTGGTACCGAGCTCGGATCCACTA-3' (SEQ ID NO: 7)

5'-ATAGGGCGAATTGGGCCCTCTAGAT-3' (SEQ ID NO: 8)

Cinquenta picomoles de cada um dos iniciadores foram usados em uma reação de PCR contendo 10 ng de pPH23, 1X Tampão de Amplificação Herculase® (Stratagene, La Jolla, CA, USA), 1 µl de mistura a 10 mM de dATP, dTTP, dGTP e dCTP e 2,5 unidades de Herculase® DNA Polymerase em um volume final de 50 µl. A amplificação foi realizada em um Robociclor® programado para 1 ciclo a 94°C por 1 minuto; e 20 ciclos cada um a 94°C por 30 segundos, 55°C por 30 segundos e 72°C por 1 minuto. O bloco de aquecimento depois foi para um ciclo de embebeção a 4°C.

O produto da reação foi isolado em um gel de agarose a 1,0% usando tampão TAE onde uma faixa de produto de 0,8 kb foi excisada a partir do gel e purificada usando um Kit de Extração em Gel QIAquick® (QIAGEN Inc., Valencia, CA, USA) de acordo com as instruções do fabricante. Uma amostra de vinte e cinco ng do fragmento foi radiorrotulada com ³²P usando um Kit Prime-It® II (Stratagene, La Jolla, CA, USA).

Aproximadamente 90.000 placas da biblioteca foram triadas pela hibridização de placa usando o fragmento de PCR rotulado como a sonda. O DNA foi reticulado sobre membranas (Hybond N+, Amersham, Arlington Heights, IL, USA) usando um UV Stratalinker® (Stratagene, La Jolla, CA, USA). O fragmento de gene radiorrotulado com ³²P foi desnaturado pela adição de hidróxido de sódio a uma concentração final de 0,1 M e adicionado a uma solução de hibridização contendo 6X SSPE, 7% de SDS a uma atividade de aproximadamente 1 x 10⁶ cpm por ml de solução de hibridização. A mistura foi incubada durante a noite a 65°C em um banho de água sob agitação. A seguir da incubação, as membranas foram lavadas três vezes por cinquenta minutos em 0,2X SSC com 0,1% de SDS a 65°C. As membranas foram secadas em mata-borrão por 15 minutos, enroladas em SaranWrap® e expostas à película de raio X durante a noite a 70°C com telas

intensificadoras (Kodak, Rochester, NY, USA).

Com base na produção de sinais de hibridização fortes com a sonda cel6b descrita acima, diversas placas foram escolhidas para outro estudo. As placas foram purificadas duas vezes em células *E. coli* Y1090ZL e os genes inseridos e o plasmídeo pZL1 foram subsequentemente excisados do vetor AZipLox como derivados de pZL1 (D'Alessio *et al.*, 1992, Focus® 14: 76) usando a excisão *in vivo* pela infecção de células *E. coli* DH10BZL (Life Technologies, Gaithersburg, MD, USA). As colônias foram inoculadas em três ml de meio LB ampicilina e cultivadas durante a noite a 37°C. O DNA Miniprep foi preparado a partir de cada uma destas culturas usando um BioRobot 9600 (QIAGEN Inc., Valencia, CA, USA). Um clone designado pPH47 foi mostrado pelo seqüenciamento de DNA conter o gene de tamanho natural para cel6b.

A cepa da *E. coli* PaHa47 contendo o plasmídeo pPH47 foi depositada com a Agricultural Research Service Patent Culture Collection, Northern Regional Research Center, 1815 University Street, Peoria, Illinois, 61604, como NRRL B-30898, com uma data de depósito de 23 de fevereiro de 2006.

Exemplo 4: Caracterização da seqüência genômica de *Thielavia terrestris* NRRL 8126 que codifica uma CEL6B endoglucanase

O seqüenciamento de DNA do clone genômico cel6b da *Thielavia terrestris* NRRL 8126 foi realizado com um Seqüenciador de DNA Automatizado Modelo 3700 da Applied Biosystems usando a versão 3.1 do química de terminador Big-Dye® (Applied Biosystems, Inc., Foster City, CA, USA) e química dGTP (Applied Biosystems, Inc., Foster City, CA, USA) e a estratégia de caminhante de iniciador. Os dados da seqüência de nucleotídeo foram escrutinizados quanto a qualidade e todas as seqüências foram comparadas entre si com a assistência do software PHRED/PHRAP (University of Washington, Seattle, WA, USA).

Um modelo de gene para a sequência de DNA genômico de *Thielavia terrestris* cel6b foi construída com base na similaridade com os genes da endoglucanase homóloga de *Fusarium oxysporum* e *Chrysosporium lucknowense* (Número de acesso XP383804 e AAQ38151,1, respectivamente).

A sequência de nucleotídeo (SEQ ID NO: 1) e a sequência de aminoácido deduzida (SEQ ID NO: 2) do gene cel6b da *Thielavia terrestris* são mostrados na Figura 2. A sequência codificadora tem 1505 pares de base incluindo o códon de parada e é interrompida pelos introns de 77, 127 e 110 pares de base. A proteína prognosticada codificada tem 396 aminoácidos, A% de G + C da sequência codificadora do gene (incluindo introns) é de 64,9% de G + C e a sequência que codifica o polipeptídeo maduro é de 64,6%. Usando o programa SignalP (Nielsen *et al.*, 1997, Protein Engineering 10: 1-6), um peptídeo de sinal de 18 resíduos foi prognosticada. A proteína madura prognosticada contém 378 aminoácidos com uma massa molecular de 40,6 kDa e um pH isoelétrico de 5,01.

Um alinhamento global aos pares comparativo de sequências de aminoácido foi determinado usando o algoritmo de Needleman-Wunsch (Needleman e Wunsch, 1970, J. Mol. Biol. 48: 443-453) como implementado no programa Needle de EMBOSS com penalidade de abertura de intervalo de 10, penalidade de extensão de intervalo de 0,5 e a matriz EBLOSUM62. O alinhamento mostrou que a sequência de aminoácido deduzida do gene da *Thielavia terrestris* que codifica o polipeptídeo CEL6B tendo atividade de endoglucanase compartilha 85% e 82% de identidade (excluindo os intervalos) com a sequência de aminoácido deduzida de duas proteínas de glicosil hidrolase Família 6 de *Chrysosporium lucknowense* e *Neurospora crassa*, respectivamente (números de acesso AAQ38151 e Q7RXI7, respectivamente).

Exemplo 5: Clonagem e expressão de um cDNA *Thielavia terrestris*

NRRL 8126 que codifica uma endoglucanase CEL6C

A *Thielavia terrestris* NRRL 8126 foi cultivada em 200 ml de meio PD com celulose a 30°C por cinco dias a 200 rpm. Micélios dos frascos de cultura agitados foram colhidos pela filtração dos conteúdos através de um funil revestido com Miracloth®. Os micélios foram depois intercalados entre dois pedaços de Miracloth® e manchados secos com toalhas de papel absorventes. A massa micelial foi depois transferida para tubos de centrífuga plásticos Falcon® 1059 e congelados em nitrogênio líquido. Os micélios congelados foram armazenados em um congelador a -80°C até o uso.

A extração de RNA total foi realizada com tiocianato de guanidínio seguido pela ultracentrifugação através de uma almofada de CsCl a 5,7 M e a isolamento de poli(A)+RNA foi realizada pela cromatografia de afinidade em oligo(dT)-celulose, usando os procedimentos descritos na WO 94/14953.

O cDNA de filamento duplo foi sintetizado a partir de 5 µg de poli(A)+ RNA pelo método da RNase H (Gubler e Hoffman, 1983, Gene 25: 263-269, Sambrook *et al.*, 1989, Molecular cloning: A laboratory manual, Cold Spring Harbor lab., Cold Spring Harbor, NY. USA). A poli(A)⁺ RNA (5 µg em 5 µl de água tratada com DEPC (pirocarbonato de dietila a 0,1%)) foi aquecida a 70°C por 8 minutos em um tubo de Eppendorf® pré siliconizado, isento de RNase, extinto em gelo e combinado em um volume final de 50 µl com composto de tampão de transcriptase reversa de 50 mM de Tris-HCl, pH 8,3, 75 mM de KCl, 3 mM de MgCl₂, 10 mM de ditioneitol (DTT) (Bethesda Research Laboratories, Bethesda, MD, USA). 1 mM de dATP, dGTP e dTTP e 0,5 mM de 5-metil-dCTP (Pharmacia, Uppsala, Suécia), 40 unidades de inibidor da ribonuclease placentária humana (RNasin, Promega, Madison, WI, USA), 1,45 µg de iniciador oligo(dT)₁₀-Not I (Pharmacia, Uppsala, Suécia) e 1000 unidades de transcriptase reversa SuperScript® II RNase H (Bethesda Research Laboratories, Bethesda, MD, USA). O cDNA do primeiro filamento

foi sintetizado incubando-se a mistura de reação a 45°C por 1 hora. Depois da síntese, a mistura híbrida de mRNA:cDNA foi filtrada em gel através de uma coluna giratória MicroSpin 5-400 HR (Pharmacia, Uppsala, Suécia) de acordo com as instruções do fabricante.

5 Depois da filtração em gel, os híbridos foram diluídos em 250 µl de tampão do segundo filamento (20 mM de Tris-HCl, pH 7,4, 90 mM de KCl, 4,6 mM de MgCl₂, 10 mM de (NH₄)₂SO₄, 0,16 mM de NAD) contendo 200 µM de cada dNTP, 60 unidades de DNA polimerase I da *E. coli* (Pharmacia, Uppsala, Suécia), 5,25 unidades de RNase H (Promega, Madison, WI, USA) e 15 unidades de DNA ligase da *E. coli* (Boehringer Mannheim, Mannheim, Alemanha). A síntese de cDNA do segundo filamento foi realizada incubando-se o tubo de reação a 16°C por 2 horas e um adicional de 15 minutos a 25°C. A reação foi interrompida pela adição de EDTA a uma concentração final de 20 mM seguido pelas extrações com fenol e

10

15 clorofórmio.

 O cDNA de filamento duplo foi precipitado a -20°C por 12 horas pela adição de 2 volumes de 96% de etanol e 0,2 volume de acetato de amônio 10 M, recuperado pela centrifugação a 13.000 x g, lavado em etanol a 70%, secado e recolocado em suspensão em 30 µl de tampão de nuclease de feijão Mung (30 mM de acetato de sódio pH 4,6, 300 mM de NaCl, 1 mM de ZnSO₄, 0,35 mM de DTT, 2% de glicerol) contendo 25 unidades de nuclease de feijão Mung (Pharmacia, Uppsala, Suécia). O DNA de grampo de cabelo de filamento único foi aparado pela incubação da reação a 30°C por 30 minutos, seguida pela adição de 70 µl de 10 mM de Tris-HCl-1 mM de EDTA

20

25 pH 7,5, extração com fenol e precipitação com 2 volumes de etanol a 96% e 0,1 volume de acetato de sódio 3 M pH 5,2 em gelo por 30 minutos.

 Os cDNAs de filamento duplo foram recuperados pela centrifugação a 13.000 x g e terminados abrupto em 30 µl de tampão de T4 DNA polimerase (20 mM de Tris-acetato, pH 7,9, 10 mM de acetato de

magnésio, 50 mM de acetato de potássio, 1 mM de DTT) contendo 0,5 mM de cada dNTP e 5 unidades de T4 DNA polimerase (New England Biolabs, Ipswich, MA, USA) pela incubação da mistura de reação a 16°C por 1 hora. A reação foi interrompida pela adição de EDTA a uma concentração final de 20 mM, seguida pelas extrações com fenol e clorofórmio e precipitação por 12 horas a -20°C pela adição 2 volumes de etanol a 96% e 0,1 volume de acetato de sódio 3 M pH 5,2.

Depois da reação de enchimento, os cDNAs foram recuperados pela centrifugação a 13.000 x g, lavados em etanol a 70% e secados. A pelota de cDNA foi recolocada em suspensão em 25 µl de tampão de ligação (30 mM de Tris-HCl, pH 7,8, 10 mM de MgCl₂, 10 mM de DTT, 0,5 mM de ATP) contendo 2,5 µg de adaptadores Bst XI não palindrômicos (Invitrogen, Carlsbad, CA, USA), mostrados abaixo e 30 unidades de T4 ligase (Promega, Madison, WI, USA) e depois incubados a 16°C por 12 horas. A reação foi interrompida aquecendo-se a 65°C por 20 minutos e depois esfriado em gelo por 5 minutos.

5'-CTTTCCAGCACA-3' (SEQ ID NO: 9)

3'-GAAAGGTC-5'

O cDNA adaptado foi digerido com *Not* I, seguido pela incubação por 2,5 horas a 37°C. A reação foi interrompida aquecendo-se a 65°C por 10 minutos. Os cDNAs foram fracionados por tamanho pela eletroforese em gel em um gel de agarose de temperatura de fusão baixa SeaPlaque® GTG a 0,8% (Cambrex Corporation, East Rutherford, NJ, USA) em 44 mM de Tris Base, 44 mM de ácido bórico, 0,5 mM de tampão de EDTA (TBE) para separar adaptadores não ligados e cDNAs pequenos. O cDNA foi selecionado por tamanho com um corte a 0,7 kb e recuperados do gel pelo uso de beta-agarose (New England Biolabs, Ipswich, MA, USA) de acordo com as instruções do fabricante e precipitado por 12 horas a -20°C pela adição de dois volumes de etanol a 96% e 0,1 volume de acetato de sódio

3 M pH 5,2.

O cDNA direcional, selecionado por tamanho foi recuperado pela centrifugação a 13.000 x g, lavado em etanol a 70%, secado e depois recolocado em suspensão em 30 µl de 10 mM de Tris-HCl-1 mM de EDTA pH 7,5. Os cDNAs foram dessalinizados pela filtração em gel através de uma
5 coluna giratória MicroSpin S300 HR de acordo com as instruções do fabricante. Três ligações de teste foram realizadas em 10 µl de tampão de ligação (30 mM de Tris-HCl, pH 7,8, 10 mM de MgCl₂, 10 mM de DTT, 0,5 mM de ATP) contendo 5 µl de cDNA de filamento duplo (tubos de reação #1
10 e #2), 15 unidades de T4 ligase (Promega, Madison, WI, USA) e 30 ng (tubo #1), 40 ng (tubo #2) e 40 ng (tubo #3, o controle de fundo do vetor) do vetor pYES2.0 clivado com *Bst* XI-Not I (Invitrogen, Carlsbad, CA, USA). As reações de ligação foram realizadas pela incubação a 16°C por 12 horas, depois aquecimento a 70°C por 20 minutos e finalmente adicionando 10 µl de
15 água a cada tubo. Um µl de cada mistura de ligação foi submetido à eletroporação em 40 µl de células de *E. coli* DH10B eletrocompetentes (Bethesda Research Laboratories, Bethesda, MD, USA) como descrito por Sambrook *et al.*, 1989, *supra*.

A biblioteca de cDNA de *Thielavia terrestris* NRRL 8126 foi
20 estabelecida como reuniões na *E. coli* DH10B. Cada reunião foi feita espalhando-se placas de *E. coli* em LB ampicilina transformadas, produzindo 15.000 a 30.000 colônias/placa depois da incubação a 37°C por 24 horas. Vinte ml de meio de LB ampicilina foram adicionados às placas e as células foram colocadas em suspensão neste. A suspensão de célula foi agitada a 100
25 rpm em um tubo de 50 ml por 1 hora a 37°C.

A biblioteca de cDNA de *Thielavia terrestris* NRRL 8126 resultante consistiu de aproximadamente 10⁶ clones individuais, com um fundo de vetor de 1%. O DNA plasmídico de algumas das reuniões da biblioteca foi isolado usando um Kit Plasmid Midi (QIAGEN Inc., Valencia,

CA. USA) de acordo com as instruções do fabricante e armazenado a -20°C.

Alíquotas de um ml de DNA plasmídico purificado (100 ng/ml) de algumas das reuniões da biblioteca (Exemplo 1) foram transformadas em *Saccharomyces cerevisiae* W3124 pela eletroporação (Becker e Guarante, 1991, Methods Enzymol. 194: 182-187) e os transformantes foram plaqueados em ágar SC contendo 2% de glicose e incubados a 30°C. No total, 50 a 100 placas contendo 250 a 400 colônias de levedura foram obtidas a partir de cada reunião.

Depois de 3 a 5 dias de incubação, as placas de ágar SC foram plaqueadas em réplica em um conjunto de placas de 0,1% de AZCL HE celulose SC URA ágar com galactose. As placas foram incubadas por 2 a 4 dias a 30°C e as colônias positivas em endoglucanase foram identificadas como colônias circundadas por um halo azul.

Exemplo 6: Caracterização da seqüência de cDNA de *Thielavia terrestris* NRRL 8126 que codifica uma CEL6C endoglucanase

As colônias de levedura que expressam a endoglucanase foram inoculadas em 20 ml de meio YPD em tubos de teste de vidro de 50 ml. Os tubos foram agitados a 200 rpm por 2 dias a 30°C. As células foram colhidas pela centrifugação por 10 minutos a 3000 rpm em uma centrífuga Heraeus Megafuge 1,0R com um rotor 75002252 (Hanau, Alemanha).

O DNA foi isolado de acordo com a WO 94/14953 e dissolvido em 50 µl de água deionizada. O DNA foi transformado em células *E. coli* DH10B pelos procedimentos padrão de acordo com Sambrook *et al.*, 1989, supra. Um transformante da *E. coli* subseqüentemente mostrado conter o gene *cel6c* da *Thielavia terrestris* NRRL 8126 foi designado pCIBG146 (Figura 3) e usado como material para depósito de material biológico. A cepa pCIBG146 da *E. coli* foi depositada como *E. coli* NRRL B-30901 em 23 de fevereiro de 2006.

O DNA plasmídico foi isolado dos transformantes de *E. coli*

usando procedimentos padrão de acordo com Sambrook *et al.*, 1989, supra. A sequência de cDNA de tamanho natural do gene cel6c da *Thielavia terrestris* NRRL 8126 foi seqüenciada com um Kit Tag DyeDeoxi Terminator Cycle Sequencing (Perkin Elmer, Wellesley, MA, USA) e iniciadores de oligonucleotídeo sintético usando um Seqüenciador de DNA Applied Biosystems ABI PRISM® 377 (ABI, Foster City, CA, USA) de acordo com as instruções do fabricante.

A seqüência de nucleotídeo (SEQ ID NO: 3) e a seqüência de aminoácido deduzida (SEQ ID NO: 4) do gene cel6c da *Thielavia terrestris* são mostrados na Figura 4. A seqüência codificadora tem 1203 pares de base incluindo o códon de parada. A proteína prognosticada codificada tem 400 aminoácidos. A% G + C da seqüência codificadora do gene é de 60,0% e da seqüência que codifica o polipeptídeo maduro é de 60,0%. Usando o programa SignalP (Nielsen *et al.*, 1997, supra), um peptídeo de sinal de 19 resíduos foi prognosticado. A proteína madura prognosticada contém 381 aminoácidos com uma massa molecular de 42,1 kDa e pI de 6,76.

Um alinhamento global aos pares comparativo de seqüências de aminoácido foi determinado usando o algoritmo de Needleman-Wunsch (Needleman e Wunsch, 1970, supra) como implementado no programa Needle de EMBOSS com penalidade de abertura de intervalo de 10, penalidade de extensão de intervalo de 0,5 e a matriz EBLOSUM62. O alinhamento mostrou que a seqüência de aminoácido deduzida do gene de *Thielavia terrestris* que codifica o polipeptídeo CEL6C compartilha 50%, 49% e 48% de identidade (excluindo os intervalos) com a seqüência de aminoácido deduzida de três proteínas da glicosil hidrolase da Família 6 de *Neurospora Grasse*, *Magnaporthe grisea* e *Agaricus bisporus*, respectivamente (números de acesso Q87B5, XP 368004 e P49075, respectivamente).

Exemplo 7: Construção do vetor de expressão pAILo2

O vetor de expressão pALlo1 foi construído modificando-se pBANE6 (Patente U.S. Nº 6.461.837), que compreende um híbrido dos promotores dos genes para a alfa-amilase neutra de *Aspergillus niger* e triose fosfato isomerase de *Aspergillus oryzae* (promotor NA2-tpi), a sequência terminadora da amiloglucosidase de *Aspergillus niger* (terminador AMG) e o gene da acetamidase de *Aspergillus nidulans* (amdS). Todas as etapas de mutagênese foram verificadas pelo seqüenciamento usando a química do terminador Big-Dye® (Applied Biosystems, Inc., Foster City, CA, USA). 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 a partir do marcador de seleção amdS pela mutagênese direcionada ao sítio. Todas as mudanças foram planejadas serem “silenciosas” deixando a sequência 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, USA) de acordo com as instruções do fabricante usando os seguintes iniciadores (o nucleotídeo sublinhado representa a base trocada):

AMDS3NcoMut (2050):

5'-GTGCCCCATGATACGCCTCCGG-3' (SEQ ID NO: 10)

AMDS2NcoMut (2721):

5'-GAGTCGTATTTCCAAGGCTCCTGACC-3' (SEQ ID NO: 11)

AMDS1 NcoMut (3396):

5-GGAGGCCATGAAGTGGACCAACGG-3' (SEQ ID NO: 12)

Um plasmídeo compreendendo todas as três mudanças de sequência esperadas foi depois submetido à mutagênese direcionada ao sítio, usando um Kit de Mutagênese direcionada ao sítio QuickChange® (Stratagene, La Jolla, CA, USA), 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 trocada) foram usados para a

mutagênese:

Iniciador Superior para mutagenizar a sequência terminadora de AMG:

5'-CACCGTGAAAGCCATGCTCTTTCCTTCGTGTAGAAGACCAG
ACAG-3' (SEQ ID NO: 13)

5 Iniciador Inferior para mutagenizar a sequência terminadora de AMG:

5'-CTGGTCTTCTACACGAAGGAMGAGCATGGCTTTCACGGTGT
CTG-3' (SEQ ID NO: 14)

10 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 5).

Iniciador Superior para mutagenizar o promotor NA2-tpi:

15 5'-CTATATACACAACTGGATTTACCATGGGCCCCGCGGCCGCAG
ATC-3' (SEQ ID NO: 15)

Iniciador Inferior para mutagenizar o promotor NA2-tpi:

5'-GATCTGCGGCCGCGGGCCCATGGTAAATCCAGTTGTGTATA
TAG-3' (SEQ ID NO: 16)

20 O gene *amdS* de pAlLo1 foi trocado com o gene *pyrG* *Aspergillus nidulans*. O plasmídeo pBANE10 (Figura 6) foi usado como uma fonte para o gene *pyrG* como um marcador de seleção. A análise da sequência de pBANE10 mostrou que o marcador *pyrG* foi contido dentro de um fragmento de restrição *Nsi* I e não contém os sítios de restrição *Nco* I ou *Pac* I. Visto que o *amdS* também é flanqueado pelos sítios de restrição *Nsi* I a
25 estratégia para mudar o marcador de seleção foi uma troca simples 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 de *Nsi* I de pBANE10 contendo o gene *pyrG* 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 com enzima de restrição para determinar que eles tiveram o inserto correto e também a sua orientação. Um clone com o gene *pyrG* transcrito na direção anti-horária foi selecionado. O novo plasmídeo foi designado pAlLo2 (Figura 7).

Exemplo 8: Expressão do gene da endoglucanase cel6b de *Thielavia terrestris* NRRL 8126 em *Aspergillus oryzae*

Dois iniciadores de oligonucleotídeo sintéticos, mostrados abaixo, foram planejados para amplificar pela PCR a matriz de leitura aberta de tamanho natural do plasmídeo pPH47 que codifica uma CEL6B endoglucanase. Um Kit de Clonagem In-Fusion (BD Biosciences, Palo Alto, CA, USA) foi usado para fabricar o fragmento diretamente no pAlLo2.

Iniciador Avançado In-Fusion:

5'-ACTGGATTACCATGAAGCTCTCGCAGTCG-3' (SEQ ID NO: 17)

Iniciador Reverso In-Fusion:

5'-AGTCACCTCTAGTTAGAACGACGGCACGGC-3' (SEQ ID NO: 18)

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

Cinquenta picomoles de cada um dos iniciadores acima foram usados em uma reação de PCR contendo 50 ng de DNA pPH47, Tampão de Amplificação 1X Pfx (Invitrogen, Carlsbad, CA, USA), 6 µl de mistura a 10 mM de dATP, dTTP, dGTP e dCTP, 2,5 unidades de Platinum Pfx DNA Polimerase (Invitrogen, Carlsbad, CA), 1 µl de MgSO₄ a 50 mM e 5 µl de solução realçadora 10X pCRx (Invitrogen, Carlsbad, CA) em um volume final de 50 µl. A amplificação foi realizada em um Eppendorf® Mastercycler® 5333 (Eppendorf Scientific, Inc., Westbury, NY, USA) programado por 1 ciclo a 98°C por 2 minutos e 35 ciclos cada um a 94°C por 30 segundos, 65°C por 30 segundos e 68°C por 1,5 minutos. Depois dos 35 ciclos, a reação foi incubada

a 68°C por 10 minutos e depois esfriada a 10°C até processada mais adiante. Um produto da reação de PCR de 1,5 kb foi isolado em um GTG a 0,8%-gel de agarose (Cambrex Bioproducts, East Rutherford, NJ, USA) usando tampão 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, USA) para evitar as mutações induzidas por UV. A faixa de DNA de 1,5 kb foi excisada com uma lâmina de barbear descartável e purificada usando um Kit de Extração em Gel QIAquick® (QIAGEN Inc., Valencia, CA, USA) de acordo com as instruções do fabricante.

O vetor pAlLo2 foi linearizado pela digestão com *Nco* I e *Pac* I (usando as condições especificadas pelo fabricante). O fragmento foi purificado pela eletroforese em gel e um Kit de Extração em Gel QIAquick como descrito acima. A clonagem do fragmento de PCR purificado no vetor pAlLo2 linearizado purificado foi realizado com um Kit de Clonagem In-Fusion. A reação (20 µl) conteve Tampão 1X In-Fusion (BD Biosciences, Palo Alto, CA, USA), 1X BSA (BD Biosciences, Palo Alto, CA, USA), 1 µl de enzima In-Fusion (diluído 1:10) (BD Biosciences, Palo Alto, CA, USA), 160 ng de pAlLo2 digerido com *Nco* I e *Pac* I e 50 ng do produto de PCR purificado de cel6b de *Thielavia terrestris*. A reação foi incubada na temperatura ambiente por 30 minutos. Uma amostra de 1 µl da reação foi 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 suplementadas com 100 µg de ampicilina por ml. As placas foram incubadas durante a noite a 37°C. Um conjunto de oito clones recombinantes putativos foi selecionado aleatoriamente a partir das placas de seleção e o DNA plasmídico foi preparado a partir de cada uma usando um BioRobot 9600. Os clones foram analisados pela digestão de restrição com *Eco* RI. Dois clones que tiveram o

padrão de digestão de restrição esperado foram depois seqüenciados para confirmar que não houve nenhuma mutação no inserto clonado. Um dos clones teve a seqüência correta e foi selecionado e designado pEJG105 (Figura 8).

Os 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 pEJG105 (assim como pAlLo2 como um controle de vetor) foram usados para transformar protoplastos JAL250 de *Aspergillus oryzae*.

A transformação de *Aspergillus oryzae* JAL250 com pEJG105 produziu cerca de 100 transformantes. Cinco transformantes foram isolados em placas de PDA individuais e incubadas por cinco dias a 34°C.

As placas de esporo confluentes foram lavadas com 5 ml de Tween 80 a 0,01% e a suspensão de esporo foi usada para inocular 25 ml de meio MDU2BP em frascos agitados de 125 ml. as culturas de transformante foram incubadas a 34°C com agitação constante a 200 rpm. No dia cinco após a inoculação, uma alíquota de cada cultura foi centrifugada a 12000 x g. Cinco µl de cada sobrenadante foram misturados com um volume igual de 2X tampão de carga (10% de beta-mercaptoetanol) e carregado em um gel de SDS-PAGE de 1,5 mm 8%-16% de Tris-Glicina e tingido com SimplyBlue® SafeStain (Invitrogen, Carlsbad, CA). Os perfis de SDS-PAGE dos caldos de cultura mostraram que quatro dos cinco transformantes tiveram uma nova faixa de proteína de aproximadamente 40 kDa. O transformante de número 1 foi selecionado para estudos adicionais e designado EJG105 de *Aspergillus oryzae*.

Exemplo 9: Expressão do gene da endoglucanase cel6c de *Thielavia terrestris* NRRL 8126 em *Aspergillus oryzae*

O gene cel6c de *Thielavia terrestris* NRRL 8126 foi excisado do vetor pYES2.0 usando *Bam* HI e *Xba* I e ligado no vetor de expressão de

Aspergillus pHD414 (EP 238 023, WO 93/11249) usando métodos padrão (Sambrook *et al.*, 1989, *supra*). O vetor de expressão de *Aspergillus* pHD414 é um derivado de p775 (EP 238 023). O plasmídeo resultante foi designado pA2BG146 (Figura 9).

5 Os protoplastos de *Aspergillus oryzae* HowB104 foram preparados como descrito na WO 95/02043. Cem microlitros de suspensão foram misturados com 5 a 25 µg de pA2BG146 em 10 µl de STC composto de 1,2 M de sorbitol, 10 mM de Tris-HCl, pH 7,5, 10 mM de CaCl₂) e ainda misturado com 5 a 25 µg de p3SR2, um gene amdS de *Aspergillus nidulans*
10 que carrega o plasmídeo (Christensen *et al.*, 1988, Bio/Technology 6: 1419-1422). A mistura foi deixada na temperatura ambiente por 25 minutos. Duzentos microlitros de 60% de PEG 4000 (BDH, Poole, Inglaterra) (polietileno glicol, peso molecular 4.000), 10 mM de CaCl₂ e 10 mM de Tris-HCl pH 7,5 foram adicionados e suavemente misturados e finalmente 0,85 ml
15 da mesma solução foi adicionado e suavemente misturado. A mistura foi deixada na temperatura ambiente por 25 minutos, centrifugada a 2.500 x g por 15 minutos e a pelota foi recolocada em suspensão em 2 ml de sorbitol 1,2 M. Este processo de sedimentação foi repetido e os protoplastos foram espalhados em placas COVE. Depois da incubação por 4 a 7 dias a 37°C os
20 esporos foram escolhidos e espalhados de modo a isolar colônias únicas. Este procedimento foi repetido e os esporos de uma única colônia depois da segunda re-isolação foram armazenados.

Cada um dos transformantes foi inoculado em 10 ml de meio YPM. Depois de 2 a 5 dias de incubação a 30°C, 200 rpm, o sobrenadante foi
25 removido. A atividade de endoglucanase foi identificada aplicando-se 20 µl de caldo de cultura a furos de 4 mm de diâmetro perfurados em uma placa a 0,1% AZCL HE celulose SC-ágar e incubação durante a noite a 30°C. A presença de atividade de endoglucanase produziu um halo azul em torno de uma colônia. Vários caldos transformantes tiveram atividade de

endoglucanase que foi significativamente maior do que o caldo de um controle de fundo de *Aspergillus oryzae* não transformado, que demonstrou expressão eficiente da CEL6C endoglucanase de *Thielavia terrestris* NRRL 8126 em *Aspergillus oryzae*.

5 **Exemplo 10: Culturas de frasco agitado grande de *Aspergillus oryzae* Jal250 contendo o gene cel6b**

Aspergillus oryzae Jal250 contendo esporos pEJG105 foram espalhados sobre uma placa de PDA e incubados por cinco dias a 34°C. A placa de esporo confluyente foi lavada duas vezes com 5 ml de 0,01% de
10 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 foi incubada a 34°C com agitação constante (200 rpm). No dia cinco após o inóculo, o caldo de cultura foi coletado pela filtração em uma unidade de filtração de náilon de 500
15 mililitros, 75 mm com um tamanho de poro de 0,45 µm com um pré filtro de fibra de vidro (Nalgene Nunc International, Rochester, NY, USA). Uma amostra de 5 µl do caldo foi analisada pela SDS-PAGE como descrito no Exemplo 9 para confirmar que o padrão de proteína foi o mesmo como aquele obtido antes. O caldo foi mostrado conter uma faixa de proteína de 40 kDa.

20 **Exemplo 11: Hidrólise de carboximetilcelulose pelas endoglucanases de *Thielavia terrestris* CEL6B e *Thielavia terrestris* CEL6C**

O sal de sódio de carboximetilcelulose (CMC, tipo 7L2) com grau médio de substituição (DS) de 0,7 foi obtido da Aqualon Division of Hercules Inc. (Wilmington, DE, USA). Uma solução de 6,25 mg por ml de
25 CMC em 50 mM de acetato de sódio pH 5,0 foi preparada adicionando-se lentamente CMC a um tampão vigorosamente agitado, seguido pelo aquecimento a aproximadamente 60°C sob agitação contínua até a dissolução completa.

As endoglucanases de *Thielavia terrestris* CEL6B e *Thielavia*

terrestris CEL6C (ambas expressadas em *Aspergillus oryzae* Jal250 como descrito nos Exemplos 8 e 9, respectivamente) foram concentradas e dessalinizadas/trocadas para 50 mM de acetato de sódio pH 5,0 usando um filtro centrífugo Centricon® Plus-20 com uma membrana 5000 NMWL Biomax-50 (Millipore, Bedford, MA, USA). As soluções de enzima concentradas e dessalinizadas foram armazenadas a -20°C. A concentração de proteína nas soluções de enzima foi determinada por um Ensaio de Microplaca BCA usando um Kit de Reagente de Ensaio de Proteína BCA de acordo com as instruções do fabricante (Pierce Chemical Co., Rockford, IL, USA).

A hidrólise de CMC foi realizada em Placas Eppendorf® DeepWell 96 (1,2 ml, Eppendorf North America, Inc., Westbury, NY, USA) seladas por um selador de placa ALPS-300® (ABgene Inc., Rochester, NY, USA). Todas as reações foram conduzidas em 50 mM de acetato de sódio pH 5,0 em um volume de 1 ml sem agitação a 50°C. A concentração de substrato inicial foi de 5 mg por ml. Ambas as enzimas foram usadas em quatro concentrações diferentes. Correspondendo a diluições de enzima em 50 mM de acetato de sódio pH 5,0 foram preparadas frescas antes da hidrólise enzimática de soluções de estoque de enzima.

Depois de 70 horas de hidrólise, 40 alíquotas foram transferidas das reações de hidrólise para uma microplaca de 96 reservatórios de fundo chato (Corning Inc., Corning, NY, USA) e misturados com 160 µl de Na₂CO₃ 128 mM-NaHCO₃ 72,5 mM para terminar as reações. As amostras diluídas foram analisadas quanto ao açúcar redutor (RS) usando um ensaio de microplaca de hidróxido do ácido p-hidroxibenzóico (PHBAH) como descrito abaixo.

Uma alíquota de 90 µl da amostra diluída foi colocada em cada reservatório de uma microplaca de fundo cônico de 96 reservatórios (Corning Inc., Corning, NY, USA). O ensaio foi iniciado pela adição de 60 µl

de PHBAH a 1,25% em 2% de hidróxido de sódio a cada reservatório. A placa não coberta foi aquecida em um bloco de aquecimento fabricado de encomenda por 10 minutos a 95°C. Depois a microplaca foi esfriada até a temperatura ambiente, 35 µl de água deionizada foram adicionados a cada reservatório. Uma alíquota de 100 µl foi removida de cada reservatório e transferida para uma microplaca de 96 reservatórios de fundo chato (Corning Inc., Corning, NY, USA). A absorbância a 410 nm (A410) foi medida usando uma Leitora de Microplaca SpectraMAX® (Molecular Devices, Sunnyvale, CA, USA). O valor A410 foi traduzido em equivalentes de glicose usando uma curva padrão.

A curva padrão foi obtida usando seis padrões de glicose (0,005, 0,010, 0,025, 0,050, 0,075 e 0,100 mg/ml), que foram tratados similarmente às amostras.

Os padrões de glicose foram preparados diluindo-se 10 mg/ml de solução de estoque de glicose com água deionizada.

O grau de conversão de carboximetilcelulose para açúcares redutores (rendimento de hidrólise,%) foi calculado usando a seguinte equação:

$$\text{Rendimento de RS (\%)} = \text{RS}_{(\text{mg/ml})} \times 100 \times 162 / (\text{CMC}_{(\text{mg/ml})} \times 180) = \text{RS}_{(\text{mg/ml})} \times 100 / (\text{CMC}_{(\text{mg/ml})} \times 1,111)$$

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

Os resultados de ensaio demonstraram que depois de 70 horas de hidrólise as endoglucanases de *Thielavia terrestris* CEL6B e *Thielavia terrestris* CEL6C foram capazes de hidrolisar de 2 a 4% das ligações glicosídicas disponíveis na carboximetilcelulose.

Depósito de Material Biológico

O seguinte material biológico foi depositado sob os termos do

Tratado de Budapest com o Agricultural Research Service Patent Culture Collection, Northern Regional Research Center, 1815 University Street, Peoria, Illinois, 61604 e dados os seguintes números de acesso:

	Depósito	Número de Acesso	Data de Depósito
5	<i>E. coli</i> PaHa47 (pPH47)	NRRL B-30898	23 de Fevereiro de 2006
	<i>E. coli</i> (pCIBG146)	NRRL B-30901	23 de Fevereiro de 2006

As cepas foram depositadas sob condições que garantem que o acesso às culturas estarão disponíveis durante a pendência deste pedido de patente a uma pessoa determinada pelo Comissário de Patentes e Marcas a ser designado para este sob 37 C.F.R. §1.14 e 35 U.S.C. §122. Os depósitos representam culturas substancialmente puras das cepas depositadas. Os depósitos estão disponíveis como requerido pelas leis de patentes estrangeiras em países em que as contrapartes do pedido objeto ou a sua progênie são depositadas. Entretanto, deve ser entendido que a disponibilidade de um depósito não constitui uma licença para a prática da invenção objeto em detrimento dos direitos de patente outorgados por ação governamental.

A invenção aqui descrita e reivindicada não deve ser limitada no escopo pelos aspectos específicos aqui divulgados, visto que estes aspectos são intencionados como ilustrações de diversos aspectos da invenção. Quaisquer aspectos equivalentes são intencionados a estar dentro do escopo desta invenção. De fato, várias modificações da invenção além daquelas mostradas e aqui descritas tornar-se-ão evidentes para aqueles de habilidade na técnica a partir da descrição precedente. Tais modificações também são intencionadas a cair dentro do escopo das reivindicações anexas. No caso de conflito, a presente divulgação incluindo as definições controlarão.

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

REIVINDICAÇÕES

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

5 (a) um polipeptídeo que compreende uma seqüência de aminoácido tendo pelo menos 90% de identidade com o polipeptídeo maduro da SEQ ID NO: 2 ou pelo menos 60% de identidade com o polipeptídeo maduro da SEQ ID NO: 4;

10 (b) um polipeptídeo que é codificado por um polinucleotídeo que hibridiza sob condições de pelo menos alta estringência com (i) a seqüência que codifica o polipeptídeo maduro da SEQ ID NO: 1. (ii) a seqüência de cDNA contida na seqüência que codifica o polipeptídeo maduro da SEQ ID NO: 1 ou (iii) um filamento complementar de (i) ou (ii) ou sob condições de pelo menos estringência média com (i) a seqüência que codifica o polipeptídeo maduro da SEQ ID NO: 3, (ii) a seqüência de DNA genômico
15 que compreende a seqüência que codifica o polipeptídeo maduro da SEQ ID NO: 3 ou (iii) um filamento complementar de (i) ou (ii);

(c) um polipeptídeo que é codificado por um polinucleotídeo que compreende uma seqüência de nucleotídeo tendo pelo menos 90% de identidade com a seqüência que codifica o polipeptídeo maduro da SEQ ID
20 NO: 1 ou 60% de identidade com a seqüência que codifica o polipeptídeo maduro da SEQ ID NO: 3; e

(d) uma variante que compreende uma substituição, deleção e/ou inserção de um ou mais aminoácidos do polipeptídeo maduro da SEQ ID NO: 2 ou SEQ ID NO: 4.

25 2. Polipeptídeo de acordo com a reivindicação 1, caracterizado pelo fato de que compreende uma seqüência de aminoácido tendo preferivelmente pelo menos 90% de identidade, mais preferivelmente pelo menos 95% de identidade e o mais preferivelmente pelo menos 97% de identidade com o polipeptídeo maduro da SEQ ID NO: 2; ou uma seqüência

de aminoácido tendo preferivelmente pelo menos 60% de identidade, mais preferivelmente pelo menos 65% de identidade, ainda mais preferivelmente pelo menos 70% de identidade, ainda mais preferivelmente pelo menos 75% de identidade, ainda mais preferivelmente pelo menos 80% de identidade, ainda mais preferivelmente pelo menos 85% de identidade, o mais preferivelmente pelo menos 90% de identidade e ainda mais preferivelmente pelo menos 95% de identidade com o polipeptídeo maduro da SEQ ID NO: 4.

3. Polipeptídeo de acordo com a reivindicação 1, caracterizado pelo fato de que compreende ou consiste da seqüência de aminoácido da SEQ ID NO: 2 ou SEQ ID NO: 4; ou um fragmento deste tendo atividade de endoglucanase.

4. Polipeptídeo de acordo com a reivindicação 3, caracterizado pelo fato de que compreende ou consiste da seqüência de aminoácido da SEQ ID NO: 2 ou SEQ ID NO: 4.

5. Polipeptídeo de acordo com a reivindicação 3, caracterizado pelo fato de que compreende ou consiste do polipeptídeo maduro da SEQ ID NO: 2 ou SEQ ID NO: 4.

6. Polipeptídeo de acordo com a reivindicação 1, caracterizado pelo fato de que é codificado por um polinucleotídeo que hibridiza sob condições de pelo menos alta estringência com (i) a seqüência que codifica o polipeptídeo maduro da SEQ ID NO: 1, (ii) a seqüência de cDNA contida na seqüência que codifica o polipeptídeo maduro da SEQ ID NO: 1 ou (iii) um filamento complementar de (i) ou (ii) ou preferivelmente sob condições de pelo menos estringência média, mais preferivelmente condições de estringência média-alta e o mais preferivelmente condições de estringência alta com (i) a seqüência que codifica o polipeptídeo maduro da SEQ ID NO: 3, (ii) a seqüência de DNA genômico que compreende a seqüência que codifica o polipeptídeo maduro da SEQ ID NO: 3 ou (iii) um filamento complementar de (i) ou (ii),

7. Polipeptídeo de acordo com a reivindicação 1, caracterizado pelo fato de que é codificado por um polinucleotídeo que compreende uma seqüência de nucleotídeo tendo preferivelmente pelo menos 90% de identidade, mais preferivelmente pelo menos 95% de identidade e o mais preferivelmente pelo menos 97% de identidade com a seqüência que codifica o polipeptídeo maduro da SEQ ID NO: 1; ou uma seqüência de nucleotídeo tendo preferivelmente pelo menos 60% de identidade, mais preferivelmente pelo menos 65% de identidade, ainda mais preferivelmente pelo menos 70% de identidade, ainda mais preferivelmente pelo menos 75% de identidade, ainda mais preferivelmente pelo menos 80% de identidade, ainda mais preferivelmente pelo menos 85% de identidade, o mais preferivelmente pelo menos 90% de identidade e ainda mais preferivelmente pelo menos 95% de identidade com a seqüência que codifica o polipeptídeo maduro da SEQ ID NO: 3.

8. Polipeptídeo de acordo com a reivindicação 1, caracterizado pelo fato de que é codificado por um polinucleotídeo que compreende ou que consiste da SEQ ID NO: 1 ou SEQ ID NO: 3; ou uma subsequência destas que codifica um fragmento de polipeptídeo tendo atividade de endoglucanase.

9. Polipeptídeo de acordo com a reivindicação 8, caracterizado pelo fato de que é codificado por um polinucleotídeo que compreende ou que consiste da SEQ ID NO: 1 ou SEQ ID NO: 3.

10. Polipeptídeo de acordo com a reivindicação 8, caracterizado pelo fato de que é codificado por um polinucleotídeo que compreende ou que consiste da seqüência que codifica o polipeptídeo maduro da SEQ ID NO: 1 ou SEQ ID NO: 3.

11. 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 de um ou mais aminoácidos do polipeptídeo maduro da SEQ ID NO: 2 ou SEQ ID NO: 4.

12. Polipeptídeo de acordo com a reivindicação 1, caracterizado pelo fato de que é codificado pelo polinucleotídeo contido no plasmídeo pPH47 que está contido na *E. coli* NRRL B-30898 ou plasmídeo pCIBG146 que está contido na *E. coli* NRRL B-30901.

5 13. Polipeptídeo de acordo com as reivindicações 1, 2, 5 ou 11, caracterizado pelo fato de que o polipeptídeo maduro tem os aminoácidos de 19 a 396 da SEQ ID NO: 2 ou os aminoácidos de 20 a 400 da SEQ ID NO: 4.

10 14. Polipeptídeo de acordo com as reivindicações 1 6, 7 ou 10, caracterizado pelo fato de que a seqüência que codifica o polipeptídeo maduro tem os nucleotídeos de 110 a 1557 da SEQ ID NO: 1 ou os nucleotídeos de 58 a 1200 da SEQ ID NO: 3.

15 15. Polipeptídeo de acordo com qualquer uma das reivindicações de 1 a 14, caracterizado pelo fato de que tem ainda atividade de enzima voltada para um ou mais substratos selecionados do grupo que consiste de xilano, xiloglicano, arabinosilano, 1,4-beta-D-manana e galactomanana.

16. Polinucleotídeo isolado, caracterizado pelo fato de que compreende uma seqüência de nucleotídeo que codifica o polipeptídeo como definido em qualquer uma das reivindicações de 1 a 15.

20 17. Polinucleotídeo isolado de acordo com a reivindicação 16, caracterizado pelo fato de que compreende pelo menos uma mutação na seqüência que codifica o polipeptídeo maduro da SEQ ID NO: 1 ou SEQ ID NO: 3, em que a seqüência de nucleotídeo mutante codifica o polipeptídeo maduro da SEQ ID NO. 2 ou SEQ ID NO: 4, respectivamente.

25 18. Polinucleotídeo isolado de acordo com a reivindicação 17, caracterizado pelo fato de que a seqüência que codifica o polipeptídeo maduro tem os nucleotídeos de 110 a 1557 da SEQ ID NO: 1 ou nucleotídeos de 58 a 1200 da SEQ ID NO: 3 e o polipeptídeo maduro tem os aminoácidos de 19 a 396 da SEQ ID NO: 2 ou aminoácidos de 20 a 400 da SEQ ID NO: 4.

19. Construção de ácido nucleico, caracterizada pelo fato de que compreende o polinucleotídeo como definido em qualquer uma das reivindicações de 16 a 18 operavelmente ligado a uma ou mais seqüências de controle que direcionam a produção do polipeptídeo em um hospedeiro de expressão.

20. Vetor de expressão recombinante, caracterizado pelo fato de que compreende a construção de ácido nucleico como definido na reivindicação 19.

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

22. Método para produzir um polipeptídeo como definido em qualquer uma das reivindicações de 1 a 15, caracterizado pelo fato de que compreende: (a) cultivar uma célula, que na sua forma do tipo selvagem produz o polipeptídeo, sob condições condutivas para a produção do polipeptídeo; e (b) recuperar o polipeptídeo.

23. Método para produzir o polipeptídeo como definido em qualquer uma das reivindicações de 1 a 15, caracterizado pelo fato de que compreende: (a) cultivar uma célula hospedeira que compreenda uma construção de ácido nucleico compreendendo uma seqüência de nucleotídeo que codifica o polipeptídeo sob condições condutivas para a produção do polipeptídeo; e (b) recuperar o polipeptídeo.

24. 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 15, que resulta no mutante produzindo menos do polipeptídeo do que a célula precursora.

25. Célula mutante, caracterizada pelo fato de que é produzida pelo método como definido na reivindicação 24.

26. Célula mutante de acordo com a reivindicação 25, caracterizada pelo fato de que compreende ainda um gene que codifica uma proteína nativa ou heteróloga.

5 27. Método para produzir uma proteína, caracterizado pelo fato de que compreende: (a) cultivar a célula mutante de acordo com a reivindicação 26 sob condições condutivas para a produção da proteína; e (b) recuperar a proteína.

10 28. Polinucleotídeo isolado de acordo com qualquer uma das reivindicações de 16 a 18, caracterizado pelo fato de que é obtido pela (a) hibridização de uma população de DNA sob condições de pelo menos alta estringência com (i) a sequência que codifica o polipeptídeo maduro da SEQ ID NO: 1, (ii) a sequência de cDNA contida na sequência que codifica o polipeptídeo maduro da SEQ ID NO: 1 ou (iii) um filamento complementar de (i) ou (ii) ou preferivelmente sob condições de pelo menos estringência média, mais preferivelmente condições de pelo menos estringência média-alta ou o mais preferivelmente condições de pelo menos estringência alta com (i) a sequência que codifica o polipeptídeo maduro da SEQ ID NO: 3, (ii) a sequência de DNA genômico que compreende a sequência que codifica o polipeptídeo maduro da SEQ ID NO: 3 ou (iii) um filamento complementar de (i) ou (ii); e (b) isolar o polinucleotídeo hibridizante, que codifica um polipeptídeo tendo atividade de endoglucanase.

25 29. Polinucleotídeo isolado de acordo com a reivindicação 28, caracterizado pelo fato de que a sequência que codifica o polipeptídeo maduro tem os nucleotídeos de 110 a 1557 da SEQ ID NO: 1 ou nucleotídeos de 58 a 1200 da SEQ ID NO: 3.

30. Método para produzir um polinucleotídeo que compreende uma sequência de nucleotídeo mutante que codifica um polipeptídeo tendo atividade de endoglucanase, caracterizado pelo fato de que compreende: (a) introduzir pelo menos uma mutação na sequência que codifica o polipeptídeo

maduro da SEQ ID NO: 1 ou SEQ ID NO: 3, em que a seqüência de nucleotídeo mutante codifica um polipeptídeo que consiste do polipeptídeo maduro da SEQ ID NO: 2 ou SEQ ID NO: 4; e (b) recuperar o polinucleotídeo que compreende a seqüência de nucleotídeo mutante.

5 31. Método de acordo com a reivindicação 30, caracterizado pelo fato de que a seqüência que codifica o polipeptídeo maduro tem os nucleotídeos de 110 a 1557 da SEQ ID NO: 1 ou nucleotídeos de 58 a 1200 da SEQ ID NO: 3 e o polipeptídeo maduro tem os aminoácidos de 19 a 396 da SEQ ID NO: 2 ou os aminoácidos de 20 a 400 da SEQ ID NO: 4.

10 32. Polinucleotídeo mutante, caracterizado pelo fato de que é produzido pelo método como definido nas reivindicações 30 ou 31.

 33. Método para produzir um polipeptídeo, caracterizado pelo fato de que compreende: (a) cultivar uma célula que compreenda o polinucleotídeo mutante como definido na reivindicação 32 que codifica o polipeptídeo sob condições condutivas para a produção do polipeptídeo; e (b) recuperar o polipeptídeo.

15 34. Construção de ácido nucleico, caracterizada pelo fato de que compreende um gene que codifica uma proteína operavelmente ligada a uma seqüência de nucleotídeo que codifica um peptídeo de sinal que
20 compreende ou que consiste dos aminoácidos de 1 a 18 da SEQ ID NO: 2 ou aminoácidos de 1 a 19 da SEQ ID NO: 4, em que o gene é estranho para a seqüência de nucleotídeo.

 35. Vetor de expressão recombinante, caracterizado pelo fato de que compreende a construção de ácido nucleico como definido na
25 reivindicação 34.

 36. Célula hospedeira recombinante, caracterizada pelo fato de que compreende a construção de ácido nucleico como definido na reivindicação 34.

 37. Método para produzir uma proteína, caracterizado pelo

fato de que compreende: (a) cultivar a célula hospedeira recombinante como definida na reivindicação 36 sob condições condutivas para a produção da proteína e (b) recuperar a proteína.

5 38. Método para produzir o polipeptídeo como definido em qualquer uma das reivindicações de 1 a 15, caracterizado pelo fato de que compreende: (a) cultivar uma planta transgênica ou uma célula de planta que compreendam um polinucleotídeo que codifica o polipeptídeo tendo atividade de endoglucanase sob condições condutivas para a produção do polipeptídeo; e (b) recuperar o polipeptídeo.

10 39. Planta transgênica, parte de planta ou célula de planta, caracterizadas pelo fato de que foram transformadas com um polinucleotídeo que codifica o polipeptídeo como definido em qualquer uma das reivindicações de 1 a 15.

15 40. Método para degradar ou converter um material celulósico, caracterizado pelo fato de que compreende: tratar o material celulósico com uma composição que compreenda uma quantidade eficaz de um polipeptídeo tendo a atividade de endoglucanase de acordo com qualquer uma das reivindicações de 1 a 15 ou a célula hospedeira como definida na reivindicação 20.

20 41. Método de acordo com a reivindicação 40, caracterizado pelo fato de que a composição compreende ainda uma quantidade eficaz de uma endo-1,4-beta-glucanase, exo-1,4-beta-D-glucanase e/ou beta-D-glucosidase.

25 42. Método de acordo com as reivindicações 40 ou 41, caracterizado pelo fato de que a composição compreende ainda uma quantidade eficaz de uma ou mais enzimas selecionadas do grupo que consiste de uma hemicelulase, esterase, protease, lacase e peroxidase.

43. Método de acordo com as reivindicações 40 a 42, caracterizado pelo fato de que o método é um processo de pré-tratamento;

uma etapa em um processo de sacarificação e fermentação simultâneo (SSF); ou uma etapa em um processo de hidrólise e fermentação híbrido (HHF).

5 44. Método de acordo com qualquer uma das reivindicações de 40 a 43, caracterizado pelo fato de que ainda compreende recuperar o material celulósico degradado ou convertido.

10 45. Método para produzir uma substância, caracterizado pelo fato de que compreende: (a) sacarificar um material celulósico com uma composição que compreenda uma quantidade eficaz de um polipeptídeo tendo atividade de endoglucanase como definido em qualquer uma das reivindicações de 1 a 15 ou a célula hospedeira como definida na reivindicação 20, (b) fermentar o material celulósico sacarificado da etapa (a) com um ou mais microorganismos de fermentação; e (c) recuperar a substância da fermentação.

15 46. Método de acordo com a reivindicação 45, caracterizado pelo fato de que a composição compreende ainda uma quantidade eficaz de endo-1,4-beta-glucanase, exo-1,4-beta-D-glucanase, e/ou beta-D-glucosidase.

20 47. Método de acordo com as reivindicações 45 ou 46, caracterizado pelo fato de que a composição compreende ainda uma quantidade eficaz de uma ou mais enzimas selecionadas do grupo que consiste de uma hemicelulase, esterase, protease, lacase e peroxidase.

48. Método de acordo com qualquer uma das reivindicações de 45 a 47, caracterizado pelo fato de que as etapas (a) e (b) são realizadas simultaneamente em uma sacarificação e fermentação simultâneas.

25 49. Método de acordo com qualquer uma das reivindicações de 45 a 48, caracterizado pelo fato de que a substância é um álcool, ácido orgânico, cetona, aminoácido ou gás.

50. Método para inibir a expressão de um polipeptídeo em uma célula, caracterizado pelo fato de que compreende administrar à célula ou expressar na célula uma molécula de RNA de filamento duplo (dsRNA), em

que o dsRNA compreende uma subsequência ou porção de um polinucleotídeo que codifica o polipeptídeo maduro da SEQ ID NO: 2 ou SEQ ID NO. 4.

51. Método de acordo com a reivindicação 50, caracterizado
5 pelo fato de que o dsRNA é de cerca de 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 ou mais nucleotídeos duplex no comprimento.

52. Molécula de RNA de filamento duplo (dsRNA) para inibir a expressão de um polipeptídeo em uma célula, caracterizada pelo fato de que a dsRNA compreende uma subsequência ou porção de um polinucleotídeo
10 que codifica o polipeptídeo maduro da SEQ ID NO: 2 ou SEQ ID NO: 4.

53. Molécula de RNA de filamento duplo (dsRNA) de acordo com a reivindicação 52, caracterizada pelo fato de que tem cerca de 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 ou mais nucleotídeos duplex no comprimento.

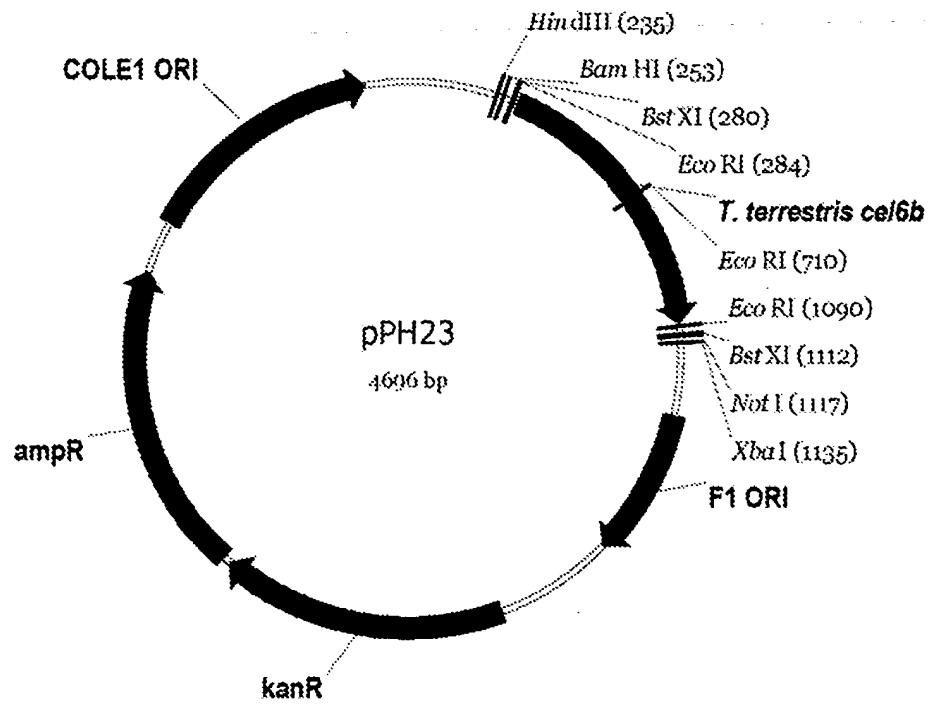


Fig. 1

Fig. 2A

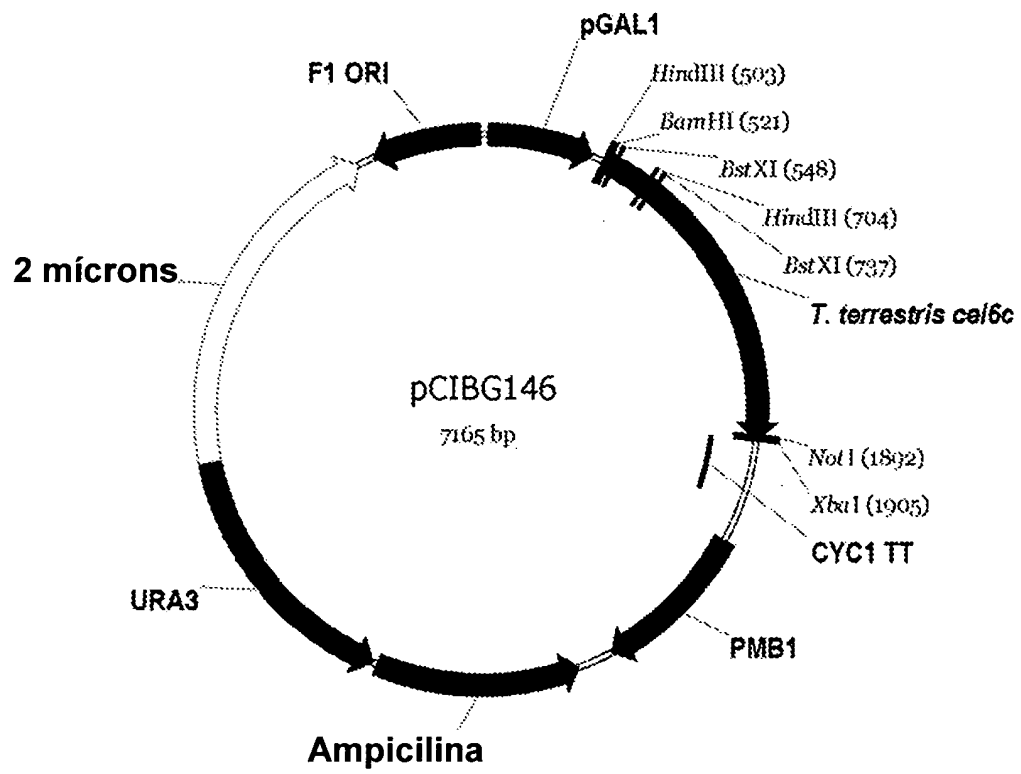


Fig. 3

V V K E G G E S D G T S D L N S N E Y D P T C R S P V

GGCGCATGTTCCCGCTCCCTGAGGCGCTGGCCAGTGGTTCAACGAGTATGTTGTTAACCTCGTTTTGAACGCTAACCCCCCT
A H V P A P E A G Q W F N E Y V V N L V L N A N P P
CTTGAGCCTACCTGGTAA 1203
L E P T W

Fig. 4B

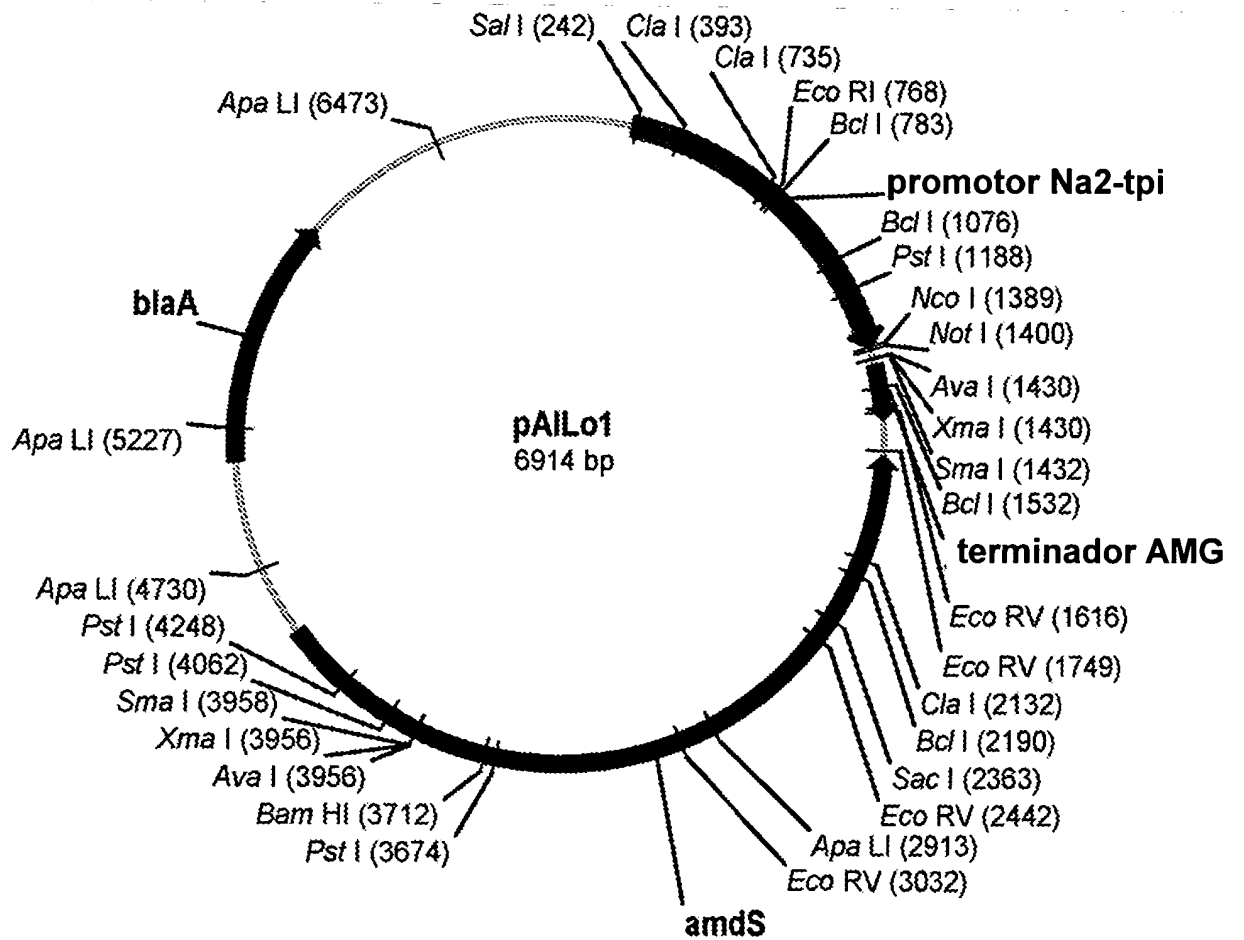


Fig. 5

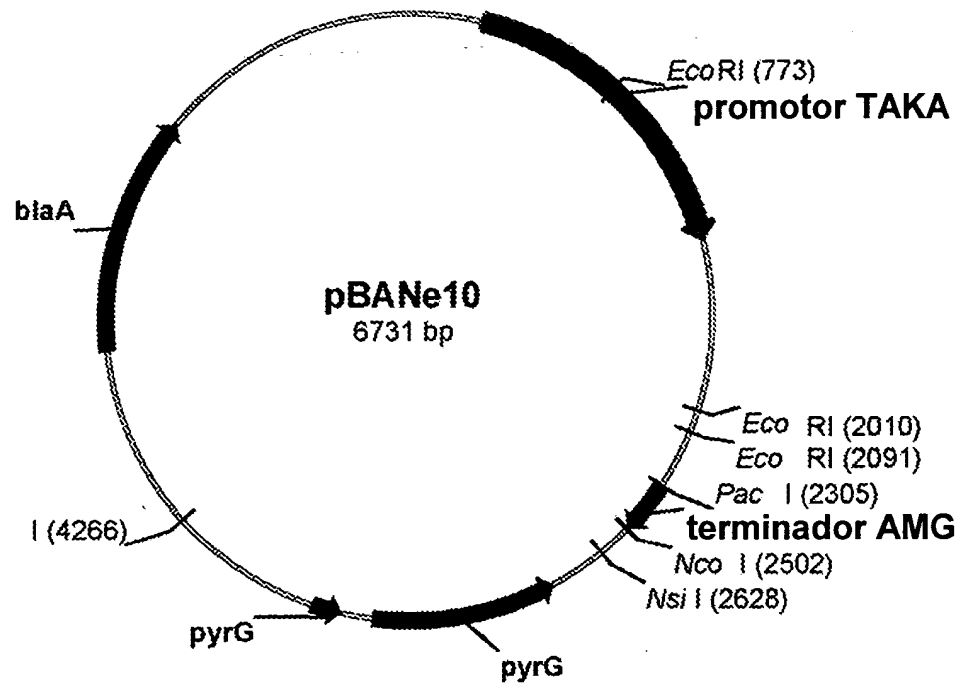


Fig. 6

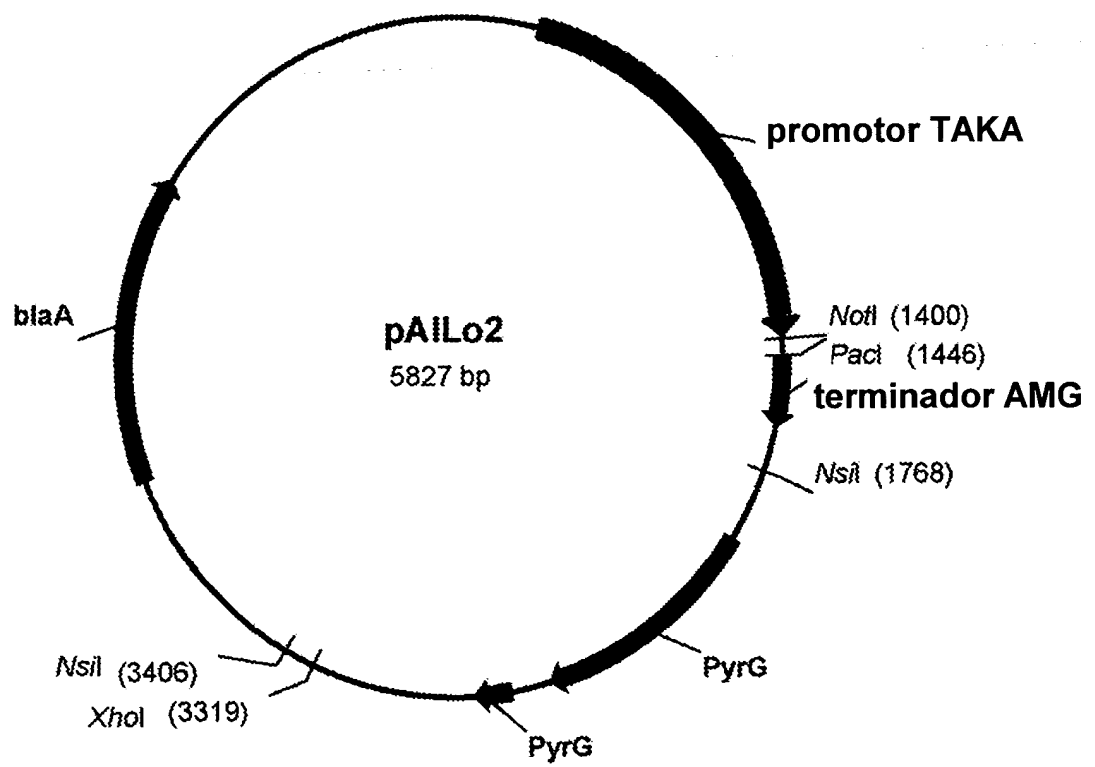


Fig. 7

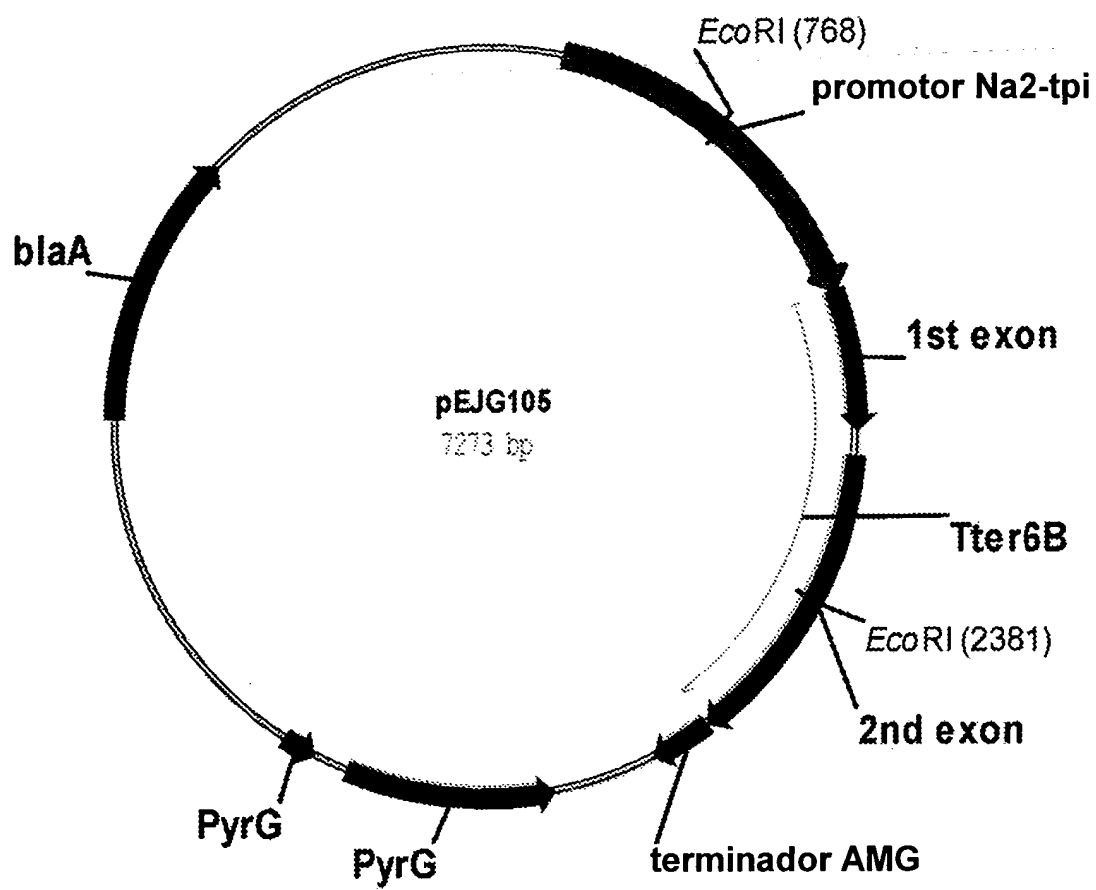


Fig. 8

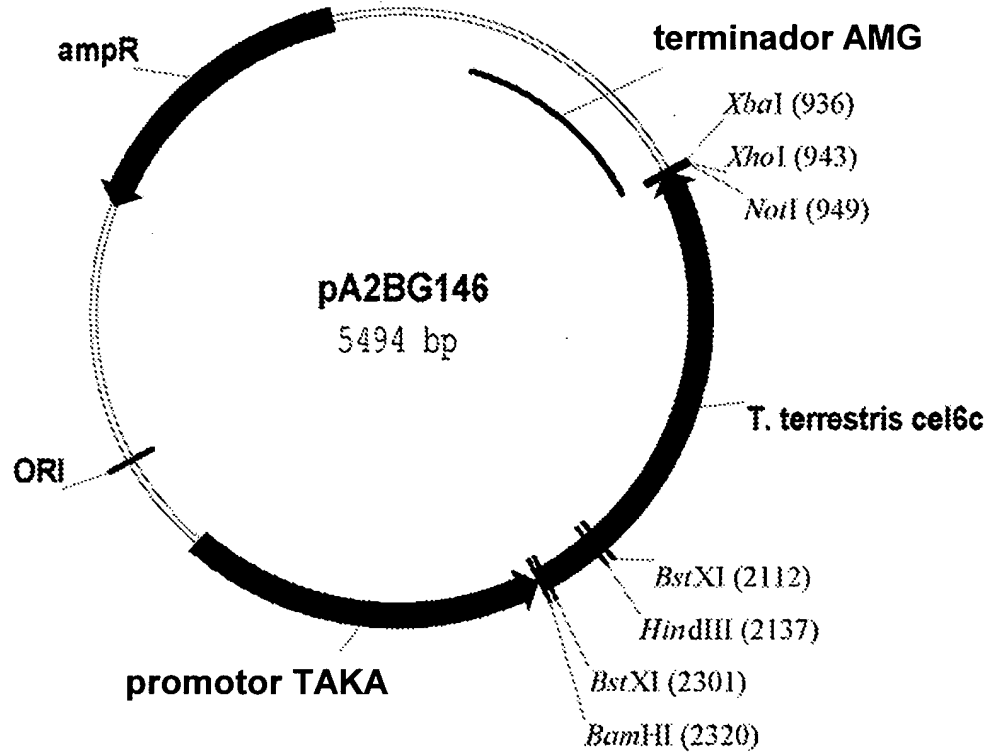


Fig. 9

RESUMO

“POLIPEPTÍDEO ISOLADO, POLINUCLEOTÍDEO, CONSTRUÇÃO DE
ÁCIDO NUCLEICO, VETOR DE EXPRESSÃO RECOMBINANTE,
CÉLULA HOSPEDEIRA RECOMBINANTE, MÉTODOS PARA
5 PRODUZIR O POLIPEPTÍDEO, PARA PRODUZIR UM MUTANTE DE
UMA CÉLULA PRECURSORA, PARA PRODUZIR UMA PROTEÍNA,
PARA PRODUZIR UM POLINUCLEOTÍDEO, PARA DEGRADAR OU
CONVERTER UM MATERIAL CELULÓSICO, PARA PRODUZIR UMA
SUBSTÂNCIA, E PARA INIBIR A EXPRESSÃO DE UM POLIPEPTÍDEO
10 EM UMA CÉLULA, CÉLULA MUTANTE, PLANTA TRANSGÊNICA,
PARTE DE PLANTA OU CÉLULA DE PLANTA, E, MOLÉCULA DE
RNA DE FILAMENTO DUPLO”

A presente invenção diz respeito a polipeptídeos isolados
tendo atividade de endoglucanase 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 métodos para produzir e usar os polipeptídeos.