



República Federativa do Brasil  
Ministério da Economia  
Instituto Nacional da Propriedade Industrial

(21) BR 112021001282-6 A2



(22) Data do Depósito: 22/07/2019

(43) Data da Publicação Nacional: 27/04/2021

(54) Título: LEVEDURA EXPRESSANDO ENZIMAS PARA PRODUÇÃO DE ETANOL

(51) Int. Cl.: C12P 7/06; C12P 7/10; C12N 9/24; C12N 9/28; C12N 9/30; (...).

(30) Prioridade Unionista: 25/07/2018 US 62/703,103.

(71) Depositante(es): NOVOZYMES A/S.

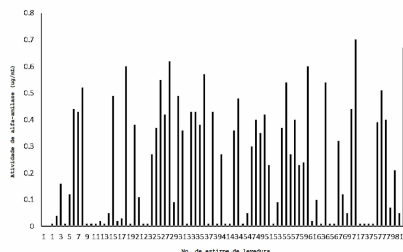
(72) Inventor(es): MICHAEL GLENN CATLETT; MONICA TASSONE; PAUL VINCENT HARRIS; ROBERT LYLE OSBORNE; SHIRO FUKUYAMA; TOMOKO MATSUI; RYOKO KATAOKA.

(86) Pedido PCT: PCT US2019042870 de 22/07/2019

(87) Publicação PCT: WO 2020/023411 de 30/01/2020

(85) Data da Fase Nacional: 22/01/2021

(57) Resumo: LEVEDURA EXPRESSANDO ENZIMAS PARA PRODUÇÃO DE ETANOL. São descritos aqui organismos fermentadores recombinantes tendo um polinucleotídeo heterólogo codificando uma alfa-amilase e/ou um polinucleotídeo heterólogo codificando uma trealase. São também descritos processos para produção de um produto de fermentação, tal como etanol, a partir de amido ou material contendo celulose com os organismos fermentadores recombinantes.



## **Relatório descritivo da patente de invenção para “LEVEDURA EXPRESSANDO ENZIMAS PARA PRODUÇÃO DE ETANOL”**

### **Referência a uma Listagem de Sequências**

**[0001]** Este pedido contém uma Listagem de Sequências em forma legível por computador, que é incorporada aqui por referência.

### **Antecedentes**

**[0002]** A produção de etanol a partir de materiais contendo amido e celulose é bem conhecida na técnica.

**[0003]** O processo comercial mais comumente usado na indústria para material contendo amido, frequentemente referido como um “processo convencional”, inclui liquefação de amido gelatinizado a elevada temperatura (cerca de 85 °C) usando tipicamente uma alfa-amilase bacteriana, seguida por sacarificação e fermentação simultâneas (SSF) levadas a cabo anaerobicamente na presença de tipicamente uma glucoamilase e uma levedura de *Saccharomyces cerevisiae*.

**[0004]** As leveduras que são usadas para produção de etanol para uso como combustível, tal como na indústria do etanol de milho, requerem várias características para assegurar produção rentável do etanol. Estas características incluem tolerância ao etanol, baixo rendimento de subprodutos, fermentação rápida e a capacidade de se limitar a quantidade de açúcares residuais permanecendo no fermento. Tais características têm um efeito pronunciado na viabilidade do processo industrial.

**[0005]** A levedura do gênero *Saccharomyces* exibe muitas das características requeridas para produção de etanol. Em particular, estirpes de *Saccharomyces cerevisiae* são amplamente usadas para a produção de etanol na indústria de combustíveis de etanol. As estirpes de *Saccharomyces cerevisiae* que são amplamente usadas na indústria de combustíveis de etanol têm a capacidade de produzir elevados rendimentos de etanol sob condições de fermentação encontradas, por exemplo, na

fermentação de mosto de milho. Um exemplo de uma tal estirpe é a levedura usada no produto de leveduras para etanol comercialmente disponível chamado ETHANOL RED®.

**[0006]** As leveduras de *Saccharomyces cerevisiae* foram geneticamente manipuladas para expressar alfa-amilase e/ou glucoamilase para melhorar o rendimento e diminuir a quantidade de enzimas exogenamente adicionadas necessárias durante SSF (p.ex., WO2018/098381, WO2017/087330, WO2017/037614, WO2011/128712, WO2011/153516, US2018/0155744). As leveduras foram também manipuladas para expressar trealase em uma tentativa de aumentar o rendimento da fermentação por desagregação da trealose residual (p.ex., WO2017/077504).

**[0007]** Apesar da melhoria significativa dos processos de produção de etanol ao longo da última década existe ainda um desejo e necessidade de se proporcionarem processos melhorados de fermentação de etanol a partir de material contendo amido e celulose em uma escala economicamente e comercialmente relevante.

### Sumário

**[0008]** São descritos aqui, *inter alia*, métodos para produção de um produto de fermentação, tal como etanol, a partir de material contendo amido ou celulose, e levedura adequada para uso em tais processos. O Requerente descobriu surpreendentemente que leveduras expressando certas alfa-amilases e/ou trealases proporcionam propriedades benéficas que podem ser úteis para a fermentação de etanol.

Um primeiro aspecto se relaciona com métodos de produção de um produto de fermentação a partir de um material contendo amido ou contendo celulose compreendendo: (a) sacarificação do material contendo amido ou contendo celulose; e (b) a fermentação do material sacarificado do passo (a) com um organismo fermentador; em que o organismo fermentador compreende um

polinucleotídeo heterólogo codificando uma alfa-amilase ou um polinucleotídeo heterólogo codificando uma trealase.

**[0009]** Em algumas modalidades dos métodos, a fermentação e a sacarificação são realizadas simultaneamente em uma sacarificação e fermentação simultâneas (SSF). Em outras modalidades, a fermentação e a sacarificação são realizadas sequencialmente (SHF).

**[0010]** Em algumas modalidades dos métodos, o método compreende recuperação do produto de fermentação a partir da fermentação (p.ex., por destilação).

**[0011]** Em algumas modalidades dos métodos, o produto de fermentação é etanol.

**[0012]** Em algumas modalidades dos métodos, a fermentação é realizada sob condições de nitrogênio reduzido (p.ex., menos do que 1000 ppm de ureia ou hidróxido de amônio, tal como menos do que 750 ppm, menos do que 500 ppm, menos do que 400 ppm, menos do que 300 ppm, menos do que 250 ppm, menos do que 200 ppm, menos do que 150 ppm, menos do que 100 ppm, menos do que 75 ppm, menos do que 50 ppm, menos do que 25 ppm ou menos do que 10 ppm).

**[0013]** Em algumas modalidades dos métodos, a alfa-amilase tem uma sequência de polipeptídeo maduro com 60%, p.ex., pelo menos 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, ou 100% de identidade de sequências com a sequência de aminoácidos de qualquer uma de SEQ ID NOs: 76-101, 121-174 e 231. Em algumas modalidades dos métodos, o polinucleotídeo heterólogo codifica uma alfa-amilase tendo uma sequência de polipeptídeo maduro que difere em não mais do que dez aminoácidos, p.ex., em não mais do que cinco aminoácidos, em não mais do que quatro aminoácidos, em não mais do que três aminoácidos, em não mais do que dois aminoácidos ou em um aminoácido da sequência de aminoácidos de qualquer uma de SEQ ID NOs: 76-101, 121-174 e 231. Em

algumas modalidades dos métodos, o polinucleotídeo heterólogo codifica uma alfa-amilase tendo uma sequência de polipeptídeo maduro compreendendo a ou consistindo na sequência de aminoácidos de qualquer uma de SEQ ID NOs: 76-101, 121-174 e 231.

**[0014]** Em algumas modalidades dos métodos, a trealase tem sequência de polipeptídeo maduro com 60%, p.ex., pelo menos 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, ou 100% de identidade de sequências com a sequência de aminoácidos de qualquer uma de SEQ ID NOs: 175-226. Em algumas modalidades dos métodos, o polinucleotídeo heterólogo codifica uma trealase tendo uma sequência de polipeptídeo maduro que difere em não mais do que dez aminoácidos, p.ex., em não mais do que cinco aminoácidos, em não mais do que quatro aminoácidos, em não mais do que três aminoácidos, em não mais do que dois aminoácidos ou em um aminoácido da sequência de aminoácidos de qualquer uma de SEQ ID NOs: 175-226. Em algumas modalidades dos métodos, o polinucleotídeo heterólogo codifica uma trealase tendo uma sequência de polipeptídeo maduro compreendendo a ou consistindo na sequência de aminoácidos de qualquer uma de SEQ ID NOs: 175-226.

**[0015]** Em algumas modalidades dos métodos, a sacarificação do passo (a) ocorre em um material contendo amido, e em que o material contendo amido é amido gelatinizado ou não gelatinizado.

**[0016]** Em algumas modalidades dos métodos, o método compreende liquefação do material contendo amido por contato do material com uma alfa-amilase antes da sacarificação.

**[0017]** Em algumas modalidades dos métodos, a liquefação do material contendo amido e/ou a sacarificação do material contendo amido são conduzidas na presença de protease exogenamente adicionada.

**[0018]** Em algumas modalidades dos métodos, o organismo fermentador compreende um polinucleotídeo heterólogo codificando uma glucoamilase, tal como uma glucoamilase tendo uma sequência de polipeptídeo maduro com 60%, p.ex., pelo menos 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, ou 100% de identidade de sequências com a sequência de aminoácidos de uma glicoamilase de *Pycnoporus* (p.ex., uma glucoamilase de *Pycnoporus sanguineus* de SEQ ID NO: 229), uma glucoamilase de *Gloeophyllum* (p.ex., uma *Gloeophyllum sepiarium* de SEQ ID NO: 8) ou uma glucoamilase de qualquer uma de SEQ ID NOs: 102-113 (p.ex., uma glucoamilase de *Saccharomycopsis fibuligera* de SEQ ID NO: 103 ou 104 ou uma glucoamilase de *Trichoderma reesei* de SEQ ID NO: 230).

**[0019]** Em algumas modalidades dos métodos, o organismo fermentador compreende um polinucleotídeo heterólogo codificando uma protease, tal como uma protease tendo uma sequência de polipeptídeo maduro de pelo menos 60%, p.ex., pelo menos 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, ou 100% de identidade de sequências com a sequência de aminoácidos de qualquer uma de SEQ ID NOs: 9-73 (p.ex., qualquer uma de SEQ ID NOs: 9, 14, 16, 21, 22, 33, 41, 45, 61, 62, 66, 67 e 69; tal como qualquer uma de SEQ NOs: 9, 14, 16 e 69).

Em algumas modalidades dos métodos, a sacarificação do passo (a) ocorre em um material contendo celulose, e em que o material contendo celulose é pré-tratado (p.ex. um pré-tratamento com ácido diluído).

Em algumas modalidades dos métodos, a sacarificação ocorre em um material contendo celulose, e em que a composição de enzimas compreende uma ou mais enzimas selecionadas de uma celulase (p.ex., endoglucanase, uma celobioidrolase ou uma beta-glucosidase), um polipeptídeo AA9, uma hemicelulase (p.ex., uma xilanase, uma acetilxilano esterase, uma feruloil esterase, uma arabinofuranosidase, uma xilosidase ou uma glucuronidase),

uma CIP, uma esterase, uma expansina, uma enzima ligninolítica, uma oxirredutase, uma pectinase, uma protease e uma swollenina.

**[0020]** Em algumas modalidades dos métodos, o organismo fermentador é uma célula de *Saccharomyces*, *Rhodotorula*, *Schizosaccharomyces*, *Kluyveromyces*, *Pichia*, *Hansenula*, *Rhodospiridium*, *Candida*, *Yarrowia*, *Lipomyces*, *Cryptococcus*, ou *Dekkera sp.* Em algumas modalidades, o organismo fermentador é uma célula de *Saccharomyces cerevisiae*.

**[0021]** Outro aspecto se relaciona com uma célula de levedura recombinante compreendendo um polinucleotídeo heterólogo codificando uma alfa-amilase e um polinucleotídeo heterólogo codificando uma trealase.

**[0022]** Em algumas modalidades, a célula de levedura recombinante é uma célula de *Saccharomyces*, *Rhodotorula*, *Schizosaccharomyces*, *Kluyveromyces*, *Pichia*, *Hansenula*, *Rhodospiridium*, *Candida*, *Yarrowia*, *Lipomyces*, *Cryptococcus*, ou *Dekkera sp.* Em algumas modalidades, a célula de levedura recombinante é uma célula de *Saccharomyces cerevisiae*.

**[0023]** Em algumas modalidades da célula de levedura, a alfa-amilase tem uma sequência de polipeptídeo maduro com 60%, p.ex., pelo menos 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, ou 100% de identidade de sequências com a sequência de aminoácidos de qualquer uma de SEQ ID NOs: 76-101, 121-174 e 231. Em algumas modalidades dos métodos, o polinucleotídeo heterólogo codifica uma alfa-amilase tendo uma sequência de polipeptídeo maduro que difere em não mais do que dez aminoácidos, p.ex., em não mais do que cinco aminoácidos, em não mais do que quatro aminoácidos, em não mais do que três aminoácidos, em não mais do que dois aminoácidos ou em um aminoácido da sequência de aminoácidos de qualquer uma de SEQ ID NOs: 76-101, 121-

174 e 231. Em algumas modalidades dos métodos, o polinucleotídeo heterólogo codifica uma alfa-amilase tendo uma sequência de polipeptídeo maduro compreendendo a ou consistindo na sequência de aminoácidos de qualquer uma de SEQ ID NOs: 76-101, 121-174 e 231.

**[0024]** Em algumas modalidades da célula de levedura, a trealase tem sequência de polipeptídeo maduro com 60%, p.ex., pelo menos 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, ou 100% de identidade de sequências com a sequência de aminoácidos de qualquer uma de SEQ ID NOs: 175-226. Em algumas modalidades dos métodos, o polinucleotídeo heterólogo codifica uma trealase tendo uma sequência de polipeptídeo maduro que difere em não mais do que dez aminoácidos, p.ex., em não mais do que cinco aminoácidos, em não mais do que quatro aminoácidos, em não mais do que três aminoácidos, em não mais do que dois aminoácidos ou em um aminoácido da sequência de aminoácidos de qualquer uma de SEQ ID NOs: 175-226. Em algumas modalidades dos métodos, o polinucleotídeo heterólogo codifica uma trealase tendo uma sequência de polipeptídeo maduro compreendendo a ou consistindo na sequência de aminoácidos de qualquer uma de SEQ ID NOs: 175-226.

**[0025]** Em algumas modalidades da célula de levedura, o organismo fermentador compreende um polinucleotídeo heterólogo codificando uma glucoamilase, tal como uma glucoamilase tendo uma sequência de polipeptídeo maduro com 60%, p.ex., pelo menos 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, ou 100% de identidade de sequências com a sequência de aminoácidos de uma glicoamilase de *Pycnoporus* (p.ex., uma glucoamilase de *Pycnoporus sanguineus* de SEQ ID NO: 229), uma glucoamilase de *Gloeophyllum* (p.ex., uma *Gloeophyllum sepiarium* de SEQ ID NO: 8) ou uma glucoamilase de qualquer uma de SEQ ID NOs: 102-113 (p.ex., uma glucoamilase de *Saccharomycopsis fibuligera*



de SEQ ID NO: 103 ou 104 ou uma glucoamilase de *Trichoderma reesei* de SEQ ID NO: 230).

**[0026]** Em algumas modalidades da célula de levedura, o organismo fermentador compreende um polinucleotídeo heterólogo codificando uma protease, tal como uma protease tendo uma sequência de polipeptídeo maduro de pelo menos 60%, p.ex., pelo menos 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, ou 100% de identidade de sequências com a sequência de aminoácidos de qualquer uma de SEQ ID NOs: 9-73 (p.ex., qualquer uma de SEQ ID NOs: 9, 14, 16, 21, 22, 33, 41, 45, 61, 62, 66, 67 e 69; tal como qualquer uma de SEQ NOs: 9, 14, 16 e 69).

#### **Breve Descrição das Figuras**

**[0027]** A Figura 1 mostra a atividade de alfa-amilase para estirpes construídas no Exemplo 1.

**[0028]** A Figura 2 mostra a atividade de trealase para estirpes construídas no Exemplo 4.

**[0029]** A Figura 3 mostra um mapa de plasmídeo para pMcTs442.

#### **Definições**

**[0030]** A não ser que definidos de outro modo ou claramente indicado pelo contexto, todos os termos técnicos e científicos usados aqui têm o mesmo significado como comumente entendido por um perito na técnica.

**[0031]** **Variante alélica:** O termo “variante alélica” significa qualquer uma de duas ou mais formas alternativas de um gene ocupando o mesmo locus cromossômico. A variação alélica surge naturalmente através de mutação e pode resultar em polimorfismo dentro de populações. As mutações gênicas podem ser silenciosas (sem mudança no polipeptídeo codificado) ou podem codificar polipeptídeos tendo sequências de

aminoácidos alteradas. Uma variante alélica de um polipeptídeo é um polipeptídeo codificado por uma variante alélica de um gene.

**[0032] Alfa-amilase:** O termo “alfa-amilase” significa uma 1,4-alfa-D-glucana glucanoidrolase, EC. 3.2.1.1, que catalisa a hidrólise de amido e outros oligo- e polissacarídeos 1,4-glucosídicos lineares e ramificados. Para propósitos da presente invenção, a atividade de alfa-amilase pode ser determinada usando um ensaio de alfa-amilase descrito na seção dos exemplos em baixo.

**[0033] Atividade Auxiliar 9:** O termo “Atividade Auxiliar 9” ou “AA9” significa um polipeptídeo classificado como uma polissacarídeo lítico mono-oxigenase (Quinlan *et al.*, 2011, *Proc. Natl. Acad. Sci. USA* 208: 15079-15084; Phillips *et al.*, 2011, *ACS Chem. Biol.* 6: 1399-1406; Lin *et al.*, 2012, *Structure* 20: 1051-1061). Os polipeptídeos AA9 era antigamente classificados na Família 61 de glicosídeo hidrolases (GH61) de acordo com Henrissat, 1991, *Biochem. J.* 280: 309-316 e Henrissat e Bairoch, 1996, *Biochem. J.* 316: 695-696.

**[0034]** Os polipeptídeos AA9 intensificam a hidrólise de um material contendo celulose por uma enzima tendo atividade celulolítica. A atividade intensificante celulolítica pode ser determinada por medição do aumento nos açúcares redutores ou do aumento do total de celobiose e glucose a partir da hidrólise de um material contendo celulose por uma enzima celulolítica sob as seguintes condições: 1-50 mg de proteína total/g de celulose em palha de milho pré-tratada (PCS), em que a proteína total é compreendida por proteína de enzima celulolítica a 50-99,5% p/p e proteína de um polipeptídeo AA9 a 0,5-50% p/p durante 1-7 dias a uma temperatura adequada, tal como 40° C-80 °C, p.ex., 50 °C, 55 °C, 60 °C, 65 °C ou 70 °C, e um pH adequado, tal como 4-9, p.ex., 4,5, 5,0, 5,5, 6,0, 6,5, 7,0, 7,5, 8,0 ou 8,5, em comparação com uma hidrólise de controle com igual carga de

proteína total sem atividade intensificante celulolítica (1-50 mg de proteína celulolítica/g de celulose em PCS).

**[0035]** A atividade intensificante do polipeptídeo AA9 pode ser determinada usando uma mistura de CELLUCLAST® 1.5L (Novozymes A/S, Bagsværd, Dinamarca) e beta-glucosidase como a fonte da atividade celulolítica, em que a beta-glucosidase está presente em um peso de pelo menos 2-5% de proteína da carga de proteína de celulase. Em uma modalidade, a beta-glucosidase é uma beta-glucosidase de *Aspergillus oryzae* (p.ex., recombinantemente produzida em *Aspergillus oryzae* de acordo com WO02/095014). Em outra modalidade, a beta-glucosidase é uma beta-glucosidase de *Aspergillus fumigatus* (p.ex., recombinantemente produzida em *Aspergillus oryzae* como descrito em WO02/095014).

**[0036]** A atividade intensificante do polipeptídeo AA9 pode ser também determinada por incubação de um polipeptídeo AA9 com celulose expandida com ácido fosfórico (PASC) a 0,5%, acetato de sódio a 100 mM a pH 5, MnSO<sub>4</sub> a 1 mM, ácido gálico a 0,1%, 0,025 mg/mL de beta-glucosidase de *Aspergillus fumigatus* e TRITON® X-100 (4-(1,1,3,3-tetrametilbutil)fenil-poli-etilenoglicol) a 0,01% durante 24-96 horas a 40 °C seguida por determinação da glucose liberada a partir da PASC.

**[0037]** A atividade intensificante do polipeptídeo AA9 pode ser também determinada de acordo com WO2013/028928 para composições a elevada temperatura.

**[0038]** Os polipeptídeos AA9 intensificam a hidrólise de um material contendo celulose catalisada por enzima tendo atividade celulolítica por redução da quantidade de enzima celulolítica requerida para se alcançar o mesmo grau de hidrólise, preferencialmente pelo menos 1,01 vezes, p.ex., pelo menos 1,05 vezes, pelo menos 1,10 vezes, pelo menos 1,25 vezes, pelo menos 1,5 vezes, pelo menos 2 vezes, pelo menos 3 vezes, pelo menos 4 vezes, pelo menos 5 vezes, pelo menos 10 vezes ou pelo menos 20 vezes.

**[0039] Beta-glucosidase:** O termo “beta-glucosidase” significa uma beta-D-glucosídeo glucoidrolase (E.C. 3.2.1.21) que catalisa a hidrólise de resíduos de beta-D-glucose não redutores terminais com a liberação de beta-D-glucose. A atividade de beta-glucosidase pode ser determinada usando *p*-nitrofenil-beta-D-glucopiranosídeo como substrato de acordo com o procedimento de Venturi *et al.*, 2002, *J. Basic Microbiol.* 42: 55-66. Uma unidade de beta-glucosidase é definida como 1,0 μmole de ânion *p*-nitrofenolato produzida por minuto a 25 °C, pH 4,8 a partir de *p*-nitrofenil-beta-D-glucopiranosídeo a 1 mM como substrato em citrato de sódio a 50 mM contendo TWEEN® 20 a 0,01%.

**[0040] Beta-xilosidase:** O termo “beta-xilosidase” significa uma beta-D-xilosídeo xiloidrolase (E.C. 3.2.1.37) que catalisa a exo-hidrólise de beta (1→4)-xilo-oligossacarídeos curtos para se removerem resíduos de D-xilose sucessivos a partir de terminais não redutores. A atividade de beta-xilosidase pode ser determinada usando *p*-nitrofenil-beta-D-xilosídeo a 1 mM como substrato em citrato de sódio a 100 mM contendo TWEEN® 20 a 0,01%, a pH 5, 40 °C. Uma unidade de beta-xilosidase é definida como 1,0 μmol de ânion *p*-nitrofenolato produzido por minuto a 40 °C, pH 5 a partir de *p*-nitrofenil-beta-D-xilosídeo a 1 mM em citrato de sódio a 100 mM contendo TWEEN® 20 a 0,01%.

**[0041] Catalase:** O termo “catalase” significa uma peróxido de hidrogênio:peróxido de hidrogênio oxidorreductase (EC 1.11.1.6) que catalisa a conversão de 2 H<sub>2</sub>O<sub>2</sub> em O<sub>2</sub> + 2 H<sub>2</sub>O. Para propósitos da presente invenção, a atividade de catalase é determinada de acordo com a Patente dos E.U.A. No. 5,646,025. Uma unidade de atividade de catalase iguala a quantidade de enzima que catalisa a oxidação de 1 μmole de peróxido de hidrogênio sob as condições de ensaio.

**[0042] Domínio catalítico:** O termo “domínio catalítico” significa a região de uma enzima contendo a maquinaria catalítica da enzima.

**[0043] Celobioidrolase:** O termo “celobioidrolase” significa uma 1,4-beta-D-glucana celobioidrolase (E.C. 3.2.1.91 e E.C. 3.2.1.176) que catalisa a hidrólise de ligações 1,4-beta-D-glicosídicas em celulose, celo-oligossacarídeos ou qualquer polímero contendo glucose ligada em beta-1,4, liberando celobiose a partir da extremidade redutora (celobioidrolase I) ou extremidade não redutora (celobioidrolase II) da cadeia (Teeri, 1997, *Trends in Biotechnology* 15: 160-167; Teeri *et al.*, 1998, *Biochem. Soc. Trans.* 26: 173-178). A atividade de celobioidrolase pode ser determinada de acordo com os procedimentos descritos por Lever *et al.*, 1972, *Anal. Biochem.* 47: 273-279; van Tilbeurgh *et al.*, 1982, *FEBS Letters* 149: 152-156; van Tilbeurgh e Claeysens, 1985, *FEBS Letters* 187: 283-288; e Tomme *et al.*, 1988, *Eur. J. Biochem.* 170: 575-581.

**[0044] Enzima celulolítica ou celulase:** O termo “enzima celulolítica” ou “celulase” significa uma ou mais (p.ex., várias) enzimas que hidrolisam um material contendo celulose. Tais enzimas incluem endoglucanase(s), celobioidrolase(s), beta-glucosidase(s) ou suas combinações. As duas abordagens básicas para medição da atividade de enzima celulolítica incluem: (1) medição da atividade de enzima celulolítica total e (2) medição das atividades de enzimas celulolíticas individuais (endoglucanases, celobioidrolases e beta-glucosidases) como revisto em Zhang *et al.*, 2006, *Biotechnology Advances* 24: 452-481. A atividade de enzima celulolítica total pode ser medida usando substratos insolúveis, incluindo papel de filtro Whatman Nº1, celulose microcristalina, celulose bacteriana, celulose de algas, algodão, lignocelulose pré-tratada, *etc.* O ensaio mais comum da atividade celulolítica total é o ensaio de papel de filtro usando papel de filtro Whatman Nº1 como o substrato. O ensaio foi estabelecido pela International Union of Pure and Applied Chemistry (IUPAC) (Ghose, 1987, *Pure Appl. Chem.* 59: 257-68).

**[0045]** A atividade de enzima celulolítica pode ser determinada por medição do aumento na produção/liberação de açúcares durante a hidrólise de um material contendo celulose por enzima(s) celulolítica(s) sob as seguintes condições: 1-50 mg de proteína de enzima celulolítica/g de celulose em palha de milho pré-tratada (PCS) (ou outro material contendo celulose pré-tratado) durante 3-7 dias a uma temperatura adequada tal como 40 °C-80 °C, p.ex., 50 °C, 55 °C, 60 °C, 65 °C ou 70 °C, e um pH adequado tal como 4-9, p.ex., 5,0, 5,5, 6,0, 6,5 ou 7,0, em comparação com uma hidrólise de controle sem adição de proteína de enzima celulolítica. Condições típicas são reações de 1 mL, PCS lavada ou não lavada, sólidos insolúveis a 5% (peso seco), acetato de sódio a 50 mM, pH 5, MnSO<sub>4</sub> a 1 mM, 50 °C, 55 °C ou 60 °C, 72 horas, análise de açúcares por cromatografia em coluna AMINEX<sup>®</sup> HPX-87H (Bio-Rad Laboratories, Inc., Hercules, CA, EUA).

**[0046] Sequência codificante:** O termo “sequência codificante” ou “região codificante” significa uma sequência de polinucleotídeos que especifica a sequência de aminoácidos de um polipeptídeo. As fronteiras da sequência codificante são geralmente determinadas por uma grelha de leitura aberta, que começa usualmente com o códon de início ATG ou códons de início alternativos tais como GTG e TTG e termina com um códon de terminação tal como TAA, TAG e TGA. A sequência codificante pode ser uma sequência de DNA genômico, cDNA, um polinucleotídeo sintético e/ou um polinucleotídeo recombinante.

**[0047] Sequência de controle:** O termo “sequência de controle” significa uma sequência de ácidos nucleicos necessária para expressão do polipeptídeo. As sequências de controle podem ser nativas ou estranhas ao polinucleotídeo codificando o polipeptídeo e nativas ou estranhas entre si. Tais sequências de controle incluem, mas não estão limitadas a, uma sequência líder, sequência de poliadenilação, sequência de

pró-peptídeo, sequência promotora, sequência de peptídeo de sinal e sequência terminadora da transcrição. As sequências de controle podem ser proporcionadas com ligantes para o propósito de se introduzirem locais de restrição específicos facilitando a ligação das sequências de controle com a região codificante do polinucleotídeo codificando um polipeptídeo.

**[0048] Ruptura:** O termo “ruptura” significa que uma região codificante e/ou sequência de controle de um gene referenciado é parcialmente ou inteiramente modificada (tal como por deleção, inserção e/ou substituição de um ou mais nucleotídeos) resultando na ausência (inativação) ou diminuição na expressão e/ou na ausência ou diminuição da atividade de enzima do polipeptídeo codificado. Os efeitos da ruptura podem ser medidos usando técnicas conhecidas na técnica tais como detecção da ausência ou diminuição da atividade de enzima usando medições de extratos isentos de células referenciadas aqui; ou pela ausência ou diminuição do mRNA correspondente (p.ex., pelo menos 25% de diminuição, pelo menos 50% de diminuição, pelo menos 60% de diminuição, pelo menos 70% de diminuição, pelo menos 80% de diminuição ou pelo menos 90% de diminuição); a ausência ou diminuição na quantidade do polipeptídeo tendo atividade de enzima correspondente (p.ex., pelo menos 25% de diminuição, pelo menos 50% de diminuição, pelo menos 60% de diminuição, pelo menos 70% de diminuição, pelo menos 80% de diminuição ou pelo menos 90% de diminuição); ou a ausência ou diminuição da atividade específica do polipeptídeo tendo atividade de enzima correspondente (p.ex., pelo menos 25% de diminuição, pelo menos 50% de diminuição, pelo menos 60% de diminuição, pelo menos 70% de diminuição, pelo menos 80% de diminuição ou pelo menos 90% de diminuição). As rupturas de um gene particular de interesse podem ser geradas por métodos conhecidos na técnica, p.ex., por recombinação homóloga dirigida (ver *Methods in Yeast Genetics* (edição de

1997), Adams, Gottschling, Kaiser e Stems, Cold Spring Harbor Press (1998)).

**[0049] Gene endógeno:** O termo “gene endógeno” significa um gene que é nativo da célula hospedeira referenciada. “Expressão de gene endógeno” significa expressão de um gene endógeno.

**[0050] Endoglucanase:** O termo “endoglucanase” significa uma 4-(1,3;1,4)-beta-D-glucana 4-glucanoidrolase (E.C. 3.2.1.4) que catalisa a endo-hidrólise de ligações 1,4-beta-D-glicosídicas na celulose, derivados da celulose (tais como celulose de carboximetila e celulose de hidroxietila), liquenina, ligações beta-1,4 em beta-1,3-1,4 glucanas mistas tais como beta-D-glucanas ou xiloglucanas de cereais e outro material vegetal contendo componentes celulósicos. A atividade de endoglucanase pode ser determinada por medição da redução na viscosidade do substrato ou aumento nas extremidades reductoras determinado por um teste de açúcares redutores (Zhang *et al.*, 2006, *Biotechnology Advances* 24: 452-481). A atividade de endoglucanase pode ser também determinada usando celulose de carboximetila (CMC) como substrato de acordo com o procedimento de Ghose, 1987, *Pure and Appl. Chem.* 59: 257-268, a pH 5, 40 °C.

**[0051] Expressão:** O termo “expressão” inclui qualquer passo envolvido na produção do polipeptídeo incluindo, mas não se limitando a, transcrição, modificação pós-transcricional, tradução, modificação pós-translacional e secreção. A expressão pode ser medida — por exemplo, para se detectar expressão aumentada — por técnicas conhecidas na técnica, tais como medição dos níveis de mRNA e/ou polipeptídeo traduzido.

**[0052] Vetor de expressão:** O termo “vetor de expressão” significa uma molécula de DNA linear ou circular que compreende um polinucleotídeo codificando um polipeptídeo e está operacionalmente ligado a sequências de controle que proporcionam sua expressão.



**[0053] Meio fermentável:** O termo “meio fermentável” ou “meio de fermentação” se refere a um meio compreendendo um ou mais (p.ex., dois, vários) açúcares, tais como glucose, frutose, sacarose, celobiose, xilose, xilulose, arabinose, manose, galactose e/ou oligossacarídeos solúveis, em que o meio é capaz, em parte, de ser convertido (fermentado) por uma célula hospedeira em um produto desejado, tal como etanol. Em alguns casos, o meio de fermentação é derivado de uma fonte natural, tal como cana-de-açúcar, amido ou celulose, e pode ser o resultado do pré-tratamento da fonte por hidrólise enzimática (sacarificação). O termo meio de fermentação é entendido aqui como se referindo a um meio antes de o organismo fermentador ser adicionado, tal como um meio resultando de um processo de sacarificação, bem como um meio usado em um processo de sacarificação e fermentação simultâneas (SSF).

**[0054] Glucoamilase:** O termo “glucoamilase” (1,4-alfa-D-glucana glucohidrolase, EC 3.2.1.3) é definido como uma enzima que catalisa a liberação de D-glucose a partir das extremidades não redutoras de amido ou moléculas de oligo- e polissacarídeos relacionadas. Para propósitos da presente invenção, a atividade de glucoamilase pode ser determinada de acordo com o procedimento descrito nos Exemplos aqui.

**[0055] Enzima hemicelulolítica ou hemicelulase:** O termo “enzima hemicelulolítica” ou “hemicelulase” significa uma ou mais (p.ex., várias) enzimas que hidrolisam um material hemicelulósico. Ver, por exemplo, Shallom e Shoham, 2003, *Current Opinion In Microbiology* 6 (3): 219-228. As hemicelulases são componentes-chave na degradação de biomassa vegetal. Exemplos de hemicelulases incluem, mas não estão limitados a, uma acetilmanana esterase, uma acetilxilana esterase, uma arabinanase, uma arabinofuranosidase, uma ácido cumárico esterase, uma feruloil esterase, uma galactosidase, uma glucuronidase, uma glucuronoil esterase, uma mananase, uma manosidase, uma xilanase e uma xilosidase.

Os substratos para estas enzimas, hemiceluloses, são um grupo heterogêneo de polissacarídeos ramificados e lineares que estão ligados através de ligações de hidrogênio às microfibrilas de celulose na parede celular vegetal, reticulando as mesmas em uma rede robusta. As hemiceluloses estão também covalentemente anexadas à lignina, formando em conjunto com a celulose uma estrutura altamente complexa. A estrutura variável e a organização das hemiceluloses requerem a ação concertada de muitas enzimas para sua degradação completa. Os módulos catalíticos de hemicelulases são glicosídeo hidrolases (GH) que hidrolisam ligações glicosídicas ou carboidrato esterases (CE), que hidrolisam ligações de éster de grupos laterais de acetato ou ácido ferúlico. Estes módulos catalíticos, com base na homologia de sua sequência primária, podem ser atribuídos às famílias GH e CE. Algumas famílias, com um dobramento similar global, podem ser adicionalmente agrupadas em clãs, marcados alfabeticamente (p.ex., GH-A). Uma classificação mais informativa e atualizada destas e outras enzimas ativas em carboidratos está disponível na base de dados Carbohydrate-Active Enzymes (CAZy). As atividades de enzima hemicelulolítica podem ser medidas de acordo com Ghose e Bisaria, 1987, *Pure & Appl. Chem.* 59: 1739-1752, a uma temperatura adequada tal como 40 °C-80 °C, p.ex., 50 °C, 55 °C, 60 °C, 65 °C ou 70 °C, e um pH adequado, tal como 4-9, p.ex., 5,0, 5,5, 6,0, 6,5 ou 7,0.

**[0056] Polinucleotídeo heterólogo:** O termo “polinucleotídeo heterólogo” é definido aqui como um polinucleotídeo que não é nativo da célula hospedeira; um polinucleotídeo nativo no qual modificações estruturais foram feitas na região codificante; um polinucleotídeo nativo cuja expressão está quantitativamente alterada em resultado de uma manipulação do DNA por técnicas de DNA recombinantes, p.ex., um promotor diferente (estranho); ou um polinucleotídeo nativo em uma célula hospedeira tendo uma ou mais cópias extra do polinucleotídeo

para alterar quantitativamente a expressão. Um “gene heterólogo” é um gene compreendendo um polinucleotídeo heterólogo.

**[0057] Condições de elevada estringência:** O termo “condições de elevada estringência” significa, para sondas com pelo menos 100 nucleotídeos em comprimento, pré-hibridação e hibridação a 42 °C em 5X SSPE, SDS a 0,3%, DNA de esperma de salmão cisalhado e desnaturado a 200 microgramas/mL e formamida a 50%, seguindo procedimentos de transferência de Southern padrão durante 12 a 24 horas. O material transportador é finalmente lavado três vezes cada uma durante 15 minutos usando 0,2X SSC, SDS a 0,2% a 65 °C.

**[0058] Célula hospedeira:** O termo “célula hospedeira” significa qualquer tipo de célula que é suscetível à transformação, transfecção, transdução e similares com um construto de ácido nucleico ou vetor de expressão compreendendo um polinucleotídeo descrito aqui (p.ex., um polinucleotídeo codificando uma alfa-amilase e/ou trealase). O termo “célula hospedeira” engloba qualquer descendência de uma célula parental que não é idêntica à célula parental devido a mutações que ocorrem durante a replicação. O termo “célula recombinante” é definido aqui como uma célula hospedeira não ocorrendo naturalmente compreendendo um ou mais (p.ex., dois, vários) polinucleotídeos heterólogos.

**[0059] Condições de baixa estringência:** O termo “condições de baixa estringência” significa, para sondas com pelo menos 100 nucleotídeos em comprimento, pré-hibridação e hibridação a 42 °C em 5X SSPE, SDS a 0,3%, DNA de esperma de salmão cisalhado e desnaturado a 200 microgramas/mL e formamida a 25%, seguindo procedimentos de transferência de Southern padrão durante 12 a 24 horas. O material transportador é finalmente lavado três vezes cada uma durante 15 minutos usando 0,2X SSC, SDS a 0,2% a 50 °C.

**[0060] Polipeptídeo maduro:** O termo “polipeptídeo maduro” é definido aqui como um polipeptídeo tendo atividade biológica que está em sua forma final após tradução e quaisquer modificações pós-translacionais, tais como processamento N-terminal, truncação C-terminal, glicosilação, fosforilação, etc. A sequência de polipeptídeo maduro não tem uma sequência de sinal, que pode ser determinada usando técnicas conhecidas na técnica (Ver, p.ex., Zhang e Henzel, 2004, *Protein Science* 13: 2819-2824).

**[0061] Condições de média estringência:** O termo “condições de média estringência” significa, para sondas com pelo menos 100 nucleotídeos em comprimento, pré-hibridação e hibridação a 42 °C em 5X SSPE, SDS a 0,3%, DNA de esperma de salmão cisalhado e desnaturado a 200 microgramas/mL e formamida a 35%, seguindo procedimentos de transferência de Southern padrão durante 12 a 24 horas. O material transportador é finalmente lavado três vezes cada uma durante 15 minutos usando 0,2X SSC, SDS a 0,2% a 55 °C.

**[0062] Condições de média-elevada estringência:** O termo “condições de média-elevada estringência” significa, para sondas com pelo menos 100 nucleotídeos em comprimento, pré-hibridação e hibridação a 42 °C em 5X SSPE, SDS a 0,3%, DNA de esperma de salmão cisalhado e desnaturado a 200 microgramas/mL e formamida a 35%, seguindo procedimentos de transferência de Southern padrão durante 12 a 24 horas. O material transportador é finalmente lavado três vezes cada uma durante 15 minutos usando 0,2X SSC, SDS a 0,2% a 60 °C.

**[0063] Construto de ácido nucleico:** O termo “construto de ácido nucleico” significa um polinucleotídeo compreendendo uma ou mais (p.ex., duas, várias) sequências de controle. O polinucleotídeo pode ter fita única ou fita dupla e pode ser isolado a partir de um gene ocorrendo

naturalmente, modificado para conter segmentos de ácidos nucleicos de uma maneira que de outro modo não existiria na natureza ou sintético.

**[0064] Operacionalmente ligado:** O termo “operacionalmente ligado” significa uma configuração na qual uma sequência de controle é colocada em uma posição apropriada em relação à sequência codificante de um polinucleotídeo tal que a sequência de controle dirija a expressão da sequência codificante.

**[0065] Palha de milho pré-tratada:** O termo “Palha de Milho Pré-tratada” ou “PCS” significa um material contendo celulose derivado de palha de milho por tratamento com calor e ácido sulfúrico diluído, pré-tratamento alcalino, pré-tratamento neutro ou qualquer pré-tratamento conhecido na técnica.

**[0066] Protease:** O termo “protease” é definido aqui como uma enzima que hidrolisa ligações de peptídeo. O termo inclui qualquer enzima pertencendo ao grupo de enzimas EC 3.4 (incluindo cada uma das suas treze subclasses). O número EC se refere à Nomenclatura de Enzimas de 1992 de NC-IUBMB, Academic Press, San Diego, Califórnia, incluindo os suplementos 1-5 publicados em *Eur. J. Biochem.* 223: 1-5 (1994); *Eur. J. Biochem.* 232: 1-6 (1995); *Eur. J. Biochem.* 237: 1-5 (1996); *Eur. J. Biochem.* 250: 1-6 (1997); e *Eur. J. Biochem.* 264: 610-650 (1999); respectivamente. O termo “subtilases” se refere a um subgrupo de serina proteases de acordo com Siezen *et al.*, 1991, *Protein Engng.* 4: 719-737 e Siezen *et al.*, 1997, *Protein Science* 6: 501-523. As serina proteases ou serina peptidases são um subgrupo de proteases caracterizadas por terem uma serina no local ativo, que forma um aduto covalente com o substrato. Adicionalmente, as subtilases (e as serina proteases) são caracterizadas por terem dois resíduos de aminoácidos de local ativo separados da serina, nomeadamente um resíduo de histidina e ácido aspártico. As subtilases podem ser divididas em 6 subdivisões, *i.e.*, a família das Subtilisinas, a família das Termitases, a

família das Proteinases K, a família das Peptidases lantibióticas, a família das Quexinas e a família das Pirolisinas. O termo “atividade de protease” significa uma atividade proteolítica (EC 3.4). A atividade de protease pode ser determinada usando métodos descritos na técnica (p.ex.0, US 2015/0125925) ou usando estojos de ensaio comercialmente disponíveis (p.ex., Sigma-Aldrich).

**[0067] Pululanase:** O termo “pululanase” significa uma enzima desramificadora de amido tendo atividade de pululana 6-glucoanidrolase (EC 3.2.1.41) que catalisa a hidrólise das ligações  $\alpha$ -1,6-glicosídicas na pululana, liberando maltotriose com extremidades de carboidrato redutoras. Para propósitos da presente invenção, a atividade de pululanase pode ser determinada de acordo com um ensaio PHADEBAS ou o ensaio de amido de batata-doce descrito em WO2016/087237.

**[0068] Identidade de Sequências:** A relação entre duas sequências de aminoácidos ou entre duas sequências de nucleotídeos é descrita pelo parâmetro “identidade de sequências”.

**[0069]** Para propósitos descritos aqui, o grau de identidade de sequências entre duas sequências de aminoácidos é determinado usando o algoritmo de Needleman-Wunsch (Needleman e Wunsch, *J. Mol. Biol.* **1970**, *48*, 443-453) como implantado no programa Needle do pacote EMBOSS (EMBOSS: The European Molecular Biology Open Software Suite, Rice *et al.*, *Trends Genet* **2000**, *16*, 276-277), preferencialmente versão 3.0.0 ou posterior. Os parâmetros opcionais usados são penalidade de lacuna aberta de 10, penalidade de extensão de lacuna de 0,5 e a matriz de substituição EBLOSUM62 (versão EMBOSS de BLOSUM62). O resultado de Needle marcado “identidade mais longa” (obtido usando a opção –nobrief) é usada como a percentagem de identidade e é calculada como se segue:

**[0070]** 
$$\text{Identidade} = \frac{(\text{Resíduos Idênticos} \times 100)}{(\text{Comprimento da Sequência Referenciada} - \text{Número Total de Lacunas no Alinhamento})}$$

**[0071]** Para propósitos descritos aqui, o grau de identidade de sequências entre duas sequências de desoxirribonucleotídeos é determinada usando o algoritmo de Needleman-Wunsch (Needleman e Wunsch, 1970, *supra*) como implementado no programa Needle do pacote EMBOSS (EMBOSS: The European Molecular Biology Open Software Suite, Rice *et al.*, 2000, *supra*), preferencialmente versão 3.0.0 ou posterior. Os parâmetros opcionais usados são penalidade de lacuna aberta de 10, penalidade de extensão de lacuna de 0,5 e a matriz de substituição EDNAFULL (versão EMBOSS de NCBI NUC4.4). O resultado de Needle marcado “identidade mais longa” (obtido usando a opção `-nobrief`) é usada como a porcentagem de identidade e é calculada como se segue:

$$\frac{(\text{Desoxirribonucleotídeos Idênticos} \times 100)}{(\text{Comprimento da Sequência Referenciada} - \text{Número Total de Lacunas no Alinhamento})}$$

**[0072] Peptídeo de sinal:** O termo “peptídeo de sinal” é definido aqui como um peptídeo ligado (fundido) em grelha ao terminal amino de um polipeptídeo tendo atividade biológica e dirige o polipeptídeo para a via secretora da célula. As sequências de sinal podem ser determinadas usando técnicas conhecidas na técnica (*Ver*, p.ex., Zhang e Henzel, 2004, *Protein Science* 13: 2819-2824). Os polipeptídeos descritos aqui podem compreender qualquer peptídeo de sinal adequado conhecido na técnica ou qualquer peptídeo de sinal descrito aqui (p.ex., o peptídeo de sinal MF $\alpha$ 1 de *S. cerevisiae* de SEQ ID NO: 7, o peptídeo de sinal EXG1 de *S. cerevisiae* de SEQ ID NO: 227 ou o peptídeo de sinal AG2 de *S. cerevisiae* de SEQ ID NO: 234 ou um peptídeo de sinal tendo pelo menos 80%, 85%, 90%, 95%, 96%, 97%, 98% ou 99% de identidade de sequências com ela).

**[0073] Trealase:** O termo “trealase” significa uma enzima que degrada a trealose em seus monossacarídeos unitários (*i.e.*, glucose). As trealases são classificadas em EC 3.2.1.28 (alfa,alfa-trealase) e EC. 3.2.1.93 (alfa,alfa-fosfotrealase). As classes de EC são baseadas nas

recomendações do *Nomenclature Committee* da *International Union of Biochemistry and Molecular Biology* (IUBMB). A descrição das classes de EC pode ser encontrada na internet, p.ex., em “<http://www.expasy.org/enzyme/>”. As trealases são enzimas que catalisam as seguintes reações:

EC 3.2.1.28: Alfa,alfa-trealose + H<sub>2</sub>O ⇌ 2 D-glucose;

EC 3.2.1. 93: Alfa,alfa-trealose 6-fosfato + H<sub>2</sub>O ⇌ D-glucose + D-glucose 6-fosfato.

**[0074]** Para propósitos da presente invenção, a atividade de trealase pode ser determinada de acordo com o ensaio de trealase descrito aqui na seção experimental.

**[0075] Condições de muito elevada estringência:** O termo “condições de muito elevada estringência” significa, para sondas com pelo menos 100 nucleotídeos em comprimento, pré-hibridação e hibridação a 42 °C em 5X SSPE, SDS a 0,3%, DNA de esperma de salmão cisalhado e desnaturado a 200 microgramas/mL e formamida a 50%, seguindo procedimentos de transferência de Southern padrão durante 12 a 24 horas. O material transportador é finalmente lavado três vezes cada uma durante 15 minutos usando 0,2X SSC, SDS a 0,2% a 70 °C.

**[0076] Condições de muito baixa estringência:** O termo “condições de muito baixa estringência” significa, para sondas com pelo menos 100 nucleotídeos em comprimento, pré-hibridação e hibridação a 42 °C em 5X SSPE, SDS a 0,3%, DNA de esperma de salmão cisalhado e desnaturado a 200 microgramas/mL e formamida a 25%, seguindo procedimentos de transferência de Southern padrão durante 12 a 24 horas. O material transportador é finalmente lavado três vezes cada uma durante 15 minutos usando 0,2X SSC, SDS a 0,2% a 45 °C.

**[0077] Xilanase:** O termo “xilanase” significa uma 1,4-beta-D-xilana-xiloidrolase (E.C. 3.2.1.8) que catalisa a endo-hidrólise de ligações 1,4-beta-D-xilosídicas em xilanas. A atividade de xilanase pode ser



determinada com AZCL-arabinoxilana a 0,2% como substrato em TRITON® X-100 a 0,01% e fosfato de sódio a 200 mM pH 6 a 37 °C. Uma unidade de atividade de xilanase é definida como 1,0 µmole de azurina produzida por minuto a 37 °C, pH 6 a partir de AZCL-arabinoxilana a 0,2% como substrato em fosfato de sódio a 200 mM e pH 6.

**[0078] Xilose Isomerase:** O termo “Xilose Isomerase” ou “XI” significa uma enzima que pode catalisar D-xilose em D-xilulose *in vivo* e converter D-glucose em D-frutose *in vitro*. A xilose isomerase é também conhecida como “glucose isomerase” e é classificada como E.C. 5.3.1.5. Como a estrutura da enzima é muito estável, a xilose isomerase é um bom modelo para estudar as relações entre a estrutura e funções da proteína (Karimaki *et al.*, *Protein Eng Des Sel*, 12004, 17 (12): 861-869). A atividade de Xilose Isomerase pode ser determinada usando técnicas conhecidas na técnica (p.ex., um ensaio de enzima acoplado usando D-sorbitol desidrogenase, como descrito por Verhoeven *et al.*, 2017, *Sci Rep* 7, 46155).

**[0079]** A referência a “cerca de” um valor ou parâmetro aqui inclui modalidades que são dirigidas a esse valor ou parâmetro *per se*. Por exemplo, uma descrição se referindo a “cerca de X” inclui a modalidade “X”. Quando usado em combinação com valores medidos, “cerca de” inclui uma gama que engloba pelo menos a incerteza associada ao método de medição do valor particular e pode incluir uma gama de mais ou menos dois desvios padrão em torno do referido valor.

**[0080]** Do mesmo modo, a referência a um gene ou polipeptídeo que é “derivado de” outro gene ou polipeptídeo X inclui o gene ou polipeptídeo X.

**[0081]** Como usadas aqui e nas reivindicações anexas, as formas singulares “um(a)”, “ou” e “o/a” incluem referentes plurais a não ser que o contexto dite claramente de outro modo.

**[0082]** É entendido que as modalidades descritas aqui incluem “consistindo” e/ou “consistindo essencialmente em” modalidades. Como usada aqui, exceto onde o contexto requeira de outro modo devido à linguagem expressa ou implicação necessária, a palavra “compreendem” ou variações tais como “compreende” ou “compreendendo” é usada em um sentido inclusivo, *i.e.*, para especificar a presença das características indicadas, mas não para excluir a presença ou a adição de características adicionais em várias modalidades.

### **DESCRIÇÃO DETALHADA**

**[0083]** São descritos aqui, *inter alia*, métodos para produção de um produto de fermentação, tal como etanol, a partir de material contendo amido ou celulose.

**[0084]** Durante a fermentação à escala industrial, as leveduras encontram vários desafios fisiológicos incluindo concentrações variáveis de açúcares, elevadas concentrações de metabolitos de leveduras tais como etanol, glicerol, ácidos orgânicos, estresse osmótico, bem como potencial competição por parte de micróbios contaminantes tais como leveduras e bactérias selvagens. Em consequência, muitas leveduras não são adequadas para uso na fermentação industrial. A estirpe industrial comercialmente disponível mais amplamente usada de *Saccharomyces* (*i.e.*, para fermentação à escala industrial) é a estirpe de *Saccharomyces cerevisiae* usada, por exemplo, no produto ETHANOL RED<sup>®</sup>. Esta estirpe é bem adequada para a produção industrial de etanol; no entanto, ainda não está claro como é que modificações à levedura irão impactar o desempenho. Em particular, a expressão funcional de enzimas heterólogas por uma levedura de *Saccharomyces cerevisiae* industrialmente relevante é incerta (Ver, por exemplo US 9,206,444 onde o requerente foi incapaz de expressar funcionalmente numerosas enzimas/classes de enzimas).

**[0085]** O Requerente descobriu surpreendentemente que leveduras expressando certas alfa-amilases e/ou trealases proporcionam propriedades benéficas que podem ser úteis para a fermentação de etanol.

**[0086]** Em um aspecto está um método de produção de um produto de fermentação a partir de material contendo amido ou contendo celulose compreendendo:

(a) sacarificação do material contendo amido ou contendo celulose; e  
(b) fermentação do material sacarificado do passo (a) com um organismo fermentador;

em que o organismo fermentador compreende um polinucleotídeo heterólogo codificando uma alfa-amilase e um polinucleotídeo heterólogo codificando uma trealase.

**[0087]** Os passos de sacarificação e fermentação são levados a cabo sequencial ou simultaneamente (SSF). Em uma modalidade, os passos de sacarificação e fermentação são levados a cabo simultaneamente (SSF). Em outra modalidade, os passos de sacarificação e fermentação são levados a cabo sequencialmente.

#### Organismo fermentador

**[0088]** O organismo fermentador descrito aqui pode ser derivado de qualquer célula hospedeira conhecida do perito capaz de produzir um produto de fermentação, tal como etanol. Como usado aqui, um “derivado” de estirpe é derivado de uma estirpe referenciada, tal como através de mutagênese, tecnologia de DNA recombinante, acoplamento, fusão celular ou citodução entre estirpes de leveduras. Os peritos na técnica entenderão que as alterações genéticas, incluindo modificações metabólicas exemplificadas aqui, podem ser descritas com referência a um organismo hospedeiro adequado e suas reações metabólicas correspondentes ou um organismo de origem adequado para material genético desejado tal como genes para uma via metabólica desejada. No entanto, dado o

sequenciamento de genoma completo de uma ampla variedade de organismos e o elevado nível de perícia na área da genômica, os peritos na técnica podem aplicar os ensinamentos e orientações proporcionados aqui a outros organismos. Por exemplo, as alterações metabólicas exemplificadas aqui podem ser prontamente aplicadas a outras espécies por incorporação do mesmo ácido nucleico codificante ou ácido nucleico codificante análogo de espécies sem ser as espécies referenciadas.

**[0089]** As células hospedeiras para preparação das células recombinantes descritas aqui podem ser de qualquer hospedeiro adequado, tal como uma estirpe de levedura, incluindo, mas não se limitando a, uma célula de *Saccharomyces*, *Rhodotorula*, *Schizosaccharomyces*, *Kluyveromyces*, *Pichia*, *Hansenula*, *Rhodosporidium*, *Candida*, *Yarrowia*, *Lipomyces*, *Cryptococcus*, ou *Dekkera sp.*. Em particular estão contempladas células hospedeiras de *Saccharomyces*, tais como células de *Saccharomyces cerevisiae*, *bayanus* ou *carlsbergensis*. Preferencialmente, a célula de levedura é uma célula de *Saccharomyces cerevisiae*. As células adequadas podem, por exemplo, ser derivadas de estirpes comercialmente disponíveis e estirpes industriais poliploides ou aneuploides, incluindo aquelas mas não se limitando àquelas de Superstart™, THERMOSACC®, C5 FUEL™, XyloFerm®, etc. (Lallemand); RED STAR e ETHANOL RED® (Fermentis/Lesaffre); FALI (AB Mauri); Baker's Best Yeast, Baker's Compressed Yeast, etc. (Fleishmann's Yeast); BIOFERM AFT, XP, CF, e XR (North American Bioproducts Corp.); Turbo Yeast (Gert Strand AB); e FERMIOL® (DSM Specialties). Outras estirpes de leveduras úteis estão disponíveis a partir de depositórios biológicos tais como a *American Type Culture Collection* (ATCC) ou a *Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH* (DSMZ), tais como, p.ex., BY4741 (p.ex., ATCC 201388); Y108-1 (ATCC PTA.10567) e NRRL YB-1952 (*ARS Culture Collection*). Ainda outras estirpes de *S. cerevisiae* são adequadas como

células hospedeiras DBY746, [Alfa][Eta]22, S150-2B, GPY55-15Ba, CEN.PK, USM21, TMB3500, TMB3400, VTT-A-63015, VTT-A-85068, VTT-c-79093 e seus derivados bem como *Saccharomyces* sp. 1400, 424A (LNH-ST), 259A (LNH-ST) e seus derivados. Em uma modalidade, a célula recombinante é um derivado de uma estirpe CIBTS1260 de *Saccharomyces cerevisiae* (depositada sob o No. de Acesso NRRL Y-50973 na *Agricultural Research Service Culture Collection* (NRRL), Illinois 61604 E.U.A.).

**[0090]** O organismo fermentador pode ser estirpe de *Saccharomyces*, p.ex., estirpe de *Saccharomyces cerevisiae* produzida usando o método descrito e em causa na patente dos EUA no. 8,257,959-BB.

**[0091]** A estirpe pode ser também um derivado da estirpe NMI V14/004037 de *Saccharomyces cerevisiae* (*Ver* WO2015/143324 e WO2015/143317 cada uma incorporada aqui por referência), nos. de estirpes V15/004035, V15/004036 e V15/004037 (*Ver* WO2016/153924 incorporada aqui por referência), nos. de estirpes V15/001459, V15/001460, V15/001461 (*Ver* WO2016/138437 incorporada aqui por referência), no. de estirpe NRRL Y67342 (*Ver* WO2017/063159 incorporada aqui por referência) ou qualquer estirpe descrita em WO2017/087330 (incorporada aqui por referência).

**[0092]** Os organismos fermentadores de acordo com a invenção foram gerados de modo a melhorarem o rendimento da fermentação e a melhorarem a economia do processo por corte dos custos com enzimas uma vez que parte das ou todas as enzimas necessárias para melhorar o desempenho do método são para serem produzidas pelo organismo fermentador.

**[0093]** Os organismos fermentadores descritos aqui podem utilizar vetores de expressão compreendendo a sequência codificante de um ou mais (p.ex., dois, vários) genes heterólogos ligados a uma ou mais sequências de controle que dirigem a expressão em uma célula adequada

sob condições compatíveis com a(s) sequência(s) de controle. Tais vetores de expressão podem ser usados em qualquer uma das células e métodos descritos aqui. Os polinucleotídeos descritos aqui podem ser manipulados em uma variedade de modos para proporcionar expressão de um polipeptídeo desejado. A manipulação do polinucleotídeo antes da sua inserção em um vetor pode ser desejável ou necessária dependendo do vetor de expressão. As técnicas para modificação de polinucleotídeos utilizando métodos de DNA recombinante são bem conhecidos na técnica.

**[0094]** Um construto ou vetor (ou múltiplos construtos ou vetores) compreendendo o um ou mais (p.ex., dois, vários) genes heterólogos pode ser introduzido em uma célula tal que o construto ou vetor seja mantido como um integrante cromossômico ou como um vetor extracromossômico autorreplicante como descrito anteriormente.

**[0095]** As várias sequências de nucleotídeos e de controle podem ser unidas para produzir um vetor de expressão recombinante que pode incluir um ou mais (p.ex., dois, vários) locais de restrição convenientes para permitir a inserção ou substituição do polinucleotídeo em tais locais. Alternativamente, o(s) polinucleotídeo(s) pode(m) ser expresso(s) por inserção do(s) polinucleotídeo(s) ou um construto de ácido nucleico compreendendo a sequência em um vetor apropriado para expressão. Na criação do vetor de expressão, a sequência codificante está localizada no vetor tal que a sequência codificante esteja operacionalmente ligada às sequências de controle apropriadas para expressão.

**[0096]** O vetor de expressão recombinante pode ser qualquer vetor (p.ex., um plasmídeo ou vírus) que possa ser convenientemente sujeito a procedimentos de DNA recombinante e possa provocar expressão do polinucleotídeo. A escolha do vetor dependerá tipicamente da compatibilidade do vetor com a célula hospedeira na qual o

vetor é para ser introduzido. O vetor pode ser um plasmídeo linear ou circular fechado.

**[0097]** O vetor pode ser um vetor autonomamente replicante, *i.e.*, um vetor que existe como uma entidade extracromossômica, cuja replicação é independente da replicação cromossômica, p.ex., um plasmídeo, um elemento extracromossômico, um minicromossomo ou um cromossomo artificial. O vetor pode conter quaisquer meios para assegurar autorreplicação. Alternativamente, o vetor pode ser um que, quando introduzido na célula hospedeira, seja integrado no genoma e replicado em conjunto com o(s) cromossomo(s) no(s) qual(ais) foi integrado. Além do mais pode ser usado um vetor ou plasmídeo único ou dois ou mais vetores ou plasmídeos que contêm em conjunto o DNA total a ser introduzido no genoma da célula ou um transpóson.

**[0098]** O vetor de expressão pode conter qualquer sequência promotora adequada que seja reconhecida por uma célula para expressão de um gene descrito aqui. A sequência promotora contém sequências de controle transcricionais que medeiam a expressão do polipeptídeo. O promotor pode ser qualquer polinucleotídeo que mostre atividade transcricional na célula de escolha incluindo promotores mutantes, truncados e híbridos e pode ser obtido de genes codificando polipeptídeos extracelulares ou intracelulares homólogos ou heterólogos para a célula.

**[0099]** Cada polinucleotídeo heterólogo descrito aqui pode estar operacionalmente ligado a um promotor que é estranho ao polinucleotídeo. Por exemplo, em uma modalidade, o polinucleotídeo heterólogo codificando o transportador de hexose está operacionalmente ligado a um promotor estranho ao polinucleotídeo. Os promotores podem ser idênticos a ou partilhar um elevado grau de identidade de sequências (p.ex., pelo menos cerca de 80%, pelo menos cerca de 85%, pelo menos cerca de 90%, pelo menos cerca de 95%, pelo menos cerca de 96%, pelo menos cerca

de 97%, pelo menos cerca de 98% ou pelo menos cerca de 99%) com um promotor nativo selecionado.

**[0100]** Exemplos de promotores adequados para dirigir a transcrição dos construtos de ácido nucleico em células de levedura incluem os, mas não estão limitados aos, promotores obtidos dos genes da enolase, (p.ex., enolase de *S. cerevisiae* ou enolase de *I. orientalis* (ENO1)), galactocinase (p.ex., galactocinase de *S. cerevisiae* ou galactocinase de *I. orientalis* (GAL1)), álcool desidrogenase/gliceraldeído-3-fosfato desidrogenase (p.ex., álcool desidrogenase/gliceraldeído-3-fosfato desidrogenase de *S. cerevisiae* ou álcool desidrogenase/gliceraldeído-3-fosfato desidrogenase de *I. orientalis* (ADH1, ADH2/GAP)), triose fosfato isomerase (p.ex., triose fosfato isomerase de *S. cerevisiae* e triose fosfato isomerase de *I. orientalis* (TPI)), metalotioneína (p.ex., metalotioneína de *S. cerevisiae* ou metalotioneína de *I. orientalis* (CUP1)), 3-fosfoglicerato cinase (p.ex., 3-fosfoglicerato cinase de *S. cerevisiae* ou 3-fosfoglicerato cinase de *I. orientalis* (PGK)), genes de PDC1, xilose redutase (XR), xilitol desidrogenase (XDH), L-(+)-lactato-citocromo c oxidorreductase (CYB2), fator de alongamento de tradução-1 (TEF1), fator de alongamento de tradução-2 (TEF2), gliceraldeído-3-fosfato desidrogenase (GAPDH) e orotidina 5'-fosfato descarboxilase (URA3). Outros promotores adequados podem ser obtidos dos genes TDH3, HXT7, PGK1, RPL18B e CCW12 de *S. cerevisiae*. Promotores úteis adicionais para células hospedeiras de levedura são descritos por Romanos *et al.*, 1992, *Yeast* 8: 423-488.

**[0101]** A sequência de controle pode ser também uma sequência terminadora da transcrição adequada, que seja reconhecida por uma célula hospedeira para terminar a transcrição. A sequência terminadora está operacionalmente ligada ao terminal 3' do polinucleotídeo codificando o polipeptídeo. Pode ser usado qualquer terminador que seja funcional na célula de levedura de escolha. O terminador pode ser idêntico a ou partilhar



um elevado grau de identidade de sequências (p.ex., pelo menos cerca de 80%, pelo menos cerca de 85%, pelo menos cerca de 90%, pelo menos cerca de 95%, pelo menos cerca de 96%, pelo menos cerca de 97%, pelo menos cerca de 98% ou pelo menos cerca de 99%) com um terminador nativo selecionado.

**[0102]** Terminadores adequados para células hospedeiras de levedura podem ser obtidos dos genes da enolase (p.ex., citocromo C de enolase de *S. cerevisiae* ou *I. orientalis* (p.ex., citocromo de *S. cerevisiae* ou *I. orientalis* (CYC1)), gliceraldeído-3-fosfato desidrogenase (p.ex., gliceraldeído-3-fosfato desidrogenase de *S. cerevisiae* ou *I. orientalis* (gpd)), PDC1, XR, XDH, transaldolase (TAL), transcetolase (TKL), ribose 5-fosfato cetol-isomerase (RKI), CYB2 e a família galactose de genes (especialmente o terminador GAL10). Outros terminadores adequados podem ser obtidos dos genes ENO2 ou TEF1 de *S. cerevisiae*. Terminadores úteis adicionais para células hospedeiras de levedura são descritos por Romanos *et al.*, 1992, *supra*.

**[0103]** A sequência de controle pode também ser uma região estabilizante de mRNA a jusante de um promotor e a montante da sequência codificante de um gene que aumenta a expressão do gene.

**[0104]** Exemplos de regiões estabilizantes de mRNA adequadas são obtidos de um gene *cryIIIA* de *Bacillus thuringiensis* (WO94/25612) e um gene SP82 de *Bacillus subtilis* (Hue *et al.*, 1995, *Journal of Bacteriology* 177: 3465-3471).

**[0105]** A sequência de controle pode ser também uma sequência líder adequada, quando transcrita é uma região não traduzida de um mRNA que é importante para tradução pela célula hospedeira. A sequência líder está operacionalmente ligada ao terminal 5' do polinucleotídeo codificando o polipeptídeo. Pode ser usada qualquer sequência líder que seja funcional na célula de levedura de escolha.

**[0106]** Líderes adequados para células hospedeiras de levedura são obtidos dos genes da enolase (p.ex., enolase de *S. cerevisiae* ou *I. orientalis* (ENO-1)), 3-fosfoglicerato cinase (p.ex., 3-fosfoglicerato cinase de *S. cerevisiae* ou *I. orientalis*), fator alfa (p.ex., fator alfa de *S. cerevisiae* ou *I. orientalis*) e álcool desidrogenase/gliceraldeído-3-fosfato desidrogenase (p.ex., álcool desidrogenase/gliceraldeído-3-fosfato desidrogenase de *S. cerevisiae* ou *I. orientalis* (ADH2/GAP)).

**[0107]** A sequência de controle pode ser também uma sequência de poliadenilação; uma sequência operacionalmente ligada ao terminal 3' do polinucleotídeo e, quando transcrita, é reconhecida pela célula hospedeira como um sinal para adicionar resíduos de poliadenosina ao mRNA transcrito. Pode ser usada qualquer sequência de poliadenilação que seja funcional na célula hospedeira de escolha. Sequências de poliadenilação úteis para células de levedura são descritas por Guo e Sherman, 1995, *Mol. Cellular Biol.* 15: 5983-5990.

**[0108]** Pode ser também desejável adicionar sequências reguladoras que permitem a regulação da expressão do polipeptídeo em relação ao crescimento da célula hospedeira. Exemplos de sistemas reguladores são aqueles que fazem com que a expressão do gene seja ligada ou desligada 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*. Nas leveduras pode ser usado o sistema ADH2 ou sistema GAL1.

**[0109]** Os vetores podem conter um ou mais (p.ex., dois, vários) marcadores selecionáveis que permitem a fácil seleção de células transformadas, transfectadas, transduzidas ou similares. Um marcador selecionável é um gene cujo produto proporciona resistência biocida ou viral, resistência a metais pesados, prototrofia a auxotróficos e similares.

Marcadores adequados para células hospedeiras de levedura incluem, mas não estão limitados a, ADE2, HIS3, LEU2, LYS2, MET3, TRP1 e URA3.

**[0110]** Os vetores podem conter um ou mais (p.ex., dois, vários) elementos que permitem a integração do vetor no genoma da célula hospedeira ou replicação autônoma do vetor na célula independente do genoma.

**[0111]** Para integração no genoma da célula hospedeira, o vetor pode se basear na sequência de polinucleotídeos codificando o polipeptídeo ou qualquer outro elemento do vetor para integração no genoma por recombinação homóloga ou não homóloga. Alternativamente, o vetor pode conter polinucleotídeos adicionais para dirigir a integração por recombinação homóloga no genoma da célula hospedeira em uma localização(ões) precisa(s) no(s) cromossomo(s). Para aumentar a probabilidade de integração em uma localização precisa, os elementos de integração devem conter um número suficiente de ácidos nucleicos, tal como 100 a 10.000 pares de bases, 400 a 10.000 pares de bases e 800 a 10.000 pares de bases, que têm um elevado grau de identidade de sequências com a sequência alvo correspondente para intensificar a probabilidade de recombinação homóloga. Os elementos de integração podem ser qualquer sequência que seja homóloga com a sequência alvo no genoma da célula hospedeira. Além do mais, os elementos de integração podem ser polinucleotídeos não codificantes ou codificantes. Por outro lado, o vetor pode ser integrado no genoma da célula hospedeira por recombinação não homóloga. Os potenciais *loci* de integração incluem aqueles descritos na técnica (p.ex., Ver US2012/0135481).

**[0112]** Para replicação autônoma, o vetor pode compreender adicionalmente uma origem de replicação permitindo que o vetor se replique autonomamente na célula de levedura. A origem de replicação pode ser qualquer replicador de plasmídeo mediando a replicação

autônoma que funciona em uma célula. O termo “origem de replicação” ou “replicador de plasmídeo” significa um polinucleotídeo que permite que um plasmídeo ou vetor se replique *in vivo*. Exemplos de origens de replicação para uso em uma célula hospedeira de levedura são a origem de replicação de 2 microns, ARS1, ARS4, a combinação de ARS1 e CEN3 e a combinação de ARS4 e CEN6.

**[0113]** Mais do que uma cópia de um polinucleotídeo descrito aqui pode ser inserida em uma célula hospedeira para aumentar a produção de um polipeptídeo. Pode ser obtido um aumento no número de cópias do polinucleotídeo por integração de pelo menos uma cópia adicional da sequência no genoma da célula de levedura ou por inclusão de 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 por cultivo das células na presença do agente selecionável apropriado.

**[0114]** Os procedimentos usados para ligar os elementos descritos acima para construir os vetores de expressão recombinantes descritos aqui são bem conhecidos de um perito na técnica (ver, p.ex., Sambrook *et al.*, 1989, *Molecular Cloning, A Laboratory Manual*, 2ª edição, Cold Spring Harbor, Nova Iorque).

**[0115]** Procedimentos e técnicas adicionais conhecidos na técnica para a preparação de células recombinantes para fermentação de etanol, são descritos, por exemplo, em WO2016/045569, o conteúdo da qual é deste modo incorporado por referência.

**[0116]** O organismo fermentador pode estar na forma de uma composição compreendendo um organismo fermentador (p.ex., uma estirpe de levedura descrita aqui) e um componente ocorrendo naturalmente e/ou não ocorrendo naturalmente.

**[0117]** O organismo fermentador descrito aqui pode estar em qualquer forma viável, incluindo forma desintegrada, seca, incluindo seca ativa e instantânea, comprimida, creme (líquido), *etc.* Em uma modalidade, o organismo fermentador (p.ex., uma estirpe de levedura de *Saccharomyces cerevisiae*) é levedura seca, tal como levedura seca ativa ou levedura instantânea. Em uma modalidade, o organismo fermentador (p.ex., uma estirpe de levedura de *Saccharomyces cerevisiae*) é uma levedura desintegrada. Em uma modalidade, o organismo fermentador (p.ex., uma estirpe de levedura de *Saccharomyces cerevisiae*) é uma levedura comprimida. Em uma modalidade, o organismo fermentador (p.ex., uma estirpe de levedura de *Saccharomyces cerevisiae*) é uma levedura em creme.

**[0118]** Em uma modalidade está uma composição compreendendo um organismo fermentador descrito aqui (p.ex., uma estirpe de levedura de *Saccharomyces cerevisiae*) e um ou mais dos componentes selecionados do grupo consistindo em: tensoativos, emulsificantes, gomas, agente de inchaço e antioxidantes e outros auxiliares de processamento.

**[0119]** As composições descritas aqui podem compreender um organismo fermentador descrito aqui (p.ex., uma estirpe de levedura de *Saccharomyces cerevisiae*) e quaisquer tensoativos adequados. Em uma modalidade, o(s) tensoativo(s) é/são (um) tensoativo(s) aniônico(s), tensoativo(s) catiônico(s) e/ou tensoativo(s) não iônico(s).

**[0120]** As composições descritas aqui podem compreender um organismo fermentador descrito aqui (p.ex., uma estirpe de levedura de *Saccharomyces cerevisiae*) e qualquer emulsificante adequado. Em uma modalidade, o emulsificante é um éster de ácidos graxos de sorbitana. Em uma modalidade, o emulsificante é selecionado do grupo de monoestearato de sorbitana (SMS), ésteres de ácido cítrico de

monodiglicerídeos, poligliceroléster, ésteres de ácidos graxos de propilenoglicol.

**[0121]** Em uma modalidade, a composição compreende um organismo fermentador descrito aqui (p.ex., uma estirpe de levedura de *Saccharomyces cerevisiae*) e Olindronal SMS, Olindronal SK ou Olindronal SPL incluindo a composição em causa na Patente Europeia No. 1,724,336 (deste modo incorporada por referência). Estes produtos estão comercialmente disponíveis a partir da Bussetti, Áustria, para leveduras secas ativas.

**[0122]** As composições descritas aqui podem compreender um organismo fermentador descrito aqui (p.ex., uma estirpe de levedura de *Saccharomyces cerevisiae*) e qualquer goma adequada. Em uma modalidade, a goma é selecionada do grupo de gomas de alfarroba, guar, tragacanto, arábica, xantana e acácia, em particular para leveduras em creme, comprimidas e secas.

**[0123]** As composições descritas aqui podem compreender um organismo fermentador descrito aqui (p.ex., uma estirpe de levedura de *Saccharomyces cerevisiae*) e qualquer agente de inchaço adequado. Em uma modalidade, o agente de dilatação é celulose de metila ou celulose de carboximetila.

**[0124]** As composições descritas aqui podem compreender um organismo fermentador descrito aqui (p.ex., uma estirpe de levedura de *Saccharomyces cerevisiae*) e qualquer antioxidante adequado. Em uma modalidade, o antioxidante é hidroxianisol butilado (BHA) e/ou hidroxitolueno butilado (BHT) ou ácido ascórbico (vitamina C), em particular para leveduras secas ativas.

#### Rupturas Gênicas

**[0125]** Os organismos fermentadores descritos aqui podem também compreender uma ou mais (p.ex., duas, várias) rupturas

gênicas, p.ex., para desviar o metabolismo de açúcares de produtos indesejados para o etanol. Em alguns aspectos, as células hospedeiras recombinantes produzem uma quantidade maior de etanol em comparação com a célula sem a uma ou mais rupturas quando cultivadas sob condições idênticas. Em alguns aspectos, um ou mais dos genes endógenos com ruptura são inativados.

**[0126]** Em certas modalidades, o organismo fermentador proporcionado aqui compreende uma ruptura de um ou mais genes endógenos codificando enzimas envolvidas na produção de produtos fermentativos alternativos tais como glicerol ou outros subprodutos tais como acetato ou dióis. Por exemplo, as células proporcionadas aqui podem compreender uma ruptura de uma ou mais de glicerol 3-fosfato desidrogenase (GPD, catalisa a reação de di-hidroxiacetona fosfato em glicerol 3-fosfato), glicerol 3-fosfatase (GPP, catalisa a conversão de glicerol-3 fosfato em glicerol), glicerol cinase (catalisa a conversão de glicerol 3-fosfato em glicerol), di-hidroxiacetona cinase (catalisa a conversão de di-hidroxiacetona fosfato em di-hidroxiacetona), glicerol desidrogenase (catalisa a conversão de di-hidroxiacetona em glicerol) e aldeído desidrogenase (ALD, p.ex., converte acetaldeído em acetato).

**[0127]** A análise de modelação pode ser usada para conceber rupturas gênicas que otimizam adicionalmente a utilização da via. Um método computacional exemplificativo para identificação e concepção de alterações metabólicas favorecendo a biossíntese de um produto desejado é a estrutura computacional OptKnock, Burgard *et al.*, 2003, *Biotechnol. Bioeng.* 84: 647-657.

**[0128]** Os organismos fermentadores compreendendo uma ruptura gênica podem ser construídos usando métodos bem conhecidos na técnica, incluindo aqueles métodos descritos aqui. Uma porção do gene pode ser submetida a ruptura tal como a região codificante ou uma sequência

de controle requerida para expressão da região codificante. Uma tal sequência de controle do gene pode ser uma sequência promotora ou uma sua parte funcional, *i.e.*, uma parte que é suficiente para afetar a expressão do gene. Por exemplo, uma sequência promotora pode ser inativada resultando em expressão nula, ou um promotor mais fraco pode ser substituído pela sequência promotora nativa para reduzir a expressão da sequência codificante. Outras sequências de controle para possível modificação incluem, mas não estão limitadas a, um líder, sequência de pró-peptídeo, sequência de sinal, terminador da transcrição e ativador transcricional.

**[0129]** Os organismos fermentadores compreendendo uma ruptura gênica podem ser construídos por técnicas de deleção gênica para eliminar ou reduzir a expressão do gene. As técnicas de deleção gênica permitem a remoção parcial ou completa do gene eliminando deste modo sua expressão. Em tais métodos, a deleção do gene é alcançada por recombinação homóloga usando um plasmídeo que foi construído para conter contiguamente as regiões 5' e 3' flanqueando o gene.

**[0130]** Os organismos fermentadores compreendendo uma ruptura gênica podem ser também construídos por introdução, substituição e/ou remoção de um ou mais (p.ex., dois, vários) nucleotídeos no gene ou uma sua sequência de controle requerida para a sua transcrição ou tradução. Por exemplo, os nucleotídeos podem ser inseridos ou removidos para a introdução de um códon de parada, a remoção do códon de início ou um desvio de grelha da grelha de leitura aberta. Uma tal modificação pode ser alcançada por mutagênese sítio-dirigida ou mutagênese gerada por PCR de acordo com métodos conhecidos na técnica. Ver, por exemplo, Botstein e Shortle, 1985, *Science* 229: 4719; Lo *et al.*, 1985, *Proc. Natl. Acad. Sci. U.S.A.* 81: 2285; Higuchi *et al.*, 1988, *Nucleic Acids Res* 16: 7351; Shimada, 1996, *Meth. Mol. Biol.* 57: 157; Ho *et al.*, 1989,



*Gene* 77: 61; Horton *et al.*, 1989, *Gene* 77: 61; e Sarkar e Sommer, 1990, *BioTechniques* 8: 404.

**[0131]** Os organismos fermentadores compreendendo uma ruptura gênica podem ser também construídos por inserção no gene de um construto de ácido nucleico de ruptura compreendendo um fragmento de ácido nucleico homólogo ao gene que criará uma duplicação da região de homologia e incorporará DNA de construto entre as regiões duplicadas. Uma tal ruptura gênica pode eliminar a expressão gênica se o construto inserido separar o promotor do gene da região codificante ou interromper a sequência codificante tal que resulte um produto gênico não funcional. Um construto de ruptura pode ser simplesmente um gene marcador selecionável acompanhado por regiões 5' e 3' homólogas ao gene. O marcador selecionável permite a identificação de transformantes contendo o gene com ruptura.

**[0132]** Os organismos fermentadores compreendendo uma ruptura gênica podem ser também construídos pelo processo de conversão gênica (ver, por exemplo, Iglesias e Trautner, 1983, *Molecular General Genetics* 189: 73-76). Por exemplo, no método de conversão gênica, uma sequência de nucleotídeos correspondendo ao gene é mutagenizada *in vitro* para produzir uma sequência de nucleotídeos defeituosa, que é depois transformada na estirpe recombinante para produzir um gene defeituoso. Por recombinação homóloga, a sequência de nucleotídeos defeituosa substitui o gene endógeno. Pode ser desejável que a sequência de nucleotídeos defeituosa compreenda também um marcador para seleção de transformante contendo o gene defeituoso.

**[0133]** Os organismos fermentadores compreendendo uma ruptura gênica podem ser adicionalmente construídos por mutagênese aleatória ou específica usando métodos bem conhecidos na técnica, incluindo, mas não se limitando a, mutagênese química (ver, por exemplo,

Hopwood, *The Isolation of Mutants in Methods in Microbiology* (J.R. Norris e D.W. Ribbons, eds.) pp. 363-433, Academic Press, Nova Iorque, 1970). A modificação do gene pode ser realizada por sujeição da estirpe parental à mutagênese e rastreamento quanto a estirpes mutantes nas quais a expressão do gene tenha sido reduzida ou inativada. A mutagênese, que pode ser específica ou aleatória, pode ser realizada, por exemplo, pelo uso de um agente de mutagenização físico ou químico adequado, uso de um oligonucleotídeo adequado ou sujeição da sequência de DNA à mutagênese gerada por PCR. Além do mais, a mutagênese pode ser realizada por uso de qualquer combinação destes métodos de mutagenização.

**[0134]** Exemplos de um agente de mutagênese físico ou químico adequado para o presente propósito incluem irradiação com ultravioleta (UV), hidroxilamina, N-metil-N'-nitro-N-nitrosoguanidina (MNNG), N-metil-N'-nitrosoguanidina (NTG) O-metil-hidroxilamina, ácido nitroso, sulfanato de etilmetano (EMS), bissulfato de sódio, ácido fórmico e análogos de nucleotídeo. Quando tais agentes são usados, a mutagênese é tipicamente realizada por incubação da estirpe parental a ser mutagenizada na presença do agente de mutagenização de escolha sob condições adequadas e seleção quanto aos mutantes exibindo expressão reduzida ou nula do gene.

**[0135]** Pode ser usada uma sequência de nucleotídeos homóloga ou complementar a um gene descrito aqui a partir de outras fontes microbianas para submeter a ruptura o gene correspondente em uma estirpe recombinante de escolha.

**[0136]** Em um aspecto, a modificação de um gene na célula recombinante não é marcada com um marcador selecionável. A remoção do gene marcador selecionável pode ser alcançada por cultivo dos mutantes em um meio de contrasseleção. Onde o gene marcador selecionável contém repetições flanqueando suas extremidades 5' e 3', as

repetições facilitarão o *looping* para fora do gene marcador selecionável por recombinação homóloga quando a estirpe mutante é submetida a contrasseleção. O gene marcador selecionável pode ser também removido por recombinação homóloga por introdução na estirpe mutante de um fragmento de ácido nucleico compreendendo regiões 5' e 3' do gene defeituoso, mas não tendo o gene marcador selecionável, seguida por seleção no meio de contrasseleção. Por recombinação homóloga, o gene defeituoso contendo o gene marcador selecionável é substituído pelo fragmento do ácido nucleico não tendo o gene marcador selecionável. Podem ser também usados outros métodos conhecidos na técnica.

#### Métodos usando um Material Contendo Amido

**[0137]** Em alguns aspectos, os métodos descritos aqui produzem um produto de fermentação a partir de um material contendo amido. Material contendo amido é bem conhecido na técnica, contendo dois tipos de homopolissacarídeos (amilose e amilopectina) e é ligado por ligações alfa-(1-4)-D-glicosídicas. Pode ser usado qualquer material de partida adequado contendo amido. O material de partida é geralmente selecionado com base no produto de fermentação desejado, tal como etanol. Exemplos de materiais de partida contendo amido incluem cereais, tubérculos ou grãos. Especificamente, o material contendo amido pode ser milho, trigo, cevada, centeio, milo, sagu, mandioca, tapioca, sorgo, aveia, arroz, ervilhas, feijão ou batatas-doces ou suas misturas. Estão também contemplados tipos cerosos e não cerosos de milho e cevada.

**[0138]** Em uma modalidade, o material de partida contendo amido é milho. Em uma modalidade, o material de partida contendo amido é trigo. Em uma modalidade, o material de partida contendo amido é cevada. Em uma modalidade, o material de partida contendo amido é centeio. Em uma modalidade, o material de partida contendo amido é milo. Em uma modalidade, o material de partida contendo amido é sagu. Em uma

modalidade, o material de partida contendo amido é mandioca. Em uma modalidade, o material de partida contendo amido é tapioca. Em uma modalidade, o material de partida contendo amido é sorgo. Em uma modalidade, o material de partida contendo amido é arroz. Em uma modalidade, o material de partida contendo amido é ervilhas. Em uma modalidade, o material de partida contendo amido é feijões. Em uma modalidade, o material de partida contendo amido é batatas-doces. Em uma modalidade, o material de partida contendo amido é aveia.

**[0139]** Os métodos usando um material contendo amido podem incluir um processo convencional (p.ex., incluindo um passo de liquefação descrito em mais detalhe em baixo) ou um processo de hidrólise de amido em bruto. Em algumas modalidades usando um material contendo amido, a sacarificação do material contendo amido é a uma temperatura acima da temperatura de gelatinização inicial. Em algumas modalidades usando um material contendo amido, a sacarificação do material contendo amido é a uma temperatura abaixo da temperatura de gelatinização inicial.

#### Liquefação

**[0140]** Em aspectos usando um material contendo amido, os métodos podem compreender adicionalmente um passo de liquefação levado a cabo por sujeição do material contendo amido a uma temperatura acima da temperatura de gelatinização inicial a uma alfa-amilase e opcionalmente uma protease e/ou uma glucoamilase. Podem também estar presentes e/ou ser adicionadas na liquefação outras enzimas tais como uma pululanase e uma fitase. Em algumas modalidades, o passo de liquefação é levado a cabo antes dos passos a) e b) dos métodos descritos.

**[0141]** O passo de liquefação pode ser levado a cabo durante 0,5-5 horas, tal como 1-3 horas, tal como tipicamente cerca de 2 horas.

**[0142]** O termo “temperatura de gelatinização inicial” significa a temperatura mais baixa à qual a gelatinização do material contendo amido começa. Em geral, o amido aquecido em água começa a gelatinizar entre cerca de 50 °C e 75 °C; a temperatura exata da gelatinização depende do amido específico e pode ser prontamente determinada pelo especialista. Assim, a temperatura de gelatinização inicial pode variar de acordo com a espécie de planta, com a variedade particular da espécie de planta, bem como com as condições de crescimento. A temperatura de gelatinização inicial de um dado material contendo amido pode ser determinada como a temperatura à qual a birrefringência é perdida em 5% dos grânulos de amido usando o método descrito por Gorinstein e Lii, 1992, *Starch/Stärke* 44 (12): 461-466.

**[0143]** A liquefação é tipicamente levada a cabo a uma temperatura na gama de 70-100 °C. Em uma modalidade, a temperatura na liquefação é entre 75-95 °C, tal como entre 75-90 °C, entre 80-90 °C ou entre 82-88 °C, tal como cerca de 85 °C.

**[0144]** O passo de cozimento a jato pode ser levado a cabo antes da liquefação no passo, por exemplo, a uma temperatura entre 110-145 °C, 120-140 °C, 125-135 °C ou cerca de 130 °C durante cerca de 1-15 minutos, durante cerca de 3-10 minutos ou cerca de 5 minutos.

**[0145]** O pH durante a liquefação pode variar entre 4 e 7, tal como um pH de 4,5-6,5, pH 5,0-6,5, pH 5,0-6,0, pH 5,2-6,2 ou cerca de 5,2, cerca de 5,4, cerca de 5,6 ou cerca de 5,8.

Em uma modalidade, o processo compreende adicionalmente, antes da liquefação, os passos de:

- i) redução do tamanho das partículas do material contendo amido, preferencialmente por moagem a seco;
- ii) formação de uma pasta semifluida compreendendo o material contendo amido e água.

**[0146]** O material de partida contendo amido, tal como grãos integrais, pode ser reduzido no tamanho das partículas, p.ex., por moagem, de modo a abrir a estrutura, para aumentar a área superficial e permitir processamento adicional. Geralmente existem dois tipos de processos: moagem a úmido e a seco. Na moagem a seco, os grãos inteiros são moídos e usados. A moagem a úmido dá uma boa separação do gérmen e farelo (grânulos de amido e proteína). A moagem a úmido é frequentemente aplicada em localizações onde o hidrolisado de amido é usado na produção de, p.ex., xaropes. Tanto a moagem a seco como a moagem a úmido são bem conhecidas na técnica de processamento de amido. Em uma modalidade, o material contendo amido é sujeito a moagem a seco. Em uma modalidade, o tamanho das partículas é reduzido até entre 0,05 e 3,0 mm, p.ex., 0,1-0,5 mm, ou tal que pelo menos 30%, pelo menos 50%, pelo menos 70% ou pelo menos 90% do material contendo amido passe através de um crivo com uma grelha de 0,05 a 3,0 mm, p.ex., grelha de 0,1-0,5 mm. Em outra modalidade, pelo menos 50%, p.ex., pelo menos 70%, pelo menos 80% ou pelo menos 90% do material contendo amido passa através de um crivo com grelha # 6.

**[0147]** A pasta semifluida aquosa pode conter 10-55% p/p de sólidos secos (DS), p.ex., 25-45% p/p de sólidos secos (DS) ou 30-40% p/p de sólidos secos (DS) de material contendo amido.

**[0148]** A alfa-amilase, opcionalmente uma protease, e opcionalmente uma glucoamilase, pode ser adicionada inicialmente à pasta semifluida aquosa para iniciar a liquefação (diluição). Em uma modalidade, somente uma porção das enzimas (p.ex., cerca de 1/3) é adicionada à pasta semifluida aquosa, enquanto o resto das enzimas (p.ex., cerca de 2/3) é adicionado durante o passo de liquefação.

**[0149]** Uma lista não exaustiva de alfa-amilases usadas na liquefação pode ser encontrada em baixo na seção “Alfa-Amilases”.

Exemplos de proteases adequadas usadas em liquefação incluem qualquer protease descrita supra na seção “Proteases”. Exemplos de glucoamilases adequadas usadas na liquefação incluem qualquer glucoamilase encontrada na seção “Glucoamilases”.

#### Sacarificação e Fermentação de Material contendo amido

**[0150]** Em aspectos usando um material contendo amido, uma glucoamilase pode estar presente e/ou ser adicionada no passo de sacarificação a) e/ou passo de fermentação b) ou sacarificação e fermentação simultâneas (SSF). A glucoamilase do passo de sacarificação a) e/ou passo de fermentação b) ou sacarificação e fermentação simultâneas (SSF) é tipicamente diferente da glucoamilase opcionalmente adicionada a qualquer passo de liquefação descrito *supra*. Em uma modalidade, a glucoamilase está presente e/ou é adicionada em conjunto com uma alfa-amilase fúngica.

**[0151]** Em alguns aspectos, o organismo fermentador compreende um polinucleotídeo heterólogo codificando uma glucoamilase, por exemplo, como descrito em WO2017/087330, o conteúdo da qual é deste modo incorporado por referência.

**[0152]** Exemplos de glucoamilases podem ser encontrados na seção “Glucoamilases”.

**[0153]** Quando se faz sacarificação e fermentação sequenciais, o passo de sacarificação a) pode ser levado a cabo sob condições bem conhecidas na técnica. Por exemplo, o passo a) de sacarificação pode durar até de cerca de 24 a cerca de 72 horas. Em uma modalidade, uma pré-sacarificação é feita. A pré-sacarificação é tipicamente feita durante 40-90 minutos a uma temperatura entre 30-65 °C, tipicamente cerca de 60 °C. A pré-sacarificação é, em uma modalidade, seguida por sacarificação durante a fermentação na sacarificação e fermentação simultâneas (SSF). A sacarificação é tipicamente levada a cabo a

temperaturas de 20-75 °C, preferencialmente de 40-70 °C, tipicamente cerca de 60 °C, e tipicamente a um pH entre 4 e 5, tal como cerca de pH 4,5.

**[0154]** A fermentação é levada a cabo em um meio de fermentação, como conhecido na técnica e, p.ex., como descrito aqui. O meio de fermentação inclui o substrato da fermentação, isto é, a fonte de carboidratos que é metabolizada pelo organismo fermentador. Com os processos descritos aqui, o meio de fermentação pode compreender nutrientes e estimulante(s) do crescimento para o(s) organismo(s) fermentador(es). Nutrientes e estimulantes do crescimento são amplamente usados na técnica de fermentação e incluem fontes de nitrogênio, tais como amônia; ureia, vitaminas e minerais ou suas combinações.

**[0155]** Geralmente, os organismos fermentadores tais como leveduras, incluindo levedura *Saccharomyces cerevisiae*, requerem uma fonte adequada de nitrogênio para propagação e fermentação. Podem ser usadas muitas fontes de nitrogênio suplementar, se necessário, e tais fontes de nitrogênio são bem conhecidas na técnica. A fonte de nitrogênio pode ser orgânica, tal como ureia, DDG, bolo úmido ou mosto de milho, ou inorgânica, tal como amônia ou hidróxido de amônio. Em uma modalidade, a fonte de nitrogênio é ureia.

**[0156]** A fermentação pode ser levada a cabo sob condições de baixo nitrogênio, p.ex., quando se usa uma levedura expressando proteases. Em algumas modalidades, o passo de fermentação é conduzido com menos do que 1000 ppm de nitrogênio suplementar (p.ex., ureia ou hidróxido de amônio), tal como menos do que 750 ppm, menos do que 500 ppm, menos do que 400 ppm, menos do que 300 ppm, menos do que 250 ppm, menos do que 200 ppm, menos do que 150 ppm, menos do que 100 ppm, menos do que 75 ppm, menos do que 50 ppm, menos do que 25 ppm ou menos do que 10 ppm de nitrogênio suplementar. Em algumas



modalidades, o passo de fermentação é conduzido sem qualquer nitrogênio suplementar.

**[0157]** O processo de sacarificação e fermentação simultâneas (“SSF”) é amplamente usado em processos de produção de produtos de fermentação à escala industrial, especialmente processos de produção de etanol. Quando se faz SSF, o passo de sacarificação a) e o passo de fermentação b) são levados a cabo simultaneamente. Não existe etapa de espera para a sacarificação, significando que um organismo fermentador, tal como levedura, e enzima(s), podem ser adicionados em conjunto. No entanto está também contemplado adicionar o organismo fermentador e a(s) separadamente. A SSF é tipicamente levada a cabo a uma temperatura de 25 °C a 40 °C, tal como de 28 °C a 35 °C, tal como de 30 °C a 34 °C ou cerca de 32 °C. Em uma modalidade, a fermentação decorre durante 6 a 120 horas, em particular 24 a 96 horas. Em uma modalidade, o pH é entre 4-5.

**[0158]** Em uma modalidade, uma composição de enzimas celulolíticas está presente e/ou é adicionada na sacarificação, fermentação ou sacarificação e fermentação simultâneas (SSF). Exemplos de tais composições de enzimas celulolíticas podem ser encontrados na seção “Enzimas e Composições Celulolíticas” em baixo. A composição de enzimas celulolíticas pode estar presente e/ou ser adicionada em conjunto com uma glucoamilase, tal como uma divulgada na seção “Glucoamilases” em baixo.

#### Alfa-Amilases

**[0159]** A alfa-amilase expressa e/ou exógena pode ser qualquer alfa-amilase que seja adequada para as células hospedeiras e/ou os métodos descritos aqui, tal como uma alfa-amilase ocorrendo naturalmente (p.ex., uma alfa-amilase nativa de outra espécie ou uma alfa-amilase endógena expressa a partir de um vetor de expressão modificado) ou uma sua variante que retenha atividade de alfa-amilase. Qualquer alfa-

amilase contemplada para expressão por um organismo fermentador descrito em baixo está também contemplada para aspectos da invenção envolvendo adição exógena de uma alfa-amilase.

**[0160]** Em algumas modalidades, o organismo fermentador compreende um polinucleotídeo heterólogo codificando uma alfa-amilase, por exemplo, como descrito em WO2017/087330, o conteúdo da qual é deste modo incorporado por referência. Qualquer alfa-amilase descrita ou referenciada aqui está contemplada para expressão no organismo fermentador.

**[0161]** Em algumas modalidades, o organismo fermentador compreendendo um polinucleotídeo heterólogo codificando uma alfa-amilase tem um nível aumentado de atividade de alfa-amilase em comparação com as células hospedeiras sem o polinucleotídeo heterólogo codificando a alfa-amilase, quando cultivados sob as mesmas condições. Em algumas modalidades, o organismo fermentador tem um nível aumentado de atividade de alfa-amilase de pelo menos 5%, p.ex., pelo menos 10%, pelo menos 15%, pelo menos 20%, pelo menos 25%, pelo menos 50%, pelo menos 100%, pelo menos 150%, pelo menos 200%, pelo menos 300% ou pelo menos 500% em comparação com o organismo fermentador sem o polinucleotídeo heterólogo codificando a alfa-amilase, quando cultivados sob as mesmas condições.

**[0162]** Alfa-amilases exemplificativas que podem ser usadas com as células hospedeiras e/ou os métodos descritos aqui incluem alfa-amilases bacterianas, de levedura ou fúngicas filamentosas, p.ex., derivadas de qualquer um dos microrganismos descritos ou referenciados aqui.

**[0163]** O termo “alfa-amilase bacteriana” significa qualquer alfa-amilase bacteriana classificada sob EC 3.2.1.1. Uma alfa-amilase bacteriana usada aqui pode, p.ex., ser derivada de uma estirpe do gênero

*Bacillus*, que é por vezes também referido como o gênero *Geobacillus*. Em uma modalidade, a alfa-amilase de *Bacillus* é derivada de uma estirpe de *Bacillus amyloliquefaciens*, *Bacillus licheniformis*, *Bacillus stearothermophilus* ou *Bacillus subtilis*, mas pode ser também derivada de outras *Bacillus sp.*

**[0164]** Exemplos específicos de alfa-amilases bacterianas incluem a alfa-amilase de *Bacillus stearothermophilus* (BSG) de SEQ ID NO: 3 em WO99/19467, a alfa-amilase de *Bacillus amyloliquefaciens* (BAN) de SEQ ID NO: 5 em WO99/19467 e a alfa-amilase de *Bacillus licheniformis* (BLA) de SEQ ID NO: 4 em WO99/19467 (todas as sequências são deste modo incorporadas por referência). Em uma modalidade, a alfa-amilase pode ser uma enzima tendo uma sequência de polipeptídeo maduro com um grau de identidade de pelo menos 60%, p.ex., pelo menos 70%, pelo menos 80%, pelo menos 90%, pelo menos 95%, pelo menos 96%, pelo menos 97%, pelo menos 98% ou pelo menos 99% com qualquer uma das sequências mostradas em SEQ ID NOs: 3, 4 ou 5 em WO99/19467.

**[0165]** Em uma modalidade, a alfa-amilase é derivada de *Bacillus stearothermophilus*. A alfa-amilase de *Bacillus stearothermophilus* pode ser um tipo selvagem maduro ou uma sua variante madura. As alfa-amilases maduras de *Bacillus stearothermophilus* podem ser naturalmente truncadas durante a produção recombinante. Por exemplo, a alfa-amilase de *Bacillus stearothermophilus* pode ser truncada no terminal C, tal que tenha de 480-495 aminoácidos em comprimento, tal como cerca de 491 aminoácidos em comprimento, p.ex., tal que não tenha um domínio funcional de ligação ao amido (em comparação com SEQ ID NO: 3 em WO99/19467).

**[0166]** A alfa-amilase de *Bacillus* pode ser também uma variante e/ou híbrido. Exemplos de uma tal variante podem ser encontrados em qualquer uma de WO96/23873, WO96/23874, WO97/41213, WO99/19467, WO00/60059 e WO02/10355 (cada uma deste modo

incorporada por referência). Variantes de alfa-amilase específicas são divulgadas nas Patentes dos E.U.A. Nos. 6,093,562, 6,187,576, 6,297,038 e 7,713,723 (deste modo incorporadas por referência) e incluem variantes de alfa-amilase de *Bacillus stearothermophilus* (frequentemente referida como alfa-amilase BSG) tendo uma deleção de um ou dois aminoácidos nas posições R179, G180, I181 e/ou G182, preferencialmente uma dupla deleção divulgada em WO96/23873 – ver, p.ex., página 20, linhas 1-10 (deste modo incorporada por referência), tal como correspondendo à deleção das posições I181 e G182 em comparação com a sequência de aminoácidos da alfa-amilase de *Bacillus stearothermophilus* indicada em SEQ ID NO: 3 divulgada em WO99/19467 ou a deleção dos aminoácidos R179 e G180 usando SEQ ID NO: 3 em WO99/19467 para numeração (cuja referência é deste modo incorporada por referência). Em algumas modalidades, as alfa-amilases de *Bacillus*, tais como alfa-amilases de *Bacillus stearothermophilus*, têm uma dupla deleção correspondendo a uma deleção das posições 181 e 182 e compreendem adicionalmente opcionalmente uma substituição N193F (também denotada I181\* + G182\* + N193F) em comparação com a sequência de aminoácidos da alfa-amilase BSG de tipo selvagem indicada em SEQ ID NO: 3 divulgada em WO99/19467. A alfa-amilase bacteriana pode ter também uma substituição em uma posição correspondendo a S239 na alfa-amilase de *Bacillus licheniformis* mostrada em SEQ ID NO: 4 em WO99/19467 ou uma variante S242 e/ou E188P da alfa-amilase de *Bacillus stearothermophilus* de SEQ ID NO: 3 em WO99/19467.

**[0167]** Em uma modalidade, a variante é uma variante S242A, E ou Q, p.ex., uma variante S242Q, da alfa-amilase de *Bacillus stearothermophilus*.

**[0168]** Em uma modalidade, a variante é uma variante da posição E188, p.ex., uma variante E188P da alfa-amilase de *Bacillus stearothermophilus*.

**[0169]** A alfa-amilase bacteriana pode, em uma modalidade, ser uma alfa-amilase de *Bacillus truncada*. Em uma modalidade, a truncação é tal que, p.ex., a alfa-amilase de *Bacillus stearothermophilus* mostrada em SEQ ID NO: 3 em WO99/19467, tem cerca de 491 aminoácidos em comprimento, tal como de 480 a 495 aminoácidos em comprimento ou, então, não tem um domínio funcional de ligação ao amido.

**[0170]** A alfa-amilase bacteriana pode ser também uma alfa-amilase bacteriana híbrida, p.ex., uma alfa-amilase compreendendo 445 resíduos de aminoácidos C-terminais da alfa-amilase de *Bacillus licheniformis* (mostrada em SEQ ID NO: 4 de WO99/19467) e os 37 resíduos de aminoácidos N-terminais da alfa-amilase derivada de *Bacillus amyloliquefaciens* (mostrada em SEQ ID NO: 5 de WO99/19467). Em uma modalidade, este híbrido tem uma ou mais das, especialmente todas as, seguintes substituições: G48A + T49I + G107A + H156Y + A181T + N190F + I201F + A209V + Q264S (usando a numeração de *Bacillus licheniformis* em SEQ ID NO: 4 de WO99/19467). Em algumas modalidades, as variantes têm uma ou mais das seguintes mutações (ou mutações correspondentes em outras alfa-amilases de *Bacillus*): H154Y, A181T, N190F, A209V e Q264S e/ou a deleção de dois resíduos entre as posições 176 e 179, p.ex., deleção de E178 e G179 (usando SEQ ID NO: 5 de WO99/19467 para numeração das posições).

**[0171]** Em uma modalidade, a alfa-amilase bacteriana é a parte madura da alfa-amilase quimérica divulgada em Richardson *et al.* (2002), *The Journal of Biological Chemistry*, Vol. 277, No 29, Edição de 19 de julho, pp. 267501-26507, referida como BD5088 ou uma sua variante. Esta alfa-amilase é a mesma que aquela mostrada em SEQ ID NO: 2 em WO2007/134207. A sequência da enzima madura começa após o aminoácido “Met” inicial na posição 1.

**[0172]** A alfa-amilase pode ser uma alfa-amilase termoestável, tal como uma alfa-amilase bacteriana termoestável, p.ex., de *Bacillus stearothermophilus*. Em uma modalidade, a alfa-amilase usada em um processo descrito aqui tem um  $T_{1/2}$  (min) a pH 4,5, 85 °C,  $\text{CaCl}_2$  a 0,12 mM de pelo menos 10 determinado como descrito no Exemplo 1 de WO2018/098381.

**[0173]** Em uma modalidade, a alfa-amilase termoestável tem um  $T_{1/2}$  (min) a pH 4,5, 85 °C,  $\text{CaCl}_2$  a 0,12 mM, de pelo menos 15. Em uma modalidade, a alfa-amilase termoestável tem um  $T_{1/2}$  (min) a pH 4,5, 85 °C,  $\text{CaCl}_2$  a 0,12 mM, de pelo menos 20. Em uma modalidade, a alfa-amilase termoestável tem um  $T_{1/2}$  (min) a pH 4,5, 85 °C,  $\text{CaCl}_2$  a 0,12 mM, de pelo menos 25. Em uma modalidade, a alfa-amilase termoestável tem um  $T_{1/2}$  (min) a pH 4,5, 85 °C,  $\text{CaCl}_2$  a 0,12 mM, de pelo menos 30. Em uma modalidade, a alfa-amilase termoestável tem um  $T_{1/2}$  (min) a pH 4,5, 85 °C,  $\text{CaCl}_2$  a 0,12 mM, de pelo menos 40.

**[0174]** Em uma modalidade, a alfa-amilase termoestável tem um  $T_{1/2}$  (min) a pH 4,5, 85 °C,  $\text{CaCl}_2$  a 0,12 mM, de pelo menos 50. Em uma modalidade, a alfa-amilase termoestável tem um  $T_{1/2}$  (min) a pH 4,5, 85 °C,  $\text{CaCl}_2$  a 0,12 mM, de pelo menos 60. Em uma modalidade, a alfa-amilase termoestável tem um  $T_{1/2}$  (min) a pH 4,5, 85 °C,  $\text{CaCl}_2$  a 0,12 mM, entre 10-70. Em uma modalidade, a alfa-amilase termoestável tem um  $T_{1/2}$  (min) a pH 4,5, 85 °C,  $\text{CaCl}_2$  a 0,12 mM, entre 15-70. Em uma modalidade, a alfa-amilase termoestável tem um  $T_{1/2}$  (min) a pH 4,5, 85 °C,  $\text{CaCl}_2$  a 0,12 mM, entre 20-70. Em uma modalidade, a alfa-amilase termoestável tem um  $T_{1/2}$  (min) a pH 4,5, 85 °C,  $\text{CaCl}_2$  a 0,12 mM, entre 25-70. Em uma modalidade, a alfa-amilase termoestável tem um  $T_{1/2}$  (min) a pH 4,5, 85 °C,  $\text{CaCl}_2$  a 0,12 mM, entre 30-70. Em uma modalidade, a alfa-amilase termoestável tem um  $T_{1/2}$  (min) a pH 4,5, 85 °C,  $\text{CaCl}_2$  a 0,12 mM, entre 40-70. Em uma modalidade, a alfa-amilase termoestável tem um  $T_{1/2}$  (min) a pH 4,5, 85 °C,

CaCl<sub>2</sub> a 0,12 mM, entre 50-70. Em uma modalidade, a alfa-amilase termoestável tem um T<sub>1/2</sub> (min) a pH 4,5, 85 °C, CaCl<sub>2</sub> a 0,12 mM, entre 60-70.

**[0175]** Em uma modalidade, a alfa-amilase é uma alfa-amilase bacteriana, p.ex., derivada do gênero *Bacillus*, tal como uma estirpe de *Bacillus stearothermophilus*, p.ex., a *Bacillus stearothermophilus* como divulgada em WO99/019467 como SEQ ID NO: 3 com um ou dois aminoácidos deletados nas posições R179, G180, I181 e/ou G182, em particular com R179 e G180 deletados ou com I181 e G182 deletados, com mutações na lista de mutações em baixo.

**[0176]** Em algumas modalidades, as alfa-amilases de *Bacillus stearothermophilus* têm uma dupla deleção I181 + G182, e substituição opcional N193F, compreendendo adicionalmente uma das seguintes substituições ou combinações de substituições:

V59A+Q89R+G112D+E129V+K177L+R179E+K220P+N224L+Q254S  
;  
V59A+Q89R+E129V+K177L+R179E+H208Y+K220P+N224L+Q254S  
;  
V59A+Q89R+E129V+K177L+R179E+K220P+N224L+Q254S+D269E  
+D281N;  
V59A+Q89R+E129V+K177L+R179E+K220P+N224L+Q254S+I270L;  
V59A+Q89R+E129V+K177L+R179E+K220P+N224L+Q254S+H274K  
;  
V59A+Q89R+E129V+K177L+R179E+K220P+N224L+Q254S+Y276F;  
V59A+E129V+R157Y+K177L+R179E+K220P+N224L+S242Q+Q254  
S;  
V59A+E129V+K177L+R179E+H208Y+K220P+N224L+S242Q+Q254  
S;  
V59A+E129V+K177L+R179E+K220P+N224L+S242Q+Q254S;

V59A+E129V+K177L+R179E+K220P+N224L+S242Q+Q254S+H274  
K;  
V59A+E129V+K177L+R179E+K220P+N224L+S242Q+Q254S+Y276  
F;  
V59A+E129V+K177L+R179E+K220P+N224L+S242Q+Q254S+D281  
N;  
V59A+E129V+K177L+R179E+K220P+N224L+S242Q+Q254S+M284  
T;  
V59A+E129V+K177L+R179E+K220P+N224L+S242Q+Q254S+G416  
V;  
V59A+E129V+K177L+R179E+K220P+N224L+Q254S;  
V59A+E129V+K177L+R179E+K220P+N224L+Q254S+M284T;  
A91L+M96I+E129V+K177L+R179E+K220P+N224L+S242Q+Q254S;  
E129V+K177L+R179E;  
E129V+K177L+R179E+K220P+N224L+S242Q+Q254S;  
E129V+K177L+R179E+K220P+N224L+S242Q+Q254S+Y276F+L427  
M;  
E129V+K177L+R179E+K220P+N224L+S242Q+Q254S+M284T;  
E129V+K177L+R179E+K220P+N224L+S242Q+Q254S+N376\*+I377\*  
;  
E129V+K177L+R179E+K220P+N224L+Q254S;  
E129V+K177L+R179E+K220P+N224L+Q254S+M284T;  
E129V+K177L+R179E+S242Q;  
E129V+K177L+R179V+K220P+N224L+S242Q+Q254S;  
K220P+N224L+S242Q+Q254S;  
M284V;  
V59A+Q89R+ E129V+ K177L+ R179E+ Q254S+ M284V; e  
V59A+E129V+K177L+R179E+Q254S+ M284V;



**[0177]** Em uma modalidade, a alfa-amilase é selecionada do grupo de variantes de alfa-amilases de *Bacillus stearothermophilus* com dupla deleção I181\*+G182\*, e opcionalmente substituição N193F, e adicionalmente uma das seguintes substituições ou combinações de substituições:

E129V+K177L+R179E;

V59A+Q89R+E129V+K177L+R179E+H208Y+K220P+N224L+Q254S

;

V59A+Q89R+ E129V+ K177L+ R179E+ Q254S+ M284V;

V59A+E129V+K177L+R179E+Q254S+ M284V; e

E129V+K177L+R179E+K220P+N224L+S242Q+Q254S (usando SEQ ID NO: 1 aqui para numeração).

**[0178]** Deve ser entendido que quando se faz referência a alfa-amilases de *Bacillus stearothermophilus* e suas variantes elas são normalmente produzidas na forma truncada. Em particular, a truncação pode ser tal que a alfa-amilase de *Bacillus stearothermophilus* mostrada em SEQ ID NO: 3 em WO99/19467, ou suas variantes, são truncadas no terminal C e têm tipicamente de 480-495 aminoácidos em comprimento, tal como cerca de 491 aminoácidos em comprimento, p.ex., tal que não tenha um domínio funcional de ligação ao amido.

**[0179]** Em uma modalidade, a variante da alfa-amilase pode ser uma enzima tendo uma sequência de polipeptídeo maduro com um grau de identidade de pelo menos 60%, p.ex., pelo menos 70%, pelo menos 80%, pelo menos 90%, pelo menos 95%, pelo menos 91%, pelo menos 92%, pelo menos 93%, pelo menos 94%, pelo menos 95%, pelo menos 96%, pelo menos 97%, pelo menos 98% ou pelo menos 99%, mas menos do que 100%, com a sequência mostrada em SEQ ID NO: 3 em WO99/19467.

**[0180]** Em uma modalidade, a alfa-amilase bacteriana, p.ex., alfa-amilase de *Bacillus*, tal como especialmente alfa-amilase de

*Bacillus stearothermophilus*, ou sua variante, é doseada para a liquefação a uma concentração de 0,01-10 KNU-A/g de DS, p.ex., entre 0,02 e 5 KNU-A/g de DS, tal como 0,03 e 3 KNU-A, preferencialmente 0,04 e 2 KNU-A/g de DS, tal como especialmente 0,01 e 2 KNU-A/g de DS. Em uma modalidade, a alfa-amilase bacteriana, p.ex., alfa-amilase de *Bacillus*, tal como especialmente alfa-amilases de *Bacillus stearothermophilus*, ou sua variante, é doseada para a liquefação a uma concentração de 0,0001-1 mg de EP (Proteína de Enzima)/g de DS, p.ex., 0,0005-0,5 mg de EP/g de DS, tal como 0,001-0,1 mg de EP/g de DS.

**[0181]** Em uma modalidade, a alfa-amilase bacteriana é derivada da alfa-amilase de *Bacillus subtilis* de SEQ ID NO: 76, da alfa-amilase de *Bacillus subtilis* de SEQ ID NO: 82, da alfa-amilase de *Bacillus subtilis* de SEQ ID NO: 83, da alfa-amilase de *Bacillus subtilis* de SEQ ID NO: 84, da alfa-amilase de *Bacillus licheniformis* de SEQ ID NO: 85, da alfa-amilase de *Clostridium phytofermentans* de SEQ ID NO: 89, da alfa-amilase de *Clostridium phytofermentans* de SEQ ID NO: 90, da alfa-amilase de *Clostridium phytofermentans* de SEQ ID NO: 91, da alfa-amilase de *Clostridium phytofermentans* de SEQ ID NO: 92, da alfa-amilase de *Clostridium phytofermentans* de SEQ ID NO: 93, da alfa-amilase de *Clostridium phytofermentans* de SEQ ID NO: 94, da alfa-amilase de *Clostridium thermocellum* de SEQ ID NO: 95, da alfa-amilase de *Thermobifida fusca* de SEQ ID NO: 96, da alfa-amilase de *Thermobifida fusca* de SEQ ID NO: 97, de *Anaerocellum thermophilum* de SEQ ID NO: 98, de *Anaerocellum thermophilum* de SEQ ID NO: 99, de *Anaerocellum thermophilum* de SEQ ID NO: 100, de *Streptomyces avermitilis* de SEQ ID NO: 101 ou de *Streptomyces avermitilis* de SEQ ID NO: 88.

**[0182]** Em uma modalidade, a alfa-amilase é derivada de *Bacillus amyloliquefaciens*, tal como a alfa-amilase de *Bacillus*

*amyloliquefaciens* de SEQ ID NO: 231 (p.ex., como descrito em WO2018/002360, ou suas variantes, como descrito em WO2017/037614).

**[0183]** Em uma modalidade, a alfa-amilase é derivada de uma alfa-amilase de levedura, tal como a alfa-amilase de *Saccharomyopsis fibuligera* de SEQ ID NO: 77, a alfa-amilase de *Debaryomyces occidentalis* de SEQ ID NO: 78, a alfa-amilase de *Debaryomyces occidentalis* de SEQ ID NO: 79, a alfa-amilase de *Lipomyces kononenkoeae* de SEQ ID NO: 80, a alfa-amilase de *Lipomyces kononenkoeae* de SEQ ID NO: 81.

**[0184]** Em uma modalidade, a alfa-amilase é derivada de uma alfa-amilase fúngica filamentosa, tal como a alfa-amilase de *Aspergillus niger* de SEQ ID NO: 86, a alfa-amilase de *Aspergillus niger* de SEQ ID NO: 87.

**[0185]** Alfa-amilases adicionais que podem ser expressas com os organismos fermentadores e usadas com os métodos descritos aqui são descritas nos exemplos e incluem, mas não estão limitados a, alfa-amilases mostradas na Tabela 1 (ou seus derivados).

Tabela 1.

<b>Organismo Dador (domínio catalítico)</b>	<b>SEQ ID NO: (polipeptídeo maduro)</b>
<i>Rhizomucor pusillus</i>	121
<i>Bacillus licheniformis</i>	122
<i>Aspergillus niger</i>	123
<i>Aspergillus tamaritii</i>	124
<i>Acidomyces richmondensis</i>	125
<i>Aspergillus bombycis</i>	126
<i>Alternaria sp</i>	127
<i>Rhizopus microsporus</i>	128

<i>Syncephalastrum racemosum</i>	129
<i>Rhizomucor pusillus</i>	130
<i>Dichotomocladium hesseltinei</i>	131
<i>Lichtheimia ramosa</i>	132
<i>Penicillium aethiopicum</i>	133
<i>Subulispora sp</i>	134
<i>Trichoderma paraviridescens</i>	135
<i>Byssoascus striatosporus</i>	136
<i>Aspergillus brasiliensis</i>	137
<i>Penicillium subspinulosum</i>	138
<i>Penicillium antarcticum</i>	139
<i>Penicillium coprophilum</i>	140
<i>Penicillium olsonii</i>	141
<i>Penicillium vasconiae</i>	142
<i>Penicillium sp</i>	143
<i>Heterocephalum aurantiacum</i>	144
<i>Neosartorya massa</i>	145
<i>Penicillium janthinellum</i>	146
<i>Aspergillus brasiliensis</i>	147
<i>Aspergillus westerdijikiae</i>	148
<i>Hamigera avellanea</i>	149
<i>Hamigera avellanea</i>	150
<i>Meripilus giganteus</i>	151
<i>Cerrena unicolor</i>	152

<i>Physalacria cryptomeriae</i>	153
<i>Lenzites betulinus</i>	154
<i>Trametes ljubarskyi</i>	155
<i>Bacillus subtilis</i>	156
<i>Bacillus subtilis subsp. subtilis</i>	157
<i>Schwanniomyces occidentalis</i>	158
<i>Rhizomucor pusillus</i>	159
<i>Aspergillus niger</i>	160
<i>Bacillus stearothermophilus</i>	161
<i>Bacillus halmapalus</i>	162
<i>Aspergillus oryzae</i>	163
<i>Bacillus amyloliquefaciens</i>	164
<i>Rhizomucor pusillus</i>	165
<i>Kionochaeta ivoriensis</i>	166
<i>Aspergillus niger</i>	167
<i>Aspergillus oryzae</i>	168
<i>Penicillium canescens</i>	169
<i>Acidomyces acidothermus</i>	170
<i>Kinochaeta ivoriensis</i>	171
<i>Aspergillus terreus</i>	172
<i>Thamnidium elegans</i>	173
<i>Meripilus giganteus</i>	174

**[0186]** Alfa-amilases adicionais contempladas para uso com a presente invenção podem ser encontradas em WO2011/153516 (o conteúdo da qual é incorporado aqui).

**[0187]** Polinucleotídeos codificando alfa-amilases adequadas adicionais podem ser obtidos de microrganismos de qualquer gênero, incluindo aqueles prontamente disponíveis dentro da base de dados UniProtKB ([www.uniprot.org](http://www.uniprot.org)).

**[0188]** Como descrito *supra*, a alfa-amilase pode ser uma alfa-amilase bacteriana. Por exemplo, a alfa-amilase pode ser derivada de uma bactéria Gram-positiva tal como uma *Bacillus*, *Clostridium*, *Enterococcus*, *Geobacillus*, *Lactobacillus*, *Lactococcus*, *Oceanobacillus*, *Staphylococcus*, *Streptococcus*, ou *Streptomyces* ou uma bactéria Gram-negativa tal como uma *Campylobacter*, *E. coli*, *Flavobacterium*, *Fusobacterium*, *Helicobacter*, *Ilyobacter*, *Neisseria*, *Pseudomonas*, *Salmonella*, ou *Ureaplasma*.

**[0189]** Em uma modalidade, a alfa-amilase é derivada de *Bacillus alkalophilus*, *Bacillus amyloliquefaciens*, *Bacillus brevis*, *Bacillus circulans*, *Bacillus clausii*, *Bacillus coagulans*, *Bacillus firmus*, *Bacillus lautus*, *Bacillus lentus*, *Bacillus licheniformis*, *Bacillus megaterium*, *Bacillus pumilus*, *Bacillus stearothermophilus*, *Bacillus subtilis*, ou *Bacillus thuringiensis*.

**[0190]** Em outra modalidade, a alfa-amilase é derivada de *Streptococcus equisimilis*, *Streptococcus pyogenes*, *Streptococcus uberis*, ou *Streptococcus equi* subsp. *Zooepidemicus*.

**[0191]** Em outra modalidade, a alfa-amilase é derivada de *Streptomyces achromogenes*, *Streptomyces avermitilis*, *Streptomyces coelicolor*, *Streptomyces griseus*, ou *Streptomyces lividans*.

**[0192]** A alfa-amilase pode ser uma alfa-amilase fúngica. Por exemplo, a alfa-amilase pode ser derivada de uma levedura tal como uma *Candida*, *Kluyveromyces*, *Pichia*, *Saccharomyces*, *Schizosaccharomyces*, *Yarrowia* ou *Issatchenkia*; ou derivada de um fungo filamentoso tal como um *Acremonium*, *Agaricus*, *Alternaria*, *Aspergillus*, *Aureobasidium*, *Botryosphaeria*, *Ceriporiopsis*, *Chaetomidium*,

*Chrysosporium*, *Claviceps*, *Cochliobolus*, *Coprinopsis*, *Coptotermes*, *Corynascus*, *Cryphonectria*, *Cryptococcus*, *Diplodia*, *Exidia*, *Filibasidium*, *Fusarium*, *Gibberella*, *Holomastigotoides*, *Humicola*, *Irpex*, *Lentinula*, *Leptosphaeria*, *Magnaporthe*, *Melanocarpus*, *Meripilus*, *Mucor*, *Myceliophthora*, *Neocallimastix*, *Neurospora*, *Paecilomyces*, *Penicillium*, *Phanerochaete*, *Piromyces*, *Poitrasia*, *Pseudoplectania*, *Pseudotriconympha*, *Rhizomucor*, *Schizophyllum*, *Scytalidium*, *Talaromyces*, *Thermoascus*, *Thielavia*, *Tolypocladium*, *Trichoderma*, *Trichophaea*, *Verticillium*, *Volvariella*, ou *Xylaria*.

**[0193]** Em outra modalidade, a alfa-amilase é derivada de *Saccharomyces carlsbergensis*, *Saccharomyces cerevisiae*, *Saccharomyces diastaticus*, *Saccharomyces douglasii*, *Saccharomyces kluyveri*, *Saccharomyces norbensis*, ou *Saccharomyces oviformis*.

**[0194]** Em outra modalidade, a alfa-amilase é derivada de *Acremonium cellulolyticus*, *Aspergillus aculeatus*, *Aspergillus awamori*, *Aspergillus foetidus*, *Aspergillus fumigatus*, *Aspergillus japonicus*, *Aspergillus nidulans*, *Aspergillus niger*, *Aspergillus oryzae*, *Chrysosporium inops*, *Chrysosporium keratinophilum*, *Chrysosporium lucknowense*, *Chrysosporium merdarium*, *Chrysosporium pannicola*, *Chrysosporium queenslandicum*, *Chrysosporium tropicum*, *Chrysosporium zonatum*, *Fusarium bactridioides*, *Fusarium cerealis*, *Fusarium crookwellense*, *Fusarium culmorum*, *Fusarium graminearum*, *Fusarium graminum*, *Fusarium heterosporum*, *Fusarium negundi*, *Fusarium oxysporum*, *Fusarium reticulatum*, *Fusarium roseum*, *Fusarium sambucinum*, *Fusarium sarcochroum*, *Fusarium sporotrichioides*, *Fusarium sulphureum*, *Fusarium torulosum*, *Fusarium trichothecioides*, *Fusarium venenatum*, *Humicola grisea*, *Humicola insolens*, *Humicola lanuginosa*, *Irpex lacteus*, *Mucor miehei*, *Myceliophthora thermophila*, *Neurospora crassa*, *Penicillium funiculosum*, *Penicillium purpurogenum*, *Phanerochaete chrysosporium*, *Thielavia*

*achromatica*, *Thielavia albomyces*, *Thielavia albopilosa*, *Thielavia australeinsis*, *Thielavia fimeti*, *Thielavia microspora*, *Thielavia ovispora*, *Thielavia peruviana*, *Thielavia setosa*, *Thielavia spededonium*, *Thielavia subthermophila*, *Thielavia terrestris*, *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei*, ou *Trichoderma viride*.

**[0195]** Será entendido que, para as espécies acima mencionadas, a invenção engloba os estados tanto perfeitos como imperfeitos e outros equivalentes taxonômicos, p.ex., anamorfos, independentemente do nome da espécie pelo qual sejam conhecidos. Os peritos na técnica reconhecerão prontamente a identidade de equivalentes apropriados.

**[0196]** Estirpes destas espécies estão prontamente acessíveis ao público em um número de coleções de cultura, tais como a *American Type Culture Collection (ATCC)*, *Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ)*, *Centraalbureau Voor Schimmelcultures (CBS)* e *Agricultural Research Service Patent Culture Collection, Northern Regional Research Center (NRRL)*.

**[0197]** As sequências codificantes de alfa-amilase descritas ou referenciadas aqui, ou uma sua subsequência, bem como a alfa-amilase descrita ou referenciada aqui, ou um seu fragmento, podem ser usadas para conceber sondas de ácidos nucleicos para identificar e clonar DNA codificando uma alfa-amilase de estirpes de diferentes gêneros ou espécies de acordo com métodos bem conhecidos na técnica. Em particular, tais sondas podem ser usadas para hibridação com o DNA genômico ou cDNA de uma célula de interesse, seguindo procedimentos de transferência de Southern padrão, de modo a identificar e isolar o gene correspondente aí. Tais sondas podem ser consideravelmente mais curtas do que a sequência inteira, mas devem ter pelo menos 15, p.ex., pelo menos 25, pelo menos 35



ou pelo menos 70 nucleotídeos em comprimento. Preferencialmente, a sonda de ácido nucleico tem pelo menos 100 nucleotídeos em comprimento, por exemplo, pelo menos 200 nucleotídeos, pelo menos 300 nucleotídeos, pelo menos 400 nucleotídeos, pelo menos 500 nucleotídeos, pelo menos 600 nucleotídeos, pelo menos 700 nucleotídeos, pelo menos 800 nucleotídeos ou pelo menos 900 nucleotídeos em comprimento. Podem ser usadas sondas tanto de DNA como de RNA. As sondas são tipicamente marcadas para detectar o gene correspondente (por exemplo, com  $^{32}\text{P}$ ,  $^3\text{H}$ ,  $^{35}\text{S}$ , biotina ou avidina).

**[0198]** Uma biblioteca de DNA genômico ou cDNA preparada a partir de tais outras estirpes pode ser rastreada quanto a DNA que hibrida com as sondas descritas acima e codifica um parente. O DNA genômico ou outro de tais outras estirpes pode ser separado por eletroforese em gel de agarose ou de poliacrilamida ou outras técnicas de separação. O DNA das bibliotecas ou o DNA separado pode ser transferido para e imobilizado em nitrocelulose ou outro material transportador adequado. Para identificar um clone ou DNA que hibride com uma sequência codificante, ou uma sua subsequência, o material transportador é usado em uma transferência de Southern.

**[0199]** Em uma modalidade, a sonda de ácido nucleico é um polinucleotídeo, ou sua subsequência, que codifica a alfa-amilase de qualquer uma de SEQ ID NOs: 76-101, 121-174 e 231 ou um seu fragmento.

**[0200]** Para propósitos das sondas descritas acima, a hibridação indica que o polinucleotídeo hibrida com uma sonda de ácido nucleico marcada, ou a sua fita complementar de comprimento total, ou uma subsequência das anteriores; sob condições de muito baixa a muito elevada estringência. As moléculas com as quais a sonda de ácido nucleico hibrida sob estas condições podem ser detectadas usando, por exemplo, filme de

raios X. As condições de estringência e lavagem são definidas como descrito *supra*.

**[0201]** Em uma modalidade, a alfa-amilase é codificada por um polinucleotídeo que hibrida sob condições de pelo menos baixa estringência, p.ex., condições de média estringência, condições de média-elevada estringência, condições de elevada estringência ou condições de muito elevada estringência com a fita complementar de comprimento total da sequência codificante para qualquer uma das alfa-amilases descritas ou referenciadas aqui (p.ex., a sequência codificante que codifica qualquer uma de SEQ ID NOs: 76-101, 121-174 e 231). (Sambrook *et al.*, 1989, *Molecular Cloning, A Laboratory Manual*, 2ª edição, Cold Spring Harbor, Nova Iorque).

**[0202]** A alfa-amilase pode ser também identificado e obtido de outras fontes incluindo microrganismos isolados a partir da natureza (p.ex., solo, adubos, água, ensilagem, *etc.*) ou amostras de DNA obtidas diretamente de materiais naturais (p.ex., solo, adubos, água, ensilagem, *etc.*) usando as sondas acima mencionadas. Técnicas para isolamento de microrganismos e DNA diretamente a partir de *habitats* naturais são bem conhecidas na técnica. O polinucleotídeo codificando uma alfa-amilase pode ser depois derivado por rastreamento similar de uma biblioteca genômica ou de cDNA de outro microrganismo ou amostra de DNA misto.

**[0203]** Logo que um polinucleotídeo codificando uma alfa-amilase tenha sido detectado com uma sonda adequada como descrito aqui, a sequência pode ser isolada ou clonada por utilização de técnicas que são conhecidas dos peritos na técnica (ver, p.ex., Sambrook *et al.*, 1989, *Molecular Cloning, A Laboratory Manual*, 2ª edição, Cold Spring Harbor, Nova Iorque). Técnicas usadas para isolar ou clonar polinucleotídeos codificando alfa-amilases incluem isolamento a partir de DNA genômico, preparação a partir de cDNA ou uma sua combinação. A clonagem dos

polinucleotídeos a partir de tal DNA genômico pode ser efetuada, p.ex., por uso da reação em cadeia da polimerase (PCR) bem conhecida ou rastreamento de anticorpos de bibliotecas de expressão para detectar fragmentos de DNA clonados com características estruturais partilhadas. Ver, p.ex., Innis *et al.*, 1990, *PCR: A Guide to Methods and Application*, Academic Press, Nova Iorque. Podem ser usados outros procedimentos de amplificação de ácidos nucleicos tais como reação em cadeia da ligase (LCR), transcrição ativada por ligação (LAT) e amplificação baseada em sequências de nucleotídeos (NASBA).

**[0204]** Em uma modalidade, a alfa-amilase tem uma sequência de polipeptídeo maduro que compreende a ou consiste na sequência de aminoácidos de qualquer uma das alfa-amilases descritas ou referenciadas aqui (p.ex., qualquer uma de SEQ ID NOs: 76-101, 121-174 e 231). Em outra modalidade, a alfa-amilase tem uma sequência de polipeptídeo maduro que é um fragmento de qualquer uma das alfa-amilases descritas ou referenciadas aqui (p.ex., qualquer uma de SEQ ID NOs: 76-101, 121-174 e 231). Em uma modalidade, o número de resíduos de aminoácidos no fragmento é pelo menos 75%, p.ex., pelo menos 80%, 85%, 90% ou 95% do número de resíduos de aminoácidos em alfa-amilase de comprimento total referenciada (p.ex., qualquer uma de SEQ ID NOs: 76-101, 121-174 e 231). Em outras modalidades, a alfa-amilase pode compreender o domínio catalítico de qualquer alfa-amilase descrita ou referenciada aqui (p.ex., o domínio catalítico de qualquer uma de SEQ ID NOs: 76-101, 121-174 e 231).

**[0205]** A alfa-amilase pode ser uma variante de qualquer uma das alfa-amilases descritas *supra* (p.ex., qualquer uma de SEQ ID NOs: 76-101, 121-174 e 231). Em uma modalidade, a alfa-amilase tem uma sequência de polipeptídeo maduro de pelo menos 60%, p.ex., pelo menos 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, ou 100% de

identidade de sequências com qualquer uma das alfa-amilases descritas *supra* (p.ex., qualquer uma de SEQ ID NOs: 76-101, 121-174 e 231).

**[0206]** Em uma modalidade, a alfa-amilase tem uma sequência de polipeptídeo maduro que difere em não mais do que dez aminoácidos, p.ex., em não mais do que cinco aminoácidos, em não mais do que quatro aminoácidos, em não mais do que três aminoácidos, em não mais do que dois aminoácidos ou em um aminoácido da sequência de aminoácidos de qualquer uma das alfa-amilases descritas *supra* (p.ex., qualquer uma de SEQ ID NOs: 76-101, 121-174 e 231). Em uma modalidade, a alfa-amilase tem uma substituição, deleção e/ou inserção de aminoácidos de uma ou mais (p.ex., duas, várias) de sequências de aminoácidos de qualquer uma das alfa-amilases descritas *supra* (p.ex., qualquer uma de SEQ ID NOs: 76-101, 121-174 e 231). Em alguns aspectos, o número total de substituições, deleções e/ou inserções de aminoácidos é não mais do que 10, p.ex., não mais do que 9, 8, 7, 6, 5, 4, 3, 2 ou 1.

**[0207]** As mudanças de aminoácidos são geralmente de uma natureza mínima, isto é, são substituições ou inserções de aminoácidos conservativas que não afetam significativamente o dobramento e/ou a atividade da proteína; pequenas deleções, tipicamente de um a cerca de 30 aminoácidos; pequenas extensões amino-terminais ou carboxil-terminais, tais como um resíduo de metionina amino-terminal; um pequeno peptídeo ligante de até cerca de 20-25 resíduos; ou uma pequena extensão que facilita a purificação por mudança da carga líquida ou outra função, tal como um trato de poli-histidina, um epítipo antigênico ou um domínio de ligação.

**[0208]** Exemplos de substituições conservativas estão dentro dos grupos 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 aminoácidos que não alteram geralmente a atividade específica são conhecidas na técnica e são descritas, por exemplo, por H. Neurath e R.L. Hill, 1979, *Em, The Proteins*, Academic Press, Nova Iorque. As permutas ocorrendo mais comumente são Ala/Ser, Val/Ile, Asp/Glu, Thr/Ser, Ala/Gly, Ala/Thr, Ser/Asn, Ala/Val, Ser/Gly, Tyr/Phe, Ala/Pro, Lys/Arg, Asp/Asn, Leu/Ile, Leu/Val, Ala/Glu e Asp/Gly.

**[0209]** Alternativamente, as mudanças de aminoácidos são de uma tal natureza que as propriedades físico-químicas dos polipeptídeos são alteradas. Por exemplo, as mudanças de aminoácidos podem melhorar a estabilidade térmica da alfa-amilase, alterar a especificidade do substrato, mudar o pH ótimo e similares.

**[0210]** Os aminoácidos essenciais podem ser identificados de acordo com procedimentos conhecidos na técnica, tais como mutagênese sítio-dirigida ou mutagênese de varrimento da alanina (Cunningham e Wells, 1989, *Science* 244: 1081-1085). Em esta última técnica são introduzidas mutações únicas de alanina em todos os resíduos na molécula, e as moléculas mutantes resultantes são testadas quanto à atividade para identificar resíduos de aminoácidos que são críticos para a atividade da molécula. Ver também Hilton *et al.*, 1996, *J. Biol. Chem.* 271: 4699-4708. O local ativo ou outra interação biológica pode ser também determinado por análise física da estrutura, como determinado por tais técnicas como ressonância magnética nuclear, cristalografia, difração de elétrons ou marcação por fotoafinidade, em conjunção com mutação de aminoácidos de locais de contato putativos. 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 podem ser também inferidas a partir da análise de identidades com outras alfa-amilases que estão relacionadas com a alfa-amilase referenciada.

**[0211]** Orientação adicional sobre a relação estrutura-atividade dos polipeptídeos aqui pode ser determinada usando técnicas de alinhamento de múltiplas sequências (MSA) bem conhecidas na técnica. Com base nos ensinamentos aqui, o perito poderia fazer alinhamentos similares com qualquer número de alfa-amilases descritas aqui ou conhecidas na técnica. Tais alinhamentos ajudam o perito a determinar domínios potencialmente relevantes (p.ex., domínios de ligação ou domínios catalíticos), bem como quais os resíduos de aminoácidos que são conservados e não conservados entre as diferentes sequências de alfa-amilases. É reconhecido na técnica que a mudança de um aminoácido que está conservado em uma posição particular entre polipeptídeos divulgados irá mais provavelmente resultar em uma mudança na atividade biológica (Bowie *et al.*, 1990, *Science* 247: 1306-1310: “*Residues that are directly involved in protein functions such as binding or catalysis will certainly be among the most conserved*”). Em contraste, a substituição de um aminoácido que não é altamente conservado entre os polipeptídeos provavelmente não irá alterar ou alterar significativamente a atividade biológica.

**[0212]** Orientação ainda adicional sobre a relação estrutura-atividade para o perito pode ser encontrada em estudos de cristalografia de raios x publicados e conhecidos na técnica.

**[0213]** Podem ser feitas e testadas substituições, deleções e/ou inserções únicas ou múltiplas de aminoácidos usando métodos conhecidos de mutagênese, recombinação e/ou embaralhamento, seguidos por um procedimento de rastreamento 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; WO95/17413; ou WO95/22625. Outros métodos que podem ser usados incluem PCR propensa a erros, exibição em fagos (por ex., Lowman *et al.*, 1991, *Biochemistry* 30: 10832-10837; Patente dos E.U.A. No. 5,223,409;

WO92/06204) e mutagênese dirigida a regiões (Derbyshire *et al.*, 1986, *Gene* 46: 145; Ner *et al.*, 1988, *DNA* 7: 127).

**[0214]** Podem ser combinados métodos de mutagênese/embaralhamento com métodos de rastreamento automatizados, de elevada produtividade para se detectar a atividade de polipeptídeos mutagenizados, clonados expressos por células hospedeiras (Ness *et al.*, 1999, *Nature Biotechnology* 17: 893-896). As moléculas de DNA mutagenizadas que codificam alfa-amilases ativas podem ser recuperadas a partir das células hospedeiras e rapidamente sequenciadas usando métodos padrão na técnica. Estes métodos permitem a determinação rápida da importância de resíduos individuais de aminoácidos em um polipeptídeo.

**[0215]** Em algumas modalidades, a alfa-amilase tem pelo menos 20%, p.ex., pelo menos 40%, pelo menos 50%, pelo menos 60%, pelo menos 70%, pelo menos 80%, pelo menos 90%, pelo menos 95%, pelo menos 96%, pelo menos 97%, pelo menos 98%, pelo menos 99% ou 100% da atividade de alfa-amilase de qualquer alfa-amilase descrita ou referenciada aqui (p.ex., qualquer uma de SEQ ID NOs: 76-101, 121-174 e 231) sob as mesmas condições.

**[0216]** Em uma modalidade, a sequência codificante de alfa-amilase híbrida sob condições de pelo menos baixa estringência, p.ex., condições de média estringência, condições de média-elevada estringência, condições de elevada estringência ou condições de muito elevada estringência com a fita complementar de comprimento total da sequência codificante de qualquer alfa-amilase descrita ou referenciada aqui (p.ex., qualquer uma de SEQ ID NOs: 76-101, 121-174 e 231). Em uma modalidade, a sequência codificante de alfa-amilase tem pelo menos 65%, p.ex., pelo menos 70%, pelo menos 75%, pelo menos 80%, pelo menos 85%, pelo menos 85%, pelo menos 90%, pelo menos 91%, pelo menos 92%, pelo menos 93%, pelo menos 94%, pelo menos 95%, pelo menos 96%, pelo

menos 97%, pelo menos 98%, pelo menos 99% ou 100% de identidade de sequências com a sequência codificante de qualquer alfa-amilase descrita ou referenciada aqui (p.ex., qualquer uma de SEQ ID NOs: 76-101, 121-174 e 231).

**[0217]** Em uma modalidade, a alfa-amilase compreende a sequência codificante de qualquer alfa-amilase descrita ou referenciada aqui (qualquer uma de SEQ ID NOs: 76-101, 121-174 e 231). Em uma modalidade, a alfa-amilase compreende uma sequência codificante que é uma subsequência da sequência codificante de qualquer alfa-amilase descrita ou referenciada aqui, em que a subsequência codifica um polipeptídeo tendo atividade de alfa-amilase. Em uma modalidade, o número de resíduos de nucleotídeos na subsequência é pelo menos 75%, p.ex., pelo menos 80%, 85%, 90%, ou 95% do número da sequência codificante referenciada.

**[0218]** A sequência codificante referenciada de qualquer aspecto ou modalidade relacionado descrito aqui pode ser a sequência codificante nativa ou uma sequência degenerada, tal como uma sequência codificante com otimização de códons concebida para uso em uma célula hospedeira particular (p.ex., otimizada para expressão em *Saccharomyces cerevisiae*).

**[0219]** A alfa-amilase pode ser um polipeptídeo fundido ou polipeptídeo de fusão clivável ao qual é fundido outro polipeptídeo no terminal N ou no terminal C da alfa-amilase. Um polipeptídeo fundido pode ser produzido por fusão de um polinucleotídeo codificando outro polipeptídeo a um polinucleotídeo codificando a alfa-amilase. Técnicas para produção de polipeptídeos de fusão são conhecidas na técnica e incluem ligação das sequências codificantes codificando os polipeptídeos tal que fiquem em grelha e que a expressão do polipeptídeo fundido esteja sob controle do(s) mesmo(s) promotor(es) e terminador. As proteínas de fusão podem ser



também construídas usando tecnologia de inteína na qual são criadas fusões pós-translacionalmente (Cooper *et al.*, 1993, *EMBO J.* 12: 2575-2583; Dawson *et al.*, 1994, *Science* 266: 776-779).

### Trealases

**[0220]** A trealase expressa e/ou exógena pode ser qualquer trealase que seja adequada para os organismos fermentadores e/ou seus métodos de uso descritos aqui, tal como uma trealase ocorrendo naturalmente ou uma sua variante que retenha atividade de trealase. Qualquer trealase contemplada para expressão por um organismo fermentador descrito em baixo está também contemplada para aspectos da invenção envolvendo adição exógena de uma trealase (p.ex., adicionada antes da, durante a ou após liquefação e/ou sacarificação).

**[0221]** Em algumas modalidades, o organismo fermentador compreendendo um polinucleotídeo heterólogo codificando uma trealase tem um nível aumentado de atividade de trealase em comparação com as células hospedeiras sem o polinucleotídeo heterólogo codificando a trealase, quando cultivados sob as mesmas condições. Em algumas modalidades, o organismo fermentador tem um nível aumentado de atividade de trealase de pelo menos 5%, p.ex., pelo menos 10%, pelo menos 15%, pelo menos 20%, pelo menos 25%, pelo menos 50%, pelo menos 100%, pelo menos 150%, pelo menos 200%, pelo menos 300% ou pelo menos 500% em comparação com o organismo fermentador sem o polinucleotídeo heterólogo codificando a trealase, quando cultivados sob as mesmas condições.

**[0222]** As trealases que podem ser expressas com os organismos fermentadores e usadas com os métodos descritos aqui incluem, mas não estão limitadas a, trealases mostradas na Tabela 2 (ou seus derivados).

Tabela 2.

Organismo Dador	SEQ ID NO:
-----------------	------------

<b>(domínio catalítico)</b>	<b>(polipeptídeo maduro)</b>
<i>Chaetomium megalocarpum</i>	175
<i>Lecanicillium psalliotae</i>	176
<i>Doratomyces sp</i>	177
<i>Mucor moelleri</i>	178
<i>Phialophora cyclaminis</i>	179
<i>Thielavia arenaria</i>	180
<i>Thielavia antarctica</i>	181
<i>Chaetomium sp</i>	182
<i>Chaetomium nigricolor</i>	183
<i>Chaetomium jodhpurensense</i>	184
<i>Chaetomium piluliferum</i>	185
<i>Myceliophthora hinnulea</i>	186
<i>Chloridium virescens</i>	187
<i>Gelasinospora cratophora</i>	188
<i>Acidobacteriaceae bacterium</i>	189
<i>Acidobacterium capsulatum</i>	190
<i>Acidovorax wautersii</i>	191
<i>Xanthomonas arborícola</i>	192
<i>Kosakonia sacchari</i>	193
<i>Enterobacter sp</i>	194
<i>Saitozyma flava</i>	195
<i>Phaeotremella skinneri</i>	196
<i>Trichoderma asperellum</i>	197

<i>Corynascus sepedonium</i>	198
<i>Myceliophthora thermophila</i>	199
<i>Trichoderma reesei</i>	200
<i>Chaetomium virescens</i>	201
<i>Rhodothermus marinus</i>	202
<i>Myceliophthora sepedonium</i>	203
<i>Moelleriella libera</i>	204
<i>Acremonium dichromosporum</i>	205
<i>Fusarium sambucinum</i>	206
<i>Phoma sp</i>	207
<i>Lentinus similis</i>	208
<i>Diaporthe nobilis</i>	209
<i>Solicoccozyma terricola</i>	210
<i>Dioszegia cryoxerica</i>	211
<i>Talaromyces funiculosus</i>	212
<i>Hamigera avellanea</i>	213
<i>Talaromyces ruber</i>	214
<i>Trichoderma lixii</i>	215
<i>Aspergillus cervinus</i>	216
<i>Rasamsonia brevistipitata</i>	217
<i>Acremonium curvulum</i>	218
<i>Talaromyces piceae</i>	219
<i>Penicillium sp</i>	220
<i>Talaromyces aurantiacus</i>	221
<i>Talaromyces pinophilus</i>	222

<i>Talaromyces leycettanus</i>	223
<i>Talaromyces variabilis</i>	224
<i>Aspergillus niger</i>	225
<i>Trichoderma reesei</i>	226

**[0223]** Polinucleotídeos codificando trealases adequadas adicionais podem ser derivados de microrganismos de qualquer gênero adequado, incluindo aqueles prontamente disponíveis dentro da base de dados UniProtKB ([www.uniprot.org](http://www.uniprot.org)).

**[0224]** As sequências codificantes de trealase podem ser também usadas para conceber sondas de ácidos nucleicos para identificar e clonar DNA codificando trealases de estirpes de gêneros ou espécies diferentes, como descrito *supra*.

**[0225]** Os polinucleotídeos codificando trealases podem ser também identificados e obtidos de outras fontes incluindo microrganismos isolados a partir da natureza (p.ex., solo, adubos, água, *etc.*) ou amostras de DNA obtidas diretamente de materiais naturais (p.ex., solo, adubos, água, *etc.*) como descrito *supra*.

**[0226]** Técnicas usadas para isolar ou clonar polinucleotídeos codificando trealases são descritas *supra*.

**[0227]** Em uma modalidade, a trealase tem uma sequência de polipeptídeo maduro que compreende a ou consiste na sequência de aminoácidos de qualquer uma das trealases descritas ou referenciadas aqui (p.ex., qualquer uma de SEQ ID NOs: 175-226). Em outra modalidade, a trealase tem uma sequência de polipeptídeo maduro que é um fragmento de qualquer uma das trealases descritas ou referenciadas aqui (p.ex., qualquer uma de SEQ ID NOs: 175-226). Em uma modalidade, o número de resíduos de aminoácidos no fragmento é pelo menos 75%, p.ex., pelo menos 80%, 85%, 90% ou 95% do número de resíduos de aminoácidos em trealase de comprimento total referenciada (p.ex., qualquer uma de SEQ ID NOs: 175-

226). Em outras modalidades, a trealase pode compreender o domínio catalítico de qualquer trealase descrita ou referenciada aqui (p.ex., o domínio catalítico de qualquer uma de SEQ ID NOs: 175-226).

**[0228]** A trealase pode ser uma variante de qualquer uma das trealases descritas *supra* (p.ex., qualquer uma de SEQ ID NOs: 175-226). Em uma modalidade, a trealase tem uma sequência de polipeptídeo maduro de pelo menos 60%, p.ex., pelo menos 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, ou 100% de identidade de sequências com qualquer uma das trealases descritas *supra* (p.ex., qualquer uma de SEQ ID Nos: 175-226).

**[0229]** Em uma modalidade, a trealase tem uma sequência de polipeptídeo maduro que difere em não mais do que dez aminoácidos, p.ex., em não mais do que cinco aminoácidos, em não mais do que quatro aminoácidos, em não mais do que três aminoácidos, em não mais do que dois aminoácidos ou em um aminoácido da sequência de aminoácidos de qualquer uma das trealases descritas *supra* (p.ex., qualquer uma de SEQ ID NOs: 175-226). Em uma modalidade, a trealase tem uma substituição, deleção e/ou inserção de aminoácidos de uma ou mais (p.ex., duas, várias) de sequências de aminoácidos de qualquer uma das trealases descritas *supra* (p.ex., qualquer uma de SEQ ID NOs: 175-226). Em alguns aspectos, o número total de substituições, deleções e/ou inserções de aminoácidos é não mais do que 10, p.ex., não mais do que 9, 8, 7, 6, 5, 4, 3, 2 ou 1.

**[0230]** Em algumas modalidades, a trealase tem pelo menos 20%, p.ex., pelo menos 40%, pelo menos 50%, pelo menos 60%, pelo menos 70%, pelo menos 80%, pelo menos 90%, pelo menos 95%, pelo menos 96%, pelo menos 97%, pelo menos 98%, pelo menos 99% ou 100% da atividade de trealase de qualquer trealase descrita ou referenciada aqui (p.ex., qualquer uma de SEQ ID NOs: 175-226 sob as mesmas condições).

**[0231]** Em uma modalidade, a sequência codificante de trealase híbrida sob condições de pelo menos baixa estringência, p.ex., condições de média estringência, condições de média-elevada estringência, condições de elevada estringência ou condições de muito elevada estringência com a fita complementar de comprimento total da sequência codificante de qualquer trealase descrita ou referenciada aqui (p.ex., qualquer uma de SEQ ID NOs: 175-226). Em uma modalidade, a sequência codificante de trealase tem pelo menos 65%, p.ex., pelo menos 70%, pelo menos 75%, pelo menos 80%, pelo menos 85%, pelo menos 85%, pelo menos 90%, pelo menos 91%, pelo menos 92%, pelo menos 93%, pelo menos 94%, pelo menos 95%, pelo menos 96%, pelo menos 97%, pelo menos 98%, pelo menos 99% ou 100% de identidade de sequências com a sequência codificante de qualquer trealase descrita ou referenciada aqui (p.ex., qualquer uma de SEQ ID NOs: 175-226).

**[0232]** Em uma modalidade, a trealase compreende a sequência codificante de qualquer trealase descrita ou referenciada aqui (qualquer uma de SEQ ID NOs: 175-226). Em uma modalidade, a trealase compreende uma sequência codificante que é uma subsequência da sequência codificante de qualquer trealase descrita ou referenciada aqui, em que a subsequência codifica um polipeptídeo tendo atividade de trealase. Em uma modalidade, o número de resíduos de nucleotídeos na subsequência é pelo menos 75%, p.ex., pelo menos 80%, 85%, 90%, ou 95% do número da sequência codificante referenciada.

**[0233]** A sequência codificante referenciada de qualquer aspecto ou modalidade relacionado descrito aqui pode ser a sequência codificante nativa ou uma sequência degenerada, tal como uma sequência codificante com otimização de códons concebida para uso em uma célula hospedeira particular (p.ex., otimizada para expressão em *Saccharomyces cerevisiae*).

**[0234]** A trealase pode também incluir polipeptídeos fundidos ou polipeptídeos de fusão clivável, como descrito *supra*.

#### Glucoamilases

**[0235]** A glucoamilase expressa e/ou exógena pode ser qualquer glucoamilase que seja adequada para os organismos fermentadores e/ou seus métodos de uso descritos aqui, tal como uma glucoamilase ocorrendo naturalmente ou uma sua variante que retenha atividade de glucoamilase. Qualquer glucoamilase contemplada para expressão por um organismo fermentador descrito em baixo está também contemplada para aspectos da invenção envolvendo adição exógena de uma glucoamilase (p.ex., adicionada antes da, durante a ou após liquefação e/ou sacarificação).

**[0236]** Em algumas modalidades, o organismo fermentador compreende um polinucleotídeo heterólogo codificando uma glucoamilase, por exemplo, como descrito em WO2017/087330, o conteúdo da qual é deste modo incorporado por referência. Qualquer glucoamilase descrita ou referenciada aqui está contemplada para expressão no organismo fermentador.

**[0237]** Em algumas modalidades, o organismo fermentador compreendendo um polinucleotídeo heterólogo codificando uma glucoamilase tem um nível aumentado de atividade de glucoamilase em comparação com as células hospedeiras sem o polinucleotídeo heterólogo codificando a glucoamilase, quando cultivados sob as mesmas condições. Em algumas modalidades, o organismo fermentador tem um nível aumentado de atividade de glucoamilase de pelo menos 5%, p.ex., pelo menos 10%, pelo menos 15%, pelo menos 20%, pelo menos 25%, pelo menos 50%, pelo menos 100%, pelo menos 150%, pelo menos 200%, pelo menos 300% ou pelo menos 500% em comparação com o organismo

fermentador sem o polinucleotídeo heterólogo codificando a glucoamilase, quando cultivados sob as mesmas condições.

**[0238]** Glucoamilases exemplificativas que podem ser usadas com as células hospedeiras e/ou os métodos descritos aqui incluem glucoamilases bacterianas, de levedura ou fúngicas filamentosas, p.ex., obtidas de qualquer um dos microrganismos descritos ou referenciados aqui, como descrito *supra* sob as seções relacionadas com alfa-amilases.

**[0239]** A glucoamilase pode ser derivada de qualquer fonte adequada, p.ex., derivada de um microrganismo ou uma planta. As glucoamilases preferenciais têm origem fúngica ou bacteriana, selecionadas do grupo consistindo em glucoamilases de *Aspergillus*, em particular glucoamilases G1 ou G2 de *Aspergillus niger* (Boel *et al.* (1984), *EMBO J.* 3 (5), p. 1097-1102), ou suas variantes, tais como aquelas divulgadas em WO92/00381, WO00/04136 e WO01/04273 (a partir da Novozymes, Dinamarca); a glucoamilase de *A. awamori* divulgada em WO84/02921, glucoamilase de *Aspergillus oryzae* (*Agric. Biol. Chem.* (1991), 55 (4), p. 941-949) ou suas variantes ou fragmentos. Outras variantes de glucoamilases de *Aspergillus* incluem variantes com estabilidade térmica intensificada: G137A e G139A (Chen *et al.* (1996), *Prot. Eng.* 9, 499-505); D257E e D293E/Q (Chen *et al.* (1995), *Prot. Eng.* 8, 575-582); N182 (Chen *et al.* (1994), *Biochem. J.* 301, 275-281); ligações de dissulfeto, A246C (Fierobe *et al.* (1996), *Biochemistry*, 35, 8698-8704); e introdução de resíduos de Pro nas posições A435 e S436 (Li *et al.* (1997), *Protein Eng.* 10, 1199-1204).

**[0240]** Outras glucoamilases incluem glucoamilase de *Athelia rolfsii* (previamente denotada *Corticium rolfsii*) (ver patente dos EUA no. 4,727,026 e (Nagasaka *et al.* (1998) "Purification and properties of the raw-starch-degrading glucoamylases from *Corticium rolfsii*", *Appl Microbiol Biotechnol* 50: 323-330), glucoamilases de *Talaromyces*, em particular derivadas de *Talaromyces emersonii* (WO99/28448), *Talaromyces*



*leycettanus* (patente dos EUA no. Re. 32,153), *Talaromyces duponti*, *Talaromyces thermophilus* (patente dos EUA no. 4,587,215). Em uma modalidade, a glucoamilase usada durante a sacarificação e/ou fermentação é a glucoamilase de *Talaromyces emersonii* divulgada em WO99/28448.

**[0241]** Glucoamilases bacterianas contempladas incluem glucoamilases do gênero *Clostridium*, em particular *C. thermoamylolyticum* (EP 135,138) e *C. thermohydrosulfuricum* (WO86/01831).

**[0242]** Glucoamilases fúngicas contempladas incluem *Trametes cingulata*, *Pachykytospora papyracea*; e *Leucopaxillus giganteus*, todas divulgadas em WO2006/069289; e *Peniophora rufomarginata* divulgada em WO2007/124285; ou uma sua mistura. Estão também contempladas glucoamilases híbridas. Exemplos incluem as glucoamilases híbridas divulgadas em WO2005/045018.

**[0243]** Em uma modalidade, a glucoamilase é derivada de uma estirpe do gênero *Pycnoporus*, em particular uma estirpe de *Pycnoporus* como descrito em WO2011/066576 (SEQ ID NO: 2, 4 ou 6 aí), incluindo a glucoamilase de *Pycnoporus sanguineus*, ou de uma estirpe do gênero *Gloeophyllum*, tal como uma estirpe de *Gloeophyllum sepiarium* ou *Gloeophyllum trabeum*, em particular uma estirpe de *Gloeophyllum* como descrito em WO2011/068803 (SEQ ID NO: 2, 4, 6, 8, 10, 12, 14 ou 16 aí). Em uma modalidade, a glucoamilase é SEQ ID NO: 2 em WO2011/068803 (*i.e.*, glucoamilase de *Gloeophyllum sepiarium*). Em uma modalidade, a glucoamilase é a glucoamilase de *Gloeophyllum sepiarium* de SEQ ID NO: 8. Em uma modalidade, a glucoamilase é a glucoamilase de *Pycnoporus sanguineus* de SEQ ID NO: 229.

**[0244]** Em uma modalidade, a glucoamilase é uma glucoamilase de *Gloeophyllum trabeum* (divulgada como SEQ ID NO: 3 em WO2014/177546). Em outra modalidade, a glucoamilase é derivada de uma

estirpe do gênero *Nigrofomes*, em particular uma estirpe de *Nigrofomes sp.* divulgada em WO2012/064351 (divulgada como SEQ ID NO: 2 aí).

**[0245]** Estão também contempladas glucoamilases com uma sequência de polipeptídeo maduro que exibem uma elevada identidade com qualquer uma das glucoamilases acima mencionadas, *i.e.*, pelo menos 60%, tal como pelo menos 70%, pelo menos 75%, pelo menos 80%, pelo menos 85%, pelo menos 90%, pelo menos 95%, pelo menos 96%, pelo menos 97%, pelo menos 98%, pelo menos 99% ou mesmo 100% de identidade com qualquer uma das sequências de polipeptídeo maduro mencionadas acima.

**[0246]** As glucoamilases podem ser adicionadas à sacarificação e/ou fermentação em uma quantidade de 0,0001-20 AGU/g de DS, tal como 0,001-10 AGU/g de DS, 0,01-5 AGU/g de DS ou 0,1-2 AGU/g de DS.

**[0247]** As glucoamilases podem ser adicionadas à sacarificação e/ou fermentação em uma quantidade de 1-1.000 µg de EP/g de DS, tal como 10-500 µg de EP/g de DS ou 25-250 µg de EP/g de DS.

**[0248]** As glucoamilases podem ser adicionadas à liquefação em uma quantidade de 0,1-100 µg de EP/g de DS, tal como 0,5-50 µg de EP/g de DS, 1-25 µg de EP/g de DS ou 2-12 µg de EP/g de DS.

**[0249]** Em uma modalidade, a glucoamilase é adicionada como uma combinação compreendendo adicionalmente uma alfa-amilase (p.ex., qualquer alfa-amilase descrita aqui). Em uma modalidade, a alfa-amilase é uma alfa-amilase fúngica, especialmente uma alfa-amilase fúngica ácida. A alfa-amilase é tipicamente uma atividade secundária.

**[0250]** Em uma modalidade, a glucoamilase é uma combinação compreendendo glucoamilase de *Talaromyces emersonii* divulgada em WO99/28448 como SEQ ID NO: 34 e glucoamilase de *Trametes cingulata* divulgada como SEQ ID NO: 2 em WO06/069289.

**[0251]** Em uma modalidade, a glucoamilase é uma combinação compreendendo glucoamilase de *Talaromyces emersonii* divulgada em WO99/28448, glucoamilase de *Trametes cingulata* divulgada como SEQ ID NO: 2 em WO06/69289 e uma alfa-amilase.

**[0252]** Em uma modalidade, a glucoamilase é uma combinação compreendendo glucoamilase de *Talaromyces emersonii* divulgada em WO99/28448, glucoamilase de *Trametes cingulata* divulgada em WO06/69289 e alfa-amilase de *Rhizomucor pusillus* com ligante de glucoamilase de *Aspergillus niger* e SBD divulgado como V039 na Tabela 5 em WO2006/069290.

**[0253]** Em uma modalidade, a glucoamilase é uma combinação compreendendo glucoamilase de *Gloeophyllum sepiarium* mostrada como SEQ ID NO: 2 em WO2011/068803 e uma alfa-amilase, em particular alfa-amilase de *Rhizomucor pusillus* com um ligante de glucoamilase de *Aspergillus niger* e domínio de ligação ao amido (SBD), divulgada como SEQ ID NO: 3 em WO2013/006756, em particular com as seguintes substituições: G128D + D143N.

**[0254]** Em uma modalidade, a alfa-amilase pode ser derivada de uma estirpe do gênero *Rhizomucor*, preferencialmente uma estirpe de *Rhizomucor pusillus*, tal como aquela mostrada em SEQ ID NO: 3 em WO2013/006756, ou do gênero *Meripilus*, preferencialmente uma estirpe de *Meripilus giganteus*. Em uma modalidade, a alfa-amilase é derivada de uma estirpe de *Rhizomucor pusillus* com um ligante de glucoamilase de *Aspergillus niger* e domínio de ligação ao amido (SBD), divulgada como V039 na Tabela 5 em WO2006/069290.

**[0255]** Em uma modalidade, a alfa-amilase de *Rhizomucor pusillus* ou a alfa-amilase de *Rhizomucor pusillus* com um ligante de glucoamilase de *Aspergillus niger* e domínio de ligação ao amido (SBD) tem pelo menos uma das seguintes substituições ou combinações de

substituições: D165M; Y141W; Y141R; K136F; K192R; P224A; P224R; S123H+Y141W; G20S + Y141W; A76G + Y141W; G128D + Y141W; G128D + D143N; P219C + Y141W; N142D + D143N; Y141W + K192R; Y141W + D143N; Y141W + N383R; Y141W + P219C + A265C; Y141W + N142D + D143N; Y141W + K192R V410A; G128D + Y141W + D143N; Y141W + D143N + P219C; Y141W + D143N + K192R; G128D + D143N + K192R; Y141W + D143N + K192R + P219C; e G128D + Y141W + D143N + K192R; ou G128D + Y141W + D143N + K192R + P219C (usando SEQ ID NO: 3 em WO2013/006756 para numeração).

**[0256]** Em uma modalidade, a combinação de glucoamilases compreende glucoamilase de *Gloeophyllum sepiarium* (p.ex., SEQ ID NO: 2 em WO2011/068803) e uma alfa-amilase de *Rhizomucor pusillus*.

**[0257]** Em uma modalidade, a combinação de glucoamilases compreende glucoamilase de *Gloeophyllum sepiarium* mostrada como SEQ ID NO: 2 em WO2011/068803 e *Rhizomucor pusillus* com um ligante de glucoamilase de *Aspergillus niger* e domínio de ligação ao amido (SBD), divulgada como SEQ ID NO: 3 em WO2013/006756 com as seguintes substituições: G128D + D143N.

**[0258]** Composições compreendendo glucoamilase comercialmente disponíveis incluem AMG 200L; AMG 300 L; SAN™ SUPER, SAN™ EXTRA L, SPIRIZYME® PLUS, SPIRIZYME® FUEL, SPIRIZYME® B4U, SPIRIZYME® ULTRA, SPIRIZYME® EXCEL, SPIRIZYME ACHIEVE® e AMG® E (da Novozymes A/S); OPTIDEX™ 300, GC480, GC417 (da DuPont-Danisco); AMIGASE™ e AMIGASE™ PLUS (da DSM); G-ZYME™ G900, G-ZYME™ e G990 ZR (da DuPont-Danisco).

**[0259]** Em uma modalidade, a glucoamilase é derivada da glucoamilase de *Debaryomyces occidentalis* de SEQ ID NO: 102. Em uma modalidade, a glucoamilase é derivada da glucoamilase de

*Saccharomycopsis fibuligera* de SEQ ID NO: 103. Em uma modalidade, a glucoamilase é derivada da glucoamilase de *Saccharomycopsis fibuligera* de SEQ ID NO: 104. Em uma modalidade, a glucoamilase é derivada da glucoamilase de *Saccharomyces cerevisiae* de SEQ ID NO: 105. Em uma modalidade, a glucoamilase é derivada da glucoamilase de *Aspergillus niger* de SEQ ID NO: 106. Em uma modalidade, a glucoamilase é derivada da glucoamilase de *Aspergillus oryzae* de SEQ ID NO: 107. Em uma modalidade, a glucoamilase é derivada da glucoamilase de *Rhizopus oryzae* de SEQ ID NO: 108. Em uma modalidade, a glucoamilase é derivada da glucoamilase de *Clostridium thermocellum* de SEQ ID NO: 109. Em uma modalidade, a glucoamilase é derivada da glucoamilase de *Clostridium thermocellum* de SEQ ID NO: 110. Em uma modalidade, a glucoamilase é derivada da glucoamilase de *Arxula adenivorans* de SEQ ID NO: 111. Em uma modalidade, a glucoamilase é derivada da glucoamilase de *Hormoconis resinae* de SEQ ID NO: 112. Em uma modalidade, a glucoamilase é derivada da glucoamilase de *Aureobasidium pullulans* de SEQ ID NO: 113.

**[0260]** Em uma modalidade, a glucoamilase é uma glucoamilase de *Trichoderma reesei*, tal como a glucoamilase de *Trichoderma reesei* de SEQ ID NO: 230.

**[0261]** Em uma modalidade, a glucoamilase tem uma Atividade Relativa estabilidade térmica a 85 °C de pelo menos 20%, pelo menos 30% ou pelo menos 35% determinada como descrito no Exemplo 4 de WO2018/098381 (estabilidade térmica).

**[0262]** Em uma modalidade, a glucoamilase tem uma atividade relativa pH ótimo a pH 5,0 de pelo menos 90%, por exemplo, pelo menos 95%, pelo menos 97% ou 100% determinada como descrito no Exemplo 4 de WO2018/098381 (pH ótimo).

**[0263]** Em uma modalidade, a glucoamilase tem uma estabilidade ao pH a pH 5,0 de pelo menos 80%, pelo menos 85%, pelo menos 90% determinada como descrito no Exemplo 4 de WO2018/098381.

**[0264]** Em uma modalidade, a glucoamilase usada na liquefação, tal como uma variante de glucoamilase de *Penicillium oxalicum*, tem uma termoestabilidade determinada como Td por DSC a pH 4,0 como descrito no Exemplo 15 de WO2018/098381 de pelo menos 70 °C, preferencialmente pelo menos 75 °C, tal como pelo menos 80 °C, tal como pelo menos 81 °C, tal como pelo menos 82 °C, tal como pelo menos 83 °C, tal como pelo menos 84 °C, tal como pelo menos 85 °C, tal como pelo menos 86 °C, tal como pelo menos 87°C, tal como pelo menos 88 °C, tal como pelo menos 89 °C, tal como pelo menos 90 °C. Em uma modalidade, a glucoamilase, tal como uma variante de glucoamilase de *Penicillium oxalicum*, tem uma termoestabilidade determinada como Td por DSC a pH 4,0 como descrito no Exemplo 15 de WO2018/098381 na gama entre 70 °C e 95 °C, tal como entre 80 °C e 90 °C.

**[0265]** Em uma modalidade, a glucoamilase, tal como uma variante de glucoamilase de *Penicillium oxalicum*, usada na liquefação tem uma termoestabilidade determinada como Td por DSC a pH 4,8 como descrito no Exemplo 15 de WO2018/098381 de pelo menos 70 °C, preferencialmente pelo menos 75 °C, tal como pelo menos 80 °C, tal como pelo menos 81 °C, tal como pelo menos 82 °C, tal como pelo menos 83 °C, tal como pelo menos 84 °C, tal como pelo menos 85 °C, tal como pelo menos 86 °C, tal como pelo menos 87°C, tal como pelo menos 88 °C, tal como pelo menos 89 °C, tal como pelo menos 90 °C, tal como pelo menos 91 °C. Em uma modalidade, a glucoamilase, tal como uma variante de glucoamilase de *Penicillium oxalicum*, tem uma termoestabilidade determinada como Td por DSC a pH 4,8 como descrito no Exemplo 15 de WO2018/098381 na gama entre 70 °C e 95 °C, tal como entre 80 °C e 90 °C.

**[0266]** Em uma modalidade, a glucoamilase, tal como uma variante de glucoamilase de *Penicillium oxalicum*, usada na liquefação tem uma atividade residual determinada como descrito no Exemplo 16 de WO2018/098381, de pelo menos 100%, tal como pelo menos 105%, tal como pelo menos 110%, tal como pelo menos 115%, tal como pelo menos 120%, tal como pelo menos 125%. Em uma modalidade, a glucoamilase, tal como uma variante de glucoamilase de *Penicillium oxalicum*, tem uma termoestabilidade determinada como atividade residual como descrito no Exemplo 16 de WO2018/098381, na gama entre 100% e 130%.

**[0267]** Em uma modalidade, a glucoamilase, p.ex., de origem fúngica tal como um fungo filamentoso, de uma estirpe do gênero *Penicillium*, p.ex., uma estirpe de *Penicillium oxalicum*, em particular a glucoamilase de *Penicillium oxalicum* divulgada como SEQ ID NO: 2 em WO2006/127802 (que é deste modo incorporada por referência).

**[0268]** Em uma modalidade, a glucoamilase tem uma sequência de polipeptídeo maduro de pelo menos 80%, p.ex, pelo menos 85%, pelo menos 90%, pelo menos 91%, pelo menos 92%, pelo menos 93%, pelo menos 94%, pelo menos 95%, pelo menos 96%, pelo menos 97%, pelo menos 98%, pelo menos 99% ou 100% de identidade com o polipeptídeo maduro mostrado em SEQ ID NO: 2 em WO2011/127802.

Em uma modalidade, a glucoamilase é uma variante da glucoamilase de *Penicillium oxalicum* divulgada como SEQ ID NO: 2 em WO2011/127802, tendo uma substituição K79V. A variante de glucoamilase K79V tem uma sensibilidade reduzida à degradação por proteases em relação ao parente como divulgado em WO2013/036526 (que é deste modo incorporada por referência).

**[0269]** Em uma modalidade, a glucoamilase é derivada de *Penicillium oxalicum*.

**[0270]** Em uma modalidade, a glucoamilase é uma variante da glucoamilase de *Penicillium oxalicum* divulgada como SEQ ID NO: 2 em WO2011/127802. Em uma modalidade, a glucoamilase de *Penicillium oxalicum* é aquela divulgada como SEQ ID NO: 2 em WO2011/127802 tendo Val (V) na posição 79.

**[0271]** Variantes de glucoamilase de *Penicillium oxalicum* contempladas são divulgadas em WO2013/053801 que é deste modo incorporada por referência.

**[0272]** Em uma modalidade, estas variantes têm sensibilidade reduzida à degradação por proteases.

**[0273]** Em uma modalidade, estas variantes têm termoestabilidade melhorada em comparação com o parente.

**[0274]** Em uma modalidade, a glucoamilase tem uma substituição K79V (usando SEQ ID NO: 2 de WO2011/127802 para numeração), correspondendo à variante PE001 e, adicionalmente, compreende uma das seguintes alterações ou combinações de alterações.

T65A; Q327F; E501V; Y504T; Y504\*; T65A + Q327F; T65A + E501V; T65A + Y504T; T65A + Y504\*; Q327F + E501V; Q327F + Y504T; Q327F + Y504\*; E501V + Y504T; E501V + Y504\*; T65A + Q327F + E501V; T65A + Q327F + Y504T; T65A + E501V + Y504T; Q327F + E501V + Y504T; T65A + Q327F + Y504\*; T65A + E501V + Y504\*; Q327F + E501V + Y504\*; T65A + Q327F + E501V + Y504T; T65A + Q327F + E501V + Y504\*; E501V + Y504T; T65A + K161S; T65A + Q405T; T65A + Q327W; T65A + Q327F; T65A + Q327Y; P11F + T65A + Q327F; R1K + D3W + K5Q + G7V + N8S + T10K + P11S + T65A + Q327F; P2N + P4S + P11F + T65A + Q327F; P11F + D26C + K33C + T65A + Q327F; P2N + P4S + P11F + T65A + Q327W + E501V + Y504T; R1E + D3N + P4G + G6R + G7A + N8A + T10D + P11D + T65A + Q327F; P11F + T65A + Q327W; P2N + P4S + P11F + T65A + Q327F + E501V + Y504T; P11F + T65A + Q327W + E501V + Y504T; T65A + Q327F



+ E501V + Y504T; T65A + S105P + Q327W; T65A + S105P + Q327F; T65A + Q327W + S364P; T65A + Q327F + S364P; T65A + S103N + Q327F; P2N + P4S + P11F + K34Y + T65A + Q327F; P2N + P4S + P11F + T65A + Q327F + D445N + V447S; P2N + P4S + P11F + T65A + I172V + Q327F; P2N + P4S + P11F + T65A + Q327F + N502\*; P2N + P4S + P11F + T65A + Q327F + N502T + P563S + K571E; P2N + P4S + P11F + R31S + K33V + T65A + Q327F + N564D + K571S; P2N + P4S + P11F + T65A + Q327F + S377T; P2N + P4S + P11F + T65A + V325T + Q327W; P2N + P4S + P11F + T65A + Q327F + D445N + V447S + E501V + Y504T; P2N + P4S + P11F + T65A + I172V + Q327F + E501V + Y504T; P2N + P4S + P11F + T65A + Q327F + S377T + E501V + Y504T; P2N + P4S + P11F + D26N + K34Y + T65A + Q327F; P2N + P4S + P11F + T65A + Q327F + I375A + E501V + Y504T; P2N + P4S + P11F + T65A + K218A + K221D + Q327F + E501V + Y504T; P2N + P4S + P11F + T65A + S103N + Q327F + E501V + Y504T; P2N + P4S + T10D + T65A + Q327F + E501V + Y504T; P2N + P4S + F12Y + T65A + Q327F + E501V + Y504T; K5A + P11F + T65A + Q327F + E501V + Y504T; P2N + P4S + T10E + E18N + T65A + Q327F + E501V + Y504T; P2N + T10E + E18N + T65A + Q327F + E501V + Y504T; P2N + P4S + P11F + T65A + Q327F + E501V + Y504T + T568N; P2N + P4S + P11F + T65A + Q327F + E501V + Y504T + K524T + G526A; P2N + P4S + P11F + K34Y + T65A + Q327F + D445N + V447S + E501V + Y504T; P2N + P4S + P11F + R31S + K33V + T65A + Q327F + D445N + V447S + E501V + Y504T; P2N + P4S + P11F + D26N + K34Y + T65A + Q327F + E501V + Y504T; P2N + P4S + P11F + T65A + F80\* + Q327F + E501V + Y504T; P2N + P4S + P11F + T65A + K112S + Q327F + E501V + Y504T; P2N + P4S + P11F + T65A + Q327F + E501V + Y504T + T516P + K524T + G526A; P2N + P4S + P11F + T65A + Q327F + E501V + N502T + Y504\*; P2N + P4S + P11F + T65A + Q327F + E501V + Y504T; P2N + P4S + P11F + T65A + S103N + Q327F + E501V + Y504T; K5A + P11F + T65A + Q327F + E501V + Y504T; P2N + P4S + P11F + T65A

+ Q327F + E501V + Y504T + T516P + K524T + G526A; P2N + P4S + P11F + T65A + V79A + Q327F + E501V + Y504T; P2N + P4S + P11F + T65A + V79G + Q327F + E501V + Y504T; P2N + P4S + P11F + T65A + V79I + Q327F + E501V + Y504T; P2N + P4S + P11F + T65A + V79L + Q327F + E501V + Y504T; P2N + P4S + P11F + T65A + V79S + Q327F + E501V + Y504T; P2N + P4S + P11F + T65A + L72V + Q327F + E501V + Y504T; S255N + Q327F + E501V + Y504T; P2N + P4S + P11F + T65A + E74N + V79K + Q327F + E501V + Y504T; P2N + P4S + P11F + T65A + G220N + Q327F + E501V + Y504T; P2N + P4S + P11F + T65A + Y245N + Q327F + E501V + Y504T; P2N + P4S + P11F + T65A + Q253N + Q327F + E501V + Y504T; P2N + P4S + P11F + T65A + D279N + Q327F + E501V + Y504T; P2N + P4S + P11F + T65A + Q327F + S359N + E501V + Y504T; P2N + P4S + P11F + T65A + Q327F + D370N + E501V + Y504T; P2N + P4S + P11F + T65A + Q327F + V460S + E501V + Y504T; P2N + P4S + P11F + T65A + Q327F + V460T + P468T + E501V + Y504T; P2N + P4S + P11F + T65A + Q327F + T463N + E501V + Y504T; P2N + P4S + P11F + T65A + Q327F + S465N + E501V + Y504T; e P2N + P4S + P11F + T65A + Q327F + T477N + E501V + Y504T.

**[0275]** Em uma modalidade, a variante de glucoamilase de *Penicillium oxalicum* tem uma substituição K79V (usando SEQ ID NO: 2 de WO2011/127802 para numeração), correspondendo à variante PE001 e, adicionalmente, compreende uma das seguintes substituições ou combinações de substituições:

P11F + T65A + Q327F;

P2N + P4S + P11F + T65A + Q327F;

P11F + D26C + K33C + T65A + Q327F;

P2N + P4S + P11F + T65A + Q327W + E501V + Y504T;

P2N + P4S + P11F + T65A + Q327F + E501V + Y504T; e

P11F + T65A + Q327W + E501V + Y504T.

**[0276]** Glucoamilases adicionais contempladas para uso com a presente invenção podem ser encontradas em WO2011/153516 (o conteúdo da qual é incorporado).

**[0277]** Polinucleotídeos codificando glucoamilases adequadas adicionais podem ser obtidos de microrganismos de qualquer gênero, incluindo aqueles prontamente disponíveis dentro da base de dados UniProtKB ([www.uniprot.org](http://www.uniprot.org)).

**[0278]** As sequências codificantes de glucoamilase podem ser também usadas para conceber sondas de ácidos nucleicos para identificar e clonar DNA codificando de estirpes de gêneros ou espécies diferentes, como descrito *supra*.

**[0279]** Os polinucleotídeos codificando glucoamilases podem ser também identificados e obtidos de outras fontes incluindo microrganismos isolados a partir da natureza (p.ex., solo, adubos, água, *etc.*) ou amostras de DNA obtidas diretamente de materiais naturais (p.ex., solo, adubos, água, *etc.*) como descrito *supra*.

**[0280]** Técnicas usadas para isolar ou clonar polinucleotídeos codificando glucoamilases são descritas *supra*.

**[0281]** Em uma modalidade, a glucoamilase tem uma sequência de polipeptídeo maduro que compreende a ou consiste na sequência de aminoácidos de qualquer uma das glucoamilases descritas ou referenciadas aqui (p.ex., qualquer uma de SEQ ID NOs: 8, 102-113, 229 e 230). Em outra modalidade, a glucoamilase tem uma sequência de polipeptídeo maduro que é um fragmento de qualquer uma das glucoamilases descritas ou referenciadas aqui (p.ex., qualquer uma de SEQ ID NOs: 8, 102-113, 229 e 230). Em uma modalidade, o número de resíduos de aminoácidos no fragmento é pelo menos 75%, p.ex., pelo menos 80%, 85%, 90% ou 95% do número de resíduos de aminoácidos em glucoamilase de comprimento total referenciada (p.ex., qualquer uma de SEQ ID NOs: 8,

102-113, 229 e 230). Em outras modalidades, a glucoamilase pode compreender o domínio catalítico de qualquer glucoamilase descrita ou referenciada aqui (p.ex., o domínio catalítico de qualquer uma de SEQ ID NOs: 8, 102-113, 229 e 230).

**[0282]** A glucoamilase pode ser uma variante de qualquer uma das glucoamilases descritas *supra* (p.ex., qualquer uma de SEQ ID NOs: 8, 102-113, 229 e 230). Em uma modalidade, a glucoamilase tem uma sequência de polipeptídeo maduro de pelo menos 60%, p.ex., pelo menos 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, ou 100% de identidade de sequências com qualquer uma das glucoamilases descritas *supra* (p.ex., qualquer uma de SEQ ID Nos: 8, 102-113, 229 e 230).

**[0283]** Em uma modalidade, a glucoamilase tem uma sequência de polipeptídeo maduro que difere em não mais do que dez aminoácidos, p.ex., em não mais do que cinco aminoácidos, em não mais do que quatro aminoácidos, em não mais do que três aminoácidos, em não mais do que dois aminoácidos ou em um aminoácido da sequência de aminoácidos de qualquer uma das glucoamilases descritas *supra* (p.ex., qualquer uma de SEQ ID NOs: 8, 102-113, 229 e 230). Em uma modalidade, a glucoamilase tem uma substituição, deleção e/ou inserção de aminoácidos de uma ou mais (p.ex., duas, várias) de sequências de aminoácidos de qualquer uma das glucoamilases descritas *supra* (p.ex., qualquer uma de SEQ ID NOs: 8, 102-113, 229 e 230). Em alguns aspectos, o número total de substituições, deleções e/ou inserções de aminoácidos é não mais do que 10, p.ex., não mais do que 9, 8, 7, 6, 5, 4, 3, 2 ou 1.

**[0284]** Em algumas modalidades, a glucoamilase tem pelo menos 20%, p.ex., pelo menos 40%, pelo menos 50%, pelo menos 60%, pelo menos 70%, pelo menos 80%, pelo menos 90%, pelo menos 95%, pelo menos 96%, pelo menos 97%, pelo menos 98%, pelo menos 99% ou 100% da atividade de glucoamilase de qualquer glucoamilase descrita ou

referenciada aqui (p.ex., qualquer uma de SEQ ID NOs: 8, 102-113, 229 e 230) sob as mesmas condições.

**[0285]** Em uma modalidade, a sequência codificante de glucoamilase híbrida sob condições de pelo menos baixa estringência, p.ex., condições de média estringência, condições de média-elevada estringência, condições de elevada estringência ou condições de muito elevada estringência com a fita complementar de comprimento total da sequência codificante de qualquer glucoamilase descrita ou referenciada aqui (p.ex., qualquer uma de SEQ ID NOs: 8, 102-113, 229 e 230). Em uma modalidade, a sequência codificante de glucoamilase tem pelo menos 65%, p.ex., pelo menos 70%, pelo menos 75%, pelo menos 80%, pelo menos 85%, pelo menos 85%, pelo menos 90%, pelo menos 91%, pelo menos 92%, pelo menos 93%, pelo menos 94%, pelo menos 95%, pelo menos 96%, pelo menos 97%, pelo menos 98%, pelo menos 99% ou 100% de identidade de sequências com a sequência codificante de qualquer glucoamilase descrita ou referenciada aqui (p.ex., qualquer uma de SEQ ID NOs: 8, 102-113, 229 e 230).

**[0286]** Em uma modalidade, a glucoamilase compreende a sequência codificante de qualquer glucoamilase descrita ou referenciada aqui (qualquer uma de SEQ ID NOs: 8, 102-113, 229 e 230). Em uma modalidade, a glucoamilase compreende uma sequência codificante que é uma subsequência da sequência codificante de qualquer glucoamilase descrita ou referenciada aqui, em que a subsequência codifica um polipeptídeo tendo atividade de glucoamilase. Em uma modalidade, o número de resíduos de nucleotídeos na subsequência é pelo menos 75%, p.ex., pelo menos 80%, 85%, 90%, ou 95% do número da sequência codificante referenciada.

**[0287]** A sequência codificante referenciada de qualquer aspecto ou modalidade relacionado descrito aqui pode ser a sequência

codificante nativa ou uma sequência degenerada, tal como uma sequência codificante com otimização de códons concebida para uso em uma célula hospedeira particular (p.ex., otimizada para expressão em *Saccharomyces cerevisiae*).

**[0288]** A glucoamilase pode também incluir polipeptídeos fundidos ou polipeptídeos de fusão clivável, como descrito *supra*.

#### Proteases

**[0289]** A protease expressa e/ou exógena pode ser qualquer protease que seja adequada para os organismos fermentadores e/ou seus métodos de uso descritos aqui, tal como uma protease ocorrendo naturalmente ou uma sua variante que retenha atividade de protease. Qualquer protease contemplada para expressão por um organismo fermentador descrito em baixo está também contemplada para aspectos da invenção envolvendo adição exógena de uma protease.

**[0290]** As proteases são classificadas com base no seu mecanismo catalítico nos seguintes grupos: Serina proteases (S), Cisteína proteases (C), Proteases aspárticas (A), Metaloproteases (M) e Proteases desconhecidas, ou ainda não classificadas, (U), ver *Handbook of Proteolytic Enzymes*, A.J. Barrett, N.D. Rawlings, J.F. Woessner (eds), Academic Press (1998), em particular a parte da introdução geral.

**[0291]** A atividade de protease pode ser medida usando qualquer ensaio adequado, no qual um substrato é empregue, que inclua ligações de peptídeos relevantes para a especificidade da protease em questão. O pH do ensaio e a temperatura do ensaio são do mesmo modo para ser adaptados à protease em questão. Exemplos de valores de pH do ensaio são pH 6, 7, 8, 9, 10 ou 11. Exemplos de temperaturas do ensaio são 30, 35, 37, 40, 45, 50, 55, 60, 65, 70 ou 80 °C.

**[0292]** Em alguns aspectos, o organismo fermentador compreendendo um polinucleotídeo heterólogo codificando uma protease

tem um nível aumentado de atividade de protease em comparação com o organismo fermentador sem o polinucleotídeo heterólogo codificando a protease, quando cultivados sob as mesmas condições. Em alguns aspectos, o organismo fermentador tem um nível aumentado de atividade de protease de pelo menos 5%, p.ex., pelo menos 10%, pelo menos 15%, pelo menos 20%, pelo menos 25%, pelo menos 50%, pelo menos 100%, pelo menos 150%, pelo menos 200%, pelo menos 300%, ou pelo menos 500% em comparação com o organismo fermentador sem o polinucleotídeo heterólogo codificando a protease, quando cultivados nas mesmas condições.

**[0293]** Proteases exemplificativas que podem ser expressas com os organismos fermentadores e usadas com os métodos descritos aqui incluem, mas não estão limitadas a, proteases mostradas na Tabela 3 (ou seus derivados).

Tabela 3.

<b>Organismo Dador (domínio catalítico)</b>	<b>SEQ ID NO: (polipeptídeo maduro)</b>	<b>Família</b>
<i>Aspergillus niger</i>	9	A1
<i>Trichoderma reesei</i>	10	
<i>Thermoascus aurantiacus</i>	11	M35
<i>Dichomitus squalens</i>	12	S53
<i>Nocardioopsis prasina</i>	13	S1
<i>Penicillium simplicissimum</i>	14	S10
<i>Aspergillus niger</i>	15	

<i>Meriphilus giganteus</i>	16	S53
<i>Lecanicillium</i> sp. WMM742	17	S53
<i>Talaromyces proteolyticus</i>	18	S53
<i>Penicillium ranomafanaense</i>	19	A1A
<i>Aspergillus oryzae</i>	20	S53
<i>Talaromyces liani</i>	21	S10
<i>Thermoascus thermophilus</i>	22	S53
<i>Pyrococcus furiosus</i>	23	
<i>Trichoderma reesei</i>	24	
<i>Rhizomucor miehei</i>	25	
<i>Lenzites betulinus</i>	26	S53
<i>Neolentinus lepideus</i>	27	S53
<i>Thermococcus</i> sp.	28	S8
<i>Thermococcus</i> sp.	29	S8
<i>Thermomyces lanuginosus</i>	30	S53
<i>Thermococcus thio reducens</i>	31	S53
<i>Polyporus arcularius</i>	32	S53
<i>Ganoderma lucidum</i>	33	S53
<i>Ganoderma lucidum</i>	34	S53
<i>Ganoderma lucidum</i>	35	S53



<i>Trametes sp. AH28-2</i>	36	S53
<i>Cinereomyces lindbladii</i>	37	S53
<i>Trametes versicolor O82DDP</i>	38	S53
<i>Paecilomyces hepiali</i>	39	S53
<i>Isaria tenuipes</i>	40	S53
<i>Aspergillus tamaraii</i>	41	S53
<i>Aspergillus brasiliensis</i>	42	S53
<i>Aspergillus iizukae</i>	43	S53
<i>Penicillium sp-72364</i>	44	S10
<i>Aspergillus denticulatus</i>	45	S10
<i>Hamigera sp. t184-6</i>	46	S10
<i>Penicillium janthinellum</i>	47	S10
<i>Penicillium vasconiae</i>	48	S10
<i>Hamigera paravellanea</i>	49	S10
<i>Talaromyces variabilis</i>	50	S10
<i>Penicillium arenicola</i>	51	S10
<i>Nocardioopsis kunsanensis</i>	52	S1

<i>Streptomyces parvulus</i>	53	S1
<i>Saccharopolyspora endophytica</i>	54	S1
<i>luteus cellwall enrichments K</i>	55	S1
<i>Saccharothrix australiensis</i>	56	S1
<i>Nocardiosis baichengensis</i>	57	S1
<i>Streptomyces sp. SM15</i>	58	S1
<i>Actinoalloteichus spitiensis</i>	59	S1
<i>Byssochlamys verrucosa</i>	60	M35
<i>Hamigera terricola</i>	61	M35
<i>Aspergillus tamaris</i>	62	M35
<i>Aspergillus niveus</i>	63	M35
<i>Penicillium sclerotiorum</i>	64	A1
<i>Penicillium bilaiae</i>	65	A1
<i>Penicillium antarcticum</i>	66	A1
<i>Penicillium sumatrense</i>	67	A1
<i>Trichoderma lixii</i>	68	A1
<i>Trichoderma brevicompactum</i>	69	A1

<i>Penicillium cinnamopurpureum</i>	70	A1
<i>Bacillus licheniformis</i>	71	S8
<i>Bacillus subtilis</i>	72	S8
<i>Trametes cf versicol</i>	73	S53

**[0294]** Polinucleotídeos codificando proteases adequadas adicionais podem ser derivados de microrganismos de qualquer gênero adequado, incluindo aqueles prontamente disponíveis dentro da base de dados UniProtKB ([www.uniprot.org](http://www.uniprot.org)).

**[0295]** Em uma modalidade, a protease é derivada de *Aspergillus*, tal como a protease de *Aspergillus niger* de SEQ ID NO: 9, a protease de *Aspergillus tamaris* de SEQ ID NO: 41 ou a protease de *Aspergillus denticulatus* de SEQ ID NO: 45. Em uma modalidade, a protease é derivada de *Dichomitus*, tal como a protease de *Dichomitus squalens* de SEQ ID NO: 12. Em uma modalidade, a protease é derivada de *Penicillium*, tal como a protease de *Penicillium simplicissimum* de SEQ ID NO: 14, a protease de *Penicillium antarcticum* de SEQ ID NO: 66 ou a protease de *Penicillium sumatrense* de SEQ ID NO: 67. Em um aspecto, a protease é derivada de *Meriphilus*, tal como a protease de *Meriphilus giganteus* de SEQ ID NO: 16. Em um aspecto, a protease é derivada de *Talaromyces*, tal como a protease de *Talaromyces liani* de SEQ ID NO: 21. Em um aspecto, a protease é derivada de *Thermoascus*, tal como a protease de *Thermoascus thermophilus* de SEQ ID NO: 22. Em um aspecto, a protease é derivada de *Ganoderma*, tal como a protease de *Ganoderma lucidum* de SEQ ID NO: 33. Em um aspecto, a protease é derivada de *Hamigera*, tal como a protease de *Hamigera terricola* de SEQ ID NO: 61. Em um aspecto, a protease é derivada de *Trichoderma*, tal como a protease de *Trichoderma brevicompactum* de SEQ ID NO: 69.

**[0296]** As sequências codificando protease podem ser também usadas para conceber sondas de ácidos nucleicos para identificar e clonar DNA codificando proteases de estirpes de gêneros ou espécies diferentes, como descrito *supra*.

**[0297]** Os polinucleotídeos codificando proteases podem ser também identificados e obtidos de outras fontes incluindo microrganismos isolados a partir da natureza (p.ex., solo, adubos, água, *etc.*) ou amostras de DNA obtidas diretamente de materiais naturais (p.ex., solo, adubos, água, *etc.*) como descrito *supra*.

**[0298]** Técnicas usadas para isolar ou clonar polinucleotídeos codificando proteases são descritas *supra*.

**[0299]** Em uma modalidade, a protease tem uma sequência de polipeptídeo maduro que compreende a ou consiste na sequência de aminoácidos de qualquer uma de SEQ ID NOs: 9-73 (p.ex., qualquer uma de SEQ ID NOs: 9, 14, 16, 21, 22, 33, 41, 45, 61, 62, 66, 67 e 69; tal como qualquer uma de SEQ NOs: 9, 14, 16, e 69). Em outra modalidade, a protease tem uma sequência de polipeptídeo maduro que é um fragmento da protease de qualquer uma de SEQ ID NOs: 9-73 (p.ex., em que o fragmento tem atividade de protease). Em uma modalidade, o número de resíduos de aminoácidos no fragmento é pelo menos 75%, p.ex., pelo menos 80%, 85%, 90%, ou 95% do número de resíduos de aminoácidos em protease de comprimento total referenciada (p.ex. qualquer uma de SEQ ID NOs: 9-73). Em outras modalidades, a protease pode compreender o domínio catalítico de qualquer protease descrita ou referenciada aqui (p.ex., o domínio catalítico de qualquer uma de SEQ ID NOs: 9-73).

**[0300]** A protease pode ser uma variante de qualquer uma das proteases descritas *supra* (p.ex., qualquer uma de SEQ ID NOs: 9-73). Em uma modalidade, a protease tem uma sequência de polipeptídeo maduro de pelo menos 60%, p.ex., pelo menos 65%, 70%, 75%, 80%, 85%, 90%,

95%, 97%, 98%, 99%, ou 100% de identidade de sequências com qualquer uma das proteases descritas *supra* (p.ex., qualquer uma de SEQ ID NOs: 9-73).

**[0301]** Em uma modalidade, a protease tem uma sequência de polipeptídeo maduro que difere em não mais do que dez aminoácidos, p.ex., em não mais do que cinco aminoácidos, em não mais do que quatro aminoácidos, em não mais do que três aminoácidos, em não mais do que dois aminoácidos, ou em um aminoácido da sequência de aminoácidos de qualquer uma das proteases descritas *supra* (p.ex., qualquer uma de SEQ ID NOs: 9-73). Em uma modalidade, a protease tem uma substituição, deleção e/ou inserção de aminoácidos de uma ou mais (p.ex., duas, várias) de sequências de aminoácidos de qualquer uma das proteases descritas *supra* (p.ex., qualquer uma de SEQ ID NOs: 9-73). Em alguns aspectos, o número total de substituições, deleções e/ou inserções de aminoácidos é não mais do que 10, p.ex., não mais do que 9, 8, 7, 6, 5, 4, 3, 2 ou 1.

**[0302]** Em uma modalidade, a sequência codificante de protease híbrida sob condições de pelo menos baixa estringência, p.ex., condições de média estringência, condições de média-elevada estringência, condições de elevada estringência ou condições de muito elevada estringência com a fita complementar de comprimento total da sequência codificante de qualquer protease descrita ou referenciada aqui (p.ex., qualquer uma de SEQ ID NOs: 9-73). Em uma modalidade, a sequência codificante de protease tem pelo menos 65%, p.ex., pelo menos 70%, pelo menos 75%, pelo menos 80%, pelo menos 85%, pelo menos 85%, pelo menos 90%, pelo menos 91%, pelo menos 92%, pelo menos 93%, pelo menos 94%, pelo menos 95%, pelo menos 96%, pelo menos 97%, pelo menos 98%, pelo menos 99% ou 100% de identidade de sequências com a

sequência codificante de qualquer protease descrita ou referenciada aqui (p.ex., qualquer uma de SEQ ID NOs: 9-73).

**[0303]** Em uma modalidade, a protease compreende a sequência codificante de qualquer protease descrita ou referenciada aqui (qualquer uma de SEQ ID NOs: 9-73). Em uma modalidade, a protease compreende uma sequência codificante que é uma subsequência da sequência codificante de qualquer protease descrita ou referenciada aqui, em que a subsequência codifica um polipeptídeo tendo atividade de protease. Em uma modalidade, o número de resíduos de nucleotídeos na subsequência é pelo menos 75%, p.ex., pelo menos 80%, 85%, 90%, ou 95% do número da sequência codificante referenciada.

**[0304]** A sequência codificante referenciada de qualquer aspecto ou modalidade relacionado descrito aqui pode ser a sequência codificante nativa ou uma sequência degenerada, tal como uma sequência codificante com otimização de códons concebida para uso em uma célula hospedeira particular (p.ex., otimizada para expressão em *Saccharomyces cerevisiae*).

**[0305]** A protease pode também incluir polipeptídeos fundidos ou polipeptídeos de fusão clivável, como descrito *supra*.

**[0306]** Em uma modalidade, a protease usada de acordo com um processo descrito aqui é uma Serina protease. Em uma modalidade particular, a protease é uma serina protease pertencendo à família 53, p.ex., uma endo-protease, tal como protease S53 de *Meripilus giganteus*, *Dichomitus squalens*, *Trametes versicolor*, *Polyporus arcularius*, *Lenzites betulinus*, *Ganoderma lucidum*, *Neolentinus lepideus* ou *Bacillus sp.* 19138, em um processo para produção de etanol a partir de um material contendo amido, o rendimento de etanol foi melhorado, quando a protease S53 estava presente/ou foi adicionada durante a sacarificação e/ou a fermentação de amido gelatinizado ou não gelatinizado. Em uma modalidade, as proteases

são selecionadas de: (a) proteases pertencendo ao grupo de enzimas EC 3.4.21; e/ou (b) proteases pertencendo ao grupo de enzimas EC 3.4.14; e/ou (c) Serina proteases da família de peptidases S53 que compreende dois tipos diferentes de peptidases: tripeptidil aminopeptidases (tipo exo) e endopeptidases; como descrito em 1993, *Biochem. J.* 290: 205-218 e na base de dados de proteases MEROPS, lançamento, 9.4 (31 de janeiro de 2011) ([www.merops.ac.uk](http://www.merops.ac.uk)). A base de dados é descrita em Rawlings, N. D., Barrett, A. J. e Bateman, A., 2010, “*MEROPS: the peptidase database*”, *Nucl. Acids Res.* 38: D227-D233.

**[0307]** Para se determinar se uma dada protease é uma Serina protease, e uma protease da família S53, é feita referência ao *Handbook* acima e aos princípios indicados aí. Tal determinação pode ser levada a cabo para todos os tipos de proteases, sejam proteases ocorrendo naturalmente ou de tipo selvagem; ou proteases geneticamente manipuladas ou sintéticas.

**[0308]** A família de peptidases S53 contém endopeptidases que atuam em ácido e tripeptidil-peptidases. Os resíduos da tríade catalítica são Glu, Asp, Ser, e existe um resíduo ácido adicional, Asp, no espaço do oxiânion. A ordem dos resíduos é Glu, Asp, Asp, Ser. O resíduo Ser é o nucleófilo equivalente a Ser na tríade Asp, His, Ser da subtilisina, e o Glu da tríade é um substituto da base geral, His, na subtilisina.

**[0309]** As peptidases da família S53 tendem a ser o mais ativas a pH ácido (ao contrário das subtilisinas homólogas), e isto pode ser atribuído à importância funcional de resíduos carboxílicos, notavelmente Asp, no espaço de oxiânion. As sequências de aminoácidos não são intimamente similares às daquelas da família S8 (*i.e.*, subtilisinas serina endopeptidases e homólogos), e isto, tomado em conjunto com os resíduos de locais ativos bastante diferentes e o pH mais baixo resultante para atividade máxima, proporciona uma diferença substancial em relação a essa

família. O dobramento de proteínas da unidade de peptidase para membros desta família se assemelha àquele da subtilisina, tendo o tipo de clã SB.

**[0310]** Em uma modalidade, a protease usada de acordo com um processo descrito aqui é uma Cisteína protease.

**[0311]** Em uma modalidade, a protease usada de acordo com um processo descrito aqui é uma Protease aspártica. As ácido aspártico proteases são descritas, por exemplo, em *Handbook of Proteolytic Enzymes*, Editado por A.J. Barrett, N.D. Rawlings e J.F. Woessner, Academic Press, San Diego, 1998, Capítulo 270). Exemplos adequados de ácido aspártico protease incluem, p.ex., aquelas divulgadas em R.M. Berka *et al. Gene*, 96, 313 (1990); (R.M. Berka *et al. Gene*, 125, 195-198 (1993)); e Gomi *et al. Biosci. Biotech. Biochem.* 57, 1095-1100 (1993), que são deste modo incorporados por referência.

**[0312]** A protease pode ser também uma metaloprotease, que é definida como uma protease selecionada do grupo consistindo em:

(a) proteases pertencendo a EC 3.4.24 (metaloendopeptidases); preferencialmente EC 3.4.24.39 (metaloproteinases ácidas);

(b) metaloproteases pertencendo ao grupo M do *Handbook* acima;

(c) metaloproteases ainda não atribuídas a clãs (designação: Clã MX) ou pertencendo a qualquer um dos clãs MA, MB, MC, MD, ME, MF, MG, MH (como definido nas pp. 989-991 do *Handbook* acima);

(d) outras famílias de metaloproteases (como definido nas pp. 1448-1452 do *Handbook* acima);

(e) metaloproteases com um motivo HEXXH;

(f) metaloproteases com um motivo HEFTH;

(g) metaloproteases pertencendo a qualquer uma das famílias M3, M26, M27, M32, M34, M35, M36, M41, M43 ou M47 (como definido nas pp. 1448-1452 do *Handbook* acima);

(h) metaloproteases pertencendo à família M28E; e



(i) metaloproteases pertencendo à família M35 (como definido nas pp. 1492-1495 do *Handbook* acima).

**[0313]** Em outras modalidades particulares, as metaloproteases são hidrolases nas quais o ataque nucleofílico em uma ligação de peptídeos é mediado por uma molécula de água, que é ativada por um cátion metálico divalente. Exemplos de cátions divalentes são zinco, cobalto ou manganês. O íon metálico pode ser mantido no lugar por ligandos de aminoácidos. O número de ligandos pode ser cinco, quatro, três, dois, um ou zero. Em uma modalidade particular, o número é dois ou três, preferencialmente três.

**[0314]** Não existem limitações quanto à origem da metaloprotease usada em um processo da invenção. Em uma modalidade, a metaloprotease é classificada como EC 3.4.24, preferencialmente EC 3.4.24.39. Em uma modalidade, a metaloprotease é uma metaloprotease estável em ácido, p.ex., uma metaloprotease fúngica estável em ácido, tal como uma metaloprotease derivada de uma estirpe do gênero *Thermoascus*, preferencialmente uma estirpe de *Thermoascus aurantiacus*, especialmente *Thermoascus aurantiacus* CGMCC No. 0670 (classificada como EC 3.4.24.39). Em outra modalidade, a metaloprotease é derivada de uma estirpe do gênero *Aspergillus*, preferencialmente uma estirpe de *Aspergillus oryzae*.

**[0315]** Em uma modalidade, a metaloprotease tem um grau de identidade de sequências com os aminoácidos -178 a 177, -159 a 177 ou, preferencialmente, aminoácidos 1 a 177 (o polipeptídeo maduro) de SEQ ID NO: 1 de WO2010/008841 (uma metaloprotease de *Thermoascus aurantiacus*) de pelo menos 80%, pelo menos 82%, pelo menos 85%, pelo menos 90%, pelo menos 95% ou pelo menos 97%; e que tem atividade de metaloprotease. Em modalidades particulares, a metaloprotease consiste em

uma sequência de aminoácidos com um grau de identidade com SEQ ID NO: 1 como mencionado acima.

**[0316]** A metaloprotease de *Thermoascus aurantiacus* é um exemplo preferencial de uma metaloprotease adequada para uso em um processo da invenção. Outra metaloprotease é derivada de *Aspergillus oryzae* e compreende a sequência de SEQ ID NO: 11 divulgada em WO2003/048353 ou os seus aminoácidos -23-353; -23-374; -23-397; 1-353; 1-374; 1-397; 177-353; 177-374; ou 177-397, e SEQ ID NO: 10 divulgada em WO2003/048353.

**[0317]** Outra metaloprotease adequada para uso em um processo da invenção é a metaloprotease de *Aspergillus oryzae* compreendendo SEQ ID NO: 5 de WO2010/008841 ou uma metaloprotease é um polipeptídeo isolado que tem um grau de identidade com SEQ ID NO: 5 de pelo menos 80%, pelo menos 82%, pelo menos 85%, pelo menos 90%, pelo menos 95%, ou pelo menos 97%; e que tem atividade de metaloprotease. Em modalidades particulares, a metaloprotease consiste na sequência de aminoácidos de SEQ ID NO: 5 de WO2010/008841.

**[0318]** Em uma modalidade particular, uma metaloprotease tem uma sequência de aminoácidos que difere em quarenta, trinta e cinco, trinta, vinte e cinco, vinte ou em quinze aminoácidos dos aminoácidos -178 a 177, -159 a 177 ou +1 a 177 das sequências de aminoácidos da metaloprotease de *Thermoascus aurantiacus* ou *Aspergillus oryzae*.

**[0319]** Em outra modalidade, uma metaloprotease tem uma sequência de aminoácidos que difere em dez ou em nove ou em oito ou em sete ou em seis ou em cinco aminoácidos dos aminoácidos -178 a 177, -159 a 177 ou +1 a 177 das sequências de aminoácidos destas metaloproteases, p.ex., em quatro, em três, em dois ou em um aminoácido.

**[0320]** Em modalidades particulares, a metaloprotease a) compreende a ou b) consiste na

i) sequência de aminoácidos dos aminoácidos -178 a 177, -159 a 177 ou +1 a 177 de SEQ ID NO:1 de WO2010/008841;

ii) sequência de aminoácidos dos aminoácidos -23-353, -23-374, -23-397, 1-353, 1-374, 1-397, 177-353, 177-374 ou 177-397 de SEQ ID NO: 3 de WO2010/008841;

iii) sequência de aminoácidos de SEQ ID NO: 5 de WO2010/008841; ou

variantes alélicas, ou fragmentos, das sequências de i), ii), e iii) que têm atividade de protease.

**[0321]** Um fragmento dos aminoácidos -178 a 177, -159 a 177 ou +1 a 177 de SEQ ID NO: 1 de WO2010/008841 ou dos aminoácidos -23-353, -23-374, -23-397, 1-353, 1-374, 1-397, 177-353, 177-374 ou 177-397 de SEQ ID NO: 3 de WO2010/008841; é um polipeptídeo tendo um ou mais aminoácidos deletados do terminal amino e/ou carboxila destas sequências de aminoácidos. Em uma modalidade, um fragmento contém pelo menos 75 resíduos de aminoácidos ou pelo menos 100 resíduos de aminoácidos, pelo menos 125 resíduos de aminoácidos, pelo menos 150 resíduos de aminoácidos, pelo menos 160 resíduos de aminoácidos, pelo menos 165 resíduos de aminoácidos, pelo menos 170 resíduos de aminoácidos ou pelo menos 175 resíduos de aminoácidos.

**[0322]** Para se determinar se uma dada protease é uma metaloprotease ou não é feita referência ao “*Handbook of Proteolytic Enzymes*” acima e aos princípios indicados aí. Tal determinação pode ser levada a cabo para todos os tipos de proteases, sejam proteases ocorrendo naturalmente ou de tipo selvagem; ou proteases geneticamente manipuladas ou sintéticas.

**[0323]** A protease pode ser uma variante de, p.ex., uma protease de tipo selvagem, tendo propriedades de termoestabilidade definidas aqui. Em uma modalidade, a protease termoestável é uma variante de uma metaloprotease. Em uma modalidade, a protease termoestável usada em um processo descrito aqui tem origem fúngica, tal como uma metaloprotease fúngica, tal como uma metaloprotease fúngica derivada de uma estirpe do gênero *Thermoascus*, preferencialmente uma estirpe de *Thermoascus aurantiacus*, especialmente *Thermoascus aurantiacus* CGMCC No. 0670 (classificada como EC 3.4.24.39).

**[0324]** Em uma modalidade, a protease termoestável é uma variante da parte madura da metaloprotease mostrada em SEQ ID NO: 2 divulgada em WO2003/048353 ou a parte madura de SEQ ID NO: 1 em WO2010/008841 adicionalmente com uma das seguintes substituições ou combinações de substituições:

S5\*+D79L+S87P+A112P+D142L;  
D79L+S87P+A112P+T124V+D142L;  
S5\*+N26R+D79L+S87P+A112P+D142L;  
N26R+T46R+D79L+S87P+A112P+D142L;  
T46R+D79L+S87P+T116V+D142L;  
D79L+P81R+S87P+A112P+D142L;  
A27K+D79L+S87P+A112P+T124V+D142L;  
D79L+Y82F+S87P+A112P+T124V+D142L;  
D79L+Y82F+S87P+A112P+T124V+D142L;  
D79L+S87P+A112P+T124V+A126V+D142L;  
D79L+S87P+A112P+D142L;  
D79L+Y82F+S87P+A112P+D142L;  
S38T+D79L+S87P+A112P+A126V+D142L;  
D79L+Y82F+S87P+A112P+A126V+D142L;  
A27K+D79L+S87P+A112P+A126V+D142L;

D79L+S87P+N98C+A112P+G135C+D142L;  
 D79L+S87P+A112P+D142L+T141C+M161C;  
 S36P+D79L+S87P+A112P+D142L;  
 A37P+D79L+S87P+A112P+D142L;  
 S49P+D79L+S87P+A112P+D142L;  
 S50P+D79L+S87P+A112P+D142L;  
 D79L+S87P+D104P+A112P+D142L;  
 D79L+Y82F+S87G+A112P+D142L;  
 S70V+D79L+Y82F+S87G+Y97W+A112P+D142L;  
 D79L+Y82F+S87G+Y97W+D104P+A112P+D142L;  
 S70V+D79L+Y82F+S87G+A112P+D142L;  
 D79L+Y82F+S87G+D104P+A112P+D142L;  
 D79L+Y82F+S87G+A112P+A126V+D142L;  
 Y82F+S87G+S70V+D79L+D104P+A112P+D142L;  
 Y82F+S87G+D79L+D104P+A112P+A126V+D142L;  
 A27K+D79L+Y82F+S87G+D104P+A112P+A126V+D142L;  
 A27K+Y82F+S87G+D104P+A112P+A126V+D142L;  
 A27K+D79L+Y82F+ D104P+A112P+A126V+D142L;  
 A27K+Y82F+D104P+A112P+A126V+D142L;  
 A27K+D79L+S87P+A112P+D142L; e  
 D79L+S87P+D142L.

**[0325]** Em uma modalidade, a protease termoestável é uma variante da metaloprotease divulgada como a parte madura de SEQ ID NO: 2 divulgada em WO2003/048353 ou a parte madura de SEQ ID NO: 1 em WO2010/008841 com uma das seguintes substituições ou combinações de substituições:

D79L+S87P+A112P+D142L;  
 D79L+S87P+D142L; e  
 A27K+ D79L+Y82F+S87G+D104P+A112P+A126V+D142L.

**[0326]** Em uma modalidade, a variante de protease tem pelo menos 75% de identidade, preferencialmente pelo menos 80%, mais preferencialmente pelo menos 85%, mais preferencialmente pelo menos 90%, mais preferencialmente pelo menos 91%, mais preferencialmente pelo menos 92%, ainda mais preferencialmente pelo menos 93%, o mais preferencialmente pelo menos 94% e, ainda o mais preferencialmente, pelo menos 95%, tal como mesmo pelo menos 96%, pelo menos 97%, pelo menos 98%, pelo menos 99%, mas menos do que 100%, de identidade com a parte madura do polipeptídeo de SEQ ID NO: 2 divulgada em WO2003/048353 ou a parte madura de SEQ ID NO: 1 em WO2010/008841.

**[0327]** A protease termoestável pode ser também derivada de qualquer bactéria desde que a protease tenha as propriedades de termoestabilidade.

**[0328]** Em uma modalidade, a protease termoestável é derivada de uma estirpe da bactéria *Pyrococcus*, tal como de uma estirpe de *Pyrococcus furiosus* (protease pfu).

**[0329]** Em uma modalidade, a protease termoestável é uma mostrada como SEQ ID NO: 1 na patente dos EUA No. 6,358,726-B1 (Takara Shuzo Company).

**[0330]** Em uma modalidade, a protease termoestável é uma protease tendo uma sequência de polipeptídeo maduro de pelo menos 80% de identidade, tal como pelo menos 85%, tal como pelo menos 90%, tal como pelo menos 95%, tal como pelo menos 96%, tal como pelo menos 97%, tal como pelo menos 98%, tal como pelo menos 99% de identidade com SEQ ID NO: 1 na patente dos EUA no. 6,358,726-B1. A protease de *Pyrococcus furiosus* pode ser adquirida a partir da Takara Bio, Japão.

**[0331]** A protease de *Pyrococcus furiosus* pode ser uma protease termoestável como descrito em SEQ ID NO: 13 de WO2018/098381. Foi descoberto que esta protease (PfuS) tem uma

termoestabilidade de 110% (80 °C/70 °C) e 103% (90 °C/70 °C) determinada a pH 4,5.

**[0332]** Em uma modalidade, uma protease termoestável usada em um processo descrito aqui tem um valor de termoestabilidade de mais do que 20% determinada como Atividade Relativa a 80 °C/70 °C determinada como descrito no Exemplo 2 de WO2018/098381.

**[0333]** Em uma modalidade, a protease tem uma termoestabilidade de mais do que 30%, mais do que 40%, mais do que 50%, mais do que 60%, mais do que 70%, mais do que 80%, mais do que 90%, mais do que 100%, tal como mais do que 105%, tal como mais do que 110%, tal como mais do que 115%, tal como mais do que 120%, determinada como Atividade Relativa a 80 °C/70 °C.

**[0334]** Em uma modalidade, a protease tem uma termoestabilidade de entre 20 e 50%, tal como entre 20 e 40%, tal como 20 e 30%, determinada como Atividade Relativa a 80 °C/70 °C. Em uma modalidade, a protease tem uma termoestabilidade entre 50 e 115%, tal como entre 50 e 70%, tal como entre 50 e 60%, tal como entre 100 e 120%, tal como entre 105 e 115%, determinada como Atividade Relativa a 80 °C/70 °C.

**[0335]** Em uma modalidade, a protease tem um valor de termoestabilidade de mais do que 10%, determinada como Atividade Relativa a 85 °C/70 °C determinada como descrito no Exemplo 2 de WO2018/098381.

**[0336]** Em uma modalidade, a protease tem uma termoestabilidade de mais do que 10%, tal como mais do que 12%, mais do que 14%, mais do que 16%, mais do que 18%, mais do que 20%, mais do que 30%, mais do que 40%, mais do que 50%, mais do que 60%, mais do que 70%, mais do que 80%, mais do que 90%, mais do que 100%, mais do que 110%, determinada como Atividade Relativa a 85 °C/70 °C.

**[0337]** Em uma modalidade, a protease tem uma termoestabilidade entre 10 e 50%, tal como entre 10% e 30%, tal como entre 10% e 25%, determinada como Atividade Relativa a 85 °C/70 °C.

**[0338]** Em uma modalidade, a protease tem mais do que 20%, mais do que 30%, mais do que 40%, mais do que 50%, mais do que 60%, mais do que 70%, mais do que 80%, mais do que 90% determinada como Atividade Restante a 80 °C; e/ou a protease tem mais do que 20%, mais do que 30%, mais do que 40%, mais do que 50%, mais do que 60%, mais do que 70%, mais do que 80%, mais do que 90% determinada como Atividade Restante a 84 °C.

**[0339]** A determinação da “Atividade Relativa” e “Atividade Restante” é feita como descrito no Exemplo 2 de WO2017/098381.

**[0340]** Em uma modalidade, a protease pode ter uma termoestabilidade acima de 90, tal como acima de 100 a 85 °C como determinada usando o teste de Zeína-BCA como divulgado no Exemplo 3 de WO2018/098381.

**[0341]** Em uma modalidade, a protease tem uma termoestabilidade acima de 60%, tal como acima de 90%, tal como acima de 100%, tal como acima de 110% a 85 °C como determinada usando o ensaio de Zeína-BCA de WO2018/098381.

**[0342]** Em uma modalidade, a protease tem uma termoestabilidade entre 60-120, tal como entre 70-120%, tal como entre 80-120%, tal como entre 90-120%, tal como entre 100-120%, tal como 110-120% a 85 °C como determinada usando o ensaio de Zeína-BCA de WO2018/098381.

**[0343]** Em uma modalidade, a protease termoestável tem pelo menos 20%, tal como pelo menos 30%, tal como pelo menos 40%, tal como pelo menos 50%, tal como pelo menos 60%, tal como pelo menos 70%, tal como pelo menos 80%, tal como pelo menos 90%, tal como pelo menos



95%, tal como pelo menos 100%, da atividade da variante da protease JTP196 ou Protease Pfu determinada pelo teste da AZCL-caseína de WO2018/098381, e descrita aqui.

**[0344]** Em uma modalidade, a protease termoestável tem pelo menos 20%, tal como pelo menos 30%, tal como pelo menos 40%, tal como pelo menos 50%, tal como pelo menos 60%, tal como pelo menos 70%, tal como pelo menos 80%, tal como pelo menos 90%, tal como pelo menos 95%, tal como pelo menos 100%, da atividade de protease da variante da Protease 196 ou Protease Pfu determinada pelo teste da AZCL-caseína de WO2018/098381, e descrita aqui.

#### Pululanases

**[0345]** Em algumas modalidades, uma pululanase está presente e/ou é adicionada no passo de liquefação e/ou passo de sacarificação, ou sacarificação e fermentação simultâneas (SSF).

**[0346]** As pululanases (E.C. 3.2.1.41, pululana 6-glucanoidrolase), são enzimas desramificadoras caracterizadas pela sua capacidade de hidrolisarem as ligações alfa-1,6-glicosídicas, por exemplo, na amilopectina e pululana.

**[0347]** Em algumas modalidades, o organismo fermentador compreende um polinucleotídeo heterólogo codificando uma pululanase. Qualquer pululanase descrita ou referenciada aqui está contemplada para expressão no organismo fermentador.

**[0348]** A pululanase pode ser qualquer pululanase que seja adequada para as células hospedeiras e/ou os métodos descritos aqui, tal como uma pululanase ocorrendo naturalmente ou uma sua variante que retenha atividade de pululanase.

**[0349]** Em algumas modalidades, o organismo fermentador compreendendo um polinucleotídeo heterólogo codificando uma pululanase tem um nível aumentado de atividade de pululanase em

comparação com as células hospedeiras sem o polinucleotídeo heterólogo codificando a pululanase, quando cultivados sob as mesmas condições. Em algumas modalidades, o organismo fermentador tem um nível aumentado de atividade de pululanase de pelo menos 5%, p.ex., pelo menos 10%, pelo menos 15%, pelo menos 20%, pelo menos 25%, pelo menos 50%, pelo menos 100%, pelo menos 150%, pelo menos 200%, pelo menos 300% ou pelo menos 500% em comparação com o organismo fermentador sem o polinucleotídeo heterólogo codificando a pululanase, quando cultivados sob as mesmas condições.

**[0350]** Pululanases exemplificativas que podem ser usadas com as células hospedeiras e/ou os métodos descritos aqui incluem pululanases bacterianas, de levedura ou fúngicas filamentosas, p.ex., obtidas de qualquer um dos microrganismos descritos ou referenciados aqui, como descrito *supra* sob as seções relacionadas com alfa-amilases.

**[0351]** As pululanases contempladas incluem as pululanases de *Bacillus amyloclavus* divulgadas na Patente dos E.U.A. 4,560,651 (deste modo incorporada por referência), a pululanase divulgada como SEQ ID NO: 2 em WO01/151620 (deste modo incorporada por referência), a *Bacillus deramificans* divulgada como SEQ ID NO: 4 em WO01/151620 (deste modo incorporada por referência) e a pululanase de *Bacillus acidopullulyticus* divulgada como SEQ ID NO: 6 em WO01/151620 (deste modo incorporada por referência) e também descrita em *FEMS Mic. Let.* (1994) 115, 97-106.

**[0352]** Pululanases contempladas adicionais incluem as pululanases de *Pyrococcus woesei*, especificamente de *Pyrococcus woesei* DSM No. 3773 divulgadas em WO92/02614.

**[0353]** Em uma modalidade, a pululanase é uma pululanase da família GH57. Em uma modalidade, a pululanase inclui um domínio X47 como divulgado em US 61/289,040 publicada como

WO2011/087836 (que aqui deste modo incorporadas por referência). Mais especificamente, a pululanase pode ser derivada de uma estirpe do gênero *Thermococcus*, incluindo *Thermococcus litoralis* e *Thermococcus hydrothermalis*, tal como a pululanase de *Thermococcus hydrothermalis* truncada no local X4 imediatamente após o domínio X47 (*i.e.*, aminoácidos 1-782). A pululanase pode ser também um híbrido das pululanases de *Thermococcus litoralis* e *Thermococcus hydrothermalis* ou uma enzima híbrida de *T. hydrothermalis/T. litoralis* com o local de truncação X4 divulgada em US 61/289,040 publicada como WO2011/087836 (que é deste modo incorporada por referência).

**[0354]** Em outra modalidade, a pululanase é uma compreendendo um domínio X46 divulgado em WO2011/076123 (Novozymes).

**[0355]** A pululanase pode ser adicionada em uma quantidade eficaz que inclui a quantidade preferencial de cerca de 0,0001-10 mg de proteína de enzima por grama de DS, preferencialmente 0,0001-0,10 mg de proteína de enzima por grama de DS, mais preferencialmente 0,0001-0,010 mg de proteína de enzima por grama de DS. A atividade de pululanase pode ser determinada como NPUN. Um Ensaio para determinação de NPUN é descrito em WO2018/098381.

**[0356]** Produtos de pululanase comercialmente disponíveis adequados incluem PROMOZYME D, PROMOZYME™ D2 (Novozymes A/S, Dinamarca), OPTIMAX L-300 (DuPont-Danisco, EUA) e AMANO 8 (Amano, Japão).

**[0357]** Em uma modalidade, a pululanase é derivada da pululanase de *Bacillus subtilis* de SEQ ID NO: 114. Em uma modalidade, a pululanase é derivada da pululanase de *Bacillus licheniformis* de SEQ ID NO: 115. Em uma modalidade, a pululanase é derivada da pululanase de *Oryza sativa* de SEQ ID NO: 116. Em uma modalidade, a pululanase é derivada da

pululanase de *Triticum aestivum* de SEQ ID NO: 117. Em uma modalidade, a pululanase é derivada da pululanase de *Clostridium phytofermentans* de SEQ ID NO: 118. Em uma modalidade, a pululanase é derivada da pululanase de *Streptomyces avermitilis* de SEQ ID NO: 119. Em uma modalidade, a pululanase é derivada da pululanase de *Klebsiella pneumoniae* de SEQ ID NO: 120.

**[0358]** Pululanases adicionais contempladas para uso com a presente invenção podem ser encontradas em WO2011/153516 (o conteúdo da qual é incorporado aqui).

**[0359]** Polinucleotídeos codificando pululanases adequadas adicionais podem ser obtidos de microrganismos de qualquer gênero, incluindo aqueles prontamente disponíveis dentro da base de dados UniProtKB ([www.uniprot.org](http://www.uniprot.org)).

**[0360]** As sequências codificantes de pululanase podem ser também usadas para conceber sondas de ácidos nucleicos para identificar e clonar DNA codificando pululanases de estirpes de diferentes gêneros ou espécies, como descrito *supra*.

**[0361]** Os polinucleotídeos codificando pululanases podem ser também identificados e obtidos de outras fontes incluindo microrganismos isolados a partir da natureza (p.ex., solo, adubos, água, *etc.*) ou amostras de DNA obtidas diretamente de materiais naturais (p.ex., solo, adubos, água, *etc.*) como descrito *supra*.

**[0362]** Técnicas usadas para isolar ou clonar polinucleotídeos codificando pululanases são descritas *supra*.

**[0363]** Em uma modalidade, a pululanase tem uma sequência de polipeptídeo maduro que compreende a ou consiste na sequência de aminoácidos de qualquer uma das pululanases descritas ou referenciadas aqui (p.ex., qualquer uma de SEQ ID NOs: 114-120). Em outra modalidade, a pululanase tem uma sequência de polipeptídeo maduro que é

um fragmento de qualquer uma das pululanases descritas ou referenciadas aqui (p.ex., qualquer uma de SEQ ID NOs: 114-120). Em uma modalidade, o número de resíduos de aminoácidos no fragmento é pelo menos 75%, p.ex., pelo menos 80%, 85%, 90% ou 95% do número de resíduos na pululanase de comprimento total referenciada. Em outras modalidades, a pululanase pode compreender o domínio catalítico de qualquer pululanase descrita ou referenciada aqui (p.ex., qualquer uma de SEQ ID NOs: 114-120).

**[0364]** A pululanase pode ser uma variante de qualquer uma das pululanases descritas *supra* (p.ex., qualquer uma de SEQ ID NOs: 114-120). Em uma modalidade, a pululanase tem uma sequência de polipeptídeo maduro de pelo menos 60%, p.ex., pelo menos 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, ou 100% de identidade de sequências com qualquer uma das pululanases descritas *supra* (p.ex., qualquer uma de SEQ ID NOs: 114-120).

**[0365]** Em uma modalidade, a pululanase tem uma sequência de polipeptídeo maduro que difere em não mais do que dez aminoácidos, p.ex., em não mais do que cinco aminoácidos, em não mais do que quatro aminoácidos, em não mais do que três aminoácidos, em não mais do que dois aminoácidos, ou em um aminoácido da sequência de aminoácidos de qualquer uma das pululanases descritas *supra* (p.ex., qualquer uma de SEQ ID NOs: 114-120). Em uma modalidade, a pululanase tem uma substituição, deleção e/ou inserção de aminoácidos de uma ou mais (p.ex., duas, várias) de sequências de aminoácidos de qualquer uma das pululanases descritas *supra* (p.ex., qualquer uma de SEQ ID NOs: 114-120). Em alguns aspectos, o número total de substituições, deleções e/ou inserções de aminoácidos é não mais do que 10, p.ex., não mais do que 9, 8, 7, 6, 5, 4, 3, 2 ou 1.

**[0366]** Em algumas modalidades, a pululanase tem pelo menos 20%, p.ex., pelo menos 40%, pelo menos 50%, pelo menos 60%, pelo

menos 70%, pelo menos 80%, pelo menos 90%, pelo menos 95%, pelo menos 96%, pelo menos 97%, pelo menos 98%, pelo menos 99% ou 100% da atividade de pululanase de qualquer pululanase descrita ou referenciada aqui sob as mesmas condições (p.ex., qualquer uma de SEQ ID NOs: 114-120).

**[0367]** Em uma modalidade, a sequência codificante de pululanase híbrida sob condições de pelo menos baixa estringência, p.ex., condições de média estringência, condições de média-elevada estringência, condições de elevada estringência ou condições de muito elevada estringência com a fita complementar de comprimento total da sequência codificante de qualquer pululanase descrita ou referenciada aqui (p.ex., qualquer uma de SEQ ID NOs: 114-120). Em uma modalidade, a sequência codificante de pululanase tem pelo menos 65%, p.ex., pelo menos 70%, pelo menos 75%, pelo menos 80%, pelo menos 85%, pelo menos 85%, pelo menos 90%, pelo menos 91%, pelo menos 92%, pelo menos 93%, pelo menos 94%, pelo menos 95%, pelo menos 96%, pelo menos 97%, pelo menos 98%, pelo menos 99% ou 100% de identidade de sequências com a sequência codificante de qualquer pululanase descrita ou referenciada aqui (p.ex., qualquer uma de SEQ ID NOs: 114-120).

**[0368]** Em uma modalidade, a pululanase compreende a sequência codificante de qualquer pululanase descrita ou referenciada aqui (p.ex., qualquer uma de SEQ ID NOs: 114-120). Em uma modalidade, a pululanase compreende uma sequência codificante que é uma subsequência da sequência codificante de qualquer pululanase descrita ou referenciada aqui, em que a subsequência codifica um polipeptídeo tendo atividade de pululanase. Em uma modalidade, o número de resíduos de nucleotídeos na subsequência é pelo menos 75%, p.ex., pelo menos 80%, 85%, 90%, ou 95% do número da sequência codificante referenciada.

**[0369]** A sequência codificante referenciada de qualquer aspecto ou modalidade relacionado descrito aqui pode ser a sequência codificante nativa ou uma sequência degenerada, tal como uma sequência codificante com otimização de códons concebida para uso em uma célula hospedeira particular (p.ex., otimizada para expressão em *Saccharomyces cerevisiae*).

**[0370]** A pululanase pode também incluir polipeptídeos fundidos ou polipeptídeos de fusão clivável, como descrito *supra*.

#### Métodos usando um Material Contendo Celulose

**[0371]** Em alguns aspectos, os métodos descritos aqui produzem um produto de fermentação a partir de um material contendo celulose. O polissacarídeo predominante na parede celular primária da biomassa é celulose, o segundo mais abundante é hemicelulose e o terceiro é pectina. A parede celular secundária, produzida após a célula ter parado de crescer, contém também polissacarídeos e é fortalecida por lignina polimérica covalentemente reticulada à hemicelulose. A celulose é um homopolímero de anidroclobiose e assim uma beta-(1-4)-D-glucana linear, enquanto que as hemiceluloses incluem uma variedade de compostos, tais como xilanas, xiloglucanas, arabinoxilanas e mananas em complexas estruturas ramificadas com um espectro de substituintes. Embora geralmente polimorfa, a celulose é encontrada em tecido de plantas principalmente como uma matriz cristalina insolúvel de cadeias de glucana paralelas. As hemiceluloses estão usualmente ligadas por ligações de hidrogênio à celulose, bem como a outras hemiceluloses, que ajudam a estabilizar a matriz da parede celular.

**[0372]** A celulose é geralmente encontrada, por exemplo, nos caules, folhas, peles, cascas e sabugos de plantas ou folhas, ramos e madeira de árvores. O material contendo celulose pode ser, mas não está limitado a, um resíduo agrícola, material herbáceo (incluindo culturas

energéticas), resíduo sólido municipal, resíduo de usinas de pasta celulósica e papel, papel residual e madeira (incluindo resíduo de silvicultura) (ver, por exemplo, Wiselogel *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 e Biotechnology* 24/25: 695-719; Mosier *et al.*, 1999, *Recent Progress in Bioconversion of Lignocellulosics*, em *Advances in Biochemical Engineering/Biotechnology*, T. Scheper, editor executivo, Volume 65, pp. 23-40, Springer-Verlag, Nova Iorque). É entendido aqui que a celulose pode estar na forma de lignocelulose, um material das paredes celulares de plantas contendo lignina, celulose e hemicelulose em uma matriz mista. Em uma modalidade, o material contendo celulose é qualquer material de biomassa. Em outra modalidade, o material contendo celulose é lignocelulose, que compreende celulose, hemiceluloses e lignina.

**[0373]** Em uma modalidade, o material contendo celulose é um resíduo agrícola, material herbáceo (incluindo culturas energéticas), resíduo sólido urbano, resíduo de usinas de pasta celulósica e papel, papel residual ou madeira (incluindo resíduo de silvicultura).

**[0374]** Em outra modalidade, o material contendo celulose é cana-do-reino, bagaço, bambu, sabugo de milho, fibra de milho, palha de milho, miscanto, palha de arroz, grama ou palha de trigo.

**[0375]** Em outra modalidade, o material contendo celulose é choupo, eucalipto, abeto, pinheiro, álamo, espruce ou salgueiro.

**[0376]** Em outra modalidade, o material contendo celulose é celulose de algas, celulose bacteriana, fibra de algodão, papel de filtro, celulose microcristalina (p.ex., AVICEL®) ou celulose tratada com ácido fosfórico.

**[0377]** Em outra modalidade, o material contendo celulose é uma biomassa aquática. Como usado aqui, o termo “biomassa aquática”



significa biomassa produzida em um ambiente aquático por um processo de fotossíntese. A biomassa aquática pode ser algas, plantas emergentes, plantas de folhas flutuantes ou plantas submersas.

**[0378]** O material contendo celulose pode ser usado como tal ou pode ser sujeito a pré-tratamento, usando métodos convencionais conhecidos na técnica, como descrito aqui. Em uma modalidade preferencial, o material contendo celulose é pré-tratado.

**[0379]** Os métodos de uso do material contendo celulose podem ser alcançados usando métodos convencionais na técnica. Além disso, os métodos podem ser implantados usando qualquer dispositivo de processamento de biomassa convencional configurado para levar a cabo os processos.

#### Pré-tratamento Celulósico

**[0380]** Em uma modalidade, o material contendo celulose é pré-tratado antes da sacarificação.

**[0381]** Na prática dos processos descritos aqui, qualquer processo de pré-tratamento conhecido na técnica pode ser usado para romper componentes da parede celular vegetal do material contendo celulose (Chandra *et al.*, 2007, *Adv. Biochem. Engin./Biotechnol.* 108: 67-93; Galbe e Zacchi, 2007, *Adv. Biochem. Engin./Biotechnol.* 108: 41-65; Hendriks e Zeeman, 2009, *Bioresource Technology* 100: 10-18; Mosier *et al.*, 2005, *Bioresource Technology* 96: 673-686; Taherzadeh e Karimi, 2008, *Int. J. Mol. Sci.* 9: 1621-1651; Yang e Wyman, 2008, *Biofuels Bioproducts and Biorefining-Biofpr.* 2: 26-40).

**[0382]** O material contendo celulose pode ser também sujeito a redução do tamanho das partículas, crivagem, pré-embebição, umedecimento, lavagem e/ou condicionamento antes do pré-tratamento usando métodos conhecidos na técnica.

**[0383]** Os pré-tratamentos convencionais incluem, mas não estão limitados a, pré-tratamento com vapor (com ou sem explosão), pré-tratamento com ácido diluído, pré-tratamento com água quente, pré-tratamento alcalino, pré-tratamento com cal, oxidação úmida, explosão úmida, explosão de fibras com amônia, pré-tratamento com organossolv e pré-tratamento biológico. Pré-tratamentos adicionais incluem pré-tratamentos de percolação com amônia, ultrassons, eletroporação, micro-ondas, CO<sub>2</sub> supercrítico, H<sub>2</sub>O supercrítica, ozônio, líquido iônico e irradiação gama.

**[0384]** Em uma modalidade, o material contendo celulose é pré-tratado antes da sacarificação (*i.e.*, hidrólise) e/ou fermentação. O pré-tratamento é preferencialmente realizado antes da hidrólise. Alternativamente, o pré-tratamento pode ser levado a cabo simultaneamente com hidrólise enzimática para liberar açúcares fermentáveis, tais como glucose, xilose e/ou celobiose. Na maioria dos casos, o próprio passo de pré-tratamento resulta em alguma conversão da biomassa em açúcares fermentáveis (mesmo na ausência de enzimas).

**[0385]** Em uma modalidade, o material contendo celulose é pré-tratado com vapor. No pré-tratamento com vapor, o material contendo celulose é aquecido para desagregar os componentes das paredes celulares vegetais, incluindo lignina, hemicelulose, e celulose para tornar a celulose e outras frações, por exemplo, hemicelulose, acessíveis a enzimas. O material contendo celulose é passado para ou através de um vaso reacional no qual vapor é injetado para aumentar a temperatura até à temperatura e pressão requeridas e aí retido durante o tempo de reação desejado. O pré-tratamento com vapor é preferencialmente realizado a 140-250 °C, p.ex., 160-200 °C ou 170-190 °C, onde a gama de temperaturas ótima depende da adição opcional de um catalisador químico. O tempo de residência para o pré-tratamento com vapor é preferencialmente 1-60 minutos, p.ex., 1-30 minutos, 1-20 minutos,

3-12 minutos ou 4-10 minutos, onde o tempo de residência ótimo depende da temperatura e adição opcional de um catalisador químico. O pré-tratamento com vapor permite cargas de sólidos relativamente elevadas, tal que o material contendo celulose seja geralmente somente umedecido durante o pré-tratamento. O pré-tratamento com vapor é frequentemente combinado com uma descarga explosiva do material após o pré-tratamento, que é conhecida como explosão a vapor, isto é, expansão rápida até à pressão atmosférica e fluxo turbulento do material para aumentar a área superficial acessível por fragmentação (Duff e Murray, 1996, *Bioresource Technology* 855: 1-33; Galbe e Zacchi, 2002, *Appl. Microbiol. Biotechnol.* 59: 618-628; Pedido de Patente dos E.U.A. No. 2002/0164730). Durante o pré-tratamento com vapor, os grupos acetila da hemicelulose são clivados e o ácido resultante autocatalisa a hidrólise parcial da hemicelulose em monossacarídeos e oligossacarídeos. A lignina é removida somente em uma extensão limitada.

**[0386]** Em uma modalidade, o material contendo celulose é sujeito a um pré-tratamento químico. O termo “tratamento químico” se refere a qualquer pré-tratamento químico que promova a separação e/ou liberação de celulose, hemicelulose e/ou lignina. Um tal pré-tratamento pode converter celulose cristalina em celulose amorfa. Exemplos de processos de pré-tratamento químico adequados incluem, por exemplo, pré-tratamento com ácido diluído, pré-tratamento com cal, oxidação úmida, expansão de fibras com amônia/congelamento (AFEX), percolação com amônia (APR), líquido iônico e pré-tratamentos com organossolv.

**[0387]** Um catalisador químico tal como  $H_2SO_4$  ou  $SO_2$  (tipicamente 0,3 a 5% p/p) é por vezes adicionado antes do pré-tratamento com vapor, o que diminui o tempo e a temperatura, aumenta a recuperação e melhora a hidrólise enzimática (Ballésteros *et al.*, 2006, *Appl. Biochem. Biotechnol.* 129-132: 496-508; Varga *et al.*, 2004, *Appl. Biochem. Biotechnol.*

113-116: 509-523; Sassner *et al.*, 2006, *Enzyme Microb. Technol.* 39: 756-762). No pré-tratamento com ácido diluído, o material contendo celulose é misturado com ácido diluído, tipicamente H<sub>2</sub>SO<sub>4</sub>, e água para formar uma pasta semifluida, aquecido por vapor até à temperatura desejada e, após um tempo de residência, expandido até à pressão atmosférica. O pré-tratamento com ácido diluído pode ser realizado com um número de desenhos de reatores, p.ex., reatores de fluxo de pistão, reatores em contracorrente, ou reatores em contracorrente contínuos com encolhimento do leite (Duff e Murray, 1996, *Bioresource Technology* 855: 1-33; Schell *et al.*, 2004, *Bioresource Technology* 91: 179-188; Lee *et al.*, 1999, *Adv. Biochem. Eng. Biotechnol.* 65: 93-115). Em uma modalidade específica, o pré-tratamento com ácido diluído de material contendo celulose é levado a cabo usando ácido sulfúrico a 4% p/p a 180 °C durante 5 minutos.

**[0388]** Podem ser também usados vários métodos de pré-tratamento sob condições alcalinas. Estes pré-tratamentos alcalinos incluem, mas não estão limitados a, pré-tratamento com hidróxido de sódio, cal, oxidação úmida, percolação com amônia (APR) e expansão de fibras com amônia/congelamento (AFEX).

**[0389]** O pré-tratamento com cal é realizado com óxido de cálcio ou hidróxido de cálcio a temperaturas de 85-150 °C e tempos de residência de 1 hora a vários dias (Wyman *et al.*, 2005, *Bioresource Technol.* 96: 1959-1966; Mosier *et al.*, 2005, *Bioresource Technology* 96: 673-686). WO2006/110891, WO2006/110899, WO2006/110900 e WO2006/110901 divulgam métodos de pré-tratamento usando amônia.

**[0390]** A oxidação úmida é um pré-tratamento térmico realizado tipicamente a 180-200 °C durante 5-15 minutos com adição de um agente oxidante tal como peróxido de hidrogênio ou sobrepressão de oxigênio (Schmidt e Thomsen, 1998, *Bioresource Technol.* 64: 139-151; Palonen *et al.*, 2004, *Appl. Biochem. Biotechnol.* 117: 1-17; Varga *et al.*,

2004, *Biotechnol. Bioeng.* 88: 567-574; Martin *et al.*, 2006, *J. Chem. Technol. Biotechnol.* 81: 1669-1677). O pré-tratamento é realizado preferencialmente com matéria seca a 1-40%, p.ex., matéria seca a 2-30% ou matéria seca a 5-20%, e frequentemente o pH inicial é aumentado pela adição de um composto alcalino como carbonato de sódio.

**[0391]** Uma modificação do método de pré-tratamento de oxidação úmida, conhecida como explosão úmida (combinação de oxidação úmida e explosão a vapor), pode tratar matéria seca até 30%. Na explosão úmida, o agente oxidante é introduzido durante o pré-tratamento após um certo tempo de residência. O pré-tratamento é depois finalizado por expansão até à pressão atmosférica (WO2006/032282).

**[0392]** A expansão de fibras com amônia (AFEX) envolve tratamento do material contendo celulose com amônia líquida ou gasosa a temperaturas moderadas tais como 90-150 °C e pressão elevada tal como 17-20 bar durante 5-10 minutos, onde o conteúdo de matéria seca pode ser tão elevado quanto 60% (Gollapalli *et al.*, 2002, *Appl. Biochem. Biotechnol.* 98: 23-35; Chundawat *et al.*, 2007, *Biotechnol. Bioeng.* 96: 219-231; Alizadeh *et al.*, 2005, *Appl. Biochem. Biotechnol.* 121: 1133-1141; Teymouri *et al.*, 2005, *Bioresource Technology* 96: 2014-2018). Durante o pré-tratamento com AFEX, a celulose e as hemiceluloses permanecem relativamente intatas. Os complexos de lignina-carboidrato são clivados.

**[0393]** O pré-tratamento com organossolv deslignifica o material contendo celulose por extração usando etanol aquoso (etanol a 40-60%) a 160-200 °C durante 30-60 minutos (Pan *et al.*, 2005, *Biotechnol. Bioeng.* 90: 473-481; Pan *et al.*, 2006, *Biotechnol. Bioeng.* 94: 851-861; Kurabi *et al.*, 2005, *Appl. Biochem. Biotechnol.* 121: 219-230). O ácido sulfúrico é usualmente adicionado como um catalisador. No pré-tratamento com organossolv, a maioria da hemicelulose e lignina é removida.

**[0394]** Outros exemplos de métodos de pré-tratamento adequados são descritos por Schell *et al.*, 2003, *Appl. Biochem. Biotechnol.* 105-108: 69-85 e Mosier *et al.*, 2005, *Bioresource Technology* 96: 673-686 e Pedido Publicado dos E.U.A. 2002/0164730.

**[0395]** Em uma modalidade, o pré-tratamento químico é levado a cabo como um tratamento de ácido diluído e, mais preferencialmente, como um tratamento de ácido diluído contínuo. O ácido é tipicamente ácido sulfúrico, mas podem ser também usados outros ácidos, tais como ácido acético, ácido cítrico, ácido nítrico, ácido fosfórico, ácido tartárico, ácido succínico, cloreto de hidrogênio ou suas misturas. O tratamento com ácido suave é conduzido na gama de pH preferencialmente de 1-5, p.ex., 1-4 ou 1-2,5. Em um aspecto, a concentração de ácido está preferencialmente na gama de 0,01 a 10% em peso de ácido, p.ex. 0,05 a 5% em peso de ácido ou 0,1 a 2% em peso de ácido. O ácido é contatado com o material contendo celulose e mantido a uma temperatura na gama de preferencialmente 140-200 °C, p.ex., 165-190 °C, durante períodos variando de 1 a 60 minutos.

**[0396]** Em outra modalidade, o pré-tratamento tem lugar em uma pasta semifluida aquosa. Em aspectos preferenciais, o material contendo celulose está presente durante o pré-tratamento em quantidades preferencialmente entre 10-80% em peso, p.ex., 20-70% em peso ou 30-60% em peso, tal como em torno de 40% em peso. O material contendo celulose pré-tratado pode não ser lavado ou ser lavado usando qualquer método conhecido na técnica, p.ex., lavado com água.

**[0397]** Em uma modalidade, o material contendo celulose é sujeito a pré-tratamento físico ou mecânico. O termo “pré-tratamento mecânico” ou “pré-tratamento físico” se refere a qualquer pré-tratamento que promove a redução do tamanho das partículas. Por exemplo, tal pré-

tratamento pode envolver vários tipos de trituração ou moagem (p.ex., moagem a seco, moagem úmida ou moagem com bolas vibratórias).

**[0398]** O material contendo celulose pode ser pré-tratado tanto fisicamente (mecanicamente) como quimicamente. O pré-tratamento mecânico ou físico pode ser acoplado com tratamento com vapor/explosão a vapor, hidrotermólise, tratamento com ácido diluído ou fraco, tratamento a elevada temperatura, elevada pressão, irradiação (p.ex., irradiação com micro-ondas), ou suas combinações. Em um aspecto, elevada pressão significa pressão na gama de preferencialmente cerca de 100 a cerca de 400 psi, p.ex., cerca de 150 a cerca de 250 psi. Em outro aspecto, elevada temperatura significa temperaturas na gama de cerca de 100 a cerca de 300 °C, p.ex., cerca de 140 a cerca de 200 °C. Em um aspecto preferencial, o pré-tratamento mecânico ou físico é realizado em um processo descontínuo usando um sistema hidrolisador de pistola a vapor que usa elevada pressão e elevada temperatura tais como definidas acima, p.ex., um Hidrolisador Sunds disponível a partir da Sunds Defibrator AB, Suécia. Os pré-tratamentos físicos e químicos podem ser levados a cabo sequencialmente ou simultaneamente, como desejado.

**[0399]** Conformemente, em uma modalidade, o material contendo celulose é sujeito a pré-tratamento químico ou físico (mecânico), ou qualquer sua combinação, para promover a separação e/ou liberação de celulose, hemicelulose e/ou lignina.

**[0400]** Em uma modalidade, o material contendo celulose é sujeito a um pré-tratamento biológico. O termo “pré-tratamento biológico” se refere a qualquer pré-tratamento biológico que promova a separação e/ou liberação de celulose, hemicelulose e/ou lignina a partir do material contendo celulose. Técnicas de pré-tratamento biológico podem envolver aplicação de microrganismos e/ou enzimas solubilizadores de lignina (ver, por exemplo, Hsu, T.-A., 1996, *Pretreatment of biomass*, em *Handbook on Bioethanol*:

*Production and Utilization*, Wyman, C. E., ed., Taylor & Francis, Washington, DC, 179-212; Ghosh e Singh, 1993, *Adv. Appl. Microbiol.* 39: 295-333; McMillan, J. D., 1994, *Pretreating lignocellulosic biomass: a review*, em *Enzymatic Conversion of Biomass for Fuels Production*, Himmel, M. E., Baker, J. O. e Overend, R. P., eds., ACS Symposium Series 566, American Chemical Society, Washington, DC, capítulo 15; 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; Olsson e Hahn-Hagerdal, 1996, *Enz. Microb. Tech.* 18: 312-331; e Vallander e Eriksson, 1990, *Adv. Biochem. Eng./Biotechnol.* 42: 63-95).

#### Sacarificação e Fermentação de Material contendo celulose

**[0401]** A sacarificação (*i.e.*, hidrólise) e fermentação, separadas ou simultâneas, incluem, mas não estão limitadas a, hidrólise e fermentação separadas (SHF); sacarificação e fermentação simultâneas (SSF); sacarificação e cofermentação simultâneas (SSCF); hidrólise e fermentação híbridas (HHF); hidrólise e cofermentação separadas (SHCF); hidrólise e cofermentação híbridas (HHCF).

**[0402]** A SHF usa passos processuais separados para hidrolisar enzimaticamente em primeiro lugar o material contendo celulose em açúcares fermentáveis, p.ex., glucose, celobiose e monômeros de pentose e, depois, fermentar os açúcares fermentáveis até etanol. Na SSF, a hidrólise enzimática do material contendo celulose e a fermentação de açúcares em etanol são combinadas em um passo (Philippidis, G. P., 1996, *Cellulose bioconversion technology*, em *Handbook on Bioethanol: Production and Utilization*, Wyman, C. E., ed., Taylor & Francis, Washington, DC, 179-212). A SSCF envolve a cofermentação de múltiplos açúcares (Sheehan e Himmel, 1999, *Biotechnol. Prog.* 15: 817-827). A HHF envolve um passo de hidrólise separado e, adicionalmente, um passo de sacarificação e hidrólise



simultâneas, que podem ser levados a cabo no mesmo reator. Os passos em um processo de HHF podem ser levados a cabo a diferentes temperaturas, *i.e.*, sacarificação enzimática a elevada temperatura seguida por SSF a uma temperatura mais baixa que o organismo de fermentação possa tolerar. É entendido aqui que qualquer método conhecido na técnica compreendendo pré-tratamento, hidrólise enzimática (sacarificação), fermentação, ou uma sua combinação, pode ser usado na prática dos processos descritos aqui.

**[0403]** Um dispositivo convencional pode incluir um reator agitado alimentado descontínuo com alimentação, um reator agitado descontínuo, um reator agitado de fluxo contínuo com ultrafiltração e/ou um reator de coluna de fluxo com pistão contínuo (de Castilhos Corazza *et al.*, 2003, *Acta Scientiarum. Technology* 25: 33-38; Gusakov e Sinitsyn, 1985, *Enz. Microb. Technol.* 7: 346-352), um reator de atrito (Ryu e Lee, 1983, *Biotechnol. Bioeng.* 25: 53-65). Tipos de reator adicionais incluem reatores de leito fluidizado, manta de fluxo ascendente, imobilizados e tipo extrusora para hidrólise e/ou fermentação.

**[0404]** No passo de sacarificação (*i.e.*, passo de hidrólise), o material contendo celulose e/ou amido, p.ex., pré-tratado, é hidrolisado para desagregar a celulose, hemicelulose e/ou amido em açúcares fermentáveis, tais como glucose, celobiose, xilose, xilulose, arabinose, manose, galactose e/ou oligossacarídeos solúveis. A hidrólise é realizada enzimaticamente, p.ex., por uma composição de enzimas celulolíticas. As enzimas das composições podem ser adicionadas simultaneamente ou sequencialmente.

**[0405]** A hidrólise enzimática pode ser levada a cabo em um ambiente aquoso adequado sob condições que podem ser prontamente determinadas por um perito na técnica. Em um aspecto, a hidrólise é realizada sob condições adequadas para a atividade da(s) enzima(s), *i.e.*, ótimas para a(s) enzima(s). A hidrólise pode ser levada a cabo como um

processo descontínuo com alimentação ou contínuo onde o material contendo celulose e/ou amido é alimentado gradualmente a, por exemplo, uma solução de hidrólise contendo enzimas.

**[0406]** A sacarificação é geralmente realizada em reatores ou fermentadores de tanque agitado sob condições controladas de pH, temperatura e mistura. Condições adequadas de tempo de processo, temperatura e pH podem ser prontamente determinadas por um perito na técnica. Por exemplo, a sacarificação pode durar até 200 horas, mas tipicamente é realizada durante preferencialmente cerca de 12 a cerca de 120 horas, p.ex., cerca de 16 a cerca de 72 horas ou cerca de 24 a cerca de 48 horas. A temperatura está na gama de preferencialmente cerca de 25 °C a cerca de 70 °C, p.ex., cerca de 30 °C a cerca de 65 °C, cerca de 40 °C a cerca de 60 °C ou cerca de 50 °C a cerca de 55 °C. O pH está na gama de preferencialmente cerca de 3 a cerca de 8, p.ex., cerca de 3,5 a cerca de 7, cerca de 4 a cerca de 6 ou cerca de 4,5 a cerca de 5,5. O conteúdo de sólidos secos está na gama de preferencialmente cerca de 5 a cerca de 50% em peso, p.ex., cerca de 10 a cerca de 40% em peso ou cerca de 20 a cerca de 30% em peso.

**[0407]** A sacarificação pode ser levada a cabo usando uma composição de enzimas celulolíticas. Tais composições enzimáticas são descritas em baixo na seção “Composição de Enzimas Celulolíticas” em baixo. As composições de enzimas celulolíticas podem compreender qualquer proteína útil na degradação do material contendo celulose. Em um aspecto, a composição de enzimas celulolíticas compreende ou compreende adicionalmente uma ou mais (p.ex., várias) proteínas selecionadas do grupo consistindo em uma celulase, um polipeptídeo AA9 (GH61), uma hemicelulase, uma esterase, uma expansina, uma enzima ligninolítica, uma oxidorreductase, uma pectinase, uma protease e uma swollenina.

**[0408]** Em outra modalidade, a celulase é preferencialmente uma ou mais (p.ex., várias) enzimas selecionadas do grupo consistindo em uma endoglucanase, uma celobioidrolase e uma beta-glucosidase.

**[0409]** Em outra modalidade, a hemicelulase é preferencialmente uma ou mais (p.ex., várias) enzimas selecionadas do grupo consistindo em uma acetilmanana esterase, uma acetilxilana esterase, uma arabinanase, uma arabinofuranosidase, uma ácido cumárico esterase, uma feruloil esterase, uma galactosidase, uma glucuronidase, uma glucuronoil esterase, uma mananase, uma manosidase, uma xilanase e uma xilosidase. Em outra modalidade, a oxirredutase é uma ou mais (p.ex., várias) enzimas selecionadas do grupo consistindo em uma catalase, uma lacase e uma peroxidase.

**[0410]** As enzimas ou composições de enzimas usadas em um processo da presente invenção podem estar em qualquer forma adequada para uso, tal como, por exemplo, uma formulação de caldo de fermentação ou uma composição de células, um lisado de células com ou sem detritos celulares, uma preparação de enzimas semipurificada ou purificada ou uma célula hospedeira como uma fonte das enzimas. A composição de enzimas pode ser um pó ou granulado seco, um granulado sem formação de poeira, um líquido, um líquido estabilizado ou uma enzima protegida estabilizada. As preparações de enzimas líquidas podem, por exemplo, ser estabilizadas por adição de estabilizantes tais como um açúcar, um álcool de açúcar ou outro poliol e/ou ácido láctico ou outro ácido orgânico de acordo com processos estabelecidos.

**[0411]** Em uma modalidade, uma quantidade eficaz de composição de enzimas celulolíticas ou hemicelulolíticas em relação ao material contendo celulose é cerca de 0,5 a cerca de 50 mg, p.ex., cerca de 0,5 a cerca de 40 mg, cerca de 0,5 a cerca de 25 mg, cerca de 0,75 a cerca

de 20 mg, cerca de 0,75 a cerca de 15 mg, cerca de 0,5 a cerca de 10 mg ou cerca de 2,5 a cerca de 10 mg por g do material contendo celulose.

**[0412]** Em uma modalidade, um tal composto é adicionado a uma razão molar entre o composto e unidades de glicosila de celulose de cerca de  $10^{-6}$  a cerca de 10, p.ex., cerca de  $10^{-6}$  a cerca de 7,5, cerca de  $10^{-6}$  a cerca de 5, cerca de  $10^{-6}$  a cerca de 2,5, cerca de  $10^{-6}$  a cerca de 1, cerca de  $10^{-5}$  a cerca de 1, cerca de  $10^{-5}$  a cerca de  $10^{-1}$ , cerca de  $10^{-4}$  a cerca de  $10^{-1}$ , cerca de  $10^{-3}$  a cerca de  $10^{-1}$  ou cerca de  $10^{-3}$  a cerca de  $10^{-2}$ . Em outro aspecto, uma quantidade eficaz de um tal composto é cerca de 0,1  $\mu\text{M}$  a cerca de 1 M, p.ex., cerca de 0,5  $\mu\text{M}$  a cerca de 0,75 M, cerca de 0,75  $\mu\text{M}$  a cerca de 0,5 M, cerca de 1  $\mu\text{M}$  a cerca de 0,25 M, cerca de 1  $\mu\text{M}$  a cerca de 0,1 M, cerca de 5  $\mu\text{M}$  a cerca de 50 mM, cerca de 10  $\mu\text{M}$  a cerca de 25 mM, cerca de 50  $\mu\text{M}$  a cerca de 25 mM, cerca de 10  $\mu\text{M}$  a cerca de 10 mM, cerca de 5  $\mu\text{M}$  a cerca de 5 mM ou cerca de 0,1 mM a cerca de 1 mM.

**[0413]** O termo "licor" significa a fase de solução, aquosa, orgânica ou uma sua combinação, tendo origem no tratamento de um material de lignocelulose e/ou hemicelulose em uma pasta semifluida, ou seus monossacarídeos, p.ex., xilose, arabinose, manose, *etc.*, sob condições como descrito em WO2012/021401, e os seus conteúdos solúveis. Um licor para intensificação celulolítica de um polipeptídeo AA9 polipeptídeo GH61) pode ser produzido por tratamento de um material (ou matéria-prima) de lignocelulose ou hemicelulose por aplicação de calor e/ou pressão, opcionalmente na presença de um catalisador, p.ex., ácido, opcionalmente na presença de um solvente orgânico e, opcionalmente, em combinação com ruptura física do material e, depois, separação da solução a partir dos sólidos residuais. Tais condições determinam o grau de intensificação celulolítica obténível através da combinação de licor e um polipeptídeo AA9 durante a hidrólise de um substrato celulósico por uma preparação de enzimas celulolíticas. O licor pode ser separado do material tratado usando um

método-padrão na técnica, tal como filtração, sedimentação ou centrifugação.

**[0414]** Em uma modalidade, uma quantidade eficaz de licor em relação à celulose é cerca de  $10^{-6}$  a cerca de 10 g por g de celulose, p.ex., cerca de  $10^{-6}$  a cerca de 7,5 g, cerca de  $10^{-6}$  a cerca de 5 g, cerca de  $10^{-6}$  a cerca de 2,5 g, cerca de  $10^{-6}$  a cerca de 1 g, cerca de  $10^{-5}$  a cerca de 1 g, cerca de  $10^{-5}$  a cerca de  $10^{-1}$  g, cerca de  $10^{-4}$  a cerca de  $10^{-1}$  g, cerca de  $10^{-3}$  a cerca de  $10^{-1}$  g ou cerca de  $10^{-3}$  a cerca de  $10^{-2}$  g por g de celulose.

**[0415]** No passo de fermentação, os açúcares, liberados a partir do material contendo celulose, p.ex., em resultado dos passos de pré-tratamento e hidrólise enzimática, são fermentados até etanol, por um organismo fermentador, tal como levedura descrita aqui. A hidrólise (sacarificação) e a fermentação podem ser separadas ou simultâneas.

**[0416]** Qualquer material contendo celulose hidrolisado adequado pode ser usado no passo de fermentação na prática dos processos descritos aqui. Tais matérias-primas incluem, mas não estão limitadas a, carboidratos (p.ex., lignocelulose, xilanas, celulose, amido, etc.). O material é geralmente selecionado com base na economia, *i.e.*, custos por potencial de açúcar equivalente, e recalcitrância para conversão enzimática.

**[0417]** A produção de etanol por um organismo fermentador usando material contendo celulose resulta do metabolismo de açúcares (monossacarídeos). A composição de açúcar do material contendo celulose hidrolisado e a capacidade do organismo fermentador de utilizar os diferentes açúcares têm um impacto direto nos rendimentos do processo. Antes da divulgação do Requerente aqui, as estirpes conhecidas na técnica utilizam glucose eficientemente mas não (ou muito limitadamente) metabolizam pentoses como xilose, um monossacarídeo comumente encontrado em material hidrolisado.

**[0418]** As composições dos meios de fermentação e as condições de fermentação dependem do organismo fermentador e podem ser facilmente determinadas por um perito na técnica. Tipicamente, a fermentação tem lugar sob condições conhecidas como sendo adequadas para gerar o produto de fermentação. Em algumas modalidades, o processo de fermentação é levado a cabo sob condições aeróbicas ou microaerofílicas (*i.e.*, onde a concentração de oxigênio é menor do que aquela no ar) ou condições anaeróbicas. Em algumas modalidades, a fermentação é conduzida sob condições anaeróbicas (*i.e.*, nenhum oxigênio detectável) ou menos do que cerca de 5, cerca de 2,5 ou cerca de 1 mmol/L/h de oxigênio. Na ausência de oxigênio, o NADH produzido na glicólise não pode ser oxidado por fosforilação oxidativa. Sob condições anaeróbicas pode ser utilizado piruvato ou um seu derivado pela célula hospedeira como um receptor de elétrons e hidrogênio de modo a gerar NAD<sup>+</sup>.

**[0419]** O processo de fermentação é tipicamente realizado a uma temperatura que é ótima para a célula fúngica recombinante. Por exemplo, em algumas modalidades, o processo de fermentação é realizado a uma temperatura na gama de cerca de 25 °C a cerca de 42 °C. Tipicamente, o processo é levado a cabo a uma temperatura que é menor do que cerca de 38 °C, menor do que cerca de 35 °C, menor do que cerca de 33 °C ou menor do que cerca de 38 °C, mas pelo menos cerca de 20 °C, 22 °C ou 25 °C.

**[0420]** Um estimulante da fermentação pode ser usado em um processo descrito aqui para melhorar adicionalmente a fermentação, e em particular, o desempenho do organismo fermentador, tal como intensificação da taxa e rendimento do produto (*p.ex.*, rendimento de etanol). Um “estimulante da fermentação” se refere a estimulantes do crescimento dos organismos fermentadores, em particular, leveduras. Estimulantes da fermentação preferenciais para crescimento incluem vitaminas e minerais.

Exemplos de vitaminas incluem multivitaminas, biotina, pantotenato, ácido nicotínico, meso-inositol, tiamina, piridoxina, ácido *para*-aminobenzoico, 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 deste modo incorporado por referência. Exemplos de minerais incluem minerais e sais minerais que podem fornecer nutrientes compreendendo P, K, Mg, S, Ca, Fe, Zn, Mn e Cu.

#### Enzimas e Composições Celulolíticas

**[0421]** Uma enzima celulolítica ou composição de enzimas celulolíticas pode estar presente e/ou ser adicionada durante a sacarificação. Uma composição de enzimas celulolíticas é uma preparação de enzimas contendo uma ou mais (p.ex., várias) enzimas que hidrolisam material contendo celulose. Tais enzimas incluem endoglucanase, celobioidrolase, beta-glucosidase e/ou suas combinações.

**[0422]** Em algumas modalidades, o organismo fermentador compreende um ou mais (p.ex., vários) polinucleotídeos heterólogos codificando enzimas que hidrolisam material contendo celulose (p.ex., uma endoglucanase, celobioidrolase, beta-glucosidase ou suas combinações). Qualquer enzima descrita ou referenciada aqui que hidrolise material contendo celulose está contemplada para expressão no organismo fermentador.

**[0423]** A enzima celulolítica pode ser qualquer enzima celulolítica que seja adequada para as células hospedeiras e/ou os métodos descritos aqui (p.ex., uma endoglucanase, celobioidrolase, beta-glucosidase), tal como uma enzima celulolítica ocorrendo naturalmente ou uma sua variante que retenha atividade de enzima celulolítica.

**[0424]** Em algumas modalidades, o organismo fermentador compreendendo um polinucleotídeo heterólogo codificando uma

enzima celulolítica tem um nível aumentado de atividade de enzima celulolítica (p.ex., endoglucanase, celobioidrolase e/ou beta-glucosidase aumentadas) em comparação com as células hospedeiras sem o polinucleotídeo heterólogo codificando a enzima celulolítica, quando cultivados sob as mesmas condições. Em algumas modalidades, o organismo fermentador tem um nível aumentado de atividade de enzima celulolítica de pelo menos 5%, p.ex., pelo menos 10%, pelo menos 15%, pelo menos 20%, pelo menos 25%, pelo menos 50%, pelo menos 100%, pelo menos 150%, pelo menos 200%, pelo menos 300% ou pelo menos 500% em comparação com o organismo fermentador sem o polinucleotídeo heterólogo codificando a enzima celulolítica, quando cultivados sob as mesmas condições.

**[0425]** Enzimas celulolíticas exemplificativas que podem ser usadas com as células hospedeiras e/ou os métodos descritos aqui incluem enzimas celulolíticas bacterianas, de leveduras ou fúngicas filamentosas, p.ex., obtidas de qualquer um dos microrganismos descritos ou referenciados aqui, como descrito *supra* sob as seções relacionadas com proteases.

**[0426]** A enzima celulolítica pode ter qualquer origem. Em uma modalidade, a enzima celulolítica é derivada de uma estirpe de *Trichoderma*, tal como uma estirpe de *Trichoderma reesei*; uma estirpe de *Humicola*, tal como uma estirpe de *Humicola insolens* e/ou uma estirpe de *Chrysosporium*, tal como uma estirpe de *Chrysosporium lucknowense*. Em uma modalidade preferencial, a enzima celulolítica é derivada de uma estirpe de *Trichoderma reesei*.

**[0427]** A composição de enzimas celulolíticas pode compreender adicionalmente um ou mais dos seguintes polipeptídeos, tais como enzimas: polipeptídeo AA9 (polipeptídeo GH61) tendo atividade



intensificante celulolítica, beta-glucosidase, xilanase, beta-xilosidase, CBH I, CBH II ou uma mistura de dois, três, quatro, cinco ou seis destes.

**[0428]** O(s) polipeptídeo(s) adicional/adicionais (p.ex., polipeptídeo AA9) e/ou enzima(s) (p.ex., beta-glucosidase, xilanase, beta-xilosidase, CBH I e/ou CBH II) pode(m) ser estranho(s) ao organismo produtor da composição de enzimas celulolíticas (p.ex., *Trichoderma reesei*).

**[0429]** Em uma modalidade, a composição de enzimas celulolíticas compreende um polipeptídeo AA9 tendo atividade intensificante celulolítica e uma beta-glucosidase.

**[0430]** Em outra modalidade, a composição de enzimas celulolíticas compreende um polipeptídeo AA9 tendo atividade intensificante celulolítica, uma beta-glucosidase e uma CBH I.

**[0431]** Em outra modalidade, a composição de enzimas celulolíticas compreende um polipeptídeo AA9 tendo atividade intensificante celulolítica, uma beta-glucosidase, uma CBH I e uma CBH II.

**[0432]** Outras enzimas, tais como endoglucanases, podem estar também compreendidas na composição de enzimas celulolíticas.

**[0433]** Tal como mencionado acima, a composição de enzimas celulolíticas pode compreender alguns polipeptídeos diferentes, incluindo enzimas.

**[0434]** Em uma modalidade, a composição de enzimas celulolíticas é uma composição de enzimas celulolíticas de *Trichoderma reesei*, compreendendo adicionalmente polipeptídeo AA9 de *Thermoascus aurantiacus* (GH61A) tendo atividade intensificante celulolítica (p.ex., WO2005/074656) e proteína de fusão de beta-glucosidase de *Aspergillus oryzae* (p.ex., uma divulgada em WO2008/057637, em particular mostrada como SEQ ID NOs: 59 e 60).

**[0435]** Em outra modalidade, a composição de enzimas celulolíticas é uma composição de enzimas celulolíticas de *Trichoderma*

*reesei*, compreendendo adicionalmente polipeptídeo AA9 (GH61A) de *Thermoascus aurantiacus* tendo atividade intensificante celulolítica (p.ex. SEQ ID NO: 2 em WO2005/074656) e beta-glicosidase de *Aspergillus fumigatus* (p.ex., SEQ ID NO: 2 de WO2005/047499).

**[0436]** Em outra modalidade, a composição de enzimas celulolíticas é uma composição de enzimas celulolíticas de *Trichoderma reesei*, compreendendo adicionalmente polipeptídeo AA9 de *Penicillium emersonii* (GH61A) tendo atividade intensificante celulolítica, em particular aquele divulgado em WO2011/041397, e beta-glicosidase de *Aspergillus fumigatus* (p.ex., SEQ ID NO: 2 de WO2005/047499).

**[0437]** Em outra modalidade, a composição de enzimas celulolíticas é uma composição de enzimas celulolíticas de *Trichoderma reesei*, compreendendo adicionalmente polipeptídeo AA9 de *Penicillium emersonii* (GH61A) tendo atividade intensificante celulolítica, em particular aquele divulgado em WO2011/041397, e beta-glicosidase de *Aspergillus fumigatus* (p.ex., SEQ ID NO: 2 de WO2005/047499) ou uma variante divulgada em WO2012/044915 (deste modo incorporada por referência), em particular uma compreendendo uma ou mais das, tal como todas as, seguintes substituições: F100D, S283G, N456E, F512Y.

**[0438]** Em uma modalidade, a composição de enzimas celulolíticas é uma composição de enzimas celulolíticas de *Trichoderma reesei* compreendendo adicionalmente polipeptídeo AA9 (GH61A) tendo atividade intensificante celulolítica, em particular aquele derivado de uma estirpe de *Penicillium emersonii* (SEQ ID NO: 2 em WO2011/041397), variante de beta-glicosidase de *Aspergillus fumigatus* (p.ex., SEQ ID NO: 2 em WO2005/047499) com uma ou mais das, em particular todas as, seguintes substituições: F100D, S283G, N456E, F512Y e divulgada em WO2012/044915; CBH1 Cel7A de *Aspergillus fumigatus*, p.ex., aquela divulgada como SEQ ID NO: 6 em WO2011/057140 e CBH II de *Aspergillus*

*fumigatus*, p.ex., aquela divulgada como SEQ ID NO: 18 em WO2011/057140.

**[0439]** Em uma modalidade preferencial, a composição de enzimas celulolíticas é uma composição de enzimas celulolíticas de *Trichoderma reesei*, compreendendo adicionalmente uma composição de hemicelulases ou enzimas hemicelulolíticas, tal como uma xilanase de *Aspergillus fumigatus* e beta-xilosidase de *Aspergillus fumigatus*.

**[0440]** Em uma modalidade, a composição de enzimas celulolíticas compreende também uma xilanase (p.ex., derivada de uma estirpe do gênero *Aspergillus*, em particular *Aspergillus aculeatus* ou *Aspergillus fumigatus*; ou uma estirpe do gênero *Talaromyces*, em particular *Talaromyces leycettanus*) e/ou uma beta-xilosidase (p.ex., derivada de *Aspergillus*, em particular *Aspergillus fumigatus*, ou uma estirpe de *Talaromyces*, em particular *Talaromyces emersonii*).

**[0441]** Em uma modalidade, a composição de enzimas celulolíticas é uma composição de enzimas celulolíticas de *Trichoderma reesei*, compreendendo adicionalmente polipeptídeo AA9 de *Thermoascus aurantiacus* (GH61A) tendo atividade intensificante celulolítica (p.ex., WO2005/074656), proteína de fusão de beta-glucosidase de *Aspergillus oryzae* (p.ex., uma divulgada em WO2008/057637, em particular como SEQ ID NOs: 59 e 60) e xilanase de *Aspergillus aculeatus* (por exemplo, Xyl II em WO94/21785).

**[0442]** Em outra modalidade, a composição de enzimas celulolíticas compreende uma preparação celulolítica de *Trichoderma reesei*, compreendendo adicionalmente polipeptídeo GH61A de *Thermoascus aurantiacus* tendo atividade intensificante celulolítica (p.ex., SEQ ID NO: 2 em WO2005/074656), variante de beta-glucosidase de *Aspergillus fumigatus* (p.ex., SEQ ID NO: 2 de WO2005/047499) e xilanase de *Aspergillus aculeatus* (Xyl II divulgada em WO94/21785).

**[0443]** Em outra modalidade, a composição de enzimas celulolíticas compreende uma composição de enzimas celulolíticas de *Trichoderma reesei*, compreendendo adicionalmente polipeptídeo AA9 (GH61A) de *Thermoascus aurantiacus* tendo atividade intensificante celulolítica (p.ex. SEQ ID NO: 2 em WO2005/074656), variante de beta-glucosidase de *Aspergillus fumigatus* (p.ex., SEQ ID NO: 2 de WO2005/047499) e xilanase de *Aspergillus aculeatus* (p.ex., Xyl II divulgada em WO94/21785).

**[0444]** Em outra modalidade, a composição de enzimas celulolíticas é uma composição de enzimas celulolíticas de *Trichoderma reesei*, compreendendo adicionalmente polipeptídeo AA9 de *Penicillium emersonii* (GH61A) tendo atividade intensificante celulolítica, em particular aquele divulgado em WO2011/041397, beta-glucosidase de *Aspergillus fumigatus* (p.ex., SEQ ID NO: 2 de WO2005/047499) e xilanase de *Aspergillus fumigatus* (p.ex., Xyl III em WO2006/078256).

**[0445]** Em outra modalidade, a composição de enzimas celulolíticas compreende uma composição de enzimas celulolíticas de *Trichoderma reesei*, compreendendo adicionalmente polipeptídeo AA9 de *Penicillium emersonii* (GH61A) tendo atividade intensificante celulolítica, em particular aquele divulgado em WO2011/041397, beta-glucosidase de *Aspergillus fumigatus* (p.ex., SEQ ID NO: 2 de WO2005/047499), xilanase de *Aspergillus fumigatus* (p.ex., Xyl III em WO2006/078256) e CBH I de *Aspergillus fumigatus*, em particular Cel7A CBH1 divulgada como SEQ ID NO: 2 em WO2011/057140.

**[0446]** Em outra modalidade, a composição de enzimas celulolíticas é uma composição de enzimas celulolíticas de *Trichoderma reesei*, compreendendo adicionalmente polipeptídeo AA9 de *Penicillium emersonii* (GH61A) tendo atividade intensificante celulolítica, em particular aquele divulgado em WO2011/041397, beta-glucosidase de *Aspergillus*

*fumigatus* (p.ex., SEQ ID NO: 2 de WO2005/047499), xilanase de *Aspergillus fumigatus* (p.ex., Xyl III em WO2006/078256), CBH I de *Aspergillus fumigatus*, em particular Cel7A CBH1 divulgada como SEQ ID NO: 2 em WO2011/057140, e CBH II derivada de *Aspergillus fumigatus*, em particular aquela divulgada como SEQ ID NO: 4 em WO2013/028928.

**[0447]** Em outra modalidade, a composição de enzimas celulolíticas é uma composição de enzimas celulolíticas de *Trichoderma reesei*, compreendendo adicionalmente polipeptídeo AA9 de *Penicillium emersonii* (GH61A) tendo atividade intensificante celulolítica, em particular aquele divulgado em WO2011/041397, beta-glucosidase de *Aspergillus fumigatus* (p.ex., SEQ ID NO: 2 de WO2005/047499) ou sua variante com uma ou mais das, em particular todas as, seguintes substituições: F100D, S283G, N456E, F512Y; xilanase de *Aspergillus fumigatus* (p.ex., Xyl III em WO2006/078256), CBH I de *Aspergillus fumigatus*, em particular Cel7A CBH I divulgada como SEQ ID NO: 2 em WO2011/057140 e CBH II derivada de *Aspergillus fumigatus*, em particular aquela divulgada em WO2013/028928.

**[0448]** Em outra modalidade, a composição de enzimas celulolíticas é uma composição de enzimas celulolíticas de *Trichoderma reesei* compreendendo a CBH I (No. de Acesso do GENSEQP AZY49536 (WO2012/103293); uma CBH II (No. de Acesso do GENSEQP AZY49446 (WO2012/103288); uma variante de beta-glucosidase (No. de Acesso do GENSEQP AZU67153 (WO2012/44915)), em particular com uma ou mais das, em particular todas as, seguintes substituições: F100D, S283G, N456E, F512Y; e AA9 (polipeptídeo GH61) (No. de Acesso do GENSEQP BAL61510 (WO2013/028912)).

**[0449]** Em outra modalidade, a composição de enzimas celulolíticas é uma composição de enzimas celulolíticas de *Trichoderma reesei* compreendendo uma CBH I (No. de Acesso do GENSEQP AZY49536 (WO2012/103293)); uma CBH II (No. de Acesso do GENSEQP AZY49446

(WO2012/103288); uma xilanase GH10 (No. de Acesso do GENSEQP BAK46118 (WO2013/019827)); e uma beta-xilosidase (No. de Acesso do GENSEQP AZI04896 (WO2011/057140)).

**[0450]** Em outra modalidade, a composição de enzimas celulolíticas é uma composição de enzimas celulolíticas de *Trichoderma reesei* compreendendo uma CBH I (No. de Acesso do GENSEQP AZY49536 (WO2012/103293)); uma CBH II (No. de Acesso do GENSEQP AZY49446 (WO2012/103288)); e um AA9 (polipeptídeo GH61; No. de Acesso do GENSEQP BAL61510 (WO2013/028912)).

**[0451]** Em outra modalidade, a composição de enzimas celulolíticas é uma composição de enzimas celulolíticas de *Trichoderma reesei* compreendendo uma CBH I (No. de Acesso do GENSEQP AZY49536 (WO2012/103293)); uma CBH II (No. de Acesso do GENSEQP AZY49446 (WO2012/103288)), um AA9 (polipeptídeo GH61; No. de Acesso do GENSEQP BAL61510 (WO2013/028912)), e uma catalase (No. de Acesso do GENSEQP BAC11005 (WO2012/130120)).

**[0452]** Em uma modalidade, a composição de enzimas celulolíticas é uma composição de enzimas celulolíticas de *Trichoderma reesei* compreendendo uma CBH I (No. de Acesso do GENSEQP AZY49446 (WO2012/103288); uma CBH II (No. de Acesso do GENSEQP AZY49446 (WO2012/103288)), uma variante de beta-glucosidase (No. de Acesso do GENSEQP AZU67153 (WO2012/44915)), com uma ou mais das, em particular todas as, seguintes substituições: F100D, S283G, N456E, F512Y; um AA9 (polipeptídeo GH61; No. de Acesso do GENSEQP BAL61510 (WO2013/028912)), uma xilanase de GH10 (No. de Acesso do GENSEQP BAK46118 (WO2013/019827)), e uma beta-xilosidase (No. de Acesso do GENSEQP AZI04896 (WO2011/057140)).

**[0453]** Em uma modalidade, a composição celulolítica é uma preparação de enzimas celulolíticas de *Trichoderma reesei*

compreendendo uma EG I (No. de Acesso do P07981), EG II (No. de Acesso da EMBL M19373), CBH I (*supra*); CBH II (*supra*); variante de beta-glucosidase (*supra*) com as seguintes substituições: F100D, S283G, N456E, F512Y; um AA9 (polipeptídeo GH61; *supra*), xilanase GH10 (*supra*); e beta-xilosidase (*supra*).

**[0454]** Todas as composições de enzimas celulolíticas divulgadas em WO2013/028928 estão também contempladas e deste modo incorporadas por referência.

**[0455]** A composição de enzimas celulolíticas compreende ou pode compreender adicionalmente uma ou mais (várias) proteínas selecionadas do grupo consistindo em uma celulase, um polipeptídeo AA9 (*i.e.*, GH61) tendo atividade intensificante celulolítica, uma hemicelulase, uma expansina, uma esterase, uma lacase, uma enzima ligninolítica, uma pectinase, uma peroxidase, uma protease e uma swollenina.

**[0456]** Em uma modalidade, a composição de enzimas celulolíticas é uma composição de enzimas celulolíticas comercial. Exemplos de composições de enzimas celulolíticas comerciais adequadas para uso em um processo da invenção incluem: CELLIC<sup>®</sup> CTec (Novozymes A/S), CELLIC<sup>®</sup> CTec2 (Novozymes A/S), CELLIC<sup>®</sup> CTec3 (Novozymes A/S), CELLUCLAST<sup>™</sup> (Novozymes A/S), SPEZYME<sup>™</sup> CP (Genencor Int.), ACCELLERASE<sup>™</sup> 1000, ACCELLERASE 1500, ACCELLERASE<sup>™</sup> TRIO (DuPont), FILTRASE<sup>®</sup> NL (DSM); METHAPLUS<sup>®</sup> S/L 100 (DSM), ROHAMENT<sup>™</sup> 7069 W (Röhm GmbH) ou ALTERNAFUEL<sup>®</sup> CMAX3<sup>™</sup> (Dyadic International, Inc.). A composição de enzimas celulolíticas pode ser adicionada em uma quantidade eficaz de cerca de 0,001 a cerca de 5,0% em peso de sólidos, p.ex., cerca de 0,025 a cerca de 4,0% em peso de sólidos ou cerca de 0,005 a cerca de 2,0% em peso de sólidos.

**[0457]** Enzimas adicionais, e suas composições, podem ser encontradas em WO2011/153516 e WO2016/045569 (os conteúdos das quais são incorporados aqui).

**[0458]** Polinucleotídeos adicionais codificando enzimas celulolíticas adequadas podem ser obtidos de microrganismos de qualquer gênero, incluindo aqueles prontamente disponíveis dentro da base de dados UniProtKB ([www.uniprot.org](http://www.uniprot.org)).

**[0459]** As sequências codificantes de enzimas celulolíticas podem ser também usadas para conceber sondas de ácidos nucleicos para identificar e clonar DNA codificando enzimas celulolíticas de estirpes de diferentes gêneros ou espécies, como descrito *supra*.

**[0460]** Os polinucleotídeos codificando enzimas celulolíticas podem ser também identificados e obtidos de outras fontes incluindo microrganismos isolados a partir da natureza (p.ex., solo, adubos, água, *etc.*) ou amostras de DNA obtidas diretamente de materiais naturais (p.ex., solo, adubos, água, *etc.*) como descrito *supra*.

**[0461]** Técnicas usadas para isolar ou clonar polinucleotídeos codificando enzimas celulolíticas são descritas *supra*.

**[0462]** Em uma modalidade, a enzima celulolítica tem uma sequência de polipeptídeo maduro de pelo menos 60%, p.ex., pelo menos 65%, pelo menos 70%, pelo menos 75%, pelo menos 80%, pelo menos 85%, pelo menos 90%, pelo menos 91%, pelo menos 92%, pelo menos 93%, pelo menos 94%, pelo menos 95%, pelo menos 96%, pelo menos 97%, pelo menos 98%, pelo menos 99%, ou 100% de identidade de sequências com qualquer enzima celulolítica descrita ou referenciada aqui (p.ex., qualquer endoglucanase, celobiohidrolase ou beta-glucosidase). Em um aspecto, a enzima celulolítica tem uma sequência de polipeptídeo maduro em que a sequência difere em não mais do que dez aminoácidos, p.ex., em não mais do que cinco aminoácidos, em não mais do que quatro aminoácidos, em não



mais do que três aminoácidos, em não mais do que dois aminoácidos ou em um aminoácido de qualquer enzima celulolítica descrita ou referenciada aqui. Em uma modalidade, a enzima celulolítica tem uma sequência de polipeptídeo maduro que compreende a ou consiste na sequência de aminoácidos de qualquer enzima celulolítica descrita ou referenciada aqui, variante alélica ou um seu fragmento tendo atividade de enzima celulolítica. Em uma modalidade, a enzima celulolítica tem uma substituição, deleção e/ou inserção de aminoácidos de um ou mais (p.ex., dois, vários) aminoácidos. Em alguns aspectos, o número total de substituições, deleções e/ou inserções de aminoácidos é não mais do que 10, p.ex., não mais do que 9, 8, 7, 6, 5, 4, 3, 2 ou 1.

**[0463]** Em algumas modalidades, a enzima celulolítica tem pelo menos 20%, p.ex., pelo menos 40%, pelo menos 50%, pelo menos 60%, pelo menos 70%, pelo menos 80%, pelo menos 90%, pelo menos 95%, pelo menos 96%, pelo menos 97%, pelo menos 98%, pelo menos 99% ou 100% da atividade de enzima celulolítica de qualquer enzima celulolítica descrita ou referenciada aqui (p.ex., qualquer endoglucanase, celobiohidrolase ou beta-glucosidase) sob as mesmas condições.

**[0464]** Em uma modalidade, a sequência codificante de enzima celulolítica híbrida sob condições de pelo menos baixa estrigência, p.ex., condições de média estrigência, condições de média-elevada estrigência, condições de elevada estrigência ou condições de muito elevada estrigência com a fita complementar de comprimento total da sequência codificante de qualquer enzima celulolítica descrita ou referenciada aqui (p.ex., qualquer endoglucanase, celobiohidrolase ou beta-glucosidase). Em uma modalidade, a sequência codificante de enzima celulolítica tem pelo menos 65%, p.ex., pelo menos 70%, pelo menos 75%, pelo menos 80%, pelo menos 85%, pelo menos 85%, pelo menos 90%, pelo menos 91%, pelo menos 92%, pelo menos 93%, pelo menos 94%, pelo

menos 95%, pelo menos 96%, pelo menos 97%, pelo menos 98%, pelo menos 99% ou 100% de identidade de sequências com a sequência codificante de qualquer enzima celulolítica descrita ou referenciada aqui.

**[0465]** Em uma modalidade, o polinucleotídeo codificando a enzima celulolítica compreende a sequência codificante de qualquer enzima celulolítica descrita ou referenciada aqui (p.ex., qualquer endoglucanase, celobioidrolase ou beta-glucosidase). Em uma modalidade, o polinucleotídeo codificando a enzima celulolítica compreende uma subsequência da sequência codificante de qualquer enzima celulolítica descrita ou referenciada aqui, em que a subsequência codifica um polipeptídeo tendo atividade de enzima celulolítica. Em uma modalidade, o número de resíduos de nucleotídeos na subsequência é pelo menos 75%, p.ex., pelo menos 80%, 85%, 90%, ou 95% do número da sequência codificante referenciada.

**[0466]** A enzima celulolítica pode também incluir polipeptídeos fundidos ou polipeptídeos de fusão clivável, como descrito *supra*.

#### Metabolismo da xilose

**[0467]** Em um aspecto, o organismo fermentador (p.ex., célula de levedura) compreende adicionalmente um polinucleotídeo heterólogo codificando uma xilose isomerase (XI). A xilose isomerase pode ser qualquer xilose isomerase que seja adequada para as células hospedeiras e os métodos descritos aqui, tal como uma xilose isomerase ocorrendo naturalmente ou uma sua variante que retenha atividade de xilose isomerase. Em uma modalidade, a xilose isomerase está presente no citosol das células hospedeiras.

**[0468]** Em algumas modalidades, o organismo fermentador compreendendo um polinucleotídeo heterólogo codificando uma xilose isomerase tem um nível aumentado de atividade de xilose isomerase

em comparação com as células hospedeiras sem o polinucleotídeo heterólogo codificando a xilose isomerase, quando cultivados sob as mesmas condições. Em algumas modalidades, os organismos fermentadores têm um nível aumentado de atividade de xilose isomerase de pelo menos 5%, p.ex., pelo menos 10%, pelo menos 15%, pelo menos 20%, pelo menos 25%, pelo menos 50%, pelo menos 100%, pelo menos 150%, pelo menos 200%, pelo menos 300% ou pelo menos 500% em comparação com as células hospedeiras sem o polinucleotídeo heterólogo codificando a xilose isomerase, quando cultivados sob as mesmas condições.

**[0469]** Xilose isomerases exemplificativas que podem ser usadas com as células hospedeiras e métodos de uso descritos aqui incluem, mas não estão limitadas a, XIs do fungo *Piromyces sp.* (WO2003/062430) ou outras fontes (Madhavan *et al.*, 2009, *Appl Microbiol Biotechnol.* 82 (6), 1067-1078) foram expressas em células hospedeiras de *S. cerevisiae*. Ainda outras XIs adequadas para expressão em leveduras foram descritas em US 2012/0184020 (uma XI de *Ruminococcus flavefaciens*), WO2011/078262 (várias XIs de *Reticulitermes speratus* e *Mastotermes darwiniensis*) e WO2012/009272 (construtos e células fúngicos contendo uma XI de *Abiotrophia defectiva*). US 8,586,336 descreve uma célula hospedeira de *S. cerevisiae* expressando uma XI obtida por fluido de rúmen bovino (mostrada aqui como SEQ ID NO: 74).

**[0470]** Polinucleotídeos codificando xilose isomerases adequadas adicionais podem ser obtidos de microrganismos de qualquer gênero, incluindo aqueles prontamente disponíveis dentro da base de dados UniProtKB ([www.uniprot.org](http://www.uniprot.org)). Em uma modalidade, a xilose isomerase é uma xilose isomerase bacteriana, de levedura ou fúngica filamentosa, p.ex., obtida de qualquer um dos microrganismos descritos ou referenciados aqui, como descrito *supra*.

**[0471]** As sequências codificantes de xilose isomerases podem ser também usadas para conceber sondas de ácidos nucleicos para identificar e clonar DNA codificando xilose isomerases de estirpes de gêneros ou espécies diferentes, como descrito *supra*.

**[0472]** Os polinucleotídeos codificando xilose isomerases podem ser também identificados e obtidos de outras fontes incluindo microrganismos isolados a partir da natureza (p.ex., solo, adubos, água, *etc.*) ou amostras de DNA obtidas diretamente de materiais naturais (p.ex., solo, adubos, água, *etc.*) como descrito *supra*.

**[0473]** Técnicas usadas para isolar ou clonar polinucleotídeos codificando xilose isomerases são descritas *supra*.

**[0474]** Em uma modalidade, a xilose isomerase tem uma sequência de polipeptídeo maduro de pelo menos 60%, p.ex., pelo menos 65%, pelo menos 70%, pelo menos 75%, pelo menos 80%, pelo menos 85%, pelo menos 90%, pelo menos 91%, pelo menos 92%, pelo menos 93%, pelo menos 94%, pelo menos 95%, pelo menos 96%, pelo menos 97%, pelo menos 98%, pelo menos 99% ou 100% de identidade de sequências com qualquer xilose isomerase descrita ou referenciada aqui (p.ex., a xilose isomerase de SEQ ID NO: 74). Em um aspecto, a xilose isomerase tem uma sequência de polipeptídeo maduro que difere em não mais do que dez aminoácidos, p.ex., em não mais do que cinco aminoácidos, em não mais do que quatro aminoácidos, em não mais do que três aminoácidos, em não mais do que dois aminoácidos ou em um aminoácido de qualquer xilose isomerase descrita ou referenciada aqui (p.ex., a xilose isomerase de SEQ ID NO: 74). Em uma modalidade, a xilose isomerase tem uma sequência de polipeptídeo maduro que compreende a ou consiste na sequência de aminoácidos de qualquer xilose isomerase descrita ou referenciada aqui (p.ex., a xilose isomerase de SEQ ID NO: 74), variante alélica ou um seu fragmento tendo atividade de xilose isomerase. Em uma modalidade, a sequência codificante

de xilose isomerase tem um substituição, deleção e/ou inserção de aminoácidos de um ou mais (p.ex., dois, vários) aminoácidos. Em alguns aspectos, o número total de substituições, deleções e/ou inserções de aminoácidos é não mais do que 10, p.ex., não mais do que 9, 8, 7, 6, 5, 4, 3, 2 ou 1.

**[0475]** Em algumas modalidades, a xilose isomerase tem pelo menos 20%, p.ex., pelo menos 40%, pelo menos 50%, pelo menos 60%, pelo menos 70%, pelo menos 80%, pelo menos 90%, pelo menos 95%, pelo menos 96%, pelo menos 97%, pelo menos 98%, pelo menos 99% ou 100% da atividade de xilose isomerase de qualquer xilose isomerase descrita ou referenciada aqui (p.ex., a xilose isomerase de SEQ ID NO: 74) sob as mesmas condições.

**[0476]** Em uma modalidade, a sequência codificante da xilose isomerase híbrida sob condições de pelo menos baixa estringência, p.ex., condições de média estringência, condições de média-elevada estringência, condições de elevada estringência ou condições de muito elevada estringência com a fita complementar de comprimento total da sequência codificante de qualquer xilose isomerase descrita ou referenciada aqui (p.ex., a xilose isomerase de SEQ ID NO: 74). Em uma modalidade, a sequência codificante de xilose isomerase tem pelo menos 65%, p.ex., pelo menos 70%, pelo menos 75%, pelo menos 80%, pelo menos 85%, pelo menos 85%, pelo menos 90%, pelo menos 91%, pelo menos 92%, pelo menos 93%, pelo menos 94%, pelo menos 95%, pelo menos 96%, pelo menos 97%, pelo menos 98%, pelo menos 99% ou 100% de identidade de sequências com a sequência codificante de qualquer xilose isomerase descrita ou referenciada aqui (p.ex., a xilose isomerase de SEQ ID NO: 74).

**[0477]** Em uma modalidade, o polinucleotídeo heterólogo codificando a xilose isomerase compreende a sequência codificante de qualquer xilose isomerase descrita ou referenciada aqui (p.ex., a xilose

isomerase de SEQ ID NO: 74). Em uma modalidade, o polinucleotídeo heterólogo codificando a xilose isomerase compreende uma subsequência da sequência codificante de qualquer xilose isomerase descrita ou referenciada aqui, em que a subsequência codifica um polipeptídeo tendo atividade de xilose isomerase. Em uma modalidade, o número de resíduos de nucleotídeos na subsequência é pelo menos 75%, p.ex., pelo menos 80%, 85%, 90%, ou 95% do número da sequência codificante referenciada.

**[0478]** As xilose isomerases podem também incluir polipeptídeos fundidos ou polipeptídeos de fusão clivável, como descrito *supra*.

**[0479]** Em um aspecto, o organismo fermentador (p.ex., célula de levedura) compreende adicionalmente um polinucleotídeo heterólogo codificando uma xilulocinase (XK). Uma xilulocinase, como usada aqui, proporciona atividade enzimática para converter D-xilulose em xilulose 5-fosfato. A xilulocinase pode ser qualquer xilulocinase que seja adequada para as células hospedeiras e os métodos descritos aqui, tal como uma xilulocinase ocorrendo naturalmente ou uma sua variante que retenha atividade de xilulocinase. Em uma modalidade, a xilulocinase está presente no citosol das células hospedeiras.

**[0480]** Em algumas modalidades, os organismos fermentadores compreendendo um polinucleotídeo heterólogo codificando uma xilulocinase têm um nível aumentado de atividade de xilulocinase em comparação com as células hospedeiras sem o polinucleotídeo heterólogo codificando a xilulocinase, quando cultivados sob as mesmas condições. Em algumas modalidades, as células hospedeiras têm um nível aumentado de atividade de xilose isomerase de pelo menos 5%, p.ex., pelo menos 10%, pelo menos 15%, pelo menos 20%, pelo menos 25%, pelo menos 50%, pelo menos 100%, pelo menos 150%, pelo menos 200%, pelo menos 300% ou pelo menos 500% em comparação com as células hospedeiras sem o

polinucleotídeo heterólogo codificando a xilulocinase, quando cultivados sob as mesmas condições.

**[0481]** Xilulocinases exemplificativas que podem ser usadas com os organismos fermentadores e métodos de uso descritos aqui incluem a, mas não estão limitados à, xilulocinase de *Saccharomyces cerevisiae* de SEQ ID NO: 75. Polinucleotídeos codificando xilulocinases adequadas adicionais podem ser obtidos de microrganismos de qualquer gênero, incluindo aqueles prontamente disponíveis dentro da base de dados UniProtKB ([www.uniprot.org](http://www.uniprot.org)). Em uma modalidade, a xilulocinase é uma xilulocinase bacteriana, de levedura ou fúngica filamentosa, p.ex., obtida de qualquer um dos microrganismos descritos ou referenciados aqui, como descrito *supra*.

**[0482]** As sequências codificantes de xilulocinases podem ser também usadas para conceber sondas de ácidos nucleicos para identificar e clonar DNA codificando xilulocinases de estirpes de gêneros ou espécies diferentes, como descrito *supra*.

**[0483]** Os polinucleotídeos codificando xilulocinases podem ser também identificados e obtidos de outras fontes incluindo microrganismos isolados a partir da natureza (p.ex., solo, adubos, água, *etc.*) ou amostras de DNA obtidas diretamente de materiais naturais (p.ex., solo, adubos, água, *etc.*) como descrito *supra*.

**[0484]** Técnicas usadas para isolar ou clonar polinucleotídeos codificando xilulocinases são descritas *supra*.

**[0485]** Em uma modalidade, a xilulocinase tem uma sequência de polipeptídeo maduro de pelo menos 60%, p.ex. pelo menos 65%, pelo menos 70%, pelo menos 75%, pelo menos 80%, pelo menos 85%, pelo menos 90%, pelo menos 91%, pelo menos 92%, pelo menos 93%, pelo menos 94%, pelo menos 95%, pelo menos 96%, pelo menos 97%, pelo menos 98%, pelo menos 99% ou 100% de identidade de sequências com

qualquer xilulocinase descrita ou referenciada aqui (p.ex., a xilulocinase de *Saccharomyces cerevisiae* de SEQ ID NO: 75). Em uma modalidade, a xilulocinase tem uma sequência de polipeptídeo maduro que difere em não mais do que dez aminoácidos, p.ex., em não mais do que cinco aminoácidos, em não mais do que quatro aminoácidos, em não mais do que três aminoácidos, em não mais do que dois aminoácidos ou em um aminoácido de qualquer xilulocinase descrita ou referenciada aqui (p.ex., a xilulocinase de *Saccharomyces cerevisiae* de SEQ ID NO: 75). Em uma modalidade, a xilulocinase tem uma sequência de polipeptídeo maduro que compreende a ou consiste na sequência de aminoácidos de qualquer xilulocinase descrita ou referenciada aqui (p.ex., a xilulocinase de *Saccharomyces cerevisiae* de SEQ ID NO: 75), variante alélica ou um seu fragmento tendo atividade de xilulocinase. Em uma modalidade, a xilulocinase tem uma substituição, deleção e/ou inserção de aminoácidos de um ou mais (p.ex., dois, vários) aminoácidos. Em alguns aspectos, o número total de substituições, deleções e/ou inserções de aminoácidos é não mais do que 10, p.ex., não mais do que 9, 8, 7, 6, 5, 4, 3, 2 ou 1.

**[0486]** Em algumas modalidades, a xilulocinase tem pelo menos 20%, p.ex., pelo menos 40%, pelo menos 50%, pelo menos 60%, pelo menos 70%, pelo menos 80%, pelo menos 90%, pelo menos 95%, pelo menos 96%, pelo menos 97%, pelo menos 98%, pelo menos 99% ou 100% da atividade de xilulocinase de qualquer xilulocinase descrita ou referenciada aqui (p.ex., a xilulocinase de *Saccharomyces cerevisiae* de SEQ ID NO: 75) sob as mesmas condições.

**[0487]** Em uma modalidade, a sequência codificante de xilulocinase híbrida sob condições de pelo menos baixa estrigência, p.ex., condições de média estrigência, condições de média-elevada estrigência, condições de elevada estrigência ou condições de muito elevada estrigência com a fita complementar de comprimento total da sequência



codificante de qualquer xilulocinase descrita ou referenciada aqui (p.ex., a xilulocinase de *Saccharomyces cerevisiae* de SEQ ID NO: 75). Em uma modalidade, a sequência codificante de xilulocinase tem pelo menos 65%, p.ex., pelo menos 70%, pelo menos 75%, pelo menos 80%, pelo menos 85%, pelo menos 85%, pelo menos 90%, pelo menos 91%, pelo menos 92%, pelo menos 93%, pelo menos 94%, pelo menos 95%, pelo menos 96%, pelo menos 97%, pelo menos 98%, pelo menos 99% ou 100% de identidade de sequências com a sequência codificante de qualquer xilulocinase descrita ou referenciada aqui (p.ex., a xilulocinase de *Saccharomyces cerevisiae* de SEQ ID NO: 75).

**[0488]** Em uma modalidade, o polinucleotídeo heterólogo codificando a xilulocinase compreende a sequência codificante de qualquer xilulocinase descrita ou referenciada aqui (p.ex., a xilulocinase de *Saccharomyces cerevisiae* de SEQ ID NO: 75). Em uma modalidade, o polinucleotídeo heterólogo codificando a xilulocinase compreende uma subsequência da sequência codificante de qualquer xilulocinase descrita ou referenciada aqui, em que a subsequência codifica um polipeptídeo tendo atividade de xilulocinase. Em uma modalidade, o número de resíduos de nucleotídeos na subsequência é pelo menos 75%, p.ex., pelo menos 80%, 85%, 90%, ou 95% do número da sequência codificante referenciada.

**[0489]** As xilulocinases podem também incluir polipeptídeos fundidos ou polipeptídeos de fusão clivável, como descrito *supra*.

**[0490]** Em um aspecto, o organismo fermentador (p.ex., célula de levedura) compreende adicionalmente um polinucleotídeo heterólogo codificando uma ribulose 5 fosfato 3-epimerase (RPE1). Uma ribulose 5 fosfato 3-epimerase, como usada aqui, proporciona atividade enzimática para converter L-ribulose 5-fosfato em L-xilulose 5-fosfato (EC 5.1.3.22). A RPE1 pode ser qualquer RPE1 que seja adequada para as

células hospedeiras e os métodos descritos aqui, tal como uma RPE1 ocorrendo naturalmente ou uma sua variante que retenha atividade de RPE1. Em uma modalidade, a RPE1 está presente no citosol das células hospedeiras.

**[0491]** Em uma modalidade, a célula recombinante compreende um polinucleotídeo heterólogo codificando uma ribulose 5-fosfato 3-epimerase (RPE1), em que a RPE1 é uma RPE1 de *Saccharomyces cerevisiae* ou uma RPE1 tendo pelo menos 60%, p.ex., pelo menos 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, ou 100% de identidade de sequências com uma RPE1 de *Saccharomyces cerevisiae*.

**[0492]** Em um aspecto, o organismo fermentador (p.ex., célula de levedura) compreende adicionalmente um polinucleotídeo heterólogo codificando uma ribulose 5-fosfato isomerase (RKI1). Uma ribulose 5-fosfato isomerase, como usada aqui, proporciona atividade enzimática para converter ribose-5-fosfato em ribulose 5-fosfato. A RKI1 pode ser qualquer RKI1 que seja adequada para as células hospedeiras e os métodos descritos aqui, tal como uma RKI1 ocorrendo naturalmente ou uma sua variante que retenha atividade de RKI1. Em uma modalidade, a RKI1 está presente no citosol das células hospedeiras.

**[0493]** Em uma modalidade, o organismo fermentador compreende um polinucleotídeo heterólogo codificando uma ribulose 5-fosfato isomerase (RKI1), em que a RKI1 é uma RKI1 de *Saccharomyces cerevisiae* ou uma RKI1 tendo uma sequência de polipeptídeo maduro de pelo menos 60%, p.ex., pelo menos 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, ou 100% de identidade de sequências com uma RKI1 de *Saccharomyces cerevisiae*.

**[0494]** Em um aspecto, o organismo fermentador (p.ex., célula de levedura) compreende adicionalmente um polinucleotídeo heterólogo codificando uma transcetolase (TKL1). A TKL1 pode ser qualquer

TKL1 que seja adequada para as células hospedeiras e os métodos descritos aqui, tal como uma TKL1 ocorrendo naturalmente ou uma sua variante que retenha atividade de TKL1. Em uma modalidade, a TKL1 está presente no citosol das células hospedeiras.

**[0495]** Em uma modalidade, o organismo fermentador compreende um polinucleotídeo heterólogo codificando uma transcetolase (TKL1), em que a TKL1 é uma TKL1 de *Saccharomyces cerevisiae* ou uma TKL1 tendo uma sequência de polipeptídeo maduro de pelo menos 60%, p.ex., pelo menos 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, ou 100% de identidade de sequências com uma TKL1 de *Saccharomyces cerevisiae*.

**[0496]** Em um aspecto, o organismo fermentador (p.ex., célula de levedura) compreende adicionalmente um polinucleotídeo heterólogo codificando uma transaldolase (TAL1). A TAL1 pode ser qualquer TAL1 que seja adequada para as células hospedeiras e os métodos descritos aqui, tal como uma TAL1 ocorrendo naturalmente ou uma sua variante que retenha atividade de TAL1. Em uma modalidade, a TAL1 está presente no citosol das células hospedeiras.

**[0497]** Em uma modalidade, o organismo fermentador compreende um polinucleotídeo heterólogo codificando uma transcetolase (TAL1), em que a TAL1 é uma TAL1 de *Saccharomyces cerevisiae* ou uma TAL1 tendo uma sequência de polipeptídeo maduro de pelo menos 60%, p.ex., pelo menos 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, ou 100% de identidade de sequências com uma TAL1 de *Saccharomyces cerevisiae*.

#### Produtos de fermentação

**[0498]** Um produto de fermentação pode ser qualquer substância derivada da fermentação. O produto de fermentação pode ser, sem limitação, um álcool (p.ex., arabinitol, *n*-butanol, isobutanol, etanol,

glicerol, metanol, etilenoglicol, 1,3-propanodiol [propilenoglicol], butanodiol, glicerina, sorbitol e xilitol); um alcano (p.ex., pentano, hexano, heptano, octano, nonano, decano, undecano e dodecano), um cicloalcano (p.ex., ciclopentano, ciclo-hexano, ciclo-heptano e ciclo-octano), um alqueno (p.ex., penteno, hexeno, hepteno e octeno); um aminoácido (p.ex., ácido aspártico, ácido glutâmico, glicina, lisina, serina e treonina); um gás (p.ex., metano, hidrogênio (H<sub>2</sub>), dióxido de carbono (CO<sub>2</sub>) e monóxido de carbono (CO)); isopreno; uma cetona (p.ex., acetona); um ácido orgânico (p.ex., ácido acético, ácido acetônico, ácido adípico, ácido ascórbico, ácido cítrico, ácido 2,5-diceto-D-glucônico, ácido fórmico, ácido fumárico, ácido glucárico, ácido glucônico, ácido glucurônico, ácido glutárico, ácido 3-hidroxipropiônico, ácido itacônico, ácido láctico, ácido málico, ácido malônico, ácido oxálico, ácido oxaloacético, ácido propiônico, ácido succínico e ácido xilônico); e policetídeo.

**[0499]** Em um aspecto, o produto de fermentação é um álcool. O termo “álcool” engloba uma substância que contém uma ou mais frações de hidroxila. O álcool pode ser, mas não está limitado a, *n*-butanol, *isobutanol*, etanol, metanol, arabinitol, butanodiol, etilenoglicol, glicerina, glicerol, 1,3-propanodiol, sorbitol, xilitol. Ver, por exemplo, Gong *et al.*, 1999, *Ethanol production from renewable resources*, em *Advances in Biochemical Engineering/Biotechnology*, Scheper, T., ed., Springer-Verlag Berlin Heidelberg, Alemanha, 65: 207-241; Silveira e Jonas, 2002, *Appl. Microbiol. Biotechnol.* 59: 400-408; Nigam e Singh, 1995, *Process Biochemistry* 30 (2): 117-124; Ezeji *et al.*, 2003, *World Journal of Microbiology and Biotechnology* 19 (6): 595-603. Em uma modalidade, o produto de fermentação é etanol.

**[0500]** Em outro aspecto, o produto de fermentação é um alcano. O alcano pode ser um alcano não ramificado ou ramificado. O alcano pode ser, mas não está limitado a, pentano, hexano, heptano, octano, nonano, decano, undecano ou dodecano.

**[0501]** Em outro aspecto, o produto de fermentação é um cicloalcano. O cicloalcano pode ser, mas não está limitado a, ciclopentano, ciclo-hexano, ciclo-heptano ou ciclo-octano.

**[0502]** Em outro aspecto, o produto de fermentação é um alqueno. O alqueno pode ser um alqueno não ramificado ou ramificado. O alqueno pode ser, mas não está limitado a, penteno, hexeno, hepteno ou octeno.

**[0503]** Em outro aspecto, o produto de fermentação é um aminoácido. O ácido orgânico pode ser, mas não está limitado a, ácido aspártico, ácido glutâmico, glicina, lisina, serina ou treonina. Ver, por exemplo, Richard e Margaritis, 2004, *Biotechnology and Bioengineering* 87 (4): 501-515.

**[0504]** Em outro aspecto, o produto de fermentação é um gás. O gás pode ser, mas não está limitado a, metano, H<sub>2</sub>, CO<sub>2</sub> ou CO. Ver, por exemplo, Kataoka *et al.*, 1997, *Water Science and Technology* 36 (6-7): 41-47; e Gunaseelan, 1997, *Biomass and Bioenergy* 13 (1-2): 83-114.

**[0505]** Em outro aspecto, o produto de fermentação é isopreno.

**[0506]** Em outro aspecto, o produto de fermentação é uma cetona. O termo “cetona” engloba uma substância que contém uma ou mais frações de cetona. A cetona pode ser, mas não está limitada a, acetona.

**[0507]** Em outro aspecto, o produto de fermentação é um ácido orgânico. O ácido orgânico pode ser, mas não está limitado a, ácido acético, ácido acetônico, ácido adípico, ácido ascórbico, ácido cítrico, ácido 2,5-diceto-D-glucônico, ácido fórmico, ácido fumárico, ácido glucárico, ácido glucônico, ácido glucurônico, ácido glutárico, ácido 3-hidroxiopropiônico, ácido itacônico, ácido láctico, ácido málico, ácido malônico, ácido oxálico, ácido propiônico, ácido succínico ou ácido xilônico. Ver, por exemplo, Chen e Lee, 1997, *Appl. Biochem. Biotechnol.* 63-65: 435-448.

**[0508]** Em outro aspecto, o produto de fermentação é policetídeo.

#### Recuperação

**[0509]** O produto de fermentação, p.ex., o etanol, pode ser opcionalmente recuperado a partir do meio de fermentação usando qualquer método conhecido na técnica incluindo, mas não se limitando a, cromatografia, procedimentos eletroforéticos, solubilidade diferencial, destilação ou extração. Por exemplo, o álcool é separado a partir do material celulósico fermentado e purificado por métodos convencionais de destilação. Pode ser obtido etanol com uma pureza até cerca de 96% por vol., que pode ser usado como, por exemplo, etanol para combustível, etanol potável, *i.e.*, bebidas alcoólicas neutras potáveis, ou etanol industrial.

**[0510]** Em alguns aspectos dos métodos, o produto de fermentação após ser recuperado é substancialmente puro. No que diz respeito aos métodos aqui, “substancialmente puro” se refere a uma preparação recuperada que contém não mais do que 15% de impurezas, em que impureza se refere a outros compostos sem ser o produto de fermentação (p.ex., etanol). Em uma variação é proporcionada uma preparação substancialmente pura em que a preparação contém não mais do que 25% de impurezas ou não mais do que 20% de impurezas ou não mais do que 10% de impurezas ou não mais do que 5% de impurezas ou não mais do que 3% de impurezas ou não mais do que 1% de impurezas ou não mais do que 0,5% de impurezas.

**[0511]** Ensaios adequados para testar a produção de etanol e contaminantes, e o consumo de açúcar, podem ser realizados usando métodos conhecidos na técnica. Por exemplo, o produto de etanol, bem como outros compostos orgânicos, pode ser analisado por métodos tais como HPLC (Cromatografia Líquida de Elevado Desempenho), GC-MS (Espectroscopia de Massa - Cromatografia Gasosa) e LC-MS

(Espectroscopia de Massa - Cromatografia Líquida) ou outros métodos analíticos adequados usando procedimentos de rotina bem conhecidos na técnica. A liberação de etanol no caldo de fermentação pode ser também testada com o sobrenadante da cultura. Os subprodutos e o açúcar residual no meio de fermentação (p.ex., glucose ou xilose) podem ser quantificados por HPLC usando, por exemplo, um detector de índice de refração para glucose e álcoois e um detector de UV para ácidos orgânicos (Lin *et al.*, *Biotechnol. Bioeng.* 90: 775-779 (2005)) ou usando outro ensaio e métodos de detecção adequados bem conhecidos na técnica.

**A invenção pode ser adicionalmente descrita nos seguintes parágrafos numerados:**

**[0512]** Parágrafo [1]. Um método de produção de um produto de fermentação a partir de material contendo amido ou contendo celulose compreendendo:

- (a) sacarificação do material contendo amido ou contendo celulose; e
- (b) fermentação do material sacarificado do passo (a) com um organismo fermentador;

em que o organismo fermentador compreende um polinucleotídeo heterólogo codificando uma alfa-amilase e um polinucleotídeo heterólogo codificando uma trealase.

**[0513]** Parágrafo [2]. O método do parágrafo [1], em que a alfa-amilase tem uma sequência de polipeptídeo maduro com 60%, p.ex., pelo menos 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, ou 100% de identidade de sequências com a sequência de aminoácidos de qualquer uma de SEQ ID NOs: 76-101, 121-174 e 231.

**[0514]** Parágrafo [3]. O método do parágrafo [1] ou [2], em que em que o polinucleotídeo heterólogo codifica uma alfa-amilase tendo uma sequência de polipeptídeo maduro que difere em não mais do que dez aminoácidos, p.ex., em não mais do que cinco aminoácidos, em não mais do

que quatro aminoácidos, em não mais do que três aminoácidos, em não mais do que dois aminoácidos ou em um aminoácido da sequência de aminoácidos de qualquer uma de SEQ ID NOs: 76-101, 121-174 e 231.

**[0515]** Parágrafo [4]. O método de qualquer um dos parágrafos [1]-[3], em que o polinucleotídeo heterólogo codifica uma alfa-amilase tendo uma sequência de polipeptídeo maduro compreendendo a ou consistindo na sequência de aminoácidos de qualquer uma de SEQ ID NOs: 76-101, 121-174 e 231.

**[0516]** Parágrafo [5]. O método de qualquer um dos parágrafos [1]-[4], em que a trealase tem uma sequência de polipeptídeo maduro com 60%, p.ex., pelo menos 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, ou 100% de identidade de sequências com a sequência de aminoácidos de qualquer uma de SEQ ID NOs: 175-226.

**[0517]** Parágrafo [6]. O método de qualquer um dos parágrafos [1]-[5], em que o polinucleotídeo heterólogo codifica uma trealase tendo uma sequência de polipeptídeo maduro que difere em não mais do que dez aminoácidos, p.ex., em não mais do que cinco aminoácidos, em não mais do que quatro aminoácidos, em não mais do que três aminoácidos, em não mais do que dois aminoácidos ou em um aminoácido da sequência de aminoácidos de qualquer uma de SEQ ID NOs: 175-226.

**[0518]** Parágrafo [7]. O método de qualquer um dos parágrafos [1]-[6], em que o polinucleotídeo heterólogo codifica uma trealase tendo uma sequência de polipeptídeo maduro compreendendo a ou consistindo na sequência de aminoácidos de qualquer uma de SEQ ID NOs: 175-226.

**[0519]** Parágrafo [8]. O método de qualquer um dos parágrafos [1]-[7], em que a sacarificação do passo (a) ocorre em um material contendo amido, e em que o material contendo amido é amido gelatinizado ou não gelatinizado.



**[0520]** Parágrafo [9]. O método do parágrafo [8], compreendendo liquefação do material contendo amido por contato do material com uma alfa-amilase antes da sacarificação.

**[0521]** Parágrafo [10]. O método do parágrafo [9], em que a liquefação do material contendo amido e/ou a sacarificação do material contendo amido são conduzidas na presença de protease exogenamente adicionada.

**[0522]** Parágrafo [11]. O método de qualquer um dos parágrafos [1]-[10], em que a fermentação é realizada sob condições de nitrogênio reduzido (p.ex., menos do que 1000 ppm de ureia ou hidróxido de amônio, tal como menos do que 750 ppm, menos do que 500 ppm, menos do que 400 ppm, menos do que 300 ppm, menos do que 250 ppm, menos do que 200 ppm, menos do que 150 ppm, menos do que 100 ppm, menos do que 75 ppm, menos do que 50 ppm, menos do que 25 ppm ou menos do que 10 ppm).

**[0523]** Parágrafo [12]. O método de qualquer um dos parágrafos [1]-[11], em que a fermentação e a sacarificação são realizadas simultaneamente em uma sacarificação e fermentação simultâneas (SSF).

**[0524]** Parágrafo [13]. O método de qualquer um dos parágrafos [1]-[11], em que a fermentação e a sacarificação são realizadas sequencialmente (SHF).

**[0525]** Parágrafo [14]. O método de qualquer um dos parágrafos [1]-[13], compreendendo recuperação do produto de fermentação a partir da fermentação.

**[0526]** Parágrafo [15]. O método do parágrafo [14], em que a recuperação do produto de fermentação a partir da fermentação compreende destilação.

**[0527]** Parágrafo [16]. O método de qualquer um dos parágrafos [1]-[15], em que o produto de fermentação é etanol.

**[0528]** Parágrafo [17]. O método de qualquer um dos parágrafos [1]-[16], em que o organismo fermentador compreende um polinucleotídeo heterólogo codificando uma glucoamilase.

**[0529]** Parágrafo [18]. O método do parágrafo [17], em que a glucoamilase tem uma sequência de polipeptídeo maduro com 60%, p.ex., pelo menos 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, ou 100% de identidade de sequências com a sequência de aminoácidos de uma glicoamilase de *Pycnopus* (p.ex., uma glucoamilase de *Pycnopus sanguineus* de SEQ ID NO: 229), uma glucoamilase de *Gloeophyllum* (p.ex., uma *Gloeophyllum sepiarium* de SEQ ID NO: 8) ou uma glucoamilase de qualquer uma de SEQ ID NOs: 102-113 (p.ex., uma glucoamilase de *Saccharomycopsis fibuligera* de SEQ ID NO: 103 ou 104 ou uma glucoamilase de *Trichoderma reesei* de SEQ ID NO: 230).

**[0530]** Parágrafo [19]. O método de qualquer um dos parágrafos [1]-[18], em que o organismo fermentador compreende um polinucleotídeo heterólogo codificando uma protease.

**[0531]** Parágrafo [20]. O método do parágrafo [19], em que a protease tem uma sequência de polipeptídeo maduro de pelo menos 60%, p.ex., pelo menos 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, ou 100% de identidade de sequências com a sequência de aminoácidos de qualquer uma de SEQ ID NOs: 9-73 (p.ex., qualquer uma de SEQ ID NOs: 9, 14, 16, 21, 22, 33, 41, 45, 61, 62, 66, 67 e 69; tal como qualquer uma de SEQ NOs: 9, 14, 16, e 69).

**[0532]** Parágrafo [21]. O método de qualquer um dos parágrafos [1]-[20], em que a sacarificação do passo (a) ocorre em um material contendo celulose, e em que o material contendo celulose é pré-tratado.

**[0533]** Parágrafo [22]. O método do parágrafo [21], em que o pré-tratamento é um pré-tratamento com ácido diluído.

**[0534]** Parágrafo [23]. O método de qualquer um dos parágrafos [1]-[20], em que sacarificação ocorre em um material contendo celulose, e em que a composição de enzimas compreende uma ou mais enzimas selecionadas de uma celulase, um polipeptídeo AA9, uma hemicelulase, uma CIP, uma esterase, uma expansina, uma enzima ligninolítica, uma oxirredutase, uma pectinase, uma protease e uma swollenina.

**[0535]** Parágrafo [24]. O método do parágrafo [23], em que a celulase é uma ou mais enzimas selecionadas de uma endoglucanase, uma celobioidrolase e uma beta-glucosidase.

**[0536]** Parágrafo [25]. O método do parágrafo [23] ou [24], em que a hemicelulase é uma ou mais enzimas selecionadas de uma xilanase, uma acetilxilana esterase, uma feruloil esterase, uma arabinofuranosidase, uma xilosidase e uma glucuronidase.

**[0537]** Parágrafo [26]. O método de qualquer um dos parágrafos [1]-[25], em que o organismo fermentador é uma célula de *Saccharomyces*, *Rhodotorula*, *Schizosaccharomyces*, *Kluyveromyces*, *Pichia*, *Hansenula*, *Rhodosporidium*, *Candida*, *Yarrowia*, *Lipomyces*, *Cryptococcus*, ou *Dekkera sp.*

**[0538]** Parágrafo [27]. O método de qualquer um dos parágrafos [1]-[26], em que o organismo fermentador é uma célula de *Saccharomyces cerevisiae*.

**[0539]** Parágrafo [28]. Uma célula de levedura recombinante compreendendo um polinucleotídeo heterólogo codificando uma alfa-amilase e um polinucleotídeo heterólogo codificando uma trealase.

**[0540]** Parágrafo [29]. A célula de levedura recombinante do parágrafo [28], em que a alfa-amilase tem uma sequência de polipeptídeo maduro com 60%, p.ex., pelo menos 65%, 70%, 75%, 80%, 85%, 90%, 95%,

97%, 98%, 99%, ou 100% de identidade de sequências com a sequência de aminoácidos de qualquer uma de SEQ ID NOs: 76-101, 121-174 e 231.

Parágrafo [30]. A célula de levedura recombinante do parágrafo [28] ou [29], em que em que o polinucleotídeo heterólogo codifica uma alfa-amilase tendo uma sequência de polipeptídeo maduro que difere em não mais do que dez aminoácidos, p.ex., em não mais do que cinco aminoácidos, em não mais do que quatro aminoácidos, em não mais do que três aminoácidos, em não mais do que dois aminoácidos ou em um aminoácido da sequência de aminoácidos de qualquer uma de SEQ ID NOs: 76-101, 121-174 e 231.

**[0541]** Parágrafo [31]. A célula de levedura recombinante de qualquer um dos parágrafos [28]-[30], em que o polinucleotídeo heterólogo codifica uma alfa-amilase tendo uma sequência de polipeptídeo maduro compreendendo a ou consistindo na sequência de aminoácidos de qualquer uma de SEQ ID NOs: 76-101, 121-174 e 231.

**[0542]** Parágrafo [32]. A célula de levedura recombinante de qualquer um dos parágrafos [28]-[31], em que a trealase tem uma sequência de polipeptídeo maduro com 60%, p.ex., pelo menos 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, ou 100% de identidade de sequências com a sequência de aminoácidos de qualquer uma de SEQ ID NOs: 175-226.

**[0543]** Parágrafo [33]. A célula de levedura recombinante de qualquer um dos parágrafos [28]-[32], em que o polinucleotídeo heterólogo codifica uma trealase tendo uma sequência de polipeptídeo maduro que difere em não mais do que dez aminoácidos, p.ex., em não mais do que cinco aminoácidos, em não mais do que quatro aminoácidos, em não mais do que três aminoácidos, em não mais do que dois aminoácidos ou em um aminoácido da sequência de aminoácidos de qualquer uma de SEQ ID NOs: 175-226.

**[0544]** Parágrafo [34]. A célula de levedura recombinante de qualquer um dos parágrafos [28]-[33], em que o polinucleotídeo heterólogo codifica uma trealase tendo uma sequência de polipeptídeo maduro compreendendo a ou consistindo na sequência de aminoácidos de qualquer uma de SEQ ID NOs: 175-226.

**[0545]** Parágrafo [35]. A célula de levedura recombinante de qualquer um dos parágrafos [28]-[34], em que o organismo fermentador compreende um polinucleotídeo heterólogo codificando uma glucoamilase.

**[0546]** Parágrafo [36]. A célula de levedura recombinante do parágrafo [35], em que a glucoamilase tem uma sequência de polipeptídeo maduro com 60%, p.ex., pelo menos 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, ou 100% de identidade de sequências com a sequência de aminoácidos de uma glicoamilase de *Pycnopus* (p.ex., uma glucoamilase de *Pycnopus sanguineus* de SEQ ID NO: 229), uma glucoamilase de *Gloeophyllum* (p.ex., uma *Gloeophyllum sepiarium* de SEQ ID NO: 8) ou uma glucoamilase de qualquer uma de SEQ ID NOs: 102-113 (p.ex., uma glucoamilase de *Saccharomyces fibuligera* de SEQ ID NO: 103 ou 104 ou uma glucoamilase de *Trichoderma reesei* de SEQ ID NO: 230). Parágrafo [37]. A célula de levedura recombinante de qualquer um dos parágrafos [28]-[36], em que o organismo fermentador compreende um polinucleotídeo heterólogo codificando uma protease.

**[0547]** Parágrafo [38]. A célula de levedura recombinante do parágrafo [37], em que a protease tem uma sequência de polipeptídeo maduro de pelo menos 60%, p.ex., pelo menos 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, ou 100% de identidade de sequências com a sequência de aminoácidos de qualquer uma de SEQ ID NOs: 9-73 (p.ex., qualquer uma de SEQ ID NOs: 9, 14, 16, 21, 22, 33, 41, 45, 61, 62, 66, 67 e 69; tal como qualquer uma de SEQ NOs: 9, 14, 16, e 69).

**[0548]** Parágrafo [39]. A levedura recombinante de qualquer um dos parágrafos [28]-[38], em que a célula é uma célula de *Saccharomyces*, *Rhodotorula*, *Schizosaccharomyces*, *Kluyveromyces*, *Pichia*, *Hansenula*, *Rhodosporidium*, *Candida*, *Yarrowia*, *Lipomyces*, *Cryptococcus*, ou *Dekkera sp.*

**[0549]** Parágrafo [40]. A levedura recombinante do parágrafo [39], em que a célula é uma célula de *Saccharomyces cerevisiae*.

**[0550]** A invenção descrita e reivindicada aqui não é para estar limitada no escopo pelos aspectos específicos aqui divulgados, uma vez que estes aspectos se pretendem como ilustrações de vários aspectos da invenção. É pretendido que quaisquer aspectos equivalentes estejam dentro do escopo desta invenção. De fato, várias modificações da invenção adicionalmente àquelas mostradas e descritas aqui se tornarão aparentes àqueles peritos na técnica a partir da descrição anterior. Tais modificações se destinam também a residir dentro do escopo das reivindicações anexas. Em caso de conflito, a presente divulgação incluindo definições prevalecerá. Todas as referências são especificamente incorporadas por referência para o que é descrito.

**[0551]** Os seguintes exemplos são oferecidos para ilustrar certos aspectos da presente invenção, mas não pretendem de modo algum limitar o escopo da invenção como reivindicada.

## EXEMPLOS

### Materiais e Métodos

**[0552]** Os químicos usados como tampões e substratos foram produtos comerciais de pelo menos grau de reagente.

**[0553]** ETHANOL RED® (“ER”): Levedura *Saccharomyces cerevisiae* disponível a partir da Fermentis/Lesaffre, EUA.

**[0554]** As placas YPD + clonAT eram compostas por 10 g de extrato de levedura, 20 g de peptona, 20 g de ágar bacto e água desionizada até 960 mL, seguido por tratamento na autoclave. 40 mL de glucose estéril a 50% e 1 mL de solução de estoque de clonNAT foram adicionados, seguidos por mistura e vertimento.

**[0555]** A solução de estoque de clonNAT era composta por 2 g de sulfato de nourseotricina e água deionizada até 20 mL.

### Exemplo 1: Construção de estirpes de Leveduras expressando uma alfa-amilase heteróloga

**[0556]** Este exemplo descreve a construção de células de levedura contendo uma alfa-amilase heteróloga sob controle de um promotor TDH3 (SEQ ID NO: 1) ou promotor ADH1 (SEQ ID NO: 5) de *S. cerevisiae*. Foram concebidos três fragmentos de DNA contendo o promotor, gene e terminador para permitir a recombinação homóloga entre os 3 fragmentos de DNA e no locus X-3 da levedura yMHCT484 (PCT/US2018/035596). A estirpe resultante tem um fragmento contendo promotor (esquerda), um fragmento contendo gene (meio) e um fragmento de terminador ENO2 (SEQ ID NO: 228) (direita) integrados no genoma de *S. cerevisiae* no locus X-3.

### Construção dos fragmentos contendo promotores (fragmentos esquerdos)

**[0557]** DNA não clonado linear sintético contendo homologia de 300 pb com o local X-3, promotor ADH1 (SEQ ID NO: 5) ou THD3 (SEQ ID NO: 1) de *S. cerevisiae* e sequência de sinal EXG1 de *S.*

*cerevisiae* (SEQ ID NO: 227) foram sintetizados pela Thermo Fisher Scientific. Os 2 DNAs lineares foram designados 17ABCK4P e 17ABCK3P para cada promotor listado acima, respectivamente. Para gerar DNA linear adicional para transformação em levedura, o DNA contendo o cassete esquerdo foi amplificado por PCR a partir de 17ABCK4P e 17ABCK3P.

Construção dos fragmentos contendo alfa-amilase (fragmentos médios)

**[0558]** DNA não clonado linear sintético contendo a sequência codificante de peptídeo de sinal EXG1 de *S. cerevisiae* (codificando o sinal de SEQ ID NO: 227), um gene de alfa-amilase com otimização de códons e 50 pb de terminador ENO2 (SEQ ID NO: 228) foram sintetizados pela Thermo Fisher Scientific.

**[0559]** Para gerar DNA linear para transformação em levedura, o DNA contendo o cassete de alfa-amilase foi amplificado por PCR a partir do DNA sintético com os iniciadores 1222985 (5'-ATGAT GAAAA AATAA GCAGA AAAGA CTAAT AATTC TTAGT TAAAA GC-3'; SEQ ID NO: 235) e 1222984 (5'-ATGCT TTCGC TTAAC ACGTT ACTGT G-3'; SEQ ID NO: 236). Foram usadas cinquenta pmoles de cada um dos iniciadores direto e reverso em uma reação de PCR contendo 5 ng de DNA de plasmídeo como modelo, cada dATP, dGTP, dCTP, dTTP a 0,1 mM, 1X Tampão HF Phusion (Thermo Fisher Scientific) e 2 unidades de DNA polimerase Phusion Hot Start em um volume final de 50 µL. A PCR foi realizada em um Termociclador T100™ (Bio-Rad Laboratories, Inc.) programado para um ciclo a 98 °C durante 1 minuto seguido por 32 ciclos cada um a 98 °C durante 10 segundos, 55 °C durante 30 segundos e 72 °C durante 2 minutos com uma extensão final a 72 °C durante 10 minutos. Após termociclagem, os produtos de reação PCR foram isolados por gel e limpos usando o estojo de limpeza NucleoSpin Gel and PCR (Machery-Nagel). Os DNAs lineares resultantes foram designados como indicado na Tabela 4.



Tabela 4. Nomes dos produtos de DNA de alfa-amilase e enzima associada

<b>Número do Produto</b>	<b>Formato do DNA</b>	<b>Peptídeo de sinal</b>	<b>Organismo Dador (domínio catalítico)</b>	<b>SEQ ID NO: (polipeptídeo maduro)</b>	<b>Fragmento Terminado</b>
17ABDQ YP	linear	EXG1	<i>Rhizomucor pusillus</i>	121	ENO2
17ABDQ XP	linear	EXG1	<i>Bacillus licheniformis</i>	122	ENO2
17ABDQ WP	linear	EXG1	<i>Aspergillus niger</i>	123	ENO2
17ABDQ VP	linear	EXG1	<i>Aspergillus tamaris</i>	124	ENO2
17ABDQ UP	linear	EXG1	<i>Acidomyces richmondensis</i>	125	ENO2
17ABDQ TP	linear	EXG1	<i>Aspergillus bombycis</i>	126	ENO2
17ABDQ SP	linear	EXG1	<i>Alternaria sp</i>	127	ENO2
17ABDQ RP	linear	EXG1	<i>Rhizopus microsporus</i>	128	ENO2
17ABDQ QP	linear	EXG1	<i>Syncephalastrum racemosum</i>	129	ENO2
17ABDQ PP	linear	EXG1	<i>Rhizomucor pusillus</i>	130	ENO2

17ABDQ OP	linear	EXG1	<i>Dichotomocla dium hesseltinei</i>	131	ENO2
17ABDQ NP	linear	EXG1	<i>Lichtheimia ramosa</i>	132	ENO2
17ABDQ MP	linear	EXG1	<i>Penicillium aethiopicum</i>	133	ENO2
17ABDQ LP	linear	EXG1	<i>Subulispora sp</i>	134	ENO2
17ABDQ KP	linear	EXG1	<i>Trichoderma paraviridesce ns</i>	135	ENO2
17ABDQJ P	linear	EXG1	<i>Byssoascus striatosporus</i>	136	ENO2
17ABDQI P	linear	EXG1	<i>Aspergillus brasiliensis</i>	137	ENO2
17ABDQ HP	linear	EXG1	<i>Penicillium subspinulosu m</i>	138	ENO2
17ABDQ GP	linear	EXG1	<i>Penicillium antarcticum</i>	139	ENO2
17ABDQ FP	linear	EXG1	<i>Penicillium coprophilum</i>	140	ENO2
17ABDQ EP	linear	EXG1	<i>Penicillium olsonii</i>	141	ENO2
17ABDQ DP	linear	EXG1	<i>Penicillium vasconiae</i>	142	ENO2
17ABDQ CP	linear	EXG1	<i>Penicillium sp</i>	143	ENO2

17ABDQ BP	linear	EXG1	<i>Heterocephal um aurantiacum</i>	144	ENO2
17ABDQ AP	linear	EXG1	<i>Neosartorya massa</i>	145	ENO2
17ABDP7 P	linear	EXG1	<i>Penicillium janthinellum</i>	146	ENO2
17ABDP6 P	linear	EXG1	<i>Aspergillus brasiliensis</i>	147	ENO2
17ABDP5 P	linear	EXG1	<i>Aspergillus westerdijkiae</i>	148	ENO2
17ABDP4 P	linear	EXG1	<i>Hamigera avellanea</i>	149	ENO2
17ABDP3 P	linear	EXG1	<i>Hamigera avellanea</i>	150	ENO2
17ABDP2 P	linear	EXG1	<i>Meripilus giganteus</i>	151	ENO2
17ABDPZ P	linear	EXG1	<i>Cerrena unicolor</i>	152	ENO2
17ABDP YP	linear	EXG1	<i>Physalacria cryptomeriae</i>	153	ENO2
17ABDP XP	linear	EXG1	<i>Lenzites betulinus</i>	154	ENO2
17ABDP WP	linear	EXG1	<i>Trametes ljubarskyi</i>	155	ENO2
17ABDP VP	linear	EXG1	<i>Bacillus subtilis</i>	156	ENO2

17ABDP UP	linear	EXG1	<i>Bacillus subtilis subsp. subtilis</i>	157	ENO2
17ABDPT P	linear	EXG1	<i>Schwanniomyces occidentalis</i>	158	ENO2
17ABDP SP	linear	EXG1	<i>Rhizomucor pusillus</i>	159	ENO2
17ABDP RP	linear	EXG1	<i>Aspergillus niger</i>	160	ENO2
17ABDP QP	linear	EXG1	<i>Bacillus stearothermophilus</i>	161	ENO2
17ABDP PP	linear	EXG1	<i>Bacillus halmapalus</i>	162	ENO2
17ABDP OP	linear	EXG1	<i>Aspergillus oryzae</i>	163	ENO2
17ABDP NP	linear	EXG1	<i>Bacillus amyloliquefaciens</i>	164	ENO2
17ABDP MP	linear	EXG1	<i>Rhizomucor pusillus</i>	165	ENO2
17ABDPL P	linear	EXG1	<i>Kionochaeta ivoriensis</i>	166	ENO2
17ABDP KP	linear	EXG1	<i>Aspergillus niger</i>	167	ENO2
17ABDPJ P	linear	EXG1	<i>Aspergillus oryzae</i>	168	ENO2

17ABDPI P	linear	EXG1	<i>Penicillium canescens</i>	169	ENO2
17ABDP HP	linear	EXG1	<i>Acidomyces acidothermus</i>	170	ENO2
17ABDQ 4P	linear	EXG1	<i>Kinochaeta ivorienensis</i>	171	ENO2
17ABDQ 3P	linear	EXG1	<i>Aspergillus terreus</i>	172	ENO2
17ABDQ 2P	linear	EXG1	<i>Thamnidium elegans</i>	173	ENO2
17ABDQ ZP	linear	EXG1	<i>Meripilus giganteus</i>	174	ENO2

Integração dos fragmentos do lado esquerdo, médio e direito para gerar estirpes de leveduras com uma alfa-amilase heteróloga

**[0560]** A levedura yMHCT484 (PCT/US2018/035596) foi transformada com fragmentos de integração esquerdos, médios e direitos descritos acima. Em cada agrupamento de transformação foram usados um fragmento esquerdo e fragmento direito fixos. O fragmento médio consistiu em um agrupamento de 19-21 fragmentos médios contendo o gene de alfa-amilase com 100-600 ng de cada fragmento (total de 1500 ng). Para auxiliar a recombinação homóloga dos fragmentos esquerdos, médios e direitos nos locais genômicos X-3 foi também usado na transformação um plasmídeo contendo Cas9 e RNA guia específico de X-3 (pMcTs442). Estes quatro componentes foram transformados na estirpe yMHCT484 de *S. cerevisiae* seguindo um protocolo de eletroporação de leveduras (Ver Thompson *et al. Yeast.* 30 abr 1998; 14 (6): 565-71). Os transformantes foram selecionados em YPD+clonNAT para selecionar transformantes que contêm o plasmídeo pMcTs442 de CRISPR/Cas9 (Figura 3). Os transformantes foram escolhidos usando um Sistema de Escolha de Colônias Q-pix (Molecular Devices) para

inocular 1 poço de placa de 96 poços contendo meio YPD+clonNAT. As placas foram cultivadas durante 2 dias, depois foi adicionado glicerol até uma concentração final de 20% e as placas foram armazenadas a -80 °C até serem necessárias. A integração de construto específico de alfa-amilase foi verificada por PCR com iniciadores específicos do lócus e subsequente sequenciamento. As estirpes geradas em este exemplo são mostradas na Tabela 5.

Tabela 5. Estirpes de *S. cerevisiae* expressando alfa-amilase.

<b>pedaço esquerdo</b>	<b>pedaço médio</b>	<b>Promotor</b>	<b>Peptídeo de sinal</b>	<b>Dador de gene de alfa-amilase (domínio catalítico)</b>	<b>SEQ ID NO: (polipeptídeo maduro)</b>	<b>ID da estirpe</b>
17ABCK 4P	17ABD QJP	ADH 1	EXG1	<i>Byssoascus striatosporus</i>	136	P110-A08
17ABCK 4P	17ABD QHP	ADH 1	EXG1	<i>Penicillium subspinosum</i>	138	P110-A09
17ABCK 4P	17ABD QUP	ADH 1	EXG1	<i>Acidomyces richmondensis</i>	125	P110-B01
17ABCK 4P	17ABD QKP	ADH 1	EXG1	<i>Trichoderma paraviridescens</i>	135	P110-B04
17ABCK 4P	17ABD QXP	ADH 1	EXG1	<i>Bacillus licheniformis</i>	122	P110-B05
17ABCK 4P	17ABD QMP	ADH 1	EXG1	<i>Penicillium aethiopicum</i>	133	P110-B08
17ABCK 4P	17ABD QSP	ADH 1	EXG1	<i>Alternaria sp</i>	127	P110-C05

17ABCK 4P	17ABD QOP	ADH 1	EXG1	<i>Dichotomocladi um hesseltinei</i>	131	P110- D01
17ABCK 4P	17ABD QTP	ADH 1	EXG1	<i>Aspergillus bombycis</i>	126	P110- D02
17ABCK 4P	17ABD QIP	ADH 1	EXG1	<i>Aspergillus brasiliensis</i>	137	P110- D10
17ABCK 4P	17ABD QVP	ADH 1	EXG1	<i>Aspergillus tamarii</i>	124	P110- F02
17ABCK 4P	17ABD QTP	ADH 1	EXG1	<i>Aspergillus bombycis</i>	126	P110- F07
17ABCK 4P	17ABD QHP	ADH 1	EXG1	<i>Penicillium subspinulosum</i>	138	P110- G03
17ABCK 4P	17ABD QUP	ADH 1	EXG1	<i>Acidomyces richmondensis</i>	125	P110- G04
17ABCK 4P	17ABD QSP	ADH 1	EXG1	<i>Alternaria sp</i>	127	P110- G06
17ABCK 4P	17ABD QLP	ADH 1	EXG1	<i>Subulispora sp</i>	134	P110- H02
17ABCK 4P	17ABD QHP	ADH 1	EXG1	<i>Penicillium subspinulosum</i>	138	P110- H05
17ABCK 4P	17ABD QGP	ADH 1	EXG1	<i>Penicillium antarcticum</i>	139	P110- H07
17ABCK 4P	17ABD PZP	ADH 1	EXG1	<i>Cerrena unicolor</i>	152	P111- C03
17ABCK 4P	17ABD QAP	ADH 1	EXG1	<i>Neosartorya massa</i>	145	P111- D10
17ABCK 4P	17ABD QDP	ADH 1	EXG1	<i>Penicillium vasconiae</i>	142	P111- F01

17ABCK 4P	17ABD QCP	ADH 1	EXG1	<i>Penicillium sp</i>	143	P111- H08
17ABCK 4P	17ABD PJP	ADH 1	EXG1	<i>Aspergillus oryzae</i>	168	P112- A03
17ABCK 4P	17ABD Q3P	ADH 1	EXG1	<i>Aspergillus terreus</i>	172	P112- A07
17ABCK 4P	17ABD Q3P	ADH 1	EXG1	<i>Aspergillus terreus</i>	172	P112- B11
17ABCK 4P	17ABD Q2P	ADH 1	EXG1	<i>Thamnidium elegans</i>	173	P112- C09
17ABCK 4P	17ABD PPP	ADH 1	EXG1	<i>Bacillus halmapalus</i>	162	P112- D05
17ABCK 4P	17ABD PJP	ADH 1	EXG1	<i>Aspergillus oryzae</i>	168	P112- D06
17ABCK 4P	17ABD PMP	ADH 1	EXG1	<i>Rhizomucor pusillus</i>	165	P112- H03
17ABCK 3P	17ABD QIP	TDH 3	EXG1	<i>Aspergillus brasiliensis</i>	137	P113- A03
17ABCK 3P	17ABD QYP	TDH 3	EXG1	<i>Rhizomucor pusillus</i>	121	P113- B05
17ABCK 3P	17ABD QXP	TDH 3	EXG1	<i>Bacillus licheniformis</i>	122	P113- B06
17ABCK 3P	17ABD QTP	TDH 3	EXG1	<i>Aspergillus bombycis</i>	126	P113- C03
17ABCK 3P	17ABD QNP	TDH 3	EXG1	<i>Lichtheimia ramosa</i>	132	P113- C06
17ABCK 3P	17ABD QVP	TDH 3	EXG1	<i>Aspergillus tamarii</i>	124	P113- C09



17ABCK 3P	17ABD QYP	TDH 3	EXG1	<i>Rhizomucor pusillus</i>	121	P113- C10
17ABCK 3P	17ABD QRP	TDH 3	EXG1	<i>Rhizopus microsporus</i>	128	P113- D07
17ABCK 3P	17ABD QVP	TDH 3	EXG1	<i>Aspergillus tamarii</i>	124	P113- D08
17ABCK 3P	17ABD QSP	TDH 3	EXG1	<i>Alternaria sp</i>	127	P113- D10
17ABCK 3P	17ABD QNP	TDH 3	EXG1	<i>Lichtheimia ramosa</i>	132	P113- F02
17ABCK 3P	17ABD QQP	TDH 3	EXG1	<i>Syncephalastr um racemosum</i>	129	P113- F05
17ABCK 3P	17ABD QJP	TDH 3	EXG1	<i>Byssoascus striatosporus</i>	136	P113- G04
17ABCK 3P	17ABD QTP	TDH 3	EXG1	<i>Aspergillus bombycis</i>	126	P113- G09
17ABCK 3P	17ABD PSP	TDH 3	EXG1	<i>Rhizomucor pusillus</i>	159	P114- A04
17ABCK 3P	17ABD P4P	TDH 3	EXG1	<i>Hamigera avellanea</i>	149	P114- B02
17ABCK 3P	17ABD PUP	TDH 3	EXG1	<i>Bacillus subtilis subsp. subtilis</i>	157	P114- B08
17ABCK 3P	17ABD PUP	TDH 3	EXG1	<i>Bacillus subtilis subsp. subtilis</i>	157	P114- C01
17ABCK 3P	17ABD P2P	TDH 3	EXG1	<i>Meripilus giganteus</i>	151	P114- C04
17ABCK 3P	17ABD PVP	TDH 3	EXG1	<i>Bacillus subtilis</i>	156	P114- C05

17ABCK 3P	17ABD QAP	TDH 3	EXG1	<i>Neosartorya massa</i>	145	P114- C06
17ABCK 3P	17ABD QEP	TDH 3	EXG1	<i>Penicillium olsonii</i>	141	P114- C07
17ABCK 3P	17ABD PTP	TDH 3	EXG1	<i>Schwanniomyc es occidentalis</i>	158	P114- D02
17ABCK 3P	17ABD PRP	TDH 3	EXG1	<i>Aspergillus niger</i>	160	P114- D07
17ABCK 3P	17ABD P3P	TDH 3	EXG1	<i>Hamigera avellanea</i>	150	P114- F06
17ABCK 3P	17ABD P6P	TDH 3	EXG1	<i>Aspergillus brasiliensis</i>	147	P114- F07
17ABCK 3P	17ABD PUP	TDH 3	EXG1	<i>Bacillus subtilis subsp. subtilis</i>	157	P114- F08
17ABCK 3P	17ABD P2P	TDH 3	EXG1	<i>Meripilus giganteus</i>	151	P114- H02
17ABCK 3P	17ABD QAP	TDH 3	EXG1	<i>Neosartorya massa</i>	145	P114- H03
17ABCK 3P	17ABD PZP	TDH 3	EXG1	<i>Cerrena unicolor</i>	152	P114- H07
17ABCK 3P	17ABD QAP	TDH 3	EXG1	<i>Neosartorya massa</i>	145	P114- H08
17ABCK 3P	17ABD PKP	TDH 3	EXG1	<i>Aspergillus niger</i>	167	P115- B03
17ABCK 3P	17ABD PMP	TDH 3	EXG1	<i>Rhizomucor pusillus</i>	165	P115- C11
17ABCK 3P	17ABD PMP	TDH 3	EXG1	<i>Rhizomucor pusillus</i>	165	P115- D09

17ABCK 3P	17ABD Q3P	TDH 3	EXG1	<i>Aspergillus terreus</i>	172	P115- F06
17ABCK 3P	17ABD PJP	TDH 3	EXG1	<i>Aspergillus oryzae</i>	168	P115- G04

**Exemplo 2: Ensaio de atividade de estirpes de leveduras**

**expressando alfa-amilase**

**[0561]** As estirpes de levedura do Exemplo 1 foram cultivadas durante a noite em meio YPD padrão contendo glucose a 2% ou 6%. O meio de levedura cultivado foi sujeito a centrifugação a 3500 rpm durante 10 min para se coletar o sobrenadante. O sobrenadante da cultura é usado para ensaios de atividade de enzima, como descrito em baixo. A levedura pode ser também cultivada usando outros meios de cultivo tais como meio YNB mínimo ou mosto de milho liquefeito industrial clarificado e filtrado.

**Ensaio de atividade de glucoamilase**

**[0562]** A atividade de glucoamilase foi medida usando maltose como substrato. A hidrólise da maltose catalisada por enzimas origina glucose como o produto de reação que pode ser detectado e quantificado usando estojos de ensaio comercialmente disponíveis tais como AUTOKIT GLUCOSE C2 da Wako Diagnostics. Os reagentes proporcionados nos estojos de ensaio reagem com a glucose resultando em uma mudança de cor com absorvância máxima a 505 nm. A intensidade da absorvância medida espectrofotometricamente é proporcional à atividade de glucoamilase. A absorvância a 505 nm pode ser ajustada à curva padrão gerada usando uma enzima glucoamilase purificada para se estimar a atividade da enzima. As condições de reação e o desenvolvimento de cor são descritos na Tabela 5 e Tabela 6, respectivamente. As Unidades de glucoamilase (AGU) para a atividade de glucoamilase são definidas como a

quantidade de enzima requerida para hidrolisar uma micromole de maltose por minuto sob as condições de reação.

Tabela 5. Condições de reação da glucoamilase

Quantidade apropriada de sobrenadante de levedura	10-200 $\mu$ L
Substrato	maltose, 10 mM
Tampão	acetato, 0,1 M
pH	5,0 $\pm$ 0,05
Temperatura de incubação	32 $^{\circ}$ C
Tempo de reação	5-20 min
Gama do ensaio de glucoamilase	0,001-0,036 AGU/mL

Tabela 6. Desenvolvimento da cor

Mistura reacional	10 $\mu$ L
Reagente de desenvolvimento AUTOKIT GLICOSE C2	200 $\mu$ L
Temperatura de incubação	temperatura ambiente ou 37 $^{\circ}$ C
Tempo de reação	10-25 min
Comprimento de onda	505 nm

Ensaio de atividade de alfa-amilase

**[0563]** A atividade de alfa-amilase foi medida usando *p*-nitrofenil-malto-heptaosídeo bloqueado (BPNPG7) como substrato, que é incluído como o reagente de amilase HR da Megazymes. A hidrólise de enzimas da ligação alfa de BPNPG7 libera um oligômero de maltossacarídeo bloqueado e um oligômero de maltossacarídeo de *p*-nitrofenila. O maltossacarídeo de *p*-nitrofenil reagirá com uma glucoamilase da Megazymes originando *p*-nitrofenol que pode ser detectado usando estojos de ensaio comercialmente disponíveis tais como R-AMHR4 da

MEGAZYMES. Os reagentes proporcionados nos estojos de ensaio reagirão especificamente com *p*-nitrofenol resultando em formação de cor. A intensidade da cor medida usando um espectrofotômetro ou leitor de microplaca é proporcional à atividade de alfa-amilase. As condições de reação e o desenvolvimento de cor são descritos na Tabela 7 e Tabela 8, respectivamente.

Tabela 7. Condições de reação de alfa-amilase

Quantidade apropriada de sobrenadante de levedura	10-200 µL
Substrato	<i>p</i> -nitrofenil-malto-heptaosídeo bloqueado (BPNPG7), 10 mM
Tampão	acetato, 0,1 M
pH	5,0 ± 0,05
Temperatura de incubação	32 °C
Tempo de reação	20 min
Gama de ensaio de alfa-amilase	5-200 ng/mL

Tabela 8. Desenvolvimento da cor

Mistura reacional	20 µL
<i>p</i> -nitrofenil-malto-heptaosídeo bloqueado (BPNPG7)	80 µL
Solução de paragem (Tris a 4%)	100 µL
Temperatura de incubação	Temperatura ambiente - 32 °C
Tempo de reação	10-25 min
Comprimento de onda	400 nm

### Resultados

**[0564]** A absorvância a 505 nm aumenta à medida que a quantidade de glucoamilase purificada adicionada para hidrolisar a maltose

em glucose aumenta. A absorvância a 400 nm aumenta à medida que a quantidade de alfa-amilase purificada adicionada aumenta. Especificamente, a alfa-amilase hidrolisa o *p*-nitrofenil-malto-heptaosídeo bloqueado (BPNPG7), liberando um oligômero de maltossacarídeo bloqueado e um oligômero de maltossacarídeo de *p*-nitrofenila. O maltossacarídeo de *p*-nitrofenila reage com uma glucoamilase da Megayzmes originando *p*-nitrofenol que absorve a 400 nm. Uma curva padrão para a glucoamilase e a alfa-amilase purificadas foi gerada e usada para se estimar a atividade de glucoamilase e alfa-amilase em sobrenadantes de levedura.

**[0565]** Os resultados para a atividade de alfa-amilase e a atividade de glucoamilase são mostrados na Tabela 9. Uma representação gráfica da atividade comparativa de alfa-amilase é mostrada na Figura 1.

Tabela 9. Atividade de alfa-amilase (AA) e glucoamilase (GA) e secreção estimada

No. da estirpe de levedura	Promotor da expressão de alfa-amilase	SEQ ID NO: (polipeptídeo maduro)	Dador de gene de AA (domínio catalítico)	Glucoamilase		Alfa-amilase	
				atividade	Conc. (ug/mL)	atividade	Conc. (ug/mL)
1	Estirpe de fundo com gene de glucoamilase, sem gene de alfa-amilase			1,03	10,6	N/A	N/A
1	Estirpe de fundo com gene de glucoamilase, sem gene de alfa-amilase			0,98	9,8	N/A	N/A

1	Estirpe de fundo com gene de glucoamilase, sem gene de alfa-amilase			1,00	10,0	N/A	N/A
2	pADH1	125	<i>Acidomyces richmondensis</i>	0,86	8,1	0,175 3	0,04
3	pADH1	125	<i>Acidomyces richmondensis</i>	1,03	10,6	0,552 9	0,16
4	pADH1	136	<i>Byssosascus striatosporus</i>	0,76	6,6	0,084	0,01
5	pADH1	162	<i>Bacillus halmapalus</i>	1,58	18,5	0,431 4	0,12
6	pADH1	172	<i>Aspergillus terreus</i>	0,91	8,8	1,415 4	0,44
7	pTDH3	129	<i>Syncephalastrium racemosum</i>	0,87	8,2	1,365 5	0,43
8	pTDH3	151	<i>Meripilus</i>	0,60	4,3	1,663 6	0,52
9	pTDH3	141	<i>Penicillium olsonii</i>	0,99	10,0	0,084 6	0,01
11	pADH1	131	<i>Dichotomocladium hesseltinei</i>	0,75	6,5	0,084	0,01
12	pADH1	122	<i>Bacillus licheniformis</i>	0,93	9,1	0,114 7	0,02
13	pADH1	133	<i>Penicillium aethiopicum</i>	1,03	10,5	0,084 2	0,01

14	pADH1	143	<i>Penicillium sp</i>	0,88	8,4	0,204 7	0,05
15	pADH1	168	<i>Aspergillus oryzae</i>	0,90	8,7	1,561	0,49
16	pTDH3	122	<i>Bacillus licheniformis</i>	0,81	7,3	0,121 9	0,02
17	pTDH3	126	<i>Aspergillus bombycis</i>	0,98	9,8	0,162 8	0,03
18	pTDH3	145	<i>Neosartorya massa</i>	0,92	8,9	1,902	0,60
19	pTDH3	160	<i>Aspergillus niger</i>	0,95	9,3	0,089 9	0,01
20	pTDH3	167	<i>Aspergillus niger</i>	0,98	9,8	1,235 7	0,38
21	pADH1	126	<i>Aspergillus bombycis</i>	0,95	9,3	0,389 1	0,11
22	pADH1	127	<i>Alternaria sp</i>	0,93	9,1	0,086 6	0,01
23	pADH1	138	<i>Penicillium subspinulosu m</i>	0,97	9,7	0,084 8	0,01
24	pADH1	145	<i>Neosartorya massa</i>	0,93	9,1	0,878 6	0,27
25	pTDH3	132	<i>Lichtheimia ramosa</i>	0,89	8,5	1,190 3	0,37
26	pTDH3	132	<i>Lichtheimia ramosa</i>	0,88	8,3	1,749 8	0,55
27	pTDH3	121	<i>Rhizomucor pusillus</i>	0,78	7,0	1,334	0,42



28	pTDH3	159	<i>Aspergillus niger</i>	0,90	8,7	1,958 2	0,62
29	pTDH3	147	<i>Aspergillus brasiliensis</i>	1,03	10,6	0,346 9	0,09
30	pTDH3	168	<i>Aspergillus oryzae</i>	0,92	9,0	1,565 5	0,49
31	pADH1	124	<i>Aspergillus tamaris</i>	0,87	8,3	1,158 3	0,36
32	pADH1	138	<i>Penicillium subspinosum</i>	0,90	8,7	0,084 8	0,01
34	pADH1	172	<i>Aspergillus terreus</i>	0,88	8,4	1,364 5	0,43
35	pTDH3	137	<i>Aspergillus brasiliensis</i>	0,92	8,9	1,218 3	0,38
36	pTDH3	169	<i>Penicillium canescens</i>	0,83	7,7	1,821 2	0,57
37	pTDH3	127	<i>Alternaria sp</i>	0,95	9,4	0,087 7	0,01
38	pTDH3	151	<i>Meripilus</i>	0,56	3,8	1,366 5	0,43
39	pTDH3	152	<i>Cerrena unicolor</i>	0,98	9,8	0,083 1	0,01
40	pTDH3	172	<i>Aspergillus terreus</i>	0,98	9,9	0,877 4	0,27
41	pADH1	134	<i>Subulispora sp</i>	0,72	6,1	0,091 6	0,01
42	pADH1	127	<i>Alternaria sp</i>	0,80	7,3	0,093 8	0,01

43	pADH1	137	<i>Aspergillus brasiliensis</i>	0,94	9,2	1,177 2	0,36
44	pADH1	168	<i>Aspergillus oryzae</i>	0,92	9,0	1,534 7	0,48
46	pTDH3	126	<i>Aspergillus bombycis</i>	1,03	10,5	0,195 5	0,05
47	pTDH3	128	<i>Rhizopus microsporus</i>	0,65	5,1	0,963 9	0,30
48	pTDH3	157	<i>Bacillus subtilis subsp. subtilis</i>	0,94	9,2	1,299 2	0,40
49	pTDH3	156	<i>Bacillus subtilis</i>	0,93	9,1	1,142 7	0,35
50	pTDH3	157	<i>Bacillus subtilis subsp. subtilis</i>	0,95	9,4	1,337 4	0,42
51	pADH1	140	<i>Penicillium coprophilum</i>	0,90	8,6	0,758 8	0,23
52	pADH1	136	<i>Byssosascus striatosporus</i>	0,94	9,2	0,085	0,01
53	pADH1	146	<i>Penicillium janthinellum</i>	0,79	7,0	0,332 5	0,09
54	pADH1	173	<i>Thamnidium elegans</i>	0,91	8,8	1,184 4	0,37
55	pADH1	163	<i>Aspergillus oryzae</i>	0,88	8,4	1,717 5	0,54
56	pTDH3	137	<i>Aspergillus brasiliensis</i>	0,93	9,0	0,901 2	0,27

57	pTDH3	123	<i>Aspergillus niger</i>	0,94	9,2	1,299 4	0,40
58	pTDH3	150	<i>Hamigera avellanea</i>	0,87	8,3	0,769 8	0,23
59	pTDH3	149	<i>Hamigera avellanea</i>	0,89	8,5	0,804 8	0,24
60	pTDH3	165	<i>Rhizomucor pusillus</i>	0,94	9,3	1,911 7	0,60
61	pADH1	138	<i>Penicillium subspinosum</i>	0,76	6,7	0,123 3	0,02
62	pADH1	126	<i>Aspergillus bombycis</i>	0,94	9,3	0,362 6	0,10
63	pADH1	142	<i>Penicillium vasconiae</i>	0,74	6,4	0,100 7	0,01
64	pADH1	165	<i>Rhizomucor pusillus</i>	0,94	9,3	1,718 3	0,54
66	pTDH3	136	<i>Byssoascus striatosporus</i>	0,96	9,6	0,084	0,01
67	pTDH3	124	<i>Aspergillus tamaris</i>	0,90	8,6	1,026 3	0,32
68	pTDH3	149	<i>Hamigera avellanea</i>	1,04	10,7	0,417 3	0,12
69	pTDH3	145	<i>Neosartorya massa</i>	0,96	9,5	0,224 9	0,05
70	pTDH3	157	<i>Bacillus subtilis subsp. subtilis</i>	0,99	9,9	1,4	0,44

71	pTDH3	165	<i>Rhizomucor pusillus</i>	0,89	8,4	2,218 1	0,70
72	pADH1	135	<i>Trichoderma paraviridescens</i>	0,92	9,0	0,087 3	0,01
73	pADH1	139	<i>Penicillium antarcticum</i>	0,78	6,9	0,092 6	0,01
74	pADH1	152	<i>Cerrena unicolor</i>	0,93	9,1	0,085 8	0,01
76	pADH1	173	<i>Thamnidium elegans</i>	0,89	8,6	1,245 8	0,39
77	pTDH3	121	<i>Rhizomucor pusillus</i>	0,80	7,2	1,632 6	0,51
78	pTDH3	124	<i>Aspergillus tamarrii</i>	0,80	7,3	1,296 4	0,40
79	pTDH3	158	<i>Schwanniomyces occidentalis</i>	1,00	10,0	0,267 9	0,07
80	pTDH3	150		0,83	7,7	0,707 9	0,21
81	pTDH3	145	<i>Neosartorya massa</i>	0,90	8,7	0,195 7	0,05
82	pTDH3	165	<i>Rhizomucor pusillus</i>	0,81	7,3	2,127 3	0,67

**Exemplo 3: Sacarificação e fermentação simultâneas (SSF)**

**de estirpes de levedura expressando alfa-amilase**

**[0566]** As estirpes de levedura foram cultivadas durante a noite em meio YPD padrão contendo glucose a 2%. O meio de levedura cultivado foi centrifugado a 3000 rpm durante 10 min para se coletar o

sobrenadante. O sobrenadante foi usado para ensaio de atividade de enzima, como descrito em baixo.

Ensaio de atividade de alfa-amilase

**[0567]** A atividade de alfa-amilase foi detectada por medição da quantidade de amido degradado através da hidrólise enzimática do amido. Os reagentes iodeto de potássio e iodo foram usados para medir o amido residual com base no desenvolvimento da cor a partir da aplicação do reagente. A intensidade da cor medida em um espectrofotômetro ou leitor de microplaca é inversamente proporcional à atividade de alfa-amilase. As condições de reação e o desenvolvimento de cor foram descritos na Tabela 11 e Tabela 12, respectivamente.

Tabela 11. Condição de reação de alfa-amilase

Quantidade de sobrenadante de levedura	20 µL
Quantidade de substrato	130 µL
Substrato	Amido a 2 mM
Tampão	Acetato de sódio, 0,1 M, Triton 100 a 0,01%
pH	5,0 ± 0,05
Temperatura de incubação	20 °C
Tempo de reação	2-3 hr

Tabela 12. Desenvolvimento da cor

Mistura reacional	150 µL
Quantidade de reagente	50 µL
Reagente	Iodeto de potássio a 14,5 mM, iodo a 0,9 mM
Temperatura de incubação	20 °C
Tempo de reação	10-15 min
Comprimento de onda	595 nm

**[0568]** A sacarificação e fermentação simultâneas (SSF) foram realizadas através de fermentações em mini-escala usando mosto de milho industrial (Avantec<sup>®</sup> Amp, Novozymes, A/S) usando condições mostradas na Tabela 13. As estirpes de levedura foram cultivadas durante a noite em meio YPD com glucose a 2% durante 24 horas a 30 °C e 300 rpm. O mosto de milho foi suplementado com 250 ppm de penicilina. Aproximadamente 0,6 mg de mosto de milho foram distribuídos por poço a placas de microtitulação de 96 poços, seguido pela adição de aproximadamente 10<sup>8</sup> células de levedura/g de mosto de milho a partir da cultura durante a noite. As placas foram incubadas a 32 °C sem agitação. Duplicados de cada estirpe foram analisados após fermentações de 48 horas. A fermentação foi parada pela adição de 100 µL de H<sub>2</sub>SO<sub>4</sub> a 8%, seguida por centrifugação a 3000 rpm durante 10 min. O sobrenadante foi analisado quanto ao etanol usando HPLC.

Tabela 13. Condições de reação de fermentação em placas de microtitulação

Substrato	Mosto de milho Avantec <sup>®</sup> Amp
Inoculação de levedura	10 <sup>8</sup> células/g de mosto de milho
Ureia suplementar	250 ppm
pH	5,0 ± 0,05
Temperatura de incubação	32 °C
Tempo de reação	48 horas

**[0569]** Como mostrado na Tabela 14 foi obtido mais etanol a partir da levedura expressando uma alfa-amilase heteróloga em comparação com a levedura não tendo expressão da alfa-amilase heteróloga. A coluna “Média (amido residual)” mostra os resultados do ensaio

de atividade de alfa-amilase com base em YPD onde o amido residual é inversamente proporcional à atividade de alfa-amilase, enquanto as colunas “Média (etanol normalizado)” mostram o etanol no ponto temporal de 48 horas de duas experiências de fermentação e sacarificação simultâneas (SSF) diferentes, normalizado para aquele da estirpe sem expressão da alfa-amilase heteróloga (yMHCT484).

Tabela 14: IDs das estirpes e dados de etanol normalizados de etanol e atividade de alfa-amilase.

Promotor	Peptídeo de sinal	Dador de gene de alfa-amilase (domínio catalítico)	SEQ ID NO: (polipeptídeo maduro)	Média (etanol normalizado) (exp 1)	Média (etanol normalizado) (exp 2)	Média (amido residual)
Estirpe de fundo com gene de glucoamilase, sem gene de alfa-amilase				1,00	1,00	0,74
ADH1	EXG1	<i>Byssoascus striatosporus</i>	136	1,18	0,69	1,007
ADH1	EXG1	<i>Penicillium subspinulosum</i>	138	1,27	0,97	1,029
ADH1	EXG1	<i>Acidomyces richmondensis</i>	125	1,10	0,88	0,414
ADH1	EXG1	<i>Trichoderma paraviridescens</i>	135	1,14	0,78	0,564
ADH1	EXG1	<i>Bacillus licheniformis</i>	122	1,28	0,85	0,453
ADH1	EXG1	<i>Penicillium aethiopicum</i>	133	1,19	0,97	0,954
ADH1	EXG1	<i>Alternaria sp</i>	127	1,17	0,97	0,460

ADH1	EXG1	<i>Dichotomocladium hesseltinei</i>	131	1,15	0,74	0,980
ADH1	EXG1	<i>Aspergillus bombycis</i>	126	1,13	0,98	0,423
ADH1	EXG1	<i>Aspergillus brasiliensis</i>	137	1,16	1,06	0,507
ADH1	EXG1	<i>Aspergillus tamarii</i>	124	1,22	1,04	0,440
ADH1	EXG1	<i>Aspergillus bombycis</i>	126	1,14	0,97	0,479
ADH1	EXG1	<i>Acidomyces richmondensis</i>	125	1,11	1,00	0,464
ADH1	EXG1	<i>Alternaria sp</i>	127	1,13	0,77	0,411
ADH1	EXG1	<i>Subulispora sp</i>	134	1,03	0,64	0,452
ADH1	EXG1	<i>Penicillium subspinulosum</i>	138	1,11	0,82	0,607
ADH1	EXG1	<i>Penicillium antarcticum</i>	139	1,17	0,71	0,476
ADH1	EXG1	<i>Cerrena unicolor</i>	152	1,07	0,95	0,533
ADH1	EXG1	<i>Neosartorya massa</i>	145	1,05	1,13	0,422
ADH1	EXG1	<i>Penicillium vasconiae</i>	142	1,15	0,72	0,394
ADH1	EXG1	<i>Penicillium sp</i>	143	1,09	0,88	0,412
ADH1	EXG1	<i>Aspergillus oryzae</i>	168	1,21	1,16	0,038
ADH1	EXG1	<i>Aspergillus terreus</i>	172	1,15	1,22	0,038



ADH1	EXG1	<i>Aspergillus terreus</i>	172	1,17	1,22	0,037
ADH1	EXG1	<i>Thamnidium elegans</i>	173	1,25	1,13	0,037
ADH1	EXG1	<i>Aspergillus oryzae</i>	168	1,10	1,16	0,044
ADH1	EXG1	<i>Rhizomucor pusillus</i>	165	1,23	1,23	0,036
TDH3	EXG1	<i>Aspergillus brasiliensis</i>	137	1,16	1,07	0,042
TDH3	EXG1	<i>Rhizomucor pusillus</i>	121	1,27	1,20	0,040
TDH3	EXG1	<i>Bacillus licheniformis</i>	122	0,91	0,93	0,046
TDH3	EXG1	<i>Aspergillus bombycis</i>	126	1,20	0,97	0,039
TDH3	EXG1	<i>Lichtheimia ramosa</i>	132	1,19	1,12	0,051
TDH3	EXG1	<i>Aspergillus tamarii</i>	124	1,21	1,11	0,048
TDH3	EXG1	<i>Rhizopus microsporus</i>	128	1,02	1,01	0,041
TDH3	EXG1	<i>Aspergillus tamarii</i>	124	1,18	1,04	0,040
TDH3	EXG1	<i>Syncephalastrum racemosum</i>	129	1,17	1,24	0,042
TDH3	EXG1	<i>Byssoascus striatosporus</i>	136	1,04	0,96	0,541

TDH3	EXG1	<i>Aspergillus bombycis</i>	126	1,05	0,97	0,042
TDH3	EXG1	<i>Rhizomucor pusillus</i>	159		1,25	0,141
TDH3	EXG1	<i>Hamigera avellanea</i>	149	1,11	1,03	0,055
TDH3	EXG1	<i>Bacillus subtilis subsp. subtilis</i>	157	1,19	1,08	0,043
TDH3	EXG1	<i>Bacillus subtilis subsp. subtilis</i>	157	1,25	1,11	0,044
TDH3	EXG1	<i>Meripilus giganteus</i>	151	1,20	1,14	0,045
TDH3	EXG1	<i>Bacillus subtilis</i>	156	1,35	1,27	0,048
TDH3	EXG1	<i>Neosartorya massa</i>	145	1,15	1,03	0,048
TDH3	EXG1	<i>Penicillium olsonii</i>	141	0,96	0,98	0,052
TDH3	EXG1	<i>Aspergillus niger</i>	160	1,06	0,98	0,053
TDH3	EXG1	<i>Hamigera avellanea</i>	150	1,11	1,04	0,046
TDH3	EXG1	<i>Bacillus subtilis subsp. subtilis</i>	157	1,16	1,13	0,045
TDH3	EXG1	<i>Meripilus giganteus</i>	151	1,11	1,14	0,039
TDH3	EXG1	<i>Cerrena unicolor</i>	152	1,02	0,96	0,546
TDH3	EXG1	<i>Neosartorya massa</i>	145	1,25	1,04	0,053
TDH3	EXG1	<i>Aspergillus niger</i>	167	1,08	1,05	0,040
TDH3	EXG1	<i>Rhizomucor pusillus</i>	165	1,30	1,23	0,039

TDH3	EXG1	<i>Aspergillus terreus</i>	172	1,13	1,11	0,043
TDH3	EXG1	<i>Aspergillus oryzae</i>	168	1,13	1,17	0,037

**Exemplo 4: Construção de estirpes de Levedura expressando uma trealase heteróloga**

**[0570]** Este exemplo descreve a construção de células de levedura contendo uma trealase heteróloga sob controle de um promotor CCW12 (SEQ ID NO: 232) ou promotor PGK1 (SEQ ID NO: 4) de *S. cerevisiae*. Foram concebidos três fragmentos de DNA contendo o promotor, gene e terminador para permitir a recombinação homóloga entre os 3 fragmentos de DNA e no locus X-3 da levedura yMHCT484 (PCT/US2018/035596). A estirpe resultante tem um fragmento contendo promotor (esquerda), um fragmento contendo gene (meio) e um fragmento de terminador TEF1 (SEQ ID NO: 233) (direita) integrados no genoma de *S. cerevisiae* no locus X-3.

**Construção dos fragmentos contendo promotores (fragmentos esquerdos)**

**[0571]** DNA não clonado linear sintético contendo homologia de 300 pb com o local X-3, promotor CCW12 (SEQ ID NO: 232) ou PGK1 (SEQ ID NO: 4) de *S. cerevisiae* e sequência de sinal AGA2 de *S. cerevisiae* (SEQ ID NO: 234) foram sintetizados pela Thermo Fisher Scientific. Os 2 DNAs lineares foram designados 17ABCK6P e 17ABCK7P para cada promotor listado acima, respectivamente.

**[0572]** Para gerar DNA linear adicional para transformação em levedura, o DNA contendo o cassete esquerdo foi amplificado por PCR a partir de 17ABCK6P e 17ABCK7P. Foram usadas cinquenta pmoles de cada um dos iniciadores direto e reverso em uma reação de PCR contendo 12,5

ng de DNA linear como modelo, cada dATP, dGTP, dCTP, dTTP a 0,1 mM, 1X Tampão HF Phusion (Thermo Fisher Scientific) e 2 unidades de DNA polimerase Phusion Hot Start em um volume final de 50 µL. A PCR foi realizada em um Termociclador T100™ (Bio-Rad Laboratories, Inc.) programado para um ciclo a 98 °C durante 1 minuto seguido por 32 ciclos cada um a 98 °C durante 10 segundos, 60 °C durante 30 segundos e 72 °C durante 1 minuto com uma extensão final a 72 °C durante 10 minutos. Após termociclagem, os produtos de reação PCR foram isolados por gel e limpos usando o estojo de limpeza NucleoSpin Gel and PCR (Machery-Nagel).

Construção dos fragmentos contendo trealase (fragmentos médios)

**[0573]** DNA não clonado linear sintético contendo a sequência codificante de peptídeo de sinal AGA2 de *S. cerevisiae* (codificando o sinal de SEQ ID NO: 234), um gene de trealase com otimização de códons e 50 pb de terminador TEF1 (SEQ ID NO: 233) foram sintetizados pela Thermo Fisher Scientific. Os DNAs lineares resultantes foram designados como indicado na Tabela 15.

**[0574]** Um subconjunto dos fragmentos contendo trealase foi ordenado como DNA de plasmídeo sintético clonado em vez de DNA não clonado linear. O DNA de plasmídeo sintético contendo sequência codificante de sinal AGA2 de *S. cerevisiae*, um gene de trealase com otimizações de códons e 50 pb de terminador TEF1 foram sintetizados pela Thermo Fisher Scientific. Os DNAs resultantes foram designados como indicado na Tabela 15.

Construção do fragmento contendo o terminador (fragmento direito)

**[0575]** DNA não clonado linear sintético contendo terminador TEF1 de *S. cerevisiae* e uma homologia de 300 pb com o local X-3 foi sintetizado pela Thermo Fisher Scientific.

Tabela 15. Nomes dos produtos de DNA de trealase e enzima associada

<b>Número do produto</b>	<b>Formato do DNA</b>	<b>Peptídeo de Sinal</b>	<b>Organismo Dador (domínio catalítico)</b>	<b>SEQ ID NO: (polipeptídeo maduro)</b>	<b>Fragmento Terminador</b>
17ABFBKP	linear	AGA2	<i>Chaetomium megalocarpum</i>	175	TEF1
17ABFBJP	linear	AGA2	<i>Lecanicillium psalliotae</i>	176	TEF1
17ABFBIP	linear	AGA2	<i>Doratomyces sp</i>	177	TEF1
17ABFBHP	linear	AGA2	<i>Mucor moelleri</i>	178	TEF1
17ABFBGP	linear	AGA2	<i>Phialophora cyclaminis</i>	179	TEF1
17ABFBFP	linear	AGA2	<i>Thielavia arenaria</i>	180	TEF1
17ABFBEP	linear	AGA2	<i>Thielavia antarctica</i>	181	TEF1
17ABFBDP	linear	AGA2	<i>Chaetomium sp</i>	182	TEF1
17ABFBGP	linear	AGA2	<i>Chaetomium nigricolor</i>	183	TEF1
17ABFBFP	linear	AGA2	<i>Chaetomium jodhpurensense</i>	184	TEF1
17ABFBAP	linear	AGA2	<i>Chaetomium piluliferum</i>	185	TEF1

17ABFA7P	linear	AGA2	<i>Myceliophthora hinnulea</i>	186	TEF1
17ABFA6P	linear	AGA2	<i>Chloridium virescens</i>	187	TEF1
17ABFA5P	linear	AGA2	<i>Gelasinospora cratophora</i>	188	TEF1
17ABFA4P	linear	AGA2	<i>Acidobacteriaceae bacterium</i>	189	TEF1
17ABFA3P	linear	AGA2	<i>Acidobacterium capsulatum</i>	190	TEF1
17ABFA2P	linear	AGA2	<i>Acidovorax wautersii</i>	191	TEF1
17ABFAZP	linear	AGA2	<i>Xanthomonas arboricola</i>	192	TEF1
17ABFAYP	linear	AGA2	<i>Kosakonia sacchari</i>	193	TEF1
17ABFAXP	linear	AGA2	<i>Enterobacter sp</i>	194	TEF1
17ABFAWP	linear	AGA2	<i>Saitozyma flava</i>	195	TEF1
17ABFAVP	linear	AGA2	<i>Phaeotremella skinneri</i>	196	TEF1
17ABFAUP	linear	AGA2	<i>Trichoderma asperellum</i>	197	TEF1
17ABFATP	linear	AGA2	<i>Corynascus sepedonium</i>	198	TEF1
17ABFASP	linear	AGA2	<i>Myceliophthora thermophila</i>	199	TEF1

17ABFARP	linear	AGA2	<i>Trichoderma reesei</i> GH37	200	TEF1
17ABFAQP	linear	AGA2	<i>Chaetomium virescens</i>	201	TEF1
17ABFAPP	linear	AGA2	<i>Rhodothermus marinus</i>	202	TEF1
17ABFAOP	linear	AGA2	<i>Myceliophthora sepedonium</i>	203	TEF1
17ABFANP	linear	AGA2	<i>Moelleriella libera</i>	204	TEF1
17ABFAMP	linear	AGA2	<i>Acremonium dichromosporum</i>	205	TEF1
17ABFALP	linear	AGA2	<i>Fusarium sambucinum</i>	206	TEF1
17ABFAKP	linear	AGA2	<i>Phoma sp</i>	207	TEF1
17ABFAJP	linear	AGA2	<i>Lentinus similis</i>	208	TEF1
17ABFAIP	linear	AGA2	<i>Diaporthe nobilis</i>	209	TEF1
17ABFAHP	linear	AGA2	<i>Solicoccozyma terricola</i>	210	TEF1
17ABFAGP	linear	AGA2	<i>Dioszegia cryoxerica</i>	211	TEF1
17ABFO6P	plasmídeo	AGA2	<i>Talaromyces funiculosus</i>	212	TEF1
17ABFO5P	plasmídeo	AGA2	<i>Hamigera avellanea</i>	213	TEF1
17ABFO4P	plasmídeo	AGA2	<i>Talaromyces ruber</i>	214	TEF1

17ABFO3P	plasmídeo	AGA2	<i>Trichoderma lixii</i>	215	TEF1
17ABFO2P	plasmídeo	AGA2	<i>Aspergillus cervinus</i>	216	TEF1
17ABFOZP	plasmídeo	AGA2	<i>Rasamsonia brevistipitata</i>	217	TEF1
17ABFOYP	plasmídeo	AGA2	<i>Acremonium curvulum</i>	218	TEF1
17ABFOXP	plasmídeo	AGA2	<i>Talaromyces piceae</i>	219	TEF1
17ABFOWP	plasmídeo	AGA2	<i>Penicillium sp</i>	220	TEF1
17ABFOVP	plasmídeo	AGA2	<i>Talaromyces aurantiacus</i>	221	TEF1
17ABFOUP	plasmídeo	AGA2	<i>Talaromyces pinophilus</i>	222	TEF1
17ABFOTP	plasmídeo	AGA2	<i>Talaromyces leycettanus</i>	223	TEF1
17ABFOSP	plasmídeo	AGA2	<i>Talaromyces variabilis</i>	224	TEF1
17ABFORP	plasmídeo	AGA2	<i>Aspergillus niger</i>	225	TEF1
17ABFOQP	plasmídeo	AGA2	<i>Trichoderma reesei</i> GH65	226	TEF1

Integração dos fragmentos do lado esquerdo, médio e direito para gerar cepas de leveduras com uma trealase heteróloga

**[0576]** A levedura yMHCT484 (PCT/US2018/035596) foi transformada com fragmentos de integração esquerdos, médios e direitos descritos acima. Em cada agrupamento de transformação foram usados um



fragmento esquerdo e fragmento direito fixos. O fragmento médio consistiu em um agrupamento de 13-21 fragmentos médios contendo o gene de trealase com 100-600 ng de cada fragmento (total de 1000 ng). Para auxiliar a recombinação homóloga dos fragmentos esquerdos, médios e direitos nos locais genômicos X-3 foi também usado na transformação um plasmídeo contendo Cas9 e RNA guia específico de X-3 (pMcTs442). Estes quatro componentes foram transformados na estirpe yMHCT484 de *S. cerevisiae* seguindo um protocolo de eletroporação de leveduras (Ver Thompson *et al. Yeast.* 30 abr 1998; 14 (6): 565-71). Os transformantes foram selecionados em YPD+clonNAT para selecionar transformantes que contêm o plasmídeo pMcTs442 de CRISPR/Cas9. Os transformantes foram escolhidos usando um Sistema de Escolha de Colônias Q-pix (Molecular Devices) para inocular 1 poço de placa de 96 poços contendo meio YPD+clonNAT. As placas foram cultivadas durante 2 dias, depois foi adicionado glicerol até uma concentração final de 20% e as placas foram armazenadas a -80 °C até serem necessárias. A integração de um construto específico de trealase foi verificada por PCR com iniciadores específicos do locus e subsequente sequenciamento.

#### **Exemplo 5: Ensaio de atividade de estirpe de leveduras expressando trealase**

**[0577]** Levedura expressando um gene de trealase de *Corynascus sepedonium* dirigido pelo promotor ADH1 (*supra*) foi cultivada durante a noite em meio YPD padrão contendo glucose a 2%. O meio de levedura cultivado foi centrifugado a 3000 rpm durante 10 min para se coletar o sobrenadante. O sobrenadante foi usado para ensaio de atividade de enzima, como descrito em baixo.

**[0578]** A atividade de trealase foi detectada por medição da quantidade de glucose liberada através da hidrólise enzimática da trealose. O reagente glucose oxidase foi usado para medir a glucose com

base no desenvolvimento da cor a partir da aplicação do reagente. A intensidade da cor medida em um espectrofotômetro ou leitor de microplacas é proporcional à atividade de trealase. As condições de reação e o desenvolvimento de cor são descritos na Tabela 16 e Tabela 17, respectivamente.

**[0579]** A Unidade Novozymes de Trealase (TNU (A)) para o padrão de ensaio de trealase é medida em relação a um padrão de enzima de atividade declarada.

Tabela 16. Condição de reação de trealase

Quantidade de sobrenadante levedura	de	20 µL
Quantidade de substrato	de	100 µL
Substrato		Trealose, 60 mM
Tampão		Acetato de sódio, 0,1 M, Triton 100 a 0,01%
pH		5,0 ± 0,05
Temperatura de incubação	de	20 °C
Tempo de reação		2-3 hr
Gama do ensaio de trealase	de	0,004-0,017 TNU(A)/mL

Tabela 17. Desenvolvimento da cor

Mistura reacional		20 µL
Reagente oxidase	glucose	200 µL
Temperatura de incubação	de	20 °C

Tempo de reação	10-15 min
Comprimento de onda	490 nm

**[0580]** Os resultados do ensaio mostraram que a expressão de trealase aumentou proporcionalmente a glucose liberada, medida como a densidade óptica a 490 nm (0,10 para a estirpe de fundo não tendo o gene de trealase em comparação com 0,88 para a estirpe expressando trealase).

**Exemplo 6: Ensaio de atividade de estirpes de leveduras expressando trealase**

**[0581]** As estirpes de levedura do Exemplo 4 foram cultivadas durante a noite em meio YPD padrão contendo glucose a 2% ou 6%. O meio de levedura cultivado foi sujeito a centrifugação a 3500 rpm durante 10 min para se coletar o sobrenadante. O sobrenadante da cultura é usado para os ensaios de atividade de enzima descritos. A levedura pode ser também cultivada usando outros meios de cultivo tais como meio YNB mínimo ou mosto de milho liquefeito industrial clarificado e filtrado.

**[0582]** A atividade de glucoamilase foi medida usando maltose como substrato como descrito *supra*.

**[0583]** A atividade de trealase foi medida usando trealose como substrato. A hidrólise de enzimas da trealose liberará glucose como produto de reação que pode ser detectado usando estojos de ensaio comercialmente disponíveis tais como AUTOKIT GLICOSE C2 da Wako Diagnostics. Os reagentes proporcionados nos estojos de ensaio reagirão especificamente com glucose resultando em formação de cor. A intensidade da cor medida em espectrofotômetro ou leitor de microplacas é proporcional à atividade de trealase. As condições de reação são descritas na Tabela 18. A Unidade de Trealase (TNU) para a trealase padrão é definida como a

quantidade de enzima que hidrolisa uma micromole de trealose por minuto sob as condições padrão.

Tabela 18. Condições de reação de trealase

Quantidade apropriada de sobrenadante de levedura	10-200 µL
Substrato	trealose, 10 mM
Tampão	acetato, 0,1 M
pH	5,0 ± 0,05
Temperatura de incubação	32 °C
Tempo de reação	5-20 min
Gama do ensaio de trealase	0,002-0,036 TNU/mL

**[0584]** A absorvância a 505 nm aumenta à medida que a quantidade de glucoamilase ou trealase purificadas adicionadas para hidrolisar a maltose ou a trealose, respectivamente, em glucose aumenta. Uma curva padrão para a glucoamilase e a trealase purificadas foi gerada e usada para se estimar a atividade de glucoamilase e trealase em sobrenadantes de levedura.

**[0585]** Os resultados para a atividade de trealase e a atividade de glucoamilase são mostrados na Tabela 19. Uma representação gráfica da atividade comparativa de trealase é mostrada na Figura 2.

Tabela 19. Atividade de trealase e glucoamilase (GA) e secreção estimada de enzima.

No. da estirpe de levedura	Promotor para expressão de trealase	SEQ ID NO: (polipeptídeo maduro)	Organismo Dador (domínio catalítico)	Glucoamilase		Trealase	
				atividade	Conc. (ug/mL)	atividade	Conc. (ug/mL)

dur a							
1	Estirpe de fundo com gene de glucoamilase, sem gene de trealase			0,35 2	10,7	N/A	N/A
1	Estirpe de fundo com gene de glucoamilase, sem gene de trealase			0,31 2	8,9	N/A	N/A
1	Estirpe de fundo com gene de glucoamilase, sem gene de trealase			0,30 8	8,7	N/A	N/A
2	pPGK1	189	<i>Acidobacteri aceae bacterium</i>	0,35 7	10,9	0,05	1,04
3	pCCW1 2	191	<i>Acidovorax wautersii</i>	0,30 1	9,1	1,8	22,79
4	pPGK1	218	<i>Acremonium curvulum</i>	0,28 7	7,8	0,07	1,08
5	pCCW1 2	218	<i>Acremonium curvulum</i>	0,26 1	6,7	0,08	1,09
6	pCCW1 2	205	<i>Acremonium dichromosp orum</i>	0,30 4	8,6	0,78	3,78
7	pCCW1 2	205	<i>Acremonium dichromosp orum</i>	0,30 9	8,8	0,89	4,63
8	pCCW1 2	216	<i>Aspergillus cervinus</i>	0,29 9	8,4	0,12	1,18

9	pPGK1	216	<i>Aspergillus cervinus</i>	0,30 9	8,8	0,13	1,2
10	pPGK1	225	<i>Aspergillus niger</i>	0,30 8	8,8	0,48	2,22
11	pCCW1 2	225	<i>Aspergillus niger</i>	0,30 7	8,7	0,67	3,11
12	pPGK1	184	<i>Chaetomium jodhpurensense</i>	0,32 7	9,6	0,15	1,26
13	pCCW1 2	184	<i>Chaetomium jodhpurensense</i>	0,30 2	8,5	1,77	21,85
14	pCCW1 2	175	<i>Chaetomium megalocarpum</i>	0,30 2	8,5	1,77	21,85
15	pCCW1 2	175	<i>Chaetomium megalocarpum</i>	0,30 3	8,5	0,4	1,95
16	pPGK1	183	<i>Chaetomium nigricolor</i>	0,38 1	11,9	0,36	1,8
17	pCCW1 2	183	<i>Chaetomium nigricolor</i>	0,31 5	9,1	0,66	3,05
18	pCCW1 2	183	<i>Chaetomium nigricolor</i>	0,31 4	9	0,63	2,9
19	pPGK1	182	<i>Chaetomium sp</i>	0,32 9	9,7	0,14	1,22

20	pCCW1 2	182	<i>Chaetomiu m sp</i>	0,29 5	8,2	0,27	1,54
21	pCCW1 2	182	<i>Chaetomiu m sp</i>	0,30 1	8,4	0,12	1,18
22	pCCW1 2	201	<i>Chaetomiu m virescens</i>	0,34 1	10,2	0,53	2,43
23	pPGK1	201	<i>Chaetomiu m virescens</i>	0,31 8	9,2	0,23	1,43
24	pPGK1	187	<i>Chloridium virescens</i>	0,31 7	9,2	0,24	1,47
25	pPGK1	187	<i>Chloridium virescens</i>	0,34 9	10,5	0,15	1,25
26	pCCW1 2	211	<i>Dioszegia cryoxerica</i>	0,29 9	8,4	0,07	1,07
27	pPGK1	211	<i>Dioszegia cryoxerica</i>	0,38 4	12,1	0,44	2,07
28	pPGK1	177	<i>Doratomyce s sp</i>	0,31 4	9	1,08	6,44
29	pCCW1 2	177	<i>Doratomyce s sp</i>	0,31 5	8,9	1,87	25,75
30	pPGK1	194	<i>Enterobacte r sp</i>	0,32 1	9,3	0,9	4,71
31	pCCW1 2	194	<i>Enterobacte r sp</i>	0,31 2	8,9	1,87	25,75
32	pPGK1	206	<i>Fusarium sambucinu m</i>	0,32 3	9,4	0,05	1,05

33	pCCW1 2	188	<i>Gelasinospora cratophora</i>	0,31 5	9,1	0,66	3,05
34	pCCW1 2	188	<i>Gelasinospora cratophora</i>	0,33	9,7	0,8	3,91
35	pPGK1	213	<i>Hamigera avellanea</i>	0,30 2	8,5	0,63	2,89
36	pCCW1 2	213	<i>Hamigera avellanea</i>	0,32 4	9,4	0,98	5,38
37	pCCW1 2	213	<i>Hamigera avellanea</i>	0,32 2	9,4	1,13	6,97
38	pPGK1	213	<i>Hamigera avellanea</i>	0,32 6	9,5	0,1	1,15
39	pPGK1	193	<i>Kosakonia sacchari</i>	0,35 6	10,8	0,12	1,19
40	pCCW1 2	193	<i>Kosakonia sacchari</i>	0,32 3	9,4	0,06	1,07
41	pPGK1	176	<i>Lecanicillium psalliotae</i>	0,34 9	10,5	0,15	1,25
42	pCCW1 2	176	<i>Lecanicillium psalliotae</i>	0,33 3	9,9	0,07	1,08
43	pPGK1	176	<i>Lecanicillium psalliotae</i>	0,44 6	14,8	0,06	1,06
44	pCCW1 2	208	<i>Lentinus similis</i>	0,31 4	9	0,05	1,05
45	pPGK1	208	<i>Lentinus similis</i>	0,32 8	9,6	0,05	1,04



46	pPGK1	204	<i>Moelleriella libera</i>	0,30 4	8,6	0,13	1,21
47	pPGK1	198	<i>Corynascus sepedonium</i>	0,32 2	9,4	0,58	2,66
48	pCCW1 2	198	<i>Corynascus sepedonium</i>	0,31 6	9,1	1,13	6,98
49	pCCW1 2	198	<i>Corynascus sepedonium</i>	0,31	8,8	1,2	7,9
50	pCCW1 2	178	<i>Mucor moelleri</i>	0,31 6	9,1	0,34	1,75
51	pPGK1	178	<i>Mucor moelleri</i>	0,30 5	8,7	0,78	3,77
52	pCCW1 2	186	<i>Myceliophthora hinnulea</i>	0,29 9	8,4	0,99	5,51
53	pPGK1	186	<i>Myceliophthora hinnulea</i>	0,34 8	10,5	0,38	1,85
54	pPGK1	203	<i>Myceliophthora sepedonium</i>	0,32 8	9,6	0,43	2,04
55	pPGK1	203	<i>Myceliophthora sepedonium</i>	0,32 8	9,6	0,41	1,97
56	pCCW1 2	203	<i>Myceliophthora sepedonium</i>	0,31	8,8	1,2	7,9
57	pPGK1	199	<i>Myceliophthora thermophila</i>	0,32 9	9,7	0,3	1,63

58	pCCW1 2	199	<i>Myceliophth ora thermophila</i>	0,31 5	9,1	0,74	3,54
59	pPGK1	220	<i>Penicillium sp</i>	0,33 4	9,9	0,05	1,04
60	pCCW1 2	220	<i>Penicillium sp</i>	0,32 6	9,6	0,05	1,04
61	pPGK1	179	<i>Phialophora cyclaminis</i>	0,34 5	10,4	0,2	1,36
62	pCCW1 2	179	<i>Phialophora cyclaminis</i>	0,31 5	9,1	0,66	3,05
63	pPGK1	207	<i>Phoma sp</i>	0,31 2	9	0,07	1,09
64	pCCW1 2	207	<i>Phoma sp</i>	0,28 3	7,7	0,09	1,12
65	pCCW1 2	217	<i>Rasamsonia brevistipitata</i>	0,29	8	0,61	2,79
66	pPGK1	217	<i>Rasamsonia brevistipitata</i>	0,32 1	9,3	0,54	2,46
67	pCCW1 2	202	<i>Rhodotherm us marinus</i>	0,31 9	9,3	0,3	1,61
68	pPGK1	202	<i>Rhodotherm us marinus</i>	0,31 8	9,2	0,12	1,18
69	pPGK1	195	<i>Saitozyma flava</i>	0,31 2	8,9	0,25	1,48
70	pCCW1 2	195	<i>Saitozyma flava</i>	0,30 9	8,8	0,39	1,9

71	pCCW1 2	221	<i>Talaromyces aurantiacus</i>	0,33 2	9,8	0,6	2,77
72	pPGK1	221	<i>Talaromyces aurantiacus</i>	0,34 5	10,4	0,57	2,63
73	pCCW1 2	223	<i>Talaromyces leycectanus</i>	0,29 3	8,1	1,13	7,02
74	pPGK1	223	<i>Talaromyces leycectanus</i>	0,31 1	8,9	0,71	3,35
75	pPGK1	219	<i>Talaromyces piceae</i>	0,30 8	8,8	0,45	2,11
76	pCCW1 2	219	<i>Talaromyces piceae</i>	0,31 3	9	0,73	3,48
77	pCCW1 2	222	<i>Talaromyces pinophilus</i>	0,36 8	11,4	1,29	9,26
78	pCCW1 2	222	<i>Talaromyces pinophilus</i>	0,30 2	8,5	0,95	5,13
79	pPGK1	222	<i>Talaromyces pinophilus</i>	0,36 5	11,3	0,62	2,87
80	pCCW1 2	224	<i>Talaromyces variabilis</i>	0,29 6	8,2	0,06	1,06
81	pPGK1	181	<i>Thielavia antarctica</i>	0,33 8	10,1	0,08	1,1
82	pCCW1 2	180	<i>Thielavia arenaria</i>	0,30 2	8,5	0,25	1,5

83	pCCW1 2	215	<i>Trichoderma lixii</i>	0,30 6	8,7	0,06	1,06
84	pPGK1	215	<i>Trichoderma lixii</i>	0,32 6	9,5	0,1	1,15
85	pPGK1	200	<i>Trichoderma reesei GH37</i>	0,31 2	8,9	0,05	1,04
86	pCCW1 2	226	<i>Trichoderma reesei GH65</i>	0,30 1	8,4	0,18	1,3
87	pCCW1 2	192	<i>Xanthomon as arboricola</i>	0,30 6	8,7	0,46	2,17
88	pPGK1	192	<i>Xanthomon as arboricola</i>	0,33	9,7	0,09	1,12

**Exemplo 7: Sacarificação e fermentação simultâneas (SSF)**

**de estirpes de levedura expressando trealase**

**[0586]** A sacarificação e fermentação simultâneas (SSF) foram realizadas através de fermentações em mini-escala para várias estirpes expressando trealase descritas *supra* usando mosto de milho industrial (Avantec® Amp, Novozymes, A/S). As estirpes de levedura foram cultivadas durante a noite em meio YPD com glucose a 2% durante 24 horas a 30 °C e 300 rpm. O mosto de milho foi suplementado com 250 ppm de ureia, doseado com 0,15 AGU/g de DS de um produto de enzima glucoamilase exógena (Spirizyme® Excel, Novozymes, A/S) e trealose a 30 mM. Aproximadamente 0,6 mg de mosto de milho foram distribuídos por poço a placas de microtitulação de 96 poços, seguido pela adição de aproximadamente 10<sup>8</sup> células de levedura/g de mosto de milho a partir da cultura durante a noite. As placas foram incubadas a 32 °C sem agitação. Duplicados de cada estirpe foram analisados após fermentações de 48

horas. A fermentação foi conduzida usando as condições mostradas na Tabela 20 e, depois, parada pela adição de 100 µL de H<sub>2</sub>SO<sub>4</sub> a 8%, seguida por centrifugação a 3000 rpm durante 10 min. O sobrenadante foi analisado quanto à trealose usando HPLC.

Tabela 20. Condições de reação de fermentação em placas de microtitulação

Substrato	Mosto de milho Avantec® Amp
Inoculação de levedura	10 <sup>8</sup> células/g de mosto de milho
Ureia suplementar	250 ppm
Trealose suplementar	30 mM
Dose de produto de glucoamilase exógena	0,15 AGU/g de DS
pH	5,0 ± 0,05
Temperatura de incubação	32 °C
Tempo de reação	48 horas

**[0587]** Como mostrado na Tabela 21 foi obtida redução mais elevada na trealose a partir da levedura expressando uma trealase heteróloga em comparação com levedura não tendo expressão de trealase heteróloga às 48 horas de fermentação e sacarificação simultâneas (SSF).

Tabela 21.

Promotor	SEQ ID NO: (polipeptídeo maduro)	Organismo Dador de Trealase	Média (% de trealase p/v)
Estirpe de fundo com gene de glucoamilase, sem gene de trealase			0,96
pCCW12	191	<i>Acidovorax wautersii</i>	0,23
pCCW12	218	<i>Acremonium curvulum</i>	0,29

pCCW12	205	<i>Acremonium dichromosporum</i>	0,12
pCCW12	216	<i>Aspergillus cervinus</i>	0,23
pCCW12	225	<i>Aspergillus niger</i>	0,22
pCCW12	184	<i>Chaetomium jodhpurensense</i>	0,09
pCCW12	175	<i>Chaetomium megalocarpum</i>	0,14
pCCW12	183	<i>Chaetomium nigricolor</i>	0,12
pCCW12	182	<i>Chaetomium sp</i>	0,19
pCCW12	201	<i>Chaetomium virescens</i>	0,13
pCCW12	211	<i>Dioszegia cryoxerica</i>	0,25
pCCW12	177	<i>Doratomyces sp</i>	0,15
pCCW12	194	<i>Enterobacter sp</i>	0,06
pCCW12	188	<i>Gelasinospora cratophora</i>	0,12
pCCW12	213	<i>Hamigera avellanea</i>	0,10
pCCW12	193	<i>Kosakonia sacchari</i>	0,23
pCCW12	176	<i>Lecanicillium psalliotae</i>	0,35
pCCW12	208	<i>Lentinus similis</i>	0,43
pCCW12	198	<i>Corynascus sepedonium</i>	0,13
pCCW12	178	<i>Mucor moelleri</i>	0,13
pCCW12	186	<i>Myceliophthora hinnulea</i>	0,14
pCCW12	203	<i>Myceliophthora sepedonium</i>	0,13
pCCW12	199	<i>Myceliophthora thermophila</i>	0,11
pCCW12	220	<i>Penicillium sp</i>	0,95
pCCW12	179	<i>Phialophora cyclaminis</i>	0,09
pCCW12	207	<i>Phoma sp</i>	0,33
pCCW12	217	<i>Rasamsonia brevistipitata</i>	0,16
pCCW12	202	<i>Rhodothermus marinus</i>	0,42
pCCW12	195	<i>Saitozyma flava</i>	0,19

pCCW12	221	<i>Talaromyces aurantiacus</i>	0,15
pCCW12	223	<i>Talaromyces leycettanus</i>	0,19
pCCW12	219	<i>Talaromyces piceae</i>	0,19
pCCW12	222	<i>Talaromyces pinophilus</i>	0,16
pCCW12	224	<i>Talaromyces variabilis</i>	0,99
pCCW12	180	<i>Thielavia arenaria</i>	0,11
pCCW12	215	<i>Trichoderma lixii</i>	0,88
pCCW12	226	<i>Trichoderma reesei GH65</i>	0,27
pCCW12	192	<i>Xanthomonas arboricola</i>	0,36
pPGK1	189	<i>Acidobacteriaceae bacterium</i>	0,30
pPGK1	218	<i>Acremonium curvulum</i>	0,40
pPGK1	216	<i>Aspergillus cervinus</i>	0,29
pPGK1	225	<i>Aspergillus niger</i>	0,11
pPGK1	184	<i>Chaetomium jodhpurensense</i>	0,17
pPGK1	183	<i>Chaetomium nigricolor</i>	0,39
pPGK1	182	<i>Chaetomium sp</i>	0,16
pPGK1	201	<i>Chaetomium virescens</i>	0,17
pPGK1	187	<i>Chloridium virescens</i>	0,20
pPGK1	211	<i>Dioszegia cryoxerica</i>	0,15
pPGK1	177	<i>Doratomyces sp</i>	0,19
pPGK1	194	<i>Enterobacter sp</i>	0,18
pPGK1	206	<i>Fusarium sambucinum</i>	0,26
pPGK1	213	<i>Hamigera avellanea</i>	0,26
pPGK1	193	<i>Kosakonia sacchari</i>	0,25
pPGK1	176	<i>Lecanicillium psalliotae</i>	0,24
pPGK1	208	<i>Lentinus similis</i>	0,15
pPGK1	204	<i>Moelleriella libera</i>	0,18
pPGK1	198	<i>Corynascus sepedonium</i>	0,18

pPGK1	178	<i>Mucor moelleri</i>	0,20
pPGK1	186	<i>Myceliophthora hinnulea</i>	0,17
pPGK1	203	<i>Myceliophthora sepedonium</i>	0,16
pPGK1	199	<i>Myceliophthora thermophila</i>	0,16
pPGK1	220	<i>Penicillium sp</i>	1,19
pPGK1	179	<i>Phialophora cyclaminis</i>	0,20
pPGK1	207	<i>Phoma sp</i>	0,32
pPGK1	217	<i>Rasamsonia brevistipitata</i>	0,11
pPGK1	202	<i>Rhodothermus marinus</i>	0,44
pPGK1	195	<i>Saitozyma flava</i>	0,19
pPGK1	221	<i>Talaromyces aurantiacus</i>	0,16
pPGK1	223	<i>Talaromyces leycettanus</i>	0,13
pPGK1	219	<i>Talaromyces piceae</i>	0,20
pPGK1	222	<i>Talaromyces pinophilus</i>	0,11
pPGK1	181	<i>Thielavia antarctica</i>	0,25
pPGK1	215	<i>Trichoderma lixii</i>	0,35
pPGK1	200	<i>Trichoderma reesei GH37</i>	0,42
pPGK1	192	<i>Xanthomonas arboricola</i>	0,42



## REIVINDICAÇÕES

1. Método de produção de um produto de fermentação a partir de um material contendo amido ou contendo celulose, caracterizado por compreender:

(a) sacarificação do material contendo amido ou contendo celulose; e

(b) fermentação do material sacarificado do passo (a) com um organismo fermentador;

em que o organismo fermentador compreende um polinucleotídeo heterólogo codificando uma alfa-amilase ou um polinucleotídeo heterólogo codificando uma trealase.

2. Método, de acordo com a reivindicação 1, caracterizado por a alfa-amilase ter uma sequência de polipeptídeo maduro com pelo menos 60%, p.ex., 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99% ou 100% de identidade de sequências, com a sequência de aminoácidos de qualquer uma de SEQ ID NOs: 76-101, 121-174 e 231.

3. Método, de acordo com qualquer uma das reivindicações 1 ou 2, caracterizado por a alfa-amilase ter uma sequência de polipeptídeo maduro compreendendo a ou consistindo na sequência de aminoácidos de qualquer uma de SEQ ID NOs: 76-101, 121-174 e 231.

4. Método, de acordo com qualquer uma das reivindicações 1 a 3, caracterizado por a trealase ter uma sequência de polipeptídeo maduro com pelo menos 60%, p.ex., 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99% ou 100% de identidade de sequências, com a sequência de aminoácidos de qualquer uma de SEQ ID NOs: 175-226.

5. Método, de acordo com qualquer uma das reivindicações 1 a 4, caracterizado por a trealase ter uma sequência de polipeptídeo maduro compreendendo a ou consistindo na sequência de aminoácidos de qualquer uma de SEQ ID NOs: 175-226.

6. Método, de acordo com qualquer uma das reivindicações 1 a 5, caracterizado por a sacarificação do passo (a) ocorrer em um material contendo amido, e em que o material contendo amido é amido gelatinizado ou não gelatinizado.

7. Método, de acordo com a reivindicação 6, caracterizado por compreender liquefação do material que contém amido por contato do material com uma alfa-amilase antes da sacarificação.

8. Método, de acordo com a reivindicação 7, caracterizado por a liquefação do material contendo amido e/ou a sacarificação do material contendo amido serem conduzidas na presença de protease exogenamente adicionada.

9. Método, de acordo com qualquer uma das reivindicações 1 a 8, caracterizado por a fermentação e a sacarificação serem realizadas simultaneamente em uma sacarificação e fermentação simultâneas (SSF).

10. Método, de acordo com qualquer uma das reivindicações 1 a 9, caracterizado por compreender recuperação do produto de fermentação a partir da fermentação.

11. Método, de acordo com qualquer uma das reivindicações 1 a 10, caracterizado por o produto de fermentação ser etanol.

12. Método, de acordo com qualquer uma das reivindicações 1 a 11, caracterizado por o organismo fermentador compreender um polinucleotídeo heterólogo codificando uma glucoamilase.

13. Método, de acordo com qualquer uma das reivindicações 1 a 12, caracterizado por o organismo fermentador compreender um polinucleotídeo heterólogo codificando uma protease.

14. Método, de acordo com qualquer uma das reivindicações 1 a 13, caracterizado por a sacarificação do passo (a) ocorrer em um material contendo celulose, e em que o material contendo celulose é pré-tratado.

15. Método, de acordo com qualquer uma das reivindicações 1 a 14, caracterizado por o organismo fermentador ser uma célula de *Saccharomyces cerevisiae*.

16. Célula de levedura recombinante caracterizada por compreender um polinucleotídeo heterólogo codificando uma alfa-amilase ou um polinucleotídeo heterólogo codificando uma trealase.

17. Célula de levedura recombinante, de acordo com a reivindicação 16, caracterizada por a alfa-amilase ter uma sequência de polipeptídeo maduro com pelo menos 60%, p.ex., 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99% ou 100% de identidade de sequências, com a sequência de aminoácidos de qualquer uma de SEQ ID NOs: 76-101, 121-174 e 231.

18. Célula de levedura recombinante, de acordo com qualquer uma das reivindicações 16 ou 17, caracterizada por a alfa-amilase ter uma sequência de polipeptídeo maduro compreendendo a ou consistindo na sequência de aminoácidos de qualquer uma de SEQ ID NOs: 76-101, 121-174 e 231.

19. Célula de levedura recombinante, de acordo com qualquer uma das reivindicações 16 a 18, caracterizada por a trealase ter sequência de polipeptídeo maduro com pelo menos 60%, p.ex., 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99% ou 100% de identidade de sequência, com a sequência de aminoácidos de qualquer uma de SEQ ID NOs: 175-226.

20. Célula de levedura recombinante, de acordo com qualquer uma das reivindicações 16 a 19, caracterizada por a trealase ter uma sequência de polipeptídeo maduro compreendendo a ou consistindo na sequência de aminoácidos de qualquer uma de SEQ ID NOs: 175-226.

21. Célula de levedura recombinante, de acordo com qualquer uma das reivindicações 16 a 20, caracterizada por o organismo fermentador compreender um polinucleotídeo heterólogo codificando uma glucoamilase.

22. Célula de levedura recombinante, de acordo com qualquer uma das reivindicações 16 a 21, caracterizada por o organismo fermentador compreender um polinucleotídeo heterólogo codificando uma protease.

23. Levedura recombinante, de acordo com qualquer uma das reivindicações 16 a 22, caracterizada por a célula ser uma célula de *Saccharomyces cerevisiae*.

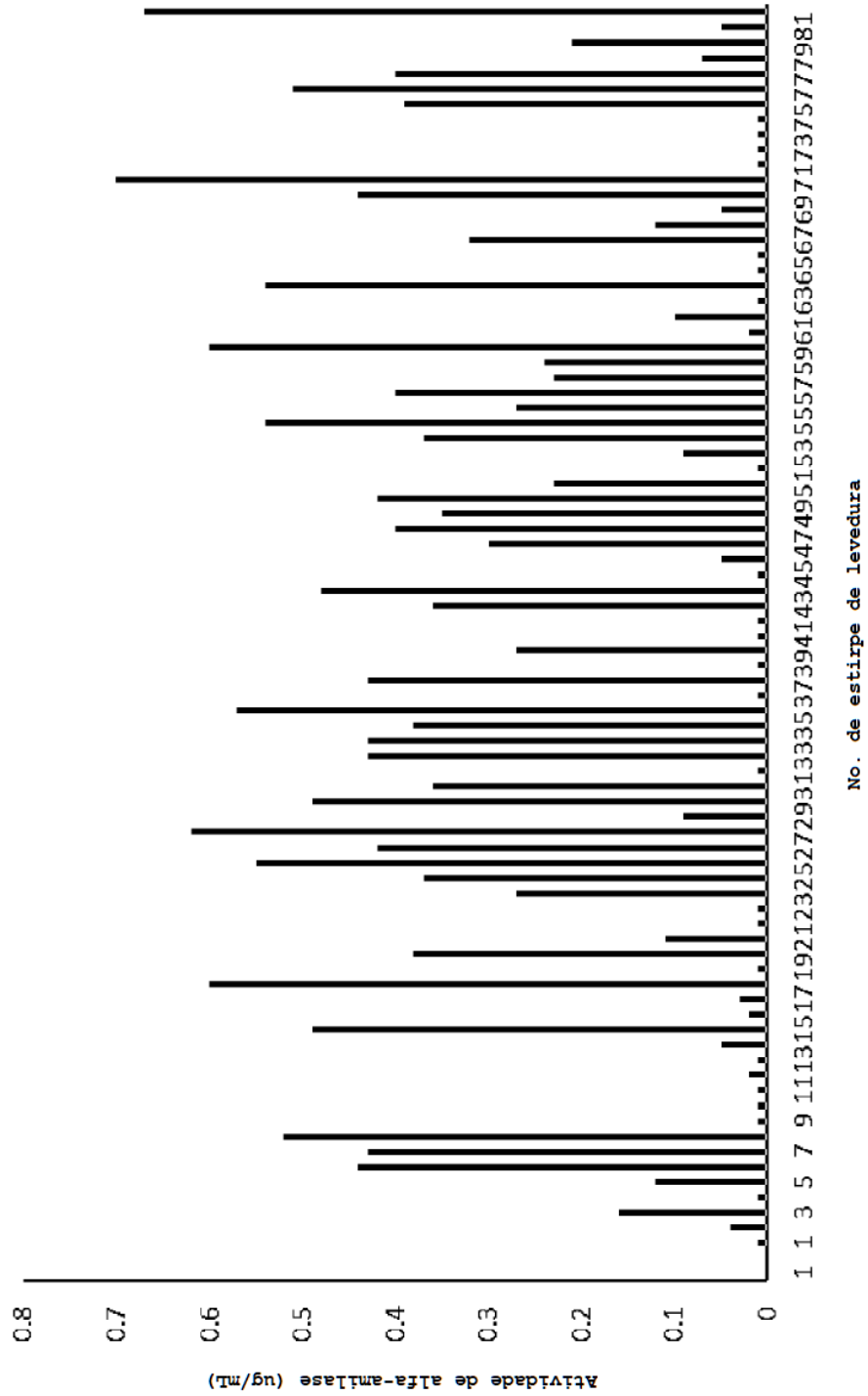


FIG. 1

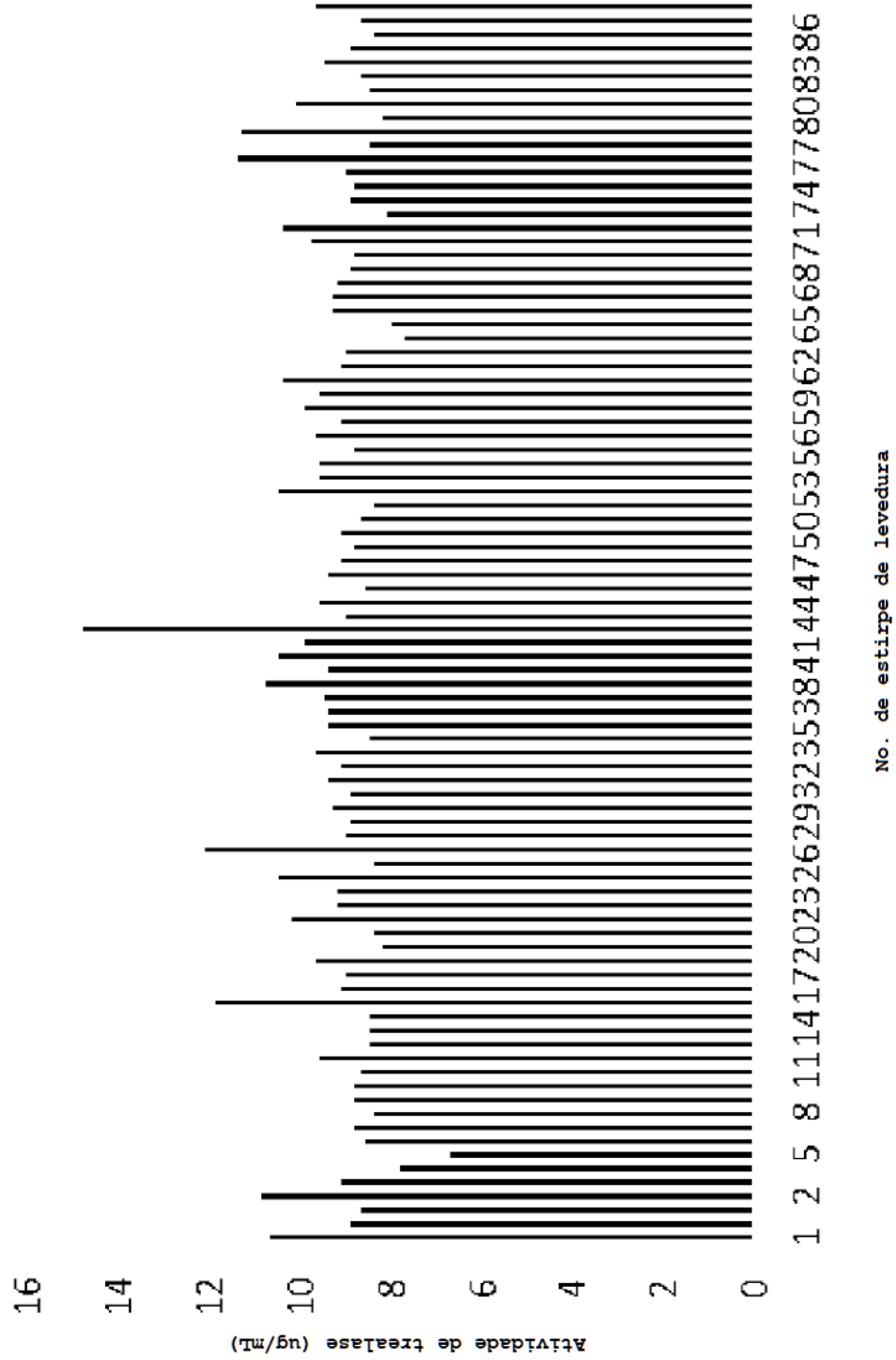


FIG. 2

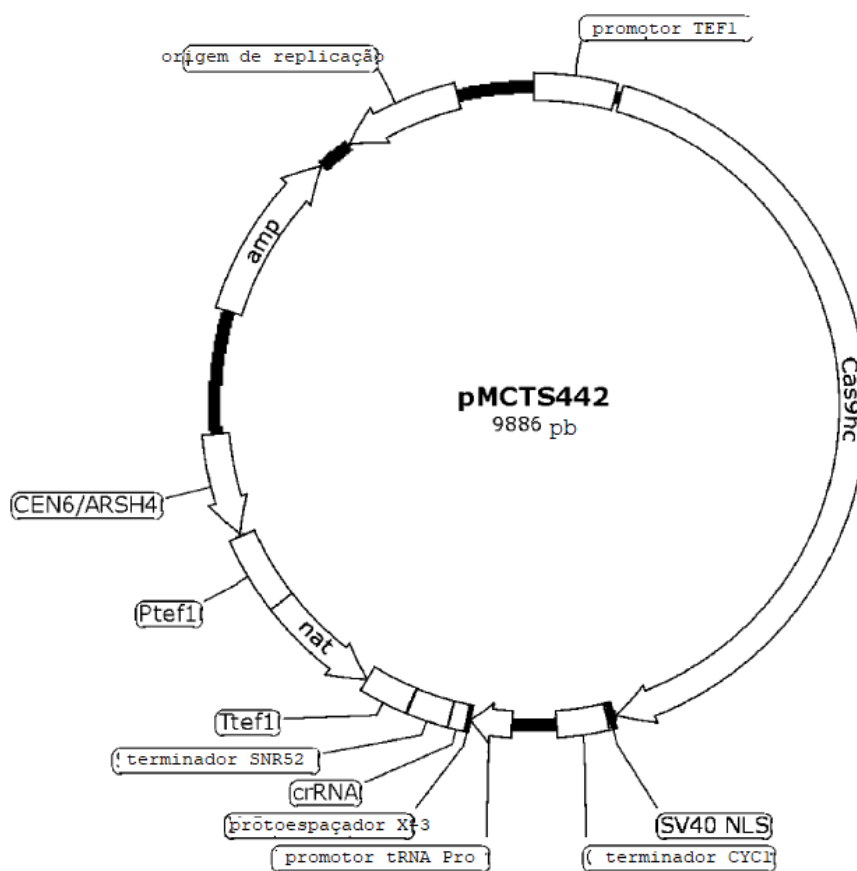


FIG. 3

## **RESUMO**

Patente de invenção para **“LEVEDURA EXPRESSANDO ENZIMAS PARA PRODUÇÃO DE ETANOL”**

São descritos aqui organismos fermentadores recombinantes tendo um polinucleotídeo heterólogo codificando uma alfa-amilase e/ou um polinucleotídeo heterólogo codificando uma trealase. São também descritos processos para produção de um produto de fermentação, tal como etanol, a partir de amido ou material contendo celulose com os organismos fermentadores recombinantes.



Este anexo apresenta o código de controle da listagem de sequências biológicas.

### Código de Controle

Campo 1



Campo 2



### Outras Informações:

- Nome do Arquivo: 1583011ENBR\_SEQUENCES.txt
- Data de Geração do Código: 22/02/2021
- Hora de Geração do Código: 17:11:32
- Código de Controle:
  - Campo 1: 46DEE0DD9D392C4C
  - Campo 2: 507E311C20543769