



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

(11) PI 0819418-1 B1



(22) Data do Depósito: 09/12/2008

(45) Data de Concessão: 10/05/2022

(54) Título: MÉTODO PARA A PRODUÇÃO DE ETANOL

(51) Int.Cl.: C12N 1/19; C12N 1/16; C12P 7/06; C12N 15/63; C12N 15/81; (...).

(30) Prioridade Unionista: 13/12/2007 US 61/007,477.

(73) Titular(es): ARCHER-DANIELS-MIDLAND COMPANY..

(72) Inventor(es): CHARLES A. ABBAS; ANDRIY A. SIBIRNY; ANDRIY Y. VORONOVSKY; OLENA P. ISHCHUK.

(86) Pedido PCT: PCT US2008086041 de 09/12/2008

(87) Publicação PCT: WO 2009/076343 de 18/06/2009

(85) Data do Início da Fase Nacional: 11/06/2010

(57) Resumo: LEVEDURA PARA A PRODUÇÃO DE ETANOL ISOLADA DA ESPÉCIE HANSENULA POLYMORPHA, MÉTODO PARA A PRODUÇÃO DE ETANOL E CEPA DE H. POLYMORPHA. Trata-se de métodos e composições para a produção do etanol a partir de matérias primas lignocelulósicas. As realizações apresentam células de levedura do gênero H.polymorpha com uma ou mais modificações, incluindo, por exemplo, um gene de trehalase ácida inativo, superexpressão de xiluloquinase , e/ou superexpressão de proteína de choque térmico 104.

MÉTODO PARA A PRODUÇÃO DE ETANOL

CAMPO DA INVENÇÃO

[001] As realizações aqui apresentadas se referem a métodos para a produção fermentativa do etanol a partir de D-xilose ao utilizar levedura.

ANTECEDENTES

[002] A engenharia metabólica de microorganismos é frequentemente um meio eficaz para produzir comercialmente uma série de produtos químicos que podem ser utilizados para múltiplas aplicações (consultar, por exemplo, Lee, S.Y., et al. Macromol. Biosci. 4:157-164 (2004)). Um produto químico que ganhou muito interesse é o etanol. Embora o etanol tenha um número de usos, ele é utilizado de maneira mais comum como combustível. Como um combustível, o etanol é um produto de baixo valor com em que uma grande parte do custo de sua produção é atribuída ao custo de matérias-primas. Seria desejável, portanto, o desenvolvimento de etanolgenos e processos de fermentação para a produção do etanol a partir de matérias primas imediatamente disponíveis e baratos.

[003] Estas matérias primas podem ser, por exemplo, lignocelulósicos. Estes lignocelulósicos podem ser derivados, por exemplo, de correntes residuais de biomassa renovável de alimentos, operações de formação de polpa de papel, resíduos agrícolas e papel reciclado de municipalidades.

[004] A lignocelulose consiste em aproximadamente 30% de D-xilose (consultar Ryabova, O. B., et al. "Xylose and Cellobiose Fermentation to Ethanol by the Thermotolerant Methylo-trophic Yeast Hansenula polymorpha",

FEMS Yeast Res. 4:157-164 (2003)). A xilose é um "açúcar da madeira" com a designação IUPAC (2S,3R,4S,5R)-oxano-2,3,4,5-tetrol.

[005] Somente um número relativamente pequeno de microorganismos do tipo selvagem pode fermentar a D-xilose. Estes microorganismos não são geralmente apropriados para a fermentação em grande escala. Esta desfavorabilidade pode ocorrer, por exemplo, em consequência da não-familiaridade com os microorganismos, a dificuldade na obtenção dos microorganismos, a fraca produtividade e/ou crescimento em lignocelulósicos pré-tratados ou o rendimento insatisfatório quando cultivados em açúcares misturados derivados de biomassa. C. Abbas, *"Lignocellulosics to ethanol: meeting ethanol demand in the future"*, The Alcohol Textbook, 4ª edição, (K.A. Jacques, T.P. Lyons and D.R. Kelsall, Eds). Nottingham University Press, Nottingham, UK, 2003, páginas 41-57; Cite 2 C. Abbas, *"Emerging biorefineries and biotechnological applications of nonconventional yeast: now and in the future"*, The Alcohol Textbook, 4ª edição. (K.A. Jacques, T.P. Lyons and D.R. Kelsall, eds). Nottingham University Press, Nottingham, UK, 2003, paginas 171-191.

[006] As leveduras são consideradas os microorganismos mais promissores para a fermentação alcoólica da xilose (consultar Ryabova, supra). Elas têm células maiores do que as bactérias, são resistentes à infecção viral, e tendem a ser mais resistentes à realimentação negativa do etanol. Além disso, o crescimento e o metabolismo da levedura foram estudados extensivamente para uma série de espécies. Uma série de leveduras é conhecida como de fermentação natural da D-xilose. Estas incluem *Pichia*

stipitis, *Candida shehatae* e *Pachysolen tannophilus* (consultar Ryabova, supra; C. Abbas 2003). A levedura de cervejaria comum, *Saccharomyces cerevisiae*, não conhecida como de fermentação natural da D-xilose, mas uma série de cepas de *S. cerevisiae* biologicamente elaboradas que fermentam a D-xilose foi relatada.

[007] Conforme mostrado na Figura 1, foi relatado que o metabolismo da D-xilose em levedura prossegue ao longo de uma rota similar àquela da glicose através da rota de fosfato de pentose. O carbono da D-xilose é processado como etanol através do ciclo glicolítico ou como CO₂ através do ciclo de TCA respiratório.

[008] Foi proposto que um obstáculo envolvido na fermentação da D-xilose é a hidrólise do xilano, que é o componente principal da hemicelulose para os monossacarídeos (consultar Ryabova, supra). Uma abordagem para superar este obstáculo é a utilização de "sacarificação e fermentação simultâneas" (SSF). Este é um processo em que a lignocelulose pré-tratada é hidrolisada por celulasas e hemicelulasas, ao passo que as hexoses e as pentoses produzidas por esta hidrólise (incluindo a xilose) são fermentadas como etanol. Isto deve permitir a conversão contínua dos açúcares em etanol, impedindo a sua acumulação no meio.

[009] Um inconveniente potencial da SSF é a diferença na temperatura ideal na qual as celulasas e as hemicelulasas são ativas (pelo menos aproximadamente 50°C) que são compatíveis com a temperatura ideal para o crescimento da levedura e a fermentação da xilose (aproximadamente 30°C). Uma solução para este inconveniente potencial consiste em executar a SSF utilizando a levedura

metilotrófica termotolerante *Hansenula polymorpha* (também conhecida como *Pichia angusta*). Foi relatado que esta levedura tem temperaturas de crescimento ideal e máxima de 37°C e 48°C, respectivamente. Estas temperaturas são mais elevadas do que aquelas toleradas pela maioria das outras leveduras produtoras de etanol (Ryabova, et al.). Além disso, Ryabova, et al. Relataram que, sob algumas condições, a *H. polymorpha* pode fermentar naturalmente a D-xilose (consultar também Voronovsky, A.Y., et al., "Expression os xylA Genes Encoding Xylose Isomerases From *Escherichia coli* and *Streptomyces coelicolor* in the Methyloleophilic Yeast *Hansenula polymorpha*" FEMS Yeast Res. 5(11):1055-62 (2005)). O comportamento da *H. polymorpha* sob altas temperaturas é relatado, por exemplo, Escalante, J., et al., "Biomass Production by a Thermotolerant Yeast: *Hansenula polymorpha*" J. Chem. Biotechnol. 48:61-70 (1990); Tsiomenko, A.B., et al., "Secretory Heat-Shock Protein of the Thermotolerant Yeast *Hansenula polymorpha*. Identification and Comparative Characteristics" Biochemistry (Moscow) 62(2):123-128 (1997); Lindquist, S. & Kim, G., "Heat-shock Protein 104 Expression is Sufficient for Thermotolerance in Yeast" Proc. Natl. Acad. Sci. USA 93:5301-5306 (1996); Guerra, E., et al. "Hypoxia Abolishes Transience of the Heat-shock Response in Methyloleophilic Yeast *Hansenula polymorpha*" Microbiology 151:805-811 (2005).

[0010] Portanto, seria vantajoso o desenvolvimento de cepas de *H. polymorpha* com uma maior capacidade de produzir o etanol a partir de matérias primas lignocelulósicas, incluindo o açúcar C5, D-xilose. Os presentes ensinamentos podem propiciar estas vantagens e/ou

outras, e podem propiciar vantagens adicionais que um elemento versado na técnica irá discernir prontamente da descrição detalhada a seguir.

DESCRIÇÃO RESUMIDA DA INVENÇÃO

[0011] O presente ensinamento descreve diversas características e aspectos diferentes da invenção com referência às várias realizações exemplificadoras. Deve ser compreendido, no entanto, que a invenção engloba numerosas realizações alternativas, as quais podem ser praticadas ao combinar qualquer uma das características e aspectos diferentes aqui descritos em qualquer combinação que um elemento versado na técnica consideraria útil.

[0012] São aqui apresentados genes e elementos genéticos úteis na modificação de células hospedeiras. Estas células hospedeiras podem incluir, por exemplo, microorganismos. Um microorganismo particularmente apropriado para o uso nas realizações da invenção é a levedura *H. polymorpha*. Os métodos e as composições da invenção são úteis para a obtenção de microorganismos com uma maior atividade de enzimas. Em uma realização, uma célula hospedeira de *H. polymorpha* superexpressa a proteína 104 de choque térmico de *H. polymorpha*. Em uma realização adicional, é apresentada uma célula de *H. polymorpha* que tem uma deleção (trehalase ácida) do gene ATH1. A deleção de ATH1 melhora a termotolerância e a tolerância ao etanol, tendo por resultado um aumento da produção do etanol. Em uma realização adicional, uma cepa de *H. polymorpha* superexpressa a xiluloquinase (XYL3). Em ainda uma outra realização, uma cepa da levedura da invenção tem pelo menos duas das seguintes propriedades: (a) a levedura superexpressa a proteína 104 de choque térmico; (b) a

levedura tem um gene ATH1 inativo; e (c) a levedura superexpressa a xiluloquinase.

[0013] Uma realização adicional apresenta processos para a produção do etanol através da fermentação de uma ou mais cepas da levedura da invenção.

BREVE DESCRIÇÃO DE VÁRIAS VISTAS DOS DESENHOS

[0014] A Figura 1 é um esquema do metabolismo da xilose e da glicose nas leveduras.

DESCRIÇÃO DETALHADA DA INVENÇÃO

[0015] Tal como aqui utilizado, as formas no singular "um" "uma," e "o/a" utilizadas no relatório descritivo e nas reivindicações incluem o singular e o plural, a menos que o conteúdo dite claramente de alguma outra maneira. Particularmente, os elementos versados na técnica irão reconhecer que embora o desenho e a criação de materiais catalíticos e de suportes catalíticos tenham sido descritos em termos de uma única célula, sistemas mais eficazes irão incluir uma ou mais células, cada uma delas expressando uma ou mais proteínas do receptor.

[0016] São aqui apresentados métodos e composições de matéria útil em uma ou em mais dentre a termotolerância crescente e a inibição de realimentação decrescente pelo etanol em leveduras produtoras de etanol, tipicamente nos membros do gênero *Hansenula (Pichia)*, e mais tipicamente em *H. polymorpha (P. angusta)*. São providas novas cepas. Os métodos de produção do etanol utilizando as cepas da invenção também são providos.

[0017] Deve ser compreendido que determinadas descrições das realizações da invenção foram simplificadas para ilustrar somente os elementos e as limitações que são

relevantes para uma compreensão clara da presente invenção, enquanto são eliminados, para finalidades de clareza, outros elementos. Os elementos versados na técnica, levando em consideração a presente descrição, irão reconhecer que outros elementos e/ou limitações podem ser desejáveis para praticar as realizações da invenção. Uma vez que tais outros elementos e/ou limitações podem ser verificados por um elemento versado na técnica ao levar em consideração a presente descrição, e não são necessários para uma compreensão completa das realizações, uma discussão de tais elementos e limitações não é aqui fornecida. Além disso, a descrição aqui apresentada não se presta a limitar o âmbito das reivindicações.

[0018] O termo "gene" significa um segmento de ácido nucléico, DNA ou RNA, que codifica e é capaz de expressar um produto de gene específico. Um gene produz frequentemente uma proteína ou um polipeptídeo como seu produto do gene, mas em seu sentido mais amplo um gene pode produzir qualquer produto desejado, se o produto for uma proteína, um polipeptídeo ou um ácido nucléico. O ácido nucléico funcional ou estrutural, tal como, sem limitação, rRNA, ribozimas, RNA de antisentido ou RNA de interferência (por exemplo, siRNA) também pode ser considerado como "produtos de genes".

[0019] Um "gene" contém uma "seqüência expressa" que pode codificar não somente uma proteína ou um polipeptídeo, mas um ácido nucléico estrutural ou funcional, tal como um de antisentido ou um siRNA. Um gene também pode conter seqüências que contêm elementos reguladores, tais como, sem limitação, promotores, intensificadores e terminadores; tais elementos reguladores podem ser

"operavelmente ligados", mais tipicamente em uma proximidade apropriada entre si. Tais promotores operam em cis (ligados entre si na mesma molécula de ácido nucléico) para causar a expressão de "um produto do gene". A escolha de constituintes do gene, tais como a combinação particular de elementos reguladores e a sequência expressa, irá ditar as condições da expressão. Por exemplo, um promotor constitutivo, tal como o promotor de TEF1 (gene 1A do fator de alongamento da translação), acoplado a uma sequência expressa, irá causar a expressão constitutiva da sequência expressa quando transferido a uma célula hospedeira apropriada. Um "promotor constitutivo" é um promotor não regulado que permite a transcrição contínua de seu gene associado. Um promotor é considerado como constitutivo se funcionar de modo a promover a transcrição de um gene sob condições de crescimento normais. Um promotor constitutivo não é tipicamente específico de substrato e não varia substancialmente em sua expressão sob condições de crescimento normais.

[0020] Um "gene" pode incluir íntrons ou outras seqüências de DNA que podem ser emendados do transcrito de RNA final. Uma sequência de DNA expressa que codifica uma proteína ou peptídeo ("seqüência de codificação de proteína") inclui um quadro de leitura aberto (ORF). A sequência de codificação da proteína pode compreender íntrons de intervenção. Além disso, o termo "gene" inclui seqüências expressas, bem como seqüências não-expressas. Deve ser compreendido que todas as seqüências de DNA aqui fornecidas incluem cordões complementares, a menos que esteja indicado de alguma outra maneira. Além disso, as seqüências de RNA podem ser preparadas a partir das seqüências de DNA ao

substituir a uracila no lugar da timina, e são incluídas no âmbito desta definição e da invenção, juntamente com cópias de RNA das seqüências de DNA da invenção isoladas das células.

[0021] O termo "oligonucleotídeo" significa um ácido nucléico de aproximadamente 7 a aproximadamente 50 bases, embora seja mais tipicamente de aproximadamente 15 a aproximadamente 35 bases. Os oligonucleotídeos são úteis como sondas ou primers para o uso em ensaios de hibridização ou de amplificação, tais como transferência Southern ou Northern; baliza molecular; reação em cadeia de polimerase (PCR); PCR transcritiva reversa (RT-PCR); RT-PCR quantitativa (QRT-PCT), por exemplo, TAQMAN; métodos de amplificação isotérmica, tais como NASBA (amplificação baseada na seqüência de ácido nucléico); e amplificação de círculo de rolagem, incluindo o uso de sondas de cadeado. Os oligonucleotídeos da invenção podem ser modificados pela adição de peptídeos, etiquetas (incluindo etiquetas fluorescentes, de ponto de quantum, ou de enzima), e outras porções químicas, e deve ser compreendido que são incluídos no âmbito desta definição e da invenção.

[0022] Tal como aqui utilizado, no contexto das novas seqüências de nucleotídeos aqui descritas, um ácido nucléico é "específico para" uma determinada seqüência, tal como o cDNA de decarboxilase de piruvato e as seqüências genômicas providas, se puder hibridizar especificamente a uma determinada seqüência sob condições restritas, tais como, sem limitação, 0,2X SSC a 65° ou em uma reação de PCR sob temperaturas de reação típicas (recozimento). Tipicamente, uma seqüência é "específica" para uma seqüência de referência

se o ácido nucléico tiver uma homologia de 90 a 100% (identidade de seqüência) com a seqüência de referência.

[0023] Os seguintes termos são utilizados para descrever as relações das seqüências entre dois ou mais ácidos nucléicos ou polinucleotídeos: (a) "seqüência de referência", (b) "janela de comparação", (c) "identidade de seqüência", (d) "porcentagem da identidade de seqüência", e (e) "identidade substancial".

[0024] Tal como aqui utilizado, a "seqüência de referência" é uma seqüência definida utilizada como uma base para a comparação da seqüência. Uma seqüência de referência pode ser um subconjunto ou a totalidade de uma seqüência específica; por exemplo, como um segmento de uma seqüência de cDNA ou gene comprimento total, ou a seqüência de DNA ou gene completa.

[0025] Tal como aqui utilizado, a "janela de comparação" faz referência a um segmento contíguo e específico de uma seqüência de polinucleotídeo, em que a seqüência de polinucleotídeo na janela de comparação pode compreender adições ou deleções (isto é, aberturas) em comparação à seqüência de referência (que não compreende adições ou deleções) para o alinhamento ideal das duas seqüências. Geralmente, a janela de comparação tem um comprimento de pelo menos 20 nucleotídeos contíguos, e opcionalmente pode ter um comprimento de 30, 40, 50, 100 ou mais. Os elementos versados na técnica entendem que, para evitar uma grande similaridade com uma seqüência de referência devido à inclusão de aberturas na seqüência de polinucleotídeo, uma penalidade da abertura é tipicamente introduzida e subtraída do número de combinações. Os métodos

de alinhamento das seqüências para a comparação são bem conhecidos no estado da técnica. Desse modo, a determinação da porcentagem de identidade de seqüência entre quaisquer duas seqüências pode ser realizada ao utilizar um algoritmo matemático. Os exemplos não-limitadores de tais algoritmos matemáticos incluem o algoritmo de Myers e Miller (1988) CABIOS 4:11- 17; o algoritmo de alinhamento local de Smith et al. (1981) Adv. Appl. Math. 2:482; o algoritmo de alinhamento global de Needleman e Wunsch (1970) J. Mol. Biol. 48:443-453; o método de busca do alinhamento local de Pearson e Lipman (1988) Proc. Natl. Acad. Sci. 85:2444-2448; o algoritmo de Karlin e Altschul (1990) Proc. Natl. Acad. Sci. USA 87:2264, tal como modificado em Karlin e Altschul (1993) Proc. Natl. Acad. Sci. USA 90:5873-5877.

[0026] As implementações de computador destes algoritmos matemáticos podem ser utilizadas para a comparação das seqüências para determinar a identidade da seqüência. Tais implementações incluem, mas sem ficar a elas limitadas: CLUSTAL no programa PC/Gene (disponível junto à Intelligenetics, Mouth View, Calif.); o programa ALIGN (versão 2.0) e GAP, BESTFIT, BLAST, FASTA e TFASTA nos CGC Wisconsin Genetics Software Packages, versão 10 (disponível junto à Accelrys Inc., 9685 Scranton Road, San Diego, California, USA). Os alinhamentos que utilizam estes programas podem ser executados ao utilizar os parâmetros padrão. O programa CLUSTAL é bem descrito por Higgins et al. (1988) Gene 73:237-244; Higgins et al. (1989) CABIOS 5:151-153; Corpet et al. (1988) Nucleic Acids Res. 16:10881-90; Huang et al. (1992) CABIOS 8:155-65; e Pearson et al. (1994) Meth. Mol. Biol. 24:307-331. O programa ALIGN é baseado no

algoritmo de Myers e Miller (1988) supra. Uma tabela de resíduo de peso PAM120, uma penalidade do comprimento da abertura de 12, e uma penalidade da abertura de 4 podem ser utilizadas com o programa ALIGN quando da comparação das seqüências de aminoácidos. Os programas BLAST de Altschul et al. (1990) J. Mol. Biol. 215:403 são baseados no algoritmo de Karlin e Altschul (1990) supra.

[0027] As buscas de nucleotídeo BLAST podem ser executadas com o programa BLASTN, classificação = 100, comprimento de palavra = 12, para obter as seqüências de nucleotídeos homólogas a uma seqüência de nucleotídeo que codifica uma proteína da invenção. As buscas de proteína BLAST podem ser executadas com o programa BLASTX, classificação = 5, comprimento de palavra = 3, para obter as seqüências de aminoácidos homólogas a uma proteína ou um polipeptídeo da invenção. Para obter alinhamentos com aberturas para finalidades de comparação, BLAST com abertura (em BLAST 2.0) pode ser utilizado tal como descrita em Altschul et al. (1997) Nucleic Acids Res. 25:3389. Alternativamente, PSI-BLAST (em BLAST 2.0) pode ser utilizado para executar uma busca iterada que detecta relações distantes entre moléculas. Consultar Altschul et al. (1997) supra. Ao utilizar BLAST, BLAST com abertura, e PSI-BLAST, os parâmetros padrão dos programas respectivos (por exemplo, BLASTN para seqüências de nucleotídeos, BLASTX para proteínas) podem ser utilizados. Consultar o site na web do National Center for Biotechnology Information na rede mundial em ncbi.nlm.nih.gov. O alinhamento também pode ser executado manualmente por inspeção.

[0028] A menos que esteja indicado de alguma

outra maneira, os valores de identidade/similaridade de sequência aqui fornecidos referem-se ao valor obtido utilizando GAP versão 10 ao utilizar os seguintes parâmetros: % de identidade e % de similaridade para uma sequência de nucleotídeo utilizando um peso GAP de 50 e um peso de comprimento de 3 e a matriz de classificação de nswgapdna.cmp; ou qualquer programa equivalente do mesmo.

[0029] "Programa equivalente" significa qualquer programa de comparação de sequência que, para quaisquer duas sequências em questão, gera um alinhamento que tem combinações idênticas de resíduos de nucleotídeos ou de aminoácidos e uma porcentagem de identidade de sequência idêntica quando comparado ao alinhamento correspondente gerado por GAP versão 10.

[0030] O GAP utiliza o algoritmo de Needleman e Wunsch (1970) supra, para encontrar o alinhamento de duas sequências completas que maximiza o número de combinações e minimiza o número de aberturas. O GAP considera todos os possíveis alinhamentos e posições da abertura e cria o alinhamento com o maior número de bases combinadas e menos aberturas. Ele permite a provisão de uma penalidade de criação da abertura e uma penalidade de extensão da abertura nas unidades de bases combinadas. O GAP deve lucrar com o número de combinações da penalidade de criação da abertura para cada abertura que insere.

[0031] Se uma penalidade de extensão da abertura maior do que zero for escolhida, o GAP deve, além disso, lucrar com cada abertura inserida do comprimento da abertura vezes a penalidade de extensão da abertura. Os valores da penalidade de criação da abertura padrão e os valores da

penalidade de extensão da abertura na Versão 10 dos GCG Wisconsin Genetics Software Packages para as seqüências de proteínas são de 8 e 2, respectivamente. Para as seqüências de nucleotídeos, a penalidade de criação da abertura padrão é de 50, ao passo que a penalidade de extensão da abertura padrão é de 3. As penalidades de criação da abertura e de extensão da abertura podem ser expressas como um número inteiro positivo selecionado do grupo de números inteiros que consistem em 0 a 200. Desse modo, por exemplo, as penalidades de criação da abertura e de extensão da abertura podem ser de 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65 ou mais. O GAP apresenta um membro da família dos melhores alinhamentos. Pode haver muitos membros desta família, mas nenhum outro membro tem uma qualidade melhor. O GAP exibe quatro valores de mérito para os alinhamentos: Qualidade, Relação, Identidade e Similaridade. A qualidade é a métrica maximizada a fim de alinhar as seqüências. A relação é a qualidade dividida pelo número de bases no segmento mais curto. A porcentagem de identidade é a porcentagem dos símbolos que combinam realmente. A porcentagem de similaridade é a porcentagem dos símbolos que são similares. Os símbolos que são transversais às aberturas são ignorados. Uma similaridade é classificada quando a classificação do valor da matriz para um par de símbolos é maior ou igual a 0,50, o limite da similaridade. A matriz de classificação utilizada na versão 10 dos GCG Wisconsin Genetics Software Packages é BLOSUM62 (consultar Henikoff e Henikoff (1989) Proc. Natl. Acad. Sci. USA 89:10915).

[0032] Tal como aqui utilizado, "identidade de seqüência" ou "identidade" no contexto de duas seqüências de

ácidos nucleicos ou polipeptídeos fazem referência aos resíduos nas duas seqüências que são os mesmos quando alinhados para a máxima correspondência em relação a uma janela especificada de comparação. Quando a porcentagem de identidade de seqüência é utilizada em referência às proteínas, é reconhecido que as posições dos resíduos que não são idênticas diferem frequentemente por substituições de aminoácidos conservadoras, onde os resíduos de aminoácido são substituídos em lugar de outros resíduos de aminoácido com propriedades químicas similares (por exemplo, carga ou capacidade hidrofóbica) e, portanto, não mudam as propriedades funcionais da molécula. Quando as seqüências diferem em substituições conservadoras, a porcentagem de identidade de seqüência pode ser ajustada para cima de modo a corrigir a natureza conservadora da substituição. As seqüências que diferem por tais substituições conservadoras são consideradas como tendo "similaridade de seqüência" ou "similaridade". Os meios para fazer este ajuste são bem conhecidos dos elementos versados na técnica. Tipicamente, isto envolve a classificação de uma substituição conservadora como uma combinação parcial e não total, aumentando desse modo a porcentagem de identidade de seqüência. Desse modo, por exemplo, onde um aminoácido idêntico recebe uma classificação igual a 1 e uma substituição não-conservadora recebe uma classificação igual a zero, uma substituição conservadora recebe uma classificação entre zero e 1. A classificação de substituições conservadoras é calculada, por exemplo, tal como implementado no programa PC/GENE (Intelligenetics, Mountain View, Calif.).

[0033] Tal como aqui utilizado, "porcentagem de

identidade de seqüência" significa o valor determinado através da comparação de duas seqüências idealmente alinhadas em uma janela de comparação, em que a parte da seqüência de polinucleotídeo na janela de comparação pode compreender adições ou deleções (isto é, aberturas) em comparação à seqüência de referência (que não compreende adições ou deleções) para o alinhamento ideal das duas seqüências. A porcentagem é calculada ao determinar o número das posições nas quais o resíduo idêntico da base de ácido nucléico ou aminoácido ocorre em ambas as seqüências para resultar no número de posições combinadas, ao dividir o número de posições combinadas pelo número total de posições na janela de comparação, e ao multiplicar o resultado por 100 para obter a porcentagem de identidade de seqüência.

[0034] A expressão "identidade substancial" de seqüências de polinucleotídeos significa que um polinucleotídeo compreende uma seqüência que tem uma identidade de seqüência de pelo menos 70%, tipicamente de pelo menos 80%, mais tipicamente de pelo menos 90%, e ainda mais tipicamente de pelo menos 95%, em comparação a uma seqüência de referência que utiliza um dos programas de alinhamento descritos ao utilizar parâmetros padrão.

[0035] No contexto das seqüências aqui fornecidas, uma seqüência é específica para essa seqüência de referência se, sob qualquer condição de reação que puder ser utilizada para distinguir uma seqüência de outra, tais como, sem limitação, PCR, transferência Southern ou transferência Northern, mas não para outras seqüências, tais como seqüências de outras espécies que incluem sem limitação aquelas de *S. cerevisiae*, *A. niger*, *A. terreus*, *P. pastoris*,

e *S. pombe*. Desse modo, em um ensaio de detecção de ácido nucléico, uma sonda/primer é "específico para" uma sequência se puder se ligar a um transcrito específico ou uma família desejada de transcritos extraídos de um espécime, para a inclusão prática (isto é, não interfere substancialmente no ensaio de detecção) de outras sequências. Em um ensaio de PCR, os primers são específicos para uma sequência de referência se amplificarem especificamente uma parte dessa sequência, para a exclusão prática de outras sequências em uma amostra.

[0036] Tal como aqui utilizado, um "primer" ou uma "sonda" para detectar uma espécie específica de ácido nucléico é qualquer primer, conjunto de primers e/ou sonda que podem ser utilizados para detectar e/ou quantificar a espécie específica de ácido nucléico. Uma "espécie de ácido nucléico" pode ser uma única espécie de ácido nucléico, que corresponde a um único gene, ou podem ser ácidos nucléico que são detectados por uma única combinação de primer e/ou sonda comum.

[0037] A expressão "célula hospedeira" significa qualquer a célula procariótica ou eucariótica onde uma sequência desejada de ácido nucléico foi introduzida na célula. Os processos metabólicos e as rotas de tal célula hospedeira podem manter, replicar e/ou expressar um vetor que contém um gene estranho ou molécula de DNA. Há uma variedade de células hospedeiras apropriadas, incluindo, mas sem ficar a elas limitadas, células bacterianas, fúngicas, de insetos, de mamíferos, e de plantas, que podem ser utilizadas de várias maneiras (por exemplo, como um veículo para manter um plasmídeo que compreende uma sequência desejada). As células

hospedeiras microbianas representativas incluem, mas sem ficar a elas limitadas, células fúngicas tais como de *Rhizopus sp.*, *Saccharomyces sp.*, *Streptomyces sp.*, *Pichia sp.*, *Aspergillus sp.*, e células bacterianas tais como de *Lactobacillus sp.*, *Escherichia sp.*, *Corynebacterium sp.*, *Brevibacterium sp.*, *Pseudomonas sp.*, *Proteus sp.*, *Enterobacter sp.*, *Citrobacter, sp.*, *Erwinia sp.*, *Xanthomonas sp.*, *Flavobacterium sp.*, *Streptococcus sp.*, *Lactococcus sp.*, *Leuconostoc sp.*, e *Enterococcus sp.* Em uma realização, a célula hospedeira é *Hansenula polymorpha* (*Pichia angusta*). Em uma outra realização, a célula hospedeira é de *Escherichia coli*. Em ainda uma outra realização, a célula hospedeira é de *Saccharomyces cerevisiae*.

[0038] O termo "polinucleotídeo" significa qualquer seqüência de nucleotídeo de um só cordão, conectada por ligações fosfodiéster, ou quaisquer seqüências de cordão duplo que compreendem duas tais seqüências de um só cordão complementares mantidas unidas por ligações de hidrogênio. A menos que esteja indicado de alguma outra maneira, cada seqüência de polinucleotídeo aqui indicada é apresentada como uma seqüência de deoxiribonucleotídeos (abreviados A, G, C e T). O termo "polinucleotídeo" engloba as moléculas de DNA ou polinucleotídeo, seqüências de deoxiribonucleotídeos, e moléculas de RNA ou poliribonucleotídeos, e as combinações destas.

[0039] O termo "promotor" significa uma seqüência de DNA dentro de uma seqüência de DNA maior que provê ou define um sítio ao qual a polimerase de RNA pode se ligar e iniciar a transcrição. Os promotores aqui descritos podem ser utilizados para superexpressar ou regular para

cima, por exemplo, e sem limitação, genes que codificam enzimas que aumentam o fluxo de carbono ao ácido láctico, fumarato, e outros metabólitos desejados durante mudanças nas condições de fermentação.

[0040] Um "equivalente" de uma determinada sequência de nucleotídeo de referência ou um elemento nela contido é uma sequência de nucleotídeo que contém, em comparação à sequência de nucleotídeo de referência, todos os elementos dessa sequência de nucleotídeo de referência, de maneira tal que a função característica desse ácido nucléico ou peptídeo de referência é mantida. Os elementos versados na técnica entendem que uma proteína funcional pode ser codificada pelas sequências de DNA equivalentes devido à degeneração no código genético. Por exemplo, um códon pode ser substituído em lugar de um outro, e contudo codifica o mesmo aminoácido, tal como, por exemplo e sem limitação, em referência ao códon de Ala, a substituição de GCG em lugar de GCA. No caso das proteínas, uma sequência pode conter os aminoácidos que representam substituições de aminoácido conservadoras, incluindo, mas sem ficar a eles limitadas, os grupos de substituição conservadora: Ser e Thr; Leu, Ile e Val; Glu e Asp; e Gln e Asn. Uma sequência tal como aqui reivindicado inclui desse modo a sequência de referência, bem como os seus equivalentes devido à degeneração no código genético. As substituições conservadoras também podem ser determinadas por outros métodos, tais como, sem limitação, aqueles utilizados pelo algoritmo BLAST (Basic Local Alignment Search Tool = Ferramenta de Busca de Alinhamento Local Básico), a Matriz de Classificação de Substituição BLOSUM, e a matriz BLOSUM 62 (também consultar, por exemplo,

Altschul et al., *Methods in Enzymology* 266:460-479 (1996)). Preponderantemente, os "equivalentes" e os "equivalentes conservados" de um ácido nucléico ou peptídeo/proteína de referência retêm substancialmente retêm ou realçam a função do ácido nucléico ou peptídeo/proteína de referência.

[0041] O termo "vetor" significa um meio para introduzir uma seqüência de nucleotídeo estranho em uma célula, incluindo, sem limitação, um plasmídio ou um vírus. Tais vetores podem operar sob o controle de um engenho de expressão de gene da célula hospedeira. Um vetor contém as seqüências que facilitam a replicação e/ou a manutenção de um segmento de ácido nucléico estranho na célula hospedeira. Geralmente, o vetor é introduzido em uma célula hospedeira para a replicação e/ou a expressão do segmento de DNA estranho ou para a entrega do DNA estranho no genoma hospedeiro. Um vetor de plasmídio típico contém: (i) uma origem de replicação, de modo que o vetor possa ser mantido e/ou replicado em uma célula hospedeira; (ii) um marcador selecionável, tal como um gene resistência a antibiótico para selecionar as células que contêm o vetor (transformantes) entre as células sem vetor, e (iii) um sítio de poliligante contendo vários sítios de reconhecimento e corte de endonuclease de restrição diferentes para facilitar a clonagem de uma seqüência de DNA estranho.

[0042] São aqui apresentados genes e elementos genéticos úteis na modificação de células hospedeiras. Estas células hospedeiras podem incluir, por exemplo, microorganismos. Um microorganismo particularmente apropriado para o uso nas realizações da invenção é a levedura de *H. polymorpha* (*P. angusta*). Os métodos e as composições da

invenção são úteis para a provisão de microorganismos com uma maior atividade de enzimas.

[0043] Os métodos e as composições aqui apresentados podem ser particularmente úteis para a produção fermentativa do etanol utilizando SSP (sacarificação e fermentação simultâneas). As cepas da invenção também podem servir particularmente para a produção fermentativa do etanol a altas temperaturas. "Altas temperaturas" significa, por exemplo, a fermentação executada a temperaturas entre 40 e 60°C, entre 40 e 55°C, entre 40 e 50°C, entre 45 e 55°C, entre 45 e 50°C, entre 48 e 52°C, e entre 48 e 50°C. As temperaturas típicas para a fermentação incluem 48°C e 50°C.

I. Cepas e métodos para a produção de etanol a partir de D-xilose utilizando *H. polymorpha* que superproduz a proteína 104 de choque térmico de *H. polymorpha*

[0044] Em uma realização da invenção é provida uma levedura produtora de etanol da espécie *H. polymorpha*, em que a levedura foi modificada para ter uma produção aumentada de etanol em um meio que tem a D-xilose como uma fonte principal de carbono ao aumentar a quantidade da expressão da proteína 104 de choque térmico de *H. polymorpha* na levedura. Nos exemplos aqui relatados, a produção do etanol é relativa àquela de uma cepa de controle que seja um transformante de *H. polymorpha* NCYC495 *leu* (*ScLEU2*). O transformante é um derivado da cepa NCYC495 *leu* 1-1 que contém o gene LEU2 de *S. cerevisiae*. Em outras realizações, a produção aumentada está relacionada àquela de uma cepa paterna de *H. polymorpha*.

[0045] A produção do etanol também pode ser aumentada ao aumentar a quantidade da expressão de uma proteína que tem uma sequência de aminoácido com uma ou mais

deleções, substituições, inserções, inversões, ou adições à seqüência de aminoácido da proteína 104 de choque térmico de *H. polymorpha*, em que a dita proteína é pelo menos 90%, pelo menos 95%, pelo menos 98%, ou pelo menos 99% idêntica à proteína 104 de choque térmico de *H. polymorpha* do tipo selvagem.

[0046] Em uma realização, a expressão da proteína 104 de choque térmico é aumentada ao aumentar o número de cópia do gene da proteína 104 de choque térmico de *H. polymorpha* (HSP104) (em uma cepa de *H. polymorpha*). Em uma realização adicional, o HSP104 é colocado sob o controle de um promotor não-nativo. O promotor não-nativo pode ser, por exemplo, mas sem ficar a ele limitado, um promotor constitutivo. Em uma realização, o promotor não-nativo é o promotor do gene GAPDH (gliceraldeído-3-fosfato deidrogenase) de *H. polymorpha*.

[0047] O número de cópia do gene HSP104 pode ser aumentado, por exemplo, ao transformar uma cepa paterna de *H. polymorpha* com um plasmídio que contém o ORF do gene HSP104, ou ao transformar uma cepa paterna de *H. polymorpha* com um plasmídio que contém o ORF do gene HSP104 sob o controle do promotor de GAPDH. O promotor é selecionado, por exemplo, entre promotores de GAPDH, PMA1 (H⁺-ATPase de membrana de plasma), TEF1 (fator de alongamento de translação 1A), e genes de PGK1 (3-fosfoglicerato quinase). Os elementos versados na técnica irão reconhecer uma série de maneiras para transformar *H. polymorpha* com o gene ou construto de gene/promotor desejado.

[0048] As leveduras com expressão aumentada da proteína 104 de choque térmico podem ser cultivadas em um

meio contendo D-xilose para produzir o etanol. Um método para obter o etanol de acordo com uma realização da invenção compreende as etapas de cultivo da levedura com expressão aumentada da proteína 104 de choque térmico em um meio de cultura contendo D-xilose para fermentar a D-xilose, acumulação do etanol no meio, e coleta do etanol do meio. A fermentação pode ser realizada na presença de 12% de xilose a uma temperatura de 48 - 50°C durante um período de tempo de um a três dias em condições semiaeróbicas (140 rpm).

[0049] A fermentação foi realizada no meio mínimo (YNB) com xilose; as concentrações da xilose eram de 4 a 12%; a melhor produção do etanol ocorreu a 12% de xilose. O pH era de aproximadamente 3 a 5 e foi verificado periodicamente. "Crescimento melhorado" significa a melhor acumulação de biomassa do meio de crescimento" à temperatura aumentada (50°C). O OD foi medido a 600 nm. Em todas as experiências apresentadas, o transformante de *H. polymorpha* NCYC495 *leu 1-1* (*ScLEU2*) foi utilizado como uma cepa de controle. Esse transformante foi utilizado como uma cepa de controle porque todas as cepas recombinantes testadas (HSP104-8, Δ ath1-36, Δ ath1-36_XYL3) foram obtidas após as transformações pelos vetores que contêm *ScLEU2*. Desse modo, as cepas testadas continham o gene (*ScLEU2*). A cepa Δ ath1-36_XYL3 foi depositada em 13 de setembro de 2007, na NRRL ARS Culture Collection, Peoria, Illinois, USA, sob os termos do tratado de Budapeste, como depósito número NRRL Y-50061.

II. Cepas e métodos para a produção do etanol a partir de D-xilose utilizando *H. polymorpha* sem um gene *ATH1* ativo (trehalase de ácido)

[0050] Em uma realização da invenção, é

apresentada uma levedura produtora de etanol da espécie *H. polymorpha*, em que a levedura foi modificada para ter a produção de etanol aumentada em D-xilose pela inativação do gene ATH1 (trehalase de ácido). O gene trehalase de ácido na levedura de *Saccharomyces cerevisiae* é discutido por Jung, Y-J. & Park, H-D. "Antisense-mediated Inhibition of Acid Trehalase (ATH1) Gene Expression Promotes Ethanol Fermentation and Tolerance in *Saccharomyces cerevisiae*" Biotech. Lett. 27:1855-1859 (2005); e Kim, J., et al. "Disruption of the Yeast ATH1 Gene Confers Better Survival after Dehydration, Freezing, and Ethanol SHock: Potential Commercial Applications" Appl. & Envir. Microbiol. 62(5):1563- 1569 (1996).

[0051] As expressões "gene ATH1 inativado", "gene ATH1 inativo", "inativação do gene ATH1", e outras ainda significam que o gene ATH1 é modificado de maneira tal que o gene codifica uma proteína mutante. A proteína mutante pode ter uma atividade diminuída ou pode ser inteiramente inativa. O gene ATH1 também pode ter sido modificado de maneira tal que é incapaz de prover a expressão natural da proteína trehalase de ácido devido à deleção de todo ou uma parte do gene ATH1, a modificação das regiões adjacentes ao gene ATH1, ou o rompimento de uma ou mais porções de um gene ATH1 por uma adição de um ou mais nucleotídeos ao gene ATH1.

[0052] Os elementos versados na técnica irão reconhecer, com o benefício desta descrição, uma série de maneiras para desativar o gene ATH1 em *H. polymorpha*. Estes incluem, por exemplo, mas sem ficar a ela limitados, a deleção do rompimento, parcial ou completo do gene. As leveduras em que o gene ATH1 foi inativado também podem ser

modificadas para intensificar a produção do etanol em condições desejadas. Por exemplo, a expressão de um ou mais genes envolvidos na produção do etanol pode ser intensificada. Um ou mais genes na levedura podem ser desativados para otimizar a alocação de recursos de carbono para a fermentação.

[0053] As leveduras nas quais o gene *ATH1* é inativo podem ser de *H. polymorpha* do tipo selvagem, ou podem ser as cepas de *H. polymorpha* nas quais a produção do etanol em D-xilose já foi aumentada além daquela de *H. polymorpha* do tipo selvagem. As leveduras também podem ser modificadas para aumentar a produção do etanol em D-xilose de acordo com outras realizações da invenção.

[0054] As leveduras com um gene *ATH1* inativo podem ser cultivadas em um meio contendo D-xilose para produzir o etanol. Um método de obtenção do etanol de acordo com uma realização da invenção compreende as etapas de cultivo da levedura com um gene *ATH1* inativo em um meio de cultura contendo D-xilose para fermentar a D-xilose, acumulação do etanol no meio, e coleta do etanol do meio. A fermentação pode ser realizada na presença de 12% de xilose a uma temperatura de 48 - 50°C durante um período de tempo de um a três dias em condições semiaeróbicas (140 rpm).

[0055] A fermentação foi realizada no meio mínimo (YNB) com xilose; as concentrações de xilose eram de 4 a 12%; a melhor produção do etanol foi com 12% de xilose. O pH era de aproximadamente 3 a 5 e foi verificado periodicamente. "Crescimento melhorado" significa melhor acumulação de biomassa do meio de crescimento à temperatura aumentada (50°C). O OD foi medido a 600 nm. Em todas as

experiências apresentadas, o transformante de *H. polymorpha* NCYC495 leu 1-1 (*ScLEU2*) foi utilizado como uma cepa de controle.

III. Cepas e métodos para a produção do etanol a partir de D-xilose utilizando *H. polymorpha* que superproduz xiluloquinase

[0056] Em uma realização da invenção, é apresentada uma levedura produtora de etanol da espécie *H. polymorpha*, em que a levedura foi modificada para ter uma produção aumentada do etanol em um meio que tem a D-xilose como uma fonte principal de carbono ao aumentar a quantidade da expressão da proteína xiluloquinase de *H. polymorpha* na levedura. Em uma realização, a produção aumentada do etanol é relativa àquela de uma cepa de controle que é uma *H. polymorpha* do tipo selvagem, e em uma outra realização a produção aumentada é relativa àquela de uma cepa paterna de *H. polymorpha*.

[0057] A produção do etanol pode também ser aumentada ao aumentar a quantidade da expressão de uma proteína que tem uma sequência de aminoácido com uma ou mais deleções, substituições, inserções, inversões, ou adições à sequência de aminoácido de xiluloquinase de *H. polymorpha*, em que a dita proteína é pelo menos 90%, pelo menos 95%, pelo menos 98%, ou pelo menos 99% idêntica à sequência de aminoácido de xiluloquinase de *H. polymorpha* do tipo selvagem.

[0058] Em uma realização, a expressão de xiluloquinase é aumentada ao aumentar o número de cópia do gene de xiluloquinase de *H. polymorpha* (XYL3) em uma cepa de *H. polymorpha*. Em uma realização adicional, o XYL3 é colocado

sob o controle de um promotor não-nativo.

[0059] O promotor não-nativo pode ser, por exemplo, mas sem ficar a ele limitado, um promotor constitutivo. Em uma realização, o promotor não-nativo é o promotor do gene GAPDH (gliceraldeído-3-fosfato deidrogenase) de *H. polymorpha*.

[0060] O promotor é selecionado, por exemplo, entre promotores de GAPDH, PMA1 (H⁺-ATPase de membrana de plasma), TEF1 (fator de alongamento de translação 1A), genes PGK1 (3- fosfoglicerato quinase).

[0061] O número de cópia do gene XYL3 pode ser aumentado, por exemplo, ao transformar uma cepa paterna de *H. polymorpha* com um plasmídio que contém o ORF do gene XYL3 sob o controle do promotor de GAPDH (GenBank # A Y550078). O número de cópia do gene HSP104 pode ser aumentado, por exemplo, ao transformar uma cepa paterna de *H. polymorpha* com um plasmídio que contém o ORF do gene HSP104 sob o controle do promotor de GAPDH (GenBank # A Y550078). Os elementos versados na técnica irão reconhecer uma série de maneiras para transformar *H. polymorpha* com o gene ou construto de gene/promotor desejado.

[0062] As leveduras com expressão aumentada de xiluloquinase podem ser cultivadas em um meio contendo D-xilose para produzir o etanol. Um método de obtenção do etanol de acordo com uma realização da invenção compreende as etapas de cultivo da levedura com expressão aumentada de xiluloquinase em um meio de cultura contendo D-xilose para fermentar a D-xilose, acumulação do etanol no meio, e coleta do etanol do meio. A fermentação pode ser realizada, por exemplo, na presença de 12% de xilose a uma temperatura de 48

- 50°C durante um período de tempo de um a três dias em condições semiaeróbicas (140 rpm).

IV. Realizações Adicionais

[0063] Em uma realização adicional da invenção, um ou mais dos aspectos descritos na seção I, na seção II, e na seção III podem ser utilizados em combinação para prover uma cepa de *H. polymorpha* com a produção aumentada do etanol em D-xilose. Por exemplo, uma realização pode prover uma cepa de *H. polymorpha* em que o gene ATH1 é inativo mas o gene XYL3 é superexpresso e/ou tem um número de cópia aumentado. Uma outra realização pode prover uma cepa de *H. polymorpha* com um gene ATH1 inativo, superexpressão e/ou número de cópia aumentado do gene XYL3, e superexpressão e/ou número de cópia aumentado do gene HSP104. Ainda uma outra realização provê uma cepa de *H. polymorpha* com superexpressão e/ou número de cópia aumentado do gene XYL3 e superexpressão e/ou número de cópia aumentado do gene HSP104. Ainda uma outra realização da invenção provê uma cepa de *H. polymorpha* com um gene ATH1 inativo e número de cópia aumentado e/ou superexpressão do gene HSP104.

[0064] Os elementos versados na técnica, com o benefício da presente descrição, irão reconhecer que modificações adicionais podem ser feitas nas células hospedeiras da invenção, o que permite requisitos nutricionais adicionais, a produção alterada de etanol e/ou de outros produtos químicos.

[0065] Por exemplo, pode ser produzida uma cepa de *H. polymorpha* apropriada para a produção de etanol em um meio em que a D-xilose é a fonte de energia principal por qualquer combinação da expressão aumentada da proteína de

choque térmico 104, da desativação de ATH1 e da expressão aumentada de xiluloquinase. As modificações que já são conhecidas dos elementos versados na técnica, assim como as modificações desenvolvidas mais tarde, podem ser utilizadas além daquelas aqui apresentadas.

Cepas e plasmídios

[0066] As cepas microbianas e os plasmídios utilizados nas realizações da invenção são apresentados na Tabela 1.

TABELA 1

Cepas	Descrição	Fonte
<i>H. polymorpha</i> NCYC495 <i>leul-1</i> (ATCC MYA-335)	<i>Leul-1</i> , deficiente em β -isopropila malato desidrogenase	ATCC, USA
HSP104-8	Derivado de NCYC495 <i>leul-1</i> com HSP104 sob controle GAPDH	Esta descrição
Δ ath1-36	Derivado de NCYC495 <i>leul-1</i> com deleção do gene ATH1 (trehalase de ácido)	Esta descrição
Δ ath1-36_XYL3	Derivado de Δ ath1-36 com deleção do gene ATH1 (trehalase de ácido) e com um gene HpXIL3	Esta descrição

	adicional sob controle GAPDH	
<u>Plasmídios</u>		
pK08+prGAP+HSP104Hp	ORF do gene <i>H. polymorpha</i> HSP104-8 (proteína de choque térmico 104) sob o controle do gene promotor GAPDH	Esta descrição
p Δ ATH1HP	Plasmídio para deleção do gene ATH1	Esta descrição
pXYL3	Plasmídio para adição do gene <i>H. polymorpha</i> xiluloquinase XYL3 dirigido pelo promotor GADPH	Esta descrição

Meio e Condições de cultura

[0067] O meio e as condições de cultura utilizados nas experiências para as realizações da invenção são tal como apresentado abaixo, a menos que sejam indicados de outra maneira em exemplos específicos. As leveduras crescem em um meio à base de nitrogênio sintético (YNB) suplementado com xilose como única fonte de carbono e energia (2%) a 37°C. Os cultivos em meio líquido foram realizados em 40 ml do meio com 12% de xilose em frascos de agitação Erlenmeyer de 125-ml em um agitador a 37 ou 48°C. As condições de oxigênio limitado foram providas por agitação a

135 - 140 rpm. A densidade da célula inicial depois da inoculação é ~ 2 mg de peso seco x ml⁻¹. Os meios são inoculados com culturas pré-desenvolvidas em 80 ml de meio YPX (1% de extrato de levedura, 2% de peptona, 8% de xilose) no cultivo em frascos de 300 ml com agitação a 220 rpm até a fase de crescimento médio-exponencial. As células para inoculação são recolhidas por meio de centrifugação, lavadas com água e concentradas para obter a densidade inicial mencionada acima.

[0068] Os elementos versados na técnica irão reconhecer que outros meios podem ser utilizados, dependendo das condições de crescimento desejadas e da composição do material lignocelulósico a ser utilizado como um material bruto para a fermentação.

Enzimas, primers e produtos químicos

[0069] Um fragmento que contém o ORF de *HpHSP104* foi isolado por PCR de DNA genômico da cepa CBS 4732 *leu2-2* utilizando os primers A29: 5'-CCCCATATGGATCAATCACAATTTACCGACAGAGC-3' e A30: 5'-GAACGGCCGCTCAGTCCAAATCTGGAG-3'. Os sítios de restrição NdeI e NotI foram incorporados nos primers A29 e A30, respectivamente, para prover uma orientação correta do fragmento de PCR isolado (o ORF de *HpHSP104*) no sítio correspondente do plasmídeo pK08-GAPpr (descrição e esquema linear do plasmídeo: consultar o artigo da autoria de Voronovsky A. Y. et al., "Expression of xylA Genes Encoding Xylose Isomerases from *Escherichia Coli* and *Streptomyces coelicolor* in the Methylophilic Yeast *Hansenula polymorpha*" FEMS Yeast Res 5(11): 1055-62 (2005)). Enzimas de restrição, enzimas de modificação de DNA e outros reagentes foram

obtidos junto à New England Biolabs, USA, Sigma, USA e Fermentas, Lituânia.

[0070] As condições de reação empregadas foram tal como recomendado pelos fornecedores. O DNA genômico da *H. polymorpha* foi isolado utilizando o kit Wizard® Genomic DNA Purification (Promega, USA). As endonucleases de restrição e a ligase de DNA (Fermentas, Lituânia, e New England, Biolabs, USA) foram utilizadas de acordo com as especificações do fabricante. O isolamento do plasmídeo de *E. coli* foi realizado com o Wizard® Plus SV Minipreps DNA Purification System (Promega, EUA). Os fragmentos de DNA foram separados em 0,8% de gel de agarose (Fisher Scientific, USA) em 1 x TAE (Sambrook et al. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1989). O isolamento dos fragmentos do gel foi realizado com o kit DNA Gel Extraction (Millipore, USA). A amplificação dos ORFs de *HSP104* e do promotor de HpGAP foi feita com Platinum® Taq DNA Polymerase High Fidelity (Invitrogen, USA), de acordo com a especificação do fabricante. Os PCRs foram realizados no termociclador GeneAmp® PCR System 9700 de (Applied Biosystems, USA). Com o benefício desta descrição, os elementos versados na técnica irão reconhecer que as transformações e os isolamentos podem ser realizados por qualquer um dentre uma variedade de materiais e de métodos conhecidos.

Transformação

[0071] Os elementos versados na técnica irão reconhecer que existem vários métodos para a transformação da *H. polymorpha*. Por exemplo, pode ser utilizado o método de eletroporação relatado em (Faber, K.N., et al., " Highly-

efficient Electrotransformation of the Yeast *Hansenula Polymorpha*" Curr. Genet. 25: 305-310 (1994)). A transformação que utiliza células intactas também pode ser eficaz (Roggenkamp R. et al., "Transformation of the methylotrophic yeast *Hansenula Polymorpha* by autonomous replication and integration vectors" Mol. Gen. Genet. 202: 302 - 308 (1986)).

Construção do Plasmídio

[0072] Os plasmídios recombinantes contendo *HSP104* ORF de *H. polymorpha* dirigidos pelo promotor de gliceraldeído-3-fosfato desidrogenase (GAPDH) de *H. polymorpha* (GenBank#A Y550078), e também incluindo o gene de *Saccharomyces Cerevisiae* *LEU2*, isolado do plasmídio *Yep 13*, foram construídos na base do plasmídio *pK08-GAPpr* (Voronovsky A. Y. et al., "Expression of *xylA* Genes Encoding Xylose Isomerases from *Escherichia Coli* and *Streptomyces coelicolor* in the Methylotrophic Yeast *Hansenula polymorpha*" FEMS Yeast Res 5(11): 1055-62 (2005)). A construção do plasmídio *pK08-GAPpr* é aqui relatada.

[0073] O plasmídio *pK08+prGAP+HSP104Hp* (Figura 1) foi linearizado por *BamH I* e utilizado para a transformação das células de *H. polymorpha* da cepa NCYC 495 *leu1-1*. Isto permitiu isolar os integrantes *Leu⁺* que contêm o recombinante *HpHSP104* ORF dirigido com o promotor de *HpGAPDH*.

[0074] O plasmídio recombinante *pΔATH1Hp* foi utilizado para a deleção do gene de *H. polymorpha* *ATH1* na cepa NCYC 495 *leu1-1*, tendo por resultado a criação da cepa *Δath1-36*.

Ensaio de etanol

[0075] kit "Alcotest" (Gonchar, M.V., Maidan, M.M., Sibimy, AA "A new oxidase-peroxidase kit "Alcotest" for

ethanol assays in alcoholic beverages" Food Technol Biotechnol. 39: 37 - 42 (2001)) foi utilizado para os ensaios de etanol.

EXEMPLOS

[0076] Os exemplos a seguir se prestam a orientar os elementos versados na técnica na prática da presente invenção. Eles não devem ser interpretados como limitadores do âmbito da invenção, o qual é definido pelas reivindicações.

Exemplo 1

[0077] O Exemplo 1 descreve a produção de transformantes de *H. polymorpha* contendo o recombinante integrado *HpHSP104*. A sequência do gene *H. polymorpha HSP104* (*HpHSP104*), que codifica a proteína de choque térmico 104, foi obtida junto ao banco de dados do genoma de *H. polymorpha* (Rhein Biotech gmbH). Um fragmento de 2,685 kb que contém o ORF do *HpHSP104* foi isolado por PCR do DNA genômico da cepa CBS 4732 *leu2-2* utilizando os primers A29 5'-CCCATATGGATCAATCACAATTTACCGACAGAGC-3' e A30 5'-GAACGGCCGCTCAGTCCAAATCTGGAG-3'.

[0078] O produto de PCR resultante (o ORF de *HpHSP104*) foi tratado com as endonucleases de restrição NdeI e NotI que flanqueiam o produto. O produto PCR de 2,685 kb *NdeI/NotI* foi ligado com o plasmídio *NdeI/NotI*-linearizado pK08-GAPpr. Isso resultou no construto pK08+prGAP+HSP104Hp. O construto contém o ORF de *HpHSP104* dirigido com o promotor *H. polymorpha* GAPDH (GenBank#AY550078). Além disso, pK08+prGAP+HSP104Hp contém o gene *Saccharomyces Cerevisiae* LEU2 (*ScLEU2*).

[0079] O plasmídio pK08+prGAP+HSP1 04Hp foi

utilizado para a transformação da cepa *H. polymorpha* NCYC 495 *leu1-1* por meio de eletroporação.

[0080] Os integrantes que contêm tanto o *ScLEU2* quanto o gene recombinante *HpHSP104* foram selecionados entre os transformantes resultantes de Leu+. Isso foi feito por PCR utilizando DNA genômico dos transformantes como um modelo e os primers correspondentes para o promotor *GAPDH* e extremidade 3' de *HpHSP104* ORF (primer anterior A58 5'-CGCGAGCTCCCAATTATCATTAATAATCAC-3' e primer posterior A30 5'-GAACGGCCGCTCAGTCCAAATCTGGAG-3'.) e *ScLEU2* (primer anterior IS25: CGGCTGCAGGAGAACTTCTAGTATATCTACATAC e primer posterior IS26: TATCTGCAGCTACGTCGTAAAGGCCGTTTCTG). Os recombinantes foram isolados em consequência do trabalho.

Exemplo 2

[0081] O Exemplo 2 relata os testes de termotolerância e de produção de etanol nos transformantes de *H. polymorpha* produzidos no Exemplo 1. Os transformantes cresceram em um meio YNB com 12% de xilose como fonte de carbono com aeração restrita (140 rpm) a 50°C. A Tabela 2 inclui uma comparação entre a termotolerância dos transformantes e aquela de NCYC 495 *leu1-1* a 50°C. A Tabela 3 inclui uma comparação entre a produção de etanol dos transformantes e aquela de NCYC 495 *leu1-1* a 50°C.

Tabela 3 - Produção de etanol (mg/ml) com os transformantes de *H. polymorpha* que contêm o gene recombinante integrado de *HpHSP104*; meio YNB, 12% de xilose.

Dia	HSP104-8	Leu (controle)
1	0,87	0,26
2	0,9	0,185
3	0,87	0,111

4	0,44	0,003
5	0,38	0,003

Exemplo 3

[0082] O Exemplo 3 descreve a produção de transformantes de *H. polymorpha* em que o gene *ATH1* (trehalase de ácido) foi parcialmente suprimido. 2,507 kb de *ATH1* ORF foram suprimidos pela transformação da cepa NCYC 495 *leu1-1* com a supressão do cassete *ATH1*; o cassete foi amplificado do plasmídeo Δ ATH1Hp utilizando os primers: A4l CCCAAGCTTATACCTTCACTAACATAACCAGTGGAC e A44CGGGTCGACTCTCTGCGACATAATAAGCTG.

Exemplo 4

[0083] O Exemplo 4 relata os testes de termotolerância e de produção de etanol nos transformantes de *H. polymorpha* produzidos no Exemplo 3. A Tabela 4 inclui uma comparação entre a termotolerância dos transformantes e aquela de NCYC 495 *leu1-1*. A Tabela 5 inclui uma comparação entre a produção de etanol dos transformantes e aquela de NCYC 495 *leu1-1* a 50°C.

Tabela 5 - Produção de etanol (mg/ml) com os transformantes de *H. polymorpha* contendo um gene *HpATH1* inativo; meio YNB, 12% de xilose; 50°C:

Dia	Δ ath1-36	Leu (controle)
1	0,9	0,26
2	0,9	0,185
3	0,79	0,111
4	0,44	0,003
5	0,44	0,003

Exemplo 5

[0084] O Exemplo 5 descreve a produção de

transformantes de *H. polymorpha* contendo um gene recombinante integrado *HpXYL3* e possuindo um gene *HpATH1* inativo.

[0085] Δ ath1-36 foi utilizado como cepa recipiente para a transformação com o plasmídio pGLG61+prGAP+XYL3Hp. Os integrantes que contém o gene recombinante *HpXYL3* foram selecionados entre os transformantes resistentes de genetecina resultantes. Isso feito por PCR utilizando o DNA genômico dos transformantes como um modelo e os primers correspondentes para o promotor GAPDH e o recombinante *HpXYL3* (anterior A58 5'-CGCGAGCTCCCAATTATCATTAATAATCAC-3' e posterior K9 5'-TTTGCGGCCGCTTAAGACTCTAATTTTGTG-3'. 2,507 kb de *ATH1* ORF foram suprimidos pela transformação da cepa NCYC 495 *leul-1* com a supressão do cassete *ATH1*; o cassete foi amplificado do plasmídio Δ ATH1Hp utilizando os primers: A41 CCCAAGCTTATACCTTCACTAACATACCAGTGGAC e A44 CGGGTCGACTCTCTGCGACATAATAAGCTG.

Exemplo 6

[0086] Considerando o fato que as realizações particulares da presente invenção foram descritas para fins da ilustração, ficará evidente aos elementos versados no estado da técnica que numerosas variações dos detalhes da presente descrição podem ser feitas sem que se desvie da invenção, tal como definido nas reivindicações anexas. As patentes e publicações aqui discutidas devem ser vistas como indicadoras do nível de habilidade no estado da técnica, embora nenhum reconhecimento seja feito de que qualquer documento é uma referência à técnica anterior.

[0087] A presente descrição apresenta diversas características e aspectos diferentes da invenção com referência às várias realizações exemplificadoras. Deve ser

compreendido, no entanto, que a invenção engloba numerosas realizações alternativas, que podem ser feitas pela combinação de qualquer um dentre as diferentes características e aspectos aqui descritos em qualquer combinação que um elemento versado no estado da técnica deve julgar útil. Todas as patentes e publicações anteriores são aqui incorporadas a título de referência. Até o ponto em que o material incorporado entra em conflito com as definições e indicações existentes, ou com outro material da descrição apresentado nesta descrição, a descrição tal como aqui determinada substitui explicitamente qualquer material conflitante incorporado a título de referência.

REIVINDICAÇÕES

1. MÉTODO PARA A PRODUÇÃO DE ETANOL, caracterizado por compreender as etapas de:

a) cultivo da levedura isolada da espécie *Hansenula polymorpha* em um meio a fim de produzir e acumular etanol no meio, sendo que a levedura compreende pelo menos uma modificação selecionada do grupo que consiste em

(a) um gene ATH1 inativo (trealase ácida) formado por uma interrupção do gene ATH1 provocada pelo homem; ou

(b) superexpressão do gene da proteína de choque térmico 104 de *H. polymorpha* quando comparada com uma cepa paterna, em que a superexpressão é realizada através da inserção de um promotor não-nativo para conduzir a expressão do referido gene de choque térmico 104 ou inserindo várias cópias do gene de choque térmico 104 na levedura; e

b) coleta do etanol do meio.

2. MÉTODO, de acordo com a reivindicação 1, caracterizado pela dita produção ocorrer a uma temperatura maior do que 48°C.

3. MÉTODO, de acordo com a reivindicação 1, caracterizado pelo dito gene ATH1 ser inativo devido à mutação, interrupção, deleção parcial e / ou deleção do gene ATH1 da referida levedura.

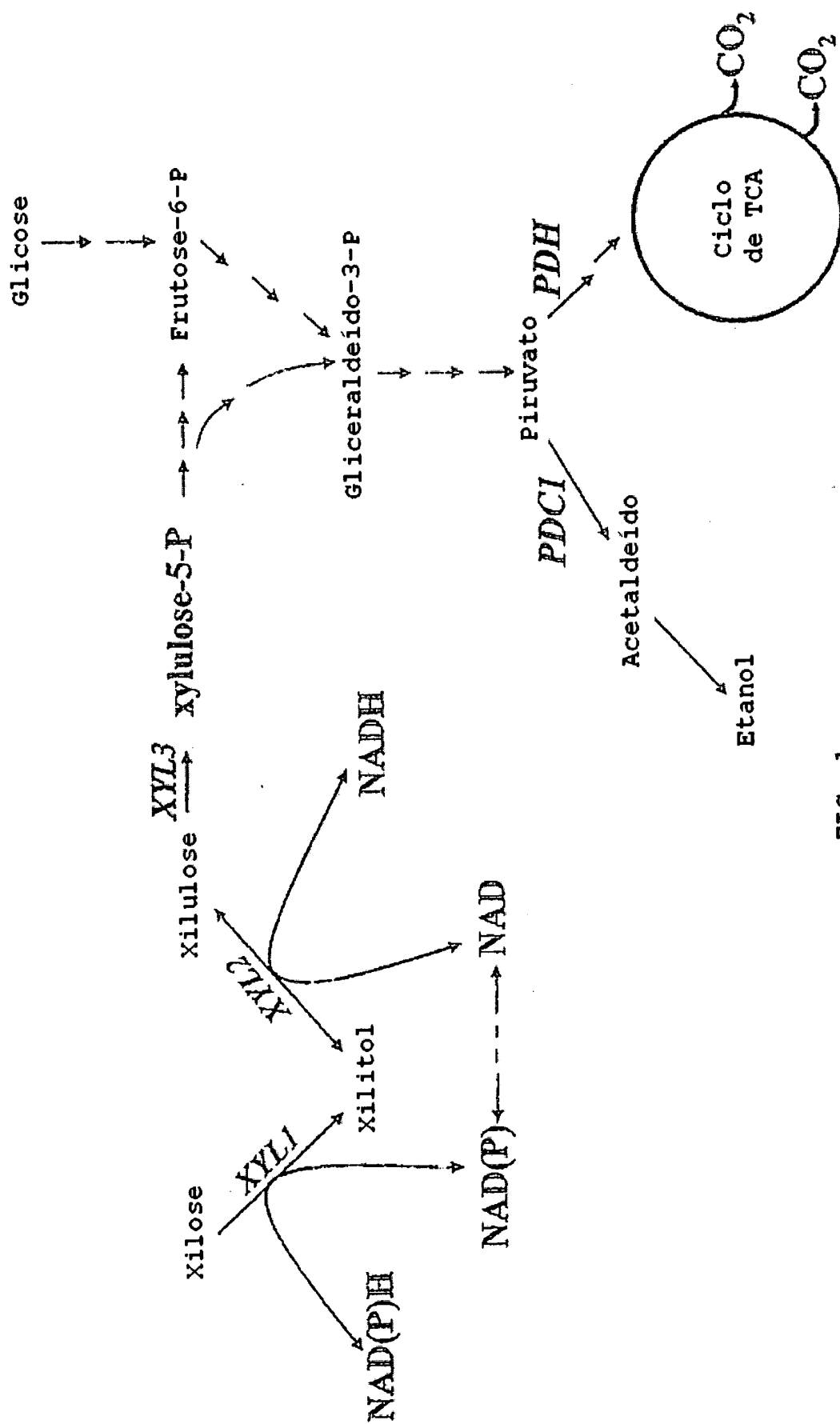


FIG. 1