



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

**(11) PI 0913552-9 B1**



**(22) Data do Depósito: 29/09/2009**

**(45) Data de Concessão: 29/09/2020**

**(54) Título:** MÉTODO PARA PRODUÇÃO DE ISOBUTANOL E MÉTODO PARA CONVERSÃO DE 2,3-DIHIDROXI ISOVALERATO

**(51) Int.Cl.:** C12P 7/16; C12N 9/88; C12N 15/09; C12R 1/645.

**(30) Prioridade Unionista:** 29/09/2008 US 61/100,806; 29/09/2008 US 61/100,801.

**(73) Titular(es):** BUTAMAX ADVANCED BIOFUELS LLC.

**(72) Inventor(es):** LARRY CAMERON ANTHONY; LORI ANN MAGGIO-HALL; STEVEN CARY ROTHMAN; JEAN-FRANCOIS TOMB.

**(86) Pedido PCT:** PCT US2009058826 de 29/09/2009

**(87) Publicação PCT:** WO 2010/037111 de 01/04/2010

**(85) Data do Início da Fase Nacional:** 14/03/2011

**(57) Resumo:** "CÉLULA DE LEVEDURA HOSPEDEIRA RECOMBINANTE, MÉTODO PARA PRODUÇÃO DE ISOBUTANOL, MÉTODO PARA CONVERSÃO DE 2,3-DIHIDROXIISOVALERATO, MÉTODO PARA PRODUÇÃO DE 2- BUTANONA E MÉTODO PARA CONVERSÃO DE 2,3-BUTANODIOL" A presente invenção refere-se a linhagens de leveduras elaboradas geneticamente para ter aumento de atividade de proteínas heterólogas que necessitam da ligação de um cluster Fe-S para sua atividade. As linhagens de leveduras têm atividade reduzida de uma proteína Fe-S endógena. As atividades de ácido dihidroxi desidratases 2Fe-2S fúngicas ou vegetais e Fe-S propanodiol desidratase reativase foram aumentadas para produção de produtos fabricados com o uso das vias biossintéticas que incluem essas enzimas, como valina, isoleucina, leucina, ácido pantotênico (vitamina 85), isobutanol, 2-butanona e 2-butanol.

**“MÉTODO PARA PRODUÇÃO DE ISOBUTANOL E MÉTODO PARA  
CONVERSÃO DE 2,3-DIHIDROXI ISOVALERATO”**

[001] Este pedido refere-se e reivindica o benefício de prioridade para o pedido provisório de patente US 61/100.801, depositado em 29 de setembro de 2008 e 61/100.806, depositado em 29 de setembro de 2008. O conteúdo destes é incorporado ao presente pedido como referência.

**CAMPO DA INVENÇÃO**

[002] A invenção refere-se ao campo da microbiologia industrial e à expressão de proteínas que necessitam de um cluster ferro-enxofre para sua atividade. Mais especificamente, a expressão da proteína Fe-S heteróloga em células de levedura é aprimorada através da inativação específica de gene de hospedeiro.

**ANTECEDENTES DA INVENÇÃO**

[003] A modificação genética para produção fermentativa de produtos comerciais é um campo ativo e em crescimento. As vias enzimáticas modificadas geneticamente para biossíntese de alguns produtos incluem enzimas que necessitam da ligação de um cluster ferro-enxofre (Fe-S) para suas atividades. Ácido dihidroxi desidratase (DHAD) é um exemplo. DHAD é parte de vias biossintéticas que ocorrem naturalmente e que produzem valina, isoleucina, leucina e ácido pantotênico (vitamina B5). O aumento da expressão da atividade de DHAD é desejado para aumentar a produção microbiana de aminoácidos de cadeia ramificada ou de ácido pantotênico. Além disso, a conversão catalisada por DHAD de 2,3-dihidroxi isovalerato para  $\alpha$ -cetoisovalerato é uma etapa comum nas vias biossintéticas múltiplas que são divulgadas na patente copendente US 20070092957 A1. É divulgado no presente pedido micro-organismos recombinantes modificados geneticamente para

produção de isobutanol, que é útil como combustível aditivo e cuja disponibilidade pode reduzir a demanda para produtos petroquímicos.

[004] A diol desidratase fornece uma atividade enzimática em uma via biossintética para produção de 2-butanona e 2-butanol, que é divulgada no pedido copendente de patente US 2007-0292927A1. É divulgada na patente US20090155870 uma butanodiol desidratase que é útil para expressão nesta via devido à sua independência à coenzima B-12. Uma diol desidratase reativase que é uma proteína do cluster Fe-S necessária para atividade da butanodiol desidratase independente de B12, também é divulgada na patente US US20090155870. 2-Butanona, também denominada metil etil cetona (MEK), é um solvente amplamente usado, extrator e ativador de reações oxidativas, bem como substrato para síntese química de 2-butanol. 2-butanol é útil como aditivo de combustível, cuja disponibilidade pode reduzir a demanda para combustíveis petroquímicos.

[005] Para aprimoramento da produção de compostos sintetizados nas vias que incluem um cluster Fe-S contendo enzima, é desejável fornecer uma célula hospedeira capaz de expressar altos níveis desta atividade enzimática na produção hospedeira de interesse. Visto que diversas bactérias e leveduras relevantes comercialmente podem expressar atividade de proteínas contendo cluster Fe-S, esta atividade está em níveis inferiores àquelas comercialmente úteis para aumento das vias biossintéticas introduzidas. Consequentemente, existe uma necessidade para a descoberta de células hospedeiras capazes de expressar atividade de proteínas contendo cluster Fe-S em níveis suficientemente altos para aumento de vias introduzidas que têm necessidade de Fe-S. Obter alta expressão funcional de clusters Fe-S heterólogos contendo enzimas é problemático devido à necessidade do cluster Fe-S, que envolve disponibilidade e carga apropriada do cluster na apo-proteína.

### **DESCRIÇÃO RESUMIDA DA INVENÇÃO**

[006] São fornecidas no presente pedido células de levedura hospedeiras recombinantes que compreendem pelo menos uma proteína do cluster Fe-S heteróloga, sendo que a levedura hospedeira tem expressão reduzida de pelo menos uma proteína do cluster Fe-S endógena.

[007] As células de levedura recombinantes podem crescer sob condições adequadas para produção de produtos que incluem isobutanol, 2-butanol e 2-butanona.

[008] Em um aspecto, a célula de levedura recombinante compreende uma ruptura no gene que codifica pelo menos uma proteína do cluster Fe-S endógena.

[009] Em outro aspecto, a proteína do cluster Fe-S endógena é selecionada do grupo que consiste em ácido-dihidroxi-desidratase, isopropilmalato-desidratase, sulfito redutase, glutamato-desidrogenase, biotina sintase, aconitase, homoaconitase, lipoato-sintase, maturação de ferredoxina, NADH-ubiquinona-oxidoredutase, succinato-desidrogenase, ubiquinol-citocromo-c redutase , proteína ABC Rli1, NTPase Nbp35, e proteína similar à hidrogenase.

[010] Em outro aspecto, a levedura é selecionada do grupo que consiste em *Saccharomyces*, *Schizosaccharomyces*, *Hansenula*, *Candida*, *Kluyveromyces*, *Yarrowia* e *Pichia*.

[011] Em outro aspecto, a proteína Fe-S endógena é expressa na mitocôndria, e em outra realização, a proteína do cluster Fe-S endógena tem uma atividade selecionada do grupo que consiste em: atividade de ácido dihidroxi desidratase e isopropilmalato desidratase.

[012] Em outro aspecto, a célula hospedeira é *Saccharomyces* que expressa um gene que codifica um polipeptídeo que tem a sequência de aminoácidos conforme apresentada na SEQ ID NO:114.

[013] Em algumas realizações, pelo menos uma proteína do cluster Fe-S heteróloga é selecionada do grupo que consiste em ácido dihidroxi desidratases 2Fe-2S fúngicas e ácido dihidroxi desidratases 2Fe-2S vegetais. Em uma realização, a ácido dihidroxi desidratase do cluster 2Fe-2S fúngica ou vegetal é expressa no citosol. Em uma realização, a ácido dihidroxi desidratase do cluster 2Fe-2S fúngica ou vegetal heteróloga é um polipeptídeo que tem uma sequência de aminoácidos que é correspondente ao perfil HMM da tabela 9 com um valor E de < 10<sup>-5</sup>, sendo que o polipeptídeo compreende adicionalmente todas as cisteínas conservadas, que correspondem às posições 56, 129 e 201 nas sequências de aminoácidos da enzima DHAD de *Streptococcus mutans* que corresponde à SEQ ID NO:179. Em uma realização, a ácido dihidroxi desidratase do cluster 2Fe-2S fúngica ou vegetal heteróloga é um polipeptídeo que tem uma sequência de aminoácidos que tem pelo menos cerca de 95% de identidade de sequência com a sequência de aminoácidos selecionada do grupo que consiste em SEQ ID NOs:46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150 e 152. Em uma realização, a ácido dihidroxi desidratase do cluster 2Fe-2S fúngica ou vegetal heteróloga é um polipeptídeo que tem uma sequência de aminoácidos que é pelo menos cerca de 90% idêntica à SEQ ID NO:114, usando-se o método de alinhamento Clustal W que usa parâmetros básicos de GAP PENALTY =10, GAP LENGTH PENALTY =0,1, e série Gonnet 250 de matriz de peso de proteína sobre o comprimento total da sequência de proteína.

[014] Em outro aspecto, é fornecido um método para conversão de 2,3-dihidroxi isovalerato para α-cetoisovalerato, o dito método que compreende:

- a) fornecer (1) uma célula hospedeira de levedura recombinante

que compreende pelo menos um gene heterólogo que codifica uma ácido dihidroxi desidratase 2Fe-2S, sendo que a célula de levedura hospedeira recombinante tem atividade reduzida de pelo menos uma proteína do cluster Fe-S endógena; e (2) uma fonte de 2,3-dihidroxi isovalerato; e

b) cultivar a célula hospedeira recombinante de (a) com a dita fonte de 2,3-dihidroxi isovalerato sob as condições onde 2,3-dihidroxi isovalerato é convertido pela célula hospedeira para α-cetoisovalerato.

[015] Em outro aspecto, é fornecido um método para conversão de 2,3-butanodiol para 2-butanona, o dito método compreende:

a) fornecer (1) uma célula de levedura hospedeira recombinante que compreende pelo menos um gene heterólogo que codifica uma propanodiol-desidratase-reativase Fe-S, sendo que célula de levedura hospedeira recombinante tem atividade reduzida de pelo menos uma proteína do cluster Fe-S endógena; e (2) uma fonte de 2,3-butanodiol; e

b) cultivar a célula hospedeira recombinante de (a) com a dita fonte de 2,3-butanodiol sob as condições onde 2,3-butanodiol é convertido pela célula hospedeira para 2-butanona.

[016] Também é fornecido um método para produção de isobutanol que compreende cultivar a célula hospedeira de levedura recombinante, conforme descrita no presente pedido, sob as condições em que o isobutanol é produzido.

[017] Em outras realizações, pelo menos uma proteína do cluster Fe-S heteróloga tem atividade de propanodiol desidratase reativase Fe-S. Em algumas realizações, pelo menos uma proteína do cluster Fe-S heteróloga que tem atividade de propanodiol desidratase reativase Fe-S é uma propanodiol desidratase reativase que tem uma sequência de aminoácidos que é pelo menos cerca 90% idêntica à sequência de aminoácidos apresentada na SEQ ID NO:44, usando-se o método de alinhamento Clustal W que usa parâmetros

básicos de GAP PENALTY=10, GAP LENGTH PENALTY =0,1, e série Gonnet 250 de matriz de peso de proteína sobre o comprimento total da sequência de proteína.

[018] Em algumas realizações, a célula produz 2-butanol, e em algumas realizações a célula produz 2-butanona. Em algumas realizações, a célula comprehende uma via biossintética de 2-butanol, e em algumas realizações a célula comprehende produz a via biossintética de 2-butanona.

#### **BREVE DESCRIÇÃO DAS FIGURAS E DESCRIÇÕES DE SEQUÊNCIAS**

[019] A invenção pode ser melhor compreendida a partir da descrição detalhada a seguir, figuras e descrição de sequências anexas, que formam uma parte deste pedido.

[020] Figura 1 mostra as vias biossintéticas para produção de isobutanol.

[021] Figura 2 mostra as vias biossintéticas para produção de 2-butanona e 2-butanol.

[022] A tabela 9 á uma tabela do perfil HMM para ácido dihidroxi desidratase, com base nas enzimas com funções testadas, preparadas conforme descrito no Exemplo 1. A tabela 9 é apresentada em anexo eletronicamente e é incorporada ao presente pedido como referência.

[023] As sequências a seguir estão em conformidade com 37 C.F.R. 1.821 - 1.825 (“Requisitos para Pedidos de Patente Contendo Divulgações de Sequências de Nucleotídeos e/ou de Aminoácidos – Os Regulamentos De Sequências”) e estão consistentes com a Norma ST.25 (1998) da Organização Mundial de Propriedade Intelectual (WIPO) e as exigências da listagem de sequências da EPO e PCT (Regulamento 5.2 e 49.5(a-bis), e Seção 208 e Anexo C das Instruções Administrativas). Os símbolos e formatos usados para os dados de sequências de nucleotídeos e aminoácidos estão em conformidade com os regulamentos estabelecidas no 37 C.F.R. §1.822.

**TABELA 1****INATIVAÇÃO ALVO DOS GENES QUE CODIFICAM A PROTEÍNA FE-S**

Organismo e gene	Ácido Nucléico SEQ ID NO:	Peptídeo SEQ ID NO:
<i>Saccharomyces cerevisiae</i> LEU 1	1	2
<i>Schizosaccharomyces pombe</i> LEU1	3	4
<i>Candida galbrata</i> CBS 138 LEU1	5	6
<i>Candida albicans</i> SC 5314 LEU1	7	8
<i>Kluyveromyces lactis</i> LEU1	9	10
<i>Yarrowia lipolytica</i> LEU1	11	12
<i>Pichia stipitis</i> LEU1	13	14
<i>Saccharomyces cerevisiae</i> YJM789 ILV3	111	112
<i>Schizosaccharomyces pombe</i> ILV3	93	94
<i>Candida galbrata</i> CBS 138 ILV3	107	108
<i>Candida albicans</i> SC5314 ILV3	101	102
<i>Kluyveromyces lactis</i> I LV3	113	114
<i>Yarrowia lipolytica</i> ILV3	105	106
<i>Pichia stipitis</i> CBS 6054 ILV3	103	104
<i>Saccharomyces cerevisiae</i> ACO1	153	154
<i>Schizosaccharomyces pombe</i> (chromosome II) ACO1	155	156
<i>Schizosaccharomyces pombe</i> (chromosome I) ACO1	157	158
<i>Kluyveromyces lactis</i> NRRL Y-1 140 ACO1	159	160
<i>Candida albicans</i> SC5314 ACO1	161	162
<i>Yarrowia lipolytica</i> CLIB122 ACO1	163	164
<i>Pichia stipitis</i> CBS 6054 ACO1	165	166
<i>Candida glabrata</i> CBS138 (cromossomo F) ACO1	167	168
<i>Candida glabrata</i> CBS138 (cromossomo D) ACO1	169	170
<i>Candida glabrata</i> CBS138 (cromossomo K) ACO1	171	172

**TABELA 2****DHADs 2Fe-2S FÚNGICAS E VEGETAIS EM ADIÇÃO ÀQUELAS NA TABELA 1**

Descrição	Ácido Nucléico SEQ ID NO:	Peptídeo SEQ ID NO:
<i>Chlamydomonas reinhardtii</i>	45	46
<i>Ostreococcus lucimarinus</i> CCE9901	47	48
<i>Vitis vinifera</i> (produto proteína não nomeado: CAO71581.1)	49	50
<i>Vitis vinifera</i> (CAN67446.1)	51	52
<i>Arabidopsis thaliana</i>	53	54

<b>Descrição</b>	<b>Ácido Nucléico SEQ ID NO:</b>	<b>Peptídeo SEQ ID NO:</b>
<i>Oryza sativa</i> (indica grupo-cultivar)	55	56
<i>Physcomitrella patens</i> subsp. <i>patens</i>	57	58
<i>Chaetomium globosum</i> CBS 148.51	59	60
<i>Neurospora crassa</i> OR74A	61	62
<i>Magnaporthe grisea</i> 70-15	63	64
<i>Gibberella zeae</i> PH-1	65	66
<i>Aspergillus niger</i>	67	68
<i>Neosartorya fischeri</i> NRRL 181 (XP_001266525.1)	69	70
<i>Neosartorya fischeri</i> NRRL 181 (XP_001262996181)	71	72
<i>Aspergillus niger</i> (An03gO4520)	73	74
<i>Aspergillus niger</i> (An14g03280)	75	76
<i>Aspergillus terreus</i> N I H2624	77	78
<i>Aspergillus clavatus</i> NRRL 1	79	80
<i>Aspergillus nidulans</i> FGSC A4	81	82
<i>Aspergillus oryzae</i>	83	84
<i>Ajellomyces capsulatus</i> NAm1	85	86
<i>Coccidioides immitis</i> RS	87	88
<i>Botryotinia fuckeliana</i> B05.10	89	90
<i>Phaeosphaeria nodorum</i> SN15	91	92
<i>Pichia guilliermondii</i> ATCC 6260	95	96
<i>Debaryomyces hansenii</i> CBS767	97	98
<i>Lodderomyces elongisporus</i> NRRL YB-4239	99	100
<i>Vanderwaltozyma polyspora</i> DSM 70294	109	110
<i>Ashbya gossypii</i> ATCC 10895	115	116

<b>Descrição</b>	<b>Ácido Nucléico SEQ ID NO:</b>	<b>Peptídeo SEQ ID NO:</b>
<i>Laccaria bicolor</i> S238N-H82	117	118
<i>Coprinopsis cinerea</i> okayama7#1 30	119	120
<i>Cryptococcus neoformans</i> var. <i>neoformans</i> JEC21	121	122
<i>Ustilago maydis</i> 521	123	124
<i>Malassezia globosa</i> CBS 7966	125	126
<i>Aspergillus clavatus</i> NRRL 1	127	128
<i>Neosartorya fischeri</i> NRRL 181 (Suposto)	129	130
<i>Aspergillus oryzae</i>	131	132
<i>Aspergillus niger</i> (An 18g04160)	133	134
<i>Aspergillus terreus</i> N I H2624	135	136
<i>Coccidioides immitis</i> RS (CIMG_04591)	137	138
<i>Paracoccidioides brasiliensis</i>	139	140
<i>Phaeosphaeria nodorum</i> SN15	141	142
<i>Gibberella zae</i> PH-1	143	144
<i>Neurospora crassa</i> OR74A	145	146
<i>Coprinopsis cinerea</i> okayama 7#130	147	148
<i>Laccaria bicolor</i> S238N-H82	149	150
<i>Ustilago maydis</i> 521	151	152

**TABELA 3****GENES DE EXPRESSÃO**

<b>Descrição</b>	<b>Ácido Nucléico SEQ ID NO:</b>	<b>Peptídeo SEQ ID NO:</b>
<i>Roseburia inulinivorans</i> (RdhtA)	15	43
<i>Roseburia inulinivorans</i> (RdhtB)	16	44
<i>Bacillus subtilis</i> (alsS)	27	28
<i>Vibrio cholerae</i> (KARI)	35	36
<i>Pseudomonas aeruginosa</i> PAO1 (KARI)	37	38
<i>Pseudomonas fluorescens</i> PF5 (KARI)	39	40

Descrição	Ácido Nucléico SEQ ID NO:	Peptídeo SEQ ID NO:
<i>Achromobacterxylosoxidans</i> (sadB)	41	42
Glicerol desidratase independente de B12 de <i>Clostridium butyricum</i>	190	191
Butanodiol desidratase reativase independente de B12 de <i>Clostridium butyricum</i>	192	193

[024] SEQ ID NO:17 é uma sequência rdhtAB sintética.

[025] SEQ ID NOs:18-21 e 30-33 são *primers* para análises usadas para PCR, clonagem ou sequenciamento, conforme descrito nos Exemplos no presente pedido.

[026] SEQ ID NO:22 é uma sequência terminadora dupla.

[027] SEQ ID NO:23 é o terminador ADH de *Saccharomyces cerevisiae*.

[028] SEQ ID NO:24 é o terminador CYC1 de *Saccharomyces cerevisiae*.

[029] SEQ ID NO:25 é o promotor FBA de *Saccharomyces cerevisiae*.

[030] SEQ ID NO:26 é o promotor GPM de *Saccharomyces cerevisiae*.

[031] SEQ ID NO:29 é o vetor pNY13.

[032] SEQ ID NO:34 é o promotor CUP1 de *Saccharomyces cerevisiae*.

[033] SEQ ID NO:173 é a região codificadora códon otimizada para ILV3 DHAD de *Kluyveromyces lactis*.

#### TABELA 4

#### DHADS VERIFICADOS FUNCIONALMENTE PARA PERFIL HMM

Organismo	Ácido Nucléico SEQ ID NO:	Peptídeo SEQ ID NO:
<i>Nitrosomonas europaea</i> ATCC 19718	174	175
<i>Synechocystis</i> sp. PCC 6803	176	177
<i>Streptococcus mutans</i> UA159	178	179

Organismo	Ácido Nucléico SEQ ID NO:	Peptídeo SEQ ID NO:
<i>Streptococcus thermophilus</i> LMG 18311	180	181
<i>Ralstonia metallidurans</i> CH34	182	183
<i>Ralstonia eutropha</i> JMP134	184	185
<i>Lactococcus lactis</i> subsp. <i>cremoris</i> SK11	186	187
<i>Flavobacterium johnsoniae</i> UW101	188	189

#### DESCRÍÇÃO DETALHADA DA INVENÇÃO

[034] É divulgada no presente pedido a descoberta de que proteínas contendo Fe-S introduzidas nas células hospedeiras de levedura têm altos níveis de atividade quando a expressão de proteínas endógenas contendo Fe-S é inibida ou interrompida. A presente invenção refere-se a células de leveduras recombinantes modificadas geneticamente para fornecer expressão de pelo menos uma proteína heteróloga, que é uma proteína do cluster Fe-S, e modificadas geneticamente para reduzir a expressão de pelo menos uma proteína do cluster Fe-S endógena. Nestas células a atividade da proteína do cluster Fe-S heteróloga é aprimorada, de modo que haja produção aprimorada de um produto produzido em uma via biossintética que inclui a atividade enzimática. Exemplos de produtos úteis comercialmente de uma via que inclui proteína Fe-S incluem valina, isoleucina, leucina, ácido pantotênico, isobutanol, 2-butanona e 2-butanol.

[035] As abreviações e definições a seguir são usadas para a interpretação das reivindicações e do relatório descriptivo.

[036] Como usado no presente pedido, os termos “compreende”, “que comprehende”, “inclui”, “que inclui”, “tem”, “que tem”, “contém”, “que contém”, ou qualquer outra variação destinam-se a cobrir uma inclusão não exclusiva. Por exemplo, uma composição, uma mistura, um processo, método, artigo ou aparelho, que comprehende uma lista de elementos não está necessariamente limitado a apenas aqueles elementos, mas pode incluir outros

elementos não expressamente listados ou inerentes a tal composição, mistura, processo, método, artigo ou aparelho. Além disso, a menos que expressamente declarado em contrário, “ou” refere-se a uma inclusão e não a uma exclusão. Por exemplo, uma condição A ou B é satisfeita por qualquer um dos seguintes: A é verdadeiro (ou presente) e B é falso (ou não presente), A é falso (ou não presente) e B é verdadeiro (ou presente), e ambos, A e B são verdadeiros (ou presentes).

[037] Também, os artigos indefinidos “um” e “uma” antes de um elemento ou componente da invenção têm a intenção de não ser restritos em relação aos exemplos (isto é, ocorrências) do elemento ou componente. Portanto “um” ou “uma” deveria ser entendido como um(a) ou pelo menos um(a), e a forma da palavra no singular do elemento ou componente também inclui o plural, a menos que o número esteja de forma óbvia no singular.

[038] O termo “invenção” ou “presente invenção”, como usado no presente pedido, é um termo não limitante e não tem a intenção de se referir a qualquer realização única da invenção particular, mas engloba todas as realizações possíveis, conforme descrito no relatório descritivo e nas reivindicações.

[039] Como usado no presente pedido, o termo empregado “cerca de” modifica a quantidade de um ingrediente ou reagente da invenção e refere-se à variação na quantidade numérica que pode ocorrer, por exemplo, através de medidas típicas e procedimentos de manuseio de líquidos usados para fazer concentrados ou uso de soluções na realidade; através de erros negligentes nestes procedimentos; através de diferenças na fabricação, fonte ou pureza dos ingredientes empregados para se fazer as composições ou realizar os métodos; e similares. O termo “cerca de” também engloba quantidades que diferem devido às diferentes condições de equilíbrio para uma composição resultante de uma mistura inicial específica. Se modificado ou não

pelo termo “cerca de”, as reivindicações incluem equivalentes às quantidades. Em uma realização, o termo “cerca de” significa dentro de uma tolerância de 10% do valor numérico relatado, preferencialmente dentro de uma tolerância de 5% do valor numérico relatado.

[040] O termo “proteína de cluster Fe-S” é uma proteína que se liga a um cluster ferro-enxofre e necessita da ligação ao cluster para sua atividade.

[041] O termo "2Fe-2S DHAD" refere-se às enzimas DHAD que necessitam de um cluster[2Fe-2S]<sup>2+</sup> para sua atividade.

[042] O termo "reativase propanodiol desidratase Fe-S" refere-se às reativases propanodiol desidratases que necessitam da ligação ao cluster Fe-S para sua atividade.

[043] O termo “via biossintética do isobutanol” refere-se a uma via enzimática para produzir isobutanol a partir de piruvato.

[044] O termo “via biossintética de 2-butanol” refere-se a uma via enzimática para produzir 2-butanol a partir de piruvato.

[045] O termo “via biossintética de 2-butanona” se refere a uma via enzimática para produzir 2-butanona a partir de piruvato.

[046] O termo "ácido-dihidroxi-desidratase", também abreviada como DHAD, refere-se a uma enzima que converte 2,3-dihidroxi isovalerato para α-cetoisovalerato.

[047] O termo “butanodiol-desidrogenase”, também conhecida como “diol-desidratase” ou “propanodiol-desidratase” refere-se a um polipeptídeo (ou polipeptídeos) que apresentam uma atividade enzimática que catalisa a conversão de 2,3-butanodiol para 2-butanona. Butanodiol desidratases que não utilizam o cofator adenosil cobalamina (também conhecido como coenzima B12 ou vitamina B12, embora a vitamina B12 possa também se referir a outras formas de cobalamina que não são coenzima B12)

são coenzimas diol desidratases independentes de B12 que necessitam da associação com uma diol desidratase reativase, que é uma proteína do cluster Fe-S. Exemplos de diol desidratases independentes de B12 incluem aquelas de *Clostridium glycolicum* (Hartmanis *et al.* (1986) *Arch. Biochem. Biophys.* 245:144-152), *Clostridium butyricum* (proteína da SEQ ID NO:191; região codificadora da SEQ ID NO:190; O'Brien *et al.* (2004) *Biochemistry* 43:4635-4645), e *Roseburia inulinivorans* (codificação: SEQ ID NO:15; proteína: SEQ ID NO:43; divulgada na patente copendente US 20090155870).

[048] O termo "propanodiol-desidratase-reactivase", também conhecida como "diol-desidratase-reactivase" ou "butanodiol-desidratase-reactivase" refere-se a um fator de reativação para diol-desidratase, uma enzima que se submete à inativação durante a catálise. Diol desidratases reativases associadas às coenzimas diol desidratases independentes de B12 podem ser proteínas do cluster Fe-S. Exemplos incluem aquelas de *Clostridium glycolicum* (Hartmanis *et al.* (1986) *Arch. Biochem. Biophys.* 245:144-152), *Clostridium butyricum* (proteína da SEQ ID NO:193; região codificadora da SEQ ID NO:192; O'Brien *et al.* (2004) *Biochemistry* 43:4635-4645), e *Roseburia inulinivorans* (codificação: SEQ ID NO:16; proteína: SEQ ID NO:44; divulgada na patente copendente de propriedade pública US 20090155870).

[049] O termo “expressão reduzida”, como aplicado à expressão de uma proteína em uma célula hospedeira, incluirá aquelas situações onde a atividade da proteína é diminuída, quando comparada com a forma tipo selvagem (como com tecnologia *antisense*, por exemplo), ou substancialmente eliminada com a ruptura, deleção ou inativação do gene, por exemplo.

[050] O termo “substrato de carbono” ou “substrato de carbono fermentável” refere-se a uma fonte de carbono capaz de ser metabolizada por organismos hospedeiros da presente invenção, e em particular fontes de carbono selecionadas do grupo que consiste em monossacarídeos,

oligossacarídeos, polissacarídeos e substratos com um carbono ou misturas destes.

[051] O termo “gene” refere-se a um fragmento de ácido nucléico que é capaz de ser expresso como uma proteína específica, opcionalmente, incluindo sequências reguladoras anteriores (sequências não codificadoras 5') e posteriores (sequências não codificadoras 3') da sequência codificadora. “Gene nativo” refere-se a um gene como encontrado na natureza, com suas próprias sequências reguladoras. “Gene quimérico” refere-se a qualquer gene não nativo, que compreende sequências reguladoras e codificadoras que não são encontradas juntas na natureza. Consequentemente, um gene quimérico pode compreender sequências reguladoras e sequências codificadoras derivadas de diferentes fontes, ou sequências reguladoras e sequências codificadoras derivadas da mesma fonte, mas organizadas de uma maneira diferente daquela encontrada na natureza. “Gene endógeno” refere-se a um gene nativo em sua localização natural no genoma de um organismo. Um “gene estranho” ou “gene heterólogo” refere-se a um gene não encontrado normalmente no organismo hospedeiro, mas que é introduzido no organismo hospedeiro por transferência gênica. “Gene heterólogo” inclui uma região codificadora nativa ou porção desta que é reintroduzida no organismo fonte sob uma forma diferente daquela do gene nativo correspondente. Por exemplo, um gene heterólogo pode incluir uma região codificadora nativa que é uma porção de um gene quimérico que inclui regiões reguladoras não-nativas que são introduzidas no hospedeiro nativo. Um genes estranho também pode compreender genes nativos inseridos em um organismo não-nativo, ou genes quiméricos. Um “transgene” é um gene que foi introduzido no genoma por um processo de transformação.

[052] Como usado no presente pedido, o termo “região codificadora” refere-se a uma sequência de DNA que codifica uma sequência

específica de aminoácido. “Sequências reguladoras adequadas” referem-se às sequências de nucleotídeos localizadas a montante (sequências não codificadoras 5’), dentro, ou a jusante (sequências não codificadoras 3’) de uma sequência codificadora, e que influenciam a transcrição, processo ou estabilidade do RNA, ou tradução da sequência codificadora associada. Sequências reguladoras podem incluir promotores, sequências líder de tradução, introns, sequências de reconhecimento de poliadenilação, sítio de processamento de RNA, sítio de ligação efetora e estrutura derivada da alça.

[053] O termo “promotor” refere-se a uma sequência de DNA capaz de controlar a expressão de uma sequência codificadora de RNA ou funcional. Em geral, a sequência codificadora está localizada 3' de uma sequência promotora. Promotores podem ser derivados em sua totalidade de um gene nativo ou podem ser compostos de diferentes elementos derivados de diferentes promotores encontrados na natureza, ou até compreender segmentos sintéticos de DNA. Os técnicos no assunto entendem que diferentes promotores podem dirigir a expressão de um gene em diferentes tecidos ou tipos celulares, em diferentes estágios do desenvolvimento, ou em resposta a diferentes condições ambientais ou psicológicas. Promotores que induzem um gene a se expressar na maioria dos tipos celulares, na maioria das vezes são comumente chamados de “promotores constitutivos”. Sabe-se, ainda, que visto que em muitos casos os limites exatos das sequências reguladoras não foram completamente definidos, fragmentos de DNA de diferentes comprimentos podem ter atividade idêntica àquela do promotor.

[054] O termo “ligado de maneira operacional” refere-se à associação de sequências de ácido nucléico com um fragmento de ácido nucléico, de forma que a função de um seja afetada pelo outro. Por exemplo, um promotor está ligado de maneira funcional a uma sequência de codificação quando este é capaz de influenciar a expressão desta sequência codificadora

(isto é, a sequência codificadora está sob o controle de transcrição do promotor). Sequências codificadoras podem estar ligadas de maneira funcional às sequências reguladoras na orientação *sense* ou *antisense*.

[055] O termo “expressão”, como usado no presente pedido, refere-se à transcrição e acúmulo estável de RNA *sense* (mRNA) ou *antisense*, derivados de um fragmento de ácido nucléico da invenção. Expressão também pode se referir à tradução de RNAm para um polipeptídeo.

[056] Como usado no presente pedido, “transformação” refere-se à transferência de um fragmento de ácido nucléico em um organismo hospedeiro, resultando em herança geneticamente estável. Organismos hospedeiros que contêm os fragmentos de ácido nucléico transformados são chamados de organismos “transgênicos”, “recombinantes” ou “transformados”.

[057] Os termos “plasmídio” e “vetor”, como usado no presente pedido, referem-se a um elemento extracromossômico que frequentemente carrega genes que não são parte do metabolismo central da célula e, usualmente, na forma de fragmentos de moléculas de DNA circular dupla fita. Esses elementos podem ser sequências de replicação autônomas, sequências de integração de genoma, sequências de nucleotídeos ou de fago, lineares ou circulares, de DNA ou RNA dupla fita ou simples, derivadas de qualquer fonte, em que várias sequências de nucleotídeos foram recombinadas ou unidas em uma única construção capaz de introduzir um fragmento promotor e uma sequência de DNA para um produto do gene selecionado, junto com a sequência não traduzida 3’ apropriada em uma célula.

[058] Como usado no presente pedido “degeneração de códon” refere-se à natureza no código genético que permite variação da sequência de nucleotídeo, sem afetar a sequência de aminoácido de um polipeptídeo codificado. O técnico no assunto está bem informado sobre a “preferência de uso de códon” exibida por uma célula hospedeira específica, no uso dos

códons de nucleotídeos para especificar um dado aminoácido. Portanto, na síntese de um gene para melhorar a expressão em uma célula hospedeira, é desejável criar o gene, de modo que sua frequência de uso de códon se aproxime da frequência preferida de uso de códon da célula hospedeira.

[059] O termo “códon otimizado” quando se referindo a genes ou regiões codificadoras de moléculas de ácido nucléico para transformação de vários hospedeiros, refere-se à alteração de códons no gene ou regiões codificadoras das moléculas de ácido nucléico, para refletir o uso típico do códon do organismo sem alterar o polipeptídeo codificado pelo DNA.

[060] Como usado no presente pedido, “fragmento isolado de ácido nucléico” ou “molécula isolada de ácido nucléico” serão usados alternadamente e significam um polímero de RNA ou DNA de fita simples ou dupla, opcionalmente contendo bases de nucleotídeos sintéticos, não naturais ou alterados. Um fragmento de ácido nucléico isolado na forma de um polímero de DNA pode ser compreendido por um ou mais segmentos de cDNA, DNA genômico, DNA sintético.

[061] Um fragmento de ácido nucléico “hibridizável” a outro fragmento de ácido nucléico, como cDNA, DNA genômico ou molécula de RNA, quando na forma de fita dupla de fragmento de ácido nucléico pode anelar a outro fragmento de ácido nucléico mediante condições apropriadas de temperatura e força iônica da solução. As condições de hibridização e lavagem são bem conhecidas e exemplificadas em Sambrook, J., Fritsch, E. F. e Maniatis, T. *Molecular Cloning: A Laboratory Manual*, 2<sup>a</sup> ed., Cold Spring Harbor Laboratory: Cold Spring Harbor, NY (1989), particularmente capítulo 11 e tabela 11.1 neste (integralmente incorporado ao presente pedido como referência). As condições de temperatura e força iônica determinam a “estringência” da hibridização. Condições de estringência podem ser ajustadas para selecionar fragmentos moderadamente similares (como sequências

homólogas de organismos relacionados de forma distante), até fragmentos altamente similares (como genes que duplicam enzimas funcionais de organismos intimamente relacionados). A pós-hibridização lava determinadas condições de estringência. Um conjunto de condições preferenciais usa uma série de lavagens iniciando com 6X SSC, SDS a 0,5% em temperatura ambiente por 15 min, e então repetido com 2X SSC, SDS a 0,5% a 45°C por 30 min, e então repetido 0,2 X SSC, SDS a 0,5% a 50°C por 30 min. Um conjunto de condições mais preferenciais usa temperaturas mais altas cujas lavagens são idênticas àquelas acima, exceto a temperatura das duas lavagens finais de 30 minutos em 0,2X SSC, SDS a 0,5% que foi aumentada para 60°C. Outro conjunto preferencial de condições altamente estringentes usa duas lavagens finais em 0,1 X SSC, SDS a 0,1 % a 65°C. Um conjunto adicional de condições estringentes de hibridização a 0,1 X SSC, SDS a 0,1 %, 65°C e lavagens com 2X SSC, SDS a 0,1 % seguido por 0,1 X SSC, SDS a 0,1 %, por exemplo.

[062] A hibridização necessita que dois ácidos nucléicos contenham sequências complementares, embora dependentes da estringência da hibridização, o emparelhamento mal sucedido entre as bases é possível. A estringência apropriada para hibridizar os ácidos nucléicos depende do comprimento dos ácidos nucléicos e do grau de complementação, variáveis bem conhecidas na técnica. Quanto maior o grau de similaridade ou homologia entre as sequências de nucleotídeos, maior o valor de  $T_m$  para híbridos de ácidos nucléicos que têm essas sequências. A estabilidade relativa (que corresponde a  $T_m$  maior) de ácidos nucléicos diminui na seguinte ordem: RNA:RNA, DNA:RNA, DNA:DNA. Para híbridos maiores que 100 nucleotídeos de comprimento, as equações para cálculo de  $T_m$  foram derivadas (consulte Sambrook *et al.*, acima, 9.50-9.51). Para hibridizações com ácidos nucléicos mais curtos, isto é, oligonucleotídeos, a posição de emparelhamentos mal sucedidos torna-se mais importante, e o comprimento do oligonucleotídeo

determina sua especificidade (consulte Sambrook *et al.*, acima, 11.7-11.8). Em uma realização, o comprimento para um ácido nucléico hibridizável é de pelo menos cerca de 10 nucleotídeos. Preferencialmente, um comprimento mínimo para um ácido nucléico hibridizável é de pelo menos cerca de 15 nucleotídeos, com mais preferência de cerca de 20 nucleotídeos, e com a máxima preferência o comprimento é de cerca de 30 nucleotídeos. Além disso, o técnico no assunto reconhecerá que a temperatura e a concentração de sal da solução de lavagem podem ser ajustadas conforme necessário, de acordo com fatores como comprimento da sonda.

[063] Uma “porção substancial” de uma sequência de aminoácido ou de nucleotídeo é aquela porção que compreende sequência suficiente de aminoácido de um polipeptídeo ou a sequência de nucleotídeo de um gene para supostamente identificar esse polipeptídeo ou gene, tanto por avaliação manual da sequência por um técnico no assunto, como por comparação e identificação automática de sequência por computador, usando-se algoritmos como BLAST (Altschul, S. F., *et al.*, *J. Mol. Biol.*, 215:403-410 (1993)). Em geral, é necessária uma sequência de dez ou mais aminoácidos adjacentes ou trinta ou mais nucleotídeos para supostamente identificar um polipeptídeo ou sequência de ácido nucléico como homólogo(a) a uma proteína ou gene conhecidos. Além disso, com relação às sequências de nucleotídeos, as sondas de oligonucleotídeos especificadas do gene que compreendem 20 a 30 nucleotídeos adjacentes podem ser usadas nos métodos de identificação do gene dependentes de sequência (por exemplo, hibridização Southern) e isolamento (por exemplo, hibridização *in situ* de colônias de bactérias ou placas de bacteriófagos). Além disso, os oligonucleotídeos curtos com 12 a 15 bases podem ser usados como *primers* para amplificação em PCR para se obter um fragmento específico de ácido nucléico que compreende os *primers*. Consequentemente, uma “porção substancial” de uma sequência de

nucleotídeo compreende sequência suficiente para identificar e/ou isolar especificamente um fragmento de ácido nucléico que compreende a sequência. O atual relatório descritivo ensina a sequência completa de aminoácidos e de nucleotídeos que codificam proteínas fúngicas específicas. O técnico no assunto, tendo o benefício das sequências relatadas no presente pedido, pode agora usar toda ou uma porção substancial das sequências divulgadas para propósitos conhecidos pelos técnicos no assunto. Consequentemente, a atual invenção compreende as sequências completas conforme relatado na Listagem de Sequências anexa, bem como porções substanciais dessas sequências conforme definido acima.

[064] O termo “complementar” é usado para descrever a relação entre bases de nucleotídeos que são capazes de hibridizar uma à outra. Por exemplo, com relação ao DNA, a adenosa é complementar à timina, e citosina é complementar à guanina.

[065] O termo “porcentagem de identidade”, como conhecido na técnica, é uma relação entre duas ou mais sequências de polipeptídeos ou duas ou mais sequências de polinucleotídeos, conforme determinado por comparação das sequências. Na técnica, “identidade” também significa o grau de relação de sequências entre as sequências de polipeptídeos ou polinucleotídeos, como pode ser o caso, conforme determinado pelo emparelhamento entre as fitas de tais sequências. (“Identidade” e “similaridade” podem ser facilmente calculadas por métodos conhecidos, incluindo, mas não se limitando àqueles descritos em: 1.) *Computational Molecular Biology* (Lesk, A. M., Ed.) Oxford University: NY (1988); 2.) *Biocomputing: Informatics and Genome Projects* (Smith, D. W., Ed.) Academic: NY (1993); 3.) *Computer Analysis of Sequence Data, Part I* (Griffin, A. M., e Griffin, H. G., Eds.) Humania: NJ (1994); 4.) *Sequence Analysis in Molecular Biology* (von Heinje, G., Ed.) Academic (1987); e 5.) *Sequence Analysis Primer* (Gribskov, M. e

Devereux, J., Eds.) Stockton: NY (1991).

[066] Métodos preferenciais para determinar a identidade são projetados para fornecer o melhor emparelhamento entre as sequências testadas. Métodos para determinar a identidade e similaridade são codificados em programas de computador publicamente disponíveis. Alinhamentos de sequências e os cálculos de identidade de sequência podem ser realizados usando-se o programa MegAlign™ do conjunto de computação de bioinformática LASERGENE (DNASTAR Inc., Madison, WI). O alinhamento múltiplo das sequências é realizado usando-se o “método de alinhamento Clustal” que engloba diversas variedades de algoritmos incluindo o “método de alinhamento Clustal V” que corresponde ao método de alinhamento designado Clustal V (descrito por Higgins e Sharp, CABIOS. 5:151-153 (1989); Higgins, D.G. et al., *Comput. Appl. Biosci* 8:189-191 (1992)) e encontrado no programa MegAlign™ do conjunto de computação de bioinformática LASERGENE (DNASTAR Inc., Madison, WI). Para múltiplos alinhamentos, os parâmetros básicos correspondem a GAP PENALTY=10 e PENALIDADE DE EXTENSÃO DE GAP=10. Os parâmetros básicos para alinhamentos de pareamento e cálculo de percentual de identidade de sequências de proteínas usando-se o método Clustal são KTUPLE=1, GAP PENALTY=3, JANELA=5 e DIAGONAIS SALVAS=5. Para os ácidos nucléicos, estes parâmetros são KTUPLE 1, GAP PENALTY=3, JANELA=5 e DIAGONAIS SALVAS=5. Após o alinhamento das sequências usando-se o programa Clustal V, é possível obter um “percentual de identidade” visualizando a tabela de “distância das sequências” no mesmo programa. Adicionalmente, o “Método de alinhamento Clustal W” está disponível e corresponde ao método de alinhamento designado Clustal V (descrito por Higgins e Sharp, CABIOS. 5:151-153 (1989); Higgins, D.G. et al., *Comput. Appl. Biosci* 8:189-191 (1992)) e encontrado no programa MegAlign™ v. 6.1 do conjunto de computação de bioinformática LASERGENE (DNASTAR

Inc., Madison, WI). Os parâmetros básicos para alinhamento múltiplo (GAP PENALTY=10, PENALIDADE DE EXTENSÃO DE GAP=0,2, Divergência no Atraso de Sequências (5)=30, Transição de Peso de DNA=0,5, Matriz de Peso de Proteína=Série Gonnet, Matriz de Peso de DNA=IUB). Após o alinhamento das sequências usando-se o programa Clustal W, é possível se obter uma “porcentagem de identidade” visualizando a tabela de “distância das sequências” no mesmo programa.

[067] É bem entendido por um técnico no assunto que muitos níveis de identidade de sequência são úteis na identificação de polipeptídeos de outras espécies, sendo que tais polipeptídios possuem a mesma função ou atividade, ou similar. Exemplos úteis de identidades de porcentagens incluem, mas não se limitam a: 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, ou 95%, ou qualquer porcentagem de número inteiro a partir de 55% a 100% podem ser úteis na descrição da presente invenção como, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% ou 99%. Fragmentos de ácidos nucléicos adequados não têm apenas as homologias acima, mas tipicamente codificam um polipeptídeo tendo pelo menos 50 aminoácidos, preferencialmente pelo menos 100 aminoácidos, com mais preferência pelo menos 150 aminoácidos, ainda com mais preferência pelo menos 200 aminoácidos e com a máxima preferência pelo menos 250 aminoácidos.

[068] O termo “programa para análise de sequência” refere-se a qualquer algoritmo ou programa de computação útil para análise de sequências de nucleotídeos ou aminoácidos. O “programa para análise de sequência” pode estar comercialmente disponível ou pode ser desenvolvido de forma independente. (O programa típico para análise de sequência incluirá, mas não está limitado a: 1.) o conjunto de programas GCG (Pacote Wisconsin Versão

9.0, Genetics Computer Group (GCG), Madison, WI); 2.) BLASTP, BLASTN, BLASTX (Altschul *et al.*, *J. Mol. Biol.*, 215:403--410 (1990)); 3.) DNASTAR (DNASTAR, Inc. Madison, WI); 4.) Sequencher (Gene Codes Corporation, Ann Arbor, MI); e 5.) o programa FASTA que incorpora o algoritmo Smith-Waterman (W. R. Pearson, *Comput. Methods Genome Res.*, [Proc. Int. Symp.] (1994), Data do Encontro 1992, 111-20. Editores: Suhai, Sandor. Plenum: Nova York, NY). Dentro do contexto desta aplicação, deve-se entender que onde os programas de análise de sequências forem utilizados para análise, os resultados das análises serão baseados nos “parâmetros básicos” do programa referido, a menos que especificado de outra forma. Como usado no presente pedido “parâmetros básicos” significa qualquer conjunto de valores ou parâmetros que originalmente são carregados no programa, quando inicializado pela primeira vez.

[069] Técnicas padrão de DNA recombinante e de biologia molecular usadas no presente pedido são bem conhecidas na técnica e são descritas de forma por Sambrook, J., Fritsch, E.F. e Maniatis, T. *Molecular Cloning: A Laboratory Manual*, Segunda Edição, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1989) (deste ponto em diante “Maniatis”); e por Silhavy, T. J., Bennan, M. L. e Enquist, L. W., *Experiments with Gene Fusions*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1984); e por Ausubel, F. M. *et al.*, *Current Protocols in Molecular Biology*, publicado por Greene Publishing Assoc. e Wiley-Interscience (1987). Métodos adicionais aqui utilizados estão em *Methods in Enzymology*, Volume 194, *Guide to Yeast Genetics and Molecular and Cell Biology* (Parte A, 2004, Christine Guthrie e Gerald R. Fink (Eds.), Elsevier Academic Press, San Diego, CA).

### **DESCOBERTA DA ATIVIDADE APRIMORADA DA PROTEÍNA DO CLUSTER FE-S EM LEVEDURAS**

[070] Proteínas que contêm uma ligação do cluster ferro-enxofre

(Fe-S), que é necessária para suas atividades, podem ter baixa atividade quando usadas nos sistemas de expressão de heterólogos. A formação de clusters Fe-S e sua transferência para apoproteínas é um processo de multietapas que envolve ao menos diversas proteínas que incluem cisteína disulfurase, uma proteína molde e uma chaperona. Dessa forma, uma proteína Fe-S heteróloga pode não ser efetivamente composta pelo sistema hospedeiro endógeno. Os requerentes descobriram uma maneira de aumentar a atividade de uma proteína Fe-S expressa como uma proteína heteróloga em uma célula hospedeira de levedura. Os requerentes descobriram que pela produção reduzida de uma proteína Fe-S endógena na célula de levedura hospedeira pode-se alcançar um aprimoramento na atividade de uma proteína expressa do cluster Fe-S heteróloga. A expressão em leveduras de outras ácido dihidroxi desidratase 2Fe-2S fúngicas ou vegetais heterólogas ou Fe-S propanodiol desidratase reativase (RdhtB) foi aprimorada quando um gene endógeno que codifica isopropilmalato desidratase (LEU1) ou um gene endógeno que codifica ácido dihidroxi desidratase (ILV3) foi inativado nas células hospedeiras de levedura.

[071] Em células de levedura hospedeiras com inativação de um gene que codifica uma proteína Fe-S endógena, a atividade da proteína Fe-S heteróloga expressa pode ser aumentada para pelo menos 1,4 vezes da atividade em uma célula de levedura hospedeira com nenhuma ativação do gene que codifica a proteína Fe-S. Por exemplo, DHAD de *Kluyveromyces lactis* tem atividade 1,4 vezes em um hospedeiro com deleção de LEU1, em comparação a um hospedeiro sem a deleção; Rdht de *Roseburia inulinivorans* tem 1,7 vezes, em comparação com a atividade um hospedeiro com deleção de LEU, conforme medido pela proteína RdhtA ativada (descrito abaixo); DHAD de *Saccharomyces cerevisiae* expressa no citosol tem 1,5 vezes da atividade comparativa em um hospedeiro com deleção de ILV3 mitocondrial; e DHAD de

*Kluyveromyces lactis* expresso no citosol tem 7,4 vezes da atividade comparativa em um hospedeiro com deleção de ILV3 mitocondrial.

### **CÉLULAS HOSPEDEIRAS DE LEVEDURA COM EXPRESSÃO REDUZIDA DE PROTEÍNA Fe-S ENDÓGENA**

[072] A expressão reduzida da proteína Fe-S endógena pode ser modificada geneticamente em qualquer célula de levedura que seja receptiva à manipulação genética. Exemplos incluem leveduras *Saccharomyces*, *Schizosaccharomyces*, *Hansenula*, *Candida*, *Kluyveromyces*, *Yarrowia* e *Pichia*. Linhagens adequadas incluem, mas não se limitam a *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Kluyveromyces lactis*, *Kluyveromyces thermotolerans*, *Candida glabrata*, *Candida albicans*, *Pichia stipitis* e *Yarrowia lipolytica*. Particularmente adequado é *Saccharomyces cerevisiae*.

[073] Em qualquer destas leveduras, qualquer proteína Fe-S endógena pode ser um alvo para expressão reduzida. Proteínas Fe-S em leveduras que podem ser direcionadas para expressão reduzida incluem, por exemplo, as proteínas a seguir (com gene codificador): aconitase (AC01), homoaconitase (LYS4), DHAD (ILV3), lipoato sintase (LIPS), biotina sintase (B102), maturação de ferredoxina (YAH1), NADH ubiquinona oxidoredutase (ND11), succinato desidrogenase (SDH2), ubiquinol-citocromo-c redutase (RIP1), isopropilmalato isomerase (LEU1), sulfito redutase (ECM17), glutamato desidrogenase (GLT1), proteína ABC R11 (RI-11), NTPase Nbp35 (NBP35), e proteína similar a hidrogenase (NAR11). Células de levedura com expressão reduzida de proteínas Fe-S individuais podem necessitar de condições especiais para crescimento, como suplementação do meio de crescimento com um nutriente específico, como de conhecimento por um técnico no assunto. Por exemplo, uma linhagem com ruptura de LEU1 é suplementada com leucina, uma linhagem com ruptura de DHAD é suplementada com leucina, isoleucina e

valina e uma linhagem com ruptura de LYS4 é suplementada com lisina. Algumas linhagens com uma ruptura não necessitam de suplementação para crescimento. Proteínas Fe-S particularmente adequadas que podem ser direcionadas para expressão reduzida incluem Isopropilmalato isomerase (LEU1), Ácido dihidroxi desidratase (ILV3), Sulfito redutase (ECM17), Glutamato desidrogenase (GLT1), e Biotina sintase (B102). A expressão reduzida é modificada geneticamente para que pelo menos uma proteína Fe-S endógena, e duas ou mais proteínas Fe-S endógenas sejam reduzidas.

[074] LEU1 codifica isopropilmalato desidratase, uma enzima que pertence a EC 4.2.1.33, que está envolvida na biossíntese de aminoácidos de cadeia ramificada, especificamente síntese de leucina. Qualquer gene que codifica uma isopropilmalato desidratase, que é uma enzima que necessita de um cluster 4Fe-4S para sua atividade, pode ser inativado em uma célula de levedura hospedeira desta divulgação. Exemplos de genes alvo de leveduras LEU1 para inativação e suas proteínas codificadas são aqueles de *Saccharomyces cerevisiae* (codificação SEQ ID NO:1; proteína SEQ ID NO:2), *Schizosaccharomyces pombe* (codificação SEQ ID NO:3; proteína SEQ ID NO:4), *Candida galbrata* strain CBS 138 (codificação SEQ ID NO:5; proteína SEQ ID NO:6), *Candida albicans* SC5314 (codificação SEQ ID NO:7; proteína 20 SEQ ID NO:8), *Kluyveromyces lactis* (codificação SEQ ID NO:9; proteína SEQ ID NO:10), *Yarrowia lipolytica* (codificação SEQ ID NO:11; proteína SEQ ID NO:12) e *Pichia stipitis* (codificação SEQ ID NO:13; proteína SEQ ID NO:14).

[075] De maneira similar, em qualquer uma destas leveduras hospedeiras, um gene ILV3 endógeno pode ser inativado para reduzir a expressão da proteína Fe-S endógena. ILV3 codifica DHAD mitocondrial que está envolvida na biossíntese de aminoácidos de cadeia ramificada. DHAD mitocondrial é codificada por um gene do núcleo, e tem uma sequência sinal

alvo mitocondrial que é transportada e está localizada na mitocôndria. Qualquer gene ILV3 pode ser inativado em uma célula de levedura hospedeira desta divulgação. Exemplos de genes alvo ILV3 de levedura para inativação e suas proteínas codificadas são aqueles de *Saccharomyces cerevisiae* YJM78 (codificação SEQ ID NO:111; proteína SEQ ID NO:112), *Schizosaccharomyces pombe* (codificação SEQ ID NO:93; proteína SEQ ID NO:94), *Candida galbrata* linhagem CBS 138 (codificação SEQ ID NO:107; proteína SEQ ID NO:108), *Candida albicans* SC5314 (codificação SEQ ID NO: 101; proteína SEQ ID NO:102), *Kluyveromyces lactis* (codificação SEQ ID NO:1 13; protein SEQ ID NO:1 14), *Yarrowia lipolytica* (codificação SEQ ID NO:105; proteína SEQ ID NO:106) e *Pichia stipitis* CBS 6054 (codificação SEQ ID NO:103; proteína SEQ ID NO:104).

[076] Devido ao fato dos genes que codificam isopropilmalato desidratases e genes que codificam enzimas DHAD serem bem conhecidos, e devido ao fato da prevalência de sequenciamento genômico, espécies adicionais adequadas destas enzimas podem ser facilmente identificadas por um técnico no assunto, com base na similaridade de sequências usando-se abordagens de bioinformática. Tipicamente o BLAST (descrito acima) realiza uma busca nos bancos de dados disponíveis publicamente com sequências de aminoácidos de isopropilmalato desidratases conhecidas, como aquelas fornecidas no presente pedido, e são usadas para identificar estas enzimas e suas sequências codificadoras que podem ser direcionadas para inativação nas linhagens presentes. Por exemplo, isopropilmalato desidratase endógena de levedura e proteínas DHAD que têm identidade de sequência de aminoácidos de pelo menos cerca de 70-75%, 75%-80%, 80-85%, 85%-90%, 90%- 95% ou 98% com qualquer proteína isopropilmalato desidratase da SEQ ID NOs:2, 4, 6 ,8, 10, 12 e 14, e proteínas DHAD de SEQ ID NOs:94, 102, 104, 106, 20 108, 112 e 114 podem ter expressão reduzida nas presentes

linhagens. Identidades são baseadas no método Clustal W de alinhamento usando-se parâmetros básicos de GAP PENALTY=10, COMPRIMENTO DE GAP =0,1, e série Gonnet 250 de matriz de peso de proteína.

[077] Adicionalmente, as sequências das regiões codificadoras de LEU1 e ILV3 fornecidas no presente pedido podem ser usadas para identificar outros homólogos na natureza. Por exemplo, cada uma das regiões codificadoras descritas no presente pedido pode ser usada para isolar genes que codificam proteínas homólogas. O isolamento de genes homólogos com o uso de protocolos dependentes de sequência é conhecido na técnica. Exemplos de protocolos dependentes de sequência incluem, mas não se limitam a: (1) métodos de hibridização de ácido nucléico; (2) métodos de amplificação de DNA e RNA, conforme exemplificado por vários usos de tecnologias de amplificação de ácido nucléico (por exemplo, reação em cadeia da polimerase (PCR), Mullis *et al.*, patente US 4.683.202; reação em cadeia da ligase (LCR), Tabor, S. *et al.*, *Proc. Acad. Sci. USA* 82:1074 (1985); ou amplificação por deslocamento de fita (SDA), Walker *et al.*, *Proc. Natl. Acad. Sci. U.S.A.*, 89:392 (1992); e (3) métodos para construção e seleção de bibliotecas por complementação.

[078] Por exemplo, genes que codificam proteínas ou polipeptídeos similares a isopropilmalato desidratase e genes que codificam DHAH fornecidos no presente pedido poderiam ser isolados diretamente pelo uso de todo ou uma porção dos atuais fragmentos de ácido nucléico como sondas de hibridização de DNA para selecionar bibliotecas a partir de qualquer organismo desejado, usando-se metodologia conhecida para os técnicos no assunto. Sondas de oligonucleotídeos específicas baseadas nas atuais sequências de ácidos nucléicos podem ser projetadas e sintetizadas por métodos conhecidos na técnica (Maniatis, acima). Além disso, as sequências inteiras podem ser usadas diretamente para sintetizar sondas de DNA pelos

métodos conhecidos pelo técnico no assunto (por exemplo, *primers* aleatórios de DNA marcados, tradução cortada, ou técnicas de marcação de terminação), ou uso de sondas de RNA disponíveis em sistemas de transcrição *in vitro*. Além disso, *primers* específicos podem ser projetados e usados para amplificar uma parte ou as atuais sequências inteiras. Os produtos resultantes da amplificação podem ser marcados diretamente durante as reações de amplificação ou marcados depois das reações de amplificação, e usados como sondas para isolar os fragmentos inteiros de DNA, por hibridização mediante condições adequadas de estringência. Tipicamente, nas técnicas de amplificação tipo PCR, os *primers* têm diferentes sequências e não são complementares entre si. Dependendo das condições de teste desejadas, as sequências dos *primers* poderiam ser projetadas para fornecer replicação eficiente e confiável do ácido nucléico alvo. Métodos de projeto de *primer* de PCR são comuns e bem conhecidos na técnica (Thein e Wallace, "The use of oligonucleotides as specific hybridization probes in the Diagnosis of Genetic Disorders", em *Human Genetic Diseases: A Practical Approach*, K. E. Davis Ed., (1986) páginas 33 a 50, IRL: Herndon, VA; e Rychlik, W., em *Methods in Molecular Biology*, White, B. A. Ed., (1993) Volume 15, páginas 31 a 39, PCR Protocols: Current Methods and Applications. Humania: Totowa, NJ).

[079] Geralmente dois segmentos curtos das sequências descritas podem ser úteis para serem usados nos protocolos da reação em cadeia da polimerase para amplificar fragmentos de ácidos nucléicos que codificam genes homólogos a partir do DNA ou RNA. A reação em cadeia da polimerase também pode ser realizada sobre uma biblioteca de fragmentos clonados de ácido nucléico, em que a sequência de um *primer* é derivada dos fragmentos de ácido nucléico descritos, e a sequência do outro *primer* leva vantagem pela presença da área de ácido poliadenílico para a extremidade 3' do RNAm precursor que codifica os genes microbianos.

[080] Alternativamente, a segunda sequência de *primer* pode estar baseada nas sequências derivadas do vetor de clonagem. Por exemplo, o técnico no assunto pode seguir o protocolo RACE (Frohman *et al.*, PNAS USA 85:8998 (1988)) para gerar cDNAs pelo uso de PCR para amplificar cópias da região entre um único ponto no transcrito e na extremidade 3' ou 5'. Os *primers* orientados nas direções 3' e 5' podem ser designados a partir das presentes sequências. Usando-se os sistemas 3' RACE ou 5' RACE disponíveis comercialmente (por exemplo, BRL, Gaithersburg, MD), fragmentos de cDNA 3' ou 5' específicos podem ser isolados (Ohara *et al.*, PNAS USA 86:5673 (1989); Loh *et al.*, Science 243:217 (1989)).

[081] Alternativamente, as sequências fornecidas que codificam a isopropilmalato desidratase e DHAD podem ser empregadas como reagentes de hibridização para a identificação de homólogos. Os componentes básicos de um teste de hibridização de ácido nucléico incluem uma sonda, uma suposta amostra que contenha o gene ou o fragmento do gene de interesse e um método de hibridização específico. As sondas são tipicamente sequências de ácido nucléico de fita simples, complementares às sequências de ácido nucléico a serem detectadas. As sondas são “hibridizáveis” à sequência de ácido nucléico a ser detectada. O comprimento da sonda pode variar de 5 bases a dezenas de centenas de bases e dependerá do teste específico ser realizado. Normalmente, um comprimento de sonda de cerca de 15 bases a cerca de 30 bases é adequado. A única parte da molécula da sonda precisa ser complementar à sequência de ácido nucléico a ser detectada. Além disso, a complementaridade entre a sonda e a sequência alvo não precisa ser perfeita. A hibridização ocorre entre as moléculas complementares de maneira imperfeita com o resultado que certa fração das bases da região hibridizada não pareia com a base complementar adequada.

[082] Os métodos de hibridização são bem definidos.

Tipicamente a sonda e a amostra devem ser misturadas mediante condições que permitam a hibridização do ácido nucléico. Isso envolve o contato da sonda e da amostra na presente de um sal inorgânico ou orgânica sob a concentração e as condições de temperatura adequadas. Os ácidos nucléicos da sonda e da amostra devem estar em contato durante um longo tempo suficiente que qualquer hibridização possível entre o ácido nucléico da sonda e da amostra pode ocorrer. A concentração da sonda e do alvo na mistura determinará o tempo necessário para a hibridização ocorrer. Quanto maior a concentração da sonda ou do alvo, menor o tempo de incubação necessário. Opcionalmente, um agente caotrópico pode ser adicionado. O agente caotrópico estabiliza dos ácidos nucléicos pela inibição da atividade da nuclease. Além disso, o agente caotrópico permite a hibridização sensível e limitada de sondas de oligonucleotídeo curtas em temperatura ambiente (Van Ness e Chen, *Nucl. Acids Res.* 19:5143-5151 (1991)). Os agentes caotrópicos adequados incluem cloreto de guanidínio, tiocianato de guanidínio, tiocianato de sódio, tetracloroacetato de lítio, perclorato de sódio, tetracloroacetato de rubídio, iodeto de potássio e trifluoroacetato de césio, entre outros. Tipicamente, o agente caotrópico estará presente a uma concentrado final de cerca de 3 M. Se desejado, pode-se adicionar formamida à mistura de hibridização, tipicamente 30 a 50% (v/v).

[083] As várias soluções de hibridização podem estar envolvidas. Tipicamente, isso compreende de cerca de 20 a 60% do volume, preferencialmente 30%, de um solvente orgânico polar. Uma solução de hibridização comum emprega cerca de 30% a 50% v/v de formamida, cloreto de sódio a cerca de 0,15 M a 1 M, tampões a cerca de 0,05 M a 0,1 M (por exemplo, citrato de sódio, Tris-HCl, PIPES ou HEPES (pH na faixa de 6 a 9)), cerca de 0,05% a 0,2% de detergente (por exemplo, dodecilsulfato de sódio), ou EDTA entre 0,5 a 20 mM, FICOLL (Pharmacia Inc.) (cerca de 300 a 500

kdal), polivinilpirrolidona (cerca de 250 a 500 kdal), e albumina do soro. Também será incluído na solução de hibridização típica um carreador de ácidos nucléicos não marcado de cerca de 0,1 a 5 mg/ml, DNA nucléico fragmentado (por exemplo, do timo de bezerro ou DNA de esperma de salmão, ou RNA de levedura), e opcionalmente de cerca de 0,5% a 2% de glicina peso/vol. Outras atividades também podem estar incluídas, como agentes de exclusão de volume que incluem uma variedade de agentes solúveis em água ou agentes de expansão (por exemplo, polietilenoglicol) e polímeros sacarídicos aniônicos (por exemplo, sulfato de dextrans).

[084] A hibridização do ácido nucléico é adaptável a uma variedade de formatos de testes. Um dos mais adequados é o formato de teste sanduíche. O teste sanduíche é particularmente adaptável à hibridização sob condições não desnaturantes. Um composto primário de um teste tipo sanduíche é um suporte sólido. Um suporte sólido tem adsorvido a ele ou ligado covalentemente a ele uma sonda de ácido nucléico imobilizado que não é marcada e complementar a uma porção da sequência.

[085] Sequências de ácido nucleico que codificam proteínas para qualquer uma das proteínas, exceto Fe-S, que podem ser direcionadas para atividade reduzida em uma célula de levedura da invenção podem ser identificadas usando-se bioinformática e outros métodos conhecidos pelos técnicos no assunto. Por exemplo, sequências de aconitase são identificadas por busca de palavra-chave nos bancos de dados de bioinformática. Sequências diversas identificadas por esse método são aquelas de *Saccharomyces cerevisiae* (codificação SEQ ID NO:153; proteína SEQ ID NO:154), *Schizosaccharomyces pombe* no cromossomo II (codificação SEQ ID NO:155; proteína SEQ ID NO: 156), *Schizosaccharomyces pombe* no cromossomo I (codificação SEQ ID NO:157; proteína SEQ ID NO:158), *Kluyveromyces lactis* (codificação SEQ ID NO:15; proteína SEQ ID NO:160),

*Candida albicans* SC5314 (codificação SEQ ID NO:161; proteína SEQ ID NO:162), *Yarrowia lipolytica* (codificação SEQ ID NO:163; proteína SEQ ID NO:164), *Pichia stipitis* CBS 6054 (codificação SEQ ID NO:165; proteína SEQ ID 25 NO:166), *Candida galbrata* CBS 138 cromossomo F (codificação SEQ ID NO:167; proteína SEQ ID NO:168), *Candida galbrata* CBS 138 cromossomo D (codificação SEQ ID NO:169; proteína SEQ ID NO:170), e *Candida galbrata* CBS 138 cromossomo K (codificação SEQ ID NO:171; proteína SEQ ID NO:172).

[086] Os genes que codificam proteínas Fe-S, por exemplo, LEU1, ILV3 ou ACO1 podem ser rompidos em qualquer célula de levedura usando-se modificação genética. Muitos métodos para modificação genética dos genes alvo são conhecidos por um técnico no assunto e podem ser usados para criar as presentes linhagens de levedura. Modificações que podem ser usadas para reduzir ou eliminar a expressão de uma proteína alvo são rupturas que incluem, mas não se limitam a, deleção do gene inteiro ou uma porção do gene, inserção de um fragmento de DNA no gene (tanto no promotor como na região codificadora), de modo que a proteína não seja expressa, ou seja expressa em níveis menores, introdução de uma mutação na região codificadora que adiciona um códon de parada ou substituição da estrutura, de modo que uma proteína funcional não seja expressa, e introdução de uma ou mais mutações na região codificadora para alterar aminoácidos, de modo que a proteína ativa não funcional ou menos ativa enzimaticamente seja expressa. Além disso, a expressão de um gene pode ser bloqueada pela expressão de um RNA *antisense* ou um RNA de interreferência, e constructos podem ser introduzidos para resultar na co-supressão. Além disso, a síntese ou estabilidade do transcrito pode ser diminuída pela mutação. Da mesma forma a eficiência pela qual uma proteína é traduzida a partir do RNAm, pode ser modulada por mutação. Todos esses métodos podem ser facilmente praticados

por um técnico no assunto, fazendo o uso de sequências codificadoras conhecidas ou identificadas, como LEU1 ou ILV3.

[087] As sequências de DNA próximas a uma sequência codificadora LEU1, ILV3, ou ACO1 também são úteis em alguns procedimentos de modificação e estão disponíveis para leveduras, como *Saccharomyces cerevisiae* na sequência completa do genoma coordenado pelo Projeto Genoma ID9518 dos Projetos Genomas coordenados pelo NCBI (National Center for Biotechnology Information) com identificação GOPID 13838. Exemplos adicionais de sequências genômicas de levedura incluem aquelas de *Yarrowia lipolytica*, GOPIC 13837, e de *Candida albicans*, que são incluídas em GPID 10771, 10701 e 16373. Genomas adicionais forma completamente sequenciados e anotados e estão publicamente disponíveis para as linhagens de levedura a seguir : *Candida glabrata* CBS 138, *Kluyveromyces lactis* NRRL Y-1 140, *Pichia stipitis* CBS 6054 e *Schizosaccharomyces pombe* 972h-.

[088] Em particular, sequências de DNA próximas a uma sequência codificadora alvo, como LEU1 ou ILV3, são úteis para métodos de modificação que utilizam recombinação de homólogos. Por exemplo, neste método as sequências adjacentes são colocadas nos limites de um gene marcador selecionável para mediar a recombinação de homólogos, assim o gene marcador substitui o gene alvo. Também, sequências genéticas alvo parciais e sequências adjacentes limítrofes a um gene marcador selecionável podem ser usadas para mediar recombinação homóloga, assim o gene marcador substitui uma porção do gene alvo. Além disso, o marcador selecionável pode ser ligado por sítios de recombinação sítio-específicos, de modo que segue a expressão da recombinase específica para o sítio correspondente, o gene para resistência é retirado do gene alvo sem reativação posterior. A recombinação sítio-específica deixa para trás um sítio de recombinação que rompe a expressão da proteína codificada pelo gene

alvo. O vetor de recombinação homólogo pode ser construído também para permitir uma deleção no gene alvo, seguindo retirada do marcador selecionável, como é conhecido por um técnico no assunto.

[089] As deleções podem ser feitas usando-se recombinação mitótica, conforme descrito em Wach *et al.* ((1994) *Yeast* 10:1793-1808). Este método envolve a preparação de um fragmento de DNA que contém um marcador selecionável entre as regiões genômicas que podem ser tão curtas quanto 20 pares de bases, e que se ligam a uma sequência de DNA alvo. O fragmento de DNA pode ser preparado por amplificação por PCR do gene marcador selecionável usando-se *primers* de oligonucleotídeos que hibridizam com as extremidades do gene marcador e que incluem regiões genômicas que podem recombinar com o genoma de levedura. O fragmento de DNA linear pode ser eficientemente transformado nas leveduras e recombinado no genoma, resultando na substituição do gene, incluindo deleção da sequência alvo de DNA (conforme descrito em *Methods in Enzymology*, volume 194, páginas 281 a 301 (1991)).

[090] Além disso, métodos para substituição do promotor podem ser usados para trocar os elementos de controle transcricionais endógenos que permitem outros meios para modular a expressão, conforme descrito em Mnaimneh *et al.* ((2004) *Cell* 118(1):31-44) e no Exemplo 12 no presente pedido.

[091] Além disso, um gene alvo em qualquer célula de levedura pode ser rompido usando-se mutagênese aleatória, que é seguido por seleção para identificar linhagens com atividade de codificação reduzida do gene alvo. Usando-se este tipo de método, a sequência de DNA de, por exemplo, LEU1, ILV3, ou qualquer outra região do genoma que afeta a expressão de uma proteína Fe-S alvo, precisa ser conhecida.

[092] Métodos para criar mutações genéticas são comuns e

conhecidos na técnica, e podem ser aplicados para praticar a criação de mutantes. Métodos para modificações genéticas aleatórias comumente utilizados (revisados em *Methods in Yeast Genetics*, 2005, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY) incluem mutagênese espontânea, mutagênese causada por genes mutantes, mutagênese química, irradiação com raios UV ou raios X, ou mutagênese transposon.

[093] A mutagênese química de leveduras comumente envolve tratamento das células de levedura com um dos DNAs mutagênicos a seguir: etil metanosulfonato (EMS), ácido nitroso, dietil sulfato, ou N-metil-N'-nitro N-nitroso-guanidina (MNNG). Esses métodos de mutagênese foram revisados em Spencer *et al* (*Mutagenesis in Yeast*, 1996, Yeast Protocols: *Methods in Cell and Molecular Biology*. Humana Press, Totowa, NJ). Mutagênese química com EMS pode ser realizada conforme descrito em *Methods in Yeast Genetics*, 2005, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. A irradiação com luz ultravioleta (UV) ou raio X também pode ser usada para produzir mutagênese aleatória nas células de levedura. O principal efeito da mutagênese por irradiação UV é a formação de dímeros de pirimidina que rompem a fidelidade de replicação do DNA. Protocolos para mutagênese UV de leveduras podem ser encontrados em Spencer *et al* (*Mutagenesis in Yeast*, 1996, Yeast Protocols: *Methods in Cell and Molecular Biology*. Humana Press, Totowa, NJ). A introdução de um fenótipo mutante também pode ser usado para gerar mutações cromossômicas aleatórias em leveduras. Os fenótipos mutantes comuns podem ser obtidos através do rompimento de um ou mais dos seguintes genes: PMS1, MAG1, RAD18 ou RAD51. O restabelecimento do fenótipo não-mutante pode ser facilmente obtido por inserção do alelo tipo selvagem. Coleções de células modificadas produzidas a partir de qualquer um desses processos de mutagênese aleatória conhecidos podem ser selecionados

para atividade reduzida da proteína Fe-S.

### **PROTEÍNAS FE-S HETERÓLOGAS**

[094] Qualquer ácido dihidroxi desidratase do cluster 2Fe-2S fúngica ou vegetal ou qualquer propanodiol desidratase reativase pode ser expresso como proteína heteróloga em uma célula hospedeira de levedura elaborada geneticamente, conforme divulgado no presente pedido para expressão de proteína endógena do cluster Fe-S, e o aumento da atividade pode ser obtido. Uma proteína heteróloga inclui aquela que é expressa de uma maneira diferente da expressão de uma proteína endógena correspondente. Por exemplo, DHAD endógena é codificada por ILV3 no núcleo e a proteína DHAD expressa tem uma sequência sinal de direcionamento mitocondrial, de modo que a proteína seja localizada na mitocôndria. Um cluster Fe-S é adicionado à proteína DHAD na mitocôndria para sua atividade na biossíntese de aminoácido de cadeia ramificada. É desejável expressar atividade DHAD no citosol para participação nas vias biossintéticas que estão localizadas no citosol. A expressão no citosol de DHAD em leveduras é expressão heteróloga, visto que a proteína nativa está localizada na mitocôndria. Por exemplo, a expressão heteróloga de DHAD de *Saccharomyces cerevisiae* em *S. cerevisiae* é obtida pela expressão da região codificadora de DHAD em *S. cerevisiae* com o sinal de direcionamento mitocondrial removido, de modo que a proteína permaneça no citosol. DHADs 2Fe-2S que podem ser usadas na presente divulgação incluem aquelas de fungos e plantas. DHADs 2Fe-2S representativas de fungos ou plantas estão listadas nas Tabelas 1 e 2. DHADs 2Fe-2S de fungos ou plantas com identidades de sequência de aminoácidos de 95% ou mais foram removidas da análise que fornece esta lista para simplificação. Entretanto, qualquer sequência com 95% ou mais de identidade de sequência de aminoácido para qualquer uma dessas sequências são úteis na presente invenção. A análise usada para se obter DHADs 2Fe-2S está

descrita no pedido de patente co-pendente US 61/100792 de propriedade comum, que é incorporado ao presente pedido como referência. A análise é como segue: Aqui um Perfil e Modelos Ocultos de Markov (HMM) foi preparado com base nas sequências de aminoácidos de oito DHADs verificadas funcionalmente. Essas DHADs são de *Nitrosomonas europaea* (DNA SEQ ID NO:174; Proteína SEQ ID NO:175), *Synechocystis sp.* PCC6803 (DNA SEQ ID:176; Proteína SEQ ID NO:177), *Streptococcus mutans* (DNA SEQ ID NO:178; Proteína SEQ ID NO:179), *Streptococcus thermophilus* (DNA SEQ ID NO:180; Proteína SEQ ID No:181), *Ralstonia metallidurans* (DNA SEQ ID NO:182; Proteína SEQ ID NO:183), *Ralstonia eutropha* (DNA SEQ ID NO:184; Proteína SEQ ID NO:185), e *Lactococcus lactis* (DNA SEQ ID NO:186; Proteína SEQ ID NO:187). Além da DHAD de *Flavobacterium johnsoniae* (DNA SEQ ID NO:188; proteína SEQ ID NO:189) descobriu-se que tem atividade de ácido dihidroxi desidratase quando expressa em *E. coli* e foi usada para fazer o Perfil. O Perfil HMM é preparado usando-se o pacote de programa de computação HMMER (A teoria para este perfil HMMs está descrita em R. Durbin, S. Eddy, A. Krogh, e G. Mitchison, *Biological sequence analysis: probabilistic models of proteins and nucleic acids*, Cambridge University Press, 1998; Krogh *et al.*, 1994; *J. Mol. Biol.* 235:1501-1531), seguindo o manual do usuário que está disponível junto a HMMER (Janelia Farm Research Campus, Ashburn, VA). A saída de dados do programa de computação HMMER é um Perfil de Modelos Ocultos de Markov (HMM) que caracteriza as sequências de entrada, fornecidas na Tabela 9.

[095] Qualquer proteína compatível com o Perfil HMM com um valor E de  $< 10^{-5}$  é uma proteína relacionada com DHAD, que inclui DHADs 4Fe-4S, DHADs 2Fe-2S, arabinato desidratases e fosfogluconato desidratases. As sequências que são compatíveis com o Perfil HMM são então analisadas para a presença de três cisteínas conservadas, que correspondem

às posições 56, 129 e 201 nos mutantes DHAD de *Streptococcus mutans*. A presença de todas as três cisteínas conservadas é característico de proteínas que têm um cluster [2Fe-2S] 2+. Proteínas que têm três cisteínas conservadas incluem arabinato desidratases e DHADs 2Fe-2S. DHADs 2Fe-2S podem ser diferenciadas das arabinato desidratases pela análise dos aminoácidos conservados para assinatura presentes nas DHADs 2Fe-2S ou nas arabinato desidratases nas posições correspondentes na sequência de aminoácido mutante para DHAD de *Streptococcus mutans*. Esses aminoácidos de assinatura estão nas DHADs 2Fe-2S ou nas arabinato desidratases, respectivamente, nas posições a seguir (com mais que 90% de ocorrência): 88 asparagina vs ácido glutâmico; 113 não conservado vs ácido glutâmico; 142 arginina ou asparagina vs não conservado; 165: não conservado vs glicina; 208 asparagina vs não conservado ; 454 leucina vs não conservado; 477 fenilalanina ou tirosina vs não conservado; e 487 glicina vs não conservado.

[096] As proteínas identificadas por esse processo que tem uma origem fúngica ou vegetal, como SEQ ID NOs:46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150 e 152 podem ser usadas na presente invenção, bem como qualquer proteína com identidade de aminoácido de pelo menos 95%, 96%, 97%, 98%, ou 99% para qualquer uma dessas sequências. Particularmente adequada é a DHAD de *Kluyveromyces lactis* (SEQ ID NO:1 14) e DHADs com pelo menos cerca de 90% de identidade de sequência de aminoácido para SEQ ID NO:1 14 que usa o método Clustal W de alinhamento que utiliza os parâmetros padrão de GAP PENALTY=10, GAP LENGTH PENALTY=0,1, e série Gonnet 250 de matriz de peso de proteína sobre o comprimento total da sequência de proteína.

[097] Além disso, DHADs 2Fe-2S fúngicas ou vegetais que

podem ser usadas na presente invenção podem ser identificadas pelas suas posições em uma ramificação DHADs 2Fe-2S fúngica ou vegetal de uma árvore filogenética de proteínas relacionadas com DHAD. Além disso, DHADs 2Fe-2S que podem ser usadas podem ser identificadas usando-se as sequências de comparação com qualquer umas das DHADs 2Fe-2S fúngicas ou vegetais cujas sequências são fornecidas no presente pedido, cuja identidade de sequência pode ser de pelo menos 80%-85%, 85%-90%, 90%-95% or 95%-99%.

[098] Adicionalmente, as sequências das das DHADs 2Fe-2S fúngicas ou vegetais fornecidas no presente pedido podem ser usadas para identificar outros homólogos na natureza. Por exemplo, cada um dos fragmentos de ácido nucléico que codifica DHAD fornecidos no presente pedido como SEQ ID NOs:45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 11, 121, 123, 125, 127, 129, 131, 133, 135, 137, 13, 141, 143, 145, 147, 149 e 151 pode ser usado para isolar genes que codidicam proteínas homólogas conforme descrito acima para a região codificadora LEU1.

[099] A coenzima propanodiol desidratase reativase independente de B12 de *Roseburia inulinivorans* é uma proteína que necessita de um cluster Fe-S para sua atividade. Esta proteína, RdhtB, é divulgada no pedido de patente US co-pendente US20090155870, que é incorporado ao presente pedido como referênciia. RdhtB reativa uma coenzima propanodiol desidratase reativase independente de B12 de *Roseburia inulinivorans*, que é denominada RdhtA e também é divulgada na patente co-pendente de propriedade comum US 20090155870. A atividade de RdhtB pode ser avaliada pelo teste da atividade de RdhtA, visto que RdhtB é necessária para atividade de RdhtA. A atividade de RdhtB, e portanto de RdhtA, é melhorada pela expressão em uma levedura hospedeira com expressão da proteína Fe-S

endógena reduzida divulgada no presente pedido. A expressão heteróloga de qualquer coenzima propanodiol desidratase reativase independente de B12 que necessita de um cluster Fe-S para atividade pode ser aprimorada em uma linhagem de levedura que tem expressão da proteína Fe-S endógena reduzida. Uma coenzima propanodiol desidratase reativase independente de B12 pode ser facilmente identificada por um técnico no assunto pela avaliação da atividade da enzima associada a propanodiol desidratase na presença ou ausência da coenzima B12. Um exemplo de uma diol desidratase reativase de *Clostridium butyricum* (região codificadora SEQ ID NO:192; proteína SEQ ID NO:193).

[0100] Outra coenzima propanodiol desidratase reativase independente de B12 que pode ser usada pode ser identificada através da análise de bioinformática de sequências, conforme comparada àquelas de RdhtB SEQ ID NO:44 por um técnico no assunto. Proteínas com coenzima propanodiol desidratase reativase independente de B12 e identidade de sequência para SEQ ID NO:44 de pelo menos 80% a 85%, 85% a 90%, 90% a 95% ou 95% a 99% podem ser usadas. Particularmente adequadas são essas que são pelo menos cerca de 90% idênticas à sequência de aminoácido conforme apresentada na SEQ ID NO:44 com o uso do método Clustal W de alinhamento que utiliza os parâmetros padrão de PENALIDADE PARA GAP = 10, PENALIDADE PARA COMPRIMENTO DE GAP = 0,1, e série Gonnet 250 de matriz de peso de proteína sobre o comprimento total da sequência de proteína.

[0101] Além disso, outra coenzima propanodiol desidratase reativase independente de B12 homóloga que pode ser usada pode ser identificada usando-se a região codificadora de RdhtB (SEQ ID NO:16) por métodos conforme descrito acima para a região codificadora de LEU1.

#### **EXPRESSÃO DE PROTEÍNAS FE-S HETERÓLOGAS**

[0102] A expressão é alcançada pela transformação com uma

sequência que codifica uma proteína Fe-S. A região codificadora a ser expressa pode ser códon-otimizada para a célula alvo hospedeira, como bem conhecido por um técnico no assunto. Métodos para expressão gênica em leveduras são conhecidos na técnica (consulte, por exemplo, *Methods in Enzymology*, Volume 194, *Guide to Yeast Genetics and Molecular and Cell Biology* (Part A, 2004, Christine Guthrie e Gerald R. Fink (Eds.), Elsevier Academic Press, San Diego, CA). A expressão dos genes nas leveduras tipicamente necessita de um promotor, ligado de modo funcional a uma região codificadora de interesse, e um terminador de transcrição. Diversos promotores de leveduras podem ser usados na construção de cassetes de expressão para genes nas leveduras, incluindo, mas não se limitando a promotores derivados dos genes a seguir: CYCI, HIS3, GAL1, GAL10, ADHI, PGK, PHO5, GAPDH, ADCI, TRPI, URA3, LEU2, ENO, TPI, CUPI, FBA, GPD, GPM e AOXI. Terminadores de transcrição adequados incluem, mas não se limitam a FBAt, GPMt, ERG10t, GAL1t, CYC1 e ADH.

[0103] Promotores adequados, terminadores de transcrição e regiões codificadoras podem ser clonadas nos vetores de transporte de leveduras em *E.coli*, e transformados nas células de leveduras. Esses vetores permitem a propagação da linhagem em *E. coli* e nas linhagens de levedura. Tipicamente, o vetor usado contém um marcador selecionável e as sequências que permitem a replicação autônoma ou a integração cromossômica do hospedeiro desejado. Tipicamente, os plasmídios usados nesta levedura foram os vetores de transporte pRS423, pRS424, pRS425 e pRS426 (Coleção Americana de Tipo de Cultura, Rockville, MD), que continham uma origem de replicação de *E. coli* (por exemplo, pMB1), uma origem de replicação levedura 2 $\mu$  de levedura e um marcador para seleção nutricional. Os marcadores de seleção para esses quatro vetores são His3 (vetor pRS423), Trp1 (cetor pRS424), Leu2 (vetor pRS425) e Ura3 (vetor pRS426). A construção dos

vetores de expressão com um gene quimérico que codifica as regiões codificadoras da proteína Fe-S pode ser realizada pelas técnicas de clonagem molecular padrão em *E. coli* ou pelo método de recombinação de reparo de Gap em leveduras.

[0104] A abordagem da clonagem de reparo de intervalo leva a vantagem da recombinação de homólogo altamente eficiente na levedura. Tipicamente, um DNA do vetor de levedura é digerido (por exemplo, neste sítio de clonagem múltipla) para criar um “gap” na sua sequência. Vários insertos de DNA de interesse são gerados para conter uma sequência maior ou igual a 21 pares de bases nas extremidades 5’ e 3’, que sequencialmente se sobrepõem uma a outra, e com as extremidades 5’ e 3’ do DNA do vetor. Por exemplo, para construir um vetor de expressão de levedura para o “Gene X, um promotor de levedura e um terminador de levedura são selecionados para o cassete de expressão. O promotor e o terminador são amplificados a partir do DNA genômico da levedura e o Gene X é amplificado em PCR a partir do organismo fonte ou obtido da clonagem do vetor que compreende a sequência do Gene X. Há, pelo menos, uma sequência de sobreposição de 21 pares de bases entre a extremidade 5’ do vetor alinhado e a sequência do promotor, entre o promotor e o Gene X, entre o Gene X e a sequência terminadora, e entre o terminador e a extremidade 3’ do vetor alinhado. O vetor com "gap" e os insertos de DNAs são então co-transfetados em uma linhagem de levedura e plaqueados no meio contendo misturas com compostos apropriados que permitem complementação dos marcadores de seleção nutricional nos plasmídios. A presença das combinações com insertos corretos podem ser confirmadas por mapeamento com PCr usando-se DNA plasmidial preparado a partir das células selecionadas. O DNA plasmidial isolado da levedura (geralmente, em concentração baixa) pode então ser transformado em uma linhagem de *E. coli*, por exemplo *TOP10*, seguido por mini preparados e pelo mapeamento de

restrição para verificar, ainda, o constructo do plasmídio. Finalmente, o constructo pode ser verificado pela análise da sequência.

[0105] Similar à técnica de reparo de gap, a integração no genoma de levedura também apresenta vantagem sobre o sistema de recombinação de homólogos em levedura. Tipicamente, um cassete contendo uma região codificadora mais elementos de controle (promotor e terminador) e marcador auxotrófico é amplificado por PCR com uma DNA polimerase de alta finalidade usando-se *primers* que hibridizam ao cassete e contém de 40 a 70 pares de bases de sequência homóloga às regiões 5' e 3' da área genômica, onde a inserção é desejada. O produto de PCR é então transformado nas leveduras e plaqueados no meio contendo a mistura com compostos apropriados que permitem seleção para o marcador auxotrófico integrado. Por exemplo, para integrar o "Gene X" na localização "Y" do cromossomo, o promotor-região codificadora X-constructo terminador é amplificado por PCR a partir de um constructo de DNA plasmidial e unido a um marcador autotrófico (como URA3), tanto por PCR SOE como por digestão por restrição comum ou clonagem. O cassete completo, contendo promotor-região codificadora X-terminador-região URA3, é amplificado por PCR com sequências *primer* que contém de 40 a 70 pares de bases de homologia para as regiões 5' e 3' de localização "Y" no cromossomo de levedura. O produto de PCR é transformado nas leveduras e selecionado no meio de crescimento sem uracila. Os transformantes podem ser verificados tanto por PCR de colônia como por sequenciamento direto de DNA cromossômico.

[0106] Qualquer região codificadora expressa nas presentes células de leveduras pode ser códon-otimizada para expressão na célula de levedura hospedeira específica sendo elaborada geneticamente, como bem conhecido por um técnico no assunto. Por exemplo, para expressão das regiões codificadoras ILV3 de *K. lactis* e *P. stipitis* em *S. cerevisiae*, cada uma

foi códon-otimizada para expressão em *S. cerevisiae* no Exemplo 1 no presente pedido.

### **BIOSSÍNTESSE DO PRODUTO COM ATIVIDADE DE PROTEÍNA Fe-S HETERÓLOGA**

#### **APRIMORADA**

[0107] A produção de qualquer produto que tem uma proteína Fe-S que contribui para sua via biossintética pode se beneficiar da atividade aprimorada divulgada no presente pedido de uma proteína Fe-S heteróloga expressa em uma levedura hospedeira com expressão de proteína Fe-S endógena reduzida. Por exemplo, DHAD fornece uma etapa nas vias para biossíntese de isobutanol, e RdhtB contribui em uma via biossintética para produzir 2-butanona ou 2-butanol.

[0108] As vias biossintéticas que incluem uma etapa realizada por DHAD para síntese de isobutanol são divulgadas no pedido de patente co-pendente e de uso US 20070092957 A1, que é integralmente incorporado ao presente pedido como referência. Um diagrama das vias biossintéticas do isobutanol divulgadas são fornecidas na Figura 1. A produção do isobutanol em uma linhagem divulgada no presente pedido se beneficia da atividade aumentada de DHAD. Conforme descrito no pedido de patente US 20070092957 A1, as etapas em um exemplo da via biossintética do isobutanol incluem a conversão de:

piruvato para acetolactato (Figura 1, etapa a da via) conforme catalisado, por exemplo, por acetolactato sintase;

acetolactato para 2,3-dihidroxi isovalerato (Figura 1, etapa b da via) conforme catalisado, por exemplo, por acetolactato isomeroreductase;

2,3-dihidroxi isovalerate para  $\alpha$ -cetoisovalerato (Figura 1, etapa c da via) conforme catalisado, por exemplo, por ácido acetohidroxi desidratase, chamada DHAD;

$\alpha$ -cetoisovalerato para isobutiraldeído (Figura 1, etapa d da via),

conforme catalisado, por exemplo, por  $\alpha$ -ceto ácido decarboxilase de cadeia ramificada; e

isobutiraldeído para isobutanol (Figura 1, etapa e da via), conforme catalisado, por exemplo, por álcool desidrogenase de cadeia ramificada.

[0109] O substrato para conversão de produtos e enzimas envolvidas nessas reações para as etapas f, g, h, I, j e k de vias alternativas são descritas na patente US 20070092957 A1.

[0110] Genes que podem ser usados para expressão da enzima para as vias do isobutanol são descritos na patente US 20070092957 A1, e genes adicionais que podem ser usados podem ser identificados por um técnico no assunto através de bioinformática ou experimentalmente, conforme descrito acima. O uso preferencial em todas as três vias de enzimas ceto-ácido redutoisomerase (KART) particularmente com alta atividade são divulgadas na patente co-pendente e de propriedade comum US 20080261230. Exemplos de KARIs com alta atividade divulgadas no presente pedido são aquelas de *Vibrio cholerae* (DNA: SEQ ID NO:35; proteína SEQ ID NO:36), *Pseudomonas aeruginosa* PAO1, (DNA: SEQ ID NO:37; proteína SEQ ID NO:38), e *Pseudomonas fluorescens* PF5 (DNA: SEQ ID NO:39; proteína SEQ ID NO:40).

[0111] Adicionalmente, são descritas na patente US 20070092957 A1 construções de genes quiméricos e genes elaborados geneticamente de leveduras, exemplificados por *Saccharomyces cerevisiae*, para produção de isobutanol usando-se as vias biossintéticas divulgadas.

[0112] Uma via biossintética que inclui propanodiol desidratase para síntese de 2-butanona e 2-butanol é divulgada no pedido de patente co-pendente e de propriedade comum US 20070292927A1, que é integralmente incorporado ao presente pedido como referência. Um diagrama da via biossintética de 2-butanona e 2 butanol é fornecida na Figura 2. 2-Butanona é o

produto feito quando a última etapa representada de conversão de 2-butanona para 2-butanol é omitida. A produção de 2-butanona ou 2-butanol em uma linhagem no presente pedido se beneficia da atividade aumentada de propanodiol desidratase reativase independente de B12. Conforme descrito na publicação de pedido de patente US 20070292927A1, as etapas na via biossintética incluem a conversão de:

piruvato para acetolactato (Figura 2, etapa a) conforme catalisado, por exemplo, por acetolactato sintase;

acetolactato para acetoína (Figura 2, etapa b) conforme catalisado, por exemplo, por acetolactato descarboxilase;

acetoína para 2,3-butanodiol (Figura 2, etapa i) conforme catalisado, por exemplo, por butanodiol desidrogenase;

2,3-butanodiol para 2-butanona (Figura 2, etapa j), conforme catalisado, por exemplo, por diol desidratase glicerol desidratase ou propanodiol desidratase; e

2-butanona para 2-butanol (Figura 2, etapa f) conforme catalisado, por exemplo, por butanol desidrogenase.

[0113] Genes que podem ser usados para expressão dessas enzimas são descritos na patente US 20070292927A1. O uso nesta via em leveduras da butanodiol desidratase de *Roseburia inulinivorans*, RdhtA, (proteína SEQ ID NO:43, região codificadora SEQ ID NO:15) é divulgado na patente co-pendente e de propriedade comum US 20090155870. Esta enzima é usada junto com a butanodiol desidratase reativase de *Roseburia inulinivorans*, RdhtB, (proteína SEQ ID NO:44, região codificadora SEQ ID NO:16). Esta butanodiol desidratase é desejada em muitos hospedeiros pois não necessita de coenzima B<sub>12</sub>.

[0114] Adicionalmente, é descrita na publicação de patente US 20090155870 a construção de genes quiméricos e genes elaborados

geneticamente de leveduras usando-se a via biossintética divulgada na patente US 20070292927A1.

#### **MEIO DE FERMENTAÇÃO**

[0115] Leveduras divulgadas no presente pedido podem ser cultivadas no meio de fermentação para produção de um produto que tem proteína Fe-S como parte da via biossintética. Meios de fermentação precisam conter substratos de carbono adequados. Os substratos adequados podem incluir, mas não se limitam a, monossacarídeos como glicose e frutose, oligossacarídeos como lactose e sacarose, polissacarídeos como amido ou celulose, ou a mistura desses e misturas não purificadas de matérias-primas renováveis como soro de queijo introduzido, bebida alcoólica de excesso de milho, melaço de açúcar de beterraba e malte de cevada. Adicionalmente, o substrato de carbono também pode ser substratos de um carbono, como dióxido de carbono ou metanol para o qual a conversão metabólica em intermediários bioquímicos essenciais foi demonstrada. Além de um ou dois organismos metilotróficos para substratos de carbono também são conhecidos por utilizar vários outros compostos contendo carbono como metilamina, glicosamina e uma variedade de aminoácidos para atividade metabólica. Por exemplo, leveduras metilotróficas são conhecidas por utilizar o carbono de metilamina para formar trealose ou glicerol (Bellion *et al.*, *Microb. Growth C1 Compd.*, [Simpósio Internacional.], 7º (1993), 415-32. Editor(es): Murrell, J. Collin; Kelly, Don P. Publisher: Intercept, Andover, Reino Unido). De maneira similar, várias espécies de *Candida* irão metabolizar alanina ou ácido oléico (Sulter *et al.*, *Arch. Microbiol.* 153:485-489 (1990)). Então, é contemplado que a fonte de carbono utilizada na presente invenção pode abranger uma ampla variedade de carbono que contém substratos e apenas será limitado pela escolha do organismo.

[0116] Embora seja considerado que todos os substratos de

carbono mencionados acima e as misturas desses sejam adequadas na presente invenção, os substratos de carbono preferidos são glicose, frutose e sacarose.

[0117] Além de uma fonte de carbono apropriada, o meio de fermentação deve conter minerais, sais, co-fatores, tampões e outros compostos adequados, conhecidos pelos técnicos no assunto, apropriados para o crescimento das culturas e para a promoção da via enzimática necessária para a produção do produto desejado.

#### CONDIÇÕES DE CULTURA

[0118] Tipicamente, as células são cultivadas em uma temperatura que varia de cerca de 20 °C a cerca de 37 °C em um médio apropriado. Meios de crescimento adequados na presente invenção são meios comuns preparados comercialmente, como caldos que incluem base de nitrogênio para leveduras, sulfato de amônia e dextrose (como fonte de carbono/energia) ou Meio YPD, uma mistura de peptona, extrato de levedura e dextrose em proporções ótimas para o crescimento da maioria das linhagens de *Saccharomyces cerevisiae*. Outro meio definido ou sintético pode também ser usado e o meio apropriado para crescimento do microorganismo específico será conhecido pelo técnico no assunto de microbiologia ou ciência da fermentação.

[0119] As faixas adequadas de pH para a fermentação estão entre pH 3,0 a pH 7,5, sendo que o pH 4,5 a pH 6,5 é preferencial como condição inicial.

[0120] As fermentações podem ser realizadas sob condições aeróbias e anaeróbias, sendo que as condições anaeróbias e microaeróbias são preferenciais.

[0121] A quantidade de butanol produzida no meio de fermentação pode ser determinada usando-se diversos métodos conhecidos na

técnica, por exemplo, cromatografia líquida de alta performance (HPLC) ou cromatografia gasosa (CG).

#### **BATELADA INDUSTRIAL E FERMENTAÇÕES CONTÍNUAS**

[0122] O presente processo emprega um método de fermentação de batelada. Uma fermentação de batelada clássica é um sistema fechado, onde a composição do meio é o conjunto no inicio da fermentação e não é sujeito a alterações artificiais durante a fermentação. Assim, no começo da fermentação o meio é inoculado com o organismo ou organismo(s) desejado(s) e permiti-se que a fermentação ocorra sem adicionar nada ao sistema. Tipicamente, entretanto, uma fermentação em “batelada” é uma batelada com respeito à adição de fonte de carbono e as tentativas são frequentemente realizadas em fatores de controle, como pH e concentração de oxigênio. Em sistemas de batelada, as composições de metabolismo e biomassa do sistema mudam constantemente até o momento que a fermentação é interrompida. Dentro das culturas em batelada, as células moderam através de uma fase de intervalo estático a uma fase de alto crescimento logarítmico e finalmente uma fase estacionária, na qual a taxa de crescimento é diminuída ou interrompida. Se não tratadas, as células na fase estacionária eventualmente morrerão. As células na fase logarítmica são responsáveis pelo volume da produção do produto final ou intermediário.

[0123] Uma variação no sistema de batelada padrão é o sistema de batelada alimentada. Os processos de fermentação de batelada alimentada também são adequados na presente invenção e compreendem um sistema de batelada típico, com exceção de que o substrato é adicionado de forma crescente enquanto a fermentação progride. Os sistemas de batelada alimentada são úteis quando a repressão do catabólito é apta para inibir o metabolismo das células e onde é desejável ter quantidades limitadas de substrato nos meios. A medição da concentração de substrato presente no

sistema de batelada alimentada é difícil e, portanto, é estimada com base nas alterações dos fatores de medição como pH, oxigênio dissolvido e a pressão parcial de uso dos gases, como CO<sub>2</sub>. Fermentações de batelada e de batelada alimentada são comuns e bem conhecidos na técnica e os exemplos podem ser encontrados em Thomas D. Brock em *Biotechnology: A Textbook of Industrial Microbiology*, Segunda Edição (1989) *Sinauer Associates, Inc., Sunderland, MA*. ou em *Deshpande, Mukund V., Appl. Biochem. Biotechnol.*, 36:227, (1992), incorporados ao presente pedido como referência.

[0124] Embora a presente invenção seja realizada no modo de batelada, está contemplado que o método seria adaptável para métodos de fermentação contínua. A fermentação contínua é um sistema aberto, no qual um meio de fermentação definido é adicionado continuamente a um biorreator e uma quantidade igual de meios condicionados é removida simultaneamente para processamento. A fermentação contínua geralmente mantém as culturas em uma alta densidade constante, na qual as células estão principalmente em crescimento de escala logarítmica.

[0125] A fermentação contínua permite que a modulação de um fator ou quaisquer números de fatores que afetam o crescimento celular ou a concentração do produto final. Por exemplo, um método manterá um nutriente limitante, como a fonte de carbono ou o nível de nitrogênio em uma taxa fixa e permite que todos os outros parâmetros sejam moderados. Em outros sistemas, diversos fatores que afetam o crescimento podem ser continuamente alterados enquanto a concentração celular, medida pela turbidez do meio, é mantida constante. Os sistemas contínuos se esforçam para manter as condições de crescimento em estado de equilíbrio e, assim, a perda celular devido ao meio sendo extraído deve ser equilibrada em relação à taxa de crescimento celular na fermentação. Os métodos de modulação de nutrientes e os fatores de crescimento para os processos de fermentação contínua bem

como as técnicas para maximizar a taxa de formação do produto são bem conhecidos na técnica de microbiologia industrial e uma variedade de métodos é detalhada por Brock, acima.

[0126] A presente invenção pode ser praticada usando-se tanto os processos de batelada, batelada alimentada como contínuo e métodos conhecidos de fermentação são adequados. Além disso, considera-se que as células podem ser imobilizadas sobre um substrato como a catálise celular completa e sujeitas às condições de fermentação para a produção de 1-butanol.

#### **MÉTODOS PARA ISOLAMENTO DE BUTANOL DO MEIO DE FERMENTAÇÃO**

[0127] O 1-butanol bioproduzido pode ser isolado a partir do meio de fermentação usando-se métodos conhecidos na técnica. Por exemplo, os sólidos podem ser removidos do meio de fermentação por centrifugação, filtração, decantação ou similares. Então, o butanol pode ser isolado do meio de fermentação, que foi tratado para remoção de sólidos conforme descrito acima, usando-se métodos como destilação, extração líquido-líquido ou separação com base na membrana. Uma vez que o butanol forma um ponto de ebulição baixo, a mistura azeotrópica com água, a destilação pode ser usada apenas para separar a mistura até sua composição azeotrópica. A destilação pode ser usada em associação com outro método de separação para obter a separação sobre o azeótropo. Os métodos que podem ser usados em associação com a destilação para isolar e purificar o butanol incluem, mas não se limitam a, decantação, extração líquido-líquido, adsorção e técnicas com base na membrana. Adicionalmente, o butanol pode ser isolado usando-se destilação azeotrópica usando-se um carregador (consulte, por exemplo, Doherty e Malone, *Conceptual Design of Distillation Systems*, McGraw Hill, Nova York, 2001).

[0128] A mistura de butanol e água forma um azeótropo heterogêneo para que a destilação possa ser usada em associação com a

decantação para isolar e purificar o butanol. Neste método, o butanol que contém o caldo de fermentação é destilado para se aproximar da composição azeotrópica. Então, a mistura azeotrópica é condensada e o butanol é separado do meio de fermentação por decantação. A fase aquosa decantada pode ser devolvida para a primeira coluna de destilação como um refluxo. A fase orgânica decantada rica em butanol pode ser ainda purificada por destilação em uma segunda coluna de destilação.

[0129] O butanol pode também ser isolado a partir do meio de fermentação usando-se extração líquido-líquido em associação com a destilação. Neste método, o butanol é extraído do caldo de fermentação usando-se a extração líquido-líquido com um solvente adequado. A fase orgânica que contém butanol é, então, destilada para separar o butanol do solvente.

[0130] A destilação em combinação com a adsorção também pode ser usada para isolar o butanol do meio de fermentação. Neste método, o caldo de fermentação contendo o butanol é destilado para próximo da composição azeotrópica e então o restante da água é removido usando-se um adsorvente, como coador molecular (Aden *et al.* *Lignocellulosic Biomass to Ethanol Process Design and Economics Utilizing Co-Current Dilute Acid Prehydrolysis and Enzymatic Hydrolysis for Corn Stover*, Relatório NREL/TP-510-32438, Laboratório Nacional De Energia Renovável, Junho 2002).

[0131] Além disso, a destilação em associação com a pervaporação pode ser usada para isolar o butanol do meio de fermentação. Neste método, a fermentação caldo de fermentação que contém o butanol é destilada para próximo da composição azeotrópica e, então, a água remanescente é removida por pervaporação através de uma membrana hidrofílica (Guo *et al.*, *J. Membr. Sci.* 245, 199-210 (2004)).

### **EXEMPLOS**

[0132] A presente invenção é ainda definida nos Exemplos a

seguir. Deve-se entender que estes Exemplos, enquanto indicando realizações preferenciais da invenção, são fornecidos somente a título ilustrativo. A partir da discussão acima e desses Exemplos, um técnico no assunto pode averiguar as características essenciais dessa invenção e sem se afastar do espírito e o escopo deste, pode fazer várias alterações e modificações da invenção para adaptá-la a vários usos e condições.

#### **MÉTODOS GERAIS**

[0133] Técnicas padrão de DNA recombinante e de clonagem molecular usadas nos Exemplos são bem conhecidas na técnica e são descritas por Sambrook, J., Fritsch, E.F. e Maniatis, T. *Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, (1989) (Maniatis) e por T. J. Silhavy, M. L. Bennan, e L. W. Enquist, *Experiments with Gene Fusions*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1984) e por Ausubel, F. M. et al., *Current Protocols in Molecular Biology*, pub. por Greene Publishing Assoc. e Wiley-Interscience (1987) e por *Methods in Yeast Genetics*, 2005, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

[0134] Os materiais e métodos adequados para a manutenção e crescimento de culturas bacterianas são conhecidos na técnica. As técnicas adequadas para o uso nos Exemplos a seguir podem ser encontradas no *Manual of Methods for General Bacteriology* (Phillipp Gerhardt, R. G. E. Murray, Ralph N. Costilow, Eugene W. Nester, Willis A. Wood, Noel R. Krieg e G. Briggs Phillips, eds), American Society for Microbiology: Washington, DC. (1994)) ou por Thomas D. Brock em *Biotechnology: A Textbook of Industrial Microbiology*, Segunda Edição, Sinauer Associates, Inc., Sunderland, MA (1989). Todos os reagentes, enzimas de restrição e materiais usados para o crescimento e manutenção das células microbianas foram obtidos junto à Aldrich Chemicals (Milwaukee, WI), BD Diagnostic Systems (Sparks, MD), Life

Technologies (Rockville, MD), ou junto à PA Sigma Chemical Company (St. Louis, MO), a menos que especificado de outro modo. As linhagens microbianas foram obtidas junto à Coleção America de Tipos de Cultura (ATCC), Manassas, VA, a menos que apontado de outra forma. Todos os *primers* de oligonucleotídeos foram sintetizados pela Sigma-Genosys (Woodlands, TX) ou Integrated DNA Technologies (Coralsville, IA).

[0135] O meio sintético completo é descrito em Amberg, Burke e Strathern, 2005, Methods in Yeast Genetics, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

#### HPLC

[0136] A análise para composição de fermentação de subproduto é conhecida pelos técnicos no assunto. Por exemplo, um método de cromatografia líquida de alto desempenho (HPLC) utiliza uma coluna Shodex SH-1011 com uma coluna de segurança Shodex SH-G (ambas adquiridas junto à Waters Corporation (Milford, MA), com detecção de índice de refração (RI). A separação cromatográfica é alcançada usando-se H<sub>2</sub>SO<sub>4</sub> a 0,01 M como a fase móvel com uma taxa de fluxo de 0,5 ml/min e uma coluna de temperatura de 50°C. O tempo de retenção do isobutanol é de 47,6 minutos.

[0137] O significado das abreviações é como a seguir: "s" significa segundo(s), "min" significa minuto(s), "h" significa hora(s), "psi" significa libras por polegadas quadradas, "nm" significa nanômetro(s), "d" significa dias, "μl" significa microlitro(s), "ml" significa mililitro(s), "l" significa litro(s), "mm" significa milímetro(s), "nm" significa nanômetros, "mM" significa milimolar, "M" significa molar, "mmol" significa milimol(es), "μmole" significa micromol(es), "g" significa grama(s), "μg" significa micrograma(s), e "ng" significa nanograma(s), "PCR" significa reação em cadeia da polimerase, "OD" significa densidade óptica, "OD<sub>600</sub>" significa densidade óptica medida em um comprimento de onda de 600 nm, "kDa" significa quilodaltons, "g" significa constante de gravitação, "bp"

significa pares de bases, "kbp" significa par(es) em quilobases, "p/v" significa porcentagem de peso/volume, "v/v" significa volume/volume, "w%" significa porcentagem em peso, "HPLC" significa cromatografia líquida de alto desempenho e "CG" significa cromatografia gasosa. O termo "seletividade molar" é o número de moles de produto produzido por mol de substrato de açúcar consumido é relatado como percentual.

### **EXEMPLO 1**

#### **EXPRESSÃO DE DHAD DE *K. LACTIS* EM LINHAGEM COM DELEÇÃO DE LEUI DE *S. CEREVISIAE***

##### **CEREVISIAE**

[0138] O gene LEU1 de levedura codifica isopropilmalato desidratase, uma enzima que necessita de um cluster Fe-S para sua função. O impacto da deleção de LEU1 sobre a atividade de DHAD expressa a partir da região codificadora de DHAD de *Kluyveromyces lactis* foi examinada neste exemplo. Para expressão gênica em leveduras, foi usado o vetor de transporte pNY13 (SEQ ID NO:29) derivado de pRS423. Este vetor de transporte contido em uma origem de replicação F1 (1423 até 1879) para manutenção em *E. coli* e uma origem de 2 micrôn (nt 7537 até 8881) para replicação em leveduras. O vetor tem um promotor FBA (nt 2111 até 3110) e terminador FBA (nt 4316 até 5315). Além disso, carrega o marcador HIS3 (nt 504 até 1163) para seleção em leveduras e marcador de resistência à ampicilina (nt 6547 até 7404) para seleção em *E. coli*. pNY9 é o mesmo vetor com um marcador URA3 que substitui o marcador HIS3.

[0139] A região codificadora ILV3 para DHAD de *Kluyveromyces lactis* foi sintetizada com otimização de códon para expressão em *S. cerevisiae* pelo DNA 2.0 (Menlo Park, CA). A sequência sintetizada clonada foi amplificada por PCR. Durante a amplificação, uma porção do peptídeo sinal mitocondrial para DHAD na terminação N foi deletada pelo uso de ilv3(K)(0) F(delet) como primer forward com ilv3(K)(0)-R como primer reverso, resultando

em uma região codificadora para expressão no citosol (SEQ ID NO:173). Além disso, um sítio SphI foi incorporado no primer *forward*, enquanto um sítio NotI foi incluído no *primer reverso*. O produto de PCR foi clonado nos vetores de transporte pNY9 e pNY13, de modo que a região codificadora estivesse mediada pelo controle do promotor FBA. Tanto o produto de PCR como cada vetor (pNY9, pNY13) foram digeridos com SphI e NotI. Após a digestão, os componentes foram ligados, e a mistura de ligação foi transformada nas células competentes TOP10 (Invitrogen). Os transformantes foram selecionados em placas de ágar LB suplementadas com 100 µg/ml de ampicilina. Os clones positivos foram selecionados por PCR com os *primers forward* e *reverso* descritos acima. Os plasmídios resultantes foram designados como pRS423::FBApILV3(KL) e pRS426::FBAp-/LV3(KL), derivados de plasmídios pNY13 e pNY9, respectivamente.

[0140] Para estudar a expressão da DHAD de *K. lactis* em *S. cerevisiae*, o vetor de expressão pRS423::FBAp-/LV3(KL) junto com um vetor vazio pRS426 foram transformados nas linhagens BY4743 e BY4743 leul::kanMX4 (ATCC 4034377). A preparação celular competente e transformação foram baseadas no Kit de Transformação de Leveduras Congeladas da Zymo Research. Os transformantes foram selecionados em placas de ágar com meio sintético de levedura sem histidina e uracila (Teknova). Para testes enzimáticos, as linhagens que carregam o constructo de expressão e o vetor vazio pRS426 foram primeiro cultivadas durante a noite em 5 ml de meio sintético de levedura completo sem histidina e uracila. 5 ml das culturas mantidas durante a noite foram transformadas em 100 ml de meio em um frasco de 250 ml. As culturas foram colhidas quando alcançaram O.D. de 1 a 2 a 600 nm. As amostras foram lavadas com 10 ml de 20 mM de Tris (pH 7,5) e então resuspensas em 1 ml do mesmo tampão Tris. As amostras foram transferidas para tubos de 2,0 ml contendo 0,1 mm de sílica (Matriz de lise B,

MP biomedicals). As células foram então rompidas em um batedor de esferas (B10101). O sobrenadante foi obtido por centrifugação em uma microcentrífuga a 13.000 rpm a 4°C por 30 minutos. Tipicamente, 0,06 a 0,1 mg de extrato de proteína cru foi usado em um teste DHAD. A proteína no extrato cru foi determinada pelo teste Bradford com coloração Coomassie.

#### **TESTE DA ENZIMA ÁCIDO-DIHIDROXI-DESIDRATASE**

[0141] O teste *in vitro* da enzima DHAD é uma variação do teste descrito em Flint *et al.* (*J. Biol. Chem.* (1993) 268:14732-14742.). O teste foi realizado em um volume total de 1,6 ml e consistiu em: 800 µl 2X tampão (100mM de Tris pH 8,0, 20mM de MgCl<sub>2</sub>), 160 µl 10X substrato (15,6 mg/ml de dihidroxi isovalerate), extrato cru (tipicamente 50 a 200 µg de proteína), e água. A reação foi incubada a 37°C. Em intervalos de 0, 30, 60 e 90 minutos, alíquotas de 350 µl da reação foram removidos e incubados com 350 µl de dinitrofenil hidrazina a 0,05% em HCl a 1N por 30 minutos a 25°C. Para interromper a reação, foi adicionado 350 µl de hidróxido de sódio a 4N na mistura de reação, e a reação foi centrifugada a 15.000 x g por 2 minutos. O sobrenadante foi transferido para um cadiño descartável e foi medida a absorbância a 540 nm em um espectrofotômetro. A quantidade de α-cetoisovalerato (KIV) produzido foi determinada pela entrada da absorbância na equação de regressão linear obtida a partir de um curva padrão de uma α-cetoisovalerato. A quantidade de KIV produzida em cada período foi plotada para determinar a taxa de produção. A inclinação da regressão linear foi então usada para calcular a atividade específica usando-se a fórmula:

[0142] O cálculo da atividade específica = (inclinação de produção de KIV production/1000) / mg proteína por 1,6 ml de reação = mmol/min \* mg

[0143] A desidratase de *K. lactis* tinha uma atividade específica na faixa de 0,2 a 0,35 µmol min<sup>-1</sup> mg<sup>-1</sup> quando expressa na linhagem de levedura BY4743 ( $\Delta$ leu1). Ao contrário, esta enzima tinha uma atividade específica na

faixa de apenas 0,14  $\mu\text{mol min}^{-1} \text{mg}^{-1}$  quando expressa na linhagem de levedura parental BY4743. As linhagens BY4743 ( $\Delta\text{leu1}$ ) e o tipo selvagem BY4743 contendo os vetores vazios pRS423 ou pRS426 tinham um antecedente de atividade na faixa de 0,03 a 0,1  $\mu\text{mol min}^{-1} \text{mg}^{-1}$ .

### **EXEMPLO 2**

#### **EXPRESSÃO DE DIOL DESIDRATASE NAS LINHAGENS COM DELEÇÃO EM LEU1 DE S.**

##### **CEREVISIAE**

[0144] Uma coenzima propanodiol desidratase independente de B12 é divulgada na patente co-pendente de propriedade comum US 20090155870. As sequências que codificam essa coenzima (S-adenosilmetionina dependente de (SAM)) propanodiol desidratase independente de B12 (SEQ ID NO:15) e esta reativase proposta associada (SEQ ID NO:16) na bactéria *Roseburia inulinivorans* [Scott et al.(2006) *J. Bacteriol.*188:4340-9], denominada a seguir como rdhtA e rdhtB, respectivamente, foram sintetizadas como um fragmento de DNA (SEQ ID NO:17) por métodos padrão e clonadas em um vetor em *E. coli* (por DNA2.0, Inc., Menlo Park, CA). Esse clone foi usado como molde para PCR para preparar fragmentos de regiões codificadoras RdhtA e RdhtB separados. A região codificadora RdhtA para a diol desidratase foi amplificada por PCR usando-se primers N695 e N696 (SEQ ID NOs:18 e 19). A região codificadora RdhtA para a diol desidratase ativase foi amplificada por PCR usando-se primers N697 e N698 (SEQ ID NOs:20 e 21). Os dois fragmentos de DNA foram combinados com um fragmento de DNA terminador duplo (SEQ ID NO:22) que tem um terminador ADH (SEQ ID NO:23) e um terminador CYC1 (SEQ ID NO:24) adjacentes entre si na orientação oposta, usando-se PCR SOE (Horton et al. (1989) *Gene* 77:61-68). O fragmento terminador duplo foi isolado como um fragmento de 0,6 kb seguido por digestão Pael de pRS426::FBA-ILV5+GPM-kivD (descrito na publicação de patente co-pendente

e de propriedade comum US 20070092957 A1, Exemplo 17). O fragmento de DNA resultante de 4 kb DNA tinha regiões codificadoras RdhtA e RdhtB na orientação oposta em ambos os lados do terminador duplo, com a extremidade 3' de cada região codificadora adjacente até a sequência terminadora dupla. Este fragmento de DNA foi então clonado pela metodologia de reparo de gap (Ma *et al.* (1987) *Genetics* 58:201-216) no vetor de transporte de *S. cerevisiae* pRS426::FBA-ILV5+GPM-kivD que foi preparado por digestão com BbvCI para remover as regiões codificadoras ILV5 e kivD e a sequência terminadora dupla entre suas extremidades 3'. O plasmídio resultante, pRS426::RdhtAB (below), continha o gene RdhtA mediante o controle do promotor FBA de *S. cerevisiae* (SEQ ID NO:25) e o gene RdhtB mediante o controle do promotor GPM de *S. cerevisiae* (SEQ ID NO:26).

[0145] Os plasmídios pRS426 e pRS426::RdhtAB foram introduzidos em linhagens de *S. cerevisiae* BY4743 (ATCC 201390) e BY4743 leul::kanMX4 (ATCC 4034377) por técnicas padrão (*Methods in Yeast Genetics*, 2005, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, páginas 201-202). As células foram plaqueadas em meio sintético completo sem uracila para selecionar os transformantes. Os transformantes foram testados para atividade de diol desidratase usando-se um teste *in vivo* como segue. As células crescidas unidas em meio sólido foram usadas para inocular em meio líquido (20 ml) em placas de petri. O meio usado foi sintetizado completo sem uracila, com ou sem a adição de 5 g/l de 1,2-propanodiol Aldrich Cat. No. 398039). As placas de petri foram transferidas para um Sistema Anaeropack™ jar (Mitsubishi Gas Chemical Co. Cat. No.50-70). Foi gerado um ambiente anaeróbico (<0,1% de oxigênio) usando-se sachês Pack-Anaero (Mitsubishi Gas Chemical Co. Cat. No. 10-01). Após 48 horas, os sobrenadantes na cultura foram amostrados, filtrados e analisados por HPLC, conforme descrito nos Métodos Gerais. Propanol,

que tem um tempo de retenção de 38,8 minutos, foi observado na cultura de sobrenadantes de linhagens que carregam pRS426::RdhtAB quando 1,2 propanodiol foi fornecido no meio. Os resultados fornecidos na Tabela 5 mostram que foi produzido mais propanol nos sobrenadantes na linhagem que também carregam a deleção de LEU1 do que na linhagem sem deleção de LEU1. As análises estatísticas forneceram uma pontuação P menor que 0,0005.

**TABELA 5****A PRODUÇÃO DE PROPANOL COM A EXPRESSÃO DE PROPANODIOL****DESIDRATASE/REATIVASE EM LEVEDURAS COM E SEM LEU1 KNOCKOUT**

Linhagem	1,2-propanediol Adicionado	Área de Pico de Propanol
BY4743 Δleu1::kanMX4/pRS426::RdhtAB	5 g/l	19472 ± 1403 (n=6)
BY4743 Δleu1::kanMX4/pRS426::RdhtAB	0 g/l	2478 (n=1)
BY4743 /pRS426::RdhtAB	5 g/l	11830 ±1963 (n=6)
BY4743 /pRS426::RdhtAB	0 g/L	2369 (n=1)
BY4743 Δleu1::kanMX4/pRS426	5 g/l	2633 (n=1)
BY4743/pRS426	5 g/l	2841 (n=1)

**EXEMPLO 3****APRIMORAMENTO DA ATIVIDADE DE ÁCIDO DIHIDROXI DESIDRATASE NO CITOSOL****(DHAD) EM S. CEREVISAIE ATRAVÉS DE UM ROMPIMENTO DE ILV3 MITOCONDRIAL****CONSTRUÇÃO DO VETOR/HOSPEDEIRO**

[0146] Em *S. cerevisiae* ILV3 codifica a ácido dihidroxi desidratase que está envolvida na biossíntese da cadeia de aminoácido ramificada. Para reduzir o antecedente da expressão de ILV3 endógeno para testes enzimáticos *in vitro* em *S. cerevisiae*, um cassete de rompimento ilv3::URA3 foi construído por amplificação por PCR do marcador URA3 a partir de pRS426 (ATCC No. 77107) com primers "ILV3::URA3 F" e "ILV3::URA3 R", fornecidos como SEQ

ID NO:30 e 31. Esses *primers* produziram um produto de PCR URA3 de 1.4 kb que continha extensões de 70 bp 5' e 3' idênticas às sequências a montante e a jusante do *locus* do cromossomo ILV3 para recombinação homóloga. O produto de PCR foi transformado nas células BY4741 (ATCC 201388) usando-se técnicas genéticas padrão (*Methods in Yeast Genetics*, 2005, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, páginas 201-202) e os transformantes resultantes foram mantidos em meio sintético completo sem uracila e suplementado com glicose a 2% a 30°C. Os transformantes foram selecionados por PCR usando-se *primers* "ILV3 F Check" e "URA3 REV Check", fornecidos como SEQ ID NOs:32 e 33, para verificar a integração no sítio correto e rompimento do locus ILV3 endógeno. Os transformantes corretos tinham o genótipo: BY4741 ilv3::URA3.

[0147] A construção dos plasmídios pRS423::FBAp-/LV3(KL) e pRS426::FBAp-/LV3(KL) foi descrita no Exemplo 1. A construção de pRS423::CUP1-alsS+FBA-ILV3 foi descrita no pedido de patente co-pendente de propriedade comum US 20070092957 A1, Exemplo 17 que é integralmente incorporado ao presente pedido como referência. pRS423::CUP1-alsS+FBA-ILV3 é o mesmo plasmídio que pRS423::CUP1p-a/sS-FBAp-ILV3. Essa construção contém um gene quimérico contendo o promotor CUP1 de *S. cerevisiae* (SEQ ID NO:34), região codificadora de alsS de *Bacillus subtilis* (SEQ ID NO:27), e terminador CYC1 (SEQ ID NO:24); e também um gene quimérico contendo o promotor FBA de *S. cerevisiae* (SEQ ID NO:25), região codificadora ILV3 de *S. cerevisiae* sem a sequência codificadora sinal de direcionamento mitocondrial (SEQ ID NO:111) e terminador ADH1 (SEQ ID NO:23).

#### **PREPARAÇÃO DE AMOSTRAS**

[0148] Os vetores plasmidiais pRS423::CUP1 p-alsS-FBAp-ILV3 e pRS423::FBAp-/LV3(KL) foram transformados nas linhagens BY4741

*ilv3::URA3* usando-se técnicas genéticas padrão (*Methods in Yeast Genetics*, 2005, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY) e mantidos em meio sintético completo sem histidina. Os vetores plasmidiais pRS423::CUP1 p-alsS-FBAp-ILV3 e pRS426::FBAp-/LV3(KL) também foram transformados na linhagem BY4741. As culturas aeróbicas foram cultivadas em frascos de 1000 ml contendo 200 ml de meio sintético completo sem histidina e suplementado com 2% de glicose em um incubador Innova4000 (New Brunswick Scientific, Edison, NJ) a 30°C e 225 rpm. As culturas foram colhidas em medidas de OD600 de 1,0 a 2,0 e transformadas em péletes por centrifugação a 6000 x g por 10 minutos. Os péletes celulares foram lavados com Tris-HCl a 10 mM, pH 8,0 e os péletes foram armazenados a -80°C até serem testados para atividade. Os extratos livres de célula foram preparados pelo método de pulsação de gotas padrão usando-se 1 ml de gotas de 0,5 mm e 1,5 ml de suspensão de célula da levedura. A concentração de proteína nos extratos foi determinada pelo teste Bradford com coloração Coomassie. Os testes da enzima DHAD e cálculos da atividade específica foram realizados conforme descrito no Exemplo 1. Os resultados fornecidos na Tabela 6 mostram que houve maior atividade de DHAD nas células com deleção de ILV3 do que nas células sem deleção de ILV3.

**TABELA 6****ATIVIDADE DE DHAD EM LEVEDURAS COM E SEM DELEÇÃO DE ILV3**

Linhagem	Atividade Específica ( $\mu\text{mol}/\text{min}^*\text{mg}$ )	Média da Atividade Específica ( $\mu\text{mol}/\text{min}^*\text{mg}$ )
BY4741	0,018	0,013
	0,014	
	0,008	
BY4741 pRS423::CUP1p-a/sS-FBAp-/LV3	0,018	0,019
	0,020	
BY4741 pRS426::FBAp-/LV3(KL)	0,040	0,038
	0,036	
BY4741 <i>ilv3::URA3</i>	0,00006	BY4741 <i>ilv3::URA3</i>

Linhagem	Atividade Específica ( $\mu\text{mol}/\text{min}^*\text{mg}$ )	Média da Atividade Específica ( $\mu\text{mol}/\text{min}^*\text{mg}$ )
BY4741 <i>ilv3::URA3</i>	0,00075	
BY4741 <i>ilv3::URA3</i> pRS423::CUP1p-a/sS-FBAp- <i>ILV3</i>	0,030 0,028	0,029
BY4741 <i>ilv3::URA3</i> pRS423::FBAp-/LV3(KL)	0,317 0,244	0,281

#### **VERIFICAÇÃO DA FORMAÇÃO DE ALFA-CETOISOVALERATO POR HPLC**

[0149] A formação de alfa-cetoisovalerato a partir dos testes *in vitro* da enzima DHAD foi realizado usando-se HPLC e derivatização de semicarbizada. O teste da enzima DHAD foi realizado em um volume total de 1,6 ml e consistiu em: 800  $\mu\text{l}$  2X tampão (100mM de Tris pH 8,0, 20mM de MgCl<sub>2</sub>), 160  $\mu\text{l}$  10X substrato (15,6 mg/ml de dihidroxi isovalerate), extrato cru (tipicamente 50 a 200  $\mu\text{g}$  de proteína), e água. As reações foram incubadas a 37°C. Em intervalos de tempo de zero e 90 minutos, alíquotas de 350  $\mu\text{l}$  das reações foram removidas, transferidas para o gelo e centrifugadas a 13,000 x g por 2 minutos a 4°C para remover a proteína precipitada. Os sobrenadantes foram transferidos para colunas com rotação no gelo Microcon YM-10 (Sigma) e centrifugadas a 13.000 x g por 20 minutos a 4°C para remover enzimas e proteínas solúveis.

[0150] Os fluxos foram misturados com 100  $\mu\text{l}$  de reagente de derivatização (hipocloreto de semicarbizada a 1% e acetato de sódio trihidratado a 1,5%) e incubados em temperatura ambiente por 15 minutos. As reações foram passadas através de filtros CoStar (CoStar, filtro com 0,22  $\mu\text{m}$ ) a 13.000 x g por 5 minutos a 4°C para remover qualquer precipitado. Os fluxos foram transferidos para frascos para análise HPLC.

[0151] A análise de alfa-cetoisovalerato derivatizada foi conduzida usando-se cromatografia de fase reversa sobre uma coluna Supelco LC-1 8

com coluna guarda Superguard LC-1 8-DB (Supelco; 25 cm x 4.6 mm, 5 µm). Os volumes das injeções eram de 10 µl. As fases móveis foram metanol (A) e 50 mM de NaOAc pH 7,2. O programa de gradiente utilizado é fornecido na Tabela 7, com detecção a 250 nm.

**TABELA 7****GRADIENTE USADO PARA TESTE HPLC DE ALFA-CETOISOVALERATO DERIVATIZADO**

Tempo (min)	Vazão (ml/min)	% de NaOAc (50mM)	% de MEOH	Curva
Inicial	1,0	95	5	
5	1,0	95	5	6
20	1,0	70	30	6
21	1,0	0	100	6
25	1,0	0	100	6
26	1,0	95	5	6
35	1,0	95	5	6

[0152] O tempo de retenção de alfa-cetoisovalerato derivado de semicarbizada foi de 11,5 minutos.

[0153] Os resultados fornecidos na Tabela 8 confirmaram que KIV foi produzido nas células, conforme detectado no teste indireto para a atividade específica acima. A quantidade de KIV listada na coluna de teste DHAD é a quantidade determinada indiretamente na amostra de 90 minutos para o testes de atividade descrito acima na determinação da atividade específica. Esta quantidade de KIV se correlaciona bem com a quantidade detectada no teste HPLC.

**TABELA 8****COMPARAÇÃO DE KIV DETECTADA NO TESTE DE ATIVIDADE DE DHAD E POR HPLC**

Linhagem	Produção de Ceto-isovalerato (µM)	
	Teste DHAD	HPLC
BY4741 <i>i/lV3::URA3 pRS423::FBAp-lLV3(KL)</i>	256	256
BY4741 <i>i/lV3::URA3 pRS423::FBAp-lLV3(KL)</i>	162	163

### REIVINDICAÇÕES

1. MÉTODO PARA PRODUÇÃO DE ISOBUTANOL, caracterizado por compreender

fornecer uma célula hospedeira de levedura recombinante que expressa uma proteína heteróloga ácido dihidroxi desidratase (DHAD) de 2Fe-2S que compreende uma sequência de aminoácidos possuindo a SEQ ID NO: 179 ou 187 e uma via biosintética de isobutanol, e em que a proteína DHAD compreende três cisteínas conservadas, correspondentes às posições 56, 129 e 201 na sequência SEQ ID NO:179; e

cultivar a célula hospedeira de levedura recombinante sob condições em que o isobutanol é produzido.

2. MÉTODO PARA CONVERSÃO DE 2,3-DIHIDROXI ISOVALERATO, para  $\alpha$ -cetoisovalerato, caracterizado por compreender:

a) fornecer (1) uma célula hospedeira de levedura recombinante que expressa uma proteína heteróloga ácido dihidroxi desidratase (DHAD) de 2Fe-2S que compreende uma sequência de aminoácidos possuindo a SEQ ID NO: 179 ou 187, e em que a proteína DHAD compreende três cisteínas conservadas, correspondentes às posições 56, 129 e 201 na sequência SEQ ID NO:179 e (2) uma fonte de 2,3-dihidroxi isovalerato; e

b) cultivar a célula hospedeira recombinante de (a) com a dita fonte de 2,3-dihidroxi isovalerato sob condições em que o 2,3-dihidroxi isovalerato é convertido pela célula hospedeira para  $\alpha$ -cetoisovalerato.

3. MÉTODO, de acordo com qualquer uma das reivindicações 1 a 2, caracterizado pela célula hospedeira de levedura recombinante ser selecionada a partir do grupo que consiste em *Saccharomyces*, *Schizosaccharomyces*, *Hansenula*, *Candida*, *Kluyveromyces*, *Yarrowia* e *Pichia*.

4. MÉTODO, de acordo com a reivindicação 3, caracterizado

pela célula hospedeira de levedura recombinante ser *Saccharomyces cerevisiae*.

5. MÉTODO, de acordo com a reivindicação 1, caracterizado pela via biossintética de isobutanol compreender as seguintes conversões de substrato para produto:

- (i) piruvato para acetolactato;
- (ii) acetolactato para 2,3-dihidroxi isovalerato;
- (iii) 2,3-dihidroxi isovalerato para  $\alpha$ -cetoisovalerato;
- (iv)  $\alpha$ -cetoisovalerato para isobutiraldeído; e
- (v) isobutiraldeído para isobutanol.

6. MÉTODO, de acordo com a reivindicação 1, caracterizado por compreender ainda a recuperação do isobutanol.

7. MÉTODO, de acordo com a reivindicação 6, caracterizado pela recuperação ser por destilação, extração líquido-líquido, adsorção, decantação, pervaporação ou combinações dos mesmos.

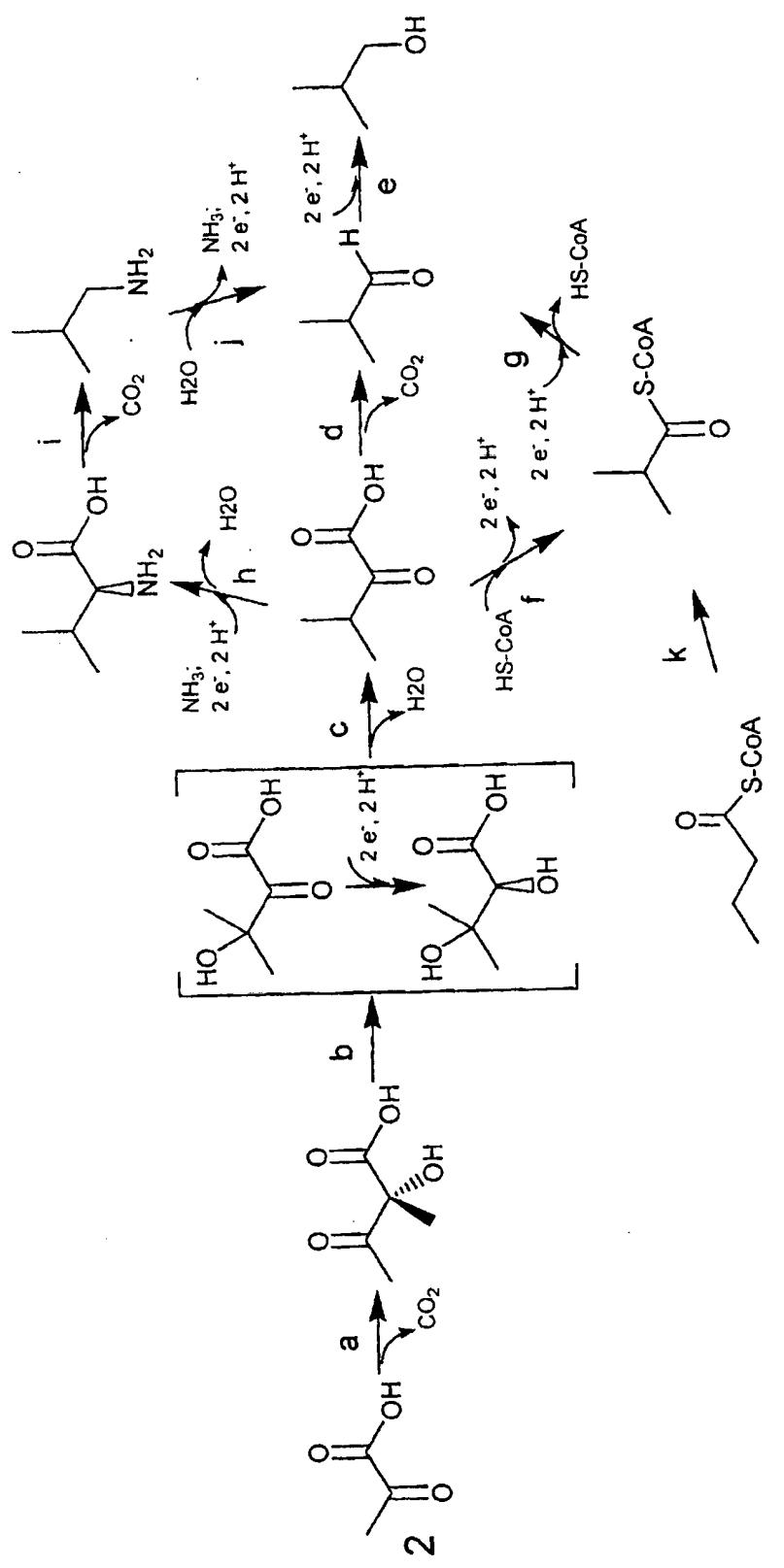
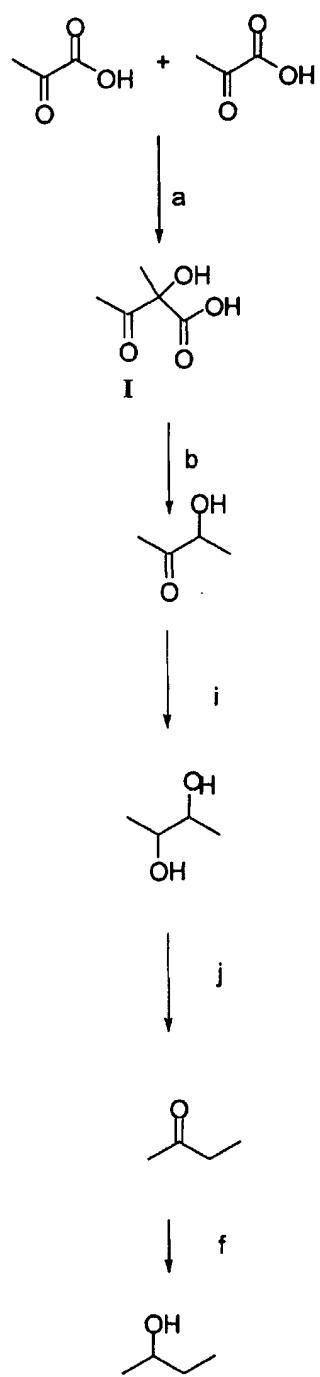


Fig. 1



*Fig. 2*