



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

(21) **PI0618937-7 A2**

(22) Data de Depósito: 17/11/2006
(43) Data da Publicação: 13/09/2011
(RPI 2123)



★ B R P I O 6 1 8 9 3 7 A 2 ★

(51) Int.Cl.:

C12N 9/04
C12N 1/20
C12N 15/00
C12P 7/56
C12Q 1/68
C07H 21/04

(54) Título: CÉLULAS DE LEVEDURA QUE PRODUZEM ÁCIDO LÁTICO TENDO L-OU D-LACTATO NÃO FUNCIONAL: GENE DE FERRICITOCROMO C OXIDOREDUTASE

(30) Prioridade Unionista: 23/11/2005 US 60/739.458, 23/11/2005 US 60/739.824

(73) Titular(es): Natureworks LLC

(72) Inventor(es): Aristos Aristidou, Benjamin Matthew Hause, Catherine Asleson Dundon, Matthew Miller, Pim Van Hoek, Pirkko Suominen

(74) Procurador(es): Alexandre Ferreira

(86) Pedido Internacional: PCT US2006044815 de 17/11/2006

(87) Publicação Internacional: WO 2007/117282 de 18/10/2007

(57) Resumo: CÉLULAS DE LEVEDURA QUE PRODUZEM ÁCIDO LÁTICO TENDO L- OU D-LACTATO NAO FUNCIONAL: GENE DE FERRICITOCROMO C OXIDOREDUTASE São descritas células de levedura tendo um gene lactato desidrogenase exógeno e modificadas reduzindo-se a atividade L- ou D-lactato:ferricitocromo c oxidoredutase na célula. Isto leva ao menor consumo de lactato pela célula e pode aumentar os rendimentos gerais de lactato em um processo de fermentação. Células tendo a menor atividade L- ou D-lactato:ferricitocromo c oxidoredutase podem ser selecionadas pela resistência a ácidos orgânicos, tais como ácido láctico ou glicólico.

Pe - 0618934-7
Pet - 020080090904
Data - 30/06/2008

"CÉLULAS DE LEVEDURA QUE PRODUZEM ÁCIDO LÁTICO
TENDO L- OU D-LACTATO NÃO FUNCIONAL: GENE DE FERRICITOCROMO
C OXIDOREDUCTASE"

Esta invenção foi feita com suporte do governo pe-
5 la Cooperative Agreement DE-FC36- 03G013145 outorgado pelo
United States Department of Energy. O governo tem certos di-
reitos nesta invenção.

Este pedido reivindica benefício para pedido pro-
visório dos Estados Unidos 60/739.458 e 60/739.824, ambos
10 depositados em 23 de novembro de 2005.

Esta invenção diz respeito a certa levedura gene-
ticamente modificada, métodos para preparar esta levedura e
processos de fermentação para produzir ácido lático usando
esta levedura geneticamente modificada.

15 Ácido lático é fabricado por um processo de fer-
mentação industrial. A fermentação é conduzida usando vários
tipos de espécies bacterianas, que consomem açúcares (prin-
cipalmente glicose) e convertem estes açúcares no ácido de-
sejado. À medida que ácido lático é produzido, o meio de
20 fermentação torna-se cada vez mais ácido. A maioria das bac-
térias que produzem estes ácidos orgânicos não tem bom desem-
penho em ambientes fortemente ácidos - elas tanto não sobre-
vivem nestas condições como de outra maneira geram o produto
tão lentamente a ponto de ser economicamente inviáveis.

25 Por este motivo, processos de fermentação do ácido
comercial são tamponados pela adição de um agente que neu-
traliza à medida que ele é formado. Isto mantém o caldo no
pH neutro ou próximo dele e permite que a bactéria cresça e

produza de maneira eficiente. Entretanto, isto converte o ácido em um sal, que subsequenteiramente pode ser dividido para obter o produto na sua forma de ácido desejado.

O agente de tamponamento mais comum é um composto de cálcio, que neutraliza o ácido orgânico para formar o sal de cálcio correspondente. Após o sal de cálcio ser recuperado do caldo de fermentação, ele é dividido pela adição de um ácido mineral, tipicamente ácido sulfúrico, para gerar o ácido orgânico e formar um sal de cálcio insolúvel do ácido mineral. Este processo então envolve despesa direta para o agente de tamponamento e ácido mineral, bem como custos para o manuseio e disposição do subproduto de sal de cálcio indesejável. Estes custos podem ser reduzidos ou eliminados se o biocatalisador puder crescer e produzir de maneira eficiente em condições de pH baixo.

Espécies de levedura foram consideradas candidatas para tais fermentações de baixo pH. Muitas espécies de levedura fermentam naturalmente açúcares de hexose em etanol, mas poucas, talvez nenhuma, produzem naturalmente ácidos orgânicos desejados tal como ácido lático. Conseqüentemente, foram feitos esforços para modificar geneticamente várias espécies de levedura para inserir um ou mais genes que permitirão à célula produzir ácido lático. A fim de desviar o metabolismo do açúcar da produção de etanol para a produção de ácido lático, estas células também foram geneticamente modificadas para interromper ou deletar o gene piruvato descarboxilase nativo (*PDC*). Este trabalho é descrito, por exemplo, em WO 99/14335, WO 00/71738 A1, WO 02/42471 A2, WO

03/049525 A2, WO 03/102152 A2 e WO 03/102201A2.

Continua a existir um desejo de fornecer biocatalisadores ainda melhores para processos de fermentação do ácido orgânico. De um modo particular, é desejável melhorar
5 as produtividades e rendimentos destes processos de fermentação, particularmente em processos de fermentação de baixo pH não tamponado.

Em um aspecto, esta invenção é uma célula de levedura geneticamente modificada tendo pelo menos um gene de
10 lactato desidrogenase (LDH) funcional exógeno cuja célula de levedura é incapaz de crescer em D-lactato, L-lactato ou tanto em D- quanto em L-lactato como uma única fonte de carbono.

Esta invenção é também uma célula de levedura geneticamente modificada tendo (a) pelo menos um gene de lactato desidrogenase (LDH) funcional exógeno integrado em seu
15 genoma e (b) uma deleção ou ruptura de pelo menos um gene de L- ou D-lactato:ferricitocromo c oxidoreductase nativo. Em modalidades preferidas, pelo menos um gene L-lactato:ferricitocromo c oxidoreductase é deletado e interrompido, quando o dito gene *LDH* é um gene *L-LDH*, e pelo menos um gene D-lactato:ferricitocromo c oxidoreductase nativo é deletado ou interrompido se o gene *LDH* for um gene *D-LDH*.
20

Esta invenção é também um processo de fermentação
25 em que uma célula de levedura geneticamente modificada da invenção é cultivada em condições de fermentação em um caldo de fermentação que inclui um açúcar fermentável para produzir ácido lático ou um sal deste.

Esta invenção é também um método para produzir uma célula de levedura geneticamente modificada da maneira descrita anteriormente, compreendendo (a) interromper ou deletar um gene de L- ou D-lactato:ferricitocromo c oxidoredutase nativo em uma célula de levedura e (b) transformar a célula de levedura tendo o gene L- ou D-lactato:ferricitocromo c oxidoredutase deletado ou interrompido com um vetor contendo um cassete do gene *LDH* exógeno funcional para integrar o cassete do gene *LDH* no genoma da célula de levedura. Prefere-se que o cassete do gene *LDH* inclua um gene L-*LDH* funcional no caso de uma ruptura ou deleção de um gene L-lactato:ferricitocromo c oxidoredutase e inclui um gene *DLDH* no caso de uma ruptura ou deleção de um gene D-lactato:ferricitocromo c oxidoredutase.

Esta invenção é um método para produzir uma célula de levedura geneticamente modificada da maneira descrita anteriormente, compreendendo a) transformar uma célula de levedura tendo um gene L-lactato:ferricitocromo c oxidoredutase nativo com uma construção contendo um L-Cassete do gene *LDH* exógeno para produzir um mutante tendo um gene L-*LDH* exógeno funcional integrado em seu genoma e em seguida (b) interromper ou deletar o gene L-lactato:ferricitocromo c oxidoredutase nativo.

Esta invenção é um método para produzir uma célula de levedura geneticamente modificada da maneira descrita anteriormente, compreendendo transformar uma célula de levedura tendo um gene L-lactato:ferricitocromo c oxidoredutase nativo com uma construção contendo (a) uma região flanquea-

dora 5' gene L-lactato:ferricitocromo c oxidoreductase nativo; (b) uma região flanqueadora 3' gene L-lactato:ferricitocromo c oxidoreductase nativo e (c) um cassete do gene L-LDH exógeno residindo no construção a jusante da região flanqueadora 5' gene L-lactato:ferricitocromo c oxidoreductase nativo e a montante da região flanqueadora 3' do gene L-lactato:ferricitocromo c oxidoreductase nativo, a construção sendo desprovida de um cassete do gene L-lactato:ferricitocromo c oxidoreductase funcional.

10 Esta invenção é ainda um método adicional para produzir uma célula de levedura geneticamente modificada do primeiro aspecto, compreendendo a) transformar uma célula de levedura tendo um gene D-lactato:ferricitocromo c oxidoreductase nativo com uma construção contendo um cassete do gene
15 D-LDH exógeno para produzir um mutante tendo um gene D-LDH exógeno funcional integrado em seu genoma e em seguida (b) interromper ou deletar o gene D-lactato:ferricitocromo c oxidoreductase nativo.

Esta invenção é, além disso, um método para produzir uma célula de levedura geneticamente modificada do primeiro aspecto, compreendendo transformar uma célula de levedura tendo um gene D-lactato:ferricitocromo c oxidoreductase nativo com um vetor contendo (a) uma região flanqueadora 5' do gene D-lactato:ferricitocromo c oxidoreductase nativo; (b)
25 uma região flanqueadora 3' do gene D-lactato:ferricitocromo c oxidoreductase nativo e (c) um cassete do gene D-LDH exógeno residindo no construção a jusante da região flanqueadora 5' do gene D-lactato:ferricitocromo c oxidoreductase nativo e

a montante da região flanqueadora 3' do gene D-lactato:ferricitocromo c oxidoredutase nativo, o vetor sendo desprovido de um cassete do gene D-lactato:ferricitocromo c oxidoredutase funcional.

5 Observou-se que a célula geneticamente modificada da invenção produz lactato em produtividades maiores e tituladores de lactato maiores, especialmente em condições de baixo pH, do que célula similar contendo o mesmo cassete do gene *LDH* exógeno, mas tendo um gene L-lactato:ferricitocromo c oxidoredutase funcional (no caso de
10 um cassete do gene *L-LDH*) ou gene D-lactato:ferricitocromo c oxidoredutase (no caso de um cassete do gene D-*LDH*). A célula transformada não tem a capacidade de crescer em um meio contendo lactato como sua única fonte de carbono, e assim a
15 deleção do gene L- ou D-lactato:ferricitocromo c oxidoredutase permite que a célula transformada seja identificada, se desejado, baseado na incapacidade dos transformantes de crescer no lactato como uma única fonte de carbono. Além disso, observou-se que a deleção do gene L-lactato:ferricitocromo c
20 oxidoredutase de acordo com a invenção melhora a resistência do ácido da célula transformada, e em particular melhora sua resistência a ácido glicólico. Isto permite que transformantes bem sucedidos sejam selecionados usando um meio ácido tal como um meio contendo ácido glicólico. A capacidade de
25 usar ácidos tal como ácido glicólico como um agente seletivo torna possível evitar o uso de antibiótico ou outros marcadores de gene de resistência durante a transformação das cepas para deletar ou interromper o gene L-

lactato:ferricitocromo c oxidoreductase.

Figura 1 é um diagrama que descreve o plasmídeo
pMM22.

Figura 2 é um diagrama que descreve o plasmídeo
5 pMM28.

Figura 3 é um diagrama que descreve o plasmídeo
pMM35.

Figura 4 é um diagrama que descreve o plasmídeo
pBH76.

Figura 5 é um diagrama que descreve o plasmídeo
10 pMM38.

Figura 6 é um diagrama que descreve o plasmídeo
pMI318.

Figura 7 é um diagrama que descreve o plasmídeo
15 pMI321.

Figura 8 é um diagrama que descreve o plasmídeo
pMI355.

Figura 9 é um diagrama que descreve o plasmídeo
pMI356.

Figura 10 é um diagrama que descreve o plasmídeo
20 pMI433.

Figura 11 é um diagrama que descreve o plasmídeo
pMI449.

Figura 12 é um diagrama que descreve o plasmídeo
25 pMI454.

Figura 13 é um diagrama que descreve o plasmídeo
pMM44.

Figura 14 é um diagrama que descreve o plasmídeo

pMM45.

Figura 15 é um diagrama que descreve o plasmídeo pCM149.

A levedura geneticamente modificada da invenção é preparada realizando certas modificações genéticas em uma célula de levedura hospedeira. A célula de levedura hospedeira é uma que, como uma cepa tipo selvagem, é capaz de metabolizar de forma nativa pelo menos um açúcar em piruvato. Ela pode ser incapaz de crescer de forma nativa no lactato como uma única fonte de carbono. *S. bulderi* é um exemplo de uma célula de levedura como essa. Em outras modalidades da invenção, a célula é geneticamente modificada para torná-la incapaz de crescer no lactato como uma única fonte de carbono. Quando transformada para introduzir um gene L-LDH exógeno da maneira aqui descrita, a célula de levedura hospedeira preferivelmente é uma que como uma cepa tipo selvagem tem um gene L- lactato:ferricitocromo c oxidoredutase funcional nativo. A célula de levedura hospedeira preferivelmente é uma que, como uma cepa tipo selvagem, contém pelo menos um gene D-lactato:ferricitocromo c oxidoredutase funcional nativo, se ela tiver que ser transformada com um gene D-LDH exógeno da maneira descrita a seguir.

Um L-lactato:ferricitocromo c oxidoredutase é um gene que codifica para uma enzima L-lactato:ferricitocromo c oxidoredutase funcional, que catalisa a metabólise de L-lactato em piruvato. Embora uma enzima como essa possa ajudar a célula a metabolizar a reação reversa de piruvato com L-lactato, ela é muito mais eficiente na catalisação da rea-

ção reversa de lactato de volta para piruvato, e assim as células tipo selvagem não produzem essencialmente nenhum L-lactato, a despeito da presença deste gene funcional. A enzima L-lactato:ferricitocromo c oxidoreductase é também conhecida pelo nome sistemático (S)-lactato:ferricitocromo-c 2-oxidoreductase ou como L-lactato deidrogenase (citocromo). O gene L-lactato:ferricitocromo c oxidoreductase é adequadamente um que tem uma região de codificação identificada como SEQ. ID. NO. 1 (o gene L-lactato:ferricitocromo c oxidoreductase *CYB2* de uma cepa *K. marxianus* tipo selvagem), SEQ. ID. NO. 79 (o gene *CYB2A* de uma cepa *I. orientalis* tipo selvagem) ou SEQ. ID. NO. 81 (o gene *CYB2B* de uma cepa *I. orientalis* tipo selvagem), ou é um gene que é pelo menos 40 %, preferivelmente pelo menos 75 %, mais preferivelmente pelo menos 90 % e ainda mais preferivelmente pelo menos 95 % homóloga a pelo menos um de SEQ. ID. NOs. 1, 79 ou 81. O gene L-lactato:ferricitocromo c oxidoreductase é adequadamente um que codifica um proteína com uma sequência de aminoácidos identificada como SEQ. ID. NO. 2, SEQ. ID. NO. 80 ou SEQ. ID. NO. 82, ou codifica para uma proteína que é pelo menos 40 %, preferivelmente pelo menos 75 %, mais preferivelmente pelo menos 90 % e ainda mais preferivelmente pelo menos 95 % homóloga a pelo menos uma de SEQ. ID. NOs. 2, 80 ou 82.

Um D-lactato:ferricitocromo c oxidoreductase é um gene que codifica para uma enzima D-lactato:ferricitocromo c oxidoreductase funcional, que catalisa o metabolismo de D-lactato com piruvato muito mais fortemente do que ele catalisa a reação reversa. O enzima D-lactato:ferricitocromo c

oxidoreductase é também conhecida pelo nome sistemático (R)-lactato:ferricitocromo-c 2-oxidoreductase ou como D-lactato deidrogenase (citocromo). O gene D-lactato:ferricitocromo c oxidoreductase é adequadamente um que tem uma região de codificação identificada como SEQ. ID. NO. 83 (o gene D-lactato:ferricitocromo c oxidoreductase (*DLD1*) de uma cepa *K. marxianus* tipo selvagem), o gene *DLD1* de *S. cerevisiae*, o gene *DLD1* de *I. orientalis*, ou é um gene que é pelo menos 40 %, preferivelmente pelo menos 75 %, mais preferivelmente pelo menos 90 % e ainda mais preferivelmente pelo menos 95 % homólogo a pelo menos um de tais genes. O gene D-lactato:ferricitocromo c oxidoreductase é adequadamente um que codifica uma proteína com uma sequência de aminoácidos identificada como SEQ. ID. NO. 84, uma proteína tendo a sequência de aminoácidos de uma proteína codificada pelo gene *DLD1* de *S. cerevisiae*, uma proteína tendo a sequência de aminoácido de uma proteína codificada pelo gene *DLD1* de *I. orientalis*, ou uma proteína que é pelo menos 40 %, preferivelmente pelo menos 75 %, mais preferivelmente pelo menos 90 % e ainda mais preferivelmente pelo menos 95 % homóloga a um desses.

Uma "construção" é uma sequência de DNA que é usada para transformar uma célula. A construção pode ser, por exemplo, na forma de um plasmídeo ou vetor circular, na forma de um plasmídeo ou vetor linearizado, pode ser uma porção de um plasmídeo ou vetor circular (tal como é obtido digerindo o plasmídeo ou vetor com uma ou mais enzimas de restrição), ou pode ser um produto PCR preparado usando um plasmídeo,

vetor ou DNA genômico como um gabarito.

O termo "nativo", quando usado aqui com relação a materiais genéticos (por exemplo, um gene, promotor, terminador ou outra sequência de DNA), refere-se a materiais genéticos que são encontrados (exceto de mutações indivíduo para indivíduo que não afetam a função) no genoma de células tipo selvagem desta espécie de levedura. "Capacidade nativa" (e suas variações tal como "nativamente capaz") indica a capacidade de células tipo selvagem realizarem a função indicada. Por exemplo, uma célula é nativamente capaz de metabolizar um açúcar em piruvato, se células tipo selvagem dessa espécie possuírem esta capacidade antes de quaisquer modificações genéticas. Um gene é considerado "funcional" dentro da célula se suas funções dentro da célula produzirem uma proteína ativa.

Nesta invenção, "exógeno" com relação a qualquer material genético significa que o material genético não é nativo para a célula hospedeira.

Células de levedura adequadas incluem aquelas de genera *Candida*, *Saccharomyces*, *Shizosaccharomyces*, *Kluyveromyces*, *Pichia*, *Issachenkia* e *Hansenula*. Célula de genera *Candida*, *Kluyveromyces*, *Saccharomyces* e *Issachenkia* são particularmente preferidas. Células especialmente preferidas são *C. sonorensis*, *K marxianus*, *K. thermotolerans*, *C. methanobosca*, *S. bulderi* e *I. orientalis*. Células mais preferidas são *K marxianus*, *S. cerevisiae*, *C. sonorensis* e *I. orientalis*. *I. orientalis* é referida algumas vezes como *Candida krusei* ou *Pichia kudriavzevii*. Cepas adequadas de *K. marxia-*

nus e *C. sonorensis* incluem aquelas descritas em WO 00/71738 A1, WO 02/42471 A2, WO 03/049525 A2, WO 03/102152 A2 e WO 03/102201A2. Uma cepa adequada de *I. orientalis* é cepa ATCC 32196.

5 A célula da invenção também contém pelo menos um funcional, gene lactato deidrogenase (LDH) exógeno integrado em seu genoma. Um gene *LDH* é um gene que codifica para uma enzima de lactato deidrogenase funcional. Genes *LDH* são específicos para a produção tanto de *L-LDH* quanto de *D-LDH*,
10 que respectivamente permitem à célula produzir tanto o enantiômero L- quanto D- ácido láctico (ou seus sais). É possível que o célula modificada da invenção contenha tanto os genes L- quanto *D-LDH*, e assim é capaz de produzir ambos enantiômeros de ácido láctico. Entretanto, prefere-se que apenas ge-
15 nes L- ou apenas *D-LDH* estejam presentes, assim a célula produz um produto de ácido láctico opticamente mais puro.

 Genes *LDH* adequados incluem aqueles obtidos a partir de fontes bacterianas, fúngicas, de levedura ou mamíferas. Exemplos de genes *L-LDH* específicos são aqueles obtidos
20 a partir de *L. helveticus*, *L. casei*, *B. megaterium*, *P. acidilactici*, e fontes bovinas. Exemplos de específicos genes *D-LDH* são aqueles obtidos a partir de *L. helveticus*, *L. johnsonii*, *L. bulgaricus*, *L. delbrueckii*, *L. plantarum* e *L. pentosus*. Genes funcionais que são idênticos ou pelo menos
25 80 % homólogos a qualquer desses genes *L-LDH* ou *D-LDH* são adequados. Os genes nativos obtidos a partir de qualquer dessas fontes podem ser submetidos a mutagênese, se necessário, para fornecer uma sequência de codificação começando

com o usual códon de partida eucariótico (ATG), ou com outros propósitos. Um gene L-LDH preferido é aquele obtido a partir de *L. helveticus* ou um que é pelo menos 80 %, 85 %, 90 % ou 95 % homólogo a tal gene. Um outro gene L-LDH preferido é aquele obtido a partir de *B. megaterium* ou um que é pelo menos 80 %, 85 %, 90 % ou 95 % homólogo a tal gene. Um gene D-LDH preferido é aquele obtido a partir de *L. helveticus* ou um que é pelo menos 80 %, 85 %, 90 % ou 95 % homóloga to tal gene.

Porcentagem de homologia de DNA, RNA ou de outro material genético e de seqüências de aminoácidos de proteína pode ser computado convenientemente usando suporte lógico BLAST versão 2.2.1 com parâmetros padrões. Seqüências tendo uma pontuação de identidades de pelo menos XX %, usando o algoritmo BLAST versão 2.2.1 com parâmetros padrões, são consideradas pelo menos XX % homólogas.

Particularmente, genes LDH adequados incluem aqueles que codificam para uma enzima com uma seqüência de aminoácidos que tem uma pontuação de identidades de pelo menos 60 %, especialmente pelo menos 80 %, 85 % ou 95 %, comparado com SEQ. ID. NO. 45 de WO 03/049525 ou comparado com SEQ. ID. NO. 49 de WO 031049525. Genes LDH particularmente adequados também incluem aqueles que codificam uma enzima tendo uma seqüência de proteína que é pelo menos 60 %, 80 %, 85 % ou 95 % homóloga a SEQ ID. NO. 46 ou 50 de WO 03/049525.

A célula transformada pode conter um único gene LDH ou múltiplos genes LDH, tal como de 1-10 genes LDH, especialmente de 1-5 genes LDH. Quando a célula transformada

contém múltiplos genes *LDH*, os genes individuais podem ser cópias do mesmo gene, ou incluir cópias de dois ou mais diferentes genes *LDH*. Múltiplas cópias do gene *LDH* exógeno podem ser integradas em um locus único (assim, eles ficam
5 adjacentes um ao outro), ou em diversos loci dentro do genoma da célula hospedeira.

O gene *LDH* exógeno está sob o controle transcripcional de um ou mais promotores e um ou mais terminadores, ambos os quais são funcionais na célula de levedura modificada. Da maneira aqui usada, o termo "promotor" refere-se a
10 uma seqüência não transcrita localizada a montante (isto é, 5') do códon de início de tradução de um gene estrutural (geralmente em cerca de 1 a 1.000 bp, preferivelmente 1-500 bp, especialmente 1-100 bp) e que controla o início de
15 transcrição do gene estrutural. Similarmente, o termo "terminador" refere-se a uma seqüência não transcrita localizada a jusante (isto é, 3') para o códon final de tradução de um gene estrutural (geralmente em cerca de 1 a 1.000 bp, mais tipicamente 1-500 pares base e especialmente 1-100 pares ba-
20 se) e que controla o final de transcrição do gene estrutural. Um promotor ou terminador é "operacionalmente ligado" a um gene estrutural se sua posição no genoma com relação à-
aquele do gene estrutural for tal que o promotor ou terminador, conforme possa ser o caso, realiza sua função transcripcional de controle.
25

Seqüências promotoras e terminadoras podem ser nativas ou exógenas para a célula hospedeira.

Um tipo adequado de promotor é pelo menos 50 %, 70

%, 90 %, 95 % ou 99 % homólogo a um promotor que é nativo a um gene de levedura. Um tipo mais adequado de promotor é pelo menos 50 %, 70 %, 90 %, 95 % ou 99 % homólogo a um promotor para um gene que é nativo para a célula hospedeira. Promotores particularmente usados incluem promotores para piruvato descarboxilase de levedura (*PDC1*), quinase de fosfoglicerato (*PGK*), xilose redutase (*XR*) xilitol deidrogenase (*XDH*), L-(+)-lactato-citocromo c oxidoreductase (*CYB2*) e genes de fator-1 de prolongamento de transcrição (*TEF1*) e fator-2 de prolongamento de transcrição (*TEF2*), especialmente a partir dos genes respectivos da célula hospedeira. Um promotor especialmente usado inclui a porção funcional de um promotor para um gene *PDC1*, *PGK*, *TEF1* ou *TEF2* da célula hospedeira ou pelo menos 80 %, 85 %, 90 % ou 95 % homóloga a um promotor como esse.

Um tipo adequado de terminador é pelo menos 50 %, 70 %, 90 %, 95 % ou 99 % homólogo a um terminador para um gene que é nativo de uma célula de levedura. O terminador pode ser pelo menos 90 %, 95 % ou 99 % homólogo a um terminador para um gene que é nativo da célula hospedeira. Terminadores particularmente usados incluem terminadores para piruvato descarboxilase de levedura (*PDC1*), xilose redutase, (*XR*), xilitol deidrogenase (*XDH*), gene L-lactato:ferricitocromo c oxidoreductase ou genes iso-2-citocromo c (*CYC*), ou um terminador a partir da família galactose de genes em levedura, particularmente o assim chamado terminador *GAL10*. Um terminador especialmente preferidas inclui uma porção funcional de um terminador para um gene

PDC1 da célula hospedeira ou é pelo menos 80 %, 85 %, 90 % ou 95 % homólogo a isto.

O uso de promotores e terminadores nativos (para a célula hospedeira), juntamente com as respectivas regiões
5 flanqueadoras a montante e a jusante, pode permitir a integração visada do gene *LDH* em loci específico do genoma da célula hospedeira, e integração simultânea do gene *LDH* e deleção ou ruptura de um outro gene nativo, tal como, por exemplo, um gene *PDC1*.

10 Quando múltiplos genes *LDH* exógenos são introduzidos na célula hospedeira, é possível para os os diferentes genes *LDH* fiquem sob o controle de diferentes tipos de promotores e/ou terminadores.

O gene *LDH* exógeno pode ser integrado aleatoriamente
15 mente no genoma da célula hospedeira ou inserido em uma ou mais localizações visadas. Exemplos de localizações visadas incluem o locus de um ou mais genes que são desejavelmente deletados ou interrompidos, tais como a de um gene *PDC1* ou a do gene L- ou D-lactato:ferricitocromo c oxidoreductase.

20 Em modalidades em que a célula contém um cassete do gene *L-LDH* funcional, a célula modificada da invenção mais preferivelmente inclui uma deleção ou ruptura de pelo menos um gene L-lactato:ferricitocromo c oxidoreductase nativo. "Deletar ou interromper" significa que toda a região de
25 codificação do gene é eliminada (deleção), ou o gene, seu promotor e/ou sua região terminadora é modificado (tal como por deleção, inserção ou mutação) de maneira que o gene tanto não produza mais a proteína ou uma versão ativa da prote-

ina, quanto produza uma enzima com atividade severamente reduzida (pelo menos 75 % reduzida, preferivelmente pelo menos 90 % reduzida). A deleção ou ruptura pode ser realizada por métodos de engenharia genética, evolução forçada ou mutagenese, seguido por seleção ou classificação apropriada para identificar os mutantes desejados. Se a célula contém gene L-lactato:ferricitocromo c oxidoredutases múltiplo (tanto cópias múltiplas do mesmo gene quanto dois ou mais diferentes L-lactato:ferricitocromo c oxidoredutase) parálogo, pelo menos um daqueles genes é deletado ou interrompido, e é preferível deletar ou interromper todos aqueles gene L-lactato:ferricitocromo c oxidoredutases.

Em modalidades onde a célula contém um cassete do gene *D-LDH* funcional, a célula modificada da invenção preferivelmente inclui uma deleção ou ruptura de pelo menos um gene D-lactato:ferricitocromo c oxidoredutase nativo. Como antes, é preferível deletar ou interromper todos aqueles genes quando múltiplas cópias ou alelos do gene estão presentes na cepa tipo selvagem.

A célula de levedura geneticamente modificada da invenção pode incluir modificações genéticas adicionais que fornecem um ou mais atributos desejáveis à célula. Uma modificação adicional de particular interesse inclui uma deleção ou ruptura de gene(s) de piruvato dicarboxilato, reduzindo dessa forma a capacidade da célula produzir etanol.

Uma outra modificação adicional de particular interesse é um (ou mais) que confere individual ou coletivamente à célula a capacidade de fermentar açúcares de pentose

em produtos de fermentação desejáveis. Entre o último tipo de modificações estão (1) inserção de um gene de xilose isomerase funcional, (2) uma deleção ou ruptura de um gene nativo que produz uma enzima que catalisa a conversão de xilose em xilitol, (3) uma deleção ou ruptura de um gene de xilitol deidrogenase funcional e/ou (4) modificações que fazem com que a célula sobreexpresse um xiluloquinase funcional. Métodos para introduzir aquelas modificações em células de levedura são descritos, por exemplo, em WO 04/099381, aqui incorporado pela referência.

Uma terceira modificação de particular interesse é a integração de uma ou mais cassetes marcadores de seleção funcional no genoma de célula hospedeira, para permitir seleção de transformantes bem sucedidos.

Deleção ou ruptura do gene L- ou D-lactato:ferricitocromo c oxidoredutase pode ser realizada em uma variedade de maneiras, incluindo por métodos de engenharia genética evolução forçado, ou mutagênese, acoplado com seleção ou classificação para identificar os mutantes desejados.

Métodos de mutagênese EMS e ultravioleta podem ser usados para deletar ou interromper o gene L- ou D-lactato:ferricitocromo c oxidoredutase. Estes métodos são bem conhecidos. Nestes métodos, células são expostas a radiação ultravioleta ou uma substância mutagênica, em condições suficientes para atingir uma alta taxa de morte (60-99 %, preferivelmente 90-99 %) da célula. Células sobreviventes são em seguida plaqueadas e selecionadas ou classificadas para

célula tendo gene L- ou D-lactato:ferricitocromo c oxidore-
dutase deletado ou interrompido. No caso de uma deleção ou
ruptura de um gene L-lactato:ferricitocromo c oxidoreductase,
isto é convenientemente feito selecionando para célula que
5 são resistentes a ácido glicólico.

Engenharia genética da célula hospedeira (para de-
letar ou interromper o gene L- ou D-lactato:ferricitocromo c
oxidoreductase ou fazer outras modificações) é conveniente-
mente realizada em uma ou mais etapas por meio do projeto e
10 construção de construções apropriados e transformação da cé-
lula hospedeira com aqueles construções. Eletroporação e/ou
substância química (tal como cloreto de cálcio - ou acetato
da base de lítio) métodos de transformação podem ser usados.
As construções em cada caso podem tanto ser cortadas com en-
15 zimas de restrição particulares, usadas como DNA circular ou
usadas como o gabarito para gerar um produto PCR.

Em um método de engenharia genética adequado para
deletar ou interromper o gene L- ou D-lactato:ferricitocromo
c oxidoreductase, uma construção de deleção é convenientemen-
20 te montado primeiro clonando duas seqüências de DNA não con-
tínuas do gene de L- ou D-lactato:ferricitocromo c oxidore-
ductase nativo e/ou suas regiões flangeadoras a montante ou
a jusante. Neste contexto, "não contínuo" significa que as
seqüências de DNA não são imediatamente adjacentes umas com
25 as outras no genoma do tipo selvagem, mas, em vez disso, são
separadas umas das outras no genoma do tipo selvagem por al-
gum outro material genético (que pode ser tão pequeno quanto
um único par base). Entre elas, estas seqüências não conti-

nuas podem incluir pelo menos uma porção da região de codificação do gene e/ou suas regiões promotoras e terminadoras. Prefere-se que uma das seqüências inclua uma região flanqueadora 5' do gene L- ou D-lactato:ferricitocromo c oxidoredutase (incluindo toda ou uma porção da região promotora e toda ou uma porção da seqüência de codificação de L- ou D-lactato:ferricitocromo c oxidoredutase), e que as outras seqüências incluam uma região flanqueadora 3' do gene L- ou D-lactato:ferricitocromo c oxidoredutase (incluindo toda ou uma porção da região terminadora e/ou toda ou uma porção da seqüência de codificação de o L- ou D-lactato:ferricitocromo c oxidoredutase). As seqüências não contínuas podem ser clonadas separadamente. Alternativamente, toda a região, incluindo os flancos 5' e 3' e o gene L- ou D-lactato:ferricitocromo c oxidoredutase podem ser clonados com deleção subsequente de pelo menos parte do gene para torná-los não funcionais. As seqüências não contínuas podem incluir toda ou uma porção das regiões promotoras e terminadoras de L-ou D-lactato:ferricitocromo c oxidoredutase, respectivamente. Uma construção de deleção é em seguida produzida contendo as duas seqüências não contínuas orientadas na mesma direção, e a construção é usada para transformar a célula hospedeira. A construção inserida no locus do gene L- ou D-lactato:ferricitocromo c oxidoredutase em algum dos transformantes. Um evento de recombinação duplo cruzado homólogo resulta em uma deleção do gene funcional L- ou D-lactato:ferricitocromo c oxidoredutase em algum de o transformantes. Transformantes bem sucedidos podem ser (1) sele-

cionados pela resistência ao ácido glicólico (no caso de uma ruptura ou deleção do gene L-lactato:ferricitocromo c oxidoredutase) ou (2) classificados pela falta de capacidade de crescer no lactato como uma única fonte de carbono ou com base nas características transmitidas para a célula através do uso de um marcador de seleção da maneira descrita mais completamente a seguir. Alternativamente, classificação pode ser realizada por PCR. Análise de PCR ou Southern pode ser usada para confirmar se ocorreu a deleção desejada.

10 A construção de deleção de L- ou D-lactato:ferricitocromo c oxidoredutase pode também incluir um ou mais genes estruturais funcionais, notavelmente um cassete do gene LDH, inserido a jusante das porções flanqueadoras 5' do gene L- ou D-lactato:ferricitocromo c oxidoredutase e a montante da porção flanqueadora 3' do gene L- ou D-lactato:ferricitocromo c oxidoredutase. Esta abordagem permite que a inserção do gene estrutural no locus do gene L- ou D-lactato:ferricitocromo c oxidoredutase. Em um caso preferido, o gene estrutural é um cassete do gene LDH, em
15 cujo caso deleção do gene L- ou D-lactato:ferricitocromo c oxidoredutase e inserção do cassete do gene *LDH* no locus do gene L- ou D-lactato:ferricitocromo c oxidoredutase pode ser realizado em uma etapa de transformação única. Transformantes bem sucedidos podem ser identificados por métodos de hibridização PCR ou Southern, ou detectando a atividade do gene estrutural inserido usando qualquer método de ensaio conveniente. No caso onde gene estrutural é um gene *LDH*, transformantes bem sucedidos podem ser identificados por sua ca-
20
25

pacidade de produzir ácido láctico. Transformantes em que um gene L-lactato:ferricitocromo c oxidoreductase deletado u interrompido pode ser selecionado com base nas sua resistência ao ácido glicólico.

5 A construção de deleção de L- ou D-lactato:ferricitocromo c oxidoreductase pode também incluir um gene marcador de seleção em vez de ou em adição ao gene estrutural. O uso de um gene marcador de seleção tem a vantagem de introduzir um meio adicional de selecionar transformantes bem sucedidos, mas não é exigido. Um "gene marcador de seleção" é um gene que codifica uma proteína necessária para a sobrevivência e/ou crescimento da célula transformada em um meio de cultura seletivo. Genes marcadores de seleção típicos codificam proteínas que (a) conferem resistência a antibióticos ou outras toxinas, (tal como zeocina (10 *Streptoalloteichus hindustanus* gene resistente a ble bleomicina), G418 (gene resistente a canamicina de Tn903) ou higromicina (gene resistente a antibiótico aminoglicosídeo de *E. coli*)), (b) complementar deficiências auxotrófica da célula (tal como deficiência leucina no aminoácido (gene *K. marxianus* LEU2) ou deficiência de uracila (por exemplo, gene *K marxianus* ou *S. cerevisiae* URA3)), (c) permite que a célula sintetize nutrientes críticos não disponíveis no meio simples ou (d) confere a capacidade da célula de crescer em (20 uma fonte de carbono particular, (tais como os genes MEL1 ou MEL5 de *S. cerevisiae*, que confere a capacidade de crescer na melibiose como uma única fonte de carbono. Marcadores de seleção preferidos incluem o gene resistente a zeocina, gene (25

resistente a G418, um genes *MEL1* e *MEL5* e gene resistente a higromicina (*hph*).

O cassete marcador de seleção incluirá adicionalmente seqüências promotoras e terminadoras, operacionalmente ligadas ao gene marcador de seleção, e que são operáveis na célula hospedeira. Promotores adequados incluem aqueles descritos anteriormente com relação ao gene *LDH*, bem como outros, tais como os que estão descritos em WO 99114335, WO 00171738, WO 02142471, WO 03/102201, WO 03/102152 e WO 031049525. Um promotor especialmente preferido é um promotor *TEF1*, *PGK* ou *PDC1* (ou porção funcional destes) de uma espécie de levedura, especialmente a cepa hospedeira, ou uma seqüência que é pelo menos 80, 85, 90 ou 95 % homóloga a um promotor *TEF1*, *PGK* ou *PDC1* como esse. Terminadores adequados incluem aqueles descritos anteriormente com relação a genes *LDH*.

O gene marcador de seleção é similarmente posicionado na construção a jusante da porção flanqueadora 5' do gene L- ou D-lactato:ferricitocromo c oxidoreductase e a montante da porção flanqueadora 3' do gene L- ou D-lactato:ferricitocromo c oxidoreductase. Em uma porção da célula transformada com esta construção, um evento de recombinação homóloga faz com que o marcador de seleção cassete (e cassete de gene estrutural, se presente) torne-se integrado no locus do gene *CYB2* ou *DLD1* respectivamente. Transformantes bem sucedidos podem ser selecionados com base nas características transmitidas a eles pelo gene marcador de seleção. Transformantes em que o gene L-lactato:ferricitocromo

c oxidoredutase torna-se deletado ou interrompido pode ser selecionado com base na sua resistência ao ácido glicólico e/ou classificado por sua incapacidade de crescer no lactato como uma única fonte de carbono. métodos de análise PCR e Southern podem ser usados como antes para caracterizar os transformantes.

Métodos para transformar cepas de levedura para inserir um gene *LDH* exógeno são descrito em WO 99114335, WO 00171738, WO 02142471, WO 03/102201, WO 03/102152 e WO 031049525; estes métodos são geralmente aplicáveis para transformar célula de acordo com esta invenção. As construções podem tanto ser cortadas com enzimas de restrição particulares, usadas como DNA circular, quanto usadas como o gabário para gerar um produto PCR.

Construções para inserção visada do cassete do gene *LDH* no locus do gene da célula hospedeira L- ou D-lactato:ferricitocromo c oxidoredutase (com simultânea deleção do gene L- ou D-lactato:ferricitocromo c oxidoredutase) são descritas anteriormente. Quando não se deseja inserir o cassete do gene *LDH* no locus do gene L- ou D-lactato:ferricitocromo c oxidoredutase, vetores projetados de forma diferente são usados.

Geralmente, um gene *LDH* exógeno é inserido preparando uma construção que contém o cassete do gene *LDH* (isto é o gene estrutural e seqüências promotoras e terminadoras associadas). A construção pode conter sítios de restrição de vários tipos para linearização ou fragmentação. A construção pode conter adicionalmente uma porção espinha da dorsal (tal

como para propagação em *E. coli*), muitas das quais são convenientemente obtidas a partir de levedura ou vetores bacterianos comercialmente disponíveis.

A construção de inserção de *LDH* preferivelmente contém um ou mais cassetes de gene marcador de seleção. Integração visada do cassete do gene *LDH* pode ser realizada criando uma construção tendo as seqüências não contínuas que são homólogas aos flancos a montante (5'-) e a jusante (3'-) de um gene visado. Essas regiões não contínuas têm cada qual adequadamente a partir de cerca de 50 a 3.000 bp em comprimento, especialmente cerca de 200-2.000 bp em comprimento. Qualquer uma ou ambas dessas seqüências não contínuas pode incluir uma porção da região de codificação do gene visado bem como uma porção ou toda as regiões promotoras ou terminadoras respectivas. O cassete do gene *LDH* é arranjado na construção entre as duas seqüências não contínuas. Um gene visado preferido (sem ser gene L- ou D-lactato:ferricitocromo c oxidoreductase) é um gene piruvato decarboxilato (*PDC*), como a célula da invenção preferivelmente tem uma ruptura ou deleção de pelo menos um gene(s) *PDC* nativo(s). Conseqüentemente, seqüências não contínuas preferidas são tomadas a partir do locus de um gene *PDC* nativo. Entretanto, outros genes nativos podem servir como alvos para inserção do cassete do gene *LDH*, usando as seqüências não contínuas a partir do locus do gene visado na construção de inserção de *LDH* de uma forma análoga.

O cassete *LDH* (incluindo promotores e terminadores associados, se diferentes daqueles do gene visado) e marca-

dor(s) de seleção (com promotores e terminadores associados, conforme possa ser necessário) residirá na construção de inserção de *LDH* entre as regiões que são homólogas aos flancos a montante e a jusante do gene visado, a jusante do
5 flanco homólogo 5' e a montante do flanco homólogo 3'.

Se for usado um cassete marcador de seleção, ele similarmente reside no vetor entre as duas seqüências não contínuas tomadas a partir do locus do gene visado, da maneira supradescrita.

10 Uma porção de células transformadas com esta construção passará por um evento de recombinação homóloga em que o gene visado é deletado ou interrompido, e o cassete *LDH* (e cassete marcador de seleção, se presente) é integrado no locus do gene visado. No caso preferido em que o gene visado
15 é um gene *PDC*, os transformantes terão uma deleção ou ruptura do gene *PDC*. Transformantes bem sucedidos podem ser selecionados de uma maneira conhecida, usando abordagens da maneira descrita anteriormente.

Quando a inserção de *LDH* e a deleção ou ruptura do
20 gene L- ou D-lactato:ferricitocromo c oxidoreductase são feitos sequencialmente, as etapas podem ser conduzidas em qualquer ordem, isto é, a deleção do gene L- ou D-lactato:ferricitocromo c oxidoreductase pode ser realizada primeiro, seguido pela inserção de *LDH*, ou a inserção de *LDH*
25 pode ser realizada primeiro, seguida pela deleção ou ruptura do gene L- ou D-lactato:ferricitocromo c oxidoreductase.

Quando um gene L-lactato:ferricitocromo c oxidoreductase é deletado primeiro, torna-se possível usar o gene L-

lactato:ferricitocromo c oxidoreductase em si como um marcador de seleção durante a inserção de *LDH* e/ou outras transformações subseqüentes. Em cada um dos casos o vetor de transformação *LDH* inclui um cassete do gene L-lactato:ferricitocromo c oxidoreductase, em que o gene o L-lactato:ferricitocromo c oxidoreductase é operacionalmente ligado a uma seqüência promotora e terminadora e terminadora. O cassete do gene L-lactato:ferricitocromo c oxidoreductase é localizado na vetor de inserção de *LDH* entre as seqüências não contínuas a partir do locus do gene visado. Uma porção dos transformantes conterão tanto o cassete *LDH* quanto o cassete de L-lactato:ferricitocromo c oxidoreductase no locus do gene visado (que neste caso é preferivelmente um gene *PDC* ou outro gene que é desejavelmente deletado ou interrompido). Transformantes bem sucedidos podem ser selecionados com base na sua capacidade de crescer em lactato como uma única fonte de carbono, ou por outros métodos da maneira descrita anteriormente.

Caso queira usar o gene L-lactato:ferricitocromo c oxidoreductase como um gene marcador de seleção uma segunda vez de acordo com esta invenção, é necessário que o primeiro gene marcador L-lactato:ferricitocromo c oxidoreductase subseqüentemente seja deletado ou interrompido. Uma maneira conveniente de realizar isto é projetar uma construção de maneira que o cassete do gene L-lactato:ferricitocromo c oxidoreductase seja flanqueado pelas seqüências repetitivas diretas. Seqüências repetitivas diretas são seqüências de DNA idênticas, que podem ser nativas, mas preferivelmente não

são nativas para a célula hospedeira, e que são orientadas na mesma direção com relação uma a outra na construção. As seqüências repetitivas diretas são vantajosamente cerca de 50-1.500 bp em comprimento. Não é necessário que as seqüências repetitivas diretas codifiquem para qualquer coisa. Em um pequeno número de transformantes, esta construção permite que ocorra um evento de recombinação homóloga, resultando em uma deleção do gene marcador L-lactato:ferricitocromo c oxidoredutase e uma das seqüências repetitivas diretas. É normalmente necessário crescer transformantes no não meio seletivo para permitir que ocorra a recombinação homóloga espontânea entre as seqüências repetitivas diretas. Também como antes, célula em que o gene L-lactato:ferricitocromo c oxidoredutase foi deletado pode ser selecionada com base na sua resistência a ácido glicólico ou classificada usando sua incapacidade de crescer em lactato como uma única fonte de carbono.

Prefere-se deletar ou interromper um gene PDC nativo na célula hospedeira, de maneira que a célula transformada tenha uma capacidade reduzida para produzir etanol. Se a célula hospedeira contiver múltiplos genes PDC, prefere-se especialmente deletar ou interromper todos os genes PDC, embora seja possível deletar menos do que todos tais genes PDC. Deleção de PDC pode ser realizada usando métodos análogos àqueles descritos em WO 99114335, WO 021424 71, WO 031049525, WO 03/102152 e WO 031102201. Deleção de PDC pode também ser realizada com inserção simultânea de um cassete do gene LDH ou outro estrutural ou cassete do gene marcador

de seleção. Em um método de interesse particular (1) seqüências não contínuas a partir do locus do(s) gene(s) *PDC* são clonados, de modo opcional juntamente com uma porção do gene *PDC* funcional; (2) uma construção contendo as seqüências não

5 contínuas é produzida (de modo opcional contendo uma porção não funcional do gene *PDC* estrutural), e (3) a célula hospedeira é transformada com a construção. Um evento de recombinação homóloga resulta em uma deleção ou ruptura do gene *PDC* funcional em uma porção dos transformantes. Isto pode ser

10 repetido se necessário para deletar ou interromper múltiplos genes *PDC*. Em alguma espécie de levedura, tal como *I. orientalis*, existem múltiplos genes *PDC* ou alelos que são bastante homólogos. Observou-se que, em tais casos, seqüências não contínuas tomadas a partir do locus de qualquer gene podem

15 ser usadas na construção para deletar ou interromper ambos os genes *PDC*. A construção usada para interromper o(s) gene(s) *PDC* pode incluir um ou mais funcional genes estruturais inseridos a jusante da porção flanqueadora 5' do gene *PDC* nativo e a montante das porções flanqueadoras 3' do gene *PDC*

20 nativo. O gene estrutural preferivelmente é parte de um cassete que inclui seqüências promotoras e terminadoras funcionais operacionalmente ligadas ao gene estrutural. Esta abordagem permite a deleção do gene *PDC* e inserção do gene funcional cassete em uma etapa de transformação única. A construção

25 pode incluir um cassete do gene marcador de seleção em substituição, ou em adição ao gene estrutural. Novamente, o cassete do gene marcador de seleção é posicionado no vetor como antes, entre as seqüências não contínuas tomadas a par-

tir do locus do gene *PDC* sendo visadas, e foi inserido no locus do gene *PDC* funcional em uma porção de transformantes.

Métodos adequados para inserir um gene de xilose isomerase funcional, deletar ou interromper um gene nativo
5 que produz uma enzima que catalisa a conversão de xilose em xilitol, deletar ou interromper um gene de xilitol deidrogenase funcional, modificar a célula que sobreexpressa um xiluloquinase funcional são descritos, por exemplo, em WO 04/099381, aqui incorporados pela referência.

10 No processo de fermentação da invenção, a célula da invenção é cultivada em um meio de fermentação que inclui um açúcar que é fermentável pela célula transformada. O açúcar pode ser um açúcar hexose tais como glicose, glicano ou outro polímero de glicose, oligômeros de glicose tais como
15 maltose, maltotriose e isomaltotriose, panose, frutose, e oligômeros de frutose. Se a célula de forma nativa tiver capacidade ou for modificada para conferir a fermentação de açúcares de pentose, o meio de fermentação pode incluir um açúcar de pentose tal como xilose, xilano ou outro oligômero
20 de xilose. Tais açúcares de pentose são adequadamente hidrolisados de uma biomassa contendo hemicelulose. No caso de açúcares oligoméricos, pode ser necessário adicionar enzimas ao caldo de fermentação a fim de digerir-los com o açúcar monomérico correspondente para fermentação pela célula.

25 O meio tipicamente conterá nutrientes exigidos pela célula particular, incluindo uma fonte de nitrogênio (tal como aminoácidos, proteínas, fontes de nitrogênio inorgânico tal como sais de amônia ou amônio, e similares), e várias

vitaminas, minerais e similares.

Outras condições de fermentação, tais como temperatura, densidade celular, seleção de substrato(s), seleção de nutrientes, e similares não são considerados críticos para a invenção e são geralmente selecionados para fornecer um processo econômico. Temperaturas durante cada fase de crescimento e a fase de produção podem variar acima da temperatura de congelamento do meio até cerca de 50 °C, embora isto dependa até certo ponto da capacidade da cepa de tolerar temperaturas elevadas. Uma temperatura preferida, particularmente durante a fase de produção, é cerca de 30-45 °C.

Durante a fase de produção, a concentração da célula no meio de fermentação é tipicamente na faixa de cerca de 0,1-20, preferivelmente cerca de 0,1-5, ainda mais preferivelmente cerca de 1-3 g células secas/litro do meio de fermentação.

A fermentação pode ser conduzida aerobicamente, microaerobicamente ou anaerobicamente. Se desejado, a taxa de captação de oxigênio pode ser usada como um controle do processo, da maneira descrita em WO 03/102200. A célula da invenção apresenta uma boa capacidade de fermentar açúcares no ácido lático ou misturas de ácido lático/etanol, em boas produtividades volumétricas e específicas ainda em condições anaeróbicas.

O meio pode ser tamponado durante a fase de produção da fermentação de maneira que o pH é mantido em uma faixa de cerca de 3,5 a cerca de 9,0, tal como de cerca de 4,5 a cerca de 7,0. Agentes de tamponamento adequados são mate-

riais básicos que neutralizam ácido lático a medida que ele é formado, e incluem, por exemplo, hidróxido de cálcio, carbonato de cálcio, hidróxido de sódio, hidróxido de potássio, carbonato de potássio, carbonato de sódio, carbonato de amônio, amônia, hidróxido de amônio e similares. De um modo geral, aqueles agentes de tamponamento que foram usados em processos de fermentação convencionais também são adequados aqui.

Em uma fermentação tamponada, produtos de fermentação ácidos tal como ácido lático são neutralizados a medida que eles são formados no sal de lactato correspondente. Recuperação do ácido envolve então regenerar o ácido livre. Isto é feito tipicamente removendo a célula e acidulando o caldo de fermentação com um ácido forte tal como ácido sulfúrico. Um subproduto do sal é formado (gesso no caso onde um sal de cálcio é o agente neutralizante e ácido sulfúrico é o agente acidulante), que é separado a partir do ácido. O ácido é em seguida recuperado através de técnicas tal como extração líquido-líquido, destilação, absorção, etc., tal como são descritos em T.B. Vickroy, Vol. 3, Chapter 38 of Comprehensive Biotechnology, (ed. M. Moo-Young), Pergamon, Oxford, 1985; R. Datta, et al., FEMS Microbiol. Rev., 1995, 16:221-231; patentes U.S. nos. 4.275.234, 4.771.001, 5.132.456, 5.420.304, 5.510.526, 5.641.406 e 5.831.122 e WO 93/00440.

Prefere-se, entretanto, permitir que o pH do meio de fermentação goteje de um pH de partida que é tipicamente 5,5 ou mais, para o pKa acima do produto de fermentação do

ácido, tal como na faixa de cerca de 1,5 a cerca de 3,5, na faixa de cerca de 1,5 a cerca de 3,0, ou na faixa de cerca de 1,5 a cerca de 2,5. A célula desta invenção tem uma capacidade inesperada de crescer e produzir bem, mesmo no meio de fermentação não tamponado onde o pH final cai abaixo de 3,5, abaixo de 3,0, abaixo de 2,5, e ainda abaixo de 2,0.

É também possível conduzir a fermentação mantendo o pH no pKa do ácido lático ou abaixo dele por todo o processo. Em um caso como este, o pH do caldo de fermentação é ajustado até o pKa de ácido lático ou abaixo dele antes do início do processo de fermentação ou nele, e mantido em tal nível como a fermentação procede. Nesta modalidade, o pH é preferivelmente mantido na faixa de cerca de 1,5 a cerca de 3,5, na faixa de cerca de 1,5 a cerca de 3,0, ou na faixa de cerca de 1,5 a cerca de 2,5.

Recuperação de ácido lático a partir de um meio de fermentação de baixo pH pode ser conduzida usando métodos tais como aqueles descritos em na patente U.S. 6.229.046.

O processo da invenção pode ser conduzido continuamente, em lotes, ou em alguma combinação destes.

A deleção ou ruptura do gene L- ou D-lactato:ferricitocromo c oxidoreductase, como pode ser o caso, pode levar a diversos benefícios. Tituladores de ácido lático são freqüentemente melhorados. Isto é observado especialmente quando a célula hospedeira é uma cepa de *Kluyveromyces*, tal como *Kluyveromyces marxianus*. Rendimentos de lactato bem como taxas de produção de lactato são igualmente aumentadas, comparado com cepas similares que não têm a de-

leção de L-lactato:ferricitocromo c oxidoreductase. Novamente, isto é especialmente observado em cepa de *Kluyveromyces* modificada. Uma vantagem adicional é que baixa produção de piruvato é observada.

Os exemplos seguintes são fornecidos para ilustrar a invenção, mas não são destinados a limitar seu escopo. Todas as partes e porcentagens são em peso, a menos que de outra forma indicada.

Exemplo 1A: Mutagênese de cepa CD607 de *K. marxianus* e seleção de cepa mutante (CD635) tendo resistência ao ácido glicólico.

Cepa CD607 de *K. marxianus* é descrita no Exemplo 3D de WO 031102152. Esta cepa tem uma deleção de seu gene piruvato descarboxilase e uma inserção de um gene de lactato deidrogenase exógeno neste locus. Célula de cepa CD607 são submetidas a mutagênese por meio de exposição a luz ultravioleta ou metanossulfonato de etila, de acordo com o seguintes protocolos:

Mutagênese UV

Células de uma placa YPD fresca (Difco #0427-17, Sparks, MD, contendo extrato de levedura 10 g/L, peptona 20 g/L, dextrose 20 g/L e ágar 15 g/L) são ré-suspensas em 2 mL de peptona de levedura + e mais 50 g/L de glicose para um OD₆₀₀ aproximado de 6. Dez alíquotas de 125 µL desta suspensão da célula são pipetadas em dez poços de um placa microtituladora de 96 poços de 300 µL. A placa microtituladora é exposta a 12.500 µJoule/cm² de luz UV até 90-99 % de morte da célula. Duas alíquotas 125 µL adicionais da suspensão da

célula são em seguida adicionadas à placa microtituladora como controles tratados simulados. As suspensões da célula de um dos doze poços são transferidas para tubos de microfuga de 1,5 mL. Os doze tubos de microfuga são em seguida incubados no escuro por toda a noite a 30 °C com agitação (225 rpm) para permitir que a célula se recupere antes de plaqueamento em placas de seleção.

Alíquotas de 100 µL de nove das suspensões da célula tratadas por UV e de uma das suspensões da célula tratada simulada são em seguida plaqueadas em uma placa de ágar de dextrose de batata (PDA, Difco Laboratories, Cat. No. 0013-17, Sparks, MD) + 15 g/L de ácido glicólico para selecionar cepas resistentes a ácido glicólico. Estas placas são incubadas a 30 °C por diversos dias até aparecerem colônias. Uma colônia única é isolada de cada uma das nove placas para análise posterior.

Mutagênese EMS

Células de uma placa YPD fresca são ré-suspensas em 2 mL de salina tamponada com fosfato com um OD₆₀₀ aproximado de 6. Doze alíquotas de 200 µL desta suspensão da célula são pipetadas em doze tubos com tampa de pressão de 14 mL, e 4 µL de metanossulfonato de etila (EMS, Sigma Chemical Co., St. Louis, MO, catalog # M0880, 1,17 g/mL de solução) são adicionados a dez dos doze tubos. Os dois tubos restantes servem como controles tratados simulados. Os tubos são em seguida incubados a 30 °C com agitação (225 rpm) por 30 minutos, novamente até 90-99 % de morte da célula. Depois da exposição a EMS, as células dos doze tubos são precipitadas,

lavadas duas vezes com Na₂S₂O₃ 5,0 % para neutralizar o EMS e lavadas uma vez com água. As células são ré-suspensas em 100 µL de meio YPD e incubadas por 6 horas a 30 °C com agitação (225 rpm) para permitir que a célula se recupere antes
5 de plaqueamento em placas de seleção.

Após a recuperação, 100 µL de nove das suspensões da célula tratadas com EMS e de uma das suspensões da célula tratada simulada são plaqueadas em uma placa PDA + 15 g/L de ácido glicólico para selecionar a resistência ao ácido gli-
10 cólico. Estas placas são incubadas a 30 °C por diversos dias até aparecerem colônias. Uma colônia única é isolada para análise posterior.

Classificação da célula mutagenizada é feita como a seguir. Aproximadamente 2×10^8 células que foram mutage-
15 nizadas em um dos processos apenas descritos são plaqueadas em (PDA) placas contendo 15 g/L de ácido glicólico e incubadas a 30 °C. Colônias que crescem nestas placas crescem por toda a noite em frascos de agitação defletidos a 30 °C e agitação 225 rpm em YP (extrato de levedura 10 g/L, peptona 20
20 g/L) + 100 g/L de glicose sem tampão. Frascos de produção são em seguida inoculados com 2 g/L de peso seco de célula desses frascos de agitação. Os frascos de produção são cultivados a 30 °C e 70 rpm agitação em YP + 50 g/L de glicose. Amostras são extraídas periodicamente para medir glicose,
25 lactato, etanol e piruvato por HPLC usando métodos tal como descrito no Exemplo 1M de WO 03/102201.

Uma cepa é designada cepa CD635. Cepa CD635 resulta do tratamento mutagênese UV descrito anteriormente. Cepa

CD635 produz cerca de 13 g/L de lactato após 88 horas, comparado com menos que 9 g/L para a cepa pai, CD607. Cepa CD635 consome levemente mais glicose do que a cepa CD607 durante este período de tempo, e obtém um rendimento significativamente maior com lactato. Cepa CD635 é capaz de crescer no lactato como uma única fonte de carbono.

Exemplo 1B: Mutagênese adicional de cepa CD635 *K. marxianus* e seleção de cepas mutantes (CD832, CD839, CD840, CD841, CD850, CD851 e CD853) tendo resistência ao ácido glicólico.

10 Células de cepa CD635 (Exemplo 1A) são submetidas a mutagênese usando os métodos descritos no Exemplo 1A, com a célula mutagenizada resultante sendo selecionada das colônias que são capazes de crescer em PDA contendo 25 g/L de ácido glicólico. A partir de um total de 23 eventos mutagê-
15 nes, aproximadamente $5,9 \times 10^9$ células são plaqueadas, das quais diversas colônias são produzidas que são resistentes ao ácido glicólico. Algumas dessas colônias resultam da mutagênese com EMS, outras resultam da mutagênese com luz UV, e ainda outras resultam de mutações espontâneas.

20 As colônias resistentes a ácido glicólico crescem separadamente por toda a noite em YB + 100 g/L de glicose em frascos de agitação a 30 °C e agitação a 250 rpm. Biomassa é coletada por centrifugação e 2 g/L de peso seco de célula são inoculados em 50 mL de YP + 50 g/L de glicose em um
25 frasco de agitação defletido. Os frascos são cultivados a 30 °C e agitados a 250 rpm por Aproximadamente 92 horas. Sete mutantes que produzem tituladores de lactato finais significativamente maiores, comparado com cepas pai CD607 e CD635,

são designados como cepas, CD832, CD839, CD840, CD841, CD850, CD851 e CD853.

Cepas CD832, CD839, CD840, CD841, CD850, CD851 e CD853 são incapazes de crescer em lactato como uma única fonte de carbono, indicando que o gene *Km CYB2* tornou-se não funcional nestes mutantes. Então, a região de codificação *KmCYB2* mais ~500 bp e a jusante da região de codificação *KmCYB2* é separadamente amplificado para cada uma dessas cepas usando PCR com enzima FailSafe de alta fidelidade e DNA genômico como o gabarito. Os produtos de PCR ~2,75 kbp resultantes são purificados por meio de purificação de coluna Qiagen e seqüenciada em toda a região de codificação *KmCYB2*.

Cepa CD635 pode crescer em lactato como uma única fonte de carbono, que indicava que ela contém um gene *KmCYB2* funcional. Seqüenciamento do gene *CYB2* a partir de CD635 confirma que ele tem a seqüência *CYB2* tipo selvagem. A seqüência de DNA do gene *KmCYB2* de cepa CD635 é dada como SEQ. ID. NO. 1. A seqüência de aminoácidos da proteína codificada por esse gene *CYB2* é dada como SEQ. ID. NO. 2.

Cepas CD832, CD839, CD840, CD841, CD850, CD851 e CD853 todas não têm a capacidade de crescer em lactato como uma única fonte de carbono, e têm uma resistência maior ao ácido glicólico comparado com a cepa CD607. Cada uma dessas características indica que estas cepas faltam um gene *KmCYB2* funcional. As seqüências de DNA nas regiões de codificação de *Km CYB* para estas cepas em cada caso apresentam mutações na porção de codificação do gene *KmCYB2*, comparado com o gene *KmCYB2* de cepa CD635, que tem a seqüência *CYB2* tipo selva-

gem. Sequência de aminoácidos para as enzimas produzidas por estas regiões de codificação também apresentam diferenças. As diferenças no DNA e sequência de aminoácidos são:

Cepa CD832: uma mutação de ponto na posição do aminoácido 457 causa uma mutação sem sentido, mudando uma arginina para um códon de parada. Esta mutação trunca a proteína em 135 aminoácidos.

Cepas CD839: uma mutação de ponto na posição do aminoácido 272 causa uma mutação de sentido incorreto, mudando uma prolina para uma treonina.

Cepa CD840: uma mutação de ponto na posição do aminoácido 222 causa uma mutação sem sentido, mudando um triptofano para um códon de parada. Esta mutação trunca a proteína em 370 aminoácidos.

Cepa CD841: uma mutação de ponto na posição do aminoácido 147 muda uma histidina para uma tirosina.

Cepa CD850: uma mutação de ponto duplo na posição do aminoácido 309 muda uma serina para uma fenilalanina.

Cepa CD851: uma deleção de dois pares base no par base 219 (posição do aminoácido 74) causa uma mutação por deslocamento do quadro de leitura, resultando em um códon de parada na posição do aminoácido 76 e truncando a proteína.

Cepa CD853: uma inserção de quarto de bases na posição do aminoácido 62 causa uma mutação por deslocamento do quadro de leitura, resultando em um códon de parada na posição do aminoácido 76 e truncando a proteína.

Exemplo 2A: Construção de um plasmídeo (pMM22, Fig. 1) contendo um cassete do gene resistente G418 entre as regiões

flanqueadoras 5' e 3' do gene *K. marxianus* CYB2.

O gene *K. marxianus* L-lactato:ferricitocromo c oxidoreductase (*KmCYB2*) região de codificação do gene é identificado dentro da sequência de DNA genômica de uma cepa
5 tipo selvagem de *K. marxianus* (designada CD21, ATCC 52486) comparando homologia com a homologia conhecida do gene *S. cerevisiae* L-lactato:ferricitocromo c oxidoreductase. Uma região flanqueadora $\approx$ 2kbp diretamente 5' da região de codificação *KmCYB2* é amplificada, com introdução de sítios de
10 restrição *MluI* e *PstI*, por PCR usando iniciadores identificados como SEQ. ID. NO. 3 e SEQ. ID. NO. 4 e DNA genômico como o gabarito. Similarmente, uma região flanqueadora - $\approx$ 2kbp diretamente 3' da região de codificação CYB2 é amplificada com introdução de sítios de restrição *ApaI* e *NgoMIV*
15 por PCR usando iniciadores identificados como SEQ. ID. NO. 5 e SEQ. ID. NO. 6 e DNA genômico como o gabarito.

Um plasmídeo designado pVR29 (descrito no Exemplo 1C e Figura 4 de WO 031102152) é digerido com *MluI* e *PstI* e o flanco *KmCYB2* 5' é ligado com o fragmento resultante para
20 produzir um plasmídeo designado como pMM21. Plasmídeo pMM21 contém o flanco *KmCYB2* 5' a montante de um cassete do gene resistente a G418.

Plasmídeo pMM21 é digerido com *ApaI* e *NgoMIV* e o fragmento resultante é ligado com o 3' *KmCYB2* flanco para
25 produzir um plasmídeo designado pMM22 (figura 1). Plasmídeo pMM22 contém o flanco *KmCYB2* 5' a montante do cassete do gene resistente a G418, e o flanco *KmCYB2* 3' a jusante do cassete do gene resistente a G418.

Exemplo 2B: Geração de um mutante *K. marxianus* (CD936) com um gene *LDH* integrado, gene *PDC* deletado e gene *KmCYB2* deletado transformando a cepa CD607 *K. marxianus* com plasmídeo pMM22 (figura 1, Ex. 2A).

5 Uma colônia *K. marxianus* correspondente a CD607 (ver Ex. 1A) é cultivada e 20 mL de células são centrifugadas. Plasmídeo pMM22 é digerido com *NheI* e *NaeI* (ambas as quais cortadas nos flancos *KmCYB2* no plasmídeo) e o fragmentos resultantes são usados para transformar cepa CD607 *K.*
10 *marxianus* usando métodos de eletroporação padrões. Transformantes são selecionados com base na resistência a G418.

Transformantes selecionados são classificadas no nitrogênio levedura base (YNB) + placas de lactato 1 % e YNB (Nitrogênio levedura base 6,7 g/L, Difco Laboratories,
15 Sparks MD, Cat. No. 0392-15) + glicose 1 %. Transformantes que apresentam crescimento reduzido no meio de lactato são selecionados. Estes transformantes são classificados usando um PCR de colônia de dentro para fora usando iniciadores designados como SEQ. ID. NO. 7 e SEQ. ID. NO. 8. O iniciador
20 5' é posicionado no flanco 5' *CYB2* fora da homologia da construção de ruptura e o iniciador 3' é posicionado dentro do cassete G418. Cada um dos transformantes selecionados apresenta uma inserção do cassete G418 no locus *KmCYB2*. DNA genômico de quatro destes transformantes é isolado e análise
25 se PCR do locus *KmCYB2* realizada usando iniciadores designados como SEQ. ID. NO. 9 e SEQ. ID. NO 10. Estes iniciadores são posicionados ~500 bp a montante e/ou a jusante da região de codificação *CYB2* para assegurar amplificação

completa do locus. Produtos de PCR tendo um tamanho maior com relação a região de codificação CYB2 nativa indicam a substituição do cassete G418 pelo gene CYB2 para cada um desses quatro transformantes. Um destes transformantes é
5 designado como cepa CD936.

Exemplo 2C: Geração de um mutante *K. marixanus* (CD998) com um gene *LDH* integrado, gene *PDC* deletado e gene *KmCYB2* deletado transformando cepa CD635 *K. marxianus* com plasmídeo pMM22 (figura 1, Ex. 2A).

10 Células de cepa CD635 (Exemplo 1A) são transformadas com plasmídeo pMM22 de uma maneira geral descrito no Exemplo 2B. Transformantes são selecionados com base na resistência a G418.

Transformantes selecionados são classificados em
15 YNB + placas de lactato 1 % e + glicose 1 %. Todos os transformantes apresentaram crescimento reduzido no meio de lactato. Estes transformantes são classificados usando uma PCR de colônia de dentro para fora da maneira descrita no Exemplo 2B, e observou-se que 11 dos transformantes têm uma
20 inserção do cassete G418 no locus *KmCYB2*. DNA genômico de quatro destes transformantes é isolado e análise PCR do locus *KmCYB2* realizado da maneira descrita no Exemplo 2B. Como antes, os iniciadores são posicionados ~500 bp a montante e/ou a jusante da região de codificação *KmCYB2* para as-
25 segurar amplificação completa do locus. Produtos de PCR tendo um tamanho maior com relação a região de codificação *KmCYB2* nativa indicam a substituição do cassete G418 pelo gene *KmCYB2* para três desses quatro transformantes. Um des-

ses três transformantes é designado como CD998.

Exemplo 3A: Construção de um plasmídeo (pMM28, Fig. 2) contendo o cassete do gene *KmCYB2* entre seqüências repetitivas diretas de *K. thermotolerans*.

5 Todo o cassete do gene *K. marxianus* CYB2 (*KmCYB2*), incluindo regiões promotoras e terminadoras, é PCR amplificado do DNA genômico de uma cepa *K. marxianus* tipo selvagem designada como CD21, com introdução de sítios de restrição *BamHI* e *SallI*, por PCR usando iniciadores iden-
10 tificados como SEQ. ID. NO. 11 e SEQ. ID. NO. 12. O produto PCR é ligado a um vetor comercial designado como pUC18 (pela Invitrogen Corp., Carlsbad, CA USA) que é digerido com *BamHI* e *SallI*. O plasmídeo resultante é designado como pMM25.

 Uma seqüência de 705 bp identificada como SEQ.
15 ID. NO. 13 é amplificada por PCR do DNA genômico de *K. thermotolerans*, com introdução de sítios de restrição *SphI* e *SallI*, usando iniciadores identificados como SEQ. ID. NO. 14 e SEQ. ID. NO. 15. Esta seqüência *K. thermotolerans* não codifica para qualquer proteína ativa. Plasmídeo pMM25 é
20 digerido com *SphI* e *SallI* e a seqüência *K. thermotolerans* é ligada a ele a montante (5') do cassete *KmCYB2* para formar um plasmídeo designado como pMM27.

 Uma seqüência *K. thermotolerans* idêntica é amplificado por PCR com adição de sítios de restrição *BamHI* e
25 *XmaI*, usando iniciadores identificados como SEQ. ID. NO. 16 e SEQ. M. NO. 17. Plasmídeo pMM2 7 é digerido com *BamHI* e *XmaI* e a seqüência *K. thermotolerans* é ligada a ele a jusante (3') do cassete *KmCYB2* para formar um plasmídeo desig-

nado como pMM28 (figura 2). Plasmídeo pMM28 contém o cas-
 sette *KmCYB2* flanqueado pelas seqüências repetitivas diretas
 de *K. thermotolerans*, ambas orientadas na mesma direção.

Exemplo 3B: construção de um plasmídeo (pMM35, Fig. 3) con-
 tendo um cassete de expressão do gene *MEL1* *S. cerevisiae*
 entre seqüência *K. thermotolerans* idênticas e entre regiões
 5 10 15 20
 flangeadoras 5' e 3' *KmCYB2*.

Um vetor designado como pNC16 é obtido a partir
 do National Research Energy Laboratories in Golden, Colora-
 do. Este plasmídeo contém o gene *MEL1* *S. cerevisiae* sob o
 controle do promotor *PDC1* *S. cerevisiae* e terminador *S. ce-*
revisiae GAL10. O cassete do gene *MEL1* é amplificado por
 PCR com adição de sítios de restrição *BglII* e *SalI* usando
 iniciadores designados como SEQ. ID. NO. 18 e SEQ. ID. NO.
 19. Plasmídeo pMM28 (figura 2, Ex. 3A) é digerido com *Ba-*
mHI e *SalI* e ligado com o cassete *MEL1*. Isto deleta simul-
 15 20
 taneamente o cassete *KmCYB2* de plasmídeo pMM28 e o substi-
 tui com o cassete *MEL1*. O plasmídeo resultante é designado
 pMM31. Ele contém o cassete *MEL1* flanqueado pelas seqüên-
 cias repetitivas diretas *K. thermotolerans*.

Uma região flangeadora $\approx$ 2kbp diretamente 3' da
 região de codificação *KmCYB2* é amplificada com introdução de
 sítios de restrição *XmaI* e *SacI* por PCR usando iniciadores
 identificados como SEQ. ID. NO. 20 e SEQ. ID. NO. 21 e DNA
 25
 genômico como o gabarito. O fragmento resultante é ligado
 ao plasmídeo pMM31 digerido por *XmaI/SacI* para inserir o
 flanco 3' *CYB2* a jusante (3') da seqüência repetitiva dire-
 ta *K. thermotolerans* que é em si a jusante do cassete *MEL1*.

O plasmídeo resultante é designado como pMM32.

Uma região flanqueadora ≈ 2 kbp diretamente 5' da região de codificação *KmCYB2* é amplificada com introdução de sítios de restrição *AatII* e *NarI* por PCR usando iniciadores identificados como SEQ. ID. NO. 22 e SEQ. ID. NO. 23 e DNA genômico como o gabarito. O fragmento resultante é ligado ao plasmídeo pMM32 digerido por *AatII/NarI*. O plasmídeo resultante (designado pMM35, Fig. 3), contém, na ordem, a região flanqueadora 5' *KmCYB2*, uma primeira sequência repetitiva direta *K. thermotolerans*, o cassete *MEL1*, a segunda sequência repetitiva direta *K. thermotolerans* e a região flanqueadora 3' *KmCYB2*.

Exemplo 3C: Geração de um mutante *K. marxianus* (CD1287) com um gene *KmCYB2* deletado transformando um cepa *K. marxianus* tipo selvagem com plasmídeo pMM35.

Plasmídeo pMM35 é digerido com *BsmI* e *NheI* e usado para transformar células de uma cepa *K. marxianus* do tipo selvagem, usando métodos de eletroporação padrões. Transformantes são selecionados para crescimento plaqueando em YNS + melibiose 2 %, com a adição de 32 mg/L *XaGal* (5-bromo-4-cloro-3-indolil-beta-D-galactosídeo, LabScientific, Inc., Livingston, NJ, Cat. No. X-566). Todos os transformantes cresceram neste meio são de cor azul. Eles são subsequentemente marcados nas placas de melibiose para análise posterior. Os transformantes são classificados por recombinação homóloga no locus *KmYB2* por meio de 5' de PCR de colônia dentro para fora usando iniciadores identificados como SEQ. M. NO. 7 e SEQ. ID. NO. 8. Colônias rendendo um produto

de PCR positivo são isoladas duas vezes da colônia única em YNB + melibiose 2 % e 32 mg/L XaGal para eliminar contaminação de fundo. DNA genômico é obtido a partir dessas cepas.

- 5 Todo o locus *KmCYB2* é amplificado por meio de PCR usando iniciadores identificados como SEQ. ID. NO. 24 e SEQ. ID. NO. 25. Quatro colônias dão um produto de PCR como esperado para uma recombinação homóloga em *KmYB2* resultando em uma inserção do cassete do gene *MEL1* no locus *KmCYB2*.
10 Uma dessas colônias é designada CD1286.

- Cepa CD1286 é marcada em um meio não seletivo para duas rodadas e plaqueada por isolados de colônia única, para classificar para a saída do laço do cassete *MEL1*. Três de Aproximadamente 1.300 colônias parecem ser brancas, indi-
15 cando a ausência do gene *MEL1*. DNA genômico é obtido a partir dessas três colônias, e PCR de todo o locus *KmCYB2* é realizado, novamente usando iniciadores identificados como SEQ. ID. NO. 24 e SEQ. ID. NO. 25. Todas as três colônias dão o produto de PCR esperado para um evento da saída do
20 laço, indicando que o cassete *MEL1* foi deletado do genoma da célula. Uma dessas cepas é designada CD1287.

- Exemplo 4A: Construção de um plasmídeo (pMM38, Fig. 5) contendo a região flanqueadora 5' *KmPDC1*, um cassete de expressão do gene *LDH L. helveticus*, um cassete de expressão
25 do gene *KmCYB2* flanqueado pela sequência *K. thermotolerans* repetitivas diretas, e a região flanqueadora 3' *KmPDC1*.

 A região promotora de *TEF2* a partir de uma cepa *K marxianus* tipo selvagem é amplificada com a adição de sí-

tios *KpnI/XbaI* usando iniciadores identificados como SEQ. ID. NO. 26 e SEQ. ID. NO. 26, usando DNA genômico como o gabarito. A ligação subsequente deste produto em um plasmídeo designado pBH76 (figura 4) rende um plasmídeo designado como pMM36. Plasmídeo pMM36 contém, na ordem, a região flangeadora 5' *KmPDC1*, o promotor *KmTEF2*, o gene *LhLDH* e o terminador *ScGAL10*. A região flangeadora 3' *KmPDC1* é em seguida amplificada por PCR com adição de sítios *XmaI/PciI* usando iniciadores identificados como SEQ. ID. NO. 28 e SEQ. ID. NO. 29 e DNA genômico como o gabarito. Este amplicon é ligado no plasmídeo pMM28 (figura 2, Ex. 3A) cortando com *XmaI/PciI* para render um plasmídeo designado pMM37. A região *KmPDC1* e *cassete de expressão LhLDH* do plasmídeo pMM36 são amplificados com a adição de sítios *NarI/NdeI* usando iniciadores designados como SEQ. ID. NO 30 e SEQ. ID. NO. 31, e esta região é ligada no pMM37 digerido por *NarI/NdeI* para produzir plasmídeo pMM38 (figura 5). Plasmídeo pMM38 é designado para inserção visada de *LhLDH* e *cassete de expressão* do genes *KmCYB2* no *PDC1* locus *K. marxianus*.

Exemplo 4B: Geração de um mutante *K. marxianus* (CD 1300) tendo um gene *LDH* exógeno, um gene *PDC* deletado e gene *CYB2* que saiu do laço, transformando cepa CD1287 com plasmídeo pMM38 (Fig 5, Ex. 4A).

Plasmídeo pMM38 é digerido com *NdeI* e usado para transformar cepa CD 1287, usando métodos de eletroporação padrões. Transformantes são selecionados para o crescimento em YNB +placas de lactato 2 %. Transformantes que crescem nas placas de lactato são marcados pelas placas de lactato e

pelas placas YPD. Estes transformantes são classificados para uma recombinação homóloga no locus *PDC1* por meio de 3' PCR de dentro para fora usando iniciadores identificados como SEQ. ID. NO. 32 e SEQ. ID. NO. 33. Transformantes rendendo um produto de PCR do tamanho esperado (2251 bp) são marcados para isolamento de colônia única. DNA genômico de três isolados de colônia única de cada transformante positivo é obtido e o locus *PDC1* é amplificado por meio de PCR usando iniciadores identificados como SEQ. ID. NO. 34 e 35. Isolados dando o produto de PCR esperado (8453 bp vs. 3 730 bp para o locus *PDC* do tipo selvagem, indicando a perda do gene *PDC* e integração do gene *LDH*) são em seguida testados pelo fenótipo *Pdc⁻* avaliando se ele produz etanol. Uma das cepas que não produz etanol é designada cepa CD1298. Cepa CD1298 contém um gene *PDC* deletado, o cassete do gene *LDH* exógeno, e o cassete do gene *CYB2* do plasmídeo pMM38.

Cepa CD1298 é marcada no meio YPD não seletivo para três rodadas e um lama de célula é plaqueada em PDA + placas de ácido glicólico 7,5 g/L. Após 72 horas a 30 °C, colônias são selecionadas e classificadas por meio de PCR de colônia para uma saída do laço do gene *KmCYB2*, usando iniciadores identificados como SEQ. ID. NO. 36 e 37. Quatro colônias dão o tamanho correto do produto para uma saída do laço do gene *CYB2* (1109 bp vs 5129 bp para o produto contendo o cassete *CYB2*). Colônias únicas são isoladas de cada uma dessas quatro cepas. DNA genômico é obtido a partir de cada uma dessas colônias únicas e o locus *PDC1* é amplificado por meio de PCR usando iniciadores identificados como SEQ. ID. NO. 38 e 36.

Duas cepas rendem um produto de PCR do tamanho esperado para uma saída do laço do cassete CYB2 (3249 bp vs 7269 bp para o produto contendo o cassete CYB2 e 2546 para o produto de PCR do tipo selvagem). Uma dessas duas cepas é designada como
5 cepa CD 1300.

Exemplo 5A: Caracterizações de frasco de agitação microaeróbico de cepas CD635 (Ex. 1A), CD853 (Ex. 1A) e CD998 (Ex. 2C)

Cepas CD635, CD853 e CD998 são separadamente cultivadas em baixa densidade (0,05-0,10 OD₆₀₀) e permitindo que a cultura produza produtos de fermentação enquanto que cresce. Vinte mL do meio (YP + dextrose 60 g/L) são pipetados em um frasco de agitação defletido de 250 mL e o meio é inoculado com a quantidade exigida de biomassa. O frasco é
10 em seguida incubado a 30 °C e orbitalmente agitado a 70 rpm por 92 horas, com amostras sendo retiradas (~ a cada 24 horas) durante o curso do experimento para testar para consumo de glicose, acúmulo de fermentação do produto, bem como crescimento de biomassa.

20 Após 92 horas de cultivo, a cepa CD635 (não um exemplo desta invenção) consumiu 17,6 g/L de glicose e produziu 11,4 g/L de lactato. No mesmo período de tempo, cepa CD853 consome 26,9 g/L de glicose e produz 23,3 g/L de lactato, levemente mais do que o dobro produzido pela cepa
25 CD635. Cepa CD998 consome 21,9 g/L de glicose e produz 18,5 g/L de lactato, mais do que 60 % mais do que cepa CD635. Rendimento para lactato é ~65 % para cepa CD635, ao passo que rendimentos para cepas CD835 e CD998 são ~86,6 % e ~84,5

%, respectivamente.

Exemplo 5B: Avaliação de resistência ao ácido glicólico para cepas CD635 (Ex. 1A), CD853 (Ex. 1A) e CD998 (Ex. 2C).

Placas PDA em duplicata são feitas contendo concentrações entre 5 g/L através de 25 g/L de ácido glicólico. Placas de controle não contendo nenhum ácido glicólico são também preparadas. Uma série de diluição de cepas CD635, CD853 e CD998 é preparada, e 5 µL de cada diluição da série de diluição é pontuada em cada uma das placas anteriores. As placas inoculadas são cultivadas a 30 °C por três dias. Cepa CD635 comparativa cresce em placas contendo ácido glicólico até uma concentração de 15 g/L, mas falta crescer em qualquer das placas contendo ácido glicólico com maior concentração. Cepa CD853 apresenta capacidade de crescer nas placas contendo ácido glicólico até 25 g/L de ácido glicólico. Cepa CD998 pode crescer nas placas contendo até 22,5 g/L de ácido glicólico. Estes testes indicam que ácido glicólico pode ser usado como um meio seletivo para cepas mutantes tendo um gene CYB2 não funcional.

Exemplo 6: caracterizações de frasco de agitação microaeróbicos de cepa CD607 comparativa (Ex. 1A), cepa CD936 (Ex. 2B) e cepa CD1300 (Ex. 3C).

Cepas CD607, CD936 e CD1300 são separadamente cultivadas nas condições descritas no Exemplo 5A. Após 93 horas de cultivo, cepa CD607 comparativa consumiu 16,5 g/kg de glicose e produziu 10,8 g/kg de lactato. No mesmo período de tempo cepa CD936 consome 21,6 g/kg de glicose e produz 19,3 g/kg de lactato, ligeiramente menos que o dobro produzido

pela cepa CD607. Cepa CD1300 consome 29,9 g/kg de glicose e produz 24,6 g/kg de lactato, mais do que 125 % a mais que a cepa CD607. Rendimento para lactato é 65,5 % para cepa CD607, ao passo que rendimentos para cepas CD936 e CD1300 são 89,4 % e 82,3 %, respectivamente. A cepa CD607 também produz 2,02 g/kg de ácido pirúvico neste intervalo de tempo, ao passo que CD936 e CD1300 produzem 0,68 g/kg e 0,83 g/kg de ácido pirúvico respectivamente. Também neste intervalo de tempo, CD607 produz 0,67 g/kg de ácido acético em comparação com a produção de ácido acético tanto para CD936 quanto para CD1300.

O crescimento é também monitorado durante este experimento. Após 93 horas, a cepa CD607 comparativa cresceu até um OD600 de cerca de 5,7, ao passo que cepas CD936 e CD1300 cresceram até um OD600 de cerca de 7,5 e 3,2, respectivamente. O crescimento muito baixo de cepa CD1300 indica que a produtividade específica desta cepa é muito maior do que da cepa CD607 comparativa e cepa CD936.

Exemplo 7A: Clonagem de região promotora *I. orientalis* PGK (*IoPGK1*); construção de um plasmídeo (pMI318, Fig. 6) tendo o gene da higromicina de *E. coli* sob o controle do promotor *IoPGK1* e o terminador *S. cerevisiae* GAL10.

Um fragmento da sonda de 920 bp de o gene *C. sonorensis* PGK1 é obtido a partir do DNA genômico de *C. sonorensis* da mesma maneira descrita no Exemplo 22 de WO 02/042471, usando iniciadores identificados como SEQ. ID. NO. 39 e SEQ. ID. NO. 40. DNA genômico é isolado de um que cresce cepa *I. orientalis* e ré-suspenso em Tris-HCl 10 mM + EDTA 1 mM (pH

8) (TE). O DNA genômico *I. orientalis* é cortado com *HindIII*, e um Southern blot é preparado e hibridizado usando métodos padrões com o gene *C. sonorensis PGK1* como uma sonda. Fragmentos de tamanho ≈ 2 kbp são isolados do gel de agarose e clonados em um plasmídeo cortado com *HindIII*. Hibridização de colônia dos Transformantes do *E. coli* com a Sonda *PGK1* resulta na isolação de um fragmento de *HindIII* contendo mais das seqüências de codificação de proteína *I. orientalis PGK1* (*IoPGK1*), mas nenhuma seqüências promotoras, conforme verificado pelo seqüenciamento.

Fragmentos genômicos contendo a região promotora de *IoPGK1* são obtidos com amplificação de PCR mediada por ligação (Mueller, P.R. and Wald, B. 1989, "In vivo footprinting of a muscle specific enhancer by ligation mediated PCR." *Science* 246:780-786). Uma mistura de um ligante identificado como SEQ. ID. NO. 41 e um ligante identificado como SEQ. ID. NO. 42 é ligada ao DNA genômico *I. orientalis* digerido pelo *HaeIII* com T4 DNA ligase (New England BioLabs). Amostras das misturas da ligação são usadas como gabarito de 50 μ L para reações contendo 0,1 pM de um iniciador identificado como SEQ. ID. NO. 43 e 1 μ M de um iniciador identificado como SEQ. ID. NO. 44. A mistura da reação é aquecida a 94 °C por 3 minutos após 2 U de Dynazyme EXT ser adicionado. As reações são cicladas 30 vezes como a seguir: 1 minuto a 94 °C, 2 minutos a 68 °C e 2 minutos a 72 °C, com uma extensão final de 10 minutos a 72 °C. Uma amostra diluída desta primeira amplificação por PCR é usada como o gabarito em uma reação de PCR aninhada (50 μ L) contendo 0,05 μ M

de um iniciador identificada como SEQ. ID. NO. 45 e 0,5 μ M de um iniciador identificado como SEQ. ID. NO. 46. A mistura da reação é aquecida a 94 °C por 3 minutos após 2 U de Dynazyme EXT ser adicionado. As reações são em seguida cicladas 30 vezes como a seguir: 1 minuto a 94 °C, 2 minutos a 67 °C e 2 minutos a 72 °C, com um extensão final de 10 minutos a 72 °C.

Um fragmento de PCR de $\approx$ 600 bp é isolado e seqüenciado. Iniciadores aninhados identificados como SEQ. ID. NO. 47 e SEQ. ID. NO. 48 são designados e usados em uma amplificação de PCR mediada por ligação juntamente com oligonucleotídeos identificados como SEQ. ID. NO. 49 e SEQ. ID. NO. 50 similarmente da maneira anterior exceto que DNA *I. orientalis* digerido por *SphI* é usado e o PCR é realizado usando uma temperatura de anelamento de 65 °C.

O promotor de *PGK1* de *I. orientalis* é PCR amplificado usando iniciadores identificados como SEQ. ID. NO. 51 e SEQ. ID. NO. 52 e o DNA genômico *I. orientalis* como o gabarito. O fragmento é cheio usando a enzima Klenow e em seguida cortado com *XbaI*. Um fragmento de 633 bp é gel isolado e ligado a um fragmento de 4428 bp obtido digerindo um plasmídeo designado como pMI270 (descrito na Fig. 4 de WO 031049525) com *XhoI*, enchendo o fragmento usando a enzima Klenow e 0,1 mM de dNTP, e digerindo com *XbaI*. Plasmídeo pMI270 contém o gene de higromicina da *E. coli* ligado com um promotor *C. sonorensis PGK1* e um terminador *S. cerevisiae GAL10*. O plasmídeo resultante é designado pMI318 (figura 6). Plasmídeo pMI318 contém o gene de higromicina da *E. co-*

li sob o controle do promotor *I. orientalis* *PGK1* e o terminador *S. cerevisiae* *GAL10*.

Exemplo 7B: Construção de um plasmídeo (pMI321, Fig. 7) contendo o gene de higromicina sob o controle do promotor de *IoPGK1* e o terminador de *S. cerevisiae* *GAL10*, e o gene *LDH* *L. helveticus* sob o controle do promotor de *IoPGK1* e terminador *S. cerevisiae* *CYC1*.

O promotor *I. orientalis* *PGK1* do Exemplo 7A é PCR amplificado usando iniciadores identificados como SEQ. ID. NO. 53 e SEQ. ID. NO. 54 e o DNA genômico *I. orientalis* como o gabarito. O fragmento é cheio usando a enzima Klenow e 0,1 mM de dNTP, e em seguida cortado com *NcoI*. Um fragmento de 633 bp é gel isolado.

Plasmídeo pVR1 (descrito em WO 03/102152 Figura 7) contém o gene *L. helveticus* *LDH* sob o controle de o *S. cerevisiae* *TEF1* promotor e o *S. cerevisiae* *CYC1* terminador. Plasmídeo pVR1 é digerido com *XhoI*, cheio em usando a enzima Klenow, e cortadas com *NcoI*. Um fragmento de 7386 bp de plasmídeo pVR1 é ligado com o fragmento promotor de *IoPGK1* de 633 bp. O plasmídeo resultante é designado pMI320. Plasmídeo pMI320 contém o gene *L. helveticus* *LDH* sob o controle do promotor de *IoPGK1* e terminador *S. cerevisiae* *CYC1*.

Plasmídeos pMI318 (figura 6, Ex. 7A) e pMI320 são digeridos com *ApaI* e *NotI*. Um fragmento de 5008 bp do plasmídeo pMI318 é ligado a um fragmento de 1995 bp do plasmídeo pMI320 para formar o plasmídeo pMI321 (figura 7).

O gene de higromicina (e seu terminador) é posicionado no plasmídeo pMI321 entre duas cópias do promotor de *IoPGK1*,

que serve como seqüências repetitivas diretas. Esta construção pode permitir uma célula transformada com plasmídeo pMI321 a engajar em uma recombinação homóloga para "saída do laço" o gene de higromicina e terminador, juntamente com
 5 uma cópia do promotor de *IoPGK1*.

Exemplo 7C: Clonagem de região promotora *I. orientalis* *PDC* (*IoPDC1A*); construção de um plasmídeo (pMI355, Fig. 8) tendo o Gene de higromicina da *E. coli* sob o controle do promotor de *IoPGK1* e o terminador *S. cerevisiae* *GAL10*, o gene
 10 *L. helveticus* *LDH* sob o controle do promotor de *IoPGK1* e terminador *S. cerevisiae* *CYC1*, e a região flanqueadora *IoPDC1A* 5'.

Uma biblioteca genômica da cepa *I. orientalis* do tipo selvagem ATCC PTA-6658 é construída no vetor cosmídeo
 15 SuperCos1 (Stratagene) de acordo com instruções fornecidas pelo fabricante. Seqüências do tipo *PDC* são amplificadas por PCR a partir do DNA genômico da cepa com iniciadores designados como SEQ. ID. NO. 55 e SEQ. ID. NO. 56. Um fragmento de 700 bp de um gene *PDC* é amplificado. A biblioteca
 20 genômica é classificada usando técnicas de hibridização com fragmentos de PCR marcados como sondas da maneira descrita em WO 03/049525 e clones de cosmídeo contendo o gene *PDC* são isolados e seqüenciados. A região *I. orientalis* *PDC1A* 5' de 1000 bp a 167 bp a montante do início de o quadro de leitura
 25 aberto é amplificado por PCR usando iniciadores identificados como SEQ. ID. NO. 57 e SEQ. ID. NO. 58 e o DNA do cosmídeo *I. orientalis* *PDC* como o gabarito. O fragmento é cortado com *Sall* e *SacI*. Um fragmento de 836 bp é gel isolado

e ligado a um fragmento de 6992 bp obtido digerindo plasmídeo pMI321 (figura 7, Exemplo 7B) com *SalI* e *SacI*. O plasmídeo resultante é nomeado pMI355 (figura 8).

Exemplo 7D: Clonagem de região terminadora *I. orientalis* PDC (*IoPDC1A*); construção de plasmídeos (pMI356 (figura 9) e pMI357) e tendo a região flanqueadora *IoPDC1A* 5', o Gene de higromicina da *E. coli* sob o controle do promotor de *IoPGK1* e o terminador *S. cerevisiae* GAL10, o gene *L. helveticus* LDH sob o controle do promotor de *IoPGK1* e terminador *S. cerevisiae* CYC1, e a região flanqueadora *IoPDC1A* 3'.

A região *I. orientalis* PDC 3' correspondente às seqüências de 233 bp a 872 bp a jusante do PDC códon de parada de tradução é amplificada por PCR usando iniciadores identificados como SEQ. ID. NO. 59 e SEQ. ID. NO. 60 e o DNA do cosmídeo *I. orientalis* PDC1A (Exemplo 7C) como o gabarito. O fragmento é cortado com *ApaI* e *SmaI*. Um fragmento de 630 bp é gel isolado e ligado a um fragmento de 7809 bp obtido digerindo plasmídeo pMI355 (figura 8, Ex. 7C) com *ApaI* e *SmaI*. O plasmídeo resultante é nomeado pMI356 (figura 9). Ele contém o higromicina e cassetes LDH do plasmídeo pMI355 entre o flanco 5' e uma porção do flanco 3' do gene de *IoPDC1A*.

Plasmídeo pMI357 é feito de uma maneira similar, exceto que ele difere do plasmídeo pMI356 com relação à porção do flanco 3' *IoPDC1A* que está presente. Estas regiões correspondem à seqüência 5 bp a montante e 216 bp a jusante do códon de parada de tradução de PDC.

Exemplo 7E: Geração de um mutante *I. orientalis* (CD1184) com genes *IoPDC1A* e *IoPDC1B* deletados e gene *LhLDH* integrado em uma etapa transformando cepa *I. orientalis* do tipo selvagem com plasmídeo pM356 (figura 9, Ex. 7D).

5 Cepa *I. orientalis* do tipo selvagem ATCC PTA 6658 é transformada com plasmídeo pMI356 usando métodos padrões. Cepas transformadas que crescem nas placas de higromicina são cultivadas. Um transformante que não produz etanol é selecionado para análise Southern, que confirma a
10 deleção tanto dos genes *IoPDC1A* e *IoPDC1B* quanto inserção de pelo menos uma cópia do gene *LhLDH*. Esta cepa é designada CD1184.

Exemplo 7F: Construção de plasmídeo pM433 (figura 10) contendo a região flanqueadora *IoPDC1A* 5', o gene *ScMEL5* sob
15 o controle do promotor de *IoPGK1*, o gene *L. helveticus LDH* sob o controle do promotor de *IoPGK1* e terminador *SeCYC1*, e a região flanqueadora *IoPDC1A* 3'.

O promotor *I. orientalis PGK1* é amplificado por PCR usando iniciadores identificados como SEQ. ID. NO. 61
20 e SEQ. ID. NO. 62 e o DNA genômico *I. orientalis* como o gabarito. O fragmento é cheio usando a enzima Klenow e 0,1 mm de dNTP, e em seguida cortado com *SphI*. Um fragmento de 669 bp é gel isolado. Um plasmídeo designado como pMI233 (descrito em Fig. 23C de WO 03/049525) é cortado com *XhoI*. O
25 fragmento é cheio com a enzima Klenow e cortado com *SphI*. Os fragmentos 4534 bp e o 669 bp são ligados e o plasmídeo resultante é nomeado pMI319. Plasmídeo pMI319 contém o gene *S. cerevisiae MEL5* (*ScMEL5*) e a região promotora de *IoPGK1*.

Plasmídeo pMI319 é cortado com *ApaI*, sua extremidade fica romba com polimerase T4, e cortado com *NotI*. Um fragmento de 2317 bp é gel isolado. Ele é ligado a um fragmento de 6498 bp obtido digerindo plasmídeo pMI357 (Exemplo 7D) com *SallI*, fazendo sua extremidade romba com a enzima Klenow e em seguida cortando com *NotI*. O plasmídeo resultante contém o gene *ScMEL5* (com seu terminador nativo) em vez de gene de higromicina de plasmídeo pMI357. O plasmídeo resultante é designado pMI433 (figura 10).

Exemplo 7G: Construção de plasmídeos pMI449 (figura 11) e pMI454 (figura 12) contendo região flanqueadora *I. orientalis* *CYB2* 5', cassete do gene *ScMEL5* entre seqüências repetitivas diretas *K. thermotolerans* e da região flanqueadora *I. orientalis* *CYB2* 3'.

Plasmídeo pMM28 (figura 2, Ex. 3A) é digerido com *BamHI*, cheio em com a enzima Klenow, e digerido com *SallI*. O fragmento de 4077 bp obtido dessa maneira é ligado a um fragmento de 2317 bp *NotI* (cheio com enzima Klenow)-*SallI* de pMI433 (figura 10, Ex. 7F). O plasmídeo resultante é designado pMI445.

A região flanqueadora 3' do gene *I. orientalis* L-lactato:ferricitocromo c oxidoreductase (*IoCYB2A*) (correspondente às seqüências de 90 a 676 bp a jusante do quadro de leitura aberto previsto) é amplificado por PCR usando iniciadores identificados como SEQ. ID. NO. 63 e SEQ. ID. NO. 64 usando um clone do cosmídeo *CYB2-2* como um gabarito. O Produto de PCR é digerido com *SacI* e *SmaI* e o fragmento de 607 bp é ligado com o fragmento de 6386 bp *SacI* - *SmaI* do plas-

mídeo pMI445. O plasmídeo resultante é designado pMI448.

A região flanqueadora IoCYB2A 5' (correspondente às seqüências de 913 a 487 bp a montante do quadro de leitura aberto previsto) é amplificado por PCR usando iniciadores identificados como SEQ. ID. NO. 65 e SEQ. M. NO. 66, novamente usando o clone do cosmídeo CYB2-2 como um gabarito. O Produto de PCR é digerido com *SphI* e o fragmento de 454 bp é ligado com o fragmento de 6993 bp *SphI* obtido digerindo parcialmente pMI448. O plasmídeo resultante é designado pMI449 (figura 11).

A região flanqueadora IoCYB2A 5' (correspondente às seqüências de 466 a 7 bp a montante do quadro de leitura aberto previsto) é amplificada por PCR usando iniciadores identificados como SEQ. ID. NO. 67 e SEQ. ID. NO. 68, uma vez novamente usando o clone do cosmídeo CYB2-2 como o gabarito. O Produto de PCR é digerido com *SphI* e o fragmento de 493 bp foi ligado com o fragmento de 6993 bp *SphI* obtido digerindo parcialmente plasmídeo pMI448. O plasmídeo resultante é designado pMI453.

A região flanqueadora IoCYB2A 3' (correspondente às seqüências de 402 bp a montante a 77 bp a jusante do códon de parada previsto) é amplificada por PCR usando iniciadores identificados como SEQ. ID. NO. 69 e SEQ. ID. NO. 70 usando o cosmídeo CYB2-2 como um gabarito. O Produto de PCR é digerido com *ApaI* e *SmaI* e o fragmento de 506 bp é ligado com o fragmento de 6886 bp *ApaI* - *SmaI* de plasmídeo pMI453. O plasmídeo resultante é designado pMI454 (figura 12).

Exemplo 7H: Geração de cepa mutante *I. orientalis* (CD1496)

para transformar sucessivamente cepa CD1184 (Ex. 7E) com plasmídeos pMI449 e pM454, seguido por mutagênese.

Cepa CD1184 é transformada com plasmídeo pMI449 usando o método de acetato de lítio e transformantes (colônias azuis) são selecionados com base na atividade de melibi-
5 ase em placas YPD X- α -gal. A substituição do gene *IoCYB2A* da cepa CD 1184 com o DNA transformado é confirmada por PCR de colônia e análise Southern em alguns dos transformantes. O marcador *MEL5* sai do laço de um destes transformantes por
10 meio de um evento de recombinação homóloga através das sequências repetitivas *K. thermotolerans*, como confirmado por análise Southern. O segundo alelo *CYB2A* é deletado deste transformante usando plasmídeo pMI454. Transformantes são analisados por PCR de colônia para a ausência de um produto
15 de PCR específico de 1000 bp *CYB2A*. O marcador *MEL5* do plasmídeo pMI454 é que saiu do laço de um transformante tendo uma deleção do segundo alelo *CYB2A* por meio de recombinação como antes. Este transformante é designado cepa CD 1436.

Cepa CD1436 é submetida a mutagênese EMS usando as
20 condições apresentadas no Exemplo 1A, exceto que as condições de exposição são 8 uL por 1 hora. Células mutagenizadas recuperam naturalmente por 6 horas em YPD e em seguida são plaqueadas em PDA + placas de ácido láctico 35 g/L e incubadas por uma semana a 30 °C. Uma cepa que produz mais lactato
25 e menos glicerol do que cepa CD 1436 é designada como cepa CD 1496. Cepa CD 1496 tem uma deleção tanto de alelos *PDC1* (com substituição por um cassete do gene *L-LDH* funcional, quanto de uma deleção de cada um de seus dois alelos *IoCYB2A* nati-

vos).

Exemplo 7I: Fermentações de frascos de agitação usando cepas CD1184, CD1436 e CD1496.

Cepas CD1184 e CD1436 são avaliadas em fermentações frasco de agitação microaeróbico de uma maneira geral descrito no Exemplo 5A, exceto a concentração de dextrose de partida ser 100 g/L e agitação ser a 100 rpm. Cada cepa pára de consumir glicose após cerca de 80 horas de produção. Após 144 horas de cultivo, titulador de lactato final é produção de lactato é cerca de 60 g/kg para cepa CD 1184 e 53 g/kg pra cepa CD1436. O rendimento de lactato é cerca de 88 % para cepa CD1184 e 86 % para cepa CD 1436. Rendimento de glicerol é cerca de 5,5 % para cepa CD 1184 e 8,5 % para cepa CD 1436.

Cepas CD 1436 e CD 1496 são similarmente avaliadas em cultivos microaeróbicos em duplicata. Após 144 horas de cultivo, titulador de lactato final é produção de lactato é cerca de 58 g/kg para cepa CD1436 e 65 g/kg para cepa CD1496. O rendimento de lactato é cerca de 68 % para cada cepa. Rendimento de glicerol é cerca de 11 % para cepa CD1436 e apenas cerca de 8 % para cepa CD 1496.

Exemplo 7J: fermentações pH 3 de cepas CD1436 (Exemplo 7H) e CD1496 (Exemplo 7H).

Cepas CD1436 e CD1496 são usadas em fermentações separadas tamponadas a um pH de 3,0. Em cada caso as células são inoculadas com um fermentador Braun-B de 5 L com um volume de trabalho de 3 L. O caldo de fermentação contém 110 g/L de dextrose no início da fermentação. Hidróxido de

potássio é adicionado como necessita de manter o pH do caldo de fermentação a 3,0. A temperatura da fermentação é 30 °C. Oxigênio é adicionado para manter uma taxa de captação de oxigênio de 5-6 mmol/L-hr.

5 Cepa CD1436 produz lactato em uma taxa de 0,77 g/l-hr, até um rendimento final de 69 %. Cepa CD 1496 produz lactato em uma taxa de 0,83 g/l-hr, até um rendimento final de 64 %. Cepas CD1436 e CD1496 produzem 0,8 e 0,0 g/L de piruvato, respectivamente. Elas produzem 18 e 23 g/L de
10 glicerol, respectivamente. Nenhuma das duas produz acetato mensurável.

Exemplo 7K: fermentações pH 3 de cepa CD1436 (Ex. 7H) e sua cepa CD1184 pai (Ex. 7E).

 Cepa CD1436 e sua cepa CD1184 pai é usada em fer-
15 mentações separadas tamponadas a um pH de 3,0. Em cada caso as células são inoculadas com um fermentador Braun-B de 5 L com um volume de trabalho de 3 L. O caldo de fermentação contém 130 g/L de dextrose no início da fermentação. Hidróxido de potássio é adicionado como necessidade de para man-
20 ter o pH do caldo de fermentação a 3,0. A temperatura da fermentação é 30 °C. Oxigênio é adicionado para manter uma taxa de captação de oxigênio de 2 mmol/L-hr.

 Cepa CD1436 produz um lactato de 66 g/L após cerca de 160 horas de cultivo, ao passo que cepa CD1184 produz a-
25 penas 56 g/L de lactato. Cepa CD1436 também produz láctico em uma taxa volumétrica maior do que cepa CD1184 (0,43 vs. 0,35 gL/h). Além disso, cepa CD1436 acumula no acetato ou piruvato ao passo que cepa CD1184 produz cerca de 1 g/L de piruva-

to e 3 g/L de acetato. Cepa CD 1436 tem um rendimento geral de lático na glicose de cerca de 67 % vs. 69 % para seu pai. Os rendimentos comparáveis são obtidos em virtude da cepa CD1436 produzir mais glicerol do que a cepa CD 1184.

- 5 Exemplo 8A: Deleção de gene *DLD1* de uma cepa *K. marxianus* contendo um cassete do gene *D-LDH* exógeno.

O locus 3' a jusante do gene *DLD1* de uma cepa *K. marxianus* do tipo selvagem é amplificado com adição de sítios de restrição *NarI/SalI* usando iniciadores identificados
 10 como SEQ. ID. NO. 71 e SEQ. ID. NO. 72 e DNA genômico como o gabarito. A região clonada é ligada ao plasmídeo pMM31 (Ex. 3B) para formar plasmídeo pMM44 (figura 13). O locus 5' a jusante do gene *DLD1* de uma cepa *K. marxianus* do tipo selvagem é amplificado com adição de sítios de restrição
 15 *XmaI/NcoI* usando iniciadores identificados como SEQ. ID. NO. 73 e SEQ. ID. NO. 74 e DNA genômico como o gabarito. A região clonada é ligada ao plasmídeo pMM31 (Ex. 3B) para formar plasmídeo pMM45 (figura 14).

Um fragmento é amplificado do plasmídeo pMM44 usando iniciadores identificados como SEQ. ID. NO. 71 e SEQ.
 20 ID. NO. 75. Um outro fragmento é amplificado do plasmídeo pMM45 usando iniciadores identificados como SEQ. ID. NO. 74 e SEQ. ID. NO. 76. Os produtos amplificados (~2,8 kbp do pMM44 e ~2,5 kbp do pMM45) são gel purificado e usados para
 25 transformar uma cepa *K. marxianus* contendo um cassete do gene *DLDH* exógeno (e tendo um gene *PDC* intacto), que é preparado de acordo com métodos similares àqueles descritos em WO 03/102152. Depois de uma recuperação de 4 horas, células

são plaqueadas em YNB 2 % de melibiose + placas XaGal 32 mg/L.

Transformantes são classificados para uma recombinação homóloga do DNA transformado no locus *DLD1* por meio de PCR de colônia de dentro para fora usando iniciadores identificados como SEQ. ID. NO. 77 e SEQ. ID. NO. 78. Os iniciadores são homólogos a uma região a montante da região 5' clonada e a região de codificação do marcador de seleção *Sc-MEL1*, respectivamente. Três dos 32 transformantes classificados dão o produto de PCR esperado de 1,8 kb e são marcados para isolados de colônia única.

Todo o locus *DLD1* é amplificado de DNA genômico dos isolados de colônia única de dois dos três transformantes positivos usando iniciadores identificados como SEQ. ID. NO. 74 e SEQ. ID. NO. 77. Todas as colônias únicas renderam o produto esperado para um nocaute do gene *DLD1*. Uma das cepas é identificada como cepa CD1603.

Cepa CD1603 e a cepa pai foram separadamente marcadas por placas YNB contendo ácido D-lático 2 % como uma única fonte de carbono. A cepa pai cresce bem neste meio, mas a cepa CD1603 apresenta crescimento desprezível.

Cepa CD1603 e a cepa pai são separadamente cultivadas de uma maneira geral descrita no Exemplo 5A. Ácido lático e produção do subproduto são medidos por meio de HPLC por todos os cultivos. Produção de ácido lático é melhorado em cepa CD1603 comparado com a cepa pai. Tanto a taxa de produção láctica quanto titulador lático são maiores em cepa CD1603 do que na cepa pai.

Exemplo 8B: fermentações pH 3 de cepa CD1603 (Exemplo 8A) e sua cepa pai.

Cepas CD1603 e seu pai são usadas em fermentações separadas tamponadas a um pH de 3,0. Em cada caso as células são inoculadas com um fermentador Braun-B de 5 L com um volume de trabalho de 3 L. O caldo de fermentação contém 90 g/L de dextrose no início da fermentação. Hidróxido de potássio é adicionado como necessidade para manter o pH do caldo de fermentação a 3,0. A temperatura da fermentação é 30 °C. Oxigênio é adicionado para manter uma taxa de captação de oxigênio de 5-6 mmol/L-hr.

Cepa CD1603 produz lactato a uma taxa de 0,58 g/L-hr, até um rendimento final de 69 %. Ela produz piruvato em um rendimento de 0,8 % e glicerol em um rendimento de 7,3 %. A cepa pai produz lactato em um taxa de apenas 0,09 g/L-hr, até um rendimento final de 68 %. O piruvato e rendimentos de glicerol da cepa pai são 21 % e 6,9 %, respectivamente. Estes valores indicam que a cepa pai consome lactato da maneira que a forma, convertendo-a em piruvato. Além disso, a cepa pai produz um rendimento de 15 % de acetato.

Exemplo 9 Produção de células de levedura *S. bulderi* tendo um gene *LhLDH* exógeno e um cassete do gene G418 exógeno.

S. bulderi é um levedura que, como uma cepa nativa, não tem capacidade de crescer no lactato como a única fonte de carbono. Iniciadores identificados como SEQ. ID. NO. 85 e SEQ. ID. NO. 86 são usados para andar juntamente com estojo AP1 da Universal Gene Walking (BD BioSciences Clontech, Palo Alto, CA, Catalog No. K1807-1) para elucidar

uma sequência de ≈ 1 kbp a montante do gene *S. bulderi* *TEF1*. Uma sequência de 850 bp da qual a sequência a montante é clonada imediatamente a montante de um gene *LhLDH*, usando sítios de clonagem *SacI/XbaI* em iniciadores identifi-
5 cados como SEQ. ID. NO. 87 e SEQ. ID. NO. 88, para produzir um plasmídeo designado como pCM149 (figura 15). Plasmídeo pCM149 contém, na ordem, o promotor *SbTEF1*, o gene *LhLDH*, um terminador *Saccharomyces exigus PG1*, um cassete do gene G418 sob o controle do promotor *ScPDC1* e o terminador *Sc-*
10 *GAL10*. O terminador *SePG1* é obtido amplificando DNA genômico de *S. exiguus* com iniciadores identificados como SEQ. ID. NO. 89 e SEQ. ID. NO. 90. O cassete do gene G418 é descrito no Exemplo 1A (figura 4) de WO03/102152.

Células de um tipo selvagem *S. bulderi* (cepa ATCC
15 MYA-403) são transformadas com plasmídeo pCM149 usando métodos padrões de acetato de lítio. Transformantes são plaqueados em YPD + 300 mg/L G418, e três transformantes produzem halos púrpuros fracos usando um ensaio LASSO, que indicam que estas cepas estão produzindo ácido láctico. Estas três
20 cepas são separadamente inoculadas em frascos de agitação YPD, e todas as três são para produzir ácido láctico. Todas as três cepas produzem de 5-7 g/L de ácido láctico após cerca de 24 horas de cultivo. A cepa pai tipo selvagem não produz qualquer ácido láctico.

Listagem de Sequência

<110> Natureworks LLC
Miller, Matthew
Suominen, Pirkko
Aristidou, Aristos
Hause, Benjamin
van Hoek, Pim
Dundon, Catherine

<120> "CÉLULAS DE LEVEDURA QUE PRODUZEM ÁCIDO LÁCTICO TENDO L- OU D-
LACTATO NÃO FUNCIONAL: GENE DE FERRICITOCROMO C OXIDOREDUCTASE"

<130> 1181A

<150> US 60/739,458

<151> 2005-11-23

<150> US 60/739,824

<151> 2005-11-23

<160> 90

<170> PatentIn versão 3.3

<210> 1

<211> 1779

<212> DNA

<213> Kluyveromyces marxianus

<400> 1
atgagatctg cagctagagt catcaacaag agctgtagcg gctctgctct ttcgagaaga 60
tgtctgagaa aatcgagctt gagcatgagt atgagatatt taagtacttc gaacatcgga 120
gtcagaaaag gtttcaacgg acagggtaaa agttcgaata aaacgatgtt gtttttggt 180
gctggggctt cagctgtagc tgggatcggg ctgctctctc aatttagcga cagcttgcaa 240
aatgctacga aagaagagct aaacaagcca aagggttccc cgctcgaggt cgccaagcac 300
tcgagtcccc atgactgctg ggtggtgata gatgggttg tctacaattt gacggaattc 360
attagcgcct atcctggtgg accagctatc atcgagaaca acgccggtaa ggatgtgacg 420
gcatcttttg gcccaattca tgcgccagat gtcattgaga agtacattgc tccggagaac 480
cggatcggtc ctctagacgg gaaaatgccc gacgacttga tctgtgcgcc attgacgccc 540
ggtgagactc ctgaggatgt tgccagaaag gaggagtgc gtcaaaatat gccagatctc 600
gactcggttg tcaacattta cgatttcgag ttcttggcgt cccagatttt gaccaaacag 660
gcatggtcct actactcttc tgctgccgat gacgaagtca cccacagaga gaaccatgct 720
gcataccacc gtatcttctt caagccaaga atcctagtca acgtcaagga ggtggacact 780
tccaccacca tgttgggtga aaaagtgggt gttccattct atgtgtctgc taccgcttta 840
tgtaaattgg gtaatccaaa ggaagggtgaa aaggatatcg ctagagggtg tggcgagagc 900
gacgtgaagc caatccagat gatttccact ttggcctctt gttctttgca agaaatcgtc 960
gaagcagcac ctccaagga ccaaatccaa tggttccaat tgtatgtcaa cagtgaccgt 1020
aagattaccg aggaattgat caaaaacgtc gaaaagctgg gcttgaaggc tatcttcgtt 1080
actgtggatg ccccatctct cggtaacaga gaaaaggatg ctaagggtcaa attcaccaac 1140

aaggacagtt ccgcaaaagc catggaaag agcaacgtca aggaatccaa ggggtgcttcc 1200
agagctcttt ccactttcat tgaccctgct ctatgctggg atgatatcgt taccttgaaa 1260
tcaaagacca agttgccaat tgtcattaag ggtgttcaat gcgtcgaaga cgtcttgaaa 1320
gccgcagaaa ttgggtgccgc cgggtgctgc ttgtccaacc acggtggtag acaactagac 1380
ttttccagag cgccaatcga agtcttgccc gaaacaatgc ctattctaaa agaaaagaaa 1440
ctagacgaca agatcgaaat cttcatcgat ggtggtgtca gaagaggtag cgatatattg 1500
aaggccttgt gtcttggtgc caaggggtgc ggtttaggta gaccattctt gtacgcaaac 1560
agttgttacg gtaaggaagg tgtcaagaaa gctatcgaat tgctaaagga cgaactagaa 1620
atgtcaatga gattgcttgg tgtcactagc atcgaccaat tgtctgagaa gtatctggat 1680
ctatctactc ttcacggtag aactgttagc gtacctcgtg acaacttgta caacggagtg 1740
tatgttccac acgaaccaac tgacttcaag gaaaattga 1779

<210> 2
<211> 592
<212> PRT
<213> Kluyveromyces marxianus

<400> 2

Met Arg Ser Ala Ala Arg Val Ile Asn Lys Ser Cys Ser Gly Ser Ala
1 5 10 15

Leu Ser Arg Arg Cys Leu Arg Lys Ser Ser Leu Ser Met Ser Met Arg
20 25 30

Tyr Leu Ser Thr Ser Asn Ile Gly Val Arg Lys Gly Phe Asn Gly Gln
35 40 45

Gly Lys Ser Ser Asn Lys Thr Met Leu Phe Leu Ala Ala Gly Ala Ser
50 55 60

Ala Val Ala Gly Ile Gly Leu Leu Ser Gln Phe Ser Asp Ser Leu Gln
65 70 75 80

Asn Ala Thr Lys Glu Glu Leu Asn Lys Pro Lys Val Ser Pro Leu Glu
85 90 95

Val Ala Lys His Ser Ser Pro Asp Asp Cys Trp Val Val Ile Asp Gly
100 105 110

Phe Val Tyr Asn Leu Thr Glu Phe Ile Ser Ala His Pro Gly Gly Pro
115 120 125

Ala Ile Ile Glu Asn Asn Ala Gly Lys Asp Val Thr Ala Ile Phe Gly
130 135 140

Pro Ile His Ala Pro Asp Val Ile Glu Lys Tyr Ile Ala Pro Glu Asn

145 150 155 160
 Arg Ile Gly Pro Leu Asp Gly Lys Met Pro Asp Asp Leu Ile Cys Ala
 165 170 175
 Pro Leu Thr Pro Gly Glu Thr Pro Glu Asp Val Ala Arg Lys Glu Glu
 180 185 190
 Leu Arg Gln Asn Met Pro Asp Leu Asp Ser Leu Val Asn Ile Tyr Asp
 195 200 205
 Phe Glu Phe Leu Ala Ser Gln Ile Leu Thr Lys Gln Ala Trp Ser Tyr
 210 215 220
 Tyr Ser Ser Ala Ala Asp Asp Glu Val Thr His Arg Glu Asn His Ala
 225 230 235 240
 Ala Tyr His Arg Ile Phe Phe Lys Pro Arg Ile Leu Val Asn Val Lys
 245 250 255
 Glu Val Asp Thr Ser Thr Thr Met Leu Gly Glu Lys Val Gly Val Pro
 260 265 270
 Phe Tyr Val Ser Ala Thr Ala Leu Cys Lys Leu Gly Asn Pro Lys Glu
 275 280 285
 Gly Glu Lys Asp Ile Ala Arg Gly Cys Gly Glu Ser Asp Val Lys Pro
 290 295 300
 Ile Gln Met Ile Ser Thr Leu Ala Ser Cys Ser Leu Gln Glu Ile Val
 305 310 315 320
 Glu Ala Ala Pro Ser Lys Asp Gln Ile Gln Trp Phe Gln Leu Tyr Val
 325 330 335
 Asn Ser Asp Arg Lys Ile Thr Glu Glu Leu Ile Lys Asn Val Glu Lys
 340 345 350
 Leu Gly Leu Lys Ala Ile Phe Val Thr Val Asp Ala Pro Ser Leu Gly
 355 360 365
 Asn Arg Glu Lys Asp Ala Lys Val Lys Phe Thr Asn Lys Asp Ser Ser
 370 375 380
 Ala Lys Ala Met Glu Lys Ser Asn Val Lys Glu Ser Lys Gly Ala Ser
 385 390 395 400
 Arg Ala Leu Ser Thr Phe Ile Asp Pro Ala Leu Cys Trp Asp Asp Ile
 405 410 415
 Val Thr Leu Lys Ser Lys Thr Lys Leu Pro Ile Val Ile Lys Gly Val

420 425 430
 Gln Cys Val Glu Asp Val Leu Lys Ala Ala Glu Ile Gly Ala Ala Gly
 435 440 445
 Val Val Leu Ser Asn His Gly Gly Arg Gln Leu Asp Phe Ser Arg Ala
 450 455 460
 Pro Ile Glu Val Leu Ala Glu Thr Met Pro Ile Leu Lys Glu Lys Lys
 465 470 475 480
 Leu Asp Asp Lys Ile Glu Ile Phe Ile Asp Gly Gly Val Arg Arg Gly
 485 490 495
 Thr Asp Ile Leu Lys Ala Leu Cys Leu Gly Ala Lys Gly Val Gly Leu
 500 505 510
 Gly Arg Pro Phe Leu Tyr Ala Asn Ser Cys Tyr Gly Lys Glu Gly Val
 515 520 525
 Lys Lys Ala Ile Glu Leu Leu Lys Asp Glu Leu Glu Met Ser Met Arg
 530 535 540
 Leu Leu Gly Val Thr Ser Ile Asp Gln Leu Ser Glu Lys Tyr Leu Asp
 545 550 555 560
 Leu Ser Thr Leu His Gly Arg Thr Val Ser Val Pro Arg Asp Asn Leu
 565 570 575
 Tyr Asn Gly Val Tyr Val Pro His Glu Pro Thr Asp Phe Lys Glu Asn
 580 585 590

<210> 3
 <211> 33
 <212> DNA
 <213> Artificial

<220>
 <223> Primer

<400> 3
 cagataacgc gtctggctcg atacttctta ttc

33

<210> 4
 <211> 33
 <212> DNA
 <213> Artificial

<220>
 <223> Primer

<400> 4
 tgcagactgc agtgctactg gtattgtatt act

33

<210> 5

<211> 34
 <212> DNA
 <213> Artificial

<220>
 <223> Primer

<400> 5
 attatagggc cccacaaaat gtatctatct ctcc

34

<210> 6
 <211> 34
 <212> DNA
 <213> Artificial

<220>
 <223> Primer

<400> 6
 tataatgccg gccagaagtc tctggatcct ctcc

34

<210> 7
 <211> 20
 <212> DNA
 <213> Artificial

<220>
 <223> Primer

<400> 7
 cctgaccact aaacgtctac

20

<210> 8
 <211> 20
 <212> DNA
 <213> Artificial

<220>
 <223> Primer

<400> 8
 cacctctttg catgagcttg

20

<210> 9
 <211> 20
 <212> DNA
 <213> Artificial

<220>
 <223> Primer

<400> 9
 ctggatctac agcgattctc

20

<210> 10
 <211> 20
 <212> DNA
 <213> Artificial

<220>
 <223> Primer

<400> 10

tgtaccggct gtccttaaac 20

<210> 11
<211> 32
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 11
tccccggctg accggaactt agcttactcg tc 32

<210> 12
<211> 32
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 12
ttgcgaggat cctgagtga gtagctgtac ac 32

<210> 13
<211> 706
<212> DNA
<213> Kluyveromyces marxianus

<400> 13
cactcgcaag ctgtgccatc gcccaacggt taattataag aaatcaacat cagccaacaa 60
ctattttcgt cccctcttt tcagtggtaa cgagcaatta cattagtaag agactatttt 120
cttcagtgat ttgtaatttt ttttcagtga tttgtaattc tttctcgaaa tatgcgggct 180
waamtaatcc ggacattcac tacatgcaag gaaaaacgag aaccgcggag atttcctcag 240
taagtaacaa tgatgatctt ttacgcttc atcatcactt tccaaagttc taagctataa 300
gttcaagcct agatacgctg aaaaactcct gaccaacaat gtaaagaaaa caattacaat 360
tgtaagggtg aaaacatcta aaaatgaaat attttattgt acatgcacac cctgatagtc 420
attctcttac ttcacccctg aaagacgtgg ctgtacaaga gttggaatcg caaggatcatg 480
aggttaaagt tagtgatctt tatgctcaaa agtgggaaggc ctcaatagac cgtgacgacw 540
wmaaaaaama aamrmaagaa gagagggtta aaatacccca agcttcttat gaagcgtatg 600
ccagaggagc attaacaaaa gacgtaaatc aggaacagga aaaacttatt tgggcggact 660
ttgtcatttt gtcgtttcct atatggtggt cttctatgcc ggctag 706

<210> 14
<211> 30
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 14
tgtcagcatg cactcgcaag ctgtgccatc 30

<210> 15
 <211> 33
 <212> DNA
 <213> Artificial

<220>
 <223> Primer

<400> 15
 aaccttgctg actagccggc atagaagacc acc 33

<210> 16
 <211> 30
 <212> DNA
 <213> Artificial

<220>
 <223> Primer

<400> 16
 tgtcaggatc cactcgcaag ctgtgccatc 30

<210> 17
 <211> 33
 <212> DNA
 <213> Artificial

<220>
 <223> Primer

<400> 17
 aaccttcccg ggtagccggc atagaagacc acc 33

<210> 18
 <211> 33
 <212> DNA
 <213> Artificial

<220>
 <223> Primer

<400> 18
 aaaaaagtcg accgataaca agctcatgca aag 33

<210> 19
 <211> 32
 <212> DNA
 <213> Artificial

<220>
 <223> Primer

<400> 19
 aaaaaaagat ctgcgtaacg aaataaatcc gc 32

<210> 20
 <211> 32
 <212> DNA
 <213> Artificial

<220>

<223> Primer

<400> 20

aaaaaacccg ggcttcctcc gctgcaggtt ca

32

<210> 21

<211> 32

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 21

aaaaaagagc tcgaagtctc tggatcctct tc

32

<210> 22

<211> 33

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 22

aaaaaagacg tcctggctcg atacttctta ttc

33

<210> 23

<211> 32

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 23

aaaaaaggcg ccgctggcag atactagttc ag

32

<210> 24

<211> 20

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 24

cgctgtgctg tgcactgtaa

20

<210> 25

<211> 32

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 25

aaaaaagagc tcgaagtctc tggatcctct tc

32

<210> 26

<211> 32

<212> DNA
 <213> Artificial

<220>
 <223> Primer

<400> 26
 aaaaaaggta ccctgcaatc acccaacgga at 32

<210> 27
 <211> 32
 <212> DNA
 <213> Artificial

<220>
 <223> Primer

<400> 27
 aaaaaatcta gataatgtta cttctcttgg ag 32

<210> 28
 <211> 32
 <212> DNA
 <213> Artificial

<220>
 <223> Primer

<400> 28
 aaaaaacccg gggagaggga gaggataaag ag 32

<210> 29
 <211> 32
 <212> DNA
 <213> Artificial

<220>
 <223> Primer

<400> 29
 aaaaaaacat gttcccatat tggatcatgc tc 32

<210> 30
 <211> 32
 <212> DNA
 <213> Artificial

<220>
 <223> Primer

<400> 30
 aaaaaacata tgtgcagggtg tcgaccagaa tg 32

<210> 31
 <211> 32
 <212> DNA
 <213> Artificial

<220>
 <223> Primer

<400> 31
 aaaaaaggcg ccgcgtaacg aaataaatcc gc 32

<210> 32
 <211> 20
 <212> DNA
 <213> Artificial

<220>
 <223> Primer

<400> 32
 acaccaaagc ctgcatattg 20

<210> 33
 <211> 20
 <212> DNA
 <213> Artificial

<220>
 <223> Primer

<400> 33
 cggtatcgtc gtctcagagt 20

<210> 34
 <211> 20
 <212> DNA
 <213> Artificial

<220>
 <223> Primer

<400> 34
 cggccacaaa gaccacaaag 20

<210> 35
 <211> 20
 <212> DNA
 <213> Artificial

<220>
 <223> Primer

<400> 35
 cggtatcgtc gtctcagagt 20

<210> 36
 <211> 20
 <212> DNA
 <213> Artificial

<220>
 <223> Primer

<400> 36
 ggctgggttg ttgttgttgt 20

<210> 37
 <211> 20
 <212> DNA
 <213> Artificial

<220>

<223> Primer

<400> 37

gtcacagcga atttcctcac

20

<210> 38

<211> 20

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 38

cggccacaaa gaccacaaag

20

<210> 39

<211> 23

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 39

aaccaaagaa ttgttgctgc ttt

23

<210> 40

<211> 25

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 40

ttcgaaaaca cctggtggac cgttc

25

<210> 41

<211> 25

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 41

gcggtgaccc gggagatctg aattc

25

<210> 42

<211> 12

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 42

gaattcagat ct

12

<210> 43

<211> 25

<212> DNA
 <213> Artificial

<220>
 <223> Primer

<400> 43
 gcggtgaccc gggagatctg aattc

25

<210> 44
 <211> 27
 <212> DNA
 <213> Artificial

<220>
 <223> Primer

<400> 44
 ccaattctgc agcaactggc tttaacg

27

<210> 45
 <211> 25
 <212> DNA
 <213> Artificial

<220>
 <223> Primer

<400> 45
 gcggtgaccc gggagatctg aattc

25

<210> 46
 <211> 21
 <212> DNA
 <213> Artificial

<220>
 <223> Primer

<400> 46
 gcctcaccat ttggtctgcc c

21

<210> 47
 <211> 25
 <212> DNA
 <213> Artificial

<220>
 <223> Primer

<400> 47
 gactccccgg agtgtcgaaa tatga

25

<210> 48
 <211> 24
 <212> DNA
 <213> Artificial

<220>
 <223> Primer

<400> 48
 gtgatagcgg gtcctttcgc tacc

24

<210> 49
 <211> 25
 <212> DNA
 <213> Artificial

<220>
 <223> Primer

<400> 49
 gcggtgaccc gggagatctg aattc

25

<210> 50
 <211> 12
 <212> DNA
 <213> Artificial

<220>
 <223> Primer

<400> 50
 gaattcagat ct

12

<210> 51
 <211> 36
 <212> DNA
 <213> Artificial

<220>
 <223> Primer

<400> 51
 gcgatctcga gatttgctgc aacggcaaca tcaatg

36

<210> 52
 <211> 36
 <212> DNA
 <213> Artificial

<220>
 <223> Primer

<400> 52
 ctagcatctg attgttgctg tcgttggttt tggttt

36

<210> 53
 <211> 36
 <212> DNA
 <213> Artificial

<220>
 <223> Primer

<400> 53
 gcgatctcga gatttgctgc aacggcaaca tcaatg

36

<210> 54
 <211> 36
 <212> DNA
 <213> Artificial

<220>

<223> Primer

<400> 54
acttgccat ggtgtgtgtt gttgtcgtt tttttg

36

<210> 55

<211> 44

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 55
ccggaattcg atatctgggc wggkaatgcc aaygarttra atgc

44

<210> 56

<211> 44

<212> DNA

<213> Artificial

<220>

<223> Primer

<220>

<221> misc_feature

<222> (21)..(21)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (33)..(33)

<223> n is a, c, g, or t

<400> 56

cgcggattca ggcctcagta ngaraawgaa ccngtrttra artc

44

<210> 57

<211> 36

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 57

actgtcgagc tcagtatatg gaattgacgg ctcatc

36

<210> 58

<211> 38

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 58

actgacgcgt cgacgtatca tttgtagccc acgccacc

38

<210> 59

<211> 35

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 59

cctcccccg gctgatagaa gggatgatg taatt

35

<210> 60

<211> 24

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 60

ccaagagtta tggggcccca gttg

24

<210> 61

<211> 36

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 61

gcgatctcga gatttgctgc aacggcaaca tcaatg

36

<210> 62

<211> 69

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 62

tggactagta catgcatgcg gtgagaaagt agaaagcaaa cattgttggtt gttgttgctg

60

ttgtttttg

69

<210> 63

<211> 42

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 63

cctcccccg ggatatcaaa gttatattat taatgattca ag

42

<210> 64

<211> 43

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 64

ggacgagagc tcgggacat gacttcagag ttgattttga gtc 43

<210> 65
 <211> 44
 <212> DNA
 <213> Artificial

<220>
 <223> Primer

<400> 65
 ggacatgcat gcgagctcaa tgcgtgacac cgccatgatg gttg 44

<210> 66
 <211> 52
 <212> DNA
 <213> Artificial

<220>
 <223> Primer

<400> 66
 ggacatgcat gctacgtacc ctgcagggca ccaacagcaa caccacctg aa 52

<210> 67
 <211> 43
 <212> DNA
 <213> Artificial

<220>
 <223> Primer

<400> 67
 ggacatgcat gcgagctcat agttgaacaa aactggcat ttg 43

<210> 68
 <211> 52
 <212> DNA
 <213> Artificial

<220>
 <223> Primer

<400> 68
 ggacatgcat gctacgtacc ctgcagggtg gtgcaactag gtttatgtgg ag 52

<210> 69
 <211> 53
 <212> DNA
 <213> Artificial

<220>
 <223> Primer

<400> 69
 cctcccccg ggatatctag ttagatagct cctcctccaa tcgaattatt agc 53

<210> 70
 <211> 42
 <212> DNA
 <213> Artificial

<220>

<223> Primer

<400> 70

ggacgagagc tcgggcccta cgtctatgta tcataaattt gg

42

<210> 71

<211> 31

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 71

aaaaagtcgc cggattttga gtcaccactg g

31

<210> 72

<211> 31

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 72

aaaaagtcga ccagccaata cagagagaga c

31

<210> 73

<211> 31

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 73

aaaaaccatg gaatgtgctt gctcgtgtgt c

31

<210> 74

<211> 31

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 74

aaaaacccgg gtgagtcgga cttgaacaac g

31

<210> 75

<211> 20

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 75

tctccagact cttgttgctg

20

<210> 76

<211> 20
 <212> DNA
 <213> Artificial

<220>
 <223> Primer

<400> 76
 gtgaggaaga agatgcacag 20

<210> 77
 <211> 20
 <212> DNA
 <213> Artificial

<220>
 <223> Primer

<400> 77
 gcctggtgga tattctctcc 20

<210> 78
 <211> 20
 <212> DNA
 <213> Artificial

<220>
 <223> Primer

<400> 78
 gaggacagct aggtttatgc 20

<210> 79
 <211> 1737
 <212> DNA
 <213> Issatchenkia orientalis

<220>
 <221> misc_feature
 <222> (546)..(546)
 <223> n is a, c, g, or t

<400> 79
 atgttactca gatcactaaa ctcttctgct cgtttgtgtca aacaaacaac cagaacaaag 60
 gttaggtatc tcagccacgt cagtgggtgca agcatggcga aacctacatt gaagaacaac 120
 tcgagagaat ccaacaaatc cagaaactat ctaattgctg ctgtgacagc attggctgta 180
 tcaacctcaa ttggagttgc cgtacatgtg aaggaccctt tgtataacga tgctaccggc 240
 agtgattctc cgagaagtat atctgttgac gagtttgtca agcataattc acaaaacgac 300
 tgttgatttg caatcaatgg caaggtttat gatttcactg attttattcc aaaccatcca 360
 ggtgggggtac ctccattagt taatcatgct gggttatgat gtactaaact ttatgagaaa 420
 ttgcatccaa aaggtacaat tgagaaattc ttgccaaagg ataagtttct ggggtgtgta 480
 gatggtgaag cgccaaaatt ggaagcagac tatttggttg acgatgatga acaagagcga 540
 ctggantatt tgggcaactt acctccttgg tcattctattc agaatgttta tgatttcgaa 600
 tacttggcca agaagatttt acctaaagat gcctgggcat attattcttg tgggtgccgat 660

gatgaaatca" caatgagaga" aadccattat" gcttatcaaa gagtttattt cagaccaaga 720
 atttgtgttg atgtcaagga agttgatact tcttatgaaa tgtaggcac taaaacctct 780
 gttccttttt atgtatctgc caccgcttg gctaaattag gccatcctga tggatgaatgc 840
 tcaattgcta gaggcgctgg taaggaaggt gtcgttcaaa tgatttcgac cctttcctca 900
 atgtcattag atgaaattgc cgctgctaga attccaggtg caaccaatg gttccaatta 960
 tacattaatg aggatagaaa tgctgctaaa ggtctggtca aacatgcaga agacttgggt 1020
 atgaaggcta tctttataac tgttgatgct ccttctctag gtaacagaga aaaggataaa 1080
 agattaaagt ttgttaatga caccgatgct gatttgggtg attccgcaga tcgaaacagt 1140
 ggtgcttcaa aggactatc ttcgttcatt gatgcttctg tctcttgga tgacgtcaaa 1200
 gcggtcaagt cgtggactaa attgcctgct ttagttaag gtgttcaaac agttgaagac 1260
 gttattgaag cttacgatgc tggttgtcaa ggtgtgttt tgtcaaacca cgggtgtagg 1320
 caactagata ctgctcctcc tccaatcgaa ttattagctg aaactgttcc aactttgaag 1380
 agattgggta aattaagacc agattttgaa attttaattg acggtggtgt caaaagaggt 1440
 accgatattt tgaaagcagt cgcaatcggg ggccaagatg tcagagtttc agttggtatg 1500
 ggtagacctt tcttatatgc caactcttgc tatggtgaag caggtgtagg aaaattaatt 1560
 caaaatctaa aggatgaatt agaaatggat atgagattgt tgggtgtcac taaaatggac 1620
 cagctatctt cgaaacatgt cgatactaaa cgtttgattg gtagagatgc gatcaactat 1680
 ttgtatgata atgtatacag cccaatcgaa accgttaaatt tcaacaatga agattga 1737

<210> 80
 <211> 577
 <212> PRT
 <213> Issatchenkia orientalis

<400> 80

Met Leu Leu Arg Ser Leu Asn Ser Ser Ala Arg Cys Val Lys Gln Thr
1 5 10 15

Thr Arg Thr Lys Val Arg Tyr Leu Ser His Val Ser Gly Ala Ser Met
20 25 30

Ala Lys Pro Thr Leu Lys Asn Asn Ser Arg Glu Ser Asn Lys Ser Arg
35 40 45

Asn Tyr Leu Ile Ala Ala Val Thr Ala Leu Ala Val Ser Thr Ser Ile
50 55 60

Gly Val Ala Val His Val Lys Asp Pro Leu Tyr Asn Asp Ala Thr Gly
65 70 75 80

Ser Asp Ser Pro Arg Ser Ile Ser Val Asp Glu Phe Val Lys His Asn
85 90 95

Ser Gln Asn Asp Cys Trp Ile Ala Ile Asn Gly Lys Val Tyr Asp Phe
100 105 110

Thr Asp Phe Ile Pro Asn His Pro Gly Gly Val Pro Pro Leu Val Asn
115 120 125

His Ala Gly Tyr Asp Gly Thr Lys Leu Tyr Glu Lys Leu His Pro Lys
130 135 140

Gly Thr Ile Glu Lys Phe Leu Pro Lys Asp Lys Phe Leu Gly Val Leu
145 150 155 160

Asp Gly Glu Ala Pro Lys Leu Glu Ala Asp Tyr Leu Val Asp Asp Asp
165 170 175

Glu Gln Glu Arg Leu Tyr Leu Gly Asn Leu Pro Pro Leu Ser Ser Ile
180 185 190

Gln Asn Val Tyr Asp Phe Glu Tyr Leu Ala Lys Lys Ile Leu Pro Lys
195 200 205

Asp Ala Trp Ala Tyr Tyr Ser Cys Gly Ala Asp Asp Glu Ile Thr Met
210 215 220

Arg Glu Asn His Tyr Ala Tyr Gln Arg Val Tyr Phe Arg Pro Arg Ile
225 230 235 240

Cys Val Asp Val Lys Glu Val Asp Thr Ser Tyr Glu Met Leu Gly Thr
245 250 255

Lys Thr Ser Val Pro Phe Tyr Val Ser Ala Thr Ala Leu Ala Lys Leu
260 265 270

Gly His Pro Asp Gly Glu Cys Ser Ile Ala Arg Gly Ala Gly Lys Glu
275 280 285

Gly Val Val Gln Met Ile Ser Thr Leu Ser Ser Met Ser Leu Asp Glu
290 295 300

Ile Ala Ala Ala Arg Ile Pro Gly Ala Thr Gln Trp Phe Gln Leu Tyr
305 310 315 320

Ile Asn Glu Asp Arg Asn Val Ala Lys Gly Leu Val Lys His Ala Glu
325 330 335

Asp Leu Gly Met Lys Ala Ile Phe Ile Thr Val Asp Ala Pro Ser Leu
340 345 350

Gly Asn Arg Glu Lys Asp Lys Arg Leu Lys Phe Val Asn Asp Thr Asp
355 360 365

Val Asp Leu Gly Asp Ser Ala Asp Arg Asn Ser Gly Ala Ser Lys Ala
370 375 380

Leu Ser Ser Phe Ile Asp Ala Ser Val Ser Trp Asn Asp Val Lys Ala
385 390 395 400

Val Lys Ser Trp Thr Lys Leu Pro Val Leu Val Lys Gly Val Gln Thr
405 410 415

Val Glu Asp Val Ile Glu Ala Tyr Asp Ala Gly Cys Gln Gly Val Val
420 425 430

Leu Ser Asn His Gly Gly Arg Gln Leu Asp Thr Ala Pro Pro Pro Ile
435 440 445

Glu Leu Leu Ala Glu Thr Val Pro Thr Leu Lys Arg Leu Gly Lys Leu
450 455 460

Arg Pro Asp Phe Glu Ile Leu Ile Asp Gly Gly Val Lys Arg Gly Thr
465 470 475 480

Asp Ile Leu Lys Ala Val Ala Ile Gly Gly Gln Asp Val Arg Val Ser
485 490 495

Val Gly Met Gly Arg Pro Phe Leu Tyr Ala Asn Ser Cys Tyr Gly Glu
500 505 510

Ala Gly Val Arg Lys Leu Ile Gln Asn Leu Lys Asp Glu Leu Glu Met
515 520 525

Asp Met Arg Leu Leu Gly Val Thr Lys Met Asp Gln Leu Ser Ser Lys
530 535 540

His Val Asp Thr Lys Arg Leu Ile Gly Arg Asp Ala Ile Asn Tyr Leu
545 550 555 560

Tyr Asp Asn Val Tyr Ser Pro Ile Glu Thr Val Lys Phe Asn Asn Glu
565 570 575

Asp

<210> 81
<211> 1698
<212> DNA
<213> Issatchenkia orientalis

<220>
<221> misc_feature
<222> (941)..(941)
<223> n i s a , c , g , o u t

<220>
 <221> mrsc_Feature
 <222> (943)..(943)
 <223> n is a, c, g, ou t

<400> 81
 atgttaagat cccagttcaa aaacattttg aaaaatgtta acaagaacca ttctctaagg 60
 agaactttta cttccagcac ctcaaaggct ggaaaaaatg cttcatacaa tgccaagatt 120
 atatctgcaa ccgtggcctc gattgttgca gcagctggct cttatatgtt ggtccagcct 180
 tcactagcta atgatgaggc acagtctgct aatccaacta ggaagatctc tgttgacgaa 240
 tttgttaaac acaaccatgc cgatgattgt tggatcactg ttaacggtaa cgtctatgac 300
 ttgactgatt tcatttcaat gcatccagggt ggtactaccc cattgattca aaatgcagggt 360
 cacgacgcaa ctgaaattta caacaagatt catccaaagg gtacaatcga gaacttctta 420
 ccaaaggaaa agcaattggg tgttttgat ggtgaagctc ctaaaatcga agttgtgctt 480
 gacgaaaagg agaaacacag attggagtgt ttgaatcatc tccctgctct tccagaatt 540
 caaaacattt atgatttcga acatattgct tctagagttt tgagcgacca agcatggaac 600
 tactattcat gtggtgccga agatgaaatc accttgaggg aaaatcatta tgcttaccaa 660
 agaatctact ttaagccaaa atgttgtgtc aatgttgag aagttgatac ctctcatgaa 720
 attttaggta caaaagcttc tgttccttc tacgtttccg cagccgcttc tgcaaagttg 780
 gggcacgagg atggtgaatg ttccattgct agagggtcag gtaagggaagg cgttattcaa 840
 atgatttctt cttctcttc caactcttg gaggaattg cagaatccag aattcctggt 900
 gcaacacaat ggtttcaatt atacgttaat gaagacaagg ntnttgtaa gaagacttta 960
 aaaagggccg aaaacttggg tatgaaggcc atctttgtca ctgtggacgc tgctagtaga 1020
 ggtaatagag aaaaagacat tacaatgaga attaccgaag atacagatga gttaatagac 1080
 gattcttctg ttagagctgg ttctacctct ggtgcattgc cagctttcat tgacaagagg 1140
 ctgacttggg atgaagttaa ggatatcatt tcatggacca agttaccagt tttgctgaag 1200
 ggtgttcaaa gaactgatga tattgagaag gcaattgata ttggttgtaa ggggtgtgtc 1260
 ttgtccaatc atggtggtag acaattagat acttctcctc ctccaataga agttatggct 1320
 gaatctgttc caatcctaaa gcaaaagggt aaactggatc caaatttcag tattttcggt 1380
 gatggtggtg ttagaagagg tacagatatt ttgaaagctt tggctattgg tggcagagac 1440
 tgtaaagttg ctgttggtct gggtagacct ttcctttatg caaatactgg ttatggtgaa 1500
 aagggtgtca gaaaggccgt gcaatttcta agagaagaat taaaggctga catgagaatg 1560
 ttgggcgtta cctctttgaa cgagctagac gactcttaca ttgacaccag aagattacta 1620
 ggtagagatg ctgttaacca catatacaac aacaactact acccaatgtc taagattcaa 1680
 ttcaaaaacg aaaaataa 1698

<210> 82
 <211> 565
 <212> PRT

<213> ~~Issatchenkia orientalis~~

<220>

<221> misc_feature

<222> (314)..(315)

<223> Xaa pode ser qualquer aminoácido de ocorrência natural

<400> 82

Met Leu Arg Ser Gln Phe Lys Asn Ile Leu Lys Asn Val Asn Lys Asn
1 5 10 15

His Ser Leu Arg Arg Thr Phe Thr Ser Ser Thr Ser Lys Ala Gly Lys
20 25 30

Asn Ala Ser Tyr Asn Ala Lys Ile Ile Ser Ala Thr Val Ala Ser Ile
35 40 45

Val Ala Ala Ala Gly Ser Tyr Met Leu Val Gln Pro Ser Leu Ala Asn
50 55 60

Asp Glu Ala Gln Ser Ala Asn Pro Thr Arg Lys Ile Ser Val Asp Glu
65 70 75 80

Phe Val Lys His Asn His Ala Asp Asp Cys Trp Ile Thr Val Asn Gly
85 90 95

Asn Val Tyr Asp Leu Thr Asp Phe Ile Ser Met His Pro Gly Gly Thr
100 105 110

Thr Pro Leu Ile Gln Asn Ala Gly His Asp Ala Thr Glu Ile Tyr Asn
115 120 125

Lys Ile His Pro Lys Gly Thr Ile Glu Asn Phe Leu Pro Lys Glu Lys
130 135 140

Gln Leu Gly Val Leu Asp Gly Glu Ala Pro Lys Ile Glu Val Val Leu
145 150 155 160

Asp Glu Lys Glu Lys His Arg Leu Glu Leu Leu Asn His Leu Pro Ala
165 170 175

Leu Ser Arg Ile Gln Asn Ile Tyr Asp Phe Glu His Ile Ala Ser Arg
180 185 190

Val Leu Ser Asp Gln Ala Trp Asn Tyr Tyr Ser Cys Gly Ala Glu Asp
195 200 205

Glu Ile Thr Leu Arg Glu Asn His Tyr Ala Tyr Gln Arg Ile Tyr Phe
210 215 220

Lys Pro Lys Cys Cys Val Asn Val Ala Glu Val Asp Thr Ser His Glu
225 230 235 240

Ile Leu Gly Thr Lys Ala Ser Val Pro Phe Tyr Val Ser Ala Ala Ala
 245 250 255
 Ser Ala Lys Leu Gly His Glu Asp Gly Glu Cys Ser Ile Ala Arg Gly
 260 265 270
 Ala Gly Lys Glu Gly Val Ile Gln Met Ile Ser Ser Phe Ser Ser Asn
 275 280 285
 Ser Leu Glu Glu Ile Ala Glu Ser Arg Ile Pro Gly Ala Thr Gln Trp
 290 295 300
 Phe Gln Leu Tyr Val Asn Glu Asp Lys Xaa Xaa Val Lys Lys Thr Leu
 305 310 315 320
 Lys Arg Ala Glu Asn Leu Gly Met Lys Ala Ile Phe Val Thr Val Asp
 325 330 335
 Ala Ala Ser Arg Gly Asn Arg Glu Lys Asp Ile Thr Met Arg Ile Thr
 340 345 350
 Glu Asp Thr Asp Glu Leu Ile Asp Asp Ser Ser Val Arg Ala Gly Ser
 355 360 365
 Thr Ser Gly Ala Leu Pro Ala Phe Ile Asp Lys Arg Leu Thr Trp Asp
 370 375 380
 Glu Val Lys Asp Ile Ile Ser Trp Thr Lys Leu Pro Val Leu Leu Lys
 385 390 395 400
 Gly Val Gln Arg Thr Asp Asp Ile Glu Lys Ala Ile Asp Ile Gly Cys
 405 410 415
 Lys Gly Val Val Leu Ser Asn His Gly Gly Arg Gln Leu Asp Thr Ser
 420 425 430
 Pro Pro Pro Ile Glu Val Met Ala Glu Ser Val Pro Ile Leu Lys Gln
 435 440 445
 Lys Gly Lys Leu Asp Pro Asn Phe Ser Ile Phe Val Asp Gly Gly Val
 450 455 460
 Arg Arg Gly Thr Asp Ile Leu Lys Ala Leu Ala Ile Gly Gly Arg Asp
 465 470 475 480
 Cys Lys Val Ala Val Gly Leu Gly Arg Pro Phe Leu Tyr Ala Asn Thr
 485 490 495
 Gly Tyr Gly Glu Lys Gly Val Arg Lys Ala Val Gln Ile Leu Arg Glu
 500 505 510

Glu Leu Lys Ala Asp Met Arg Met Leu Gly Val Thr Ser Leu Asn Glu
 515 520 525

Leu Asp Asp Ser Tyr Ile Asp Thr Arg Arg Leu Leu Gly Arg Asp Ala
 530 535 540

Val Asn His Ile Tyr Asn Asn Asn Tyr Tyr Pro Met Ser Lys Ile Gln
 545 550 555 560

Phe Lys Asn Glu Lys
 565

<210> 83
 <211> 1778
 <212> DNA
 <213> Kluyveromyces marxianus

<400> 83
 atgctacca ggttcgtggt tagatctggt gccgcaggaa gaaacctggg cttcagtttc 60
 actcgtaa at gcgaccatac ttttctaagt aagactgtgc gcaatgactt gtcgcacaga 120
 ccctattcta ccggaactaa cggtaatggc agtgccaatg ccggtaaaag ccaaagcggt 180
 ttgttggttg gtgttttttg aggtactttg atcgggtggtg ggctgggttg gtactttttg 240
 gggtcgaagt tcgaccagca acagagctca tctcagcaag tcagcgattt gtccattgct 300
 agactagagg atttggtact gcccaagtac tgcgacaaga agacgttttg tactgctgtc 360
 gaggagtga aacagggtgt ggataacaac ccagagaact ttcgggacgc caagagcgac 420
 ttagactcgc actcagacac atatttcaac tcgcaccatg ccacgccaga acagagaccc 480
 gaaatcggtt tgttcccccg taacacggaa gacgtttcca agttactcaa gatatgccac 540
 aagtactcca ttcctgtcat tccattctcc ggtgggacct cgctcgaggg ccacttcattg 600
 cccacgagac cgggctcctg cgctggtttg gacatctcca agtacatgaa ccaattatcc 660
 agttgaacaa ggaagacctc gacgtggtcg tgcaagggtg tgttccatgg gaagacttga 720
 acgactactt gaacgaccac gggttgttgt tcggttgcca tcctgggtcca ggtgcccaga 780
 tcgctgggtg tattgctaac tcttggtccg gaaccaatgc gtaccgttac ggtactatga 840
 aagaaaacgt cgtcaacatc accatgtgtc ttgccgacgg taccatcatc aagaccaaga 900
 gaagaccaag aaaatcctct gctgggttaca acttgaacgg tttgatcatc ggaagcgagg 960
 gtaccttggg tattgtcacc gaggcaacaa taaagtgtca cgtgagatct aactttgaaa 1020
 ccgtcgtgtg tgttccgttc ccagcgtgg ccgatgctgc ctcgtgctcc tcccacttga 1080
 tccaggccgg tatccagctg aatgccatgg aattgttga cgataacatg atgaagatca 1140
 tcaacaagag cgggtgtact tcaagaacaa actgggtcga atcaccaacc ttattcttca 1200
 agattggtgg gagaagcgaa aaggccatca aagaagtgg taaagaggtc gagaagatcg 1260
 cctctcaaca caacaacagc aacttcgaat ttgcatctga tgaagaaact aaattggaat 1320

tgtggaagc tagaaaagtc gctgttgggt ctacgattga tgcaggtaag aagttggatc 1380
 caaacgtcaa cgtttgacc accgatgttg ccgttcctat ctccaaattc gcccagggtta 1440
 tcaacgatac aaaggaggaa atgaacgctt ctggattatt gacctctctt gttgggtcatg 1500
 ccggtgacgg taatttccat gctttcatca tttaaatgc cgaacaacgt aagaccgcag 1560
 aaaccattgt cgaaaatatg gtcaagagag ccattgacgc agaaggtacc tgtaccggtg 1620
 aacacggtgt tggattgggt aagagagagt ttttggttga agagttgggc gaagatacaa 1680
 ttgctgtcat gagaaaattg aaacttgctt tggatccaaa gaggatcttg aaccctgaca 1740
 aggtcttcaa gattgaccct aacgatcacc aacactga 1778

<210> 84
 <211> 592
 <212> PRT
 <213> Kluyveromyces marxianus

<400> 84

Met Leu Pro Arg Phe Val Val Arg Ser Gly Ala Ala Gly Arg Asn Leu
1 5 10 15

Gly Phe Ser Phe Thr Arg Lys Cys Asp His Thr Phe Leu Ser Lys Thr
20 25 30

Val Arg Asn Asp Leu Ser His Arg Pro Tyr Ser Thr Gly Thr Asn Gly
35 40 45

Asn Gly Ser Ala Asn Ala Gly Lys Ser Gln Ser Gly Leu Leu Phe Gly
50 55 60

Val Phe Gly Gly Thr Leu Ile Gly Gly Gly Leu Val Ala Tyr Phe Leu
65 70 75 80

Gly Ser Lys Phe Asp Gln Gln Gln Ser Ser Ser Gln Gln Val Ser Asp
85 90 95

Leu Ser Ile Ala Arg Leu Glu Asp Leu Asp Ser Pro Lys Tyr Cys Asp
100 105 110

Lys Lys Thr Phe Ala Thr Ala Val Glu Glu Leu Lys Gln Val Leu Asp
115 120 125

Asn Asn Pro Glu Asn Phe Ser Asp Ala Lys Ser Asp Leu Asp Ser His
130 135 140

Ser Asp Thr Tyr Phe Asn Ser His His Ala Thr Pro Glu Gln Arg Pro
145 150 155 160

Glu Ile Val Leu Phe Pro Arg Asn Thr Glu Asp Val Ser Lys Leu Leu
165 170 175

Lys Ile Cys His Lys Tyr Ser Ile Pro Val Ile Pro Phe Ser Gly Gly
180 185 190

Thr Ser Leu Glu Gly His Phe Met Pro Thr Arg Pro Gly Ser Cys Val
195 200 205

Val Leu Asp Ile Ser Lys Tyr Met Asn Gln Ile Ile Gln Leu Asn Lys
210 215 220

Glu Asp Leu Asp Val Val Val Gln Gly Gly Val Pro Trp Glu Asp Leu
225 230 235 240

Asn Asp Tyr Leu Asn Asp His Gly Leu Leu Phe Gly Cys Asp Pro Gly
245 250 255

Pro Gly Ala Gln Ile Ala Gly Cys Ile Ala Asn Ser Cys Ser Gly Thr
260 265 270

Asn Ala Tyr Arg Tyr Gly Thr Met Lys Glu Asn Val Val Asn Ile Thr
275 280 285

Met Cys Leu Ala Asp Gly Thr Ile Ile Lys Thr Lys Arg Arg Pro Arg
290 295 300

Lys Ser Ser Ala Gly Tyr Asn Leu Asn Gly Leu Ile Ile Gly Ser Glu
305 310 315 320

Gly Thr Leu Gly Ile Val Thr Glu Ala Thr Ile Lys Cys His Val Arg
325 330 335

Ser Asn Phe Glu Thr Val Ala Val Val Pro Phe Pro Ser Val Ala Asp
340 345 350

Ala Ala Ser Cys Ser Ser His Leu Ile Gln Ala Gly Ile Gln Leu Asn
355 360 365

Ala Met Glu Leu Leu Asp Asp Asn Met Met Lys Ile Ile Asn Lys Ser
370 375 380

Gly Ala Thr Ser Arg Thr Asn Trp Val Glu Ser Pro Thr Leu Phe Phe
385 390 395 400

Lys Ile Gly Gly Arg Ser Glu Lys Ala Ile Lys Glu Val Val Lys Glu
405 410 415

Val Glu Lys Ile Ala Ser Gln His Asn Asn Ser Asn Phe Glu Phe Ala
420 425 430

Ser Asp Glu Glu Thr Lys Leu Glu Leu Trp Glu Ala Arg Lys Val Ala
435 440 445

Leu Trp Ser Thr Ile Asp Ala Gly Lys Lys Leu Asp Pro Asn Val Asn
450 455 460

Val Trp Thr Thr Asp Val Ala Val Pro Ile Ser Lys Phe Ala Gln Val
465 470 475 480

Ile Asn Asp Thr Lys Glu Glu Met Asn Ala Ser Gly Leu Leu Thr Ser
485 490 495

Leu Val Gly His Ala Gly Asp Gly Asn Phe His Ala Phe Ile Ile Tyr
500 505 510

Asn Ala Glu Gln Arg Lys Thr Ala Glu Thr Ile Val Glu Asn Met Val
515 520 525

Lys Arg Ala Ile Asp Ala Glu Gly Thr Cys Thr Gly Glu His Gly Val
530 535 540

Gly Ile Gly Lys Arg Glu Phe Leu Val Glu Glu Leu Gly Glu Asp Thr
545 550 555 560

Ile Ala Val Met Arg Lys Leu Lys Leu Ala Leu Asp Pro Lys Arg Ile
565 570 575

Leu Asn Pro Asp Lys Val Phe Lys Ile Asp Pro Asn Asp His Gln His
580 585 590

<210> 85
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 85
gcgatatcaa tggatgatacc tc

22

<210> 86
<211> 21
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 86
gaaccgaagc gtacttgaaa g

21

<210> 87
<211> 32
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 87
 attaatgagc tctatgtgag tgcgtcttea gg

32

<210> 88
 <211> 34
 <212> DNA
 <213> Artificial

<220>
 <223> Primer

<400> 88
 ataatttcta gatgttgatg atgtgttgaa ttg

34

<210> 89
 <211> 33
 <212> DNA
 <213> Artificial

<220>
 <223> Primer

<400> 89
 gaaggatccg tgttgaatta ggtaagggtc tag

33

<210> 90
 <211> 33
 <212> DNA
 <213> Artificial

<220>
 <223> Primer

<400> 90
 gaaactagtc ttccagtaac catatctttc ctg

33

REIVINDICAÇÕES

1. Célula de levedura geneticamente modificada, **CARACTERIZADA** pelo fato de que tem (a) pelo menos um gene de lactato desidrogenase (LDH) exógeno funcional integrado no seu genoma e (b) uma deleção ou ruptura de (i) pelo menos um gene nativo de L- ou D-lactato:ferricitocromo c oxidoredutase.

2. Célula de levedura geneticamente modificada, de acordo com a reivindicação 1, **CARACTERIZADA** pelo fato de que tem (a) uma deleção ou ruptura de pelo menos um gene L-lactato:ferricitocromo c oxidoredutase e (b) pelo menos um gene de L-lactato desidrogenase exógeno funcional integrado no seu genoma.

3. Célula de levedura geneticamente modificada, de acordo com a reivindicação 2, **CARACTERIZADA** pelo fato de que é uma célula do gêneros *Candida*, *Saccharomyces*, *Shizosaccharomyces*, *Kluyveromyces*, *Pichia*, *Issachenkia* ou *Hansenula*.

4. Célula de levedura geneticamente modificada, de acordo com a reivindicação 3, **CARACTERIZADA** pelo fato de que é uma célula das espécies *K. marxianus*, *S. cerevisiae*, *C. sonorensis*, *S. bulderi* ou *I. orientalis*.

5. Célula de levedura geneticamente modificada, de acordo com a reivindicação 2, **CARACTERIZADA** pelo fato de que adicionalmente tem uma deleção ou ruptura de pelo menos um gene piruvato descarboxilase.

6. Célula de levedura geneticamente modificada, de acordo com a reivindicação 2, **CARACTERIZADA** pelo fato de que o gene deletado ou interrompido L-lactato:ferricitocromo c

oxidoreductase é um gene tendo uma sequência identificada como SEQ ID NO. 1, SEQ. ID. NO. 79 ou SEQ. ID. NO. 81, ou é pelo menos 40 % homóloga a qualquer um de tais genes.

7. Célula de levedura geneticamente modificada, de acordo com a reivindicação 2, **CARACTERIZADA** pelo fato de que o gene deletado ou interrompido L-lactato:ferricitocromo c oxidoreductase codifica para uma enzima tendo uma sequência de aminoácidos identificada como SEQ. ID. NO. 2, SEQ. ID. NO. 80 ou SEQ. ID. NO. 82, ou uma enzima que tem uma sequência de aminoácidos que é pelo menos 40 % homóloga a qualquer uma destas.

8. Célula de levedura geneticamente modificada, de acordo com a reivindicação 1, **CARACTERIZADA** pelo fato de que tem (a) uma deleção ou ruptura de pelo menos um gene D-lactato:ferricitocromo c oxidoreductase e (b) pelo menos um gene de D-lactato desidrogenase exógeno funcional integrado no seu genoma.

9. Célula de levedura geneticamente modificada, de acordo com a reivindicação 8, **CARACTERIZADA** pelo fato de que é uma célula do gênero *Candida*, *Saccharomyces*, *Shizosaccharomyces*, *Kluyveromyces*, *Pichia*, *Issachenkia* ou *Hansenula*.

10. Célula de levedura geneticamente modificada, de acordo com a reivindicação 9, **CARACTERIZADA** pelo fato de que é uma célula das espécies *K. marxianus*, *S. cerevisiae*, *C. sonorensis*, *S. bulderi* ou *I. orientalis*.

11. Célula de levedura geneticamente modificada, de acordo com a reivindicação 8, **CARACTERIZADA** pelo fato de que adicionalmente tem uma deleção ou ruptura de pelo menos

um gene piruvato descarboxilase.

12. Célula de levedura geneticamente modificada, de acordo com a reivindicação 8, **CARACTERIZADA** pelo fato de que o gene deletado ou interrompido D-lactato:ferricitocromo c oxidoreductase é um gene tendo uma sequência identificada como SEQ. ID. NO. 83 (o gene D-lactato:ferricitocromo c oxidoreductase (*DLD1*) de uma cepa de *K. marxianus* tipo selva-gem), o gene *DLD1* de *S. cerevisiae*, o gene *DLD1* de *I. orientalis*, ou é um gene que é pelo menos 40 % homólogo a pelo menos um de tais genes.

13. Célula de levedura geneticamente modificada, de acordo com a reivindicação 8, **CARACTERIZADA** pelo fato de que o gene deletado ou interrompido D-lactato:ferricitocromo c oxidoreductase codifica para uma enzima tendo uma sequência de aminoácidos identificada como SEQ. ID. NO. 84, uma proteína tendo a sequência de aminoácidos de uma proteína codificada pelo gene *DLD1* de *S. cerevisiae*, uma proteína tendo a sequência de aminoácidos de uma proteína codificada pelo gene *DLD1* de *I. orientalis*, ou uma proteína que é pelo menos 40 % homóloga a uma destas.

14. Processo de fermentação, **CARACTERIZADO** pelo fato de que uma célula de levedura geneticamente modificada, conforme definida em qualquer uma das reivindicações 1-13 é cultivada em condições de fermentação em um caldo de fermentação que inclui um açúcar fermentável para produzir ácido lático ou um sal deste.

Fig. 1

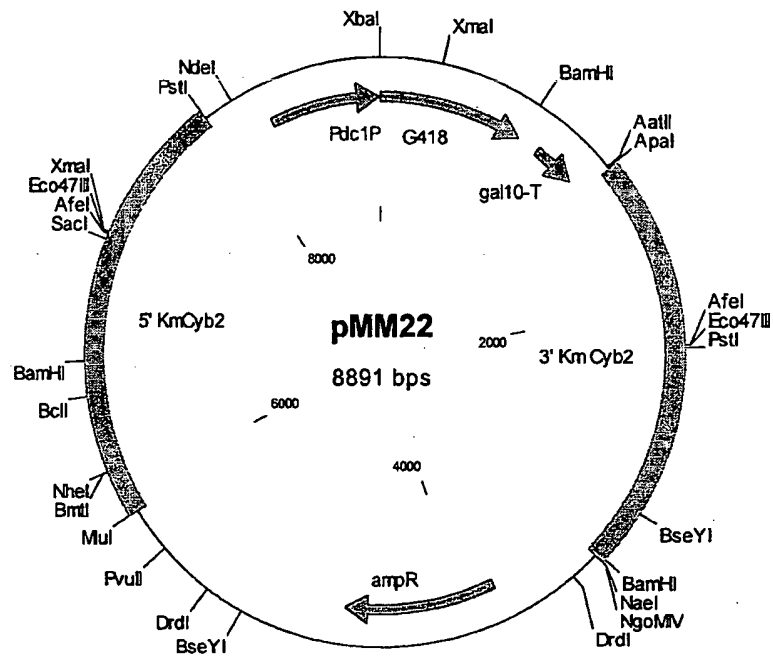


Fig. 2

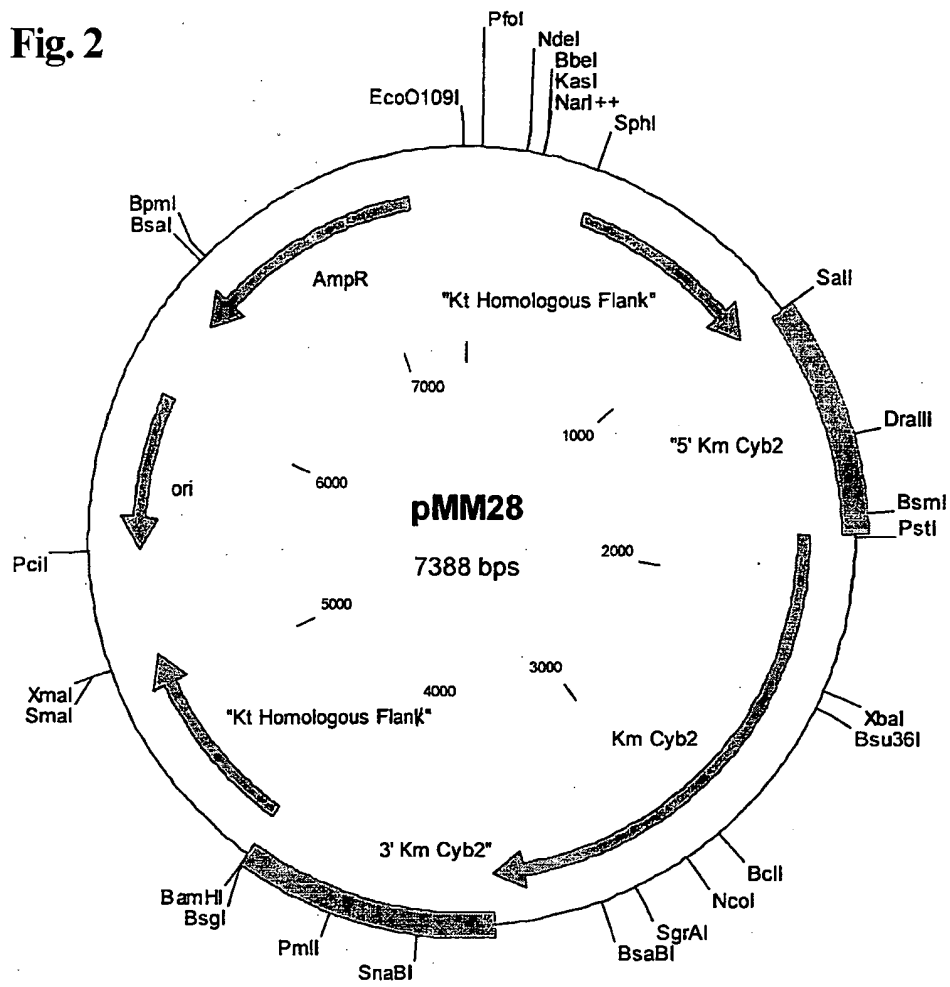


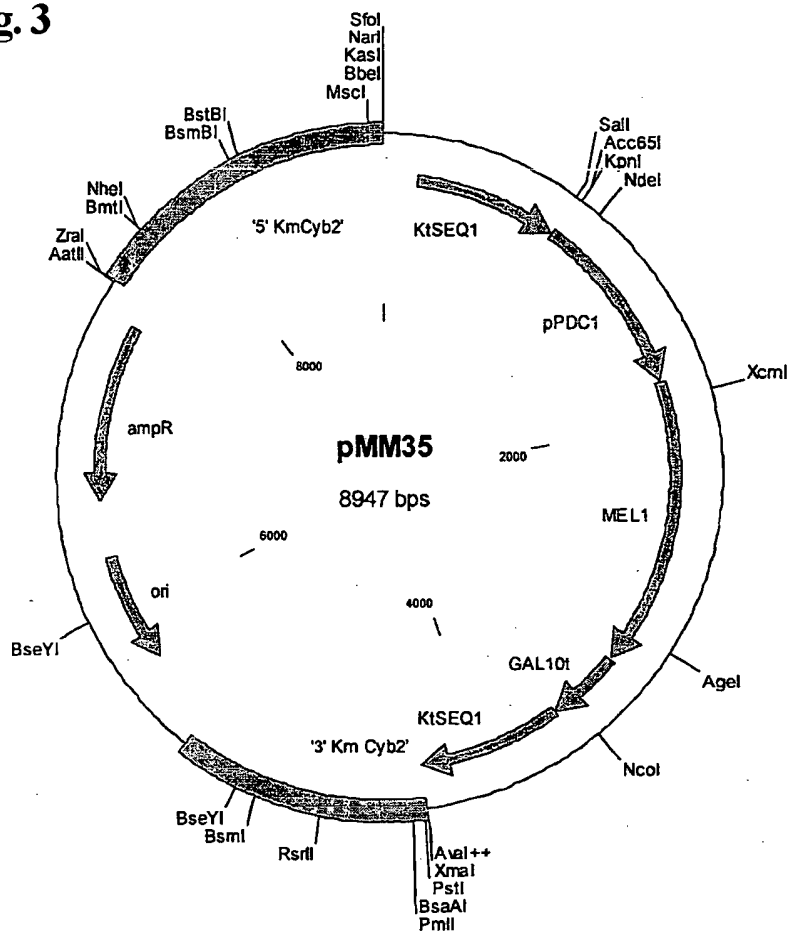
Fig. 3

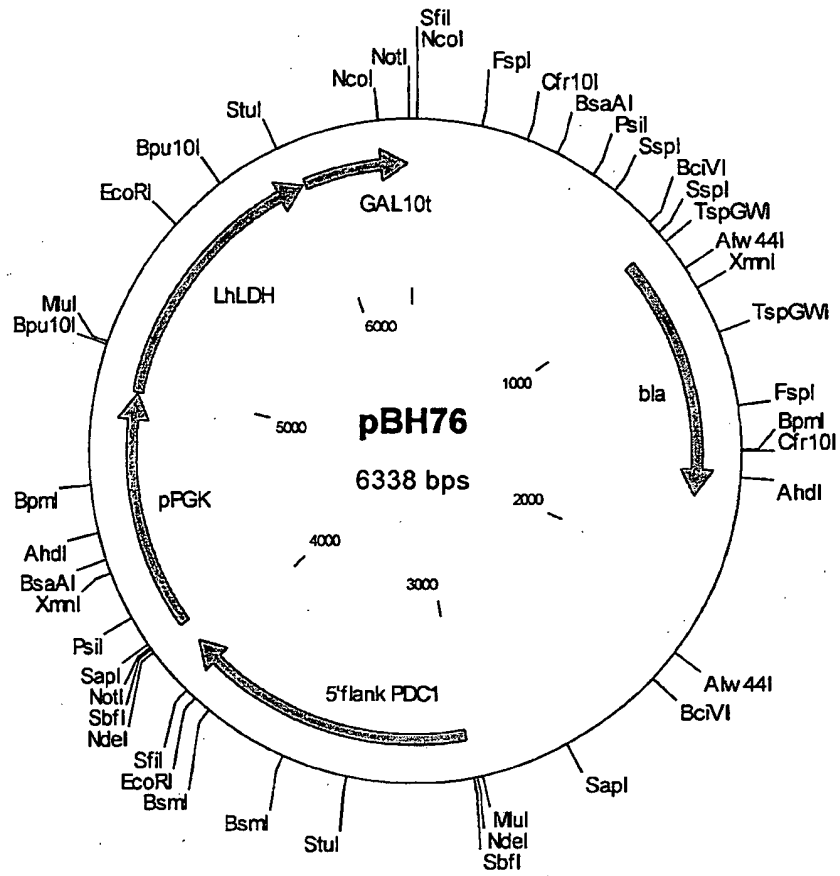
Fig. 4

Fig. 5

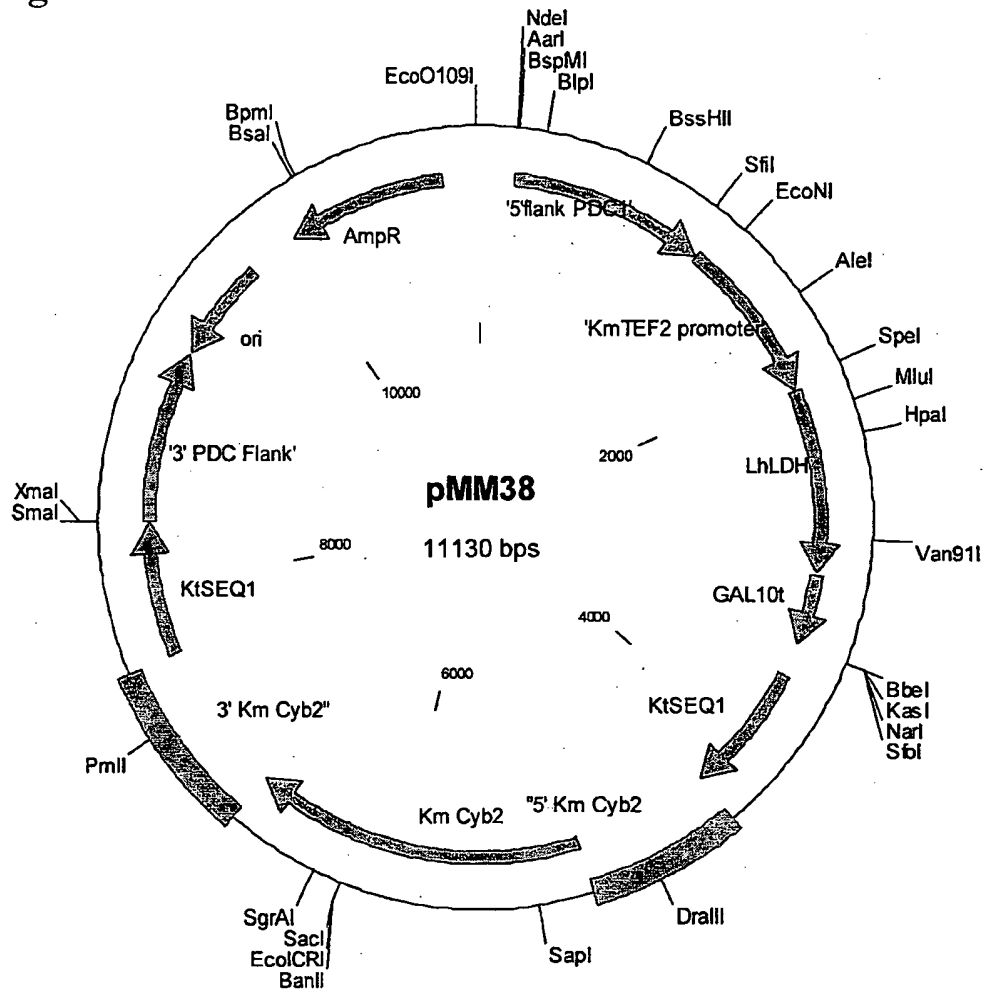


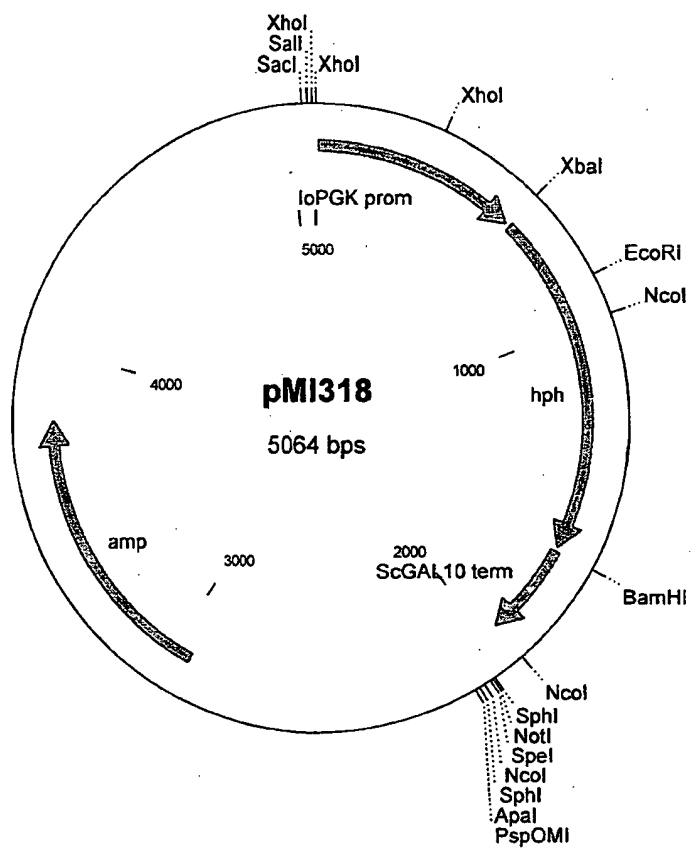
Fig. 6

Fig. 7

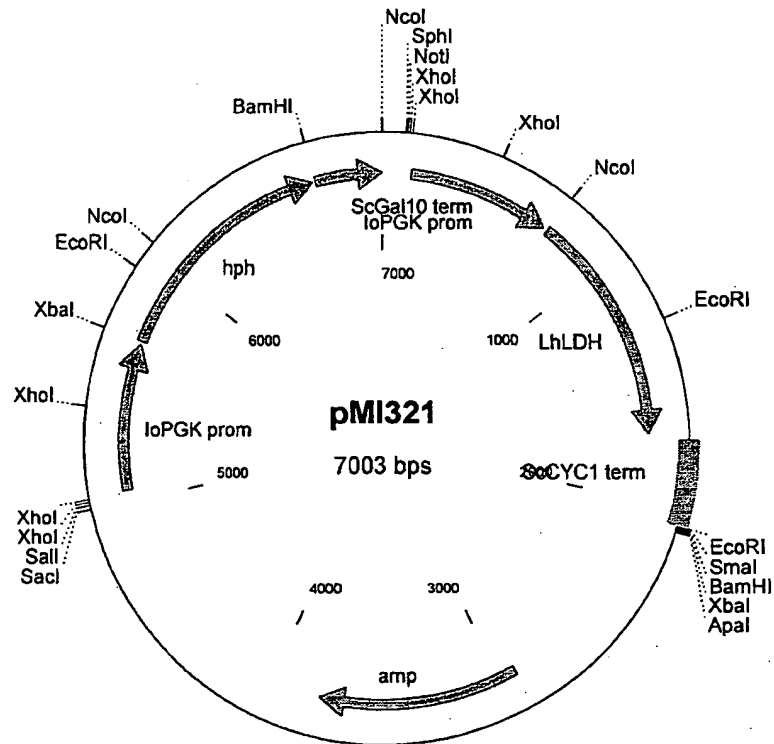


Fig. 8

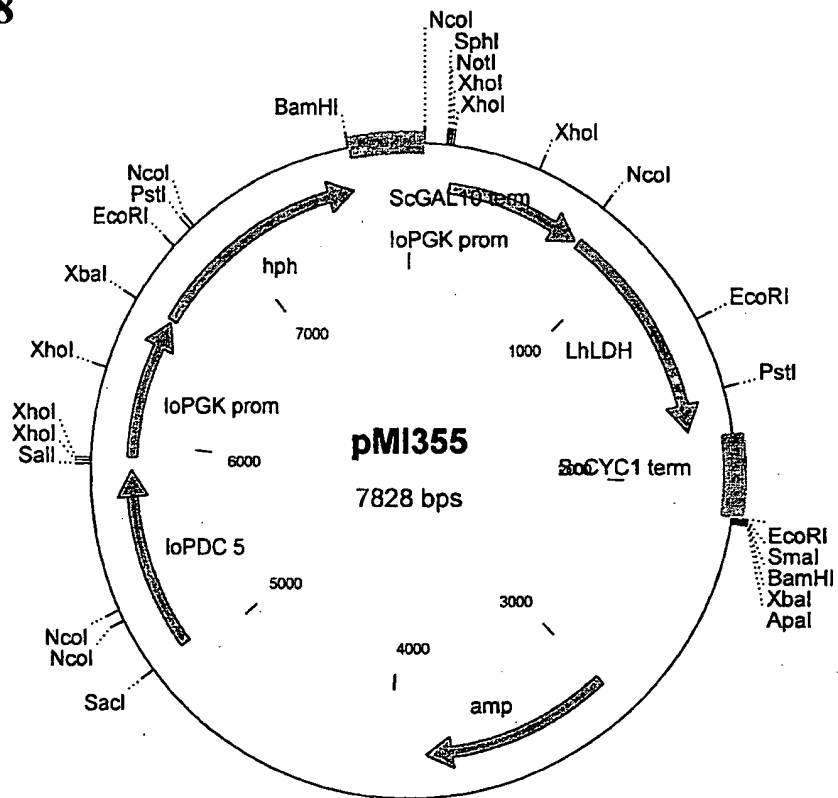


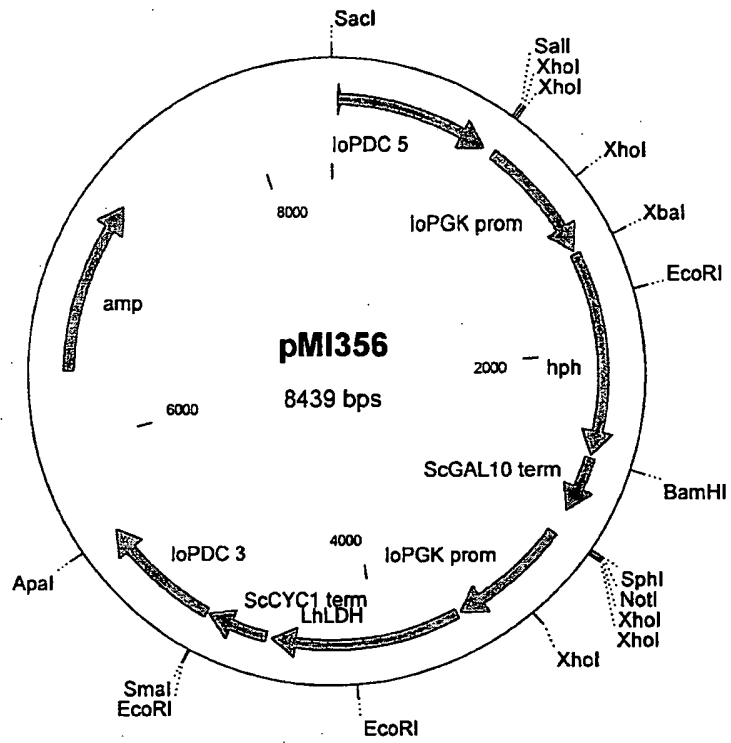
Fig. 9

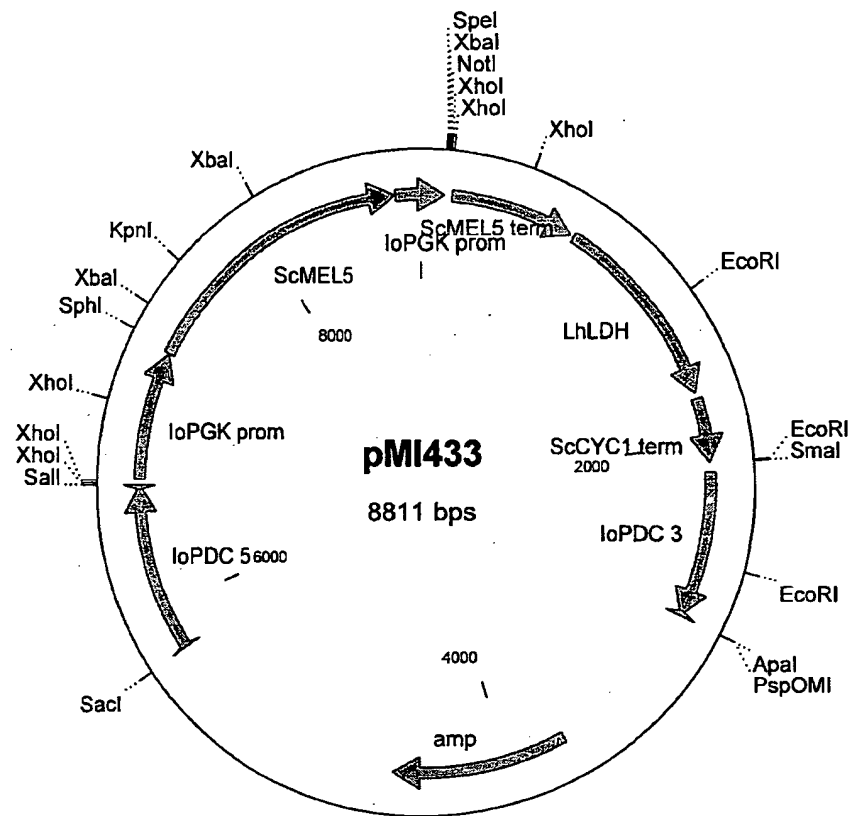
Fig. 10

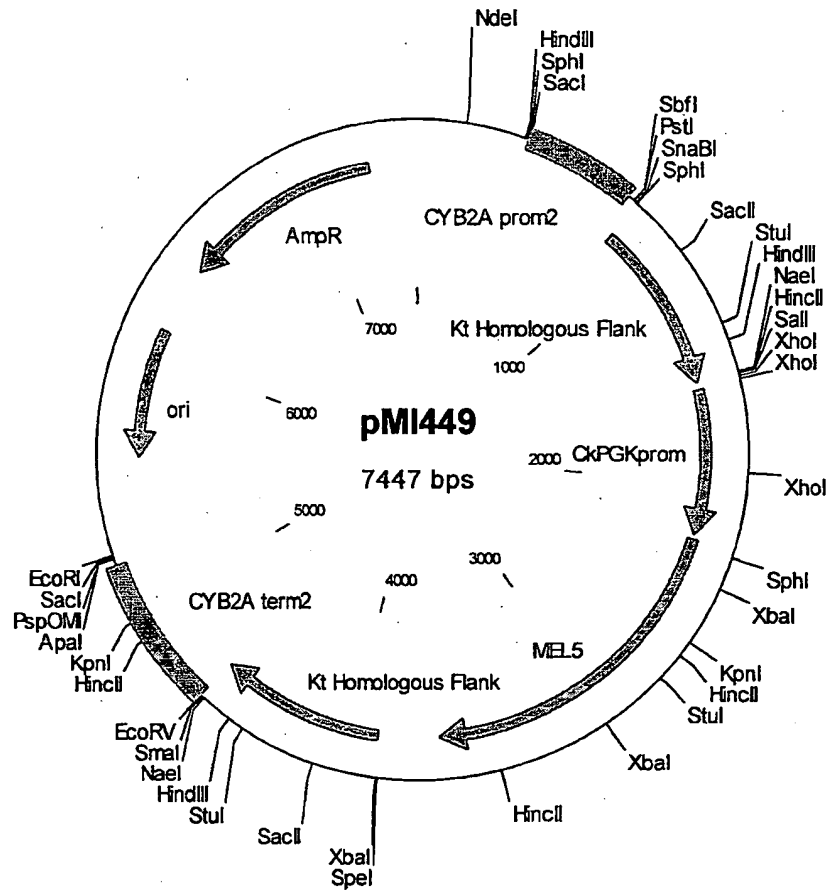
Fig. 11

Fig. 12

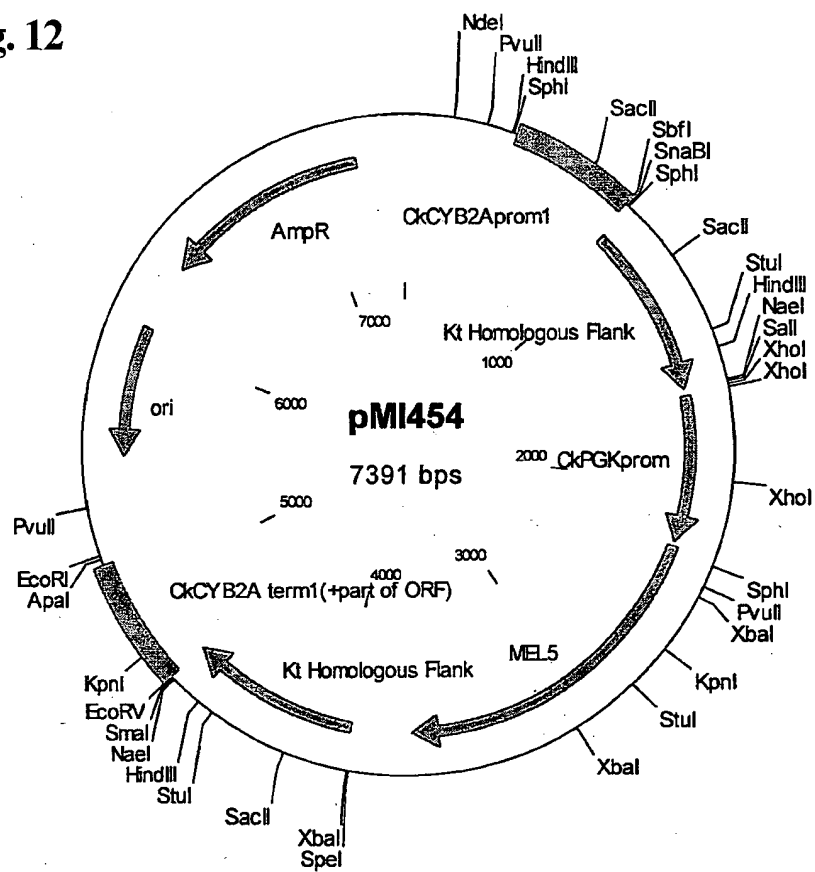


Fig. 13

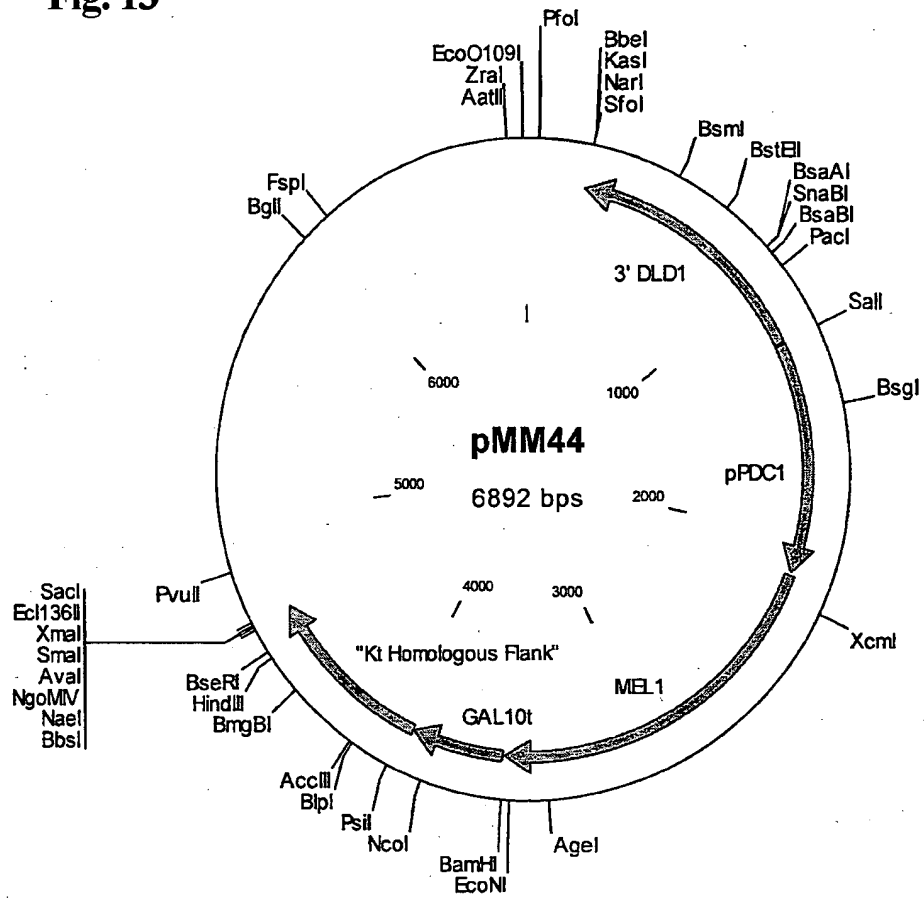


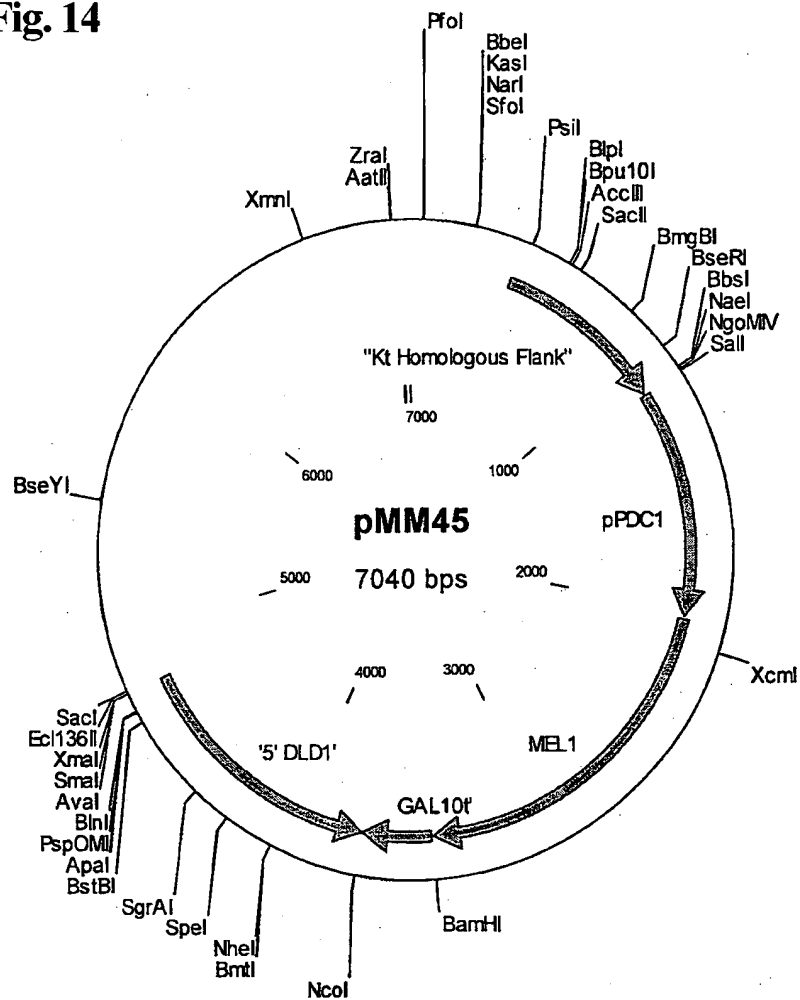
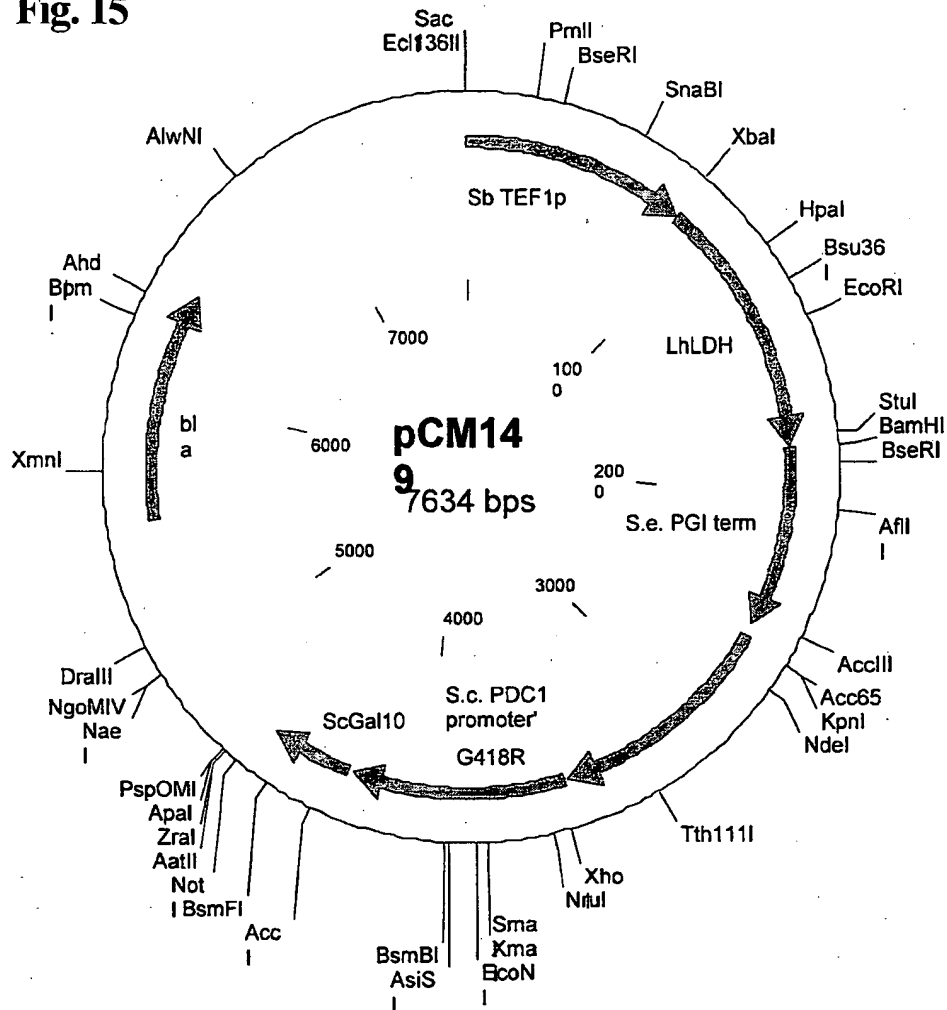
Fig. 14

Fig. 15



RESUMO

"CÉLULAS DE LEVEDURA QUE PRODUZEM ÁCIDO LÁTICO
TENDO L- OU D-LACTATO NÃO FUNCIONAL: GENE DE FERRICITOCROMO
C OXIDOREDUTASE"

5 São descritas células de levedura tendo um gene
lactato desidrogenase exógeno e modificadas reduzindo-se a
atividade L- ou D-lactato:ferricitocromo c oxidoredutase na
célula. Isto leva ao menor consumo de lactato pela célula e
pode aumentar os rendimentos gerais de lactato em um proces-
10 so de fermentação. Células tendo a menor atividade L- ou D-
lactato:ferricitocromo c oxidoredutase podem ser seleciona-
das pela resistência a ácidos orgânicos, tais como ácido lá-
tico ou glicólico.