



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

(21) **PI0615617-7 A2**

(22) Data de Depósito: 07/09/2006
(43) Data da Publicação: 24/05/2011
(RPI 2107)



(51) *Int.Cl.:*
C12N 9/02 2006.01
C12P 17/04 2006.01

(54) Título: **GENE SMS 27**

(30) Prioridade Unionista: 08/09/2005 EP 05019524.7

(73) Titular(es): DSM IP ASSETS B.V

(72) Inventor(es): Bastien Chevreux, Masako Shinjoh, Nigel Mouncey

(74) Procurador(es): ORLANDO DE SOUZA

(86) Pedido Internacional: PCT EP2006008727 de 07/09/2006

(87) Publicação Internacional: WO 2007/028609 de 15/03/2007

(57) **Resumo:** GENE SMS 27 A presente invenção relaciona-se aos genes recentemente identificados que codificam as proteínas que são envolvidas na síntese do ácido L-ascórbico (em seguida também referido como vitamina C). A invenção também apresenta os polinucleotídeos que compreendem as sequências de extensão completa do polinucleotídeo dos novos genes e os fragmentos destes, os novos polipeptídeos codificados pelos polinucleotídeos e fragmentos destes, assim como seus equivalentes funcionais. A invenção atual relaciona-se também ao uso de tais polinucleotídeos e de polipeptídeos como ferramentas biotecnológicas na produção da vitamina C dos microorganismos, por meio de uma modificação dos tais polinucleotídeos e/ou de polipeptídeos codificados tem um impacto direto ou indireto no rendimento, na produção, e/ou na eficiência da produção do produto de fermentação em tal microorganismo. São incluídos também métodos/processos de usar os polinucleotídeos e as sequências modificadas do polinucleotídeo para transformar microorganismos hospedeiros. A invenção relaciona-se também aos microorganismos geneticamente projetados e ao seu uso para a produção direta da vitamina C.

GENE SMS 27

A presente invenção refere-se a genes recentemente identificados que codificam proteínas que são envolvidas na síntese de ácido L-ascórbico (doravante também mencionado
5 como vitamina C) e/ou ácido 2-ceto-L-gulônico (doravante também mencionado como 2-KGA). A invenção também apresenta polinucleotídeos compreendendo as seqüências de polinucleotídeo de comprimento total dos genes novos e fragmentos dos mesmos, os polipeptídeos novos codificados
10 pelos polinucleotídeos e fragmentos dos mesmos, bem como seus equivalentes funcionais. A presente invenção também se refere ao uso dos polinucleotídeos e polipeptídeos como ferramentas biotecnologias na produção de vitamina C e/ou 2-KGA a partir de microorganismos, pelo que uma modificação
15 dos polinucleotídeos e/ou polipeptídeos codificadas tem um impacto direto ou indireto sobre o rendimento, produção e/ou eficiência de produção do produto de fermentação no microorganismo. São também incluídos métodos/processos de utilizar os polinucleotídeos e seqüências de
20 polinucleotídeo modificadas para transformar microorganismos hospedeiros. A invenção também se refere a microorganismos geneticamente construídos e seu uso para a produção direta de vitamina C e/ou 2-KGA.

A vitamina C é um dos fatores nutrientes muito
25 importantes e indispensáveis para os seres humanos. A vitamina C também é utilizada em ração de animais embora alguns animais de fazendas possam sintetizar a mesma em seu próprio corpo.

Durante os últimos 70 anos, a vitamina C tem sido
30 produzida industrialmente a partir de D-glicose pelo método

bem conhecido de Reichstein. Todas as etapas nesse processo são químicas exceto uma (a conversão de D-sorbitol em L-sorbose) que é realizada por conversão microbiana. Desse sua implementação inicial para a produção industrial de
5 vitamina C, várias modificações químicas e técnicas foram utilizadas para aperfeiçoar a eficiência do método Reichstein. Os recentes desenvolvimentos da produção de vitamina C são resumidos na Encyclopedia of Industrial Chemistry de Ullmann, 5ª edição, vol. A27 (1996), pág.
10 547ff.

Diferentes etapas intermediárias da produção de vitamina C foram executadas com o auxílio de microorganismos ou enzimas isoladas a partir da mesma. Desse modo, ácido 2-ceto-L gulônico (2-KGA), um composto
15 intermediário que pode ser quimicamente convertido em vitamina C por intermédio de uma reação de rearranjo alcalino, pode ser produzido por um processo de fermentação que se inicia a partir de L-sorbose ou D-sorbitol, por intermédio de cepas que pertencem, por exemplo, aos gêneros
20 *Ketogulonicigenium* ou *Gluconobacter*, ou por um processo de fermentação alternativa se iniciando a partir de D-glicose, por intermédio de cepas recombinantes que pertencem aos gêneros *Gluconobacter* ou *Pantoea*.

Os métodos atuais de produção química para Vitamina C
25 têm algumas características indesejáveis como elevado consumo de energia e uso de grandes quantidades de solventes orgânicos e inorgânicos. Portanto, durante as últimas décadas, outras abordagens para fabricar Vitamina C utilizando conversões microbiais, que seriam mais
30 econômicas como também ecológicas, têm sido pesquisadas.

A produção direta de vitamina C a partir de diversos substratos incluindo D-sorbitol, L-sorbose e L-sorbose foi reportada em vários microorganismos, como algas, levedura e bactérias de ácido acético, utilizando métodos de cultivo diferentes. Os exemplos de bactérias conhecidas capazes de produzir diretamente a Vitamina C incluem, por exemplo, cepas dos gêneros de *Gluconobacter*, *Gluconacetobacter*, *Acetobacter*, *Ketogulonicigenium*, *Pantoea*, *Pseudomonas* ou *Escherichia*. Os exemplos de levedura ou algas conhecidas incluem, por exemplo, *Candida*, *Saccharomyces*, *Zygosaccharomyces*, *Schizosaccharomyces*, *Kluyveromyces* ou *Chlorella*.

Microorganismos capazes de assimilar D-sorbitol para crescimento possuem, normalmente, enzimas capazes de oxidar esse composto em um substrato de assimilação universal como D-frutose. Além disso, microorganismos capazes de crescer em L-sorbose possuem uma enzima, reductase L-sorbose dependente de NAD(P)H, que é capaz de reduzir esse composto em D-sorbitol, o qual é então adicionalmente oxidado em D-frutose. D-frutose é um excelente substrato para o crescimento de muitos microorganismos, após ter sido fosforilado por intermédio de D-frutose cinase.

Por exemplo, no caso de bactérias de ácido acético, as quais são microorganismos gram-negativos de aeróbio forçados pertencendo ao gênero *Acetobacter*, *Gluconobacter* e *Gluconacetobacter*, esses microorganismos são capazes de transportar D-sorbitol para citosol e converter o mesmo em D-frutose por intermédio de uma desidrogenase D-sorbitol dependente de NAD citosólico. Algumas cepas individuais, como *Gluconobacter oxydans* IFO 3292, e IFO 3293, são

capazes também de transportar L-sorbose para o citosol e reduzir o mesmo em D-sorbitol por intermédio de uma reductase L-sorbose dependente de NAD(P)H, que é então adicionalmente oxidada em D-frutose. Nessas bactérias, a via Embden-Meyerhof-Parnas, bem como o ciclo de ácido tricarboxílico não é totalmente ativo, e a via principal que canaliza açúcares para dentro do metabolismo central é a via de fosfato pentose. D-frutose-6-fosfato, obtido a partir de D-frutose por uma reação de fosforilação entra na via de fosfato de pentose, sendo adicionalmente metabolizado e produzindo energia de redução na forma de compostos tricarboxílicos e NAD(P)H necessários para crescimento e manutenção.

Bactérias de ácido acético são bem conhecidas por sua capacidade em oxidar incompletamente diferentes substratos como álcoois, açúcares, álcoois de açúcar e aldeídos. Esses processos são genericamente conhecidos como fermentações oxidativas ou oxidações incompletas, e foram bem estabelecidos por um longo tempo na indústria de alimentos e química, especialmente na produção de vinagre e na produção de L-sorbose. Um produto útil conhecido como sendo obtido a partir de oxidações incompletas de D-sorbitol ou L-sorbose utilizando cepas que pertencem ao gênero *Gluconobacter* é 2-KGA.

As bactérias de ácido acético realizam essas reações de oxidação incompleta por intermédio de diferentes desidrogenases localizadas no espaço periplásmico, na membrana periplásmica bem como no citoplasma. Co-fatores diferentes são empregados pelas diferentes desidrogenases, o mais comum sendo PQQ e FAD para enzimas periplásmicas ou

ligadas à membrana, e NAD/NADP para enzimas citoplásmicas.

Embora todos os produtos dessas reações de oxidação difundam de volta no ambiente aquoso externo através da membrana externa, alguns deles podem ser passiva ou
5 ativamente transportados para dentro da célula e ser adicionalmente utilizados em vias metabólicas responsáveis por crescimento e formação de energia. Dentro da célula, produtos oxidados podem, muitas vezes, ser reduzidos de volta ao seu substrato original por intermédio de
10 reductases, e então ser canalizados de volta para o metabolismo central.

Proteínas, em particular enzimas e transportadores, que são ativos na metabolização de D-sorbitol ou L-sorbose são aqui referidas como estando envolvidas no Sistema de
15 Metabolização de Sorbitol/Sorbose. Tais proteínas são abreviadas aqui como proteínas SMS e funcionam na metabolização direta de D-sorbitol ou L-sorbose.

A metabolização de D-sorbitol ou L-sorbose inclui em um lado a assimilação desses compostos no citosol e
20 conversão adicional em metabólitos úteis para vias de assimilação como a via Embden-Meyerhof-Parnas, a via de fosfato pentose, a via Entner-Doudoroff, e o ciclo de ácido tricarboxíclico, todos envolvidos em reações anabólicas e de formação de energia totalmente vitais necessárias para
25 crescimento e manutenção de células vivas. Por outro lado, a metabolização de D-sorbitol ou L-sorbose também inclui a conversão desses compostos em produtos oxidados adicionais como L-sorbose, 2-KGA e Vitamina C pelos denominados processos de oxidação incompleta.

30 Um objetivo da presente invenção é melhorar os

rendimentos e/ou produtividade da produção de Vitamina C.

Surpreendentemente, verificou-se agora que proteínas SMS ou subunidades dessas proteínas tendo atividade em direção à ou que estão envolvidas na assimilação ou
5 conversão de D-sorbitol, L-sorbose ou L-sorbose desempenham um papel importante na produção biotecnológica da vitamina C e/ou 2-KGA.

Em uma modalidade, as proteínas SMS da presente invenção são selecionadas a partir de oxidorreductases [EC
10 1], preferivelmente oxidorreductases atuando sobre o grupo CH-OH de doadores [EC 1.1], mais preferivelmente oxidorreductases com NAD^+ ou NADP^+ como aceitador [EC 1.1.1] e oxidorreductases com outros aceitadores [EC 1.1.99], mais preferivelmente selecionados de oxidorreductases que
15 pertencem a classes de enzima [EC 1.1.1.1], [EC 1.1.1.15] ou [EC 1.2.1.-], ou preferivelmente oxidorreductases atuando sobre o grupo de doadores aldeído ou oxo [EX 1.2], mais preferivelmente oxidorreductases com NAD^+ ou NADP^+ como aceitador [EC 1.2.1].

20 Além disso, as proteínas SMS da presente invenção podem ser selecionadas do grupo que consiste em desidrogenase de D-sorbitol dependente de PQQ ligada por membrana, desidrogenase de L-sorbose ligada por membrana, desidrogenase de L-sorbose ligada por membrana,
25 desidrogenase de D-sorbitol dependente de FAD ligada por membrana, desidrogenase de D-sorbitol dependente de NAD citosólico, desidrogenase de D-sorbitol dependente de NAD(P) (também chamada de reductase de sorbose dependente de NADPH), desidrogenase de xilitol dependente de NAD,
30 desidrogenase de álcool dependente de NAD, desidrogenase de

L-sorbose ligada por membrana, reductase de L-sorbose dependente de NAD(P)H, desidrogenase de sorbosone dependente de NADP citosólico, reductase de L-sorbosone dependente de NAD(P)H citosólise, desidrogenase de aldeído
5 ligado por membrana, desidrogenase de aldeído citosólico, desidrogenase de glicerol-3-fosfato, desidrogenase de gliceraldeído-3-fosfato, e outros envolvidos em SMS incluindo proteínas envolvidas na regulação dos genes que codificam essas enzimas, como por exemplo, fatores de
10 transcrição, em particular repressores.

Em particular, verificou-se agora que as proteínas SMS codificadas por polinucleotídeos tendo uma seqüência de nucleotídeos que hibridiza, preferivelmente, sob condições altamente rigorosas em uma seqüência mostrada na SEQ ID NO:
15 1, desempenham um papel importante na produção biotecnológica da Vitamina C e/ou 2-KGA. Verificou-se agora, que pela alteração genética do nível de expressão de nucleotídeos, de acordo com a invenção em um microorganismo capaz de produzir diretamente a Vitamina C, como por
20 exemplo, *Gluconobacter*, a fermentação direta da Vitamina C e/ou 2-KGA pelo microorganismo pode ser aperfeiçoada ainda mais.

Conseqüentemente, a invenção refere-se a um polinucleotídeo selecionado do grupo que consiste em:

25 (a) polinucleotídeos que codificam um polipeptídeo compreendendo a seqüência de aminoácido, de acordo com a SEQ ID NO:2,

(b) polinucleotídeos compreendendo a seqüência de nucleotídeo de acordo com a SEQ ID NO: 1,

30 (c) polinucleotídeos compreendendo uma seqüência de

nucleotídeo que pode ser obtido por amplificação de ácido
nucléico como reação de cadeia de polimerase, utilizando
DNA genômico a partir de um microorganismo como padrão e um
conjunto iniciador (primer) de acordo com SEQ ID NO: 3 e
5 SEQ ID NO: 4;

(d) polinucleotídeos compreendendo uma seqüência de
nucleotídeos codificando um fragmento ou derivado de um
polipeptídeo codificado por um polinucleotídeo de qualquer
um de (a) a (c) onde no derivado um ou mais resíduos de
10 aminoácido são substituídos de forma conservativa em
comparação com o polipeptídeo, e o fragmento ou derivado
tem a atividade de um regulador de transcrição,
preferivelmente um repressor de desidrogenase L-sorbose
(SDH) e/ou desidrogenase L-sorbose (SNDH) (SMS 27);

(e) polinucleotídeos, cuja fita complementar
hibridiza sob condições rigorosas em um polinucleotídeo
como definido em qualquer um de (a) a (d) e que codifica um
regulador de transcrição, preferivelmente um repressor de
desidrogenase L-sorbose (SDH) e/ou desidrogenase L-
20 sorbose (SNDH) (SMS 27); e

(f) polinucleotídeos que são pelo menos 60%, como 70,
85, 90 ou 95% idênticos a um polinucleotídeo como definido
em qualquer um de (a) a (d) e que codificam um regulador de
transcrição, preferivelmente um repressor de desidrogenase
25 L-sorbose (SDH) e/ou desidrogenase L-sorbose (SNDH) (SMS
27);

ou

a fita complementar de um tal polinucleotídeo.

Verificou-se que a proteína de SMS como isolada a
30 partir de *Gluconobacter oxydans* IFO 3293, mostrada na SEQ

ID NO: 2 e descrita aqui é uma proteína de SMS particularmente útil, uma vez que pareceu que executa uma função crucial na produção direta de Vitamina C em microorganismos, em particular bactérias, como bactérias de ácido acético, como *Gluconobacter*, *Acetobacter* e *Gluconacetobacter*. Por conseguinte, a invenção se refere a um polinucleotídeo codificando um polipeptídeo de acordo com a SEQ ID NO: 2. Essa proteína pode ser codificada por uma seqüência de nucleotídeo como mostrado na SEQ ID NO: 1.

10 A invenção portanto se refere também a polinucleotídeos compreendendo a seqüência de nucleotídeos de acordo com SEQ ID NO: 1.

As seqüências de aminoácidos e nucleotídeos determinadas acima foram utilizadas como uma "seqüência de consulta" para executar uma busca com programa Blast2 (versão 2 ou BLAST a partir do National Center for Biotechnology [NCBI] contra o banco de dados PRO SW-SwissProt (liberação completa mais atualizações incrementais)). A partir das buscas, o polinucleotídeo SMS

15 27, de acordo com a SEQ ID NO: 1 foi anotado como codificando um regulador de transcrição que pertence à família de TetR/AcrR. A proteína como codificada pela SEQ ID NO: 2 atua como um repressor por ligar diretamente à região de promotor dos genes respectivos, incluindo os

25 genes que codificam para desidrogenase L-sorbose (SDH), como, por exemplo, mostrado na SEQ ID NO: 11 codificando uma proteína como da SEQ ID NO: 12, desidrogenase L-sorbose (SNDH), como por exemplo mostrado na SEQ ID NO: 13 codificando uma proteína como de SEQ ID NO: 14, e

30 exportador L-sorbose, como por exemplo mostrado na SEQ ID

NO: 15 codificando uma proteína como de SEQ ID NO: 16.

Um ácido nucléico, de acordo com a invenção pode ser obtido por amplificação de ácido nucléico utilizando cDNA, mRNA ou alternativamente, DNA genômico, como padrão e iniciadores de oligonucleotídeo apropriados como os iniciadores de nucleotídeo, de acordo com a SEQ ID NO: 3 e SEQ ID NO: 4, de acordo com técnicas de amplificação PCR padrão. O ácido nucléico desse modo amplificado pode ser clonado em um vetor apropriado e caracterizado por análise de seqüência de DNA.

O padrão para a reação pode ser cDNA obtido por transcrição inversa de mRNA preparado a partir de cepas conhecidas ou que se suspeita compreenderem um polinucleotídeo, de acordo com a invenção. O produto de PCR pode ser subclonado e seqüenciado para assegurar que as seqüências amplificadas representam as seqüências de uma seqüência de ácido nucléico nova como descrito aqui, ou um equivalente funcional da mesma.

O fragmento de PCR pode ser então utilizado para isolar um clone de cDNA de comprimento total por uma variedade de métodos conhecidos. Por exemplo, o fragmento amplificado pode ser rotulado e utilizado para classificar uma biblioteca de cDNA de cosmídio ou bacteriófago. Alternativamente, o fragmento rotulado pode ser utilizado para classificar uma biblioteca genômica.

Por conseguinte, a invenção se refere a polinucleotídeos compreendendo uma seqüência de nucleotídeo que pode ser obtido por amplificação de ácido nucléico como reação de cadeia polimerase, utilizando DNA como DNA genômico a partir de um microorganismo como um padrão e um

conjunto iniciador de acordo com SEQ ID NO: 3 e SEQ ID NO: 4.

A invenção também se refere a polinucleotídeos compreendendo uma seqüência de nucleotídeos codificando um
5 fragmento ou derivado de um polipeptídeo codificado por um polinucleotídeo como descrito aqui, onde no derivado um ou mais resíduos de aminoácido são substituídos de forma conservativa em comparação com o polipeptídeo, e o fragmento ou derivado tem a atividade de um polipeptídeo de
10 SMS, preferivelmente um polipeptídeo de SMS 27.

A invenção também se refere a polinucleotídeos cuja fita complementar hibridiza sob condições rigorosas com um polinucleotídeo como definido aqui, e que codificam um polipeptídeo de SMS, preferivelmente um polipeptídeo de SMS
15 27.

A invenção também se refere a polinucleotídeos que são pelo menos 60% idênticos a um polinucleotídeo, como definido aqui, e que codificam um polipeptídeo de SMS; e a invenção também se refere a polinucleotídeos sendo a fita
20 complementar de um polinucleotídeo como definido acima.

A invenção também se refere a iniciadores, sondas e fragmentos que podem ser utilizados para amplificar ou detectar um DNA de acordo com a invenção e para identificar espécies relacionadas ou famílias de microorganismos também
25 contendo tais genes.

A presente invenção também se refere a vetores que incluem polinucleotídeos da invenção e microorganismos que são geneticamente construídos com os polinucleotídeos ou vetores.

30 A invenção também se refere a processos para produzir

microorganismos capazes de expressar um polipeptídeo codificado pelo polinucleotídeo acima definido e um polipeptídeo codificado por um polinucleotídeo como definido acima.

5 A invenção também se refere a microorganismos, onde a atividade de um polipeptídeo de SMS, preferivelmente um polipeptídeo de SMS 27, é reduzida ou abolida de modo que o rendimento de Vitamina C e/ou 2-KGA que é diretamente produzida a partir de D-sorbitol ou L-sorbose, é aumentado.

10 A pessoa versada saberá como reduzir ou abolir a atividade de uma proteína de SMS, preferivelmente uma proteína de SMS 27. Isso pode ser, por exemplo, realizado por modificação genética do organismo hospedeiro de tal modo que produza menos cópias ou nenhuma cópia da proteína
15 de SMS, preferivelmente a proteína de SMS 27, do que o organismo do tipo selvagem ou pela diminuição ou abolição da atividade específica da proteína de SMS, preferivelmente a proteína de SMS 27.

Na descrição a seguir, os procedimentos são detalhados
20 para obter esse objetivo, isto é, o aumento no rendimento e/ou produção da vitamina C que é diretamente produzida a partir de D-sorbitol ou L-sorbose pela redução ou abolição da atividade de uma proteína de SMS 27. Esses procedimentos se aplicam mutatis mutandis para outras proteínas de SMS.

25 As modificações para que o organismo produza menos ou nenhuma cópia do gene e/ou proteína de SMS 27 podem incluir o uso de um promotor fraco, ou a mutação (por exemplo, inserção, deleção ou mutação de pontos) de (partes de) gene de SMS 27 ou seus elementos reguladores. A diminuição ou
30 abolição da atividade específica de uma proteína de SMS 27

também pode ser realizada por métodos conhecidos na técnica. Tais métodos podem incluir a mutação (por exemplo, inserção, deleção ou mutação de pontos) do (partes do) gene de SMS 27. Isso pode por exemplo, afetar a interação com DNA que é mediado pela região N-terminal de SMS 27 ou interação com outras moléculas efetoras.

São também conhecidos na técnica métodos de reduzir ou abolir a atividade de uma dada proteína pelo contato da proteína de SMS 27 com inibidores específicos ou outras substâncias que interagem especificamente com a proteína de SMS 27. Para identificar tais inibidores específicos, a proteína de SMS 27 pode ser expressa e testada em relação à atividade na presença de compostos que se suspeita inibirem a atividade da proteína de SMS 27. Compostos de inibição em potencial podem, por exemplo, ser anticorpos monoclonais ou policlonais contra a proteína de SMS 27. Tais anticorpos podem ser obtidos por protocolos de imunização de rotina de animais de laboratório apropriados.

A invenção pode ser executada em qualquer microorganismo contendo um gene de SMS 27 ou homólogo do mesmo. Microorganismos apropriados podem ser selecionados a partir do grupo que consiste em levedura, algas e bactérias, como cepas do tipo selvagem, cepas mutantes derivadas por mutagênese clássica e métodos de seleção ou como cepas recombinantes. Os exemplos de tais bactérias podem ser, por exemplo, *Gluconobacter*, *Gluconacetobacter*, *Acetobacter*, *Ketogulonicigenium*, *Pantoea*, *Pseudomonas*, *Rhizobium*, *Sinorhizobium*, como, *Sinorhizobium meliloti*, *Bradyrhizobium*, como *Bradyrhizobium japonicum*, *Roseobacter*, *Ralstonia*, *Pseudomonas*, como por exemplo *Pseudomonas*

putida, e *Escherichia*, como, por exemplo, *Escherichia coli*. São preferidos *Gluconobacter* ou *Acetobacter aceti*, como por exemplo, *G. oxydans*, *G. cerinus*, *G. frateurii*, *A. aceti* subsp. *Xylinum* ou *A. aceti* subsp. *Orleanus*, preferivelmente
5 *G. oxydans* IFO 3293, e seus derivados contendo genes envolvidos nas vias de produção de vitamina C e as regiões adjacentes onde o gene SMS 27 ou seu equivalente poderia estar localizado.

Microorganismos que podem ser utilizados para a
10 presente invenção podem ser publicamente disponíveis de fontes diferentes, por exemplo, Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ), Mascheroder Weg 1B, D-38124 Braunschweig, Alemanha, American Type Culture Collection (ATCC), P.O. Box 1549, Manassas, VA 20108 EUA ou
15 Culture Collection Division, NITE Biological Resource Center, 2-5-8, Kazusakamatari, Kisarazu-shi, Chiba, 292-0818, Japão (anteriormente: Institute for Fermentation, Osaka (IFO), 17-85, Juso-honmachi 2-chome, Yodogawa-Ku, Osaka 532-8686, Japão). Os exemplos de bactérias preferidas
20 depositadas junto ao IFO são, por exemplo, *Gluconobacter oxydans* (anteriormente conhecido como *G. melanogenus*) IFO 3293, *Gluconobacter oxydans* (anteriormente conhecido como *G. melanogenus*) IFO 3292, *Gluconobacter oxydans* (anteriormente conhecido como *G. rubiginosus*) IFO 3244,
25 *Gluconobacter frateurii* (anteriormente conhecido como *G. industrius*) IFO 3260, *Gluconobacter cerinus* IFO 3266, *Gluconobacter oxydans* IFO 3287, e *Acetobacter aceti* subsp. *Orleanus* IFO 3259, que foram todos depositados em 05 de abril de 1954; *Acetobacter aceti* subsp. *Xylinum* IFO 13693
30 depositado em 22 de outubro de 1975, e *Acetobacter aceti*

subsp. Xylinum IFO 13773 depositado em 08 de dezembro de 1977. A cepa *Acetobacter* sp. ATCC 15164, que também é um exemplo de uma bactéria preferida, foi depositada junto ao ATCC.

5 Um microorganismo como da presente invenção, pode conter modificações adicionais no nível de DNA ou de proteína (vide acima), desde que essa modificação tenha um impacto direto sobre o rendimento, produção e/ou eficiência da produção direta de Vitamina C e/ou 2-KGA a partir dos
10 substratos, como, por exemplo, D-sorbitol ou L-sorbose. Tais modificações adicionais podem, por exemplo, afetar outros genes que codificam proteínas de SMS como descrito acima, em particular genes que codificam desidrogenase de L-sorbose ligadas por membrana, como desidrogenase de L-
15 sorbose SNDHai, desidrogenase de D-sorbitol ligadas por PQQ ligadas por membrana e/ou outros genes que codificam proteínas envolvidas no transporte de açúcar e/ou álcoois de açúcar, como por exemplo exportadores, em particular exportadores de sorbose, preferivelmente um gene como
20 mostrado na SEQ ID NO: 15. Os métodos de executar tais modificações são conhecidos na técnica, com alguns exemplos adicionalmente descritos aqui. Para o uso de SNDHai para produção direta de vitamina C bem como a sequência de aminoácidos e nucleotídeos da mesma consulte WO 2005/017159
25 que é aqui incorporado a título de referência.

De acordo com um objetivo adicional da presente invenção é fornecido o uso de um polinucleotídeo como definido acima ou um microorganismo que é geneticamente construído utilizando tais polinucleotídeos na produção de
30 vitamina C e/ou 2-KGA.

A invenção também se refere a processos para a expressão de genes endógenos em um microorganismo, a processos para a produção de polipeptídeos como definido acima em um microorganismo e a processos para a produção de microorganismos capazes de produzir Vitamina C e/ou 2-KGA. Todos esses processos podem compreender a etapa de alterar um microorganismo, onde "alterar" como utilizado aqui abrange o processo para "alterar geneticamente" ou "alterar a composição do meio de cultura de célula e/ou métodos utilizados para cultivar" de tal modo que o rendimento e/ou produtividade do produto de fermentação possa ser aperfeiçoado em comparação com o organismo do tipo selvagem. Como utilizado aqui, "rendimento aperfeiçoado de Vitamina C" significa um aumento de pelo menos 5%, 10%, 25%, 30%, 40%, 50%, 75%, 100%, 200% ou ainda mais do que 500%, em comparação com um microorganismo do tipo selvagem, isto é, um microorganismo que não seja geneticamente alterado.

O termo "geneticamente construído" ou "geneticamente alterado" significa a alteração científica da estrutura do material genético em um organismo vivo. Envolve a produção e uso de DNA recombinante. Mais particularmente é utilizado para delinear o organismo geneticamente construído ou modificado a partir do organismo de ocorrência natural. A engenharia genética pode ser feita por diversas técnicas conhecidas na arte, como por exemplo, substituição de gene, amplificação de gene, ruptura de gene, transfecção, transformação utilizando plasmídeos, vírus, ou outros vetores. Um organismo geneticamente modificado, por exemplo, microorganismo geneticamente modificado, também é

freqüentemente mencionado como um organismo recombinante, por exemplo, microorganismo recombinante.

De acordo ainda com outro aspecto da invenção é fornecido um processo para a produção de Vitamina C e/ou 2-KGA por fermentação direta.

Vários substratos podem ser utilizados como uma fonte de carbono em um processo da presente invenção, isto é, um processo para conversão direta de um dado substrato em vitamina C como, por exemplo, mencionado acima. Fontes de carbono particularmente adequadas são aquelas que são facilmente obteníveis a partir de via de metabolização de D-glicose ou D-sorbitol como, por exemplo, D-glicose, D-sorbitol, L-sorbose, L-sorbose, 2-ceto-L-gulonato, D-gluconato, 2-ceto-D-gluconato ou 2,5-diceto-gluconato. Preferivelmente, o substrato é selecionado, por exemplo, de D-glicose, D-sorbitol, L-sorbose ou L-sorbose, mais preferivelmente de D-glicose, D-sorbitol ou L-sorbose, e mais preferivelmente de D-sorbitol, L-sorbose ou L-sorbose. O termo "substrato" e "substrato de produção" com relação ao processo acima utilizando um microorganismo é utilizado de forma intercambiável aqui.

Um meio como utilizado aqui para o processo acima utilizando um microorganismo pode ser qualquer meio apropriado para a produção de Vitamina C e/ou 2-KGA. Tipicamente, o meio é um meio aquoso compreendendo por exemplo, sais, substrato(s) e um certo pH. O meio no qual o substrato é convertido em Vitamina C e/ou 2-KGA também é mencionado como o meio de produção.

"Fermentação" ou "produção" ou "processo de fermentação" como utilizado aqui, pode ser o uso de células

de crescimento utilizando meios, condições e procedimentos conhecidos por pessoas versadas, ou o uso de células não em crescimento denominadas em repouso, após terem sido cultivadas pelo uso de meios, condições e procedimentos conhecidos por pessoas versadas, sob condições apropriadas para a conversão de substratos apropriados em produtos desejados como Vitamina C e/ou 2-KGA. Preferivelmente, células em repouso são utilizadas para a produção de Vitamina C. Um exemplo de tal processo para a produção de vitamina C é descrito em WO 2005/017159 (como incorporado aqui a título de referência). Preferivelmente, 2-KGA é produzido utilizando células em crescimento (vide, por exemplo, EP 0 518 136 B1).

O termo "fermentação direta", "produção direta", "conversão direta" e similares pretende significar que um microorganismo é capaz da conversão de um certo substrato no produto especificado por intermédio de uma ou mais etapas de conversão biológica, sem a necessidade de qualquer etapa adicional de conversão química. Por exemplo, o termo "conversão direta de D-sorbitol em Vitamina C" pretende descrever um processo onde um microorganismo está produzindo Vitamina C e onde D-sorbitol é oferecido como uma fonte de carbono sem a necessidade de uma etapa intermediária de conversão química. Um único microorganismo capaz de fermentar diretamente a Vitamina C é preferido. O microorganismo é cultivado sob condições que permitem tal conversão a partir do substrato como definido acima.

Com relação ao processo acima utilizando um microorganismo é entendido que os microorganismos acima mencionados também incluem sinônimos ou basônimos de tais

espécies tendo as mesmas propriedades fisiológicas, como definido pelo Código Internacional de Nomenclatura de procariotes. A nomenclatura dos microorganismos, como utilizado aqui, é a oficialmente aceita (na data de depósito do pedido de prioridade) pelo International Committee on Systematics of Prokaryotes e Bacteriology and Applied Microbiology Division da International Union of Microbiological Societies, e publicado por seu veículo de publicação oficial International Journal of Systematic and Evolutionary Microbiology (IJSEM). Uma referência específica é feita a Urbance e outros, IJSEM (2001) vol. 51: 1059-1070, com uma notificação corretiva em IJSEM (2001) vol. 51:1231-1233, descrevendo a reclassificação taxonômica de *G. oxydans* DSM 4025 como *Ketogulonicigenium vulgare*.

Como utilizado aqui, células em repouso se referem a células de um microorganismo que são, por exemplo, viáveis porém não em crescimento ativo, ou que estão crescendo em taxas de crescimento baixas específicas, por exemplo, taxas de crescimento que são mais baixas do que $0,02 \text{ h}^{-1}$, preferivelmente mais baixas do que $0,01 \text{ h}^{-1}$. Diz-se que as células que mostram as taxas de crescimento acima estão em um "modo de célula em repouso".

O processo da presente invenção como acima utilizando um microorganismo pode ser executado em diferentes etapas ou fases: preferivelmente, o microorganismo é cultivado em uma primeira etapa (também mencionada como etapa (a) ou fase de crescimento) sob condições que permitem crescimento. Essa fase é terminada pela mudança das condições de tal modo que a taxa de crescimento do

microorganismo é reduzida levando a células em repouso, também mencionado como etapa (b), seguido pela produção da vitamina C a partir do substrato utilizando (b), também mencionado como a fase de produção.

5 A fase de crescimento e produção como executada no processo acima, utilizando um microorganismo, pode ser realizada no mesmo recipiente, isto é, somente um recipiente, ou em dois ou mais recipientes diferentes, com uma etapa de separação de células opcional entre as duas
10 fases. A Vitamina C produzida pode ser recuperada das células por qualquer meio apropriado. Os meios de recuperação por exemplo em que a Vitamina C produzida possa ser separada do meio de produção. Opcionalmente, a Vitamina C desse modo produzida pode ser adicionalmente processada.

15 Para fins da presente invenção referente ao processo acima utilizando um microorganismo, os termos "fase de crescimento", "etapa em crescimento", "etapa de crescimento" e "período de crescimento" são utilizados de forma intercambiável na presente invenção. O mesmo se
20 aplica aos termos "fase de produção", "etapa de produção", "período de produção".

 Um modo de executar o processo acima utilizando um microorganismo da presente invenção, pode ser um processo onde o microorganismo é cultivado em um primeiro
25 recipiente, o denominado recipiente de crescimento, como uma fonte para as células em repouso, e pelo menos parte das células são transferidas para um segundo recipiente, o denominado recipiente de produção. As condições no recipiente de produção podem ser tais que as células
30 transferidas a partir do recipiente de crescimento se

tornam células em repouso como definido acima. A Vitamina C é produzida no segundo recipiente e recuperada a partir do mesmo.

Com relação ao processo acima utilizando um
5 microorganismo, em um aspecto, a etapa de crescimento pode ser executada em um meio aquoso, isto é, o meio de crescimento, suplementado com nutrientes apropriados para crescimento sob condições aeróbicas. O cultivo pode ser realizado, por exemplo, em modo de batelada, alimentado em
10 batelada, semicontínuo ou contínuo. O período de cultivo pode variar dependendo, por exemplo, do hospedeiro, pH, temperatura e meio nutriente a ser utilizado, e pode ser, por exemplo, aproximadamente 10 h até aproximadamente 10 dias, preferivelmente aproximadamente 1 até aproximadamente
15 10 dias, mais preferivelmente aproximadamente 1 até aproximadamente 5 dias quando realizado no modo de batelada ou alimentado em batelada, dependendo do microorganismo. Se as células são cultivadas no modo contínuo, o tempo de permanência pode ser, por exemplo, de aproximadamente 2 a
20 aproximadamente 100 h, preferivelmente de aproximadamente 2 h a aproximadamente 50 h, dependendo do microorganismo. Se o microorganismo for selecionado de bactérias, o cultivo pode ser conduzido por exemplo, em um pH de aproximadamente 3,0 a aproximadamente 9,0, preferivelmente aproximadamente
25 4,0 a aproximadamente 9,0, mais preferivelmente aproximadamente 4,0 a aproximadamente 8,0, ainda mais preferivelmente aproximadamente 5,0 a aproximadamente 8,0. Se algas ou levedura forem utilizadas, o cultivo pode ser realizado, por exemplo, em um pH abaixo de aproximadamente
30 7,0, preferivelmente abaixo de aproximadamente 6,0, mais

preferivelmente abaixo de aproximadamente 5,5, e mais preferivelmente abaixo de aproximadamente 5,0. Uma faixa de temperatura apropriada para realizar o cultivo utilizando bactérias pode ser, por exemplo, de aproximadamente 13°C a aproximadamente 40°C, preferivelmente de aproximadamente 18°C a aproximadamente 37°C, mais preferivelmente de aproximadamente 13°C a aproximadamente 36°C, e mais preferivelmente ainda de aproximadamente 18°C a aproximadamente 33°C. Se algas ou levedura forem utilizados, uma faixa de temperatura apropriada para realizar o cultivo pode ser, por exemplo, de aproximadamente 15°C a aproximadamente 40°C, preferivelmente de aproximadamente 20°C a aproximadamente 45°C, mais preferivelmente de aproximadamente 25°C a aproximadamente 40°C, ainda mais preferivelmente de aproximadamente 25°C a aproximadamente 38°C, e mais preferivelmente ainda de aproximadamente 30°C a aproximadamente 38°C. O meio de cultura para crescimento normalmente pode conter tais nutrientes como fontes de carbono assimiláveis, por exemplo, glicerol, D-manitol, D-sorbitol, L-sorbose, eritritol, ribitol, xilitol, arabitól, inositol, dulcitol, D-ribose, D-frutose, D-glicose, sacarose, e etanol, preferivelmente L-sorbose, D-glicose, D-sorbitol, D-manitol, glicerol e etanol; e fontes de nitrogênio digeríveis como substâncias orgânicas, por exemplo, peptona, extrato de levedura e aminoácidos. Os meios podem ser com ou sem uréia e/ou corpo fluido de infusão de milho e/ou fermento de pão. Várias substâncias inorgânicas podem ser também utilizadas como fontes de nitrogênio, por exemplo, nitratos e sais de amônio. Além

disso, o meio de crescimento, pode conter normalmente sais inorgânicos, por exemplo, sulfato de magnésio, sulfato de manganês, fosfato de potássio e carbonato de cálcio. As células obtidas utilizando os procedimentos descritos acima 5 podem ser então adicionalmente incubadas essencialmente nos mesmos modos, temperatura e condições de pH como descrito acima, na presença de substratos como D-sorbitol, L-sorbose, ou D-glicose, de tal modo que convertem esses substratos diretamente em vitamina C e/ou 2-KGA. A 10 incubação pode ser feita em um meio rico em nitrogênio, contendo, por exemplo, fontes de nitrogênio orgânicas, por exemplo, peptona, extrato de levedura, fermento de pão, uréia, aminoácidos, e corpo fluido de infusão de milho, ou fontes de nitrogênio inorgânicas, por exemplo, nitratos e 15 sais de amônio, em cujo caso as células serão capazes de crescer ainda mais enquanto produzem vitamina C e/ou 2-KGA. Alternativamente, a incubação pode ser feita em um meio deficiente em nitrogênio, em cujo caso as células não crescerão substancialmente, e estarão em um modo de célula 20 em repouso, ou modo de biotransformação. Em todos os casos, o meio de incubação pode conter também sais inorgânicos, por exemplo, sulfato de magnésio, sulfato de manganês, fosfato de potássio, e cloreto de cálcio.

Com relação ao processo acima utilizando um 25 microorganismo, na fase de crescimento as taxas específicas de crescimento são, por exemplo, pelo menos $0,02 \text{ h}^{-1}$. Para células crescendo no modo de batelada, alimentado por batelada ou semicontínuo, a taxa de crescimento depende, por exemplo, da composição do meio de crescimento, pH, 30 temperatura e similares. Em geral, as taxas de crescimento

podem estar, por exemplo, em uma faixa de aproximadamente 0,05 a aproximadamente $0,2 \text{ h}^{-1}$, preferivelmente de aproximadamente $0,06$ a aproximadamente $0,15 \text{ h}^{-1}$, e mais preferivelmente de aproximadamente $0,07$ a aproximadamente $0,13 \text{ h}^{-1}$.

Em outro aspecto do processo acima utilizando um microorganismo, células em repouso podem ser fornecidas por cultivo do microorganismo respectivo em placas agar desse modo servindo como recipiente de crescimento, utilizando essencialmente as mesmas condições, por exemplo, período de cultivo, pH, temperatura, meio nutriente como descrito acima, com a adição de agar agar.

Com relação ao processo acima utilizando um microorganismo, se a fase de crescimento e produção forem executadas em dois recipientes separados, então as células a partir da fase de crescimento podem ser colhidas ou concentradas e transferidas para um segundo recipiente, o denominado recipiente de produção. Esse recipiente pode conter um meio aquoso suplementado com qualquer substrato de produção aplicável que pode ser convertido em Vitamina C pelas células. As células a partir do recipiente de crescimento podem ser colhidas ou concentradas por qualquer operação apropriada, como por exemplo centrifugação, microfiltração ou ultrafiltração de fluxo cruzado de membrana, filtração, decantação, floculação. As células desse modo obtidas podem ser também transferidas para o recipiente de produção na forma do caldo original a partir do recipiente de crescimento, sem serem colhidas, concentradas ou lavadas, isto é na forma de uma suspensão de células. Em uma modalidade preferida, as células são

transferidas a partir do recipiente de crescimento para o recipiente de produção na forma de uma suspensão de células sem qualquer etapa de lavagem ou isolamento intermediária.

Desse modo, em uma modalidade preferida do processo acima utilizando um microorganismo, a etapa (a) e a (c) do processo da presente invenção, como descrito acima, não são separadas por nenhuma etapa de lavagem e/ou separação.

Com relação ao processo acima utilizando um microorganismo, se a fase de crescimento e produção forem executadas no mesmo recipiente, as células podem ser crescidas sob condições apropriadas até a densidade desejada de células seguido por uma substituição do meio de crescimento com o meio de produção contendo o substrato de produção. Tal substituição pode ser, por exemplo, a alimentação do meio de produção para o recipiente ao mesmo tempo e taxa que a retirada ou colheita de sobrenadante a partir do recipiente. Para manter as células em repouso no recipiente, as operações para reciclagem ou retenção de células podem ser utilizadas, como por exemplo, etapas de reciclagem de células. Tais etapas de reciclagem, por exemplo, incluem porém não são limitadas a métodos utilizando centrífugas, filtros, etapas de micro-filtração ou ultrafiltração de fluxo cruzado de membrana, reatores de membrana, floculação ou imobilização de células em matrizes porosas, não porosas ou poliméricas apropriadas. Após uma fase de transição, o recipiente é levado a processar condições sob as quais as células estão em um modo de célula em repouso como definido acima, e o substrato de produção é eficientemente convertido em Vitamina C.

O meio aquoso no recipiente de produção, como

utilizado para a etapa de produção com relação ao processo acima utilizando um microorganismo, doravante denominado meio de produção, pode conter somente o(s) substrato(s) de produção a ser convertido em vitamina C, ou pode conter, por exemplo, sais inorgânicos adicionais, por exemplo, cloreto de sódio, cloreto de cálcio, sulfato de magnésio, sulfato de manganês, fosfato de potássio, fosfato de cálcio, e carbonato de cálcio. O meio de produção também pode conter fontes de nitrogênio digeríveis como, por exemplo, substâncias orgânicas, por exemplo, peptona, extrato de levedura, uréia, aminoácidos, e corpo fluido de infusão de milho, e substâncias inorgânicas, por exemplo, amônia, sulfato de amônio, e nitrato de sódio, em tais concentrações que as células são mantidas em um modo de célula em repouso como definido acima. O meio pode ser com ou sem uréia e/ou corpo fluido de infusão de milho e/ou fermento de pão. A etapa de produção pode ser realizada, por exemplo, em modo de batelada, alimentado por batelada, semicontínuo ou contínuo. No caso de modo alimentado por batelada, semicontínuo ou contínuo, as duas células a partir do recipiente de crescimento e meio de produção podem ser alimentadas contínua ou intermitentemente para o recipiente de produção em taxas de alimentação apropriadas. Alternativamente, somente o meio de produção pode ser alimentado contínua ou intermitentemente para o recipiente de produção, enquanto as células vindo do recipiente de crescimento são transferidas imediatamente para o recipiente de produção. As células vindo do recipiente de crescimento podem ser utilizadas como uma suspensão de células dentro do recipiente de produção ou podem ser

utilizadas, por exemplo, como células floculadas ou imobilizadas em qualquer fase sólida como matrizes porosas ou poliméricas. O período de produção, definido como o período decorrido entre a entrada do substrato no recipiente de produção e a colheita do sobrenadante contendo Vitamina C, o denominado fluxo de colheita, pode variar dependendo, por exemplo, do tipo e concentração de células, pH, temperatura e meio nutriente a ser utilizado, e é preferivelmente aproximadamente 2 a aproximadamente 100 h. O pH e temperatura podem ser diferentes do pH e temperatura da etapa de crescimento, porém são essencialmente iguais aos para a etapa de crescimento.

Em uma modalidade preferida do processo acima utilizando um microorganismo, a etapa de produção é realizada em modo contínuo, significando que um primeiro fluxo de alimentação contendo as células a partir do recipiente de crescimento e um segundo fluxo de alimentação contendo o substrato é alimentado contínua ou intermitentemente para o recipiente de produção. O primeiro fluxo pode conter somente as células isoladas/separadas a partir do meio de crescimento ou uma suspensão de células, vindo diretamente da etapa de crescimento, isto é, células suspensas no meio de crescimento, sem qualquer etapa intermediária de separação de células, lavagem e/ou isolamento. O segundo fluxo de alimentação, como definido aqui, pode incluir todos os outros fluxos de alimentação necessários para a operação da etapa de produção, por exemplo, o meio de produção compreendendo o substrato na forma de um ou vários fluxos diferentes, água para diluição, e base para controle de pH.

Com relação ao processo acima utilizando um microorganismo, quando os dois fluxos são alimentados continuamente, a razão da taxa de alimentação do primeiro fluxo para a taxa de alimentação do segundo fluxo pode variar entre aproximadamente 0,01 e aproximadamente 10, preferivelmente entre aproximadamente 0,01 e aproximadamente 5, mais preferivelmente entre aproximadamente 0,02 e aproximadamente 2. Essa razão depende da concentração de células e substrato no primeiro e segundo fluxo, respectivamente.

Outro modo de executar o processo como acima utilizando um microorganismo da presente invenção, pode ser um processo utilizando uma certa densidade de células de células em repouso no recipiente de produção. A densidade de células é medida como unidades de absorvência (densidade óptica) em 600 nm por métodos conhecidos pelas pessoas versadas. Em uma modalidade preferida, a densidade de células na etapa de produção é pelo menos aproximadamente 10, mais preferivelmente entre aproximadamente 10 e aproximadamente 200, ainda mais preferivelmente entre aproximadamente 15 e aproximadamente 200, ainda mais preferivelmente entre aproximadamente 15 e aproximadamente 120, e mais preferivelmente ainda entre aproximadamente 20 e aproximadamente 120.

Com relação ao processo acima utilizando um microorganismo, para manter as células no recipiente de produção na densidade desejada de células durante a fase de produção como executado, por exemplo, no modo contínuo ou semicontínuo, qualquer meio conhecido na técnica pode ser utilizado, como por exemplo, reciclagem de células por

centrifugação, filtração, ultrafiltração ou microfiltração de fluxo cruzado de membrana, decantação, floculação, retenção de células no recipiente por dispositivos de membrana ou imobilização de células. Além disso, caso a
5 etapa de produção seja executada no modo contínuo ou semicontínuo e as células seja contínua ou intermitentemente alimentadas a partir do recipiente de crescimento, a densidade de células no recipiente de produção pode ser mantida em um nível constante por
10 exemplo, por colheita de uma quantidade de células a partir do recipiente de produção correspondendo à quantidade de células sendo alimentadas a partir do recipiente de crescimento.

Com relação ao processo acima utilizando um
15 microorganismo, a vitamina C produzida contida no denominado fluxo de colheita é recuperada/colhida a partir do recipiente de produção. O fluxo de colheita pode incluir, por exemplo, solução aquosa contendo células ou isenta de células vindo do recipiente de produção, que
20 contém Vitamina C como resultado da conversão do substrato de produção pelas células em repouso no recipiente de produção. As células ainda presentes no fluxo de colheita podem ser separadas da Vitamina C por quaisquer operações conhecidas na técnica, como por exemplo, filtração,
25 centrifugação, decantação, ultrafiltração ou micro-filtração de fluxo cruzado de membrana, ultrafiltração ou micro-filtração de fluxo tangencial ou filtração final. Após essa operação de separação de células, o fluxo de colheita está essencialmente isento de células.

30 Em um aspecto adicional, o processo da presente

invenção pode ser combinado com etapas adicionais de separação e/ou purificação da Vitamina C produzida a partir de outros componentes contidos no fluxo de colheita, isto é, as etapas de processamento denominadas a jusante. Essas etapas podem incluir qualquer meio conhecido por uma pessoa versada, como, por exemplo, concentração, cristalização, precipitação, adsorção, permuta de íons, eletrodiálise, eletrodiálise de membrana bipolar e/ou osmose inversa. A Vitamina C pode ser adicionalmente purificada como a forma de ácido livre ou qualquer uma de suas formas de sal conhecidas por intermédio de operações como, por exemplo, tratamento com carvão ativado, permuta de íons, adsorção e eluição, concentração, cristalização, filtração e secagem. Especificamente, uma primeira separação de Vitamina C a partir de outros componentes no fluxo de colheita poderia ser executada por qualquer combinação ou repetição apropriada, por exemplo, dos seguintes métodos: eletrodiálise de dois ou três compartimentos, eletrodiálise de membrana bipolar, osmose inversa ou adsorção, por exemplo, em resinas de permuta de íons ou resinas não iônicas. Se a forma resultante da Vitamina C for um sal de ácido L-ascórbico, a conversão da forma de sal na forma de ácido livre pode ser realizada, por exemplo, por eletrodiálise de membrana bipolar, permuta de íons, técnicas cromatográficas de leito em movimento simulado, e similares. A combinação das etapas mencionadas, por exemplo, eletrodiálise e eletrodiálise de membrana bipolar em uma etapa poderia ser também utilizada como combinação das etapas mencionadas, por exemplo, várias etapas de permuta de íons utilizando métodos cromatográficos de leito

em movimento simulado. Qualquer um desses procedimentos individualmente ou em combinação constitui um meio conveniente para isolar e purificar o produto, isto é, Vitamina C. O produto desse modo obtido pode ser
5 adicionalmente isolado em um modo como, por exemplo por concentração, cristalização, precipitação, lavagem e secagem dos cristais e/ou purificados adicionalmente, por exemplo, por tratamento com carvão ativado, permuta de íons e/ou recristalização.

10 Em uma modalidade preferida, a vitamina C é purificada a partir do fluxo de colheita por uma série de etapas de processamento a jusante, como descrito acima, sem ter de ser transferida para uma solução não aquosa em qualquer momento desse processamento, isto é, todas as etapas são
15 executadas em um ambiente aquoso. Tal procedimento de processamento a jusante preferido pode incluir, por exemplo, a concentração do fluxo de colheita vindo do recipiente de produção por intermédio de eletrodiálise de dois ou três compartimentos, conversão de vitamina C em sua
20 forma de sal presente na solução concentração em sua forma de ácido por intermédio de eletrodiálise de membrana bipolar e/ou permuta de íons, purificação pelos métodos como, por exemplo, tratamento com carvão ativado, resinas não iônicas ou de permuta de íon, seguido por uma etapa de
25 concentração adicional e cristalização. Esses cristais podem ser separados, lavados e secos. Se necessário, os cristais podem ser novamente resolubilizados em água, tratados com carvão ativado e/ou resinas de permuta de íons e recristalizadas. Esses cristais podem ser então
30 separados, lavados e secos.

Modalidades vantajosas da invenção tornar-se-ão evidentes a partir das reivindicações dependentes. Esses e outros aspectos e modalidades da presente invenção devem ser evidentes para aqueles versados na técnica a partir dos ensinamentos da presente invenção.

A seqüência do gene compreendendo uma seqüência de nucleotídeos de acordo com SEQ ID NO: 1 codificando uma proteína de SMS 27 foi determinada pelo sequenciamento de um clone genômico obtido a partir de *Gluconobacter oxydans* IFO 3293.

A invenção também se refere a um polinucleotídeo codificando pelo menos um fragmento biologicamente ativo ou derivado de um polipeptídeo de SMS 27 como mostrado na SEQ ID NO: 2.

Como utilizado aqui, "derivado ou fragmento biologicamente ativo" significa um polipeptídeo que retém essencialmente a mesma função ou atividade biológica que o polipeptídeo mostrado na SEQ ID NO: 2. Os exemplos de atividade biológica podem, por exemplo, ser atividade enzimática, atividade de sinalização ou reatividade de anticorpo. O termo "mesma função biológica" ou "equivalente funcional" como utilizado aqui, significa que a proteína tem essencialmente a mesma atividade biológica, por exemplo, reatividade enzimática, sinalização ou de anticorpo, como um polipeptídeo mostrado na SEQ ID NO:2.

Os polipeptídeos e polinucleotídeos da presente invenção são preferivelmente fornecidos em uma forma isolada, e preferivelmente são purificados para homogeneidade.

O termo "isolado" quer dizer que o material é removido

de seu ambiente original (por exemplo, o ambiente natural se estiver ocorrendo naturalmente). Por exemplo, um polipeptídeo ou polinucleotídeo de ocorrência natural presente em um microorganismo vivo não é isolado, porém o mesmo polinucleotídeo ou polipeptídeo, separado de alguns ou todos dos materiais coexistentes no sistema natural, é isolado. Tais polinucleotídeos poderiam ser parte de um vetor e/ou tais polinucleotídeos ou polipeptídeos poderiam ser parte de uma composição e ainda ser isolados em que tal vetor ou composição não faça parte de seu ambiente natural.

Um polinucleotídeo ou ácido nucléico isolado como utilizado aqui, pode ser um DNA ou RNA que não é imediatamente contínuo com as duas seqüências de codificação com as quais está imediatamente contínua (uma na extremidade-5' e uma na extremidade-3') no genoma de ocorrência natural do organismo do qual é derivado. Desse modo, em uma modalidade, um ácido nucléico inclui algumas ou todas as seqüências de não codificação 5' (por exemplo, promotor) que estão imediatamente contíguos à seqüência de codificação. O termo "polinucleotídeo isolado" inclui, portanto, por exemplo, um DNA recombinante que é incorporado em um vetor, em um vírus ou plasmídeo de replicação autônoma, ou no DNA genômico de um procariota ou eucariota, ou que existe como uma molécula separada (por exemplo, um cDNA ou um fragmento de DNA genômico produzido por PCR ou tratamento de endonuclease de restrição) independente de outras seqüências. Também inclui um DNA recombinante que faz parte de um gene híbrido codificando um polipeptídeo adicional que é substancialmente livre de material celular, material viral ou meio de cultura (quando

produzido por técnicas de DNA recombinante), ou precursores químicos ou outros produtos químicos (quando quimicamente sintetizado). Além disso, um "fragmento de ácido nucléico isolado" é um fragmento de ácido nucléico que não ocorre naturalmente como um fragmento e não seria encontrado no estado natural.

Como utilizado aqui, os termos "polinucleotídeo", "gene" e "gene recombinante" se referem a moléculas de ácido nucléico que podem ser isoladas do DNA cromossômico, que incluem um quadro de leitura aberto codificando uma proteína, por exemplo, proteínas de SMS *G. oxydans* IFO 3293. Um polinucleotídeo pode incluir uma seqüência de polinucleotídeo como mostrado na SEQ ID NO: 1 ou fragmentos do mesmo e regiões a montante e a jusante das seqüências de gene que podem incluir, por exemplo, regiões promotoras, regiões reguladoras e regiões terminadoras importantes para a expressão apropriada e estabilização do polipeptídeo derivado das mesmas.

Um gene pode incluir seqüências de codificação, seqüências não de codificação como, por exemplo, seqüências não traduzidas localizadas nas extremidades 3' e 5' da região de codificação de um gene, e seqüências reguladoras. Além disso, um gene se refere a uma molécula de ácido nucléico isolado, como definido aqui. É adicionalmente reconhecido pela pessoa versada que os polimorfismos de seqüência de DNA que levam a mudanças nas seqüências de aminoácido de proteínas de SMS podem existir em uma população, por exemplo, a população de *Gluconobacter oxydans*. Tal polimorfismo genético no gene de SMS 27 pode existir entre indivíduos em uma população devido à variação

natural ou em células de populações diferentes. Tais variações naturais podem resultar, tipicamente, em 1-5% de variância na seqüência de nucleotídeos do gene de SMS 27. Todas e quaisquer tais variações de nucleotídeo e o polimorfismo de aminoácido resultante em SMS 27 são resultado de variação natural e que não alteram a atividade funcional de proteínas de SMS, pretendem estar compreendidas no escopo da invenção.

Como utilizado aqui, os termos "polinucleotídeo" ou "molécula de ácido nucléico" pretendem incluir moléculas de DNA (por exemplo, cDNA ou DNA genômico) e moléculas de RNA (por exemplo, mRNA) e análogos do DNA ou RNA gerado utilizando análogos de nucleotídeo. A molécula de ácido nucléico pode ser de fita única ou fita dupla, porém preferivelmente é DNA de fita dupla. O ácido nucléico pode ser sintetizado utilizando análogos ou derivados de oligonucleotídeo (por exemplo, nucleotídeos de fosforotioato ou inosina). Tais oligonucleotídeos podem ser utilizados, por exemplo, para preparar ácidos nucléicos que têm capacidades de pares de base alteradas ou resistência aumentada a nucleases.

A informação de seqüência como fornecida aqui, não deve ser tão estreitamente construída de modo a exigir inclusão de bases erroneamente identificadas. As seqüências específicas reveladas aqui podem ser facilmente utilizadas para isolar o gene completo a partir de um microorganismo recombinante ou não recombinante capaz de converter uma dada fonte de carbono diretamente em Vitamina C e/ou 2-KGA, em particular *Gluconobacter oxydans*, preferivelmente *Gluconobacter oxydans* IFO 3293 que por sua vez pode ser

facilmente submetido a análises de seqüência adicionais desse modo identificando erros de seqüência.

A menos que de outro modo indicado, todas as seqüências de nucleotídeo determinadas pelo sequenciamento de uma molécula de DNA da presente invenção foram determinadas utilizando um seqüenciador de DNA automatizado e todas as seqüências de aminoácidos de polipeptídeos codificados por moléculas de DNA determinadas aqui foram previstas por translação de uma seqüência de DNA determinada como acima. Portanto, como sabido na técnica para qualquer seqüência de DNA determinada por essa abordagem automatizada, qualquer seqüência de nucleotídeo determinada aqui pode conter alguns erros. Seqüências de nucleotídeo determinadas por automação são tipicamente pelo menos aproximadamente 90% idênticas, mais tipicamente pelo menos aproximadamente 95% a pelo menos aproximadamente 99,9% idênticas à seqüência de nucleotídeo efetiva da molécula de DNA seqüenciada. A seqüência efetiva pode ser mais precisamente determinada por outras abordagens incluindo métodos de sequenciamento de DNA manual bem conhecidos na técnica. Como é também conhecida na técnica, uma única inserção ou deleção em uma seqüência de nucleotídeos determinada em comparação com a seqüência efetiva causará um deslocamento de quadro em translação da seqüência de nucleotídeo de tal modo que a seqüência de aminoácidos prevista codificada por uma seqüência de nucleotídeo determinada será totalmente diferente da seqüência de aminoácidos efetivamente codificada pela molécula de DNA seqüenciada, iniciando no ponto dessa inserção ou deleção.

A pessoa versada na técnica é capaz de identificar essas bases erroneamente identificadas e sabe como corrigir esses erros.

Uma molécula de ácido nucléico, de acordo com a
5 invenção, pode compreender somente uma porção ou um
fragmento da seqüência de ácido nucléico fornecida pela
presente invenção, como, por exemplo, a seqüência mostrada
na SEQ ID NO: 1, por exemplo, um fragmento que pode ser
utilizado como uma sonda ou iniciador como, por exemplo,
10 SEQ ID NO: 3 ou SEQ ID NO: 4 ou um fragmento codificando
uma porção de uma proteína, de acordo com a invenção. A
seqüência de nucleotídeo determinada a partir da clonagem
do gene de SMS 27 permite a geração de sondas e iniciadores
projetados para uso na identificação e/ou clonagem de
15 outros membros da família SMS 27, bem como homólogos de SMS
27 de outras espécies. A sonda/iniciador compreende,
tipicamente, oligonucleotídeos substancialmente purificados
que compreende, tipicamente, uma região de seqüência de
nucleotídeo que hibridiza preferivelmente sob condições
20 altamente rigorosas a pelo menos aproximadamente 12 ou 15,
preferivelmente aproximadamente 18 ou 20, mais
preferivelmente aproximadamente 22 ou 25, ainda mais
preferivelmente aproximadamente 30, 35, 40, 45, 50, 55, 60,
65 ou 75 ou mais nucleotídeos consecutivos de uma seqüência
25 de nucleotídeos mostrada na SEQ ID NO: 1 ou um fragmento ou
derivado da mesma.

Uma molécula de ácido nucléico abrangendo toda ou uma
porção da seqüência de ácido nucléico da SEQ ID NO: 1 pode
ser também isolada pela reação de cadeia de polimerase
30 (PCR) utilizando iniciadores de oligonucleotídeo sintético

projetados com base nas informações de sequência contidas aqui.

Um ácido nucléico da invenção pode ser amplificado utilizando cDNA, mRNA ou alternativamente, DNA genômico, como padrão e iniciadores de oligonucleotídeo apropriados, de acordo com técnicas de amplificação de PCR padrão. O ácido nucléico desse modo amplificado pode ser clonado em um vetor apropriado e caracterizado pelo fato de análise de sequência de DNA.

Os fragmentos de um polinucleotídeo, de acordo com a invenção, podem compreender também polinucleotídeos não codificando polipeptídeos funcionais. Tais polinucleotídeos podem funcionar como sondas ou iniciadores para uma reação de PCR.

Ácidos nucléicos, de acordo com a invenção, independente de se codificam polipeptídeos funcionais ou não funcionais, podem ser utilizados como sondas de hibridização ou iniciadores de reação de cadeia de polimerase (PCR). Os usos das moléculas de ácido nucléico da presente invenção que não codificam um polipeptídeo tendo uma atividade de SMS 27 incluem, entre outras coisas, (1) isolar o gene codificando a proteína da presente invenção, ou variantes alélicos do mesmo a partir de uma biblioteca de cDNA, por exemplo, de outros organismos do que *Gluconobacter oxydans* e (2) análise de Northern blot para detectar expressão de mRNA da proteína em células específicas ou (3) uso em aumentar e/ou aperfeiçoar a função ou atividade de genes de SMS 27 homólogos nos outros organismos.

Sondas baseadas nas seqüências de nucleotídeo

fornecidas aqui, podem ser utilizadas para detectar transcritos ou seqüências genômicas codificando proteínas iguais ou homólogas, por exemplo, em outros organismos. Moléculas de ácido nucléico correspondendo a variantes naturais e homólogos de não *G. oxydans* do DNA *G. oxydans* SMS 27 da invenção que são também abrangidas pela presente invenção, podem ser isoladas com base em sua homologia ao ácido nucléico *G. oxydans* SMS 27 revelado aqui utilizando o DNA de *G. oxydans*, ou uma porção do mesmo, como uma sonda de hibridização de acordo com técnicas de hibridização padrão, preferivelmente sob condições de hibridização altamente rigorosas.

Em modalidades preferidas, a sonda compreende ainda um grupo de rótulo fixado à mesma, por exemplo, o grupo de rótulo pode ser um radioisótopo, um composto fluorescente, uma enzima, ou um co-factor de enzima.

As seqüências de gene homólogo podem ser isoladas, por exemplo, pela execução de PCR utilizando dois grupos de iniciador de oligonucleotídeo de degenerar projetados com base em seqüências de nucleotídeo como ensinado aqui.

O padrão para a reação pode ser cDNA obtido por transcrição inversa de mRNA preparado de cepas conhecidas ou que se suspeita expressarem um polinucleotídeo de acordo com a invenção. O produto de PCR pode ser subclonado e seqüenciado para assegurar que as seqüências amplificadas representam as seqüências de uma nova seqüência de ácido nucléico como descrito aqui, ou um equivalente funcional da mesma.

O fragmento de PCR pode ser então utilizado para isolar um clone de cDNA de comprimento completo por uma

variedade de métodos conhecidos. Por exemplo, o fragmento amplificado pode ser rotulado e utilizado para classificar uma biblioteca de cDNA de cosmídio ou bacteriófago. Alternativamente, o fragmento rotulado pode ser utilizado para classificar uma biblioteca genômica.

A tecnologia de PCR também pode ser utilizada para isolar seqüências de cDNA de comprimento total a partir de outros organismos. Por exemplo, RNA pode ser isolado, seguindo procedimentos padrão, a partir de uma fonte de tecido ou celular apropriada. Uma reação de transcrição inversa pode ser executada no RNA utilizando um iniciador de oligonucleotídeo específico para a extremidade mais 5' do fragmento amplificado para a iniciação da síntese de primeira fita.

O híbrido de RNA/DNA resultante pode então ser "caudado" (por exemplo, com guaninas) utilizando uma reação de transferase terminal padrão, o híbrido pode ser digerido com RNaseH, e a síntese de segunda fita pode ser então iniciada (por exemplo, com iniciador poli-C). Desse modo, seqüências de cDNA a montante do fragmento amplificado podem ser facilmente isolados. Para um exame de estratégias de clonagem úteis, vide, por exemplo, Sambrook e outros, supra; e Ausubel e outros, supra.

Além disso, ácidos nucléicos codificando outros membros da família SMS 27, que desse modo têm uma seqüência de nucleotídeo que difere de uma seqüência de nucleotídeo de acordo com SEQ ID NO:1, estão compreendidos no escopo da invenção. Além disso, ácidos nucléicos codificando proteínas de SMS 27 a partir de espécies diferentes que desse modo podem ter uma seqüência de nucleotídeo que

difere de uma sequência de nucleotídeo mostrada na SEQ ID NO: 1 estão compreendidas no escopo da invenção.

A invenção também se refere a um polinucleotídeo isolado hibridizável sob condições rigorosas, preferivelmente sob condições altamente rigorosas, em um polinucleotídeo como o da presente invenção, como por exemplo, um polinucleotídeo mostrado na SEQ ID NO: 1. Vantajosamente, tal polinucleotídeo pode ser obtido de um microorganismo capaz de converter uma dada fonte de carbono diretamente em Vitamina C, em particular *Gluconobacter oxydans*, preferivelmente *Gluconobacter oxydans* IFO 3293.

Como utilizado aqui, o termo "hibridizar" pretende descrever condições para hibridização e lavagem sob as quais as seqüências de nucleotídeo pelo menos aproximadamente 50%, pelo menos aproximadamente 60%, pelo menos aproximadamente 70%, mais preferivelmente pelo menos aproximadamente 80%, ainda mais preferivelmente pelo menos aproximadamente 85% a 90%, mais preferivelmente ainda pelo menos 95% homólogas entre si permanecem tipicamente hibridizadas entre si.

Em uma modalidade, um ácido nucléico da invenção é pelo menos 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% ou mais homólogo a uma sequência de ácido nucléico mostrada na SEQ ID NO: 1 ou complemento da mesma.

Um exemplo não limitador preferido de condições de hibridização rigorosas é hibridização em 6x de cloreto de sódio/citrato de sódio (SSC) em aproximadamente 45°C, seguido por uma ou mais lavagens em 1x SSC, 0,1% SDS a 50°C, preferivelmente a 55°C, mais preferivelmente a 60°C e

ainda mais preferivelmente a 65°C.

Condições altamente vigorosas incluem incubações a 42°C por um período de vários dias, como 2-4 dias, utilizando uma sonda de DNA rotulada, como sonda de DNA rotulada com digoxigenin (DIG), seguido por uma ou mais lavagens em 2x SSC, 0,1% SDS em temperatura ambiente e uma ou mais lavagens em 0,5x SSC, 0,1% SDS ou 0,1x SSC, 0,1% SDS a 65-68°C. Em particular, condições altamente rigorosas incluem, por exemplo, 2 h a 4 dias de incubação a 42°C utilizando uma sonda de DNA rotulada DIG (preparada, por exemplo, utilizando um sistema de rotulação DIG; Roche Diagnostics GmbH, 68298 Mannheim, Alemanha) em uma solução como solução DigEasyHyb (Roche Diagnostics GmbH) com ou sem 100 µg/ml de DNA de esperma de salmão, ou uma solução compreendendo formamida a 50%, 5x SSC (150 mM NaCl, 15 mM de citrato trissódico), 0,02% de sulfato de dodecil de sódio, 0,1% de N-lauroíl sarcosina, e 2% de reagente de bloqueio (Roche Diagnostics GmbH) seguido por lavagem dos filtros duas vezes por 5 a 15 minutos em 2x SSC e 0,1% SDS em temperatura ambiente e então lavagem duas vezes por 15-30 minutos em 0,5x SSC e 0,1% SDS ou 0,1x SSC e 0,1% SDS a 65-68°C.

Preferivelmente, uma molécula de ácido nucléico isolada da invenção que hibridiza preferivelmente sob condições altamente rigorosas em uma seqüência de nucleotídeo da invenção corresponde a uma molécula de ácido nucléico de ocorrência natural. Como utilizado aqui, uma molécula de ácido nucléico de "ocorrência natural" se refere a uma molécula de RNA ou DNA tendo uma seqüência de nucleotídeo que ocorre na natureza (por exemplo, codifica

uma proteína natural). Em uma modalidade, o ácido nucléico codifica uma proteína de *G. oxydans* SMS 27 natural.

O técnico especializado saberá quais condições aplicar para condições de hibridização rigorosa e altamente
5 rigorosa. A orientação adicional em relação a tais condições é prontamente disponível na técnica, por exemplo, em Sambrook e outros, 1989, *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Press, N.Y.; e Ausubel e outros (eds.), 1995, *Current Protocols in Molecular Biology* (John
10 Wiley & Sons, N.Y.). Evidentemente, um polinucleotídeo que hibridiza somente a uma seqüência poli(A) (como o trato poli(A) terminal-3' de mRNAs), ou a uma extensão complementar de resíduos T (ou U), não seria incluído em um polinucleotídeo da invenção utilizado para hibridizar
15 especificamente a uma porção de um ácido nucléico da invenção, uma vez que tal polinucleotídeo hibridizaria a qualquer molécula de ácido nucléico contendo uma extensão de poli (A) ou complemento do mesmo (por exemplo, praticamente qualquer clone de cDNA de fita dupla).

20 Em uma abordagem típica, bibliotecas de cDNA ou DNA genômico construído de outros organismos, por exemplo, microorganismos capazes de converter uma dada fonte de carbono diretamente em Vitamina C, em particular outra espécie de *Gluconobacter* pode ser classificada.

25 Por exemplo, as cepas de *Gluconobacter* podem ser classificadas em relação a polinucleotídeos homólogos por análise de Southern e/ou Northern blot. Após detecção de transcritos homólogos em polinucleotídeos, de acordo com a invenção, bibliotecas de DNA podem ser construídas de RNA
30 isolado a partir da cepa apropriada, utilizando técnicas

padrão bem conhecidas por aqueles versados na técnica. Alternativamente, uma biblioteca de DNA genômico total pode ser classificada utilizando uma sonda hibridizável em um polinucleotídeo, de acordo com a invenção.

5 Uma molécula de ácido nucléico da presente invenção, como por exemplo, uma molécula de ácido nucléico mostrada na SEQ ID NO: 1 ou um fragmento ou derivado da mesma, pode ser isolado utilizando técnicas de biologia molecular padrão e as informações de seqüência fornecidas aqui. Por
10 exemplo, utilizando toda ou uma parte da seqüência de ácido nucléico mostrada na SEQ ID NO: 1 como uma sonda de hibridização, moléculas de ácido nucléico de acordo com a invenção podem ser isoladas utilizando técnicas de clonagem e hibridização padrão (por exemplo, como descrito em
15 Sambrook, J., Fritsh, E.F., e Maniatis, T. Molecular Cloning: A Laboratory Manual. 2ª ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).

Além disso, oligonucleotídeos correspondendo a ou
20 hibridizáveis com seqüências de nucleotídeo, de acordo com a invenção, podem ser preparados por técnicas sintéticas padrão, por exemplo, utilizando um sintetizador de DNA automatizado.

Os termos "homologia" ou "percentagem de identidade"
25 são utilizados de forma intercambiável aqui. Para fins da presente invenção, é definido aqui que para determinar a percentagem de identidade de duas seqüências de aminoácido ou de duas seqüências de ácido nucléico, as seqüências são alinhadas para fins de comparação ótima (por exemplo,
30 intervalos podem ser introduzidos na seqüência de uma

primeira seqüência de ácido nucléico ou aminoácido para alinhamento ótimo com uma segunda seqüência de ácido nucléico ou aminoácido). Os resíduos de aminoácido ou nucleotídeos em posições de aminoácido correspondentes ou posições de nucleotídeo são então comparados. Quando uma posição na primeira seqüência é ocupada pelo mesmo resíduo de aminoácido ou nucleotídeo que a posição correspondente na segunda seqüência, então as moléculas são idênticas naquela posição. A percentagem de identidade entre as duas seqüências é uma função do número de posições idênticas compartilhadas pelas seqüências (isto é, % de identidade = número de posições idênticas/número total de posições (isto é, posições de sobreposição) x 100). Preferivelmente, as duas seqüências são do mesmo comprimento.

A pessoa versada estará ciente do fato de que vários programas de computador diferentes são disponíveis para determinar a homologia entre duas seqüências. Por exemplo, uma comparação de seqüências e determinação de percentagem de identidade entre duas seqüências pode ser realizada utilizando um algoritmo matemático. Em uma modalidade preferida, a percentagem de identidade entre duas seqüências de aminoácido é determinada utilizando o algoritmo de Needleman e Wunsch (J. Mol. Bio. (48): 444-453 (1970)) que foi incorporado no programa GAP no pacote de software GCG (disponível em <http://www.accelrys.com>), utilizando uma matriz Blossom 62 ou uma matriz PAM250, e um peso de intervalo de 16, 14, 12, 10, 8, 6 ou 4, e um peso de comprimento de 1, 2, 3, 4, 5 ou 6. A pessoa versada reconhecerá que todos esses parâmetros diferentes fornecerão resultados levemente diferentes porém que a

percentagem de identidade geral de duas seqüências não é significativamente alterada ao utilizar diferentes algoritmos.

Ainda em outra modalidade, a percentagem de identidade entre duas seqüências de nucleotídeo é determinada utilizando o programa GAP no pacote de software GCG (disponível em <http://www.accelrys.com>) utilizando uma matriz NWSgapdna.CMP e um peso de intervalo de 40, 50, 60, 70 ou 80 e um peso de comprimento de 1, 2, 3, 4, 5 ou 6. Em outra modalidade, a percentagem de identidade entre duas seqüências de nucleotídeo ou aminoácido é determinada utilizando o algoritmo de E. Meyers e W. Miller (CABIOS 4:11-17 (1989) que foi incorporado no programa ALIGN (versão 2.0) (disponível em <http://vega.igh.cnrs.fr/bin/align-guess.cgi>) utilizando uma tabela de resíduo de peso PAM120, uma penalidade de comprimento de intervalo de 12 e uma penalidade de intervalo de 4.

As seqüências de proteína e ácido nucléico da presente invenção podem ser adicionalmente utilizadas como uma "seqüência de consulta" para executar uma busca contra bancos de dados públicos para, por exemplo, identificar outros membros da família ou seqüências relacionadas. Tais buscas podem ser executadas utilizando os programas BLASTN e BLASTX (versão 2.0) de Altschul, e outros (1990) J. Mol. Biol. 215:403-10. Buscas de nucleotídeo BLAST podem ser executadas com o programa BLASTN, marcação = 100, comprimento de palavra = 12 para obter seqüências de nucleotídeo homólogas às moléculas de ácido nucléico da presente invenção. Buscas de proteína BLAST podem ser

executadas com o programa BLASTX, marcação = 50, comprimento de palavra = 3 para obter seqüências de aminoácido homólogas às moléculas de proteína da presente invenção. Para obter alinhamentos com intervalo para fins de comparação, Gapped BLAST pode ser utilizado como descrito em Altschul e outros (1997) Nucleic Acids Res. 25 (17): 3389-3402. Ao utilizar programas BLAST e Gapped BLAST, os parâmetros padrão dos respectivos programas (por exemplo, BLASTX e BLASTN) podem ser utilizados. Vide <http://www.ncbi.nlm.nih.gov>.

Em outra modalidade preferida, uma molécula de ácido nucléico isolada da invenção compreende uma molécula de ácido nucléico que é o complemento de uma seqüência de nucleotídeo como da presente invenção, como por exemplo, a seqüência mostrada na SEQ ID NO: 1. Uma molécula de ácido nucléico, que é complementar a uma seqüência de nucleotídeo revelada aqui, é uma que é suficientemente complementar a uma seqüência de nucleotídeo mostrada na SEQ ID NO: 1 de tal modo que pode hibridizar com a seqüência de nucleotídeo desse modo formando um dúplex estável.

Em uma modalidade preferida adicional, um ácido nucléico da invenção como mostrado na SEQ ID NO: 1 ou o complemento do mesmo contém pelo menos uma mutação que leva a um produto de gene com atividade/função modificada. Pelo menos uma mutação pode ser introduzida por métodos descritos aqui. Em um aspecto, pelo menos uma mutação leva a uma proteína de SMS 27 cuja função comparada com a cópia do tipo selvagem é total ou parcialmente destruída. Os métodos para introduzir essas mutações são bem conhecidos na técnica.

O termo "redução" de atividade, como utilizado aqui, abrange diminuir a atividade de um ou mais polipeptídeos no organismo de produção, que por sua vez são codificados pelos polinucleotídeos correspondentes descritos aqui. Há diversos métodos disponíveis na técnica para realizar a redução de atividade de uma dada proteína, nesse caso a proteína de SMS 27. Em geral, a atividade específica de uma proteína pode ser diminuída ou o número de cópia da proteína pode ser diminuído.

Para facilitar essa diminuição, o número de cópia dos genes correspondendo aos polinucleotídeos descritos aqui pode ser diminuído, como por exemplo, por subexpressão ou ruptura de um gene. Diz-se que um gene é "sub-expresso" se o nível de transcrição do gene for reduzido em comparação com o gene do tipo selvagem. Isso pode ser medido, por exemplo, por análise de Northern blot quantificando a quantidade de mRNA como uma indicação para expressão de gene. Como utilizado aqui, um gene é sub-expresso se a quantidade de mRNA gerado for diminuída em pelo menos 1%, 2%, 5%, 10%, 25%, 50%, 75%, 100%, 200% ou até mesmo mais do que 500%, em comparação com a quantidade de mRNA gerado de um gene do tipo selvagem. Alternativamente, um promotor fraco pode ser utilizado para orientar a expressão do polinucleotídeo. Em outra modalidade, o promotor, região reguladora e/ou o sítio de ligação de ribossomo a montante do gene podem ser alterados para obter a expressão descendente. A expressão também pode ser reduzida pela diminuição da meia-vida relativa do RNA mensageiro. Em outra modalidade, a atividade do próprio polipeptídeo pode ser diminuída empregando uma ou mais mutações na seqüência

de aminoácido de polipeptídeo, que diminui a atividade. Por exemplo, a alteração da afinidade do polipeptídeo para seu substrato correspondente pode resultar em atividade reduzida. De modo semelhante, a meia vida relativa do polipeptídeo pode ser diminuída. Em qualquer cenário, seja expressão de gene reduzida ou atividade reduzida, a redução pode ser obtida pela alteração da composição do meio de cultura de célula e/ou métodos utilizados para cultivo. "Expressão reduzida" ou "atividade reduzida", como utilizado aqui, significa uma diminuição de pelo menos 5%, 10%, 25%, 50%, 75%, 100%, 200% ou mesmo mais do que 500%, em comparação com uma proteína do tipo selvagem, polinucleotídeo, gene; ou a atividade e/ou a concentração da proteína presente antes dos polinucleotídeos ou polipeptídeos serem reduzidos. A atividade da proteína de SMS 27 pode ser também reduzida pelo contato da proteína com um inibidor geral ou específico de sua atividade. Os termos "atividade reduzida", "atividade diminuída ou abolida" são utilizados de forma intercambiável aqui.

Outro aspecto da invenção se refere a vetores, contendo um ácido nucléico codificando uma proteína de acordo com a invenção ou um equivalente funcional ou porção do mesmo. Como utilizado aqui, o termo "vetor" se refere a uma molécula de ácido nucléico capaz de transportar outro ácido nucléico ao qual foi ligado. Um tipo de vetor é um "plasmídeo", que se refere a um loop de DNA de fita dupla, circular no qual segmentos adicionais de DNA podem ser ligados. Outro tipo de vetor é um vetor viral, onde segmentos adicionais de DNA podem ser ligados no genoma viral. Certos vetores são capazes de replicação autônoma em

uma célula hospedeira na qual são introduzidos (por exemplo, vetores bacterianos tendo uma origem de replicação bacteriana). Outros vetores são integrados no genoma de uma célula hospedeira após introdução na célula hospedeira, e desse modo são replicados juntamente com o genoma hospedeiro.

Os vetores recombinantes da invenção compreendem um ácido nucléico da invenção em uma forma apropriada para expressão do ácido nucléico em uma célula hospedeira, o que significa que o vetor de expressão recombinante inclui uma ou mais seqüências reguladoras, selecionadas com base nas células hospedeiras a serem utilizadas para expressão, que é operativamente ligado à seqüência de ácido nucléico a ser expressa. Em um vetor de expressão recombinante, "operativamente ligado" pretende significar que a seqüência de nucleotídeo de interesse é ligada à(s) seqüência(s) reguladora(s) em um modo que permite expressão da seqüência de nucleotídeo (por exemplo, em um sistema de transcrição/translação in vitro ou em uma célula hospedeira quando o vetor é introduzido na célula hospedeira). O termo "seqüência reguladora" pretende incluir promotores, intensificadores e outros elementos de controle de expressão (por exemplo, atenuador). Tais seqüências reguladoras são descritas, por exemplo, em Goeddel; Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, CA (1990). Seqüências reguladoras incluem aquelas que orientam expressão induzível ou constitutiva de uma seqüência de nucleotídeo em muitos tipos de células hospedeiras e aquelas que orientam a expressão da seqüência de nucleotídeo somente em uma certa célula hospedeira (por

exemplo, seqüências reguladoras específicas de tecido). Será reconhecido por aqueles versados na técnica que o desenho do vetor de expressão pode depender de tais fatores como a escolha da célula hospedeira a ser transformada, o nível de expressão de proteína desejado, etc. Os vetores de expressão da invenção podem ser introduzidos nas células hospedeiras para desse modo produzir proteínas ou peptídeos, codificados por ácidos nucleicos como descrito aqui, incluindo, porém não limitado a, proteínas mutantes, fragmentos das mesmas, variantes ou equivalentes funcionais das mesmas, e proteínas de fusão, codificadas por um ácido nucleico como descrito aqui, por exemplo, proteínas de SMS 27, formas mutantes de proteínas de SMS 27, proteínas de fusão e similares.

Os vetores de expressão recombinante da invenção podem ser projetados para expressão de proteínas de SMS 27 em um microorganismo apropriado. Por exemplo, uma proteína de acordo com a invenção pode ser expressa em células bacterianas como cepas que pertence aos gêneros *Gluconobacter*, *Gluconacetobacter* ou *Acetobacter*. Vetores de expressão úteis na presente invenção incluem vetores derivados de cromossomo, episssomo e vírus, por exemplo, vetores derivados de plasmídeos bacterianos, bacteriófagos, e vetores derivados de combinações dos mesmos como aqueles derivados de elementos genéticos de bacteriófago e plasmídeo, como cosmídios e fagemídios.

O inserto de DNA pode ser operativamente ligado a um promotor apropriado, que pode ser um promotor constitutivo ou induzível. A pessoa versada saberá como selecionar promotores apropriados. As construções de expressão podem

conter sítios para iniciação de transcrição, terminação e, na região transcrita, um sítio de ligação de ribossomo para translação. A porção de codificação dos transcritos maduros expressos pelas construções pode incluir, preferivelmente, um códon de iniciação no início e um códon de terminação apropriadamente posicionado na extremidade do polipeptídeo a ser traduzida.

O DNA de vetor pode ser introduzido em células hospedeiras apropriadas através de técnicas de transfecção ou transformação, convencionais. Como utilizado aqui, os termos "transformação", "transconjugação" e "transfecção" pretendem se referir a uma variedade de técnicas reconhecidas na arte para introduzir ácido nucléico estranho (por exemplo, DNA) em uma célula hospedeira, incluindo co-precipitação de fosfato de cálcio ou cloreto de cálcio, transfecção mediada por dextrano-DEAE, transdução, infecção, lipofecção, transfecção mediada por lipídeo catiônico ou eletroporação. Métodos apropriados para transformar ou transfectar células hospedeiras podem ser encontrados em Sambrook, e outros, (supra), Davis e outros, Basic Methods in Molecular Biology (1986) e outros manuais de laboratório.

Para identificar e selecionar células que integraram o DNA estranho em seu genoma, um gene que codifica um marcador selecionável (Por exemplo, resistência a antibiótico) é geneticamente introduzido nas células hospedeiras juntamente com o gene de interesse. Marcadores selecionáveis preferidos incluem aqueles que conferem resistência a drogas, como canamicina, tetraciclina, ampicilina e estreptomicina. Um ácido nucléico codificando

um marcador selecionável é preferivelmente introduzido em uma célula hospedeira no mesmo vetor que aquele que codifica uma proteína de acordo com a invenção ou pode ser introduzido em um vetor separador como, por exemplo, um vetor suicide, que não pode replicar nas células hospedeiras. As células estavelmente transfectadas com o ácido nucléico introduzido podem ser identificadas por seleção de droga (por exemplo, células que incorporaram o gene marcador selecionável sobreviverão, enquanto as outras células morrerão).

A invenção provê também um polipeptídeo isolado tendo a seqüência de aminoácido mostrada na SEQ ID NO: 2 ou uma seqüência de aminoácido que pode ser obtido por expressão de um polinucleotídeo da presente invenção, como por exemplo, uma seqüência de polinucleotídeo mostrada na SEQ ID NO:1 em um hospedeiro apropriado.

Polipeptídeos de acordo com a invenção, podem conter somente substituições conservativas de um ou mais aminoácidos na seqüência de aminoácidos representada pela SEQ ID NO: 2 ou substituições, inserções ou deleções de aminoácidos não essenciais. Por conseguinte, um aminoácido não essencial é um resíduo que pode ser alterado nas seqüências de aminoácidos mostradas na SEQ ID NO: 2 sem alterar substancialmente a função biológica. Por exemplo, resíduos de aminoácido que são conservados entre as proteínas da presente invenção, são previstos como sendo particularmente não sensíveis à alteração. Além disso, aminoácidos conservados entre as proteínas, de acordo com a presente invenção e outras proteínas de SMS 27 não são provavelmente sensíveis à alteração.

O termo "substituição conservativa" pretende significar que uma substituição na qual o resíduo de aminoácido é substituído com um resíduo de aminoácido tendo uma cadeia lateral similar. Essas famílias são conhecidas na técnica e incluem aminoácidos com cadeias laterais básicas (por exemplo, lisina, arginina e histidina), cadeias laterais acídicas (por exemplo, ácido aspártico, ácido glutâmico), cadeias laterais polares não carregadas (por exemplo, glicina, asparagina, glutamina, serina, treonina, tirosina, cisteína), cadeias laterais não polares (por exemplo, alanina, valina, leucina, isoleucina, prolina, fenil alanina, metionina, triptofano), cadeias laterais ramificadas-beta (por exemplo, treonina, valina, isoleucina), e cadeias alterais aromáticas (por exemplo, tirosina, fenil alanina, triptofano, histidina).

Como mencionado acima, os polinucleotídeos da invenção podem ser utilizados na engenharia genética de uma célula hospedeira apropriada para tornar a mesma melhor e mais eficiente na fermentação, por exemplo, em um processo de fermentação direta para Vitamina C.

De acordo com a invenção, uma célula hospedeira geneticamente construída/recombinantemente produzida (também mencionada como uma célula recombinante ou célula transformada) contendo tal polinucleotídeo modificado onde a função da proteína ligada é significativamente modificada em comparação com uma célula do tipo selvagem de tal modo que o rendimento, produção e/ou eficiência de produção de um ou mais produtos de fermentação como Vitamina C é aperfeiçoada. A célula hospedeira pode ser selecionada de um microorganismo capaz de produzir diretamente um ou mais

produtos de fermentação como, por exemplo, Vitamina C a partir de uma fonte de carbono dada, em particular, *Gluconobacter oxydans*, preferivelmente *G. oxydans* IFO 3293.

Uma "célula transformada" ou "célula recombinante" é uma célula na qual (ou em um ancestral da qual) foi introduzido, por intermédio de técnicas de DNA recombinante, um ácido nucléico de acordo com a invenção, ou onde a atividade da proteína de SMS 27 foi diminuída ou abolida. Células hospedeiras apropriadas incluem células de microorganismos capazes de produzir um produto de fermentação dado, por exemplo, converter uma fonte de carbono dada diretamente em vitamina C. Particularmente, essas incluem cepas dos gêneros *Pseudomonas*, *Pantoea*, *Escherichia*, *Carynebacterium*, *ketogulonicigenium* e bactérias de ácido acético como por exemplo, *Gluconobacter*, *Acetobacter* ou *Gluconoacetobacter*, preferivelmente *Acetobacter* sp., *Acetobacter* aceti, *Gluconobacter frateurii*, *Gluconobacter cerinus*, *Gluconobacter thailandicus*, *Gluconobacter oxydans*, mais preferivelmente *G. oxydans*, mais preferivelmente *G. oxydans* IFO 3293.

Para melhorar a produção de Vitamina C e/ou 2-KGA de uma certa célula hospedeira recombinante, a expressão de gene SMS 27 pode ser inibida naquele organismo, por exemplo, por alvejar seqüências de nucleotídeo complementares à região reguladora de uma seqüência de nucleotídeo de SMS 27 (por exemplo, um promotor de SMS 27 e/ou intensificadores) para formar estruturas helicoidais triplas que evitam transcrição de um gene de SMS 27 em células alvo. Vide, genericamente, Helene, C. (1991) *Anticancer Drug Des.* 6 (6): 569-84; Helene, C. e outros

(1992) Ann. N.Y. Acad. Sci. 660: 27-36; e Maher, L.J.

(1992) Bioassays 14 (12): 807 - 15.

A inibição ou prevenção de expressão de gene também pode ser obtida pela modificação do gene de SMS 27, por exemplo, pela introdução de uma ou mais mutações no gene de SMS 27 onde a modificação leva a uma proteína de SMS 27 com uma função que é significativamente diminuída em comparação com a proteína do tipo selvagem.

Portanto, em uma outra modalidade, o polinucleotídeo contendo pelo menos uma mutação é derivado de um polinucleotídeo como representado pela SEQ ID NO: 1 ou equivalentes do mesmo.

Uma mutação, como utilizada aqui, pode ser qualquer mutação que leva a um polipeptídeo menos funcional ou instável, por exemplo produtos de gene de SMS 27 menos funcionais ou instáveis. Isso pode incluir, por exemplo, uma alteração no genoma de um microorganismo, que interfere na síntese de SMS 27 ou leva à expressão de uma proteína de SMS 27 com uma seqüência de aminoácido alterada cuja função comparada com a cópia do tipo selvagem tendo uma seqüência de aminoácido não alterada é total ou parcialmente destruída. A interferência pode ocorrer no nível de transcrição, translação ou pós-translação.

A alteração no genoma do microorganismo pode ser obtida, por exemplo, por substituição através de uma recombinação de cruzamento única ou dupla de uma seqüência de DNA do tipo selvagem por uma seqüência de DNA contendo a alteração.

Para seleção conveniente de transformantes do microorganismo com a alteração em seu genoma a alteração

pode, por exemplo, ser uma seqüência de DNA codificando um marcador de resistência a antibiótico ou um gene complementando uma auxotrofia possível do microorganismo. Mutações incluem, porém não são limitados a, mutações de inserção de deleção. Um exemplo de tal alteração inclui uma ruptura de gene, isto é, uma perturbação de um gene de tal modo que o produto que é normalmente produzido a partir desse gene não é produzido em uma forma funcional. Isso poderia ser devido a uma deleção completa, uma deleção e inserção de um marcador seletivo, uma inserção de um marcador seletivo, uma mutação de deslocamento de quadro, uma deleção em quadro, ou uma mutação de ponto que leva à terminação prematura. Em alguns desses casos o mRNA inteiro para o gene está ausente, em outros a quantidade de mRNA produzido varia. Em todos os casos o polipeptídeo codificado pelo gene não é produzido em uma forma funcional, quer ausente ou em uma forma mudada, como por exemplo, uma proteína tendo atividade reduzida como definido aqui.

Uma alteração no genoma do microorganismo levando a um polipeptídeo menos funcional ou não funcional também pode ser obtida por realizar mutagênese aleatoriamente o genoma do microorganismo utilizando, por exemplo, mutágenos químicos, radiação ou elementos transponíveis e selecionar ou classificar em relação a mutantes que são produtores melhores ou mais eficientes de um ou mais produtos de fermentação. Métodos padrão para classificação e seleção são conhecidos por uma pessoa versada na técnica.

Em uma modalidade específica, é desejável nocautear o gene de SMS 27 da presente invenção, isto é, onde sua

expressão de gene é artificialmente suprimida para melhorar o rendimento, produtividade, e/ou eficiência de produção do produto de fermentação quando introduzido em uma célula hospedeira apropriada. Os métodos de fornecer nocautes bem como microorganismos que contêm esses genes suprimidos são bem conhecidos na técnica. A supressão do gene de SMS 27 endógeno pode ser induzida por deleção de pelo menos uma parte do gene ou a região reguladora do mesmo. Como utilizado aqui, "supressão da expressão de gene" inclui supressão completa e parcial, bem como supressão sob condições específicas e também supressão da expressão de qualquer um dos dois alelos.

Para criar um microorganismo de nocaute no qual a expressão do gene de SMS 27 é artificialmente suprimida, primeiramente o gene de SMS 27 pode ser clonado e então um vetor para recombinação homóloga pode ser construído utilizando o gene para inativar o gene de SMS 27 endógeno no microorganismo alvo. O vetor para recombinação homóloga contém então uma sequência de ácido nucléico projetada para inativar o gene de SMS 27 endógeno no microorganismo alvo. Tal ácido nucléico pode ser, por exemplo, uma sequência de ácido nucléico do gene de SMS 27 ou a região reguladora do mesmo, como a região de flanquear existente do gene a ser inativado (em cis), ou existindo separadamente (em trans), contendo pelo menos uma deleção parcial, ou alternativamente pode ser uma sequência de ácido nucléico do gene de SMS 27 ou a região reguladora do mesmo contendo outros genes. Um gene que também pode funcionar como um marcador é preferivelmente selecionado como o gene a ser inserido no gene de SMS 27 ou a região reguladora do mesmo.

Os genes de inserto a serem utilizados incluem, por exemplo, genes de resistência à droga como definido acima. Não há limitação específica sobre a posição onde os genes podem ser inseridos no gene de SMS 27, desde que a inserção
5 naquela posição resulte na supressão da expressão do gene de SMS 27 endógeno no alvo. Para evitar efeitos polares da inserção, deleções silenciosas em quadro podem ser introduzidas utilizando, por exemplo, o sistema sacB ou PCR de homologia de flanquear longa. Essas técnicas são bem
10 conhecidas pelas pessoas versadas na arte.

As estratégias de mutagênese acima mencionadas para proteínas de SMS 27 podem resultar em rendimentos aumentados de um composto desejado em particular, vitamina C e/ou 2-KGA. Essa lista não pretende ser limitadora;
15 variações nessas estratégias de mutagênese serão prontamente evidentes para uma pessoa com conhecimentos comuns na técnica. Por esses mecanismos, moléculas de proteína e ácido nucléico da invenção podem ser utilizadas para gerar microorganismos como *Gluconobacter oxydans* ou
20 cepas relacionadas de bactérias expressando moléculas de proteína e ácido nucléico de SMS 27 mudadas de tal modo que o rendimento, produtividade, e/ou eficiência de produção de um composto desejado como vitamina C e/ou 2-KGA são aperfeiçoados.

25 Com relação ao processo acima utilizando um microorganismo, em um aspecto, o processo da presente invenção leva a rendimentos de Vitamina C que são pelo menos aproximadamente maiores do que 5,7 g/l, como 10 g/l, 20 g/l, 50 g/l, 100 g/l, 200 g/l, 300 g/l, 400 g/l ou maior
30 do que 600 g/l. Em uma modalidade, o rendimento da Vitamina

C produzida pelo processo da presente invenção está na faixa de aproximadamente mais do que 5,7 a aproximadamente 600 g/l. O rendimento da Vitamina C se refere à concentração de vitamina C no fluxo de colheita saindo diretamente do recipiente de produção, isto é, o sobrenadante livre de células compreendendo a Vitamina C.

Em um aspecto da invenção, microorganismos (em particular dos gêneros de *Gluconobacter*, *Gluconacetobacter* e *Acetobacter*) são fornecidos que são capazes de produzir diretamente a Vitamina C a partir de uma fonte de carbono apropriada como D-sorbitol e/ou L-sorbose. Quando medido, por exemplo, em um método de célula em repouso após um período de incubação de 20 horas, verificou-se que esses organismos são capazes de produzir a Vitamina C diretamente do D-sorbitol ou L-sorbose, mesmo até um nível de 280 mg/l e 670 mg/l respectivamente. Em outro aspecto da invenção, um microorganismo é fornecido capaz de produzir diretamente a Vitamina C em quantidades de 300 mg/l ao começar a partir de D-sorbitol ou mais ou 800 mg/l ou mais ao começar a partir de L-sorbose, respectivamente quando por exemplo medido em um método de célula em repouso após um período de incubação de 20 horas. Isso pode ser obtido pela diminuição ou abolição da atividade de um polipeptídeo de SMS, preferivelmente um polipeptídeo de SMS 27. O rendimento da Vitamina C produzida a partir de D-sorbitol pode ser até mesmo tão elevado quanto 400, 600, 1000 mg/l ou mesmo exceder 1,5, 2, 4, 10, 20, 50, g/l. O rendimento da Vitamina C produzida a partir de L-sorbose pode ser até mesmo tão elevado quanto 1000 mg/l ou mesmo exceder 1,5, 2, 4, 10, 20, 50 g/l. Preferivelmente, essas quantidades de

Vitamina C podem ser obtidas quando medidas pelo método de célula em repouso após um período de incubação de 20 horas.

Como utilizado aqui, a medição em um "método de célula em repouso" compreende (i) cultivar as células por intermédio de qualquer método bem conhecido pela pessoa versada na técnica, (ii) colher as células a partir do caldo de crescimento, e (iii) incubar as células colhidas em um meio contendo o substrato que deve ser convertido no produto desejado, por exemplo, Vitamina C, sob condições onde as células não crescem mais, isto é, não há aumento na quantidade de biomassa durante essa etapa denominada conversão. Uma descrição mais geral do método de célula em repouso é descrita, por exemplo, em WO 2005/017159 e nos parágrafos a seguir.

Em um aspecto da invenção, microorganismos (em particular a partir dos gêneros de *Gluconobacter*, *Gluconacetobacter* e *Acetobacter*) são fornecidos que são capazes de produzir diretamente 2-KGA a partir de uma fonte de carbono apropriada como D-sorbitol e/ou L-sorbose. Quando medido, por exemplo, pelo método como do Exemplo 4 verificou-se que esses organismos são capazes de produzir 2-KGA diretamente a partir de D-sorbitol ou L-sorbose em quantidades de aproximadamente 0,5 a 0,7 g/l. Em outro aspecto da invenção, um microorganismo é fornecido capaz de produzir diretamente 2-KGA em quantidades de aproximadamente 7, 8, 9, 10 g/l ou mais ou mesmo aproximadamente 50, 60, 70, 80, 90, 100 g/l ou mais ao iniciar a partir de L-sorbose. Isso pode ser obtido por diminuir ou mesmo abolir a atividade de um polipeptídeo SMS, preferivelmente um polipeptídeo SMS 27 em *G. oxydans*

IFO 3293.

O microorganismo recombinante contendo, por exemplo, um gene de SMS 27 modificado e que é capaz de produzir o produto de fermentação em rendimento, produtividade, e/ou
5 eficiência significativamente mais elevados, pode ser cultivado em um meio aquoso suplementado com nutrientes apropriados sob condições aeróbicas como descrito acima.

As moléculas de ácido nucléico, polipeptídeos, vetores, iniciadores e microorganismos recombinantes
10 descritos aqui, podem ser utilizados em um ou mais dos seguintes métodos: identificação de *Gluconobacter oxydans* e organismos relacionados; mapeamento de genomas de organismos relacionados a *Gluconobacter oxydans*; identificação e localização de seqüências de *Gluconobacter*
15 *oxydans*, de interesse; estudos de evolução; determinação de regiões de proteína de SMS 27 exigidas para função; modulação de uma atividade ou função de proteína de SMS 27; modulação da atividade de uma via de SMS; e modulação de produção celular de um composto desejado, como a Vitamina C
20 e/ou 2-KGA.

A invenção provê métodos para classificar moléculas que modulam a atividade de uma proteína de SMS 27, por interação com a própria proteína ou um substrato ou parceiro de ligação da proteína de SMS 27, ou por modulação
25 da transcrição ou translação de uma molécula de ácido nucléico de SMS 27 da invenção. Em tais métodos, um microorganismo expressando uma ou mais proteínas de SMS 27 da invenção é contatado com um ou mais dos compostos de teste, e o efeito de cada composto de teste sobre a
30 atividade ou nível de expressão da proteína de SMS 27 é

avaliado.

Em geral, a atividade biológica, enzimática ou outra de proteínas de SMS pode ser medida por métodos bem conhecidos por uma pessoa versada, como, por exemplo, pela
5 incubação de uma fração de célula contendo a proteína de SMS na presença de seu substrato, aceitador(es) ou doador(es) de elétrons incluindo metossulfato de fenazina (PMS), diclorofenol-indofenol (DCIP), NAD, NADH, NADP, NADPH, cujo consumo pode ser direta ou indiretamente medido
10 por métodos fotométricos, colorimétricos ou fluorimétricos, e outros componentes inorgânicos que poderiam ser relevantes para o desenvolvimento da atividade. Desse modo, por exemplo, a atividade de desidrogenase de D-sorbitol ligado por membrana pode ser medida em um ensaio onde
15 frações de membrana contendo essa enzima são incubadas na presença de tampão de fosfato em pH 6, D-sorbitol e os aceitadores de elétron artificiais DCIP e PMS. A taxa de consumo de DCIP pode ser medida em 600 nm, e é diretamente proporcional à atividade de desidrogenase de D-sorbitol
20 presente na fração de membrana.

A atividade biológica, enzimática ou outra das proteínas SMS, em particular a proteína SMS 27, pode ser medida por métodos bem conhecidos por uma pessoa versada na técnica, como por exemplo, por determinar a expressão de
25 genes conhecidos como estando sob o controle da proteína SMS 27 por métodos conhecidos por aqueles versados na técnica, como por exemplo Northern Blot, análise de fusão de transcrição, análise de microconjunto, etc. A ligação da proteína SMS 27 a regiões promotoras alvo pode ser
30 demonstrada por experimentos de pegada de DNA, ensaios de

deslocamento de gel e similares.

Pode ser evidente a partir da descrição acima que o produto de fermentação dos métodos de acordo com a invenção pode não ser limitado somente à Vitamina C. O "composto
5 desejado" ou "produto de fermentação" como utilizado aqui pode ser qualquer produto natural de *Gluconobacter oxydans*, que inclui os produtos finais e intermediários de vias de biossíntese, como por exemplo, L-sorbose, L-sorbose, D-gluconato, 2-ceto-D-gluconato, 5-ceto-D-gluconato, 2,5-
10 diceto-D-gluconato e 2-ceto-L-gulonato (2-KGA), em particular a geração biossintética de Vitamina C.

Desse modo, a presente invenção é dirigida ao uso de um polinucleotídeo, polipeptídeo, vetor, iniciador e microorganismo recombinante como descrito aqui na produção
15 de Vitamina C e/ou 2-KGA, isto é, a conversão direta de uma fonte de carbono em Vitamina C e/ou 2-KGA. Em uma modalidade preferida, um polinucleotídeo, polipeptídeo, vetor e microorganismo recombinante modificado, como descrito aqui, é utilizado para melhorar o rendimento,
20 produtividade e/ou eficiência da produção de Vitamina C e/ou 2-KGA.

Os termos "produção" ou "produtividade" são reconhecidos na técnica e incluem a concentração do produto de fermentação (por exemplo, Vitamina C e/ou 2-KGA) formado
25 em um tempo dado e um dado volume de fermentação (por exemplo, kg de produto por hora por litro). O termo "eficiência de produção" inclui o tempo necessário para um que nível específico de produção seja obtido (por exemplo, quanto tempo demora para que a célula obtenha uma taxa
30 específica de saída de um produto de fermentação). O termo

"rendimento" é reconhecido na técnica e inclui a eficiência da conversão da fonte de carbono no produto (isto é, Vitamina C). Isso é genericamente escrito como, por exemplo, kg de produto por kg de fonte de carbono. Por 5 "aumentar o rendimento e/ou produção/produtividade" do composto quer se dizer que a quantidade de moléculas recuperadas, ou de moléculas recuperadas úteis daquele composto em uma dada quantidade de cultura durante uma quantidade dada de tempo é aumentada. Os termos 10 "biossíntese" ou uma "via biossintética" são reconhecidos na técnica e incluem a síntese de um composto, preferivelmente um composto orgânico, por uma célula dos compostos intermediários no que pode ser um processo altamente regulado e de multi-etapas. A linguagem 15 "metabolismo" é reconhecida na técnica e inclui a totalidade das reações bioquímicas que ocorrem em um organismo. O metabolismo de um composto específico, então, (por exemplo, o metabolismo de um aminoácido como glicina) compreende as vias biossintética, de modificação e 20 degradação gerais na célula relacionada a esse composto. A linguagem "transporte" ou "importar" é reconhecida na técnica e inclui o movimento facilitado de uma ou mais moléculas através de uma membrana celular através da qual a molécula seria de outro modo incapaz de passar ou ser 25 passada ineficientemente.

A Vitamina C como utilizado aqui, pode ser qualquer forma química de ácido L-ascórbico encontrado em soluções aquosas, como por exemplo, não dissociadas, em sua forma de ácido livre ou dissociada como um ânion. A forma de sal 30 solubilizado de ácido L-ascórbico pode ser caracterizada

como o ânion na presença de qualquer tipo de cátions normalmente encontrados em sobrenadantes de fermentação, como por exemplo, potássio, sódio, amônio ou cálcio. Também podem ser incluídos cristais isolados da forma de ácido livre de ácido L-ascórbico. Por outro lado, cristais isolados de uma forma de sal de ácido L-ascórbico são denominados pelo nome de seu sal correspondente, isto é, ascorbato de sódio, ascorbato de potássio, ascorbato de cálcio e similares.

10 Em uma modalidade preferida, a presente invenção é relacionada a um processo para a produção de Vitamina C onde um nucleotídeo de acordo com a invenção ou uma sequência de polinucleotídeo modificado, como descrito acima, é introduzido em um microorganismo apropriado, o
15 microorganismo recombinante é cultivado sob condições que permitem a produção de Vitamina C em produtividade elevada, rendimento e/ou eficiência, o produto de fermentação produzido é isolado do meio de cultura e opcionalmente adicionalmente purificado.

20 Essa invenção é adicionalmente ilustrada pelos exemplos a seguir os quais não devem ser interpretados como limitadores. O teor de todas as referências, pedidos de patente, patentes e pedidos de patente publicados, citados em todo esse pedido são incorporados pela presente a título
25 de referência.

Exemplos

Exemplo 1: Preparação de DNA cromossômico e amplificação de fragmento de DNA por PCR

DNA cromossômico de *Gluconobacter oxydans* IFO 3293 foi
30 preparado a partir das células cultivadas a 30°C por 1 dia

em meio líquido de caldo de manitol (MB) consistindo em 25 g/l de manitol, 5 g/l de extrato de levedura (Difco), e 3 g/l de Bactopeptone (Difco) pelo método descrito por Sambrook e outros (1989) "Molecular Cloning: A Laboratory Manual/Second edition", Cold Spring Harbor Laboratory Press.

Um fragmento de DNA foi preparado por PCR com o DNA cromossômico preparado acima e um conjunto de iniciadores, Pf (SEQ ID NO:3) e Pr (SEQ ID NO:4). Para a reação, o Kit de PCR Expand High Fidelity (Roche Diagnostics) e 10 ng do DNA cromossômico foram utilizados em volume total de 100 µl de acordo com a instrução do fornecedor, para ter o produto de PCR contendo sequência de DNA de SMS 27 (SEQ ID NO: 1). O produto de PCR foi recuperado a partir da reação e sua
15 sequência correta confirmada.

Exemplo 2: Ruptura do gene de SMS 27 em *G. oxydans* IFO 3293

Para manipular um mutante de nocaute do gene de SMS 27, PCR de Long-Flanking Homology (LFH) foi utilizado para
20 manipular uma mutação de inserção-deleção limpa. Em primeiro lugar, as regiões a montante e a jusante flanqueando o gene SMS 27 foram amplificadas por PCR utilizando os respectivos pares de iniciador SMS 27KmLFH+1 (SEQ ID NO: 7) / SMS 27LFH-1 (SEQ ID NO:8). O DNA genômico
25 *G. oxydans* IFO 3293 foi utilizado como padrão e as condições de reação consistiam em 30 ciclos de desnaturação a 95°C por 1 min, com recozimento a 50°C por 1 min. e extensão a 72°C por 1,5 min. Nos dois casos, o polimerase de DNA Herculase (Stratagene) foi utilizado para minimizar
30 erros gerados por PCR. O cassete de resistência à

canamicina foi amplificado utilizando DNA de plasmídeo pUC4K (Amersham Bioscience, no. de acesso XO6404) como um padrão e par de iniciadores SMS 27Km-1 (SEQ ID NO: 9) / SMS 27Km+1 (SEQ ID NO: 10) para gerar um fragmento de 1,3-kb.

5 As condições de reação foram como acima. Os três produtos foram purificados por gel, misturados e utilizados na reação de PCR de segundo round com os iniciadores de flanquear SMS 27LFH+1 / SMS 27LFH-1 para gerar um produto de 2,6-kb. As condições de reação para a reação de segundo
10 round consistiam em 94°C, 2 min., a seguir 10 ciclos de [94°C, 30 s, 63°C, 30 s, 68°C, 6 min.], seguido por 20 ciclos de [94°C, 30 s, 63°C, 30 s, 68°C, 6 min. com um período adicional de 20 s por ciclo] e uma extensão final de 68°C por 10 min.

15 O produto de PCR foi clonado no vetor pGEM-TEasy (Promega) para fornecer plasmídeo pGEM-SMS 27. Esse plasmídeo foi transformado em células de *G. oxydans* IFO 3293 competentes selecionando transformantes em meio de agar MB contendo canamicina em uma concentração final de 50
20 $\mu\text{g ml}^{-1}$, resultando em cepa mutante *G. oxydans* IFO 3293-SMS 27::Km. PCR utilizando os iniciadores de flanquear SMS 27LFH+1 / SMS 27LFH-1 foi usado para confirmar que a mutação SMS 27::Km tinha integrado através de um cruzamento duplo.

25 Exemplo 3: Produção de Vitamina C a partir de D-sorbitol utilizando células em repouso

As células de *G. oxydans* IFO 3293 e *G. oxydans* IFO 3293-SMS 27::Km1 foram revestidas primeiramente em meio MB por três dias. As células foram raspadas dessas placas e
30 espalhadas sobre meio No. 3BD contendo 7% de sorbitol e

cultivadas por 3 dias a 30°C. Essas células foram utilizadas em reações de células em repouso com 2% de sorbitol como substrato e uma densidade de células de OD₆₀₀=10.

5 Exemplo 4: Produção de 2-KGA a partir de D-sorbitol utilizando células em repouso

De acordo com a montagem experimental do exemplo 3, as células das cepas *G. oxydans* IFO 3293 e *G. oxydans* IFO 3293-SMS 27::Km, foram testadas em relação a sua capacidade para
10 produzir 2-KGA a partir de D-sorbitol.

Após 20 h de tempo de incubação, *G. oxydans* IFO 3293 produziu 700 mg/l de 2-KGA a partir de D-sorbitol. Em comparação, a cepa *G. oxydans* IFO 3293-SMS 27::Km produziu mais de 7 g/l de 2-KGA.

15 Exemplo 5: Presença do gene de SMS 27 e equivalentes em outros organismos

A presença da SEQ ID NO:1 e/ou equivalentes mostrando similaridade/identidade com a SEQ ID NO: 1 em outros organismos do que aqueles revelados anteriormente, por
20 exemplo, organismos como mencionado na Tabela 1, pode ser determinada por um experimento de hibridização de DNA simples.

As cepas de *Acetobacter aceti* subsp. *Xylinum* IFO 13693 e IFO 13773 são cultivadas a 27°C por 3 dias no meio no.
25 350 contendo 5 g/l de Bactopeptone (Difco), 5 g/l de extrato de levedura (Difco), 5 g/l de glicose, 5 g/l de manitol, 1 g/l de MgSO₄.7H₂O, 5 ml/l de etanol, e 15 g/l de agar. Todas as outras cepas de *Acetobacter*, *Gluconacetobacter* e todas *Gluconobacter* são cultivadas a
30 27°C por 3 dias em meio de agar de caldo de manitol (MB)

contendo 25 g/l de manitol, 5 g/l de extrato de levedura (Difco), 3 g/l de Bactopeptone (Difco) e 18 g/l de agar (Difco). E- coli K-12 é cultivado no meio de agar de Caldo Luria. As outras cepas/células são cultivadas no meio recomendado por fornecedores ou de acordo com métodos conhecidos na técnica. O DNA genômico é extraído como descrito por, por exemplo, Sambrook e outros, 1989, "Molecular Cloning: A laboratory Manual/second edition", Cold Spring Harbor Laboratory Press a partir de um organismo apropriado, como por exemplo mencionado na Tabela 1.

Preparações de DNA genômico são digeridas com enzimas de restrição como EcoRI e HindIII, e 1 µg dos fragmentos de DNA é separado por eletroforese de gel agarose (1% de agarose). O gel é tratado com 0,25 N HCl por 15 min. e então 0,5 N NaOH por 30 min., e então manchado sobre nitrocelulose ou uma membrana de náilon com Vacuum Blotter modelo 785 (BIO-RAD Laboratories AG, Suíça) de acordo com a instrução do fornecedor. A mancha resultante é então colocada em contato/hibridizado com uma solução onde a sonda, como, por exemplo, um fragmento de DNA com a sequência SEQ ID NO:1 ou um fragmento de DNA contendo a parte ou toda a sequência de SEQ ID NO: 1 para detectar fragmento(s) de DNA positivo(s) a partir de um organismo de teste. Uma sonda rotulada DIG, por exemplo, SEQ ID NO: 1, pode ser preparada de acordo com o Exemplo 1 utilizando o kit de rotulação PCR-DIG (Roche Diagnostics) e um conjunto de iniciadores, SEQ ID NO: 3 e SEQ ID NO: 4. Um resultado dessa mancha é representado na Tabela 1.

A hibridização pode ser executada sob condições

rigorosas ou altamente rigorosas. Um exemplo não limitador, preferido dessas condições é a hibridização em 6x de cloreto de sódio/citrato de sódio (SSC) em aproximadamente 45°C, seguido por uma ou mais lavagens em 1x SSC, 0,1% SDS a 50°C, preferivelmente a 55°C, mais preferivelmente a 60°C e ainda mais preferivelmente a 65°C. Condições altamente rigorosas incluem, por exemplo, 2h a 4 dias de incubação em 42°C em uma solução como solução DigEasyHyb (Roche Diagnostics GmbH) com ou sem 100 µg/ml de DNA de esperma de salmão, ou uma solução compreendendo formamida a 50%, 5x SSC (150 mM NaCl, 15 mM de citrato trissódico), 0,02% de sulfato de dodecil de sódio, 0,1% de N-lauroíl sarcosina, e 2% de reagente de bloqueio (Roche Diagnostics GmbH) seguido por lavagem dos filtros duas vezes por 5 a 15 minutos em 2x SSC e 0,1% SDS em temperatura ambiente e então lavagem duas vezes por 15-30 minutos em 0,5x SSC e 0,1% SDS ou 0,1x SSC e 0,1% SDS a 65-68°C. Para detectar fragmentos de DNA com identidade inferior ao DNA de sonda, as etapas de lavagem finais podem ser feitas em temperaturas mais baixas como 50-65°C e por tempo de lavagem mais curto como 1-15 min.

Os genes correspondendo aos sinais positivos nos respectivos organismos mostrados na Tabela 1 podem ser clonados por um método de PCR bem conhecido na técnica utilizando DNA genômico de tal organismo juntamente com um conjunto de iniciadores apropriado, como, por exemplo, SEQ ID NO: 3 e SEQ ID NO: 4 sob condições como descrito no Exemplo 1 ou como a seguir: 5 a 100 ng de DNA genômico são utilizados por reação (volume total 50 µl). O sistema de PCR Expand High Fidelity (Roche Diagnostics) pode ser utilizado com condições de reação consistindo em 94°C por 2

min.; 30 ciclos de (i) etapa de desnaturação a 94°C por 15 s., (ii) etapa de recozimento a 60°C por 30 s., (iii) etapa de síntese a 72°C por 0,5 a 5 min. dependendo do comprimento de DNA alvo (1 min/1kb); extensão a 72°C por 7 min. Alternativamente, pode-se executar um PCR com iniciadores de degenerar, que podem ser sintetizados com base na SEQ ID NO: 2 ou seqüências de aminoácidos como seqüências de consenso selecionadas por alinhamento de várias seqüências de aminoácidos obtidas por um programa de busca de seqüência como BLASTP (ou BLASTX quando a seqüência de nucleotídeo é utilizada como uma "seqüência de consulta") para encontrar proteínas tendo uma similaridade com a proteína da SEQ ID NO: 2. Para PCR utilizando iniciadores de degenerar, a temperatura da segunda etapa de recozimento (vide acima) pode ser diminuída para 55°C ou mesmo para 50-45°C. Um resultado de um tal experimento é mostrado na Tabela 1.

As amostras das reações de pCR são separadas por eletroforese de agarose gel e as faixas são visualizadas com um transiluminador após colorir com, por exemplo, brometo de etídio, isolado do gel e a seqüência correta é confirmada.

As seqüências de consenso mencionadas acima poderiam ser seqüências de aminoácido que pertencem a certas categorias de vários bancos de dados de família/domínio de proteína como PROSITE (banco de dados de domínios e famílias de proteína), COGs (Cluster de grupos ortólogos), CDD (Bancos de dados de domínio conservado), pfam (coleção grande de múltiplos alinhamentos de seqüência e modelos Markov ocultos cobrindo muitas famílias e domínios de

proteína comuns). Após se poder selecionar certa proteína com função idêntica/similar à proteína da presente invenção a partir das proteínas contendo domínio ou família de tais bancos de dados, DNA correspondente codificando a proteína
5 pode ser amplificado por PCR utilizando a seqüência de proteína ou sua seqüência de nucleotídeos quando disponível em bancos de dados públicos.

Exemplo 6: Ruptura do gene de SMS 27 e equivalentes em outros organismos para produção de Vitamina C e/ou 2-KGA

10 Para melhorar a produção de Vitamina C em um microorganismo apropriado que é capaz de produzir diretamente Vitamina C a partir de um dado substrato, o gene de SMS 27 e equivalentes como, por exemplo, um produto de PCR obtido no Exemplo 4, mencionado doravante como gene X,
15 pode ser rompido de acordo com o gene de SMS 27 em *G. oxydans* IFO 3293 (vide o Exemplo 2) para gerar um mutante de nocaute contendo gene equivalente de SMS 27:Km. Cepas de hospedeiro apropriadas para geração desses mutantes de nocaute podem ser selecionadas de, por exemplo, cepas de
20 *Gluconobacter* listadas na Tabela 1, em particular, por exemplo, *G. oxydans* IFO 3292, *G. oxydans* ATCC 621H, *G. oxydans* IFO 12528, *G. oxydans* IFO 3291, *G. oxydans* IFO 3255, *G. oxydans* IFO 3244, *G. cerinus* IFO 3266, *G. frateurii* IFO 3260, *G. oxydans* IFO 3287, *Acetobacter aceti*
25 subsp. *Orleanus* IFO 3259, *Acetobacter aceti* subsp. *Xylinum* IFO 13693, *Acetobacter aceti* subsp. *Xylinum* IFO 13773 e *Acetobacter* sp. ATCC 15164 e seus derivados contendo genes envolvidos nas vias de produção de vitamina C e as regiões adjacentes onde o gene SMS 27 ou seu equivalente poderia
30 ser localizado.

O mutante de nocaute, como, por exemplo, mutante de nocaute *G. oxydans* IFO 3293-SMS 27 gene equivalente: Km pode ser gerado como a seguir: o produto de PCR obtido a partir de *G. oxydans* IFO 3293 descrito no Exemplo 5 é clonado em um vetor de *E. coli* pCR2.1-TOPO e utilizado para transformar *E. coli* TG1 para ter um transformante de Apr contendo gene pCR2.1 X. A seguir, o cassete Kmr isolado de pUC-4K (Amersham Bioscience, no. de acesso X06404) é inserido em um do sítio de restrição do gene alvo com ligase e o produto de ligação resultante é utilizado para transformar *E. coli* TG1 a fim de ter transformante Apr Kmr contendo pCR2.1-gene X::Km. O plasmídeo pCR2.1-gene X:Km preparado a partir do transformante é digerido por duas enzimas de restrição selecionadas do sítio de clonagem múltipla da parte de vetor para isolar um fragmento de DNA contendo gene X::Km. O fragmento de DNA resultante é utilizado para transformar a cepa hospedeira contendo o gene equivalente de SMS 27 por eletroporação para ter o meio de ruptura de gene contendo gene equivalente SMS 27::Km.

Modificações adicionais incluindo genes envolvidos na conversão de D-sorbitol, L-sorbose e/ou L-sorbose em Vitamina C nas cepas podem ser geradas para melhorar a produção de vitamina C por tais cepas.

A produção de vitamina C utilizando as células do mutante de nocaute, por exemplo *G. oxydans* IFO 3292-gene equivalente SMS 27::Km, e a cepa do tipo selvagem correspondente, por exemplo, *G. oxydans* IFO 3293, são executadas de acordo com o Exemplo 3.

A produção de 2-KGA utilizando as células do mutante

de nocaute, por exemplo *G. oxydans* IFO 3292- gene equivalente SMS 27::Km, e a cepa do tipo selvagem correspondente, por exemplo *G. oxydans* IFO 3292, são executadas de acordo com o Exemplo 4.

- 5 Na reação de células em repouso com L-sorbose a 1% como substrato, a cepa mutante pode produzir pelo menos mais de 20% de Vitamina C e pelo menos mais de 10% de 2-KGA em comparação com a cepa do tipo selvagem.

Tabela 1: Equivalentes do gene de SMS 27 em outros

10 organismos

Cepa	Sinal 1	Sinal 2	Sinal 3
<i>G. oxydans</i> IFO 3293	+	+	+
<i>G. oxydans</i> DSM 17078	+	+	+
<i>G. oxydans</i> IDO 3292	+	+	+
<i>G. oxydans</i> ATCC 621H	+	+	+
<i>G. oxydans</i> IFO 12528	+	+	+
<i>G. oxydans</i> G 624	+	+	+
<i>G. oxydans</i> T-100	+	+	+
<i>G. oxydans</i> IFO 3291	+	+	+
<i>G. oxydans</i> IFO 3255	+	+	+
<i>G. oxydans</i> ATCC 9937	+	+	+
<i>G. oxydans</i> IFO 3244	+	+	+
<i>G. cerynus</i> IFO 3266	+	+	+
<i>G. frateurii</i> IFO 3260	+	+	+
<i>G. oxydans</i> IFO 3287	+	+	+
<i>Acetobacter aceti</i> subsp. <i>Orleanus</i> IFO 3259	+	+	+
<i>Acetobacter aceti</i> subsp. <i>Xylinum</i> IFO 13693	+	+	+
<i>Acetobacter aceti</i> subsp.	+	+	+

<i>Xylinum</i> IFO 13773			
<i>Acetobacter</i> sp. ATCC 15164	+	+	+
<i>G. thailandicus</i> NBRC 100600	+	+	+
<i>Gluconacetobacter liquefaciens</i> ATCC 14835	+	+	+
<i>Gluconacetobacter polyoxogenes</i> NBI 1028	+	+	+
<i>Gluconacetobacter</i> <i>diazotrophicus</i> ATCC 49037	+	+	+
<i>Gluconacetobacter europaeus</i> DSM 6160	+	+	+
<i>Acetobacter aceti</i> 1023	+	-	+
<i>Acetobacter pasteurianus</i> NCI 1193	+	-	+
<i>E. coli</i>	-	-	-
<i>Saccharomyces cerevisiae</i>	-	-	-
<i>Aspergillus niger</i>	-	-	-
Camundongo	-	-	-

Sinal 1: Detecção de DNA em uma mancha com DNA genômico de cepas diferentes e SEQ ID NO:1 como sonda rotulada. Sinal 2: detecção de DNA de cepas diferentes em uma reação de PCR utilizando par de iniciadores SEQ ID NO: 3 e SEQ ID NO: 4. Sinal 3: detecção de DNA de cepas diferentes em uma reação de PCR utilizando iniciadores de degenerar. Para maior explicação, consulte o texto.

LISTAGEM DE SEQUÊNCIA

<110> DSM IP Assets B.V.

5 <120> Novo Gene SMS 27

<130> 24998

<160> 16

10

<170> PatentIn versão 3.2

<210> 1

<211> 495

15 <212> DNA

<213> Gluconobacter oxydans IFO 3293

<400> 1

atgccccttg agaaaagcag agcgtcctcc tctacgaaa agcgtgtcag ggacgctgag

20 60

gccacgaaat tacgcattct cgaagcggcg aagaaggagt ttgcccgaaa cgggcttgaa

120

ggcgcgcggg tcgatacgat cgcctcgcg gccaaaggcca acaagaggat gctctaccac

180

25 tatttcaaaa gtaaggacgg ccttttccgc accgtcctgg agcgggaata tctcaacatc

240

cgggaggaag agaccagact ggatctggag aacctcccc cccaggacag ccctcgaaac

300

gctcgtccgt ttcacatggg aatactatat cgccaatccg gagttcatca cccttgtaga

30 360

cagcgagaac ctctgtcagg cgcgccatat ccggaattca cccacactga agacgatcag

420

tcagaaattc gtccgccgca tccagtcctt tctggatcgg ggcgttcggg aaaatgtctt

480

5 ccgcgatggc attga

495

<210> 2

<211> 164

10 <212> PRT

<213> Gluconobacter oxydans IFO 3293

<400> 2

Met Pro Leu Glu Lys Ser Arg Ala Ser Ser Ser Thr Lys Lys Arg Val

15 1 5 10 15

Arg Asp Ala Glu Ala Thr Lys Leu Arg Ile Leu Glu Ala Ala Lys Lys

20 25 30

Glu Phe Ala Arg Asn Gly Leu Glu Gly Ala Arg Val Asp Thr Ile Ala

35 40 45

20 Leu Asp Ala Lys Ala Asn Lys Arg Met Leu Tyr His Tyr Phe Lys Ser

50 55 60

Lys Asp Gly Leu Phe Arg Thr Val Leu Glu Arg Glu Tyr Leu Asn Ile

65 70 75 80

Arg Glu Glu Glu Thr Arg Leu Asp Leu Glu Asn Leu Pro Pro Gln Asp

25 85 90 95

Ser Pro Arg Asn Ala Arg Pro Phe His Met Gly Ile Leu Tyr Arg Gln

100 105 110

Ser Gly Val His His Pro Cys Glu Gln Arg Glu Pro Leu Ser Gly Arg

115 120 125

30 Pro Tyr Pro Glu Phe Thr His Thr Glu Asp Asp Gln Ser Glu Ile Arg

	130	135	140
	Pro Pro His Pro Val Pro Ser Gly Ser Gly Arg Ser Gly Lys Cys Leu		
	145	150	155 160
	Pro Arg Trp His		

5

<210> 3

<211> 20

<212> DNA

<213> Artificial

10

<220>

<223> primer

<400> 3

15 atgccccttg agaaaagcag

20

<210> 4

<211> 20

20 <212> DNA

<213> Artificial

<220>

<223> primer

25

<400> 4

tcaatgccat cgcggaagac

20

30 <210> 5

<211> 22

<212> DNA

<213> Artificial

5 <220>

<223> primer

<400> 5

gggtgacggg aacatcatga gc

10 22

<210> 6

<211> 43

<212> DNA

15 <213> Artificial

<220>

<223> primer

20 <400> 6

cacgaggcag acctcagcgc cgcgtaattt cgtggcctca gcg

43

<210> 7

25 <211> 47

<212> DNA

<213> Artificial

<220>

30 <223> primer

<400> 7

cagagatttt gagacacaac gtggcgatat gccgcctggt ctgcaag

47

5

<210> 8

<211> 21

<212> DNA

<213> Artificial

10

<220>

<223> primer

<400> 8

15 gggcggtcttc ctgcacatcg g

21

<210> 9

<211> 43

20 <212> DNA

<213> Artificial

<220>

<223> primer

25

<400> 9

cgctgaggcc acgaaattac gcggcgctga ggtctgcctc gtg

43

30 <210> 10

<211> 47

<212> DNA

<213> Artificial

5 <220>

<223> primer

<400> 10

cttgcagacc aggcggcata tcgccacggt gtgtctcaaa atctctg

10 47

<210> 11

<211> 1596

<212> DNA

15 <213> Gluconobacter oxydans IFO3293

<400> 11

atgacgagcg gttttgatta catcgttgtc ggtggcggtt cggetggctg tgttctcgca

60

20 gcccgccttt ccgaaaatcc ttccgtccgt gtctgcctca tcgaggcggg ccggcgggac

120

acgcatcccc tgatccatat gccggtcggg ttgcggaaga tgaccacggg gccgcatacc

180

tgggatcttc tgacggagcc gcagaaacat gcgaacaacc gccagatccc ctatgtgcag

25 240

ggccggattc tgggcggcgg atcgtccatc aacgcggaag tcttcacgcg gggacaccct

300

tccgatttcg accgctgggc ggcggaaggt gcggatggct ggagcttccg ggatgtccag

360

30 aagtacttca tccgttccga aggcaatgcc gtgttttcgg gcacctggca tggcacgaac

420
 gggccgctcg ggggtgtccaa cctcgcagat ccgaacccga ccagccgtgc cttcgtgcag
 480
 agctgtcagg aaatggggct gccctacaac cctgacttca atggcgcac gcaggaaggg
 5 540
 gctggcatct accagatgac catccggaac aaccggcgct gctcgacggc tgtgggggtat
 600
 ctgcgtccgg ccctggggcg gaagaacctg acggttgtga cgcgggcgct ggtcctgaag
 660
 10 atcgtcttca acgggacgcg ggcgacgggc gtgcagtata tcgccaacgg caccctgaat
 720
 accgccgaag cgagccagga aatcgttgtg acggccggag cgatcggaac gccgaagctg
 780
 atgatgctgt cgggcgtcgg gcctgccgcg catcttcgcg aaaatggat cccggtcgtg
 840
 caggatctgc cgggcgtggg cgagaacctt caggaccatt tcggtgtgga tatcgtagcc
 900
 gagctcaaga cggatgagag cttcgacaag taccggaaac tgcaactggat gctgtgggca
 960
 20 ggtcttgaat acaccatgtt cagatccggc cccgtcgcgt ccaacgtggt tgagggcggc
 1020
 gcgttctggt actcggaccc gtcacgggt gttcctgac tccagttcca ttttcttgcg
 1080
 ggggcagggg ctgaggcagg ggtgacgtcc gttcccaagg gcgcgtcggg gattacgctg
 1140
 25 aacagctatg tgctgcgtcc gaagtctcgc ggtaccgttc ggctgcgttc ggcagatcca
 1200
 aggggtcaatc cgatggtcga tcccaatttc cttggagacc cggccgacct tgagacgtct
 1260
 30 gcggaaggtg tgcggctgag ctacgagatg ttctcccagc cttccttgca gaagcacatc

1320

aaggaaacat gcttcttttag cggtaaacag ccgacgatgc agatgtatcg ggactatgcg

1380

cgggaacatg gccggacctc ctatcatccg acatgcacct gcaagatggg gcgggatgac

5 1440

atgtccgtcg tcgatccgcg tctgaagggt catggccttg agggcatcag gatctgtgac

1500

agctcgggtca tgccgtcgct gctcgggtcc aacaccaatg ccgcgacgat catgatcagt

1560

10 gagcgggcag cggatttcat tcaggggaac gcctga

1596

<210> 12

<211> 531

15 <212> PRT

<213> Gluconobacter oxydans IFO 3293

<400> 12

20 Met Thr Ser Gly Phe Asp Tyr Ile Val Val Gly Gly Gly Ser Ala Gly

1 5 10 15

Cys Val Leu Ala Ala Arg Leu Ser Glu Asn Pro Ser Val Arg Val Cys

20 25 30

Leu Ile Glu Ala Gly Arg Arg Asp Thr His Pro Leu Ile His Met Pro

25 35 40 45

Val Gly Phe Ala Lys Met Thr Thr Gly Pro His Thr Trp Asp Leu Leu

50 55 60

Thr Glu Pro Gln Lys His Ala Asn Asn Arg Gln Ile Pro Tyr Val Gln

65 70 75 80

30 Gly Arg Ile Leu Gly Gly Gly Ser Ser Ile Asn Ala Glu Val Phe Thr

	85	90	95
	Arg Gly His Pro Ser Asp Phe Asp Arg Trp Ala Ala Glu Gly Ala Asp		
	100	105	110
	Gly Trp Ser Phe Arg Asp Val Gln Lys Tyr Phe Ile Arg Ser Glu Gly		
5	115	120	125
	Asn Ala Val Phe Ser Gly Thr Trp His Gly Thr Asn Gly Pro Leu Gly		
	130	135	140
	Val Ser Asn Leu Ala Asp Pro Asn Pro Thr Ser Arg Ala Phe Val Gln		
	145	150	155
10	Ser Cys Gln Glu Met Gly Leu Pro Tyr Asn Pro Asp Phe Asn Gly Ala		
	165	170	175
	Ser Gln Glu Gly Ala Gly Ile Tyr Gln Met Thr Ile Arg Asn Asn Arg		
	180	185	190
	Arg Cys Ser Thr Ala Val Gly Tyr Leu Arg Pro Ala Leu Gly Arg Lys		
15	195	200	205
	Asn Leu Thr Val Val Thr Arg Ala Leu Val Leu Lys Ile Val Phe Asn		
	210	215	220
	Gly Thr Arg Ala Thr Gly Val Gln Tyr Ile Ala Asn Gly Thr Leu Asn		
	225	230	235
20	Thr Ala Glu Ala Ser Gln Glu Ile Val Val Thr Ala Gly Ala Ile Gly		
	245	250	255
	Thr Pro Lys Leu Met Met Leu Ser Gly Val Gly Pro Ala Ala His Leu		
	260	265	270
	Arg Glu Asn Gly Ile Pro Val Val Gln Asp Leu Pro Gly Val Gly Glu		
25	275	280	285
	Asn Leu Gln Asp His Phe Gly Val Asp Ile Val Ala Glu Leu Lys Thr		
	290	295	300
	Asp Glu Ser Phe Asp Lys Tyr Arg Lys Leu His Trp Met Leu Trp Ala		
	305	310	315
30	Gly Leu Glu Tyr Thr Met Phe Arg Ser Gly Pro Val Ala Ser Asn Val		

```

      325              330              335
Val Glu Gly Gly Ala Phe Trp Tyr Ser Asp Pro Ser Ser Gly Val Pro
      340              345              350
Asp Leu Gln Phe His Phe Leu Ala Gly Ala Gly Ala Glu Ala Gly Val
5      355              360              365
Thr Ser Val Pro Lys Gly Ala Ser Gly Ile Thr Leu Asn Ser Tyr Val
      370              375              380
Leu Arg Pro Lys Ser Arg Gly Thr Val Arg Leu Arg Ser Ala Asp Pro
385              390              395              400
10 Arg Val Asn Pro Met Val Asp Pro Asn Phe Leu Gly Asp Pro Ala Asp
      405              410              415
Leu Glu Thr Ser Ala Glu Gly Val Arg Leu Ser Tyr Glu Met Phe Ser
      420              425              430
Gln Pro Ser Leu Gln Lys His Ile Lys Glu Thr Cys Phe Phe Ser Gly
15      435              440              445
Lys Gln Pro Thr Met Gln Met Tyr Arg Asp Tyr Ala Arg Glu His Gly
      450              455              460
Arg Thr Ser Tyr His Pro Thr Cys Thr Cys Lys Met Gly Arg Asp Asp
465              470              475              480
20 Met Ser Val Val Asp Pro Arg Leu Lys Val His Gly Leu Glu Gly Ile
      485              490              495
Arg Ile Cys Asp Ser Ser Val Met Pro Ser Leu Leu Gly Ser Asn Thr
      500              505              510
Asn Ala Ala Thr Ile Met Ile Ser Glu Arg Ala Ala Asp Phe Ile Gln
25      515              520              525
Gly Asn Ala
      530

<210> 13
30 <211> 1497

```

<212> DNA

<213> *Gluconobacter oxydans* IFO 3293

<400> 13

```

5  atgaatgttg tctcaaagac tgtatcttta ccgttaaagc cgcgtgagtt cggattctat
   60
   attgatggag aatggcgcgc aggtaaggat ttcttcgatc gttcctcgcc ggctcatgat
  120
   gttcccgctca cccgtattcc acgctgcacc cgtgaggacc ttgatgaggc agtcgctgct
10 180
   gcacgtcgtg ctttcgagaa cggaagctgg tcgggcctgg cggccgcgga tcgtgcgggc
  240
   gttcttctga aagccgcggg ccttctgcgc gagcgccgcg atgacatcgc ttactgggaa
  300
15 gttctcgaaa acgggaagcc catcagccag gcgaagggcg agatcgatca ctgtatcgcc
  360
   tgtttcgaga tggcggccgg cgctgcgcgg atgctgcatg gtgatacggt caacaatctg
  420
   ggcgaggggc tgttcggcat ggtcctgcgg gagcccatcg gtgtcgtcgg tctgattacg
20 480
   ccgtggaact tcccgttcat gatcctgtgt gagegcgcgc ctttcattct cgcacccggc
  540
   tgcacgctgg tcgtcaagcc tgccgaagtc acgagtgcc aacgcttct tctggcggaa
  600
25 gtgctggcgg atgcggggct gccgaagggt gtcttcaatg ttgtgacggg cacggggcgc
  660
   acggtcgggc aggccatgac cgagcatcag gacatcgaca tgctgtcctt cacgggctcc
  720
   acgggcgctc gcaagtctg catccatgcg gcggctgaca gcaacctgaa gaaactgggc
30 780

```

cttgagcttg gcggaagaa cccgatcgtc gtgttcgctg acagcaacct tgaggatgcg

840

gccgatgcgg tagctttcgg gattagtttc aacaccgggc agtgctgtgt gtcgtcgagc

900

5 cgcctgattg tagagcggtc cgtggccgag aagttcgagc gtctcgttgt ggcgaaaatg

960

gagaagatcc gcgtcggcga tccgttcgat cctgagacgc agatcggcgc cattacgacg

1020

gaagcgcaga acaagaccat tctggactat atcgccaagg gcaaggccga gggcgccagg

10 1080

ctgctctgcg gtggcgggat cgtcgatttc ggcaaagggc agtatatcca gccgacgctt

1140

ttcacggatg tgaagccctc gatgggcacg gcgctgacg agatttttgg gccggctcctg

1200

15 gcgtccttcc acttcgatac cgtcgatgag gcgatcgca ttgccaatga cacgggtttac

1260

ggcctggccg catcggctcg gagcaaggat atcgacaagg cgcttgccgt gaccgcgtcg

1320

gttcgcgctg gccgcttctg ggtgaacacc atcatgagcg gtggtcccga gacgccgctg

20 1380

ggtgggtttca agcagtcggg ctggggccgt gaggccggtc tgtacggcgt tgaggaatat

1440

acgcagatca aatctgtcca tatcgaaact ggcaaacgtt cgactggat ttcgtaa

1497

25

<210> 14

<211> 498

<212> PRT

<213> Gluconobacter oxydans IFO 3293

30

<400> 14

```

Met Asn Val Val Ser Lys Thr Val Ser Leu Pro Leu Lys Pro Arg Glu
1           5           10           15
5  Phe Gly Phe Tyr Ile Asp Gly Glu Trp Arg Ala Gly Lys Asp Phe Phe
           20           25           30
Asp Arg Ser Ser Pro Ala His Asp Val Pro Val Thr Arg Ile Pro Arg
           35           40           45
Cys Thr Arg Glu Asp Leu Asp Glu Ala Val Ala Ala Ala Arg Arg Ala
10      50           55           60
Phe Glu Asn Gly Ser Trp Ser Gly Leu Ala Ala Ala Asp Arg Ala Ala
65           70           75           80
Val Leu Leu Lys Ala Ala Gly Leu Leu Arg Glu Arg Arg Asp Asp Ile
           85           90           95
15  Ala Tyr Trp Glu Val Leu Glu Asn Gly Lys Pro Ile Ser Gln Ala Lys
           100          105          110
Gly Glu Ile Asp His Cys Ile Ala Cys Phe Glu Met Ala Ala Gly Ala
           115          120          125
Ala Arg Met Leu His Gly Asp Thr Phe Asn Asn Leu Gly Glu Gly Leu
20      130          135          140
Phe Gly Met Val Leu Arg Glu Pro Ile Gly Val Val Gly Leu Ile Thr
145          150          155          160
Pro Trp Asn Phe Pro Phe Met Ile Leu Cys Glu Arg Ala Pro Phe Ile
           165          170          175
25  Leu Ala Ser Gly Cys Thr Leu Val Val Lys Pro Ala Glu Val Thr Ser
           180          185          190
Ala Thr Thr Leu Leu Leu Ala Glu Val Leu Ala Asp Ala Gly Leu Pro
           195          200          205
Lys Gly Val Phe Asn Val Val Thr Gly Thr Gly Arg Thr Val Gly Gln
30      210          215          220

```

```

Ala Met Thr Glu His Gln Asp Ile Asp Met Leu Ser Phe Thr Gly Ser
225                230                235                240
Thr Gly Val Gly Lys Ser Cys Ile His Ala Ala Ala Asp Ser Asn Leu
                245                250                255
5  Lys Lys Leu Gly Leu Glu Leu Gly Gly Lys Asn Pro Ile Val Val Phe
                260                265                270
Ala Asp Ser Asn Leu Glu Asp Ala Ala Asp Ala Val Ala Phe Gly Ile
                275                280                285
Ser Phe Asn Thr Gly Gln Cys Cys Val Ser Ser Ser Arg Leu Ile Val
10 290                295                300
Glu Arg Ser Val Ala Glu Lys Phe Glu Arg Leu Val Val Ala Lys Met
305                310                315                320
Glu Lys Ile Arg Val Gly Asp Pro Phe Asp Pro Glu Thr Gln Ile Gly
                325                330                335
15 Ala Ile Thr Thr Glu Ala Gln Asn Lys Thr Ile Leu Asp Tyr Ile Ala
                340                345                350
Lys Gly Lys Ala Glu Gly Ala Arg Leu Leu Cys Gly Gly Gly Ile Val
                355                360                365
Asp Phe Gly Lys Gly Gln Tyr Ile Gln Pro Thr Leu Phe Thr Asp Val
20 370                375                380
Lys Pro Ser Met Gly Ile Ala Arg Asp Glu Ile Phe Gly Pro Val Leu
385                390                395                400
Ala Ser Phe His Phe Asp Thr Val Asp Glu Ala Ile Ala Ile Ala Asn
                405                410                415
25 Asp Thr Val Tyr Gly Leu Ala Ala Ser Val Trp Ser Lys Asp Ile Asp
                420                425                430
Lys Ala Leu Ala Val Thr Arg Arg Val Arg Ala Gly Arg Phe Trp Val
                435                440                445
Asn Thr Ile Met Ser Gly Gly Pro Glu Thr Pro Leu Gly Gly Phe Lys
30 450                455                460

```

Gln Ser Gly Trp Gly Arg Glu Ala Gly Leu Tyr Gly Val Glu Glu Tyr

465 470 475 480

Thr Gln Ile Lys Ser Val His Ile Glu Thr Gly Lys Arg Ser His Trp

485 490 495

5 Ile Ser

<210> 15

<211> 1401

<212> DNA

10 <213> Gluconobacter oxydans IFO 3293

<400> 15

atgagtgcac ttggcgctgt agagagcggt ccgcgcgttt cactcaacgg cagaagtgcc

60

15 ctggtagcca ctctggctgc ggtggcaggt ctgatgttcg ggctggatct ggggtgtcata

120

tccggagcgc tgaaattcat cggatcgcag tttcaggcca atgatttcgg caaggagtgc

180

attgtctccg gcatgatggt cggagcagcg ctccggggccc tggtcgggtgg gcggctgtcc

20 240

tatattctgg ggcgcaaaaa gctgctgac atctcgccac tgctcttcat ttccgggtcg

300

gcactctgcg cgctggcacc ctccgttccg gtctgatcg cggggcgtgt ggtccttgga

360

25 ctggcgatcg gcgtcgcgag ctttaccgct ccactctata ttccggagat tgcgcaccag

420

cagcatcgcg gggcgctgat ttccgtctat cagctcatga ttacgggtggg cattctggtc

480

gcgttcatct ccgatgcgct tctggcctat tcgggatcat ggcgctggat gctggggatc

30 540

gtgggtgtcc ccggcgtgct gttcctgac ggtgtcctgt tccttcgga cagcccgcg
 600
 tggctggttc tgcgggggcg gcatgatgaa gcgttcgatg tcctgcacgg gctgcgtggc
 660
 5 aatgatgcgc atgcgcggga cgagatcgtc gaaatccgtg atcagcttgc acaggccgat
 720
 ggcggatata cgctcttcaa gaaaaacccg aacttcgcc gcgcggtcta tctgggcgtg
 780
 gcgcttcagg cgatccagca gttcacccgc atcaacgtgg tcatgtacta cgcgccgat
 10 840
 atcttcgaaa tggccggttt cggcaccagt gccgccctgt ggggcacggg gatcgtggga
 900
 ctggtcaatg tgctctccac gttcatcgcc attggatatg tcgatcgctg gggccggcgg
 960
 15 cccatgctga tcgcgggctt catcatgatg gcgctcgga tgtttacggg cggcatgctg
 1020
 atgtatttcg gcacgggcaa cagcgatctg gcgcgttacg cggcagttgc gatgctgctg
 1080
 tgcttcatcg tcgggtttgc attttcgcc gggccgctgg tgtggatcct gtgctctgaa
 20 1140
 gttcagccca tgaaggggcg tgacttcggc atcgggtgct cgaccttcac caactggggc
 1200
 accaacaatga ttgttggtct gaccttctc tcgcttctga cgacactggg aagtgccag
 1260
 25 accttctggc tctatgccct gcttaatctc atcttcattt tcgtgacgat gaagtctgtg
 1320
 ccggaacccc gtggcgtgct gcttgagcag atcgagcgga atctcatggc cggtaaaccc
 1380
 ctcaatcgta tcggcctgta a
 30 1401

<210> 16

<211> 466

<212> PRT

5 <213> Gluconobacter oxydans IFO 3293

<400> 16

Met Ser Asp Val Gly Val Val Glu Ser Val Pro Arg Val Ser Leu Asn

10 1 5 10 15

Gly Arg Ser Ala Leu Val Ala Thr Leu Ala Ala Val Ala Gly Leu Met

20 25 30

Phe Gly Leu Asp Leu Gly Val Ile Ser Gly Ala Leu Lys Phe Ile Gly

35 40 45

15 Ile Glu Phe Gln Val Asn Asp Phe Gly Lys Glu Cys Ile Val Ser Gly

50 55 60

Met Met Val Gly Ala Ala Leu Gly Ala Leu Phe Gly Gly Arg Leu Ser

65 70 75 80

Tyr Ile Leu Gly Arg Lys Lys Leu Leu Ile Ile Ser Ala Leu Leu Phe

20 85 90 95

Ile Ser Gly Ser Ala Leu Cys Ala Leu Ala Pro Ser Val Pro Val Leu

100 105 110

Ile Ala Gly Arg Val Val Leu Gly Leu Ala Ile Gly Val Ala Ser Phe

115 120 125

25 Thr Ala Pro Leu Tyr Ile Ser Glu Ile Ala His Gln Gln His Arg Gly

130 135 140

Ala Leu Ile Ser Val Tyr Gln Leu Met Ile Thr Val Gly Ile Leu Val

145 150 155 160

Ala Phe Ile Ser Asp Ala Leu Leu Ala Tyr Ser Gly Ser Trp Arg Trp

30 165 170 175

Met Leu Gly Ile Val Gly Val Pro Gly Val Leu Phe Leu Ile Gly Val
 180 185 190
 Leu Phe Leu Pro Asp Ser Pro Arg Trp Leu Val Leu Arg Gly Arg His
 195 200 205
 5 Asp Glu Ala Phe Asp Val Leu His Gly Leu Arg Gly Asn Asp Ala His
 210 215 220
 Ala Arg Asp Glu Ile Val Glu Ile Arg Asp Gln Leu Ala Gln Ala Asp
 225 230 235 240
 Gly Gly Tyr Thr Leu Phe Lys Lys Asn Pro Asn Phe Arg Arg Ala Val
 10 245 250 255
 Tyr Leu Gly Val Ala Leu Gln Ala Ile Gln Gln Phe Thr Gly Ile Asn
 260 265 270
 Val Val Met Tyr Tyr Ala Pro His Ile Phe Glu Met Ala Gly Phe Gly
 275 280 285
 15 Thr Ser Ala Ala Leu Trp Gly Thr Val Ile Val Gly Leu Val Asn Val
 290 295 300
 Leu Ser Thr Phe Ile Ala Ile Gly Tyr Val Asp Arg Trp Gly Arg Arg
 305 310 315 320
 Pro Met Leu Ile Ala Gly Phe Ile Met Met Ala Leu Gly Met Phe Thr
 20 325 330 335
 Val Gly Met Leu Met Tyr Phe Gly Thr Gly Asn Ser Asp Leu Ala Arg
 340 345 350
 Tyr Ala Ala Val Ala Met Leu Leu Cys Phe Ile Val Gly Phe Ala Phe
 355 360 365
 25 Ser Ala Gly Pro Leu Val Trp Ile Leu Cys Ser Glu Val Gln Pro Met
 370 375 380
 Lys Gly Arg Asp Phe Gly Ile Gly Cys Ser Thr Phe Thr Asn Trp Gly
 385 390 395 400
 Thr Asn Met Ile Val Gly Leu Thr Phe Leu Ser Leu Leu Thr Thr Leu
 30 405 410 415

Gly Ser Ala Gln Thr Phe Trp Leu Tyr Ala Leu Leu Asn Leu Ile Phe

420

425

430

Ile Phe Val Thr Met Lys Phe Val Pro Glu Thr Arg Gly Val Ser Leu

435

440

445

5 Glu Gln Ile Glu Arg Asn Leu Met Ala Gly Lys Pro Leu Asn Arg Ile

450

455

460

Gly Leu

465

REIVINDICAÇÕES

1. Polinucleotídeo caracterizado pelo fato de ser selecionado do grupo que consiste em:

(a) polinucleotídeos que codificam um polipeptídeo compreendendo a seqüência de aminoácido, de acordo com a SEQ ID NO:2,

(b) polinucleotídeos compreendendo a seqüência de nucleotídeo de acordo com a SEQ ID NO: 1,

(c) polinucleotídeos compreendendo uma seqüência de nucleotídeo que pode ser obtido por amplificação de ácido nucléico tal como reação de cadeia de polimerase, utilizando DNA genômico a partir de um microorganismo como padrão e um conjunto iniciador de acordo com SEQ ID NO: 3 e SEQ ID NO: 4;

(d) polinucleotídeos compreendendo uma seqüência de nucleotídeos codificando um fragmento ou derivado de um polipeptídeo codificado por um polinucleotídeo de qualquer um de (a) a (c) onde no derivado um ou mais resíduos de aminoácido são substituídos de forma conservativa em comparação com o polipeptídeo, e o fragmento ou derivado tem a atividade de um regulador de transcrição, preferivelmente um repressor de desidrogenase L-sorbose (SDH) e/ou desidrogenase L-sorbose (SNDH);

(e) polinucleotídeos, cuja fita complementar hibridiza sob condições rigorosas em um polinucleotídeo como definido em qualquer um de (a) a (d) e que codifica um regulador de transcrição, preferivelmente um repressor de desidrogenase L-sorbose (SDH) e/ou desidrogenase L-sorbose (SNDH); e

(f) polinucleotídeos que são pelo menos 60%, tal como

70, 85, 90 ou 95% idênticos a um polinucleotídeo como definido em qualquer um de (a) a (d) e que codificam um regulador de transcrição, preferivelmente um repressor de desidrogenase L-sorbose (SDH) e/ou desidrogenase L-sorbose (SNDH);

ou

a fita complementar de um tal polinucleotídeo.

2. Vetor caracterizado pelo fato de conter o polinucleotídeo da reivindicação 1.

10 3. Vetor, de acordo com a reivindicação 2, caracterizado pelo fato de que o polinucleotídeo é operativamente ligado a seqüências de controle de expressão permitindo a expressão em células hospedeiras procarióticas ou eucarióticas.

15 4. Microorganismo caracterizado pelo fato de ser manipulado geneticamente com um polinucleotídeo da reivindicação 1 ou com o vetor da reivindicação 2 ou 3.

5. Microorganismo, de acordo com a reivindicação 4, caracterizado pelo fato de ser capaz de produzir
20 diretamente Vitamina C a partir de D-sorbitol em quantidades de 300 mg/l ou mais quando medido em um método de célula em repouso após um período de incubação de 20 horas.

6. Microorganismo, de acordo com a reivindicação 5,
25 caracterizado pelo fato de ser capaz de produzir diretamente a Vitamina C de L-sorbose em quantidades de 800 mg/l ou mais.

7. Polipeptídeo caracterizado pelo fato de ser codificado por um polinucleotídeo de acordo com a
30 reivindicação 1.

8. Processo para produzir células capazes de expressar um polipeptídeo da reivindicação 7 caracterizado pelo fato de compreender a etapa de manipular geneticamente células com o vetor da reivindicação 2 ou 3 ou com um polinucleotídeo da reivindicação 1.

9. Uso de um polinucleotídeo rompido da reivindicação 1 caracterizado pelo fato de ser para a produção de Vitamina C e/ou 2-KGA.

10. Microorganismo da reivindicação 4 ou um microorganismo contendo um gene endógeno compreendendo um polinucleotídeo da reivindicação 1 caracterizado pelo fato de ser geneticamente alterado de tal modo que leva a um rendimento e/ou eficiência de produção, aperfeiçoados, de vitamina C e/ou 2-KGA produzido pelo microorganismo.

11. Microorganismo, de acordo com a reivindicação 10, caracterizado pelo fato de produzir um polipeptídeo da reivindicação 7, com atividade de um regulador de transcrição diminuída ou abolida, preferivelmente um repressor de gene de desidrogenase L-sorbose (SDH) e/ou desidrogenase L-sorbose (SNDH).

12. Microorganismo, de acordo com qualquer uma das reivindicações 4, 5, 6, 10 ou 11, caracterizado pelo fato de que o polinucleotídeo, da reivindicação 1, é rompido.

13. Microorganismo, de acordo com qualquer uma das reivindicações 4, 5, 6, 10, 11 ou 12, caracterizado pelo fato de ser selecionado do grupo que consiste em *Pseudomonas*, *Pantoea*, *Escherichia*, *Ketogulonicigenium* e bactérias de ácido acético como por exemplo, *Gluconobacter*, *Acetobacter* ou *Gluconoacetobacter*, preferivelmente *Acetobacter* sp., *Acetobacter aceti*, *Gluconobacter*

frateurii, *Gluconobacter cerinus*, *Gluconobacter thailandicus*, *Gluconobacter oxydans*, mais preferivelmente *G. oxydans*, mais preferivelmente *G. oxydans* IFO 3293.

14. Processo para a produção de um gene regulador de transcrição endógeno rompido, preferivelmente um gene de desidrogenase L-sorbose (SDH) e/ou desidrogenase L-sorbose (SNDH) em um microorganismo, o microorganismo compreendendo um polinucleotídeo da reivindicação 1, o processo sendo caracterizado pelo fato de compreender a etapa de alterar o polinucleotídeo de tal modo que leva a um rendimento e/ou eficiência de produção aperfeiçoados de Vitamina C e/ou 2-KGA produzida pelo microorganismo.

15. Processo para a produção de um microorganismo capaz de produzir Vitamina C e/ou 2-KGA caracterizado pelo fato de compreender a etapa de alterar o microorganismo de modo que o microorganismo produz um polipeptídeo com atividade reduzida ou abolida de um regulador de transcrição, preferivelmente um repressor de gene de desidrogenase L-sorbose (SDH) e/ou desidrogenase L-sorbose (SNDH) levando a um rendimento e/ou eficiência de produção aperfeiçoados de Vitamina C e/ou 2-KGA produzida pelo microorganismo.

16. Processo para a produção de um microorganismo contendo um gene endógeno compreendendo um polinucleotídeo da reivindicação 1 caracterizado pelo fato de compreender a etapa de alterar o microorganismo de modo que o gene endógeno seja sub-expresso ou rompido, levando a um rendimento e/ou eficiência de produção aperfeiçoados da Vitamina C e/ou 2-KGA produzida pelo microorganismo.

17. Processo, de acordo com a reivindicação 15 ou 16,

caracterizado pelo fato de ser para a produção de um microorganismo de qualquer uma das reivindicações 10, 11, 12 ou 13.

18. Microorganismo caracterizado pelo fato de que pode ser obtido por um processo de qualquer uma das reivindicações 14, 15 ou 16.

19. Processo para a produção de Vitamina C e/ou 2-KGA com um microorganismo de qualquer uma das reivindicações 10, 11, 12 ou 13 ou reivindicações 4, 5 ou 6 caracterizado pelo fato de que o microorganismo é incubado/cultivado em um meio aquoso sob condições que permitem a produção direta de Vitamina C e/ou 2-KGA a partir de D-sorbitol ou L-sorbose e em que opcionalmente Vitamina C e/ou 2-KGA é isolada como o produto de fermentação.

GENE SMS 27

A presente invenção relaciona-se aos genes recentemente identificados que codificam as proteínas que são envolvidas na síntese do ácido L-ascórbico (em seguida também referido como vitamina C). A invenção também apresenta os polinucleotídeos que compreendem as seqüências de extensão completa do polinucleotídeo dos novos genes e os fragmentos destes, os novos polipeptídeos codificados pelos polinucleotídeos e fragmentos destes, assim como seus equivalentes funcionais. A invenção atual relaciona-se também ao uso de tais polinucleotídeos e de polipeptídeos como ferramentas biotecnológicas na produção da vitamina C dos microorganismos, por meio de uma modificação dos tais polinucleotídeos e/ou de polipeptídeos codificados tem um impacto direto ou indireto no rendimento, na produção, e/ou na eficiência da produção do produto de fermentação em tal microorganismo. São incluídos também métodos/processos de usar os polinucleotídeos e as seqüências modificadas do polinucleotídeo para transformar microorganismos hospedeiros. A invenção relaciona-se também aos microorganismos geneticamente projetados e ao seu uso para a produção direta da vitamina C.