



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

(11) BR 122020013701-0 B1



(22) Data do Depósito: 18/05/2009

(45) Data de Concessão: 01/06/2021

(54) Título: MICROORGANISMO GENETICAMENTE MODIFICADO PARA PRODUÇÃO DE UM HIDROCARBONETO, MÉTODO DE PRODUÇÃO DE UM ALCANO OU ALCENO E CULTURA DE CÉLULAS

(51) Int.Cl.: C12N 9/02; C12N 9/88; C12P 7/64; C10L 1/02.

(30) Prioridade Unionista: 16/05/2008 US 61/053,955.

(73) Titular(es): GENOMATICA, INC..

(72) Inventor(es): ANDREAS SCHIRMER; MATHEW RUDE; SHANE BRUBAKER.

(86) Pedido PCT: PCT US2009044403 de 18/05/2009

(87) Publicação PCT: WO 2009/140695 de 19/11/2009

(85) Data do Início da Fase Nacional: 03/07/2020

(62) Pedido Original do Dividido: PI0912684-8 - 18/05/2009

(57) Resumo: Composições e métodos para a produção de aldeídos, alcanos e alcenos são descritos neste documento. Os aldeídos, os alcanos e os alcenos podem ser usados em biocombustíveis.

**MICROORGANISMO GENETICAMENTE MODIFICADO PARA
PRODUÇÃO DE UM HIDROCARBONETO, MÉTODO DE PRODUÇÃO DE UM
ALCANO OU ALCENO E CULTURA DE CÉLULAS**

(Dividido do PI 0912684-8 depositado em 18/05/2009)

REFERÊNCIA A PEDIDOS RELACIONADOS

[001] Este pedido reivindica o benefício do Pedido Provisório norte-americano nº 61/053.955, depositado em 16 de maio de 2008, cujo conteúdo está incorporado neste documento em sua totalidade.

ANTECEDENTES DA INVENÇÃO

[002] O petróleo é um recurso natural limitado encontrado na Terra nas formas líquida, gasosa e sólida. O petróleo é primeiramente composto de hidrocarbonetos, os quais são formados principalmente por carbono e hidrogênio. Ele também contém quantidades significativas de outros elementos, tais como nitrogênio, oxigênio ou enxofre em diferentes formas.

[003] O petróleo é um recurso valioso, mas os produtos do petróleo são desenvolvidos a custos consideráveis, tanto financeiros quanto ambientais. Primeiro, as fontes de petróleo devem ser descobertas. A exploração de petróleo é um empreendimento caro e arriscado. O custo da exploração de poços em águas profundas pode exceder os US\$ 100 milhões. Além do mais, não há garantia de que esses poços contenham petróleo. Estima-se que apenas 40% dos poços perfurados conduzam a poços produtivos gerando hidrocarbonetos comerciais. Além do custo econômico, a exploração de petróleo carrega em si um elevado custo ambiental. Por exemplo, a exploração em alto mar perturba o ambiente marinho nas imediações.

[004] Após um poço produtivo ser descoberto, o petróleo deve ser extraído da Terra a um grande custo. Durante a recuperação primária, a pressão natural do subsolo é suficiente para extrair cerca de 20% do petróleo no poço. Enquanto essa pressão natural cai, métodos secundários de recuperação são empregados, se forem econômicos. Geralmente, a recuperação secundária envolve o aumento da pressão do poço através de, por exemplo, injeção de água, injeção de gás natural ou por elevação a gás. Usando métodos secundários de recuperação, um adicional de 5% a 15% de petróleo é recuperado. Uma vez que os métodos secundários de recuperação estejam esgotados, métodos terciários de recuperação podem ser usados, se for econômico. Métodos terciários envolvem a redução da viscosidade do petróleo de forma a facilitar a extração. Usando métodos terciários de recuperação, um adicional de 5% a 15% de petróleo é recuperado. Portanto, mesmo sob as melhores circunstâncias, apenas 50% do petróleo em um poço pode ser extraído. A extração de petróleo ainda carrega um custo ambiental. Por exemplo, a extração de petróleo pode resultar em grandes infiltrações de petróleo ascendendo até a superfície. Além disso, a perfuração em alto mar envolve a dragagem do solo oceânico, o que perturba ou destrói o ambiente marinho nas imediações.

[005] Como os depósitos de petróleo não são encontrados uniformemente por toda a Terra, o petróleo deve ser transportado por grandes distâncias das regiões produtoras de petróleo até as regiões consumidoras de petróleo. Além dos custos de transporte, há também o risco ambiental de derramamentos de óleo devastadores.

[006] Em sua forma natural, o petróleo cru extraído da

Terra possui poucos usos comerciais. Ele é uma mistura de hidrocarbonetos (por exemplo, parafinas (ou alcanos), olefinas (ou alcenos), alcinos, naftenos (ou cicloalcanos), compostos alifáticos, compostos aromáticos etc.) de variados comprimentos e complexidades. Além disso, o petróleo cru contém outros compostos orgânicos (por exemplo, compostos orgânicos contendo nitrogênio, oxigênio, enxofre etc.) e impurezas (por exemplo, enxofre, sal, ácido, metais etc.).

[007] Portanto, o petróleo cru deve ser refinado e purificado antes que possa ser usado comercialmente. Devido à sua alta densidade energética e seu transporte fácil, a maioria do petróleo é refinada em combustíveis, tais como combustíveis de transportes (por exemplo, gasolina, *diesel*, combustível de avião etc.), óleo de aquecimento, gás líquido de petróleo etc.

[008] O petróleo cru é igualmente uma fonte primária de matérias-primas para a produção de petroquímicos. As duas classes principais de matérias-primas derivadas a partir do petróleo são olefinas de cadeia pequena (por exemplo, etileno e propileno) e aromáticos (por exemplo, benzeno e isômeros de xileno). Essas matérias-primas são derivadas a partir de cadeias de hidrocarbonetos maiores no petróleo cru através de seu craqueamento a um custo considerável usando uma variedade de métodos, tais como craqueamento catalítico, craqueamento a vapor ou reforma catalítica. Essas matérias-primas são usadas para fazer petroquímicos, os quais não podem ser diretamente refinados a partir do petróleo cru, tais como monômeros, solventes, detergentes ou adesivos.

[009] Um exemplo de uma matéria-prima derivada a partir do petróleo cru é o etileno. O etileno é usado para produzir

petroquímicos tais como polietileno, etanol, óxido de etileno, etileno glicol, poliéster, éter glicol, etoxilato, acetato de vinila, 1,2-dicloroetano, tricloroetileno, tetracloroetileno, cloreto de vinila e cloreto de polivinila. Outro exemplo de matéria-prima é o propileno, que é usado para produzir álcool isopropílico, acrilonitrila, polipropileno, óxido de propileno, propileno glicol, éteres de glicol, butileno, isobutileno, 1,3-butadieno, elastômeros sintéticos, poliolefinas, alfaolefinas, álcoois graxos, ácido acrílico, polímeros acrílicos, cloreto de alila, epicloroidrina e resinas de epóxi.

[010] Esses petroquímicos podem, então, ser utilizados para fazer especialidades químicas, tais como plásticos, resinas, fibras, elastômeros, fármacos, lubrificantes ou géis. Especialidades químicas particulares que podem ser produzidas a partir de matérias-primas petroquímicas são: ácidos graxos, hidrocarbonetos (por exemplo, de cadeia longa, de cadeia ramificada, saturado, insaturado etc.), álcoois graxos, ésteres, aldeídos graxos, cetonas, lubrificantes etc.

[011] As especialidades químicas possuem muitos usos comerciais. Os ácidos graxos são usados comercialmente como surfactantes, por exemplo, em detergentes ou sabões. Eles também podem ser usados como aditivos em combustíveis, óleos lubrificantes, tintas, lacas, velas, óleo de salada, encurtamento, cosméticos e emulsificantes. Além disso, os ácidos graxos são usados como ativadores de aceleração em produtos de borracha. Os ácidos graxos também podem ser usados como uma matéria-prima para produzir metil ésteres,

amidas, aminas, cloretos ácidos, anidridos, dímeros cetenos e ácidos peróxidos e ésteres.

[012] Os hidrocarbonetos possuem muitos usos comerciais. Por exemplo, alcanos de cadeia mais curta são usados como combustíveis. Metano e etano são os principais constituintes do gás natural. Alcanos de cadeia mais longa (por exemplo, de cinco a dezesseis carbonos) são usados como combustíveis de transporte (por exemplo, gasolina, *diesel* ou combustível de aviação). Alcanos contendo mais do que dezesseis átomos de carbono são componentes importantes de óleos combustíveis e óleos lubrificantes. Alcanos ainda mais longos, os quais são sólidos à temperatura ambiente, podem ser usados, por exemplo, como uma cera de parafina. Alcanos que contenham aproximadamente trinta e cinco carbonos são encontrados no betume, o qual é usado para pavimentação de estradas. Além disso, alcanos de cadeia mais longa podem ser craqueados para produzir hidrocarbonetos de cadeia mais curta, comercialmente úteis.

[013] Assim como os alcanos de cadeia mais curta, os alcenos de cadeia mais curta são usados em combustíveis de transportes. Alcenos de cadeia mais longa são usados em plásticos, lubrificantes e lubrificantes sintéticos. Além do mais, alcenos são usados como uma matéria-prima para produzir álcoois, ésteres, plastificantes, surfactantes, aminas terciárias, agentes de recuperação de óleo melhorados, ácidos graxos, tióis, anidridos alcenilsucínicos, epóxidos, alcanos clorados, alcenos clorados, ceras, aditivos de combustível e redutores de fluxo de arraste.

[014] Álcoois graxos possuem muitos usos comerciais. Os álcoois graxos de cadeia mais curta são usados nas indústrias

de cosméticos e de alimentos como emulsificantes, emolientes e espessantes. Devido à sua natureza anfifílica, os álcoois graxos se comportam como surfactantes não iônicos, os quais são úteis como detergentes. Além disso, os álcoois graxos são usados em ceras, gomas, resinas, pomadas farmacêuticas e loções, aditivos de óleos lubrificantes, têxteis antiestáticos e agentes finalizadores, plastificantes, cosméticos, solventes industriais e solventes para gorduras.

[015] Os ésteres possuem muitos usos comerciais. Por exemplo, o biodiesel, um combustível alternativo, é composto de ésteres (por exemplo, metil éster de ácido graxo, etil éster de ácido graxo etc.). Alguns ésteres com baixo peso molecular são voláteis com um odor agradável, o que faz com que eles sejam usados como fragrâncias ou agentes flavorizantes. Além disso, os ésteres são usados como solventes para lacas, tintas e vernizes. Além do mais, algumas substâncias que ocorrem naturalmente, tais como ceras, gorduras e óleos, são compostas de ésteres. Os ésteres são também usados como agentes de amolecimento em resinas e plásticos, plastificantes, retardadores de chamas e aditivos em gasolina e óleo. Além disso, ésteres podem ser usados na produção de polímeros, filmes, têxteis, corantes e fármacos.

[016] Aldeídos são usados para produzir muitas especialidades químicas. Por exemplo, aldeídos são usados para produzir polímeros, resinas (por exemplo, *Bakelite*), corantes, flavorizantes, plastificantes, perfumes, fármacos e outros químicos. Alguns são usados como solventes, conservantes ou desinfetantes. Alguns compostos naturais e sintéticos, tais como vitaminas e hormônios, são aldeídos. Além disso, muitos açúcares contêm grupos aldeídos.

[017] Cetonas são usadas comercialmente como solventes. Por exemplo, a acetona é frequentemente usada como um solvente, mas também é uma matéria-prima para a produção de polímeros. Cetonas são igualmente usadas em lacas, tintas, explosivos, perfumes e processamento têxtil. Além disso, cetonas são usadas para produzir álcoois, alcenos, alcanos, iminas e enaminas.

[018] Ademais, o petróleo cru é uma fonte de lubrificantes. Os lubrificantes derivados de petróleo são tipicamente compostos de olefinas, particularmente poliolefinas e alfaolefinas. Os lubrificantes podem tanto ser refinados a partir do petróleo cru quanto serem manufaturados usando matérias-primas refinadas a partir do petróleo cru.

[019] A obtenção dessas especialidades químicas a partir do petróleo cru exige um investimento financeiro significativo bem como uma grande quantidade de energia. Isso também é um processo ineficiente porque frequentemente os hidrocarbonetos de cadeia longa no petróleo cru são craqueados para produzir monômeros menores. Esses monômeros são, então, usados como a matéria-prima para a produção de especialidades químicas mais complexas.

[020] Além dos problemas com a exploração, a extração, o transporte e o refino do petróleo, o petróleo é uma fonte limitada e em encolhimento. Uma estimativa de consumo de petróleo no mundo é de 30 bilhões de barris por ano. Por algumas estimativas, prevê-se que, de acordo com os níveis atuais de produção, as reservas mundiais de petróleo podem ser esgotadas antes do ano 2050.

[021] Finalmente, a queima de combustíveis à base de

petróleo libera gases do efeito estufa (por exemplo, dióxido de carbono) e outras formas de poluição do ar (por exemplo, monóxido de carbono, dióxido de enxofre etc.). Conforme a demanda mundial por combustível aumenta, a emissão de gases do efeito estufa e outras formas de poluição do ar também aumentam. A acumulação de gases do efeito estufa na atmosfera conduz a um aumento do aquecimento global. Portanto, além de danificar o meio ambiente de forma local (por exemplo, derramamentos de óleo, dragagem dos ambientes marinhos etc.), a queima de petróleo também danifica o meio ambiente globalmente.

[022] Devido aos desafios inerentes colocados pelo petróleo, há uma necessidade por uma fonte renovável ao petróleo a qual não necessite ser explorada, extraída, transportada por longas distâncias ou substancialmente refinada como o petróleo. Também há uma necessidade por uma fonte renovável ao petróleo que possa ser produzida economicamente sem criar o tipo de dano ao meio ambiente produzido pela indústria do petróleo e pela queima dos combustíveis à base de petróleo. Por razões parecidas, há também uma necessidade por uma fonte renovável de químicos que são tipicamente derivados a partir do petróleo.

SUMÁRIO DA INVENÇÃO

[023] A invenção é baseada, pelo menos em parte, na identificação de genes de cianobactérias que codificam polipeptídeos de hidrocarbonetos biossintéticos. De acordo, em um aspecto, a invenção destaca um método de produção de hidrocarboneto, o método compreendendo a produção em uma célula hospedeira de um polipeptídeo composto da sequência de aminoácido de SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18,

20, 22, 24, 26, 28, 30, 32, 34 ou 36, ou uma variável desta, e isolando o hidrocarboneto da célula hospedeira.

[024] Em algumas concretizações, o polipeptídeo compreende uma sequência de aminoácido contendo pelo menos cerca de 70%, pelo menos cerca de 80%, pelo menos cerca de 85%, pelo menos cerca de 90%, pelo menos cerca de 91%, pelo menos cerca de 92%, pelo menos cerca de 93%, pelo menos cerca de 94%, pelo menos cerca de 95%, pelo menos cerca de 96%, pelo menos cerca de 97%, pelo menos cerca de 98% ou pelo menos cerca de 99% de identificação com a SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34 ou 36.

[025] Em algumas concretizações, o polipeptídeo compreende a sequência de aminoácido de SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34 ou 36 com uma ou mais substituições, adições, inserções ou supressões de aminoácidos. Em algumas concretizações, o polipeptídeo possui atividade de descarboxilase. Em ainda outras concretizações, o polipeptídeo compreende a sequência de aminoácido de SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34 ou 36, com uma ou mais substituições conservadoras de aminoácidos. Por exemplo, o polipeptídeo compreende uma ou mais das seguintes substituições conservadoras de aminoácidos: reposição de um aminoácido alifático, tal como alanina, valina, leucina e isoleucina, com outro aminoácido alifático; reposição de uma serina com uma treonina; reposição de uma treonina com uma serina; reposição de um resíduo ácido, tal com ácido aspártico e ácido glutâmico, com outro resíduo ácido; reposição de um resíduo carregando um grupo amida, tal como

asparagina e glutamina, com outro resíduo carregando um grupo amida; troca de um resíduo básico, tal como lisina e arginina, com outro resíduo básico; e reposição de um resíduo aromático, tal como fenilalanina e tirosina, com outro resíduo aromático. Em algumas concretizações, o polipeptídeo possui cerca de 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90, 100 ou mais substituições, adições, inserções ou supressões de aminoácidos. Em algumas concretizações, o polipeptídeo possui atividade descarboxilase.

[026] Em outras concretizações, o polipeptídeo compreende a sequência de aminoácido de: (i) SEQ ID NO: 37 ou SEQ ID NO: 38 ou SEQ ID NO: 39; ou (ii) SEQ ID NO: 40 e qualquer uma de (a) SEQ ID NO: 37, (b) SEQ ID NO: 38 e (c) SEQ ID NO: 39; ou (iii) SEQ ID NO: 41 ou SEQ ID NO: 42 ou SEQ ID NO: 43 ou SEQ ID NO: 44. Em certas concretizações, o polipeptídeo possui atividade descarboxilase.

[027] Em outro aspecto, a invenção apresenta um método de produção de um hidrocarboneto, o método compreendendo a expressão em uma célula hospedeira de um polinucleotídeo compreendendo uma sequência de nucleotídeo contendo pelo menos cerca de 70%, pelo menos cerca de 75%, pelo menos cerca de 80%, pelo menos cerca de 85%, pelo menos cerca de 90% pelo menos cerca de 91%, pelo menos cerca de 92%, pelo menos cerca de 93%, pelo menos cerca de 94%, pelo menos cerca de 95%, pelo menos cerca de 96%, pelo menos cerca de 97%, pelo menos cerca de 98% ou pelo menos cerca de 99% de identificação de sequência com SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33 ou 35. Em algumas concretizações, a sequência de nucleotídeo é SEQ ID NO: 1,

3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33 ou 35. Em algumas concretizações, o método ainda compreende isolar o hidrocarboneto da célula hospedeira.

[028] Em outras concretizações, a sequência de nucleotídeo hibridiza-se a um complemento de SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33 ou 35, ou a um fragmento desta, por exemplo, sob condições de baixa estringência, média estringência, alta estringência ou muito alta estringência.

[029] Em outras concretizações, a sequência de nucleotídeo codifica um polipeptídeo contendo: (i) a sequência de aminoácido de SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34 ou 36; ou a sequência de aminoácido de SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34 ou 36 com uma ou mais substituições, adições, inserções ou supressões de aminoácidos. Em algumas concretizações, o polipeptídeo compreende a sequência de aminoácido de SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34 ou 36 com uma ou mais substituições conservadoras de aminoácidos. Em algumas concretizações, o polipeptídeo possui atividade descarboxilase.

[030] Em outras concretizações, a sequência de nucleotídeo codifica um polipeptídeo contendo a mesma atividade biológica que um polipeptídeo compreendendo a sequência de aminoácido de SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34 ou 36. Em algumas concretizações, a sequência de nucleotídeo é SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33 ou 35 ou um fragmento desta. Em outras concretizações, a

sequência de nucleotídeo hibridiza-se a um complemento de SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33 ou 35 ou a um fragmento desta, por exemplo, sob condições de baixa estringência, média estringência, alta estringência ou muito alta estringência. Em algumas concretizações, a atividade biológica é atividade descarboxilase.

[031] Em algumas concretizações, o método compreende a transformação de uma célula hospedeira com um vetor recombinante compreendendo uma sequência de nucleotídeo contendo pelo menos cerca de 70%, pelo menos cerca de 75%, pelo menos cerca de 80%, pelo menos cerca de 85%, pelo menos cerca de 90% pelo menos cerca de 91%, pelo menos cerca de 92%, pelo menos cerca de 93%, pelo menos cerca de 94%, pelo menos cerca de 95%, pelo menos cerca de 96%, pelo menos cerca de 97%, pelo menos cerca de 98% ou pelo menos cerca de 99% de identificação de sequência com SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33 ou 35. Em algumas concretizações, o vetor recombinante ainda compreende um promotor operacionável ligado à sequência de nucleotídeos. Em algumas concretizações, o promotor é um promotor de desenvolvimento regulado, um específico de organela, um específico de tecido, um induzível, um constitutivo ou um específico de célula. Em concretizações particulares, o vetor recombinante compreende pelo menos uma sequência selecionada a partir do grupo composto de (a) uma sequência regulatória operacionalmente acoplada à sequência de nucleotídeos; (b) um marcador de seleção operacionável acoplado à sequência de nucleotídeo; (c) uma sequência de marcador operacionável acoplada à sequência de nucleotídeo;

(d) uma fração de purificação operacionável acoplada à sequência de nucleotídeo; (e) uma sequência de secreção operacionável acoplada à sequência de nucleotídeo; e (f) uma sequência alvo operacionável acoplada à sequência de nucleotídeo. Em certas concretizações, a sequência de nucleotídeo é estavelmente incorporada dentro do DNA genômico da célula hospedeira, e a expressão da sequência de nucleotídeo está sob controle de uma região promotora regulada.

[032] Em qualquer um dos aspectos descritos neste documento, a célula hospedeira pode ser selecionada a partir do grupo composto de uma célula de mamífero, uma célula vegetal, uma célula de inseto, uma célula de levedura, uma célula de fungo, célula de fungos filamentosos ou célula bacteriana.

[033] Em algumas concretizações, a célula hospedeira é uma célula bacteriana gram-positiva. Em outras concretizações, a célula hospedeira é uma célula bacteriana gram-negativa.

[034] Em algumas concretizações, a célula hospedeira é selecionada a partir do gênero *Escherichia*, *Bacillus*, *Lactobacillus*, *Rhodococcus*, *Pseudomonas*, *Aspergillus*, *Trichoderma*, *Neurospora*, *Fusarium*, *Humicola*, *Rhizomucor*, *Kluyveromyces*, *Pichia*, *Mucor*, *Myceliophthora*, *Penicillium*, *Phanerochaete*, *Pleurotus*, *Trametes*, *Chrysosporium*, *Saccharomyces*, *Stenotrophomonas*, *Schizosaccharomyces*, *Yarrowia*, ou *Streptomyces*.

[035] Em concretizações particulares, a célula hospedeira é uma célula de *Bacillus lentus*, uma célula de *Bacillus brevis*, uma célula de *Bacillus stearothermophilus*,

uma célula de *Bacillus licheniformis*, uma célula de *Bacillus alkalophilus*, uma célula de *Bacillus coagulans*, uma célula de *Bacillus circulans*, uma célula de *Bacillus pumilis*, uma célula de *Bacillus thuringiensis*, uma célula de *Bacillus clausii*, uma célula de *Bacillus megaterium*, uma célula de *Bacillus subtilis* ou uma célula de *Bacillus amyloliquefaciens*.

[036] Em outras concretizações, a célula hospedeira é uma célula de *Trichoderma koningii*, uma célula de *Trichoderma viride*, uma célula de *Trichoderma reesei*, uma célula de *Trichoderma longibrachiatum*, uma célula de *Aspergillus nidulans*, uma célula de *Aspergillus niger*, uma célula de *Aspergillus oryzae*, uma célula de *Humicola insolens*, uma célula de *Humicola lanuginosa*, uma célula de *Rhodococcus opacus*, uma célula de *Rhizomucor miehei* ou uma célula de *Mucor michei*.

[037] Ainda em outras concretizações, a célula hospedeira é uma célula de *Streptomyces lividans* ou uma célula de *Streptomyces murinus*. Em outras concretizações, a célula hospedeira é uma célula de *Actinomycetes*.

[038] Em algumas concretizações, a célula hospedeira é uma célula CHO, uma célula COS, uma célula VERO, uma célula BHK, uma célula HeLa, uma célula Cv1, uma célula MDCK, uma célula 293, uma célula 3T3 ou uma célula PC12.

[039] Em concretizações particulares, a célula hospedeira é uma célula de *E. coli*, tais como uma de linhagem B, uma de linhagem C, uma de linhagem K ou uma célula de *E. coli* de linhagem W.

[040] Em outras concretizações, a célula hospedeira é uma célula hospedeira cianobacteriana. Em concretizações

particulares, a célula hospedeira cianobacteriana é uma célula listada na Tabela 1.

[041] Em algumas concretizações, o hidrocarboneto é secretado pela célula hospedeira.

[042] Em determinadas concretizações, a célula hospedeira sobrepresa um substrato descrito neste documento. Em algumas concretizações, o método ainda inclui a transformação da célula hospedeira com um ácido nucleico que codifique uma enzima descrita neste documento, e a célula hospedeira sobrepresse um substrato descrito aqui. Em outras concretizações, o método ainda inclui a cultura da célula hospedeira na presença de pelo menos um substrato descrito neste documento. Em algumas concretizações, o substrato é um derivado de ácido graxo, um acil-ACP, um ácido graxo, um acil-CoA, um aldeído graxo, um álcool graxo ou um éster graxo.

[043] Em algumas concretizações, o substrato derivado de ácido graxo é um substrato insaturado de derivado de ácido graxo, um substrato monoinsaturado de derivado de ácido graxo ou um substrato saturado de derivado de ácido graxo. Em outras concretizações, o substrato derivado de ácido graxo é uma cadeia linear de substrato derivado de ácido graxo, uma cadeia ramificada de substrato derivado de ácido graxo ou um substrato derivado de ácido graxo que inclua uma fração cíclica.

[044] Em certas concretizações dos aspectos descritos neste documento, o hidrocarboneto produzido é um alcano. Em algumas concretizações, o alcano é um alcano C₃-C₂₅. Por exemplo, o alcano é um alcano C₃, C₄, C₅, C₆, C₇, C₈, C₉, C₁₀, C₁₁, C₁₂, C₁₃, C₁₄, C₁₅, C₁₆, C₁₇, C₁₈, C₁₉, C₂₀, C₂₁, C₂₂, C₂₃, C₂₄

ou C₂₅. Em algumas concretizações, o alcano é tridecano, metiltridecano, nonadecano, metilnonadecano, heptadecano, metil-heptadecano, pentadecano ou metilpentadecano.

[045] Em algumas concretizações, o alcano é um alcano de cadeia linear, um alcano de cadeia ramificada ou um alcano cíclico.

[046] Em determinadas concretizações, o método ainda compreende a cultura da célula hospedeira na presença de um derivado de ácido graxo saturado, e o hidrocarboneto produzido é um alcano. Em certas concretizações, o derivado de ácido graxo saturado é um substrato derivado de ácido graxo C₆-C₂₆. Por exemplo, o substrato derivado de ácido graxo é um substrato derivado de ácido graxo C₆, C₇, C₈, C₉, C₁₀, C₁₁, C₁₂, C₁₃, C₁₄, C₁₅, C₁₆, C₁₇, C₁₈, C₁₉, C₂₀, C₂₁, C₂₂, C₂₃, C₂₄, C₂₅ ou C₂₆. Em concretizações particulares, o substrato derivado de ácido graxo é 2-metilicosanal, icosanal, octadecanal, tetradecanal, 2-metiloctadecanal, estearaldeído ou palmitaldeído.

[047] Em algumas concretizações, o método ainda inclui o isolamento do alcano a partir da célula hospedeira ou do meio de cultura. Em outras concretizações, o método ainda inclui o craqueamento ou o refino do alcano.

[048] Em certas concretizações dos aspectos descritos neste documento, o hidrocarboneto produzido é um alceno. Em algumas concretizações, o alceno é um alceno C₃-C₂₅. Por exemplo, o alceno é um alceno C₃, C₄, C₅, C₆, C₇, C₈, C₉, C₁₀, C₁₁, C₁₂, C₁₃, C₁₄, C₁₅, C₁₆, C₁₇, C₁₈, C₁₉, C₂₀, C₂₁, C₂₂, C₂₃, C₂₄ ou C₂₅. Em algumas concretizações, o alceno é pentadeceno, heptadeceno, metilpentadeceno, ou metil-heptadeceno.

[049] Em algumas concretizações, o alceno é um alceno de

cadeia linear, um alceno de cadeia ramificada ou um alceno cíclico.

[050] Em certas concretizações, o método ainda compreende a cultura da célula hospedeira na presença de um derivado de ácido graxo insaturado, e o hidrocarboneto produzido é um alceno. Em certas concretizações, o derivado de ácido graxo insaturado é um substrato derivado de ácido graxo C₆-C₂₆. Por exemplo, o substrato derivado de ácido graxo é um substrato derivado de ácido graxo C₆, C₇, C₈, C₉, C₁₀, C₁₁, C₁₂, C₁₃, C₁₄, C₁₅, C₁₆, C₁₇, C₁₈, C₁₉, C₂₀, C₂₁, C₂₂, C₂₃, C₂₄, C₂₅ ou C₂₆. Em concretizações particulares, o substrato derivado de ácido graxo é octadecenal, hexadecenal, metilhexadecenal ou metiloctadecenal.

[051] Em outro aspecto, a invenção destaca um microrganismo geneticamente modificado compreendendo uma sequência de controle exógena estavelmente incorporada dentro do DNA genômico do microrganismo. Em uma concretização, a sequência de controle é integrada à montante a um polinucleotídeo compreendendo uma sequência de nucleotídeo contendo pelo menos cerca de 70% de identificação de sequência com SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33 ou 35. Em algumas concretizações, a sequência de nucleotídeo possui pelo menos cerca de 75%, pelo menos cerca de 80%, pelo menos cerca de 85%, pelo menos cerca de 90%, pelo menos cerca de 91%, pelo menos cerca de 92%, pelo menos cerca de 93%, pelo menos cerca de 94%, pelo menos cerca de 95%, pelo menos cerca de 96%, pelo menos cerca de 97%, pelo menos cerca de 98% ou pelo menos cerca de 99% de identificação de sequência com SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29,

31, 33 ou 35. Em algumas concretizações, a sequência de nucleotídeo é SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33 ou 35.

[052] Em algumas concretizações, o polinucleotídeo é endógeno ao microrganismo. Em algumas concretizações, o microrganismo expressa um nível elevado de um hidrocarboneto em relação a um tipo selvagem de microrganismo. Em algumas concretizações, o microrganismo é uma cianobactéria.

[053] Em outro aspecto, a invenção apresenta um método de preparo de um hidrocarboneto, o método compreendendo a cultura de um microrganismo geneticamente modificado descrito neste documento sob condições adequadas para a expressão genética e o isolamento do hidrocarboneto.

[054] Em outro aspecto, a invenção destaca um método de preparo de um hidrocarboneto, compreendendo o contato de um substrato com (i) um polipeptídeo contendo pelo menos 70% de identificação com a sequência de aminoácido de SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34 ou 36, ou um variante desta; (ii) um polipeptídeo codificado por uma sequência de nucleotídeo contendo pelo menos 70% de identificação com SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33 ou 35, ou uma variante desta; (iii) um polipeptídeo compreendendo a sequência de aminoácido de SEQ ID NO: 37, 38 ou 39; (iv) um polipeptídeo compreendendo a sequência de aminoácido de SEQ ID NO: 40 e qualquer uma das (a) SEQ ID NO: 37, (b) SEQ ID NO: 38, e (c) SEQ ID NO: 29; ou (v) SEQ ID NO: 41, 42, 43 ou 44. Em algumas concretizações, o polipeptídeo possui atividade descarboxilase.

[055] Em algumas concretizações, o polipeptídeo tem pelo

menos cerca de 80%, pelo menos cerca de 85%, pelo menos cerca de 90%, pelo menos cerca de 91%, pelo menos cerca de 92%, pelo menos cerca de 93%, pelo menos cerca de 94%, pelo menos cerca de 95%, pelo menos cerca de 96%, pelo menos cerca de 97%, pelo menos cerca de 98% ou pelo menos cerca de 99% de identificação de sequência com SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34 ou 36. Em algumas concretizações, o polipeptídeo possui a sequência de aminoácido de SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34 ou 36.

[056] Em algumas concretizações, o polipeptídeo é codificado por uma sequência de nucleotídeo contendo pelo menos cerca de 75%, pelo menos cerca de 80%, pelo menos cerca de 85%, pelo menos cerca de 90%, pelo menos cerca de 91%, pelo menos cerca de 92%, pelo menos cerca de 93%, pelo menos cerca de 94%, pelo menos cerca de 95%, pelo menos cerca de 96%, pelo menos cerca de 97%, pelo menos cerca de 98% ou pelo menos cerca de 99% de identificação de sequência com SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33 ou 35. Em algumas concretizações, o polipeptídeo é codificado por uma sequência de nucleotídeo contendo SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33 ou 35.

[057] Em algumas concretizações, o substrato biológico é um derivado de ácido graxo, um acil-ACP, um ácido graxo, um acil-CoA, um aldeído graxo, um álcool graxo ou um éster graxo.

[058] Em algumas concretizações, o substrato é derivado de ácido graxo saturado, e o hidrocarboneto é um alceno, por exemplo, um alceno C₃-C₂₅. Por exemplo, o alceno é um alceno

C₃, C₄, C₅, C₆, C₇, C₈, C₉, C₁₀, C₁₁, C₁₂, C₁₃, C₁₄, C₁₅, C₁₆, C₁₇, C₁₈, C₁₉, C₂₀, C₂₁, C₂₂, C₂₃, C₂₄ ou C₂₅. Em algumas concretizações, o alcano é tridecano, metiltridecano, nonadecano, metilnonadecano, heptadecano, metilheptadecano, pentadecano ou metilpentadecano.

[059] Em algumas concretizações, o alcano é um alcano de cadeia linear, um alcano de cadeia ramificada ou um alcano cíclico.

[060] Em algumas concretizações, o substrato derivado de ácido graxo é 2-metilicosanal, icosanal, octadecanal, tetradecanal, 2-metiloctadecanal, estearaldeído ou palmitaldeído.

[061] Em outras concretizações, o substrato biológico é um derivado de ácido graxo insaturado, e o hidrocarboneto é um alceno, por exemplo, um alceno C₃-C₂₅. Por exemplo, o alceno é um alceno C₃, C₄, C₅, C₆, C₇, C₈, C₉, C₁₀, C₁₁, C₁₂, C₁₃, C₁₄, C₁₅, C₁₆, C₁₇, C₁₈, C₁₉, C₂₀, C₂₁, C₂₂, C₂₃, C₂₄ ou C₂₅. Em algumas concretizações, o alceno é pentadeceno, heptadeceno, metilpentadeceno, ou metil-heptadeceno.

[062] Em algumas concretizações, o alceno é um alceno de cadeia linear, um alceno de cadeia ramificada ou um alceno cíclico.

[063] Em algumas concretizações, o derivado de ácido graxo insaturado é octadecenal, hexadecenal, metilhexadecenal ou metiloctadecenal.

[064] Em outro aspecto, a invenção destaca um hidrocarboneto produzido através de qualquer um dos métodos ou microrganismos descritos neste documento. Em concretizações particulares, o hidrocarboneto é um alcano ou um alceno contendo um $\delta^{13}\text{C}$ de cerca de -15,4 ou maior. Por

exemplo, o alceno ou o alceno tem um $\delta^{13}\text{C}$ de cerca de -15,4 a cerca de -10,9, por exemplo, de cerca de -13,92 a cerca de -13,84. Em outras concretizações, o alceno ou o alceno tem um $f_{\text{M}}^{14}\text{C}$ de pelo menos cerca de 1,003. Por exemplo, o alceno ou o alceno tem um $f_{\text{M}}^{14}\text{C}$ de pelo menos cerca de 1,01 ou pelo menos cerca de 1,5. Em algumas concretizações, o alceno ou o alceno tem um $f_{\text{M}}^{14}\text{C}$ de cerca de 1,111 a cerca de 1,124.

[065] Em outro aspecto a invenção destaca um biocombustível que inclua um hidrocarboneto produzido por qualquer um dos métodos ou microrganismos descritos neste documento. Em concretizações particulares, o hidrocarboneto é um alceno ou um alceno contendo um $\delta^{13}\text{C}$ de cerca de -15,4 ou maior. Por exemplo, o alceno ou alceno possui um $\delta^{13}\text{C}$ de cerca de -15,4 a cerca de -10,9, por exemplo, de cerca de -13,92 a cerca de -13,84. Em outras concretizações, o alceno ou alceno possui um $f_{\text{M}}^{14}\text{C}$ de pelo menos cerca de 1,003. Por exemplo, o alceno ou alceno possui um $f_{\text{M}}^{14}\text{C}$ de pelo menos cerca de 1,01 ou pelo menos cerca de 1,5. Em algumas concretizações, o alceno ou alceno possui um $f_{\text{M}}^{14}\text{C}$ de cerca de 1,111 a cerca de 1,124. Em algumas concretizações, o biocombustível é *diesel*, gasolina ou combustível de jatos.

[066] Em outro aspecto, a invenção destaca um ácido nucleico isolado consistindo de não mais do que cerca de 500 nucleotídeos de SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33 ou 35. Em algumas concretizações, o ácido nucleico consiste de não mais do que cerca de 300 nucleotídeos, não mais do que cerca de 350 nucleotídeos, não mais do que cerca de 400 nucleotídeos, não mais do que cerca de 450 nucleotídeos, não mais do que cerca de 550 nucleotídeos, não mais do que cerca de 600

nucleotídeos, ou não mais do que cerca de 650 nucleotídeos de SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33 ou 35. Em algumas concretizações, o ácido nucleico codifica um polipeptídeo contendo atividade descarbonilase.

[067] Em outro aspecto, a invenção destaca um ácido nucleico isolado consistindo em não mais do que cerca de 99%, não mais do que cerca de 98%, não mais do que cerca de 97%, não mais do que cerca de 96%, não mais do que cerca de 95%, não mais do que cerca de 94%, não mais do que cerca de 93%, não mais do que cerca de 92%, não mais do que cerca de 91%, não mais do que cerca de 90%, não mais do que cerca de 85%, ou não mais do que cerca de 80% de nucleotídeos de SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33 ou 35. Em algumas concretizações, o ácido nucleico codifica um polipeptídeo contendo atividade descarbonilase.

[068] Em outro aspecto, a invenção destaca um polipeptídeo isolado consistindo em não mais do que cerca de 200, não mais do que cerca de 175, não mais do que cerca de 150 ou não mais do que cerca de 100 dos aminoácidos de SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34 ou 36. Em algumas concretizações, o polipeptídeo possui atividade descarbonilase.

[069] Em outro aspecto, a invenção destaca um polipeptídeo isolado consistindo em não mais do que cerca de 99%, não mais do que cerca de 98%, não mais do que cerca de 97%, não mais do que cerca de 96%, não mais do que cerca de 95%, não mais do que cerca de 94%, não mais do que cerca de 93%, não mais do que cerca de 92%, não mais do que cerca de 91%, não mais do que cerca de 90%, não mais do que cerca de

85%, ou não mais do que cerca de 80% dos aminoácidos de SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34 ou 36. Em algumas concretizações, o polipeptídeo possui atividade descarbonilase.

DEFINIÇÕES

[070] Ao longo do relatório descritivo, uma referência pode ser feita usando um nome de gene ou de polipeptídeo abreviado, mas entende-se que tal nome abreviado de gene ou de polipeptídeo representa ao gênero dos genes ou dos polipeptídeos. Tais nomes de genes incluem todos os genes que codificam o mesmo polipeptídeo e os polipeptídeos homólogos contendo a mesma função fisiológica. Os nomes de polipeptídeos incluem todos os polipeptídeos que possuem a mesma atividade (por exemplo, que catalisem a mesma reação química essencial).

[071] Os números de acesso referidos neste documento são derivados a partir da base de dados do NCBI (*National Center for Biotechnology Information*), mantida pelo *National Institute of Health*, EUA. Salvo indicação em contrário, os números de acesso são como fornecidos na base de dados de abril de 2009.

[072] Os números EC são estabelecidos pelo Comitê de Nomenclatura da União Internacional de Bioquímica e Biologia Molecular (NC-IUBMB) (disponível em <http://www.chem.qmul.ac.uk/iubmb/enzyme/>). Os números EC referidos neste documento são derivados a partir da base de dados do *KEGG Ligand*, mantida pela Enciclopédia Quioto de Genes e Genômica, patrocinada em parte pela Universidade de Tóquio. Salvo indicação em contrário, os números EC são como fornecidos na base de dados de março de 2008.

[073] Os artigos "um" e "uma" são usados neste documento para se referirem a um ou mais de um (*i.e.*, a pelo menos um) dos objetos gramaticais do artigo. A título de exemplo, "um elemento" significa um elemento ou mais de um elemento.

[074] O termo "cerca de" é usado neste documento para significar um valor de $\pm 20\%$ de um valor numérico dado. Assim, "cerca de 60%" significa um valor entre $60 \pm (20\% \text{ de } 60)$ (*i.e.*, entre 48 e 70).

[075] Como usado neste documento, o termo "aldeído" significa um hidrocarboneto contendo a fórmula RCHO caracterizada por um grupo carbonila insaturado (C=O). Em uma concretização preferida, o aldeído é qualquer aldeído feito a partir de um ácido graxo ou de um derivado de ácido graxo. Em uma concretização, o grupo R é de pelo menos cerca de 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 ou 20 carbonos de comprimento.

[076] Como usado neste documento, um "gene biossintético de aldeído" ou um "polinucleotídeo biossintético de aldeído" é um ácido nucleico que codifica um polipeptídeo biossintético de aldeído.

[077] Como usado neste documento, um "polipeptídeo biossintético de aldeído" é um polipeptídeo que seja uma parte da via biossintética de um aldeído. Tais polipeptídeos podem agir sobre um substrato biológico para produzir um aldeído. Em algumas instâncias, o polipeptídeo biossintético de aldeído possui atividade redutase.

[078] Como usado neste documento, o termo "alcano" significa um hidrocarboneto contendo apenas ligações simples de carbono-carbono.

[079] Como usado neste documento, um "gene biossintético

de alceno" ou um "polinucleotídeo biossintético de alceno" é um ácido nucleico que codifica um polipeptídeo biossintético de alceno.

[080] Como usado neste documento, um "polipeptídeo biossintético de alceno" é um polipeptídeo que seja uma parte da via biossintética de um alceno. Tais polipeptídeos podem agir sobre um substrato biológico para produzir um alceno. Em algumas instâncias, o polipeptídeo biossintético de alceno possui atividade descarboxilase.

[081] Como usado neste documento, um "gene biossintético de alceno" ou um "polinucleotídeo biossintético de alceno" é um ácido nucleico que codifica um polipeptídeo biossintético de alceno.

[082] Como usado neste documento, um "polipeptídeo biossintético de alceno" é um polipeptídeo que seja uma parte da via biossintética de um alceno. Tais polipeptídeos podem agir sobre um substrato biológico para produzir um alceno. Em algumas instâncias, o polipeptídeo biossintético de alceno possui atividade descarboxilase.

[083] Como usado neste documento, o termo "atenuar" significa enfraquecer, reduzir ou diminuir. Por exemplo, um polipeptídeo pode ser atenuado modificando-se o polipeptídeo para reduzir sua atividade (por exemplo, modificando uma sequência de nucleotídeo que codifique o polipeptídeo).

[084] Como usado neste documento, o termo "biodiesel", significa um biocombustível que pode ser um substituto do *diesel*, o qual é derivado do petróleo. O biodiesel pode ser usado em motores de combustão interna a *diesel* tanto de uma forma pura, a qual é referida como biodiesel "puro", quanto como uma mistura em qualquer concentração com *diesel* à base

de petróleo. O biodiesel pode incluir ésteres ou hidrocarbonetos, tais como aldeídos e alcanos.

[085] Como usado neste documento, o termo "biocombustível" refere-se a qualquer combustível derivado a partir de biomassa. Os biocombustíveis podem ser substituídos por combustíveis à base de petróleo. Por exemplo, biocombustíveis são, inclusive, combustíveis de transportes (por exemplo, gasolina, *diesel*, combustível de jatos etc.), combustíveis de aquecimento e combustíveis geradores de eletricidade. Os biocombustíveis são uma fonte renovável de energia.

[086] Como usado neste documento, o termo "biomassa" refere-se a uma fonte de carbono derivada a partir de material biológico. A biomassa pode ser convertida em biocombustível. Uma fonte exemplar de biomassa é matéria de planta. Por exemplo, milho, cana-de-açúcar ou *switchgrass* podem ser usados como biomassa. Outro exemplo não limitante de biomassa é matéria animal, por exemplo, estrume de vaca. A biomassa também inclui produtos resíduos a partir da indústria, da agricultura, da silvicultura, das famílias. Exemplos de tais produtos resíduos que podem ser usados como biomassa são resíduos de fermentação, palha, madeira serrada, esgoto, lixo e restos de comida. A biomassa também inclui fontes de carbono, tais como carboidratos (por exemplo, monossacarídeos, dissacarídeos ou polissacarídeos).

[087] Como usado neste documento, o sintagma "fonte de carbono" refere-se a um substrato ou composto adequado para ser usado como uma fonte de carbono para o crescimento de célula procariótica ou eucariótica simples. As fontes de carbono podem estar em diversas formas, incluindo, mas não

se limitando a, polímeros, carboidratos, ácidos, álcoois, aldeídos, cetonas, aminoácidos, peptídeos e gases (por exemplo, CO e CO₂). Estas incluem, por exemplo, diversos monossacarídeos, tais como glicose, frutose, manose e galactose; oligossacarídeos, tais como fruto-oligossacarídeo e galacto-oligossacarídeo; polissacarídeos, tais como xilose e arabinose; dissacarídeos, tais como sacarose, maltose e turanose; material celulósico, tal como metil celulose e carboximetil celulose de sódio; ésteres de ácido graxo saturados ou insaturados; tais como succinato, lactato e acetato; álcoois, tais como etanol e suas misturas. A fonte de carbono pode também ser um produto da fotossíntese, incluindo, mas não se limitando a, glicose. Uma fonte de carbono preferida é a biomassa. Outra fonte de carbono preferida é a glicose.

[088] Como usado neste documento, um "aditivo de redução em ponto de turvação" é um aditivo adicionado a uma composição para diminuir ou baixar o ponto de turvação de uma solução.

[089] Como usado neste documento, o sintagma "ponto de turvação de um fluido" significa a temperatura na qual sólidos dissolvidos não estão mais completamente solúveis. Abaixo desta temperatura, os sólidos começam a precipitar como uma segunda fase, dando ao fluido uma aparência turva. Na indústria do petróleo, o ponto de turvação se refere à temperatura abaixo a qual um material solidificado ou outro hidrocarboneto pesado se cristaliza em óleo cru, óleo refinado ou combustível para formar uma aparência turva. A presença de materiais solidificados influencia no comportamento seguinte do fluido, na tendência de o fluido

entupir os filtros de combustíveis, injetores, etc., na acumulação de materiais solidificados sobre superfícies frias (por exemplo, incrustar um gasoduto ou permutador de aquecimento), e nas características de emulsão do fluido com água.

[090] Uma sequência de nucleotídeo é "complementar" à outra sequência de nucleotídeo se cada uma das bases das duas sequências combinarem (*i.e.*, for capaz de formar pares de base Watson Crick). O termo "filamento complementar" é usado neste documento alternadamente com o termo "complemento". O complemento de um filamento de ácido nucleico pode ser o complemento de um filamento de codificação ou o complemento de um filamento de não codificação.

[091] Como usado neste documento, o termo "condições suficientes para permitir expressão" significa quaisquer condições que permitam a uma célula hospedeira produzir um produto desejado, tal como um polipeptídeo, um aldeído ou um alceno descrito neste documento. Condições adequadas incluem, por exemplo, condições de fermentação. Condições de fermentação podem compreender muitos parâmetros, tais como faixas de temperaturas, níveis de aeração e meios de composição. Cada uma dessas condições, individualmente e em combinação, permite à célula hospedeira crescer. O meio de cultura exemplar inclui caldos e géis. Geralmente, o meio inclui uma fonte de carbono, tal como glicose, frutose, celulose, ou semelhantes, que pode ser metabolizado diretamente pela célula hospedeira. Além disso, enzimas podem ser usadas no meio para facilitar a mobilização (por exemplo, a despolimerização do amido ou da celulose de

açúcares fermentáveis) e o subsequente metabolismo da fonte de carbono.

[092] Para determinar se as condições são suficientes para permitir expressão, a célula hospedeira pode ser cultivada, por exemplo, por cerca de 4, 8, 12, 24, 36 ou 48 horas. Durante e/ou após a cultura, amostras podem ser obtidas e analisadas para determinar se as condições permitem expressão. Por exemplo, as células hospedeiras nas amostras ou no meio no qual as células hospedeiras cresceram podem ser testadas para a presença de um produto desejado. Durante o teste para a presença de um produto, análises, tais como, mas não limitadas a, TLC, HPLC, GC/FID, GC/MS, LC/MS e MS podem ser usadas.

[093] Entende-se que os polipeptídeos descritos neste documento devem conter conservante adicional ou substituições de aminoácido não essencial, o qual não possui um efeito substancial sobre as funções do polipeptídeo. Se uma substituição particular for tolerada ou não (*i.e.*, não adversamente afetará as propriedades biológicas desejadas, tais como a atividade descarbonilase) pode ser determinada como descrito em Bowie *et al.*, *Science* (1990) 247:1306-1310. Uma "substituição conservadora de aminoácido" é uma na qual o resíduo de aminoácido é substituído com um resíduo de aminoácido contendo uma cadeia lateral semelhante. Famílias de resíduos de aminoácidos contendo cadeias laterais semelhantes foram definidas no estado da técnica. Essas famílias incluem aminoácidos com cadeias laterais básicas (por exemplo, lisina, arginina, histidina), cadeias laterais ácidas (por exemplo, ácido aspártico, ácido glutâmico), cadeias laterais de polos descarregados (por exemplo,

glicina, asparagina, glutamina, serina, treonina, tirosina, cisteína), cadeias laterais não polares (por exemplo, alanina, valina, leucina, isoleucina, prolina, fenilalanina, metionina, triptofano), cadeias laterais betarramificadas (por exemplo, treonina, valina, isoleucina), e cadeias laterais aromáticas (por exemplo, tirosina, fenilalanina, triptofano, histidina).

[094] Como usado neste documento, "elemento de controle" significa um elemento de controle transcricional. Elementos de controle incluem promotores e potenciadores. O termo "elemento promotor", "promotor" ou "sequência promotora" refere-se a uma sequência de DNA que funcione com um interruptor que ative a expressão de um gene. Se o gene é ativado, diz-se que é transcrito ou participa de transcrição. A transcrição envolve a síntese de mRNA a partir do gene. Um promotor, portanto, serve como um elemento regulatório transcricional e também fornece um local para a iniciação da transcrição do gene em mRNA. Elementos de controle interagem especificamente com proteínas celulares envolvidas na transcrição (MANIATIS *et al.*, *Science* 236:1237, 1987).

[095] Como usado neste documento, o termo "sintase de éster" significa um peptídeo capaz de produzir ésteres graxos. Mais especificamente, uma sintase de éster é um peptídeo que converte um tioéster em um éster graxo. Em uma concretização preferida, a sintase de éster converte um tioéster (por exemplo, acil-CoA) em um éster graxo.

[096] Em uma concretização alternativa, uma sintase de éster utiliza um tioéster e um álcool como substratos para produzir um éster graxo. Sintases de éster são capazes de usar tioésteres de cadeias curtas e longas como substratos.

Além disso, sintases de éster são capazes de usar álcoois de cadeias curtas e longas como substratos.

[097] Exemplos não limitantes de sintases de éster são sintases de cera, sintases de éster de cera, acil CoA:transacilases de álcool, aciltransferases e acil coenzima A graxo:aciltransferases de álcool graxo. Sintases de éster exemplares são classificadas em número de classificação enzimática EC 2,3,1,75. Exemplares de número de acesso GenBank são fornecidos na Tabela A.

Tabela A [Fig. 40]. Identifica os diversos genes que podem ser expressados, sobreexpressados ou atenuados para aumentar a produção de substratos particulares.

Números de Acesso são do NCBI, GenBank, Edição 159.0 de 15 de abril de 2007

Números EC são de KEGG, Edição 42.0 de abril de 2007 (mais atualizações diárias até março de 2008)

<u>CATEGORIA</u>	<u>GENE</u>	<u>NOME</u>	<u>ACESSO</u>	<u>NÚMERO EC</u>	<u>MODIFICAÇÃO</u>	<u>USO</u>	<u>ORGANISMO</u>
1. Aumento na produção de ácido graxo/aumento na produção de produto							
aumenta acil-CoA							
reduz catabolismo de derivados e intermediários							
reduz inibição retroalimentada							
atenua outras vias que consomem ácidos graxos							
	accA	Acetil-CoA carboxilase, subunidade A (carboxiltransferase alfa)	AAC73296, NP 414727	6.4.1.2	Sobreexpressa	aumenta produção de Malonil-CoA	<i>Escherichia coli</i> , <i>Lactococci</i>
	accB	Acetil-CoA carboxilase, subunidade B (BCCP: proteína transportadora da biotina carboxilada)	NP 417721	6.4.1.2	Sobreexpressa	aumenta produção de Malonil-CoA	<i>Escherichia coli</i> , <i>Lactococci</i>
	accC	Acetil-CoA carboxilase, subunidade C (biotina carboxilase)	NP 417722	6.4.4.2, 6.3.4.14	Sobreexpressa	aumenta produção de Malonil-CoA	<i>Escherichia coli</i> , <i>Lactococci</i>
	accD	Acetil-CoA carboxilase, subunidade D (carboxiltransferase beta)	NP 416819	6.4.1.2	Sobreexpressa	aumenta produção de Malonil-CoA	<i>Escherichia coli</i> , <i>Lactococci</i>

aceE	piruvato desidrogenase, subunidade E1	NP_414656, AAC73226	1.2.4.1	Sobreexpressa	aumenta produção de Acetil-CoA	<i>Escherichia coli</i>
aceF	piruvato desidrogenase, subunidade E2	NP_414657	2.3.1.12	Sobreexpressa	aumenta produção de Acetil-CoA	<i>Escherichia coli</i>
ackA	acetato quinase	AAC75356, NP_416799	2.7.2.1	Elimina ou reduz	aumenta produção de Acetil-CoA	<i>Escherichia coli</i>
ackB	acetato quinase AckB	BAB81430	2.7.2.1	Elimina ou reduz	aumenta produção de Acetil-CoA	<i>Escherichia coli</i>
acpP	proteína transportadora de acila	AAC74178	NENHUM	Sobreexpressa	aumenta produção de Acetil-CoA	<i>Escherichia coli</i>
fadD	acil-CoA sintase	AP_002424	2.3.1.86, 6.2.1.3	Sobreexpressa	aumenta produção de ácido graxo	<i>Escherichia coli</i> W3110
adhE	álcool desidrogenase	CAA47743	1.1.1.1, 1.2.1.10	Elimina ou reduz	aumenta produção de Acetil-CoA	<i>Escherichia coli</i> W3111
cer1	aldeído descarboxilase	BAA11024	4.1.99.5	Sobreexpressa	aumenta produção de Acetil-CoA	<i>Arabidopsis thaliana</i>
fabA	beta-hidroxi-ácido de canoil tioéster desidrase	NP_415474	4.2.1.60	expressa	produção de acil-CoA graxo	<i>E. coli</i> K12
fabD	[proteína transportadora de acila] S-maloniltransferase	AAC74176	2.3.1.39	Sobreexpressa	aumenta produção de Acetil-CoA	<i>E. coli</i> K12
fabF	3-oxoacil-[proteína transportadora de acila] sintase II	AAC74179	2.3.1.179	Elimina ou Sobreexpressa	aumenta produção de Acetil-CoA	<i>E. coli</i> K12
fabG	3-oxoacil-[proteína transportadora de acila] redutase	AAC74177	1.1.1.100	Sobreexpressa	aumenta produção de Acetil-CoA	<i>E. coli</i> K12

	fabH	3-oxoacil-[proteína transportadora de acila] sintase III	AAC74175	2.3.1.180	Sobreexpressa	aumenta produção de Acetil-CoA	<i>E. coli</i> K12, <i>lactococci</i>
	fabI	enoil-[proteína transportadora de acila] redutase, dependente de NADH	NP_415804	1.3.1.9	expressa	produção de acil-CoA graxo	<i>E. coli</i> K12, <i>lactococci</i>
	fabR	repressor transcricional	NP_418398	NENHUM	Elimina ou reduz	modula produção de ácido graxo insaturado	<i>E. coli</i> K12
	fabZ	desidratase de proteína transportadora de (3R)-hidroximiristol acila	NP_414722	4.2.1.-			<i>E. coli</i> K12
	fadE	acil-CoA desidrogenase	AAC73325	1.3.99.3, 1.3.99.-	Elimina ou reduz	aumenta produção de Acetil-CoA	
	acr1	Acil-CoA graxo redutase	YP_047869, AAC45217	1.2.1.42	Sobreexpressa	para produção de álcool graxo	<i>Acinetobacter sp. i.e. calcoaceticus</i>
	GST, gshB	Glutaciona sintase	P04425	6.3.2.3	Elimina ou reduz	aumenta Acil-CoA	<i>E. coli</i> K12
	gpsA	sn-glicerol 3-fosfato desidrogenase biosintética	AAC76632, NP_418065	EC: 1.1.1.94	Elimina ou reduz	aumenta produção de Acetil-CoA	<i>E. coli</i> K12
	ldhA	lactato desidrogenase	AAC74462, NP_415898	EC: 1.1.1.27, 1.1.1.28	Elimina ou reduz	aumenta produção de Acetil-CoA	<i>E. coli</i> K12
	Lipase	Triglicerídeo lipase	CAA89087, CAA98876	3.1.1.3	expressa	aumenta produção de ácido graxo	<i>Saccharomyces cerevisiae</i>
		Malonil-CoA descarboxilase	AAA26500	4.1.1.9, 4.1.1.41	Sobreexpressa		<i>Saccharopolyspora erithraea</i>
	panD	aspartato 1-descarboxilase	BAB96708	4.1.1.11	Sobreexpressa	aumenta Acil-CoA	<i>Escherichia coli</i> W3110

	panK a.k.a. coaA	pantotenato quinase	AAC76952	2.7.1.33	Sobreexpressa	aumenta produção de Acetil-CoA	<i>E. coli</i>
	panK a.k.a. coaA, R106K	pantotenato quinase	AAC76952	2.7.1.33	Expressa, Sobreexpressa, mutação R106K	aumenta produção de Acetil-CoA	<i>E. coli</i>
	pdh	piruvato desidrogenase	BAB34380, AAC73226, NP 415392	1.2.4.1	Sobreexpressa	aumenta produção de Acetil-CoA	
	pflB	formato acetiltransferase (piruvato formato liase)	AAC73989, P09373	EC: 2.3.1.54	Elimina ou reduz	aumenta produção de Acetil-CoA	
	plsB	aciltransferase	AAC77011	2.3.1.15	mutação D311E	reduz limites no pool Acil-CoA	<i>E. coli K12</i>
	poxB	piruvato oxidase	AAC73958, NP 415392	1.2.2.2	Elimina ou reduz	aumenta produção de Acetil-CoA	
	pta	fosfotransacetilase	AAC75357, NP 416800	2.3.1.8	Elimina ou reduz	aumenta produção de Acetil-CoA	
	udhA	piridina nucleotídeo transidrogenase	CAA46822	1.6.1.1	Sobreexpressa	conversão NADH a NADPH ou vice-versa	
	fadB	3-hidroxi-butiril-CoA epimerase / delta(3)-cis-delta(2)-trans-enoil-CoA isomerase / enoil-CoA hidratase fundidas e 3-hidroxiacil-CoA desidrogenase	AP 003956	4.2.1.17, 5.1.2.3, 5.3.3.8, 1.1.1.35	Elimina ou reduz	Bloqueia a degradação de ácidos graxos	<i>E. coli</i>

	fadJ	3-hidroxiacil-CoA desidrogenase; K01692 enoil-CoA hidratase; K01782 3-hidroxibutiril-CoA epimerase	AAC75401	1.1.1.35, 4.2.1.17, 5.1.2.3	Elimina ou reduz	Bloqueia a degradação de ácidos graxos	<i>E. coli</i>
	fadA	3-cetoacil-CoA tiolase	BAE77458	2.3.1.16	Elimina ou reduz	Bloqueia a degradação de ácidos graxos	<i>E. coli</i>
	fadI	beta-cetoacil-CoA tiolase	AAC75402	2.3.1.16	Elimina ou reduz	Bloqueia a degradação de ácidos graxos	<i>E. coli</i>
	YdiO	acil-coA desidrogenase	YP_852786	1.3.99.-	Elimina ou reduz	Bloqueia a degradação de ácidos graxos	<i>E. coli</i>
2. Controle de Estrutura							
2A. Controle de Comprimento de Cadeia							
2	tesA	tioesterase	P0ADA1	3.1.2.-, 3.1.1.5	Elimina e/ou expressa	comprimento de cadeia C18	
	tesA sem sequência líder	tioesterase	AAC73596, NP_415027	3.1.2.-, 3.1.1.5	expressa ou sobreexpressa	C18:1	<i>E. coli</i>
	tesA sem sequência líder: L109P	tioesterase	P0ADA1	3.1.2.-, 3.1.1.5	Expressa e/ou sobreexpressa mutação L109P	comprimento de cadeia <C18	<i>E. coli</i>
	fatB1 (umbellularia)	tioesterase	Q41635	3.1.2.14	expressa ou sobreexpressa	C12:0	<i>Umbellularia californica</i>
	fatB2 (umbellularia)D ELETEumbelluria)	tioesterase	AAC49269	3.1.2.14	expressa ou sobreexpressa	C8:0 - C10:0	<i>Cuphea hookeriana</i>
	fatB3	tioesterase	AAC72881	3.1.2.14	expressa ou sobreexpressa	C14:0 - C16:0	<i>Cuphea hookeriana</i>
	fatB (cinnamomum)	tioesterase	Q39473	3.1.2.14	expressa ou sobreexpressa	C14:0	<i>Cinnamomum camphora</i>
	fatB[M141T]*	tioesterase	CAA85388	3.1.2.14	expressa ou sobreexpressa	C16:1	<i>Arabidopsis thaliana</i>
	fatA1 (Helianthus)	tioesterase	AAL79361	3.1.2.14	expressa ou sobreexpressa	C18:1	<i>Helianthus annuus</i>

	atfata (ARABIDOPSIS FATA ACYL-ACPTHTIOESTERASE)	tioesterase	NP_189147, NP_193041	3.1.2.14	expressa ou sobreexpressa	C18:1	<i>Arabidopsis thaliana</i>
	fatA	tioesterase	CAC39106	3.1.2.14	expressa ou sobreexpressa	C18:1	<i>Brassica juncea</i>
	fatA (cuphea)	tioesterase	AAC72883	3.1.2.14	expressa ou sobreexpressa	C18:1	<i>Cuphea hookeriana</i>
2B. Controle de Ramificação							
	atenua <i>FabH</i> expressa <i>FbH</i> de <i>S. glaucescens</i> ou <i>S. coelicolor</i> e knock out <i>FabH</i> endógeno					aumenta derivados de ácidos graxos de cadeia ramificada	
	expressa <i>FabH</i> de <i>B. subtilis</i> e knock out <i>FabH</i> endógeno						
	bdk - E3 - subunidade dihidroplipoil desidrogenase			EC 1.2.4.4			
	bkd - E1 - subunidade alfa/beta	subunidades descarboxilase do complexo cadeia-ramificada a-cetoácido desidrogenase		EC 1.2.4.4			
	bkd - E2 - subunidade dihidrolipoil transacilase			EC 1.2.4.4			
	bkdA1	a-subunidade de cadeia-ramificada a-cetoácido desidrogenase (E1a)	NP_628006	EC 1.2.4.4	expressa ou sobreexpressa	produz precursores de acil-CoA de cadeia ramificada	<i>Streptomyces coelicolor</i>

	bkdB1	a-subunidade de cadeia-ramificada a-cetoácido desidrogenase (E1b)	NP_628005	EC 1.2.4.4	expressa ou sobreexpressa	produz precursores de acil-CoA de cadeia ramificada	<i>Streptomyces coelicolor</i>
	bkdC1	diidrolipoil transacetilase (E2)	NP_628004	EC 2.3.1.168	expressa ou sobreexpressa	produz precursores de acil-CoA de cadeia ramificada	<i>Streptomyces coelicolor</i>
	bkdA2	a-subunidade de cadeia-ramificada a-cetoácido desidrogenase (E1a)	NP_733618	EC 1.2.4.4	expressa ou sobreexpressa	produz precursores de acil-CoA de cadeia ramificada	<i>Streptomyces coelicolor</i>
	bkdB2	b-subunidade de cadeia-ramificada a-cetoácido desidrogenase (E1b)	NP_628019	EC 1.2.4.4	expressa ou sobreexpressa	produz precursores de acil-CoA de cadeia ramificada	<i>Streptomyces coelicolor</i>
	bkdC2	diidrolipoil transacetilase (E2)	NP_628018	EC 2.3.1.168	expressa ou sobreexpressa	produz precursores de acil-CoA de cadeia ramificada	<i>Streptomyces coelicolor</i>
	bkdA	a-subunidade de cadeia-ramificada a-cetoácido desidrogenase (E1a)	BAC72074	EC 1.2.4.4	expressa ou sobreexpressa	produz precursores de acil-CoA de cadeia ramificada	<i>Streptomyces avermitilis</i>
	bkdB	b-subunidade de cadeia-ramificada a-cetoácido desidrogenase (E1b)	BAC72075	EC 1.2.4.4	expressa ou sobreexpressa	produz precursores de acil-CoA de cadeia ramificada	<i>Streptomyces avermitilis</i>

	bkdC	diidrolipoil transacetilase (E2)	BAC72076	EC 2.3.1.168	expressa ou sobreexpressa	produz precursores de acil-CoA de cadeia ramificada	<i>Streptomyces avermitilis</i>
	bkdF	a-subunidade de cadeia-ramificada a-cetoácido desidrogenase (E1a)	BAC72088	EC 1.2.4.4	expressa ou sobreexpressa	produz precursores de acil-CoA de cadeia ramificada	<i>Streptomyces avermitilis</i>
	bkdG	b-subunidade de cadeia-ramificada a-cetoácido desidrogenase (E1b)	BAC72089	EC 1.2.4.4	expressa ou sobreexpressa	produz precursores de acil-CoA de cadeia ramificada	<i>Streptomyces avermitilis</i>
	bkdH	diidrolipoil transacetilase (E2)	BAC72090	EC 2.3.1.168	expressa ou sobreexpressa	produz precursores de acil-CoA de cadeia ramificada	<i>Streptomyces avermitilis</i>
	bkdAA	a-subunidade de cadeia-ramificada a-cetoácido desidrogenase (E1a)	NP_390285	EC 1.2.4.4	expressa ou sobreexpressa	produz precursores de acil-CoA de cadeia ramificada	<i>Bacillus subtilis</i>
	bkdAB	b-subunidade de cadeia-ramificada a-cetoácido desidrogenase (E1b)	NP_390284	EC 1.2.4.4	expressa ou sobreexpressa	produz precursores de acil-CoA de cadeia ramificada	<i>Bacillus subtilis</i>
	bkdB	diidrolipoil transacetilase (E2)	NP_390283	EC 2.3.1.168	expressa ou sobreexpressa	produz precursores de acil-CoA de cadeia ramificada	<i>Bacillus subtilis</i>

	bkdA1	a-subunidade de cadeia-ramificada a-cetoácido desidrogenase (E1a)	AAA65614	EC 1.2.4.4	expressa ou sobreexpressa	produz precursores de acil-CoA de cadeia ramificada	<i>Pseudomonas putida</i>
	bkdA2	b-subunidade de cadeia-ramificada a-cetoácido desidrogenase (E1b)	AAA65615	EC 1.2.4.4	expressa ou sobreexpressa	produz precursores de acil-CoA de cadeia ramificada	<i>Pseudomonas putida</i>
	bkdC	diidrolipoil transacetilase (E2)	AAA65617	EC 2.3.1.168	expressa ou sobreexpressa	produz precursores de acil-CoA de cadeia ramificada	<i>Pseudomonas putida</i>
	lpd	diidrolipoamida desidrogenase (E3)	NP_414658	1.8.1.4	expressa ou sobreexpressa	produz precursores de acil-CoA de cadeia ramificada	<i>Escherichia coli</i>
	IlvE	aminotransferas e de aminoácido de cadeia ramificada	YP_026247	2.6.1.42	expressa ou sobreexpressa	produz a-cetoácidos ramificados	<i>Escherichia coli</i>
	IlvE	aminotransferas e de aminoácido de cadeia ramificada	AAF34406	2.6.1.42	expressa ou sobreexpressa	produz a-cetoácidos ramificados	<i>Lactococcus lactis</i>
	IlvE	aminotransferas e de aminoácido de cadeia ramificada	NP_745648	2.6.1.42	expressa ou sobreexpressa	produz a-cetoácidos ramificados	<i>Pseudomonas putida</i>
	IlvE	aminotransferas e de aminoácido de cadeia ramificada	NP_629657	2.6.1.42	expressa ou sobreexpressa	produz a-cetoácidos ramificados	<i>Streptomyces coelicolor</i>
	ccr	crotonil-CoA redutase	NP_630556	1.6.5.5, 1.1.1.1	expressa ou sobreexpressa	converte crotonil-CoA em butiril-CoA	<i>Streptomyces coelicolor</i>

	ccr	crotonil-CoA redutase	AAD53915	1.6.5.5, 1.1.1.1	expressa ou sobreexpressa	converte crotonil-CoA em butiril-CoA	<i>Streptomyces cinnamomensis</i>
	IcmA, isobutiril-CoA mutase	isobutiril-CoA mutase, subunidade A	NP_629554	5.4.99.2	expressa ou sobreexpressa	converte butiril-CoA em isobutiril-CoA	<i>Streptomyces coelicolor</i>
	IcmA, isobutiril-CoA mutase	isobutiril-CoA mutase, subunidade A	AAC08713	5.4.99.2	expressa ou sobreexpressa	converte butiril-CoA em isobutiril-CoA	<i>Streptomyces cinnamomensis</i>
	IcmB, isobutiril-CoA mutase	isobutiril-CoA mutase, subunidade B	NP_630904	5.4.99.2	expressa ou sobreexpressa	converte butiril-CoA em isobutiril-CoA	<i>Streptomyces coelicolor</i>
	IcmB, isobutiril-CoA mutase	isobutiril-CoA mutase, subunidade B	CAB59633	5.4.99.2	expressa ou sobreexpressa	converte butiril-CoA em isobutiril-CoA	<i>Streptomyces cinnamomensis</i>
	FabH, ACPs e genes fabF com especificidade para acil-CoAs de cadeia ramificada						
	IlvE	aminotransferase de aminoácido de cadeia ramificada	CAC12788	EC 2.6.1.42	sobreexpressa	aminotransferase de aminoácido de cadeia ramificada	<i>Staphylococcus carnosus</i>
	FabH1	beta-cetoacil-ACP sintase III	MP_626634	2.3.1.180	expressa ou sobreexpressa	iniciação da biosíntese de ácido graxo de cadeia ramificada	<i>Streptomyces coelicolor</i>

	ACP	proteína transportadora de acila	NP 626635	NENHUM	expressa ou sobreexpressa	iniciação e alongamento da biosíntese de ácido graxo de cadeia ramificada	<i>Streptomyces coelicolor</i>
	FabF	beta-cetoacil-ACP sintase II	NP 626636	2.3.1.179	expressa ou sobreexpressa	alongamento da biosíntese de ácido graxo de cadeia ramificada	<i>Streptomyces coelicolor</i>
	FabH3	beta-cetoacil-ACP sintase III	NP 823466	2.3.1.180	expressa ou sobreexpressa	iniciação da biosíntese de ácido graxo de cadeia ramificada	<i>Streptomyces avermitilis</i>
	FabC3 (ACP)	proteína transportadora de acila	NP 823467	NENHUM	expressa ou sobreexpressa	iniciação e alongamento da biosíntese de ácido graxo de cadeia ramificada	<i>Streptomyces avermitilis</i>
	FabF	beta-cetoacil-ACP sintase II	NP 823468	2.3.1.179	expressa ou sobreexpressa	alongamento da biosíntese de ácido graxo de cadeia ramificada	<i>Streptomyces avermitilis</i>
	FabH A	beta-cetoacil-ACP sintase III	NP 389015	2.3.1.180	expressa ou sobreexpressa	iniciação da biosíntese de ácido graxo de cadeia ramificada	<i>Bacillus subtilis</i>
	FabH B	beta-cetoacil-ACP sintase III	NP 388898	2.3.1.180	expressa ou sobreexpressa	iniciação da biosíntese de ácido graxo de cadeia ramificada	<i>Bacillus subtilis</i>

	ACP	proteína transportadora de acila	NP_389474	NENHUM	expressa ou sobreexpressa	iniciação e alongamento da biosíntese de ácido graxo de cadeia ramificada	<i>Bacillus subtilis</i>
	FabF	beta-cetoacil-ACP sintase II	NP_389016	2.3.1.179	expressa ou sobreexpressa	alongamento da biosíntese de ácido graxo de cadeia ramificada	<i>Bacillus subtilis</i>
	SmalDRAFT_0818	beta-cetoacil-ACP sintase III	ZP_01643059	2.3.1.180	expressa ou sobreexpressa	iniciação da biosíntese de ácido graxo de cadeia ramificada	<i>Stenotrophomonas maltophilia</i>
	SmalDRAFT_0821	proteína transportadora de acila	ZP_01643063	NENHUM	expressa ou sobreexpressa	iniciação e alongamento da biosíntese de ácido graxo de cadeia ramificada	<i>Stenotrophomonas maltophilia</i>
	SmalDRAFT_0822	beta-cetoacil-ACP sintase II	ZP_01643064	2.3.1.179	expressa ou sobreexpressa	alongamento da biosíntese de ácido graxo de cadeia ramificada	<i>Stenotrophomonas maltophilia</i>
	FabH	beta-cetoacil-ACP sintase III	YP_123672	2.3.1.180	expressa ou sobreexpressa	iniciação da biosíntese de ácido graxo de cadeia ramificada	<i>Legionella pneumophila</i>
	ACP	proteína transportadora de acila	YP_123675	NENHUM	expressa ou sobreexpressa	iniciação e alongamento da biosíntese de ácido graxo de cadeia ramificada	<i>Legionella pneumophila</i>

	FabF	beta-cetoacil-ACP sintase II	YP 123676	2.3.1.179	expressa ou sobreexpressa	alongamento da biosíntese de ácido graxo de cadeia ramificada	<i>Legionella pneumophila</i>
	FabH	beta-cetoacil-ACP sintase III	NP 415609	2.3.1.180	elimina ou reduz	iniciação da biosíntese de ácido graxo de cadeia ramificada	<i>Escherichia coli</i>
	FabF	beta-cetoacil-ACP sintase II	NP 415613	2.3.1.179	elimina ou reduz	alongamento da biosíntese de ácido graxo de cadeia ramificada	<i>Escherichia coli</i>
<i>Para produzir ácidos graxos cíclicos</i>							
	AnsJ	desidratase (suposto)	não disponível	não disponível	expressa ou sobreexpressa	biosíntese da ciclohexilcarbonil-CoA	<i>Streptomyces collinus</i>
	AnsK	CoA ligase (suposto)	não disponível	não disponível	expressa ou sobreexpressa	biosíntese da ciclohexilcarbonil-CoA	<i>Streptomyces collinus</i>
	AnsL	desidrogenase (suposto)	não disponível	não disponível	expressa ou sobreexpressa	biosíntese da ciclohexilcarbonil-CoA	<i>Streptomyces collinus</i>
	ChcA	enoil-CoA redutase	U72144	EC 1.3.1.34	expressa ou sobreexpressa	biosíntese da ciclohexilcarbonil-CoA	<i>Streptomyces collinus</i>
	AnsM	oxidoreductase (suposto)	não disponível	não disponível	expressa ou sobreexpressa	biosíntese da ciclohexilcarbonil-CoA	<i>Streptomyces collinus</i>
	PlmJ	desidratase (suposto)	AAQ84158	não disponível	expressa ou sobreexpressa	biosíntese da ciclohexilcarbonil-CoA	<i>Streptomyces sp. HK803</i>
	PlmK	CoA ligase (suposto)	AAQ84158	não disponível	expressa ou sobreexpressa	biosíntese da ciclohexilcarbonil-CoA	<i>Streptomyces sp. HK803</i>
	PlmL	desidrogenase (suposto)	AAQ84159	não disponível	expressa ou sobreexpressa	biosíntese da ciclohexilcarbonil-CoA	<i>Streptomyces sp. HK803</i>

	ChcA	enoil-CoA redutase	AAQ84160	EC 1.3.1.34	expressa ou sobreexpressa	biosíntese da ciclohexilcar bonil-CoA	<i>Streptomyces sp. HK803</i>
	PlmM	oxidoreductase (suposto)	AAQ84161	não disponível	expressa ou sobreexpressa	biosíntese da ciclohexilcar bonil-CoA	<i>Streptomyces sp. HK803</i>
	ChcB	enoil-CoA isomerase	AF268489	não disponível	expressa ou sobreexpressa	biosíntese da ciclohexilcar bonil-CoA	<i>Streptomyces collinus</i>
	ChcB/CaiD	enoil-CoA isomerase	NP 629292	4.2.1.-	expressa ou sobreexpressa	biosíntese da ciclohexilcar bonil-CoA	<i>Streptomyces coelicolor</i>
	ChcB/CaiD	enoil-CoA isomerase	NP 824296	4.2.1.-	expressa ou sobreexpressa	biosíntese da ciclohexilcar bonil-CoA	<i>Streptomyces avermitilis</i>
2C. Controle do Nível de Saturação							
	Sfa	Supressor de FabA	AAN79592, AAC44390	NENHUM	sobreexpressa	aumenta ácidos graxos monoinsaturados	<i>E. coli</i>
	ver também FabA na sec. 1				expressa	produz ácidos graxos insaturados	
	GnsA	supressores da mutação secG nulo	ABD18647.1	NENHUM	sobreexpressa	aumenta ácidos graxos insaturados	<i>E. coli</i>
	GnsB	supressores da mutação secG nulo	AAC74076.1	NENHUM	sobreexpressa	aumenta ácidos graxos insaturados	<i>E. coli</i>
	ver também seção 2A - itens com :0 são insaturados (nenhuma dupla ligação) e com :1 são saturadas (1 dupla ligação)						

	fabB	3-oxoacil-[proteína-transportadora-acila] sintase I	BAA16180	EC:2.3.1.41	sobreexpressa	modula produção de ácido graxo insaturado	<i>Escherichia coli</i>
	fabK	trans-2-enoil-ACP redutase II	AAF98273	1.3.1.9	expressa	modula produção de ácido graxo insaturado	<i>Streptococcus pneumoniae</i>
	fabL	enoil-(proteína transportadora de acila) redutase	AAU39821	1.3.1.9	expressa	modula produção de ácido graxo insaturado	<i>Bacillus licheniformis</i> DSM 13
	fabM	trans-2, cis-3-decenoil-ACP isomerase	DAA05501	4.2.1.17	sobreexpressa	modula produção de ácido graxo insaturado	<i>Streptococcus mutans</i>
3. Rendimento Final de Produto							
3A. Rendimento de ceras							
	AT3G51970	cadeia-longa-álcool O-graxo-aciltransferase	NP_190765	2.3.1.26	expressa	produção de ceras	<i>Arabidopsis thaliana</i>
		tioesterase (ver seção de controle de comprimento de cadeia)			expressa	aumenta produção de ácido graxo	
		formadora de álcool graxo acil-CoA redutase		1.1.1.*	expressa	converte acil-coa em álcool graxo	
	acr1	acil-CoA redutase (ACR1)	YP_047869	1.2.1.42	expressa	converte acil-coa em álcool graxo	<i>Acinetobacter sp. ADP1</i>
	yqhD	álcool desidrogenase	AP_003562	1.1.-.-	expressa	aumenta	<i>E. coli</i> W3110
	ELO1	ácido graxo elongase	BAD98251	2.3.1.-	expressa	produz ácidos graxos de comprimento de cadeia muito longo	<i>Pichia angusta</i>

	plsC	aciltransferase	AAA16514	2.3.1.51	expressa		<i>Saccharomyces cerevisiae</i>
	DAGAT/DGAT	diacilglicerol aciltransferase	AAF19262	2.3.1.20	expressa	produção de ceras	<i>Arabidopsis thaliana</i>
	hWS	acil-CoA cera álcool aciltransferase	AAX48018	2.3.1.20	expressa	produção de ceras	<i>Homo sapiens</i>
	aft1	cera ester sintase bifuncional/acil-CoA:diacilglicerol aciltransferase	AAO17391	2.3.1.20, 2.3.1.75	expressa	produção de ceras	<i>Acinetobacter sp. ADP1</i>
	mWS	cera ester sintase (simmondsia)	AAD38041	2.3.1.-, 2.3.1.75	expressa	produção de ceras	<i>Simmondsia chinensis</i>
<u>3B. Rendimento de Álcóois Graxos</u>							
		várias tioesterases (ver Sec. 2A)			expressa	produz	
	acr1	acil-CoA redutase	YP_047869	1.2.1.42	expressa	produz	<i>Acinetobacter sp. ADP1</i>
	yqhD	álcool desidrogenase	AP_003562	1.1.-.-	expressa	produz	<i>Escherichia coli W3110</i>
	BmFAR	FAR (formadora de álcool graxo acil-CoA redutase)	BAC79425	1.1.1.*	expressa	reduz acil-CoA graxo a álcool graxo	<i>Bombyx mori</i>
	Akrla4	redutase de aldeído microsomal de mamífero	NP_067448	1.1.1.2	expressa	produz	<i>Mus musculus</i>
	GTNG_1865	desidrogenase de aldeído de cadeia longa	YP_001125970	1.2.1.3	expressa	produz	<i>Geobacillus thermodenitrificans NG80-2</i>
	FadD	acil-CoA sintase	NP_416319	EC 6.2.1.3	expressa	produz mais	<i>E. coli K12</i>
<u>Para produzir Butanol</u>							
	atoB	acetil-CoA acetiltransferase	YP_049388	2.3.1.9	expressa	produz	<i>Erwinia carotovora</i>

	hbd	Beta-hidroxi-butiril-CoA desidrogenase	BAD51424	1.1.1.157	expressa	produz	<i>Butyrivibrio fibriosolvens</i>
	CPE0095	crotonase	BAB79801	4.2.1.55	expressa	produz	<i>Clostridium perfringens</i>
	bcd	butiril-CoA desidrogenase	AAM14583	1.3.99.2	expressa	produz	<i>Clostridium beijerinckii</i>
	ALDH	coenzima A-acilante aldeído desidrogenase	AAT66436	1.2.1.3	expressa	produz	<i>Clostridium beijerinckii</i>
	AdhE	aldeído-álcool desidrogenase	AAN80172	1.1.1.1 1.2.1.10	expressa	produz	<i>Escherichia coli</i> CFT073
3C. Rendimento de Éster de Ácido Graxo							
	tioesterase	ver seção de controle de comprimento de cadeia			expressa	produz	
	acr1	acil-CoA redutase	YP_047869	1.2.1.42	expressa	produz	<i>Acinetobacter sp.</i> ADP1
	yqhD	álcool desidrogenase	AP_003562	1.1.-.-	expressa	produz	<i>E. coli</i> K12
	AAT	álcool O-acetiltransferase	AAG13130	2.3.1.84	expressa	produz	<i>Fragaria x ananassa</i>
4. Exportação							
	Exportador cera éster (família FATP, proteína transportadora de ácido graxo (cadeia longa)		NP_524723	NENHUM	expressa	exporta cera	<i>Drosophila melanogaster</i>
	proteína de transporte ABC	suposto transportador de alcãno	AAN73268	NENHUM	expressa	exporta produtos	<i>Rhodococcus erythropolis</i>
	CER5	transportador de cera	Atlg51500, AY734542, At3g21090, Atlg51460	NENHUM	expressa	exporta produtos	<i>Arabidopsis thaliana</i>

	AtMRP5	Arabidopsis thaliana resistência multidrogas-associada	NP_171908	NENHUM	expressa	exporta produtos	<i>Arabidopsis thaliana</i>
	AmiS2	transportador ABC AmiS2	JC5491	NENHUM	expressa	exporta produtos	<i>Rhodococcus sp.</i>
	AtPGP1	ARABIDOPSIS THALIANA P GLICOPROTEÍNA1	NP_181228	NENHUM	expressa	exporta produtos	<i>Arabidopsis thaliana</i>
	AcrA	suposta proteína de transporte efluxo multidrogas acrA	CAF23274	NENHUM	expressa	exporta produtos	<i>Candidatus Protochlamydia amoebophila UWE25</i>
	AcrB	provável proteína de transporte efluxo multidrogas acrB	CAF23275	NENHUM	expressa	exporta produtos	<i>Candidatus Protochlamydia amoebophila UWE25</i>
	TolC	proteína de membrana externa [biogênese do envelope celular,	ABD59001	NENHUM	expressa	exporta produtos	<i>Francisella tularensis subsp. novicida</i>
	AcrE	proteína transmembrana afeta formação de septo e permeabilidade da membrana celular	YP_312213	NENHUM	expressa	exporta produtos	<i>Shigella sonnei Ss046</i>
	AcrF	proteína de resistência acriflavina F	P24181	NENHUM	expressa	exporta produtos	<i>Escherichia coli</i>
	tll1618	transportador de efluxo multidrogas	NP_682408.1	NENHUM	expressa	exporta produtos	<i>Thermosynechococcus elongatus BP-1J</i>

	tll1619	transportador de efluxo multidrogas	NP_682409.1	NENHUM	expressa	exporta produtos	<i>Thermosynechococcus elongatus BP-1J</i>
	tll10139	transportador de efluxo multidrogas	NP_680930.1	NENHUM	expressa	exporta produtos	<i>Thermosynechococcus elongatus BP-1J</i>
5. Fermentação							
	genes <i>checkpoint</i> de replicação					aumenta eficiência de produção	
	umuD	DNA polimerase V, subunidade	YP_310132	3.4.21.-	sobreexpressa	aumenta eficiência de produção	<i>Shigella sonnei Ss046</i>
	umuC	DNA polimerase V, subunidade	ABC42261	2.7.7.7	sobreexpressa	aumenta eficiência de produção	<i>Escherichia coli</i>
	NADH:NADPH transidrogenase (subunidades alfa e beta) (pntA, pntB)		P07001, P0AB70	1.6.1.2	expressa	aumenta eficiência de produção	<i>Shigella flexneri</i>

[098] Como usado neste documento, o termo "ácido graxo" significa um ácido carboxílico contendo a fórmula RCOOH. R representa um grupo alifílico, preferencialmente um grupo alquila. R pode compreender entre cerca de 4 e cerca de 22 átomos de carbono. Ácidos graxos podem ser saturados, monoinsaturados ou polinsaturados. Em uma concretização preferencial, o ácido graxo é feito a partir de uma via biossintética de ácido graxo.

[099] Como usado neste documento, o termo "via biossintética de ácido graxo" significa uma via biossintética que produza ácidos graxos. A via biossintética de ácido graxo inclui enzimas de ácido graxo que podem ser projetadas, como descrito neste documento, para produzir ácidos graxos e, em algumas concretizações, podem ser expressadas com enzimas adicionais para produzir ácidos graxos contendo características desejadas de cadeia de carbono.

[100] Como usado neste documento, o termo "derivado de ácido graxo" significa produtos feitos em parte a partir da via biossintética de ácido graxo do organismo hospedeiro de produção. O "derivado de ácido graxo" também inclui produtos feitos em parte a partir de acil-ACP ou de derivados de acil-ACP. A via biossintética de ácido graxo inclui enzimas de sintase de ácido graxo que podem ser projetadas como descrito neste documento para produzir derivados de ácido graxo e, em alguns exemplos, podem ser expressadas com enzimas adicionais para produzir derivados de ácidos graxos contendo características desejadas de cadeia de carbono. Derivados de ácido graxo exemplares incluem, por exemplo, ácidos graxos, acil-CoA, aldeído graxo, álcoois de cadeias longas e curtas,

hidrocarbonetos, álcoois graxos e ésteres (por exemplo, ceras, ésteres de ácido graxos ou ésteres graxos).

[101] Como usado neste documento, o termo "enzimas de derivados de ácidos graxos" significa todas as enzimas que podem ser expressadas ou sobreexpressadas na produção de derivados de ácido graxo. Essas enzimas são coletivamente referidas neste documento como enzimas de derivados de ácidos graxos. Essas enzimas devem ser parte da via biossintética de ácido graxo. Exemplos não limitantes de enzimas de derivados de ácido graxo incluem sintases de ácido graxo, tioesterases, sintases de acil-CoA, redutases de acil-CoA, desidrogenases de álcool, aciltransferases de álcool, redutase de acil-CoA formadora de álcool graxo, sintases de éster, polipeptídeos biossintéticos de aldeído e polipeptídeos biossintéticos de alcano. Enzimas de derivados de ácido graxo convertem um substrato em um derivado de ácido graxo. Em alguns exemplos, o substrato pode ser um derivado de ácido graxo o qual a enzima de derivado de ácido graxo converta em um derivado de ácido graxo diferente.

[102] Como usado neste documento, o termo "peptídeos formadores de álcool graxo" significa um peptídeo capaz de catalisar a conversão de acil-CoA para álcool graxo, incluindo redutase de acil-CoA formadora de álcool graxo (FAR, EC 1,1,1,*), redutase de acyl-CoA (EC 1,2,1,50) ou desidrogenase de álcool (EC 1,1,1,1). Adicionalmente, uma habilidade comum no estado da arte apreciará que alguns peptídeos formadores de álcool graxo catalisem outras reações também. Por exemplo, alguns peptídeos de redutase de acil-CoA aceitarão outros substratos além dos ácidos graxos. Tais peptídeos não específicos estão, portanto, também

incluídos. Sequências de ácido nucleico codificando peptídeos formadores de álcool graxo são conhecidos no estado da técnica, e tais peptídeos estão publicamente disponíveis. Exemplos de número de acesso GenBank são fornecidos na Tabela A.

[103] Como usado neste documento, "enzima de ácido graxo" significa qualquer enzima envolvida na biossíntese de ácido graxo. Enzimas de ácido graxo podem ser expressadas ou sobreexpressadas nas células hospedeiras para produzir ácidos graxos. Exemplos não limitantes de enzimas de ácido graxo incluem sintases de ácido graxo e tioesterases.

[104] Como usado neste documento, o termo "éster graxo" significa um éster. Em uma concretização preferida, um éster graxo é qualquer éster feito a partir de um ácido graxo, por exemplo, um éster de ácido graxo. Em uma concretização, um éster graxo contém um lado A (*i.e.*, a cadeia de carbono anexada ao oxigênio carboxilado) e um lado B (*i.e.*, a cadeia de carbono compreendendo o carboxilato fonte). Em uma concretização preferida, quando o éster graxo é derivado a partir da via biossintética de ácido graxo, o lado A é contribuído por um álcool e o lado B é contribuído pelo um ácido graxo. Qualquer álcool pode ser usado para formar o lado A dos ésteres graxos. Por exemplo, o álcool pode ser derivado a partir da via biossintética de ácido graxo. Alternativamente, o álcool pode ser produzido através de vias biossintéticas de ácidos não graxos. Além disso, o álcool pode ser fornecido exogenamente. Por exemplo, o álcool pode ser fornecido na fermentação de caldo em instâncias nas quais o éster graxo seja produzido por um organismo. Alternativamente, um ácido carboxílico, tal com um ácido

graxo ou ácido acético, pode ser fornecido exogenamente em instâncias nas quais o éster graxo seja produzido por um organismo que possa também produzir álcool.

[105] As cadeias de carbono que compreendem o lado A ou o lado B podem ser de qualquer comprimento. Em uma concretização, o lado A do éster é de pelo menos cerca de 1, 2, 3, 4, 5, 6, 7, 8, 10, 12, 14, 16 ou 18 carbonos de comprimento. O lado B do éster é de pelo menos cerca de 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24 ou 26 carbonos de comprimento. O lado A e/ou o lado B podem ser cadeia linear ou ramificada. As cadeias ramificadas devem ter um ou mais pontos de ramificação. Além disso, as cadeias ramificadas devem incluir ramificações cíclicas. Ademais, o lado A e/ou o lado B pode ser saturado ou insaturado. Se insaturado, o lado A e/ou o lado B pode ter um ou mais pontos de insaturação.

[106] Em uma concretização, o éster graxo é produzido biossinteticamente. Nessa concretização, primeiro o ácido graxo é "ativado". Exemplos não limitantes de ácidos graxos "ativados" são acil-CoA, acil-ACP e acil fosfato. Acil-CoA pode ser um produto direto da biossíntese ou da degradação de ácido graxo. Além disso, acil-CoA pode ser sintetizado a partir de um ácido graxo livre, de uma CoA ou de um nucleotídeo adenosina trifostado (ATP). Em exemplo de uma enzima que produz acil-CoA é sintase de acil-CoA.

[107] Após o ácido graxo ser ativado, ele pode ser prontamente transferido para um recipiente nucleófilo. Exemplos nucleófilos são álcoois, tióis ou fosfatos.

[108] Em uma concretização, o éster graxo é uma cera. A cera pode ser derivada a partir de um álcool de cadeia longa

e de um ácido graxo de cadeia longa. Em uma outra concretização, o éster graxo pode ser derivado a partir de um acil-tioéster graxo e um álcool. Em outra concretização, o éster graxo é um tioéster de ácido graxo, por exemplo, acil Coenzima A (CoA) graxo. Em outras concretizações, o éster graxo é um acil pantotenato graxo, uma proteína transportadora de acila (ACP) ou um éster fosfato graxo. Ésteres graxos possuem muitos usos. Por exemplo, ésteres graxos podem ser usados como biocombustível.

[109] Como usado neste documento, "fração de carbono moderno" ou " f_M " possui o mesmo significado como definido pelo *National Institute of Standards and Technology* (NIST) Materiais e Referência Padrão (SRMs) 4990B e 4990C, conhecido como padrões de ácidos oxálicos HOxI e HOxII, respectivamente. A definição fundamental refere-se a 0,95 vezes a razão do isótopo $^{14}\text{C}/^{12}\text{C}$ HOxI (referido a AD 1950). Isso é grosseiramente equivalente ao decaimento corrigido da madeira pré-Revolução Industrial. Para a biosfera de vida atual (material de planta), f_M é de aproximadamente 1,1.

[110] Cálculos de "homologia" entre duas sequências podem ser realizados como se segue. As sequências são alinhadas para fins de comparação ideal (por exemplo, lacunas podem ser introduzidas em um ou ambos de uma primeira e uma segunda sequência aminoácido ou ácido nucleico para melhor alinhamento, e sequências não homólogas podem ser ignoradas para fins de comparação). Em uma concretização preferida, o comprimento de uma sequência de referência que esteja alinhada para fins de comparação é de pelo menos cerca de 30%, preferencialmente de pelo menos 40%, mais preferencialmente de cerca de 50%, ainda mais

preferencialmente de cerca de 60%, e ainda mais preferencialmente de cerca de 70%, de pelo menos cerca de 80%, de pelo menos cerca de 90% ou cerca de 100% do comprimento da sequência de referência. Os resíduos de aminoácido ou nucleotídeos nas posições correspondentes ao aminoácido ou nucleotídeos são então comparados. Quando uma posição na primeira sequência é ocupada pelo menos resíduo de aminoácido ou nucleotídeo como a posição correspondente na segunda sequência, então as moléculas são idênticas naquela posição (como usado neste documento, a "identidade" do aminoácido ou do ácido nucleico é equivalente à "homologia" do aminoácido ou do ácido nucleico). A identidade percentual entre as duas sequências é uma função do número de posições idênticas compartilhadas pelas sequências, levando-se em conta o número de lacunas e o comprimento de cada lacuna, o que precise ser introduzido para o melhor alinhamento das duas sequências.

[111] A comparação de sequências e a determinação de homologia percentual entre as duas sequências podem ser adquiridas usando algoritmo matemático. Em uma concretização preferida, a homologia percentual entre duas sequências de aminoácidos é determinada usando-se o Needleman e Wunsch (1970), *J. Mol. Biol.* 48:444-453, algoritmo que foi incorporado ao programa GAP na embalagem do *software* GCG, usando tanto uma matriz Blossum 62 quanto uma matriz PAM250, e um peso de lacuna de 16, 14, 12, 10, 8, 6 ou 4 e um peso de comprimento de 1, 2, 3, 4, 5 ou 6. Em ainda outra concretização preferida, a homologia percentual entre duas sequências de nucleotídeos é determinada usando o programa GAP no pacote do *software* GCG, usando uma matriz

NWSgapdna.CMP e um peso de lacuna de 40, 50, 60, 70 ou 80 e um peso de comprimento de 1, 2, 3, 4, 5 ou 6. Um conjunto particularmente preferido de parâmetros (e o único que deve ser usado se o praticante estiver incerto sobre quais parâmetros devem ser aplicados para determinar se uma molécula está dentro de uma limitação de homologia das reivindicações) é uma matriz de pontuação Blossum 62 com uma penalidade de lacuna de 12, uma penalidade de lacuna ampla de 4 e uma penalidade de lacuna de *frameshift* de 5.

[112] Como usado neste documento, uma "célula hospedeira" é uma célula usada para produzir um produto descrito neste documento (por exemplo, um aldeído ou alcano descrito neste documento). Uma célula hospedeira pode ser modificada para expressar ou sobreexpressar genes selecionados ou para ter atenuada a expressão de genes selecionados. Exemplos não limitantes de células hospedeiras incluem células de planta, animal, humana, bactéria, levedura ou de fungos filamentosos.

[113] Como usado neste documento, o termo "hibridiza sob condições de baixa estringência, média estringência, alta estringência ou muito alta estringência" descreve condições para hibridização e lavagem. A orientação para realizar reações de hibridização pode ser encontrada em *Current Protocols in Molecular Biology*, John Wiley & Sons, NY (1989), 6,3,1-6,3,6. Métodos aquosos e não aquosos são descritos nessa referência e tanto um quanto outro pode ser usado. Condições específicas de hibridização referidas neste documento são como se segue: 1) condições de hibridização de baixa estringência em 6X de cloreto de sódio/citrato de sódio (SSC) a cerca de 45°C, seguido por duas lavagens em 0,2X

SSC, 0,1% SDS a pelo menos 50°C (a temperatura das lavagens pode ser aumentada para 55°C para condições de baixa estringência); 2) condições de hibridização de média estringência em 6X SSC a cerca de 45°C, seguida por uma ou mais lavagens em 0,2X SSC, 0,1% SDS a 60°C; 3) condições de hibridização em elevada estringência em 6X SSC a cerca de 45°C, seguida de uma ou mais lavagens em 0,2X SSC, 0,1% SDS a 65°C; e preferencialmente 4) condições de hibridização em muito elevada estringência são 0,5M de fosfato de sódio, 7% SDS a 65°C, seguido de uma ou mais lavagens a 0,2X SSC, 1% SDS a 65°C. Condições muito elevadas de estringência (4) são as condições preferidas salvo em contrário especificado.

[114] O termo "isolado" como usado neste documento com respeito aos ácidos nucleicos, tais como DNA ou RNA, refere-se a moléculas separadas a partir de outros DNAs ou RNAs, respectivamente, que são apresentados na fonte natural de ácido nucleico. Além disso, "um ácido nucleico isolado" inclui fragmentos de ácido nucleico, tais como fragmentos que não ocorrem naturalmente. O termo "isolado" é também usado neste documento para se referir a polipeptídeos, que são isolados a partir de outras proteínas celulares, e engloba tanto os polipeptídeos edógenos purificados quanto os polipeptídeos recombinantes. O termo "isolado" como usado neste documento também refere-se a um ácido nucleico ou polipeptídeo que seja substancialmente livre de material celular, material viral ou meio de cultura quando produzido por técnicas de DNA recombinante. O termo "isolado" como usado neste documento também se refere a um ácido nucleico ou polipeptídeo que seja substancialmente livre de precursores químicos ou outras químicas quando sintetizado

quimicamente.

[115] Como usado neste documento, o "nível de expressão de um gene em uma célula" refere-se ao nível de mRNA, às transcrições nascentes de pré-mRNA, aos processos de transcrição intermediários, ao(s) mRNA(s) maduro(s) e/ou aos produtos da degradação codificados pelo gene na célula.

[116] Como usado neste documento, o termo "microrganismo" significa espécies microbiais procarióticas e eucarióticas a partir de domínios Archea, Bacteria e Eucarya, o último incluindo levedura e fungos filamentosos, protozoa, algas ou alto protista. O termo "célula microbiana", como usado neste documento, significa uma célula a partir de um microrganismo.

[117] Como usado neste documento, o termo "ácido nucleico" refere-se a polinucleotídeos, tais como ácido desoxirribonucleico (DNA), e, quando apropriado, ácido ribonucleico (RNA). O termo ainda inclui análogos tanto do RNA quanto do DNA feitos a partir de nucleotídeos análogos e, como aplicável à concretização sendo descrita, polinucleotídeos simples (senso e antissenso) e de filamento duplo, ESTs, cromossomos, cDNAs, mRNAs e rRNAs.

[118] Como usado neste documento, o termo "operacionalmente ligado" significa que a sequência de nucleotídeo selecionada (por exemplo, codificando um polipeptídeo descrito neste documento) esteja em proximidade com um promotor para permitir que o promotor regule a expressão da sequência de nucleotídeo selecionada. Além disso, o promotor está localizado à montante da sequência de nucleotídeo selecionada em termos da direção de transcrição e tradução. Por "operacionalmente ligado" destina-se que uma

sequência de nucleotídeo e uma sequência(s) regulatória sejam conectadas de tal forma que permita a expressão do gene quando as moléculas apropriadas (por exemplo, proteínas ativadoras transcricionais) estiverem ligadas à sequência(s) regulatória.

[119] O termo "ou" é usado neste documento para significar, e é usado alternadamente com, o termo "e/ou", salvo em contexto claramente indicando o contrário.

[120] Como usado neste documento, "sobrexpressar" significa expressar ou provocar que seja expressado um ácido nucleico, polipeptídeo ou hidrocarboneto em um célula a um concentração maior do que seja normalmente expressado em uma célula de tipo selvagem correspondente. Por exemplo, um polipeptídeo pode ser "sobrexpresso" em uma célula hospedeira recombinante quando o polipeptídeo está presente em uma concentração maior na célula hospedeira recombinante comparada a sua concentração em uma célula hospedeira não recombinante da mesma espécie.

[121] Como usado neste documento, "coeficiente de partição" ou "P" é definido como a concentração de equilíbrio de um composto em uma fase orgânica dividida pela concentração em equilíbrio em uma fase aquosa (por exemplo, caldo de fermentação). Em uma concretização de um sistema bifásico descrito neste documento, a fase orgânica é realizada pelo aldeído ou alceno durante o processo de produção. Entretanto, em alguns exemplos, uma fase orgânica pode ser fornecida, tal como pelo fornecimento de uma camada de octano, para facilitar a separação do produto. Durante a descrição de um sistema de duas fases, as características de partição de um composto podem ser descritas como logP. Por

exemplo, um composto com um logP de 1 seria a partição 10:1 para a fase orgânica. Um composto com um logP de -1 seria a partição 1:10 para a fase orgânica. Escolhendo-se um caldo de fermentação apropriado e a fase orgânica, um aldeído ou alcano com um valor de logP elevado pode separar dentro da fase orgânica mesmo que em concentrações muito baixas no vaso de fermentação.

[122] Como usado neste documento, o termo "purificar", "purificado" ou "purificação" significa a remoção ou o isolamento de uma molécula a partir de seu ambiente por, por exemplo, isolamento ou separação. Molécula "substancialmente purificadas" são pelo menos cerca de 60% livres, preferencialmente pelo menos cerca de 75% livres, e mais preferencialmente pelo menos cerca de 90% livres de outros componentes com os quais elas são associadas. Como usado neste documento, esses termos também se referem à remoção de contaminantes de uma amostra. Por exemplo, a remoção de contaminantes pode resultar em um aumento na porcentagem de aldeídos ou alcanos em uma amostra. Por exemplo, quando aldeídos ou alcanos são produzidos em uma célula hospedeira, os aldeídos ou alcanos podem ser purificados pela remoção de proteínas da célula hospedeira. Após a purificação, a porcentagem de aldeídos ou alcanos na amostra é aumentada.

[123] Os termos "purificar", "purificado" ou "purificação" não exigem pureza absoluta. Eles são termos relativos. Assim, por exemplo, quando aldeídos ou alcanos são produzidos em células hospedeiras, um aldeído purificado ou alcano purificado é aquele que é substancialmente separado dos outros componentes celulares (por exemplo, ácidos nucleicos, polipeptídeos, lipídeos, carboidratos ou outros

hidrocarbonetos). Em outro exemplo, uma preparação de aldeído purificado ou alceno purificado é aquela na qual o aldeído ou alceno é substancialmente livre de contaminantes, tais como aqueles que podem estar presentes na fermentação seguinte. Em algumas concretizações, um aldeído ou um alceno é purificado quando pelo menos cerca de 50% de peso de uma amostra for composta do aldeído ou do alceno. Em outras concretizações, um aldeído ou um alceno é purificado quando pelo menos cerca de 60%, 70%, 80%, 85%, 90%, 92%, 95%, 98% ou 99% ou mais do peso de uma amostra for composto pelo aldeído ou alceno.

[124] Como usado neste documento, o termo "polipeptídeo recombinante" refere-se a um polipeptídeo que seja produzido por técnicas de DNA recombinante, onde geralmente o DNA codifica o polipeptídeo expressado ou o RNA é inserido dentro de um vetor de expressão adequado e que seja, por sua vez, usado para transformar a célula hospedeira para produzir o polipeptídeo ou RNA.

[125] Como usado neste documento, o termo "substancialmente idêntico" (ou "substancialmente homólogo") é usado para se referir a uma primeira sequência de aminoácido ou nucleotídeo que contenha um número suficiente de resíduos de aminoácido ou nucleotídeos idêntico ou equivalente (por exemplo, com uma cadeia lateral semelhante, por exemplo, substituições de aminoácido conservadas) a uma segunda sequência de aminoácidos ou nucleotídeos tal que a primeira e a segunda sequências de aminoácidos ou nucleotídeos tenham atividades similares.

[126] Como usado neste documento, o termo "sintase" significa uma enzima que catalisa um processo de síntese.

Como usado neste documento, o termo sintase inclui sintases, sintetases ou ligases.

[127] Como usado neste documento, o termo "transfecção" significa a introdução de um ácido nucleico (por exemplo, por meio de um vetor de expressão) dentro de uma célula recipiente pela transferência gênica mediada de ácido nucleico.

[128] Como usado neste documento, "transformação" refere-se a um processo no qual o genótipo da célula é modificado como resultado da absorção nuclear de ácido nucleico exógeno. Isso deve resultar na célula transformada expressando uma forma recombinante de um RNA ou polipeptídeo. No caso de expressão antissenso a partir do gene transferido, a expressão de uma forma que ocorra naturalmente do polipeptídeo é interrompida.

[129] Como usado neste documento, uma "proteína de transporte" é um polipeptídeo que facilita o movimento de um ou mais compostos dentro e/ou fora de uma organela celular e/ou uma célula.

[130] Como usado neste documento, uma "variante" de polipeptídeo X refere-se a um polipeptídeo contendo a sequência de aminoácido de polipeptídeo X na qual um ou mais resíduos de aminoácidos seja alterado. A variante deve conter mudanças conservantes ou mudanças não conservantes. A orientação em determinar quais resíduos de aminoácidos devem ser substituídos, inseridos ou apagados sem afetar a atividade celular pode ser encontrada usando-se programas de computadores bem-conhecidos no estado da técnica, por exemplo, o *software* LASERGENE (DNASTAR).

[131] O termo "variante", quando usado no contexto de

uma sequência de polinucleotídeo, deve englobar uma sequência de polinucleotídeo relacionada àquela de um gene ou à sequência de codificação deste. Essa definição deve ainda incluir, por exemplo, variantes "alélicas", "de encaixe", de "espécies" ou "polimórficas". Uma variante de encaixe deve conter uma identificação significativa a uma polinucleotídeo de referência, mas geralmente terá um número maior ou menor de polnucleotídeos devido ao encaixe alternativo de éxons durante o processamento de mRNA. O polipeptídeo correspondente deve possuir domínios funcionais adicionais ou uma ausência de domínios. Variantes de espécies são sequências de polinucleotídeos que variam de uma espécie para outro. Os polipeptídeos resultantes geralmente possuirão identificação de aminoácido significativa relativa uma com a outra. Uma variante polimórfica é uma variação na sequência de polinucleotídeos de um gene particular entre indivíduos de uma dada espécie.

[132] Como usado neste documento, o termo "vetor" refere-se a uma molécula de ácido nucleico capaz de transportar outro ácido nucleico para o qual este tenha sido ligado. Um tipo de vetor útil é um episoma (*i.e.*, um ácido nucleico capaz de replicação extracromossomial). Vetores úteis são aqueles capazes de replicações autônomas e/ou expressões de ácidos nucleicos aos quais eles estão ligados. Vetores capazes de direcionar a expressão de genes para os quais eles estão operacionalmente ligado são referidos neste documento como "vetores de expressão. Em geral, vetores de expressão úteis em técnicas de DNA recombinante estão frequentemente na forma de "plasmídeos", o que se refere geralmente a voltas circulares de DNA de filamento duplo

que, em sua forma de vetor, não são ligados ao cromossomo. No presente relatório descritivo, "plasmídeos" e "vetor" são usados alternadamente, já que o plasmídeo é a forma de vetor mais comumente usada. Entretanto, também incluídas estão outras formas de vetores de expressão que servem de funções equivalentes e que se tornaram conhecidas no estado da técnica subsequentemente neste documento.

[133] Salvo em definição contrária, todos os termos técnicos e científicos usados neste documento possuem o mesmo significado como entendidos comumente por aquele de habilidade ordinária no estado da técnica à qual esta invenção pertence. Embora os métodos e os materiais semelhantes ou equivalentes àqueles descritos neste documento possam ser usados na prática ou no teste da presente invenção, métodos e materiais adequados são descritos abaixo. Todas as publicações, os pedidos de patentes, as patentes ou outras referências mencionadas neste documento estão incorporadas para referência em sua totalidade. No caso de conflito, o presente relatório descritivo, incluindo as definições, controlará. Além disso, os materiais, os métodos e os exemplos são ilustrativos apenas e não têm a intenção de serem limitantes.

[134] Outras características e vantagens da invenção serão aparentes a partir da seguinte descrição detalhada e das reivindicações.

BREVE DESCRIÇÃO DAS FIGURAS

[135] FIG. 1A é um traço GC/MS de hidrocarbonetos produzidos pelas células *Prochlorococcus marinus* CCMP1986. FIG. 1B é um padrão de fragmentação de massa do pico a 7,55 min da FIG. 1A.

[136] FIG. 2A é um traço GC/MS de hidrocarbonetos produzidos pelas células *Nostoc punctiforme* PCC73102. FIG. 2B é um padrão de fragmentação de massa do pico a 8,73 min da FIG. 2A.

[137] FIG. 3A é um traço GC/MS de hidrocarbonetos produzidos pelas células *Gloeobaceter violaceus* ATCC29082. FIG. 3B é um padrão de fragmentação de massa do pico a 8,72 min da FIG. 3A.

[138] FIG.4A é um traço GC/MS de hidrocarbonetos produzidos pelas células *Synechocystis sp.* PCC6803. FIG. 4B é um padrão de fragmentação de massa do pico a 7,36 min da FIG. 4A.

[139] FIG. 5A é um traço GC/MS de hidrocarbonetos produzidos pelas células *Synechocystis sp.* PCC6803. FIG. 5B é um traço GC/MS de hidrocarbonetos produzidos pelas células *Synechocystis sp.* PCC6803 com a exclusão dos genes *sll0208* e *sll0209*.

[140] FIG. 6A é um traço GC/MS de hidrocarbonetos produzidos pelas células do tipo selvagem de *E. coli* MG1655. FIG. 6B é um traço GC/MS de hidrocarbonetos produzidos pelas células de *E. coli* MG1655 expressando *Synechococcus elongatus* PCC7942 YP_400611 (Synpcc7942_1594) (SEQ ID NO:65).

[141] FIG. 7 é um traço GC/MS de hidrocarbonetos produzidos pelas células de *E. coli* expressando *Cyanothece sp.* ATCC51142 cce_1430 (YP_001802846) (SEQ ID NO:69).

[142] FIG. 8A é um traço GC/MS de hidrocarbonetos produzidos pelas células de *E. coli* expressando *Synechococcus elongatus* PCC7942 YP_400611 (Synpcc7942_1594) (SEQ ID NO:65). FIG. 8A retrata um padrão de fragmentação de massa do pico a 6,98 min da FIG. 8A e de

8-heptadeceno.

[143] FIG. 9 é um traço GC/MS de hidrocarbonetos produzidos pelas células de *E. coli* MG1655 expressando *Synechococcus elongatus* PCC7942 YP_400611 (Synpcc7942_1594) (SEQ ID NO:65) e *Nostoc punctiforme* PCC73102 Npun02004178 (ZP_00108838) (SEQ ID NO:5).

[144] FIG. 10 é um traço GC/MS de hidrocarbonetos produzidos pelas células de *E. coli* MG1655 expressando *Synechococcus elongatus* PCC7942 YP_400611 (Synpcc7942_1594) (SEQ ID NO:65) e *Synechocystis* sp. PCC6803 sl10208 (NP_442147) (SEQ ID NO:3).

[145] FIG. 11 é um traço GC/MS de hidrocarbonetos produzidos pelas células de *E. coli* MG1655 expressando *Synechococcus elongatus* PCC7942 YP_400611 (Synpcc7942_1594) (SEQ ID NO:65) e *Nostoc* sp. PCC7210 alr5283 (NP_489323) (SEQ ID NO:7).

[146] FIG. 12 é um traço GC/MS de hidrocarbonetos produzidos pelas células de *E. coli* MG1655 expressando *Synechococcus elongatus* PCC7942 YP_400611 (Synpcc7942_1594) (SEQ ID NO:65) e *Acaryochloris marina* MBIC11017 AM1_4041 (YP_001518340) (SEQ ID NO:46) otimizado de códon.

[147] FIG. 13 é um traço GC/MS de hidrocarbonetos produzidos pelas células de *E. coli* MG1655 expressando *Synechococcus elongatus* PCC7942 YP_400611 (Synpcc7942_1594) (SEQ ID NO:65) e *Thermosynechococcus elongatus* BP-1 tl11313 (NP_682103) (SEQ ID NO:47) otimizado de códon.

[148] FIG. 14 é um traço GC/MS de hidrocarbonetos produzidos pelas células de *E. coli* MG1655 expressando *Synechococcus elongatus* PCC7942 YP_400611 (Synpcc7942_1594) (SEQ ID NO:65) e *Synechococcus* sp. JA-3-3Ab CYA_0415

(YP_473897) (SEQ ID NO:48) otimizado de códon.

[149] FIG. 15 é um traço GC/MS de hidrocarbonetos produzidos pelas células de *E. coli* MG1655 expressando *Synechococcus elongatus* PCC7942 YP_400611 (Synpcc7942_1594) (SEQ ID NO:65) e *Gloeobacter violaceus* PCC7421 gll3146 (NP_926092) (SEQ ID NO:15).

[150] FIG. 16 é um traço GC/MS de hidrocarbonetos produzidos pelas células de *E. coli* MG1655 expressando *Synechococcus elongatus* PCC7942 YP_400611 (Synpcc7942_1594) (SEQ ID NO:65) e *Prochlorococcus marinus* MIT9313 PMT1231 (NP_895059) (SEQ ID NO:49) otimizado de códon.

[151] FIG. 17 é um traço GC/MS de hidrocarbonetos produzidos pelas células de *E. coli* MG1655 expressando *Synechococcus elongatus* PCC7942 YP_400611 (Synpcc7942_1594) (SEQ ID NO:65) e *Prochlorococcus marinus* CCMP1986 PMM0532 (NP_892650) (SEQ ID NO:19).

[152] FIG. 18 é um traço GC/MS de hidrocarbonetos produzidos pelas células de *E. coli* MG1655 expressando *Synechococcus elongatus* PCC7942 YP_400611 (Synpcc7942_1594) (SEQ ID NO:65) e *Prochlorococcus marinus* NATL2A PMN2A_1863 (YP_293054) (SEQ ID NO:51) otimizado de códon.

[153] FIG. 19 é um traço GC/MS de hidrocarbonetos produzidos pelas células de *E. coli* MG1655 expressando *Synechococcus elongatus* PCC7942 YP_400611 (Synpcc7942_1594) (SEQ ID NO:65) e *Synechococcus sp.* RS9917 RS9917_09941 (ZP_01079772) (SEQ ID NO:52) otimizado de códon.

[154] FIG. 20 é um traço GC/MS de hidrocarbonetos produzidos pelas células de *E. coli* MG1655 expressando *Synechococcus elongatus* PCC7942 YP_400611 (Synpcc7942_1594) (SEQ ID NO:65) e *Synechococcus sp.* RS9917 RS9917_12945

(ZP_01080370) (SEQ ID NO:53) otimizado de códon.

[155] FIG. 21 é um traço GC/MS de hidrocarbonetos produzidos pelas células de *E. coli* MG1655 expressando *Synechococcus elongatus* PCC7942 YP_400611 (Synpcc7942_1594) (SEQ ID NO:65) e *Cyanothece sp.* ATCC51142 cce_0778 (YP_001802195) (SEQ ID NO:27).

[156] FIG. 22 é um traço GC/MS de hidrocarbonetos produzidos pelas células de *E. coli* MG1655 expressando *Synechococcus elongatus* PCC7942 YP_400611 (Synpcc7942_1594) (SEQ ID NO:65) e *Cyanothece sp.* PCC7425 Cyan7425_0398 (YP_002481151) (SEQ ID NO:29).

[157] FIG. 23 é um traço GC/MS de hidrocarbonetos produzidos pelas células de *E. coli* MG1655 expressando *Synechococcus elongatus* PCC7942 YP_400611 (Synpcc7942_1594) (SEQ ID NO:65) e *Cyanothece sp.* PCC7425 Cyan7425_2986 (YP_002483683) (SEQ ID NO:31).

[158] FIG. 24A é um traço GC/MS de hidrocarbonetos produzidos pelas células de *E. coli* MG1655 expressando *Prochlorococcus marinus* CCMP1986 PMM0533 (NP_892651) (SEQ ID NO:71). FIG. 24B é um traço GC/MS de hidrocarbonetos produzidos pelas células de *E. coli* MG1655 expressando *Prochlorococcus marinus* CCMP1986 PMM0533 (NP_892651) (SEQ ID NO:71) e *Prochlorococcus marinus* CCMP1986 PMM0532 (NP_892650) (SEQ ID NO:19).

[159] FIG. 25A é um traço GC/MS de hidrocarbonetos produzidos pelas células de *E. coli* MG1655 $\Delta fadE lacZ::P_{trc}$ 'tesA-fadD. FIG. 25B é um traço GC/MS de hidrocarbonetos produzidos pelas células de *E. coli* MG1655 $\Delta fadE lacZ::P_{trc}$ 'tesA-fadD expressando *Synechococcus elongatus* PCC7942 YP_400611 (Synpcc7942_1594) (SEQ ID NO:65) e *Acaryochloris*

marina MBIC11017 AM1_4041 (YP_001518340) (SEQ ID NO:9).

[160] FIG. 26A é um traço GC/MS de hidrocarbonetos produzidos pelas células de *E. coli* MG1655 $\Delta fadE lacZ::P_{trc}$ 'tesA-fadD expressando *Synechocystis sp.* PCC6803 sll0209 (NP_442146) (SEQ ID NO:67). FIG. 26B é um traço GC/MS de hidrocarbonetos produzidos pelas células de *E. coli* MG1655 $\Delta fadE lacZ::P_{trc}$ 'tesA-fadD expressando *Synechocystis sp.* PCC6803 sll0208 (NP_442147) (SEQ ID NO:3).

[161] FIG. 27A é um traço GC/MS de hidrocarbonetos produzidos pelas células de *E. coli* MG1655 $\Delta fadE lacZ::P_{trc}$ 'tesA-fadD expressando a descendência *M. smegmatis* MC2 155 MSMEG_5739 (YP_889972) (SEQ ID NO:85). FIG. 27B é um traço GC/MS de hidrocarbonetos produzidos pelas células de *E. coli* MG1655 $\Delta fadE lacZ::P_{trc}$ 'tesA-fadD expressando a descendência *M. smegmatis* MC2 155 MSMEG_5739 (YP_889972) (SEQ ID NO:85) e *Nostoc punctiforme* PCC73102 Npun02004178 (ZP_00108838) (SEQ ID NO:5).

[162] FIG. 28 é uma representação gráfica de hidrocarbonetos produzidos pelas células de *E. coli* MG1655 $\Delta fadE lacZ::P_{trc}$ 'tesA-fadD expressando a descendência *M. smegmatis* MC2 155 MSMEG_5739 (YP_889972) (SEQ ID NO:85) tanto sozinha quanto em combinação com *Nostoc sp.* PCC7120 alr5283 (SEQ ID NO:7), *Nostoc punctiforme* PCC73102 Npun02004178 (SEQ ID NO:5), *P. mariunus* CCMP1986 PMM0532 (SEQ ID NO:19), *G. violaceus* PCC7421 gl13146 (SEQ ID NO:15), *Synechococcus sp.* RS9917_09941 (SEQ ID NO:23), *Synechococcus sp.* RS9917-12945 (SEQ ID NO:25) ou *A. marina* MBIC11017 AM1_4041 (SEQ ID NO:9).

[163] FIG. 29A é uma representação da estrutura tridimensional de uma proteína subunidade β ribonuclease redutase classe I, RNR β . FIG. 29B é uma representação da

estrutura tridimensional do lugar ativo de *Prochlorococcus marinus* MIT9313 PMT1231 (NP_895059) (SEQ ID NO:17). FIG.29C é uma representação da estrutura tridimensional do lugar ativo de *Prochlorococcus marinus* MIT9313 PMT1231 (NP_895059) (SEQ ID NO:17).

[164] FIG. 30A é um traço GC/MS de hidrocarbonetos produzidos pelas células de *E. coli* MG1655 expressando *Nostoc punctiforme* PCC73102 Npun02004178 (ZP_00108838) (SEQ ID NO:5). FIG. 30B é um traço GC/MS de hidrocarbonetos produzidos pelas células de *E. coli* MG1655 expressando a variante *Nostoc punctiforme* PCC73102 Npun02004178 (ZP_00108838) Y123F. FIG. 30C é um traço GC/MS de hidrocarbonetos produzidos pelas células de *E. coli* MG1655 expressando a variante *Nostoc punctiforme* PCC73102 Npun02004178 (ZP_00108838) Y126F.

[165] FIG. 31 retrata traços GC/MS de hidrocarbonetos produzidos *in vitro* usando *Nostoc punctiforme* PCC73102 Npun02004178 (ZP_00108838) (SEQ ID NO:6) e octadenal (A); Npun02004178 (ZP_00108838) (SEQ ID NO:6), octadenal, redutase de ferredoxina de espinafre e NADPH (B); octadenal, ferredoxina de espinafre, redutase de ferredoxina de espinafre e NADPH (C); ou Npun02004178 (ZP_00108838) (SEQ ID NO:6), ferredoxina de espinafre e ferredoxina de espinafre (D).

[166] FIG. 32 retrata traços GC/MS de hidrocarbonetos produzidos *in vitro* usando *Nostoc punctiforme* PCC73102 Npun02004178 (ZP_00108838) (SEQ ID NO:6), NADPH, octadenal e tanto (A) ferredoxina de espinafra quanto redutase de ferredoxina de espinafre; (B) *N. punctiforme* PCC73102 Npun02003626 (ZP_00109192) (SEQ ID NO:88) e *N. punctiforme*

PCC73102 Npun02001001 (ZP_00111633) (SEQ ID NO:90); (C) Npun02003626 (ZP_00109192) (SEQ ID NO:88) e *N. punctiforme* PCC73102 Npun02003530 (ZP_00109422) (SEQ ID NO:92); ou (D) Npun02003626 (ZP_00109192) (SEQ ID NO:88) e *N. punctiforme* PCC73102 Npun02003123 (ZP_00109501) (SEQ ID NO:94).

[167] FIG. 33A é um traço GC/MS de hidrocarbonetos produzidos *in vitro* usando octadecanoil-CoA, *Synechococcus elongatus* PCC7942 YP_400611 (Synpcc7942_1594) (SEQ ID NO:66), NADH e Mg²⁺. FIG. 33B é um traço GC/MS de hidrocarbonetos produzidos *in vitro* usando octadecanoil-CoA, *Synechococcus elongatus* PCC7942 YP_400611 (Synpcc7942_1594) (SEQ ID NO:66), NADPH e Mg²⁺. FIG. 33C é um traço GC/MS de hidrocarbonetos produzidos *in vitro* usando octadecanoil-CoA, *Synechococcus elongatus* PCC7942 YP_400611 (Synpcc7942_1594) (SEQ ID NO:66) e NADPH.

[168] FIG. 34A é um traço GC/MS de hidrocarbonetos produzidos *in vitro* usando octadecanoil-CoA, NADPH marcado, *Synechococcus elongatus* PCC7942 YP_400611 (Synpcc7942_1594) (SEQ ID NO:66) e NADPH não marcado. FIG. 34B é um traço GC/MS de hidrocarbonetos produzidos *in vitro* usando octadecanoil-CoA, NADPH marcado, *Synechococcus elongatus* PCC7942 YP_400611 (Synpcc7942_1594) (SEQ ID NO:66) e S-(4-²H)NADPH. FIG. 34C é um traço GC/MS de hidrocarbonetos produzidos *in vitro* usando octadecanoil-CoA, NADPH marcado, *Synechococcus elongatus* PCC7942 YP_400611 (Synpcc7942_1594) (SEQ ID NO:66) e R-(4-²H)NADPH.

[169] FIG. 35 é um traço GC/MS de hidrocarbonetos no sobrenadante livre de células produzidos pelas células de *E. coli* MG1655 Δ *fadE* em meio Che-9 expressando *Synechococcus elongatus* PCC7942 YP_400611 (Sunpcc7942_1594) (SEQ ID

NO:65).

[170] FIG. 36 é um traço GC/MS de hidrocarbonetos no sobrenadante livre de células produzidos pelas células de *E. coli* MG1655 $\Delta fadE$ em meio Che-9 expressando *Synechococcus elongatus* PCC7942 YP_400611 (Sunpcc7942_1594) (SEQ ID NO:65) e *Nostoc punctiforme* PCC73102 Npun02004178 (ZP_00108838) (SEQ ID NO:5).

[171] FIG. 37 é um traço GC/MS de hidrocarbonetos produzidos pelas células de *E. coli* MG1655 expressando *Nostoc sp.* PCC7120 alr5283 (NP_489323) (SEQ ID NO:7) e *Nostoc sp.* PCC7120 alr5284 (NP_489324) (SEQ ID NO:81).

DESCRIÇÃO DETALHADA DA INVENÇÃO

[172] A invenção fornece composições e métodos de produção de aldeídos, álcoois graxos e hidrocarbonetos (tais como alcanos, alcenos e alcinos) a partir de substratos, por exemplo, de um acil-ACP, de um ácido graxo, de uma acil-CoA, de um aldeído graxo, ou de um substrato de álcool graxo (por exemplo, como descrito em PCT/US08/058788, especificamente incorporado para referência neste documento). Tais aldeídos, alcanos e alcenos são úteis como biocombustíveis (por exemplo, substitutos para gasolina, *diesel*, combustível de jatos etc.), especialidades químicas (por exemplo, lubrificantes, aditivos de combustíveis etc.) ou matéria-prima para outras conversões químicas (por exemplo, combustível, polímeros, plásticos, têxteis, solventes, adesivos etc.). A invenção é baseada, em parte, na identificação de genes que estejam envolvidos na biossíntese de aldeído, alcano e alceno.

[173] Tais genes biossintéticos de alcano e alceno incluem, por exemplo, *Synechococcus elongatus* PCC7942

Synpcc7942_1593 (SEQ ID NO:1), *Synechocystis* sp. PCC6803 s110208 (SEQ ID NO:3), *Nostoc punctiforme* PCC73102 Npun02004178 (SEQ ID NO:5), *Nostoc* sp. PCC 7120 alr5283 (SEQ ID NO:7), *Acariochloris marina* MBIC11017 AM1_4041 (SEQ ID NO:9), *Thermosynechococcus elongatus* BO-1 t111313 (SEQ ID NO:11), *Synechococcus* sp. JA-3-3A CYA_0415 (SEQ ID NO:13), *Gloeobacter violaceus* PCC 7421 g113146 (SEQ ID NO:15), *Prochlorococcus marinus* MIT9313 PM123 (SEQ ID NO:17), *Prochlorococcus marinus* subespécie *pastoris* str. CCMP1986 PMM0532 (SEQ ID NO:19), *Prochlorococcus marinus* str. NATL2A PMN2A_1863 (SEQ ID NO:21), *Synechococcus* sp. RS9917 RS9917_09941 (SEQ ID NO:23), *Synechococcus* sp. RS9917 RS9917_09945 (SEQ ID NO:25), *Cyanothece* sp. ATCC51142 cce_0778 (SEQ ID NO:27), *Cyanothece* sp. PCC7245 Cyan7425DRAFT_1220 (SEQ ID NO:29), *Cyanothece* sp. PCC7245 cce_0778 (SEQ ID NO:31), *Anabaena variabilis* ATCC29413 YP_323043 (Ava_2533) (SEQ ID NO:33) e *Synechococcus elongatus* PCC6301 YP_170760 (syc0050_d) (SEQ ID NO:35). Outros genes biossintéticos de alcanos e alcenos são listados na Tabela 1 e na Tabela B.

[174] Genes biossintéticos de aldeído incluem, por exemplo, *Synechococcus elongatus* PCC7942 Synpcc7942_1594 (SEQ ID NO:65), *Synechocystis* sp. PCC6803 s110209 (SEQ ID NO:67), *Cyanothece* sp. ATCC51142 cce_1430 (SEQ ID NO:69), *Prochlorococcus marinus* subespécie *pastoris* str. CCMP1986 PMM0533 (SEQ ID NO:71), *Prochlorococcus marinus* str. NATL2A PMN2A_1863 (SEQ ID NO:21), *Gloeobacter violaceus* PCC7421 NP_96091 (g113145) (SEQ ID NO:73), *Nostoc punctiforme* PCC73102 ZP_00108837 (Npun02004176) (SEQ ID NO:75), *Anabaena variabilis* ATCC29413 YP_323044 (Ava_2534) (SEQ ID NO:77),

Synechococcus elongatus PCC6301 YP_170761 (syc0051_d) (SEQ ID NO:79) e *Nostoc sp.* PCC 7120 alr5284 (SEQ ID NO:81). Outros genes biossintéticos de aldeídos são listados na Tabela 1 e na Tabela C.

[175] Usando os métodos descritos neste documento, aldeídos, álcoois graxos, alcanos e alcenos podem ser preparados usando-se um ou mais genes biossintéticos de aldeído, alcano e/ou alceno ou polipeptídeos descritos neste documento, ou variantes destes, utilizando células hospedeiras métodos livres de células.

Tabela 1: Genes biossintéticos homólogos de aldeídos e alcanos em genomas cianobacteriais

Cianobactéria	Gene de Alcano Biosint.		Gene de Aldeído Biosint.	
	número de acesso	%ID	número de acesso	%ID
<i>Synechococcus elongatus</i> PCC 7942	YP_400610	100	YP_400611	100
<i>Synechococcus elongatus</i> PCC 6301	YP_170760	100	YP170761	100
<i>Microcoleus chthonoplastes</i> PCC 7420	EDX75019	77	EDX74978	70
<i>Arthrospira maxima</i> CS-328	EDZ94963	78	EDZ94968	68
<i>Lyngbya sp.</i> PCC 8106	ZP_01619575	77	ZP_01619574	69
<i>Nodularia spumigena</i> CCY9414	ZP_01628096	77	ZP_01628095	70
<i>Trichodesmium erithraeum</i> IMS101	YP_721979	76	YP_721978	69
<i>Microcystis aeruginosa</i> NIES-843	YP_001660323	75	YP_001660322	68
<i>Microcystis aeruginosa</i> PCC 7806	CAO90780	74	CAO90781	67
<i>Nostoc sp.</i> PCC 7120	NP_489323	74	NP_489324	72
<i>Nostoc azollae</i> 0708	EEG05692	73	EEG05693	70
<i>Anabaena variabilis</i> ATCC 29413	YP_323043	74	YP_323044	73
<i>Crocospaera watsonii</i> WH 8501	ZP_00514700	74	ZP_00516920	67
<i>Synechocystis sp.</i> PCC 6803	NP_442147	72	NP_442146	68
<i>Synechococcus sp.</i> PCC 7335	EDX86803	73	EDX87870	67
<i>Cyanothece sp.</i> ATCC 51142	YP_001802195	73	YP_001802846	67
<i>Cyanothece sp.</i> CCY0110	ZP_01728578	72	ZP_01728620	68
<i>Nostoc punctiforme</i>	ZP_00108838	72	ZP_00108837	71

PCC 73102				
<i>Acaryochloris</i>				
<i>marina</i> MBIC11017	YP_001518340	71	YP_001518341	66
<i>Cyanothece</i> sp. PCC 7425	YP_002481151	71	YP_002481152	70
<i>Cyanothece</i> sp. PCC 8801	ZP_02941459	70	ZP_02942716	69
<i>Thermosynechococcus elongatus</i> BP-1	NP_682103	70	NP_682102	70
<i>Synechococcus</i> sp. JA-2-3B'a (2-13)	YP_478639	68	YP_478638	63
<i>Synechococcus</i> sp. RCC307	YP_001227842	67	YP_001227841	64
<i>Synechococcus</i> sp. WH 7803	YP_001224377	68	YP_001224378	65
<i>Synechococcus</i> sp. WH 8102	NP_897829	70	NP_897828	65
<i>Synechococcus</i> sp. WH 7805	ZP_01123214	68	ZP_01123215	65
<i>Synechococcus</i> não cultivado marinho tipo-A GOM 3012	ABD96376	70	ABD96375	65
<i>Synechococcus</i> sp. JA-3-3Ab	YP_473897	68	YP_473896	62
<i>Synechococcus</i> não cultivado marinho tipo-A GOM 306	ABD96328	70	ABD96327	65
<i>Synechococcus</i> não cultivado marinho tipo-A GOM 3M9	ABD96275	68	ABD96274	65
<i>Synechococcus</i> sp. CC9311	YP_731193	63	YP_731192	63
<i>Synechococcus</i> não cultivado marinho tipo-A 5B2	ABB92250	69	ABB92249	64
<i>Synechococcus</i> sp. WH 5701	ZP_01085338	66	ZP_01085337	67
<i>Gloeobacter violaceus</i> PCC 7421	NP_926092	63	NP_926091	67
<i>Synechococcus</i> sp. RS9916	ZP_01472594	69	ZP_01472595	66
<i>Synechococcus</i> sp. RS9917	ZP_01079772	68	ZP_01079773	65
<i>Synechococcus</i> sp. CC9605	YP_381055	66	YP_381056	66
<i>Cyanobium</i> sp. PCC 7001	EDY39806	64	EDY38361	64
<i>Prochlorococcus marinus</i> str. MIT 9303	YP_001016795	63	YP_001016797	66
<i>Prochlorococcus marinus</i> str. MIT9313	NP_895059	63	NP_895058	65
<i>Synechococcus</i> sp. CC9902	YP_377637	66	YP_377636	65
<i>Prochlorococcus marinus</i> str. MIT 9301	YP_001090782	62	YP_001090783	62
<i>Synechococcus</i> sp. BL107	ZP_01469468	65	ZP_01469469	65
<i>Prochlorococcus</i>	YP_001008981	62	YP_001008982	61

<i>marinus</i> str. AS9601 <i>Prochlorococcus</i> <i>marinus</i> str. MIT9312	YP_397029	62	YP_397030	61
<i>Prochlorococcus</i> <i>marinus</i> subsp. <i>pastoris</i> str. CCMP1986	NP_892650	60	NP_892651	63
<i>Prochlorococcus</i> <i>marinus</i> str. MIT 9211	YP_001550420	61	YP_001550421	63
<i>Cyanothece</i> sp. PCC 7425	YP_002483683	59	-	
<i>Prochlorococcus</i> <i>marinus</i> str. NATL2A	YP_293054	59	YP_293055	62
<i>Prochlorococcus</i> <i>marinus</i> str. NATL1A	YP_001014415	59	YP_00104416	62
<i>Prochlorococcus</i> <i>marinus</i> subsp. <i>marinus</i> str. CCMP1375	NP_874925	59	NP_874926	64
<i>Prochlorococcus</i> <i>marinus</i> str. MIT 9515_05961	YP_001010912	57	YP_001010913	63
<i>Prochlorococcus</i> <i>marinus</i> str. MIT 9215_06131	YP_001483814	59	YP_001483815	62
<i>Synechococcus</i> sp. RS9917	ZP_01080370	43	-	
<i>Synechococcus</i> não cultivado marinho tipo-A GOM 5D20			ABD96480	65

Tabela B [FIG. 38]. Lista de exemplos de homólogos de *Synechococcus elongatus* PCC7942 YP_400610 (Synpcc7942_1593) (SEQ ID NO:1) a partir de uma base de dados metagenômica.

Números de Acesso de 10 de abril de 2009

Número de Acesso	% Identidade	% Simila- ridade	Comprimento de Alinhamento
gi 135970898 gb EBL05614.1	59,5	74,3	237
gi 134964254 gb EBE59803.1	57,6	74,2	229
gi 142528845 gb ECY73505.1	60,4	77,9	222
gi 135713677 gb EBJ38387.1	61,1	78,7	221
gi 141225813 gb ECQ49060.1	59,7	77,8	221
gi 144115151 gb EDI97334.1	67,7	80,9	220
gi 142133005 gb ECV83152.1	67,3	80,9	220
gi 137965371 gb EBX01252.1	67,3	80,9	220
gi 134786157 gb EBD42319.1	67,3	80,9	220
gi 136216894 gb EBM66672.1	62,3	77,7	220

gi 143271262 gb EDE04654.1	63,0	78,5	219
gi 135973786 gb EBL07573.1	63,0	78,1	219
gi 140222739 gb ECK35865.1	63,0	77,6	219
gi 139710482 gb ECG93903.1	62,6	78,1	219
gi 140109767 gb ECJ60294.1	62,6	78,5	219
gi 137939755 gb EBW86789.1	62,6	78,1	219
gi 140086977 gb ECJ44922.1	62,6	78,1	219
gi 143729007 gb EDG48476.1	62,6	78,1	219
gi 143217179 gb EDD66368.1	62,6	78,1	219
gi 138728422 gb ECB60007.1	62,6	78,1	219
gi 143540790 gb EDF53461.1	62,6	78,1	219
gi 143580323 gb EDF73830.1	62,6	78,1	219
gi 137317024 gb EBT41871.1	62,6	78,1	219
gi 143567212 gb EDF67415.1	62,6	78,1	219
gi 140726723 gb ECN09681.1	62,6	78,1	219
gi 136249402 gb EBM88688.1	62,6	78,1	219
gi 141156650 gb ECQ02018.1	62,6	78,1	219
gi 143550472 gb EDF58473.1	62,6	78,1	219
gi 139581046 gb ECG04925.1	62,6	78,1	219
gi 141904835 gb ECU06854.1	62,6	78,1	219
gi 143596624 gb EDF78559.1	62,1	78,1	219
gi 142808717 gb EDA73967.1	62,1	78,1	219
gi 137639134 gb EBV19468.1	62,1	78,1	219
gi 140970945 gb ECO75236.1	62,1	77,6	219
gi 137724335 gb EBV66164.1	62,1	78,1	219
gi 143095952 gb EDC78454.1	66,1	80,3	218
gi 135919851 gb EBK71172.1	66,1	80,3	218
gi 143083445 gb EDC69308.1	66,1	80,3	218
gi 137949431 gb EBW92260.1	63,1	77,6	214
gi 136336883 gb EBN48108.1	62,2	78,0	214
gi 136008042 gb EBL28916.1	62,2	78,0	214
gi 134796061 gb EBD49256.1	60,8	76,6	214
gi 143142117 gb EDD12240.1	60,8	76,6	214
gi 141812906 gb ECT63492.1	60,3	77,6	214
gi 137436109 gb EBU09046.1	60,3	77,1	214

gi 142796298 gb EDA64685.1	71,0	83,3	210
gi 138143480 gb EBX98705.1	70,5	83,3	210
gi 136004890 gb EBL27104.1	68,6	82,4	210
gi 142206956 gb ECW39307.1	67,1	82,4	210
gi 135813588 gb EBK00762.1	62,4	78,6	210
gi 137008728 gb EBR69250.1	62,7	79,0	209
gi 141443295 gb ECR86509.1	63,0	78,9	208
gi 135999440 gb EBL23890.1	63,1	79,1	206
gi 136001501 gb EBL25083.1	62,6	77,7	206
gi 137008499 gb EBR69124.1	62,3	77,5	204
gi 135813080 gb EBK00444.1	60,8	77,5	204
gi 139947991 gb ECI56786.1	63,6	79,8	198
gi 136255251 gb EBM92608.1	61,1	77,3	198
gi 141717528 gb ECS91623.1	58,6	76,8	198
gi 141153056 gb ECP99448.1	59,3	75,3	194
gi 140654921 gb ECM59517.1	67,9	81,9	193
gi 140993407 gb ECO91072.1	63,9	79,6	191
gi 143171676 gb EDD33295.1	63,9	79,6	191
gi 143659341 gb EDG12240.1	59,5	75,8	190
gi 138539256 gb ECA29247.1	68,1	81,9	188
gi 141964470 gb ECU48335.1	64,5	80,3	183
gi 139227663 gb ECE28885.1	63,9	79,2	183
gi 135926500 gb EBK75672.1	58,2	76,4	182
gi 140708786 gb ECM97174.1	68,4	81,4	177
gi 139523141 gb ECF65392.1	62,6	79,3	174
gi 137874181 gb EBW49523.1	61,5	76,4	174
gi 143221750 gb EDD69688.1	60,9	75,9	174
gi 140086962 gb ECJ44914.1	63,4	80,2	172
gi 142781070 gb EDA53384.1	61,8	77,1	170
gi 139775004 gb ECH37282.1	61,0	76,3	169
gi 136260534 gb EBM96169.1	61,0	75,7	169
gi 137441185 gb EBU11854.1	58,1	75,5	167
gi 136330550 gb EBN43814.1	65,1	80,7	166
gi 139650149 gb ECG51660.1	67,3	80,6	165
gi 143638003 gb EDF99869.1	63,6	80,0	165

gi 137949739 gb EBW92432.1	61,7	75,9	162
gi 143382653 gb EDE68551.1	63,9	81,0	158
gi 138989189 gb ECC70595.1	63,9	80,4	158
gi 138408887 gb EBZ46853.1	64,1	78,5	156
gi 137230040 gb EBS93199.1	61,3	76,1	155
gi 141605381 gb ECS53894.1	65,6	81,8	154
gi 137858747 gb EBW40699.1	64,9	81,8	154
gi 140209383 gb ECK27191.1	66,0	79,1	153
gi 142753984 gb EDA33411.1	60,7	75,3	150
gi 137242084 gb EBS99775.1	64,9	82,4	148
gi 136229422 gb EBM75188.1	64,6	82,3	147
gi 140311369 gb ECK89744.1	73,3	85,6	146
gi 140866197 gb ECO03641.1	65,1	82,2	146
gi 139229558 gb ECE29833.1	61,4	76,6	145
gi 141659030 gb ECS68172.1	61,4	76,6	145
gi 139580852 gb ECG04786.1	65,3	81,9	144
gi 138338712 gb EBZ05758.1	65,3	81,9	144
gi 136204827 gb EBM58548.1	61,3	76,8	142
gi 139095530 gb ECD38154.1	64,0	81,3	139
gi 136351648 gb EBN58190.1	60,1	76,1	138
gi 138155154 gb EBY06350.1	67,9	80,3	137
gi 137644530 gb EBV22059.1	63,1	80,8	130
gi 143775710 gb EDG72409.1	61,5	79,2	130
gi 143500330 gb EDF32920.1	63,1	77,9	122
gi 139709584 gb ECG93249.1	71,1	85,1	121
gi 142537519 gb ECY79816.1	62,8	81,0	121
gi 137944410 gb EBW89433.1	64,2	78,3	120
gi 137387955 gb EBT81682.1	63,9	78,2	119
gi 139955976 gb ECI62054.1	71,2	84,8	118
gi 137251843 gb EBT05348.1	62,1	77,6	116
gi 138442523 gb EBZ70326.1	62,1	76,7	116
gi 141590592 gb ECS49420.1	60,7	78,6	112
gi 143187997 gb EDD45026.1	60,7	77,7	112
gi 143655969 gb EDG10472.1	61,5	78,0	109
gi 139459255 gb ECF24788.1	62,9	79,1	105

gi 141976584 gb ECU56751.1	62,9	79,1	105
gi 139233104 gb ECE31190.1	69,6	83,3	102
gi 139233107 gb ECE31193.1	70,3	82,2	101
gi 138582711 gb ECA59616.1	61,4	77,2	101
gi 138442855 gb EBZ70514.1	61,4	75,3	101
gi 137662676 gb EBV31757.1	60,4	76,2	101
gi 139846062 gb ECH87248.1	62,6	74,8	99
gi 136935327 gb EBR27657.1	60,6	74,5	94
gi 137466252 gb EBU25703.1	67,0	80,2	91
gi 137820604 gb EBW18665.1	63,7	80,2	91
gi 138539295 gb ECA29276.1	62,6	79,1	91
gi 136239262 gb EBM81844.1	64,7	81,2	85
gi 136294902 gb EBN19768.1	64,7	81,2	85
gi 137413136 gb EBT96003.1	61,2	76,5	85
gi 137641041 gb EBV20354.1	60,0	72,9	85
gi 142508710 gb ECY58869.1	64,3	82,1	84
gi 140096399 gb ECJ51008.1	62,2	81,7	82
gi 137938664 gb EBW86178.1	64,0	81,3	75
gi 137275448 gb EBT18729.1	58,1	70,3	74
gi 141955842 gb ECU42610.1	63,0	80,8	73
gi 139221707 gb ECE24659.1	66,2	83,1	71
gi 142508709 gb ECY58868.1	52,9	68,6	70
gi 137523719 gb EBU55323.1	65,2	82,6	69
gi 140781524 gb ECN46583.1	65,6	82,8	64
gi 137627577 gb EBV13553.1	60,0	73,3	60
gi 141951833 gb ECU39722.1	59,7	73,7	57
gi 137232510 gb EBS94613.1	59,7	73,7	57

Corte utilizado: >50% de identidade e >25% de comprimento do
synpcc7942_1593

Tabela C [FIG. 39]. Lista de exemplos de homólogos de *Synechococcus elongatus* PCC7942 YP_400611 (Synpcc7942_1594) (SEQ ID NO:65) a partir de uma base de dados metagenômica.

Números de Acesso de 10 de abril de 2009

Número de Acesso	% Identidade	% Similari- dade	Comprimento de Alinhamento
gi 143288250 gb EDE13503.1	71,3	80,5	87

gi 142342310 gb ECX39602.1	71,3	80,5	87
gi 137959588 gb EBW92346.1	70,7	85,9	92
gi 139984340 gb ECI81897.1	70,6	78,8	85
gi 140249046 gb ECK54318.1	69,7	82,0	267
gi 142111437 gb ECV67406.1	69,2	81,3	182
gi 142994709 gb EDC04737.1	68,8	81,2	138
gi 143066602 gb EDC56955.1	67,7	81,5	248
gi 138840827 gb ECC11022.1	67,1	78,8	146
gi 137829071 gb EBW23606.1	66,7	77,1	96
gi 142133008 gb ECV83155.1	66,5	79,5	337
gi 143095956 gb EDC78458.1	66,4	79,1	339
gi 140732156 gb ECN13587.1	66,1	79,5	254
gi 136241230 gb EBM83170.1	66,0	78,6	103
gi 140001769 gb ECI93451.1	65,9	76,9	91
gi 139305662 gb ECE48752.1	65,9	80,0	205
gi 137634503 gb EBV17219.1	65,8	80,7	114
gi 138584841 gb ECA61142.1	65,5	81,9	116
gi 135919849 gb EBK71170.1	65,4	76,6	107
gi 141153057 gb ECP99449.1	65,4	79,6	280
gi 141976585 gb ECU56752.1	65,3	77,6	98
gi 142206955 gb ECW39306.1	65,3	78,5	340
gi 141804802 gb ECT36785.1	65,3	76,8	95
gi 138931154 gb ECC47219.1	65,1	76,7	86
gi 138408888 gb EBZ46854.1	65,1	78,0	255
gi 134743188 gb EBD14908.1	64,8	78,8	307
gi 138168794 gb EBY16028.1	64,8	77,6	304
gi 135749749 gb EBJ60721.1	64,8	79,1	105
gi 138338711 gb EBZ05757.1	64,6	80,0	175
gi 142827948 gb EDA88477.1	64,5	77,7	121
gi 138361576 gb EBZ15968.1	64,4	78,7	267
gi 140517919 gb ECM08416.1	64,4	77,4	115
gi 135813081 gb EBK00455.1	64,3	79,7	143
gi 137627576 gb EBV13552.1	64,3	80,1	171
gi 141161845 gb ECQ05757.1	64,2	77,1	109
gi 140992134 gb ECO90156.1	64,2	76,8	95

gi 137796334 gb EBW04596.1	64,2	78,9	279
gi 137619413 gb EBV08950.1	64,1	77,6	223
gi 140517917 gb ECM08414.1	64,1	79,6	181
gi 137232509 gb EBS94612.1	64,1	79,0	181
gi 135811491 gb EBJ99446.1	64,0	77,7	314
gi 141167040 gb ECQ09480.1	64,0	78,8	250
gi 143441820 gb EDE97777.1	64,0	78,2	261
gi 140311368 gb ECK89743.1	64,0	79,3	261
gi 140970943 gb ECO75234.1	64,0	76,6	111
gi 136817739 gb EBQ60666.1	63,8	78,8	250
gi 141717529 gb ECS91624.1	63,7	78,8	146
gi 137632337 gb EBV16047.1	63,7	77,7	256
gi 137662677 gb EBV31758.1	63,6	79,7	143
gi 140091056 gb ECJ47190.1	63,6	78,2	280
gi 143217178 gb EDD66367.1	63,5	78,8	137
gi 139984339 gb ECI81896.1	63,5	76,3	156
gi 139382506 gb ECE73591.1	63,4	78,9	194
gi 140096397 gb ECJ51006.1	63,4	79,7	153
gi 134606350 gb EBC34611.1	63,4	77,9	131
gi 140705175 gb ECM95033.1	63,4	77,0	191
gi 139846064 gb ECH87250.1	63,1	76,0	179
gi 137953535 gb EBW94572.1	63,1	78,0	241
gi 143738737 gb EDG53066.1	63,1	76,6	111
gi 141951832 gb ECU39721.1	63,1	76,6	111
gi 143271261 gb EDE04653.1	63,1	76,6	111
gi 139846065 gb ECH87251.1	63,0	78,8	146
gi 137251844 gb EBT05349.1	63,0	78,8	146
gi 136249401 gb EBM88687.1	63,0	78,8	146
gi 134628580 gb EBC48074.1	63,0	77,0	100
gi 136312048 gb EBN31461.1	62,9	76,2	143
gi 143221751 gb EDD69689.1	62,9	78,9	194
gi 141955844 gb ECU42612.1	62,9	75,7	140
gi 143395654 gb EDE73119.1	62,8	77,5	218
gi 142781071 gb EDA53385.1	62,8	79,3	164
gi 136303394 gb EBN25555.1	62,7	77,7	319

gi 143557688 gb EDF62238.1	62,6	77,9	131
gi 136008043 gb EBL28917.1	62,6	77,6	294
gi 143596625 gb EDF78560.1	62,6	77,0	318
gi 137641042 gb EBV20355.1	62,6	79,1	163
gi 136231267 gb EBM76426.1	62,5	76,7	339
gi 143175604 gb EDD36054.1	62,5	76,7	339
gi 142508708 gb ECY58867.1	62,5	76,7	339
gi 135926501 gb EBK75673.1	62,5	76,6	320
gi 141955884 gb ECU42641.1	62,4	77,8	189
gi 142821119 gb EDA83282.1	62,4	78,7	202
gi 134609411 gb EBC36492.1	62,4	75,8	194
gi 142885864 gb EDB27722.1	62,3	75,4	207
gi 136204828 gb EBM58549.1	62,3	77,0	318
gi 143580324 gb EDF73831.1	62,2	76,7	339
gi 143766375 gb EDG67769.1	62,2	77,0	339
gi 143500332 gb EDF32922.1	62,2	76,4	339
gi 139233105 gb ECE31191.1	62,1	76,5	132
gi 143738779 gb EDG53089.1	62,1	76,8	314
gi 134964255 gb EBE59804.1	62,1	77,1	153
gi 140863545 gb ECO01751.1	62,1	77,0	269
gi 137944409 gb EBW89432.1	62,0	78,5	158
gi 143411619 gb EDE81261.1	62,0	76,7	339
gi 142753988 gb EDA33415.1	62,0	77,0	339
gi 139580853 gb ECG04787.1	61,9	75,7	202
gi 141227933 gb ECQ50606.1	61,9	75,1	173
gi 143659340 gb EDG12239.1	61,8	78,3	157
gi 136935328 gb EBR27658.1	61,8	77,7	157
gi 137235449 gb EBT18730.1	61,8	77,5	204
gi 138585243 gb ECA61437.1	61,8	76,5	136
gi 139195947 gb ECE06889.1	61,7	76,2	269
gi 139424973 gb ECF02640.1	61,7	76,3	274
gi 141380828 gb ECR42772.1	61,7	75,8	227
gi 136351647 gb EBN58189.1	61,7	77,0	339
gi 136304410 gb EBN26254.1	61,6	75,7	185
gi 139948037 gb ECI56814.1	61,6	76,8	224

gi 135970899 gb EBL05615.1	61,6	76,8	211
gi 138627165 gb ECA90647.1	61,5	75,5	143
gi 137395720 gb EBT86160.1	61,5	74,8	143
gi 140086960 gb ECJ44912.1	61,5	77,1	249
gi 141024916 gb ECP11582.1	61,4	76,7	223
gi 139095531 gb ECD38155.1	61,4	76,7	210
gi 141659029 gb ECS68171.1	61,4	77,2	127
gi 139969430 gb ECI71470.1	61,4	76,7	215
gi 136986729 gb EBR56775.1	61,4	75,7	202
gi 143634197 gb EDF97600.1	61,3	76,7	313
gi 135973785 gb EBL07572.1	61,3	76,3	279
gi 143600944 gb EDD54508.1	61,3	74,6	173
gi 137787263 gb EBV99371.1	61,2	76,3	232
gi 139204136 gb ECE12313.1	61,2	76,7	219
gi 136001500 gb EBL25082.1	61,2	76,3	219
gi 141874476 gb ECT85572.1	61,2	76,4	237
gi 137905325 gb EBW67375.1	61,1	74,9	175
gi 140089341 gb ECJ46519.1	61,1	75,4	203
gi 140855194 gb ECN95754.1	61,1	75,9	203
gi 134965622 gb EBE60718.1	61,1	75,8	339
gi 141527125 gb ECS15588.1	61,1	76,0	208
gi 136218988 gb EBM68086.1	61,0	73,4	154
gi 142364499 gb ECX54765.1	60,9	76,1	330
gi 141603393 gb ECS53340.1	60,8	74,8	143
gi 136216893 gb EBM66671.1	60,8	76,3	245
gi 143743653 gb EDG56305.1	60,8	73,7	148
gi 140222741 gb ECK35867.1	60,7	73,0	163
gi 142389887 gb ECX71634.1	60,6	74,8	198
gi 139315697 gb ECE51398.1	60,6	76,4	241
gi 136255250 gb EBM92607.1	60,5	76,0	291
gi 137387954 gb EBT81681.1	60,5	76,6	124
gi 139229559 gb ECE29834.1	60,4	75,3	202
gi 140866196 gb ECO03646.1	60,1	74,6,	303
gi 139229561 gb ECE29836.1	60,0	74,0	100
gi 140957440 gb ECO66006.1	59,9	75,9	274

gi 143567213 gb EDF67416.1	59,8	74,8	286
gi 139955973 gb ECI62051.1	59,8	72,2	97
gi 140726724 gb ECN09682.1	59,7	73,4	154
gi 139775003 gb ECH37281.1	59,6	73,7	99
gi 137949740 gb EBW92733.1	58,7	71,7	92
gi 139650150 gb ECG51661.1	57,7	73,2	97
gi 142528844 gb ECY73504.1	56,7	70,0	90

Corte utilizado: >50% de identidade e >25% do comprimento do
synpcc7942_1594

Genes biossintéticos de aldeído, alceno e alceno e variantes

[176] Os métodos e as composições descritos neste documento incluem, por exemplo, genes biossintéticos de alceno ou alceno contendo a sequência de nucleotídeo de SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33 ou 35, assim como as variantes de polinucleotídeos desta. Em algumas instâncias, o gene biossintético de alceno ou alceno codifica um ou mais motivos de aminoácidos descritos neste documento. Por exemplo, o gene biossintético de alceno ou alceno pode codificar um polipeptídeo compreendendo a SEQ ID NO: 37, 38, 39, 41, 42, 43 ou 44. O gene biossintético de alceno ou alceno pode também incluir um polipeptídeo compreendendo a SEQ ID NO: 40 e também qualquer uma de SEQ ID NO: 37, 38 ou 39.

[177] Os métodos e as composições descritas neste documento também incluem, por exemplo, genes biossintéticos de aldeído contendo a sequência de nucleotídeos de SEQ ID NO: 65, 67, 69, 71, 73, 75, 77, 79 ou 81, assim como variantes de polinucleotídeos desta. Em algumas instância, o gene biossintético de aldeído codifica um ou mais motivos de aminoácido descrito neste documento. Por exemplo, o gene biossintético de aldeído pode codificar um polipeptídeo compreendendo a SEQ ID NO: 54, 55, 56, 57, 58, 59, 60, 61,

62, 63 ou 64.

[178] As variantes podem ocorrer naturalmente ou ser criadas *in vitro*. Em particular, tais variantes podem ser criadas usando técnicas de engenharia genética, tais como mutagênese direcionada local, mutagênese química aleatória, procedimentos de supressão Exonuclease III e técnicas de clonagem padrão. Alternativamente, tais variantes, fragmentos, análogos ou derivados podem ser criados usando síntese química ou procedimentos de modificação.

[179] Os métodos de produção de variantes são bem-conhecidos no estado da técnica. Eles incluem procedimentos nos quais sequências de ácido nucleico obtidos a partir de isolados naturais são modificados para gerar ácidos nucleicos que codifiquem polipeptídeos contendo características que melhorem seus valores nas aplicações industriais ou laboratoriais. Em tais procedimentos, um grande número de sequências de variantes contendo uma ou mais diferenças de nucleotídeos com respeito à sequência obtida a partir do isolado natural são geradas e caracterizadas. Tipicamente, essas diferenças de nucleotídeo resultam em mudanças de aminoácidos em relação aos polipeptídeos codificados pelos ácidos nucleicos a partir dos isolados naturais.

[180] Por exemplo, variantes podem ser criadas usando-se PCR propensa a erro (ver, por exemplo, LEUNG *et al. Technique* 1:11-15, 1989; e CALDWELL *et al., PCR Methods Applic.* 2:28-33, 1992). Em PCR propensa a erro, PCR é realizada sob condições nas quais a fidelidade de cópia da polimerase do DNA é baixa, tal que uma elevada taxa de pontos de mutações é obtida ao longo de todo o comprimento do

produto da PCR. Brevemente, em tais procedimentos, os ácidos nucleicos a serem mutagenizados (por exemplo, uma sequência de polinucleotídeos biossintéticos de aldeído ou alceno) são misturados com os detonadores de PCR, o tamponador de reação, $MgCl_2$, $MnCl_2$, Taq polimerase e uma concentração apropriada de dNTPs para alcançar uma taxa elevada de ponto de mutação ao longo de todo o comprimento do produto da PCR. Por exemplo, a reação pode ser realizada usando 20 fmoles de ácido nucleico a ser mutagenizado (por exemplo, uma sequência de polinucleotídeos biossintéticos de aldeído ou alceno), 30 pmoles de cada detonador de PCR, um tamponador de reação compreendendo 50 mM de KCl, 10 mM de Tris HCl (pH 8,3) e 0,01% de gelatina, 7 mM de $MgCl_2$, 0,5 mM de $MnCl_2$, 5 unidades de Taq polimerase, 0,2 mM de dGTP, 0,2 mM de dATP, 1 mM de dCTP e 1 mM de dTTP. A PCR pode ser realizada por 30 ciclos de 94°C por 1 min, 45°C por 1 min e 72°C por 1 min. Entretanto, apreciar-se-á que esses parâmetros possam ser variados como apropriado. Os ácidos nucleicos mutagenizados são então clonados dentro de um vetor apropriado e as atividades do polipeptídeos codificadas pelo ácido nucleico mutagenizado são avaliadas.

[181] As variantes podem também ser criadas usando-se mutagênese direta de oligonucleotídeos para gerar mutações em lugares específicos em qualquer DNA clonado de interesse. A mutagênese de oligonucleotídeo é descrita em, por exemplo, Reidhaar-Olson *et al.*, *Science* 241:53-57, 1988. Brevemente, em tais procedimentos uma pluralidade de oligonucleotídeos de filamento duplo carregando uma ou mais mutações para serem introduzidas dentro do DNA clonado são sintetizadas e inseridas dentro do DNA clonado para serem mutagenizadas

(por exemplo, uma sequência de polinucleotídeos biossintéticos de aldeído ou alceno). Os clones contendo o DNA mutagenizado são recuperados, e as atividades dos polipeptídeos que eles codificam são avaliadas.

[182] Outro método para a geração de variantes é a montagem por PCR. A montagem por PCR envolve a montagem de um produto de PCR a partir de uma mistura de pequenos fragmentos de DNA. Um grande número de reações de PCR diferentes ocorre em paralelo no mesmo frasco, com os produtos de uma reação detonando os produtos de outra reação. A montagem por PCR é descrita em, por exemplo, patente norte-americana nº 5.965.408.

[183] Ainda outro método de geração de variantes é a mutagênese sexual da PCR. Na mutagênese sexual da PCR, a recombinação homóloga forçada ocorre entre moléculas de DNA de diferente, mal altamente relacionadas, sequência de DNA *in vitro* como resultado da fragmentação aleatória da molécula de DNA baseada na homologia de sequência. Isso é seguido pela fixação do cruzamento pela extensão de detonação em uma reação de PCR. A mutagênese sexual de PCR é descrita em, por exemplo, Stemmer, *PNAS, USA* 91:10747-10751, 1994.

[184] Variantes também podem ser criadas por mutagêneses *in vitro*. Em algumas concretizações, mutações aleatórias em uma sequência de ácido nucleico são geradas pela propagação da sequência em um filamento bacteriano, tais como um filamento de *E. coli*, que carrega mutações em uma ou mais das vias de reparação de DNA. Tais filamentos "mutadores" possuem uma taxa de mutação aleatória mais alta que aquela de um filamento do tipo selvagem. A propagação de uma sequência de DNA (por exemplo, uma sequência de

polinucleotídeos biossintéticos de aldeído ou alceno) em um desses filamentos eventualmente gerará mutações aleatórias dentro do DNA. Filamentos mutadores adequados para o uso para mutagênese *in vivo* são descritas, por exemplo, na publicação de PCT n° WO 91/16427.

[185] Variantes também podem ser geradas usando mutagênese de cassete. Em mutagênese de cassete, uma pequena região de uma molécula de DNA duplamente filamentada é substituída com um "cassete" de oligonucleotídeo sintético que difere da sequência nativa. O oligonucleotídeo frequentemente contém uma sequência nativa completamente e/ou parcialmente randomizada.

[186] A mutagênese de conjunto recursivo também pode ser usada para gerar variantes. A mutagênese de conjunto recursivo é um algoritmo para a engenharia de proteína (*i.e.*, mutagênese de proteína) desenvolvida para produzir populações diversas de mutantes fenotipicamente relacionados cujos membros difiram em sequência de aminoácidos. Esse método utiliza um mecanismo de resposta para controlar as sucessivas rodadas de mutagênese de cassete combinatória. Mutagênese de conjunto recursivo é descrito em, por exemplo, Arkin *et al.* *PNAS, USA* 89:7811-7815, 1992.

[187] Em algumas concretizações, as variantes são criadas usando-se mutagênese de conjunto exponencial. Mutagênese de conjunto exponencial é um processo para gerar bibliotecas combinatórias com uma elevada porcentagem de mutantes únicos e funcionais, onde grupos pequenos de resíduos sejam randomizados em paralelo para identificar, em cada posição alterada, os aminoácidos que conduzem às proteínas funcionais. Mutagênese de conjunto exponencial é

descrita em, por exemplo, Delegrave *et al.*, *Biotech. Res.* 11:1548-1552, 1993. Mutagênese aleatória e direcionada de local são descritas em, por exemplo, Arnold, *Curr. Opin. Biotech.* 4:450-455, 1993.

[188] Em algumas concretizações, as variantes são criadas usando-se procedimentos de embaralhamento, nos quais porções de uma pluralidade de ácidos nucleicos que codifiquem polipeptídeos distintos são fundidas juntas para criar sequências quiméricas de ácido nucleico que codifiquem polipeptídeos quiméricos, como descrito em, por exemplo, patentes norte-americanas nº 5.965.408 e 5.939.250.

[189] Variantes de polinucleotídeos ainda incluem análogos de ácido nucleico. Análogos de ácido nucleico podem ser modificados na metade de base, na metade de açúcar ou no suporte principal de fosfato para melhorar, por exemplo, a estabilidade, a hibridização ou a solubilidade do ácido nucleico. Modificações na metade de base incluem desoxiuridina para desoxitimidina e 5-metil-2'-desoxicitidina ou 5-bromo-2'-desoxicitidina para desoxicitidina. Modificações na metade de açúcar incluem a modificação da 2' hidroxila do açúcar de ribose para formar os açúcares 2'-O-metil ou 2'-O-alila. O suporte principal de fosfato de desoxirribose pode ser modificado para produzir ácidos nucleicos morfolinos, nos quais cada metade de base está ligada e um anel morfolino de 6 membros ou a ácido nucleico de peptídeo, no qual o suporte principal de desoxifosfato é substituído por um suporte principal de pseudopeptídeo e as quatro bases são retidas. (Ver, por exemplo, Summerton *et al.*, *Antisense Nucleic Acid Drug Dev.* (1997) 7:187-195; e Hyrup *et al.*, *Bioorgan. Med. Chem.* (1996)

4:5-23.) Além disso, o suporte principal de desoxifosfato pode ser substituído com, por exemplo, um suporte principal de fosforotioato ou fosforoditioato, uma fosforoamidita ou um suporte principal de fosfotriéster de alquila.

[190] Os polipeptídeos biossintéticos de aldeído e alceno Synpcc7942_1594 (SEQ ID NO: 66) e Synpcc7942_1593 (SEQ ID NO:2) possuem homólogos em outra cianobactéria (exemplos não limitantes são retratados na Tabela 1). Assim, qualquer sequência de polinucleotídeo codificando um homólogo listado na Tabela 1, ou uma variante deste, pode ser usada como polinucleotídeo biossintético de aldeído ou alceno nos métodos descritos neste documento. Cada cianobactéria listada na Tabela 1 possui cópias de ambos os genes. O nível de identificação da sequência do gene produz taxas a partir de 61% a 73% para Synpcc7942_1594 (SEQ ID NO:66) e a partir de 43% a 78% para Synpcc7942_1593 (SEQ ID NO:2).

[191] Outros homólogos do polipeptídeo biossintético de aldeído Synpcc7942_1594 (SEQ ID NO: 66) estão listados na Tabela C, e qualquer sequência de polinucleotídeo codificando um homólogo listado na Tabela C, ou uma variante desta, pode ser usada como um polinucleotídeo biossintético de aldeído nos métodos descritos neste documento. Outros homólogos do polipeptídeo biossintético de alceno Synpcc7942_1593 (SEQ ID NO:2) estão listados na Tabela B, e qualquer sequência de polinucleotídeo codificando um homólogo listado na Tabela B, ou uma variante desta, pode ser usada como um polinucleotídeo biossintético de alceno nos métodos descritos neste documento.

[192] Em certas instâncias, um gene biossintético de

aldeído, alceno e/ou alceno é otimizado de códon para a expressão em uma célula hospedeira. Por exemplo, para a expressão em *E. coli*, um ou mais códons podem ser otimizados como descritos em, por exemplo, Grosjean *et al.*, *Gene* 18:199-209 (1982).

Polipeptídeos biossintéticos de aldeído, alceno e alceno e variantes

[193] Os métodos e as composições descritos neste documento também incluem polipeptídeos biossintéticos de alceno ou alceno contendo a sequência de aminoácido de SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34 ou 36, assim com variantes de polipeptídeos desta. Em algumas instâncias, um polipeptídeo biossintético de alceno ou alceno é aquele que inclui um ou mais dos motivos de aminoácidos descritos neste documento. Por exemplo, o polipeptídeo biossintético de alceno ou alceno pode incluir a sequência de aminoácido de SEQ ID NO: 37, 38, 39, 41, 42, 43 e 44. O polipeptídeo biossintético de alceno ou alceno pode também incluir a sequência de aminoácido de SEW ID N°: 40 e também qualquer uma de SEQ ID NO: 37, 38 ou 39.

[194] Os métodos e as composições descritas neste documento também incluem polipeptídeo biossintéticos de aldeído contendo a sequência de aminoácidos de SEQ ID NO: 66, 68, 70, 72, 74, 76, 78, 80 ou 82, assim como variantes de polipeptídeos desta. Em algumas instâncias, um polipeptídeo biossintético de aldeído é aquele que inclui um ou mais dos motivos de aminoácidos descritos neste documento. Por exemplo, o polipeptídeo biossintético de aldeído pode incluir uma sequência de aminoácido de SEQ ID NO: 54, 55, 56, 57, 58, 59, 60, 61, 62, 63 ou 64.

[195] As variantes de polipeptídeo biossintético de aldeído, alceno e alceno podem ser variantes nas quais um ou mais resíduos sejam substituídos com um resíduo de aminoácido conservado ou não conservado (preferencialmente um resíduo conservado de aminoácido). Tal resíduo substituído de aminoácido pode ou não pode ser aquele codificado pelo código genético.

[196] Substituições conservadoras são aquelas que substituem um aminoácido dado em um polipeptídeo por outro aminoácido de características semelhantes. Substituições conservadoras típicas são as seguintes substituições: substituições de um aminoácido alifático, tal como alanina, valina, leucina e isoleucina, com outro aminoácido alifático; substituição de uma serina com uma treonina, ou vice-versa; substituição de um resíduo ácido, tal como ácido aspártico e ácido glutâmico, com outro resíduo ácido; substituição de um resíduo carregando um grupo amida, tal como asparagina e glutamina, com outro resíduo carregando um grupo amida; mudança de um resíduo básico, tal como lisina e arginina, com outro resíduo básico; e substituição de um resíduo aromático, tal como fenilalanina e tirosina, com outro resíduo aromático.

[197] Outras variantes de polipeptídeo são aquelas nas quais um ou mais resíduos de aminoácido incluem um grupo substituto. Ainda outras variantes de polipeptídeo são aquelas nas quais o polipeptídeo esteja associado com outro composto, tal como um composto para aumentar a meia-vida do polipeptídeo (por exemplo, polietileno glicol).

[198] Variantes de polipeptídeo adicionais são aquelas nas quais aminoácidos adicionais são fundidos aos

polipeptídeos, tais como uma sequência líder, uma sequência secretória, uma sequência de proteína ou uma sequência que facilite a purificação, o enriquecimento ou a estabilização do polipeptídeo.

[199] Em algumas instâncias, uma variante de polipeptídeo biossintético de alcanos ou alcenos retém a mesma função biológica que um polipeptídeo contendo uma sequência de aminoácido de SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34 ou 36 (por exemplo, retém a atividade biossintética de alcanos ou alcenos) e possui uma sequência de aminoácido substancialmente idêntica a esta.

[200] Em outras instâncias, variantes de polipeptídeo biossintético de alcanos ou alcenos possuem pelo menos cerca de 50%, pelo menos cerca de 55%, pelo menos cerca de 60%, pelo menos cerca de 65%, pelo menos cerca de 70%, pelo menos cerca de 75%, pelo menos cerca de 80%, pelo menos cerca de 85%, pelo menos cerca de 90%, pelo menos cerca de 95% ou mais que cerca de 95% de homologia à sequência de aminoácido de SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34 ou 36. Em outra concretização, as variantes de polipeptídeo incluem um fragmento compreendendo pelo menos cerca de 5, 10, 15, 20, 25, 30, 35, 40, 50, 75, 100 ou 150 aminoácidos consecutivos deste.

[201] Em algumas instâncias, uma variante de polipeptídeo biossintético de aldeído retém a mesma função biológica que um polipeptídeo contendo uma sequência de aminoácido de SEQ ID NO: 66, 68, 70, 72, 74, 76, 78, 80 ou 82 (por exemplo, retém a atividade biossintética de aldeído) e possui uma sequência de aminoácido substancialmente idêntica a esta.

[202] Ainda em outras instâncias, variantes de polipeptídeo biossintético de aldeído possuem pelo menos cerca de 50%, pelo menos cerca de 55%, pelo menos cerca de 60%, pelo menos cerca de 65%, pelo menos cerca de 70%, pelo menos cerca de 75%, pelo menos cerca de 80%, pelo menos cerca de 85%, pelo menos cerca de 90%, pelo menos cerca de 95% ou mais que cerca de 95% de homologia à sequência de aminoácido de SEQ ID NO: 66, 68, 70, 72, 74, 76, 78, 80 e 82. Em outra concretização, as variantes de polipeptídeo incluem um fragmento compreendendo pelo menos cerca de 5, 10, 15, 20, 25, 30, 35, 40, 50, 75, 100 ou 150 aminoácidos consecutivos deste.

[203] As variantes de polipeptídeo ou os fragmentos deste podem ser obtidos isolando-se os ácidos nucleicos codificando-os usando as técnicas descritas neste documento ou expressando ácidos nucleicos sintéticos os codificando. Alternativamente, variantes de polipeptídeos ou fragmentos destes podem ser obtidos por meio de enriquecimento bioquímico ou por procedimentos de purificação. A sequência de variantes de polipeptídeos ou fragmentos podem ser determinados pela digestão proteolítica, pela eletroforese de gel e/ou pelo microssequenciamento. A sequência das variantes de polipeptídeo biossintético de alceno ou alceno ou fragmentos podem ser, então, comparados à sequência de aminoácido de SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34 ou 36 usando-se qualquer um dos programas descritos neste documento. A sequência de variantes de polipeptídeo biossintético de aldeído ou fragmentos podem ser comparados à sequência de aminoácido de SEQ ID NO: 66, 68, 70, 72, 74, 76, 78, 80 ou 82 usando-se

qualquer um dos programas descritos neste documento.

[204] As variantes de polipeptídeo e os fragmentos destas podem ser testados para a produção de atividade aldeído, álcool graxo, alceno e/ou alceno usando métodos rotineiros. Por exemplo, as variantes de polipeptídeo ou os fragmentos podem entrar em contato com um substrato (por exemplo, um substrato derivado de ácido graxo ou outro substrato descrito neste documento) sob condições que permitam à variante de polipeptídeo funcionar. Um decréscimo no nível do substrato ou um aumento no nível de um aldeído, alceno ou alceno pode ser medido para determinar a atividade de produção de aldeído, álcool graxo, alceno ou alceno, respectivamente.

Anticorpos de polipeptídeo biossintético de antialdeído, antiálcool graxo, antialceno e antialceno

[205] Os polipeptídeos biossintéticos de aldeído, álcool graxo, alceno e alceno descritos neste documento podem também ser usados para produzir anticorpos dirigidos contra polipeptídeos biossintéticos de aldeído, álcool graxo, alceno e alceno. Tais anticorpos podem ser usados, por exemplo, para detectar a expressão de um polipeptídeo biossintético de aldeído, álcool graxo, alceno e alceno usando métodos conhecidos no estado da arte. O anticorpo pode ser, por exemplo, um anticorpo policlonal; um anticorpo monoclonal ou fragmento de ligação de antígeno deste; um anticorpo modificado tal como um anticorpo quimérico, um anticorpo remodelado, um anticorpo humanizado ou um fragmento destes (por exemplo, Fab1, Fab, F(ab1)₂); ou um anticorpo biossintético, por exemplo, um anticorpo de cadeia simples, um anticorpo de domínio simples (DAB), Fv, um Fv de

cadeia simples (scFv), ou semelhantes.

[206] Os métodos para o preparo e o uso de anticorpos policlonais e monoclonais são descritos, por exemplo, em Harlow *et al.*, *Using Antibodies: A Laboratory Manual: Portable Protocol I*. Cold Spring Harbor Laboratory (1º de dezembro de 1998). Métodos para o preparo de anticorpos modificados e fragmentos de anticorpo (por exemplo, anticorpos quiméricos, anticorpos remodelados, anticorpos humanizados ou fragmentos destes, por exemplo, fragmentos de Fab', Fab, F(ab')₂); ou anticorpos biossintéticos (por exemplo, anticorpos de cadeia simples, anticorpos de domínio simples (DABs), Fv, Fv de cadeia simples (scFv), e semelhantes), são conhecidos no estado da técnica e podem ser encontrados, por exemplo, em Zola, *Monoclonal Antibodies: Preparation and Use of Monoclonal Antibodies and Engineered Antibody Derivatives*, Springer Verlag (15 de dezembro de 2000, 1ª edição).

Substratos

[207] As composições e os métodos descritos neste documento podem ser usados para produzir aldeídos, alcoóis graxos, alcanos e/ou alcenos a partir de um substrato apropriado. Com quanto não se deseja ser limitado por uma teoria particular, acredita-se que os polipeptídeos biossintéticos de alcano ou alceno descritos neste documento produzam alcanos ou alcenos a partir de substratos por meio de um mecanismo de descarbonilação. Em algumas instâncias, o substrato é um derivado de ácido graxo, por exemplo, um aldeído graxo, e um alcano contendo padrões de ramificação particulares e comprimento de cadeia de carbono pode ser produzido a partir de um derivado de ácido graxo, por

exemplo, um aldeído graxo, contendo aquelas características particulares. Em outras instâncias, o substrato é um derivado insaturado de ácido graxo, por exemplo, um aldeído graxo insaturado, e um alceno contendo padrões de ramificação particulares e comprimento de cadeia de carbono que podem ser produzidos a partir de um derivado insaturado de ácido graxo, por exemplo, um aldeído graxo insaturado, contendo aquelas características particulares.

[208] Conquanto não se deseja ser limitado por uma teoria particular, acredita-se que os polipeptídeos biossintéticos de aldeído descritos neste documento produzam aldeídos a partir de substratos por meio de um mecanismo de redução. Em certas instâncias, o substrato é um acil-ACP.

[209] Conquanto não se deseja ser limitado por uma teoria particular, acredita-se que os alcoóis graxos descritos neste documento sejam produzidos a partir de substratos por meio de um mecanismo de redução. Em certas instâncias, o substrato é um aldeído graxo.

[210] Em conformidade, cada etapa dentro de uma via biossintética que conduza à produção desses substratos pode ser modificada para produzir ou sobreproduzir o substrato de interesse. Por exemplo, os genes conhecidos envolvidos nas vias biossintéticas do ácido graxo, do aldeído graxo e do álcool graxo podem ser expressados, sobreexpressados ou atenuados na célula hospedeira para produzir um substrato desejado (ver, por exemplo, PCT/US08/058788, especificamente incorporado como referência neste documento). Exemplares de genes são fornecidos na Tabela A.

Síntese dos Substratos

[211] A sintase de ácido graxo (FAS) é um grupo de

polipeptídeos que catalisa a iniciação e o alongamento de cadeias acila (MARRAKCHI *et al.* *Biochemical Society*, 30:1050-1055, 2002). A proteína transportadora de acila (ACP) juntamente com as enzimas na via FAS controlam o comprimento, o grau de saturação e a ramificação dos derivados de ácido graxo produzidos. A vida biossintética de ácido graxo envolve os acetil-CoA e malonil-CoA precursores. As etapas nesta via são catalisadas por enzimas da biossíntese de ácido graxo (*FAB*) e família de gene de carboxila de acetil-CoA (*acc*) (ver, por exemplo, Heath *et al.* *Prog. Lipid Res.* 40(6):467-97 (2001)).

[212] Células hospedeiras podem ser projetadas para expressar substratos de derivados de ácido graxo expressando ou sobreexpressando recombinantemente os genes sintases de acetil-CoA e/ou malonil-CoA. Por exemplo, para aumentar a produção de acetil-CoA, um ou mais dos seguintes genes pode ser expressado na célula hospedeira: *pdh*, *panK*, *aceEF* (codificando o componente E1p desidrogenase e o componente E2p di-hidrolipoamida aciltransferase dos complexos piruvato e 2-oxoglutarato desidrogenase), *fabH*, *fabD*, *fabG*, *acpP*, e *fabF*. Exemplos de número de adesão GenBank para esses genes são: *pdh* (BAB34380, AAC73227, AAC73226), *panK* (também conhecido como *coaA*, AAC76952), *aceEF* (AAC73227, AAC73226), *fabH* (AAC74175), *fabD* (AAC74176), *fabG* (AAC74177), *acpP* (AAC74178), *fabF* (AAC74179). Adicionalmente, os níveis de expressão de *fadE*, *gspA*, *ldhA*, *pflb*, *adhE*, *pta*, *poxB*, *ackA* e/ou *ackB* podem ser atenuados ou desligados em um célula hospedeira manipulada pela transformação com plasmídeos condicionalmente replicativos ou não replicativos contendo mutações nulas ou mutações de supressão dos genes

correspondentes ou substituindo as sequências promotoras ou melhoradoras. Exemplos de número de adesão GenBank para esses genes são: *fadE* (AAC73325), *gspA* (AAC76632), *ldhA* (AAC74462), *pflb* (AAC73989), *adhE* (AAC74323), *pta* (AAC75357), *poxB* (AAC73958), *ackA* (AAC75356) e *ackB* (BAB81430). As células hospedeiras resultantes terão aumentado os níveis de produção de acetil-CoA quando crescidas em um ambiente apropriado.

[213] A sobreexpressão de malonil-CoA pode ser efetuada introduzindo-se *accABCD* (por exemplo, número de adesão AAC73296, EC 6,4,1,2) dentro de uma célula hospedeira. Os ácidos graxos podem ser, então, sobreexpressados em células hospedeiras introduzindo-se dentro da célula hospedeira uma sequência de DNA codificando uma lipase (por exemplo, números de adesão CAA89087, CAA98876).

[214] Além disso, a inibição de *PlsB* pode levar a um aumento dos níveis de acil-ACP de cadeia longa, o que inibirá as primeiras etapas na via (por exemplo, *accABCD*, *fabH* e *fabI*). A mutação *plsB* (por exemplo, número de adesão AAC77011) D311E pode ser usada para aumentar a quantidade de acil-CoA disponível.

[215] Além disso, uma célula hospedeira pode ser projetada para sobreexpressar um gene *sfa* (supressor de *fabA*, por exemplo, de número de adesão AAN79592) para aumentar a produção de ácidos graxos monoinsaturados (ROCK *et al.* *J. Bacteriology* 178:5382-5387, 1996).

[216] Em algumas instâncias, as células hospedeiras podem ser projetadas para expressar, sobreexpressar ou atenuar a expressão de uma tioesterase para aumentar a produção de substrato de ácido graxo. O comprimento da

cadeira de um substrato de ácido graxo é controlado pela tioesterase. Em algumas instâncias, um gene *tes* ou *fat* pode ser sobreexpressado. Em outras instâncias, os ácidos graxos C₁₀ podem ser produzidos atenuando-se a tioesterase C₁₈ (por exemplo, números de adesão AAC73596 e P0ADA1), que utiliza C_{18:1}-ACP, e expressando a tioesterase C₁₀ (por exemplo, número de adesão Q39513), que utiliza C₁₀-ACP. Isso resulta em uma população de ácidos graxos relativamente homogênea que possua uma cadeia de carbono de 10 de comprimento. Em ainda outras instâncias, ácidos graxos de C₁₄ podem ser produzidos atenuando-se as tioesterases endógenas que produzem ácidos graxos sem C₁₄ e expressando as tioesterases, que usam C₁₄-ACP (por exemplo, número de adesão Q39473). Em algumas situações, os ácidos graxos de C₁₂ podem ser produzidos expressando-se tioesterases que usam C₁₂-ACP (por exemplo, número de adesão Q41635) e atenuando-se tioesterases que produzam ácidos graxos sem C₁₂. A sobreprodução de acetil-CoA, malonil-CoA e ácido graxo pode ser verificada usando-se métodos conhecidos no estado da técnica, por exemplo, usando precursores radioativos, HPLC e GC-MS subsequente à lise [quebra/separação] celular. Exemplos não limitantes de tioesterases que podem ser usadas nos métodos descritos neste documento estão listados na Tabela 2.

Tabela 2: Tioesterases

Número de adesão	Organismo de fonte	Gene	Produto preferencial produzido
AAC73596	<i>E. coli</i>	<i>tesA</i> sem sequência líder	C _{18:1}
AAC73555	<i>E. coli</i>	<i>tesB</i>	

Q41635, AAA34215	<i>Umbellularia californica</i>	<i>fatB</i>	C _{12:0}
Q39513; AAC49269	<i>Cuphea hookeriana</i>	<i>fatB2</i>	C _{8:0} -C _{10:0}
AAC49269; AAC72881	<i>Cuphea hookeriana</i>	<i>fatB3</i>	C _{14:0} -C _{16:0}
Q39473; AAC49151	<i>Cinnamomum camphorum</i>	<i>fatB</i>	C _{14:0}
CAA85388	<i>Arabidopsis thaliana</i>	<i>fatB</i> [M141T]*	C _{16:1}
NP 189147; NP 193041	<i>Arabidopsis thaliana</i>	<i>fatA</i>	C _{18:1}
CAC39106	<i>Bradyrhizobium japonicum</i>	<i>fatA</i>	C _{18:1}
AAC72883	<i>Cuphea hookeriana</i>	<i>fatA</i>	C _{18:1}
AAL79361	<i>Helianthus annus</i>	<i>FatA1</i>	

*MAYER *et al.*, *BMC Plant Biology* 7:1-11, 2007

Formação de aldeídos, álcoois graxos, alcanos e alcenos ramificados

[217] Aldeídos, álcoois graxos, alcanos e alcenos podem ser produzidos de forma a conter pontos de ramificação usando-se derivados de ácido graxo como substratos. Por exemplo, embora o *E. coli* produza naturalmente cadeias lineares de derivados de ácido graxo, (sFAs), o *E. coli* pode ser manipulado para produzir derivados de ácido graxo de cadeia ramificada (brFAs) pelas introdução e expressão ou sobreexpressão de genes que forneçam precursores ramificados no *E. coli* (por exemplo, as famílias de genes *bkd*, *ilv*, *icm* e *fab*). Além disso, uma célula hospedeira pode ser manipulada para expressar ou sobreexpressar genes que codifiquem

proteínas para o alongamento de brFAs (por exemplo, ACP, FabF, etc.) e/ou para suprimir ou atenuar os genes correspondentes da célula hospedeira que normalmente conduzem aos sFAs.

[218] A primeira etapa na formação de brFAs é a produção de ácidos alfa-ceto correspondentes pela aminotransferase de um aminoácido de cadeia ramificada. As células hospedeiras podem, endogenamente, incluir genes codificando tais enzimas ou tais genes podem ser introduzidos recombinantemente. *E. coli*, por exemplo, endogenamente expressa tal enzima IlvE (EC 2,6,1,42; acesso GenBank YP_026247). Em algumas células hospedeiras, uma aminotransferase de aminoácido heterólogo de cadeia ramificada pode não ser expressada. Entretanto, *E. coli* IlvE ou qualquer outra aminotransferase de aminoácido de cadeia ramificada (por exemplo, IlvE a partir de *Lactococcus lactis* (acesso no GenBank AAF34406), IlvE a partir de *Pseudomonas putida* (acesso no GenBank NP_745648) ou IlvE a partir de *Streptomyces coelicolor* (acesso no GenBank NP_629657), se não endógena, pode ser introduzida e expressada recombinantemente.

[219] A segunda etapa é a descarboxilação oxidativa dos α -cetoácidos aos acil-CoA de cadeia ramificada correspondentes. Essa reação pode ser catalisada por uma desidrogenada complexa de um α -cetoácido de cadeia ramificada (*bkd*; EC 1,2,4,4) (DENOYA *et al. J. Bacteriol.* 177:3504, 1995), que consiste em subunidades E1 α / β (descarboxilase), E2 (diidrolipoil transacilase) e E3 (diidrolipoil desidrogenase). Essas desidrogenases complexas de α -cetoácidos de cadeia ramificada são semelhantes às desidrogenases complexas de α -cetoglutaratatos. Qualquer

microrganismo que possua brFAs e/ou cresça em aminoácidos de cadeia ramificada pode ser usado como uma fonte para isolar genes *bkd* para a expressão em células hospedeiras, por exemplo, a *E. coli*. Além disso, o *E. coli* possui o componente E3 como parte de sua desidrogenase complexa de piruvato (*lpd*, EC 1,8,1,4, adesão no GenBak NP_414658). Assim, ele pode ser suficiente para expressar apenas os genes *E1 α / β* e *E2 bkd*. A Tabela 3 lista exemplos não limitante de genes *bkd* a partir de vários microrganismos que podem ser recombinantemente introduzidos e expressados em uma célula hospedeira para fornecer precursores de acil-CoA de cadeia ramificada.

Tabela 3: Genes *bkd* a partir de microrganismos selecionados

Organismo	Gene	# de acesso no GenBank
<i>Streptomyces coelicolor</i>	<i>bkdA1</i> (E1 α)	NP_628006
	<i>bkdB1</i> (E1 β)	NP_628005
	<i>bkdC1</i> (E2)	NP_638004
<i>Streptomyces coelicolor</i>	<i>bkdA2</i> (E1 α)	NP_733618
	<i>bkdB2</i> (E1 β)	NP_628019
	<i>bkdC2</i> (E2)	NP_628018
<i>Streptomyces avermitilis</i>	<i>bkdA</i> (E1a)	BAC72074
	<i>bkdB</i> (E1b)	BAC72075
	<i>bkdC</i> (E2)	BAC72076
<i>Streptomyces avermitilis</i>	<i>bkdF</i> (E1 α)	BAC72088
	<i>bkdG</i> (E1 β)	BAC72089
	<i>bkdH</i> (E2)	BAC72090
<i>Bacillus subtilis</i>	<i>bkdAA</i> (E1 α)	NP_390288
	<i>bkdAB</i> (E1 β)	NP_390288
	<i>bkdB</i> (E2)	NP_390288
<i>Pseudomonas putida</i>	<i>bkdA1</i> (E1 α)	AAA65614
	<i>bkdA2</i> (E1 β)	AAA65615
	<i>bkdC</i> (E2)	AAA65617

[220] Em outro exemplo, isobutiril-CoA pode ser feito em uma célula hospedeira, por exemplo, em *E. coli*, através da coexpressão de uma redutase de crotonil-CoA (*Ccr*, EC 1,6,5,5, 1,1,1,1) e de mutase de isobutiril-CoA (grande subunidade *IcmA*, EC 5,4,99,2; pequena subunidade *IcmB*, EC 5,4,99,2) (HAN e REYNOLDS, *J. Bacteriol.* 179:5157, 1997). Crotonil-CoA

é um intermediário na biossíntese de ácido graxo em *E. coli* e outros microrganismos. Exemplos não limitantes de genes *ccr* e *ICM* a partir de microrganismos selecionados estão listados na Tabela 4.

Tabela 4: Genes *ccr* e *ICM* a partir de microrganismos selecionados

Organismo	Gene	# de acesso no GenBank
<i>Sreptomyces coelicolor</i>	<i>Ccr</i>	NP_630556
	<i>icmA</i>	NP_629554
	<i>icmB</i>	NP_630904
<i>Streptomyces cinnamomencis</i>	<i>ccr</i>	AAD53915
	<i>icmA</i>	AAC08713
	<i>icmB</i>	AJ246005

[221] Além da expressão de genes *bkd*, a iniciação da biossíntese de brFA utiliza β -cetoacil proteína de transporte sintase III (FabH, EC 2,3,1,41) com especificidade para acil-CoAs de cadeia ramificada (LI et al., *J. Bacteriol.* 187:3795-3799, 2005). Exemplos não limitantes de tais enzimas FabH estão listados na tabela 5. Genes *fabH* que são envolvidos na biossíntese de ácido graxo de qualquer microrganismo contendo brFA podem ser expressado na célula hospedeira. As enzimas BKD e FabH das células hospedeiras que naturalmente não produzem brFA podem não sustentar a produção de brFA. Além disso, *bkd* e *fabH* podem ser expressados recombinantemente. Vetores contendo os genes *bkd* e *fabH* podem ser inseridos dentro de uma célula hospedeira. Igualmente, o nível endógeno da produção de Bkd e FabH pode não ser suficiente para produzir brFA. Nesse caso, eles podem ser sobreexpressados. Além disso, outros componentes da via de biossíntese de ácido graxo pode ser expressados ou sobreexpressados, tal como proteínas transportadoras de acil (ACPs) e β -cetoacil proteína de

transporte sintase II (*fabF*, EC 2,3,1,41) (exemplos não limitantes de candidatos estão listados na Tabela 5). Além de expressar esses genes, alguns genes na via biossintética de ácido graxo endógeno podem ser atenuados na célula hospedeira (por exemplo, os genes *E. coli fabH* (número de acesso no GenBank NP_415609) e/ou *fabF* (número de acesso no GenBank NP_415613)).

Tabela 5: Genes *fabH*, *ACP* e *fabF* a partir de microrganismos selecionados com bFFAs

Organismo	Gene	Número de acesso no GenBank
<i>Streptomyces coelicolor</i>	<i>fabH1</i>	NP_626634
	<i>ACP</i>	NP_626635
	<i>fabF</i>	NP_626636
<i>Streptomyces avermitilis</i>	<i>fabH3</i>	NP_823466
	<i>fabC3 (ACP)</i>	NP_823467
	<i>fabF</i>	NP_823468
<i>Bacillus subtilis</i>	<i>fabH_A</i>	NP_389015
	<i>fabH_B</i>	NP_388898
	<i>ACP</i>	NP_389474
	<i>fabF</i>	NP_389016
<i>Stenotrophomonas maltophilia</i>	SmalDRAFT_0818 (<i>FabH</i>)	ZP_01643059
	SmalDRAFT_0821 (<i>ACP</i>)	ZP_01643063
	SmalDRAFT_0822 (<i>FabF</i>)	ZP_01643064
<i>Legionella pneumophila</i>	<i>FabH</i>	YP_123672
	<i>ACP</i>	YP_123675
	<i>fabF</i>	YP_123676

Formação de aldeídos cíclicos, álcoois graxos, alcanos e alcenos

[222] Aldeídos cíclicos, álcoois graxos, alcanos e alcenos podem ser produzidos usando-se derivados de ácido graxo cíclico como substratos. Para produzir substratos de derivados de ácido graxo cíclico, os genes que fornecem precursores cíclicos (por exemplo, as famílias de gene *ans*, *chc* e *plm*) podem ser introduzidos dentro da célula hospedeira

e expressados para permitir a iniciação da biossíntese de ácido graxo a partir dos precursores cíclicos. Por exemplo, para converter uma célula hospedeira, tal como *E. coli*, em uma capaz de sintetizar derivados de ácido graxo ω -cíclico (cyFA), um gene que forneça o precursor cíclico ciclohexilcarbonil-CoA (CHC-CoA) (CROPP *et al.*, *Nature Biotech.* 18:980-983, 2000) pode ser introduzido e expressado na célula hospedeira. Exemplos não limitantes de genes que forneçam CHC-CoA em *E. coli* incluem: *ansJ*, *ansK*, *ansL*, *chcA* e *ansM* a partir do conjunto de gene ansatrienina de *Streptomyces collinus* (CHEN *et al.* *Eur. J. Biochem.* 261:98-107, 1999) ou *plmJ*, *plmK*, *plmL*, *chcA* e *plmM* a partir do conjunto de gene foslactomicina B de *Streptomyces* SP. HK803 (PALANIAPPAN *et al.* *J. Biol. Chem.*, 278:35552-35557, 2003) juntos com o gene *chcB* (PATTON *et al.*, *Biochem.* 39:7595-7604, 2000) a partir de *S. collinus*, *S. avermitilis* ou *S. coelicolor* (ver Tabela 6). Os genes listados na Tabela 5 podem, então, ser expressados para permitir a iniciação e o alongamento de ácidos graxos ω -cíclicos. Alternativamente, os genes homólogos podem ser isolados a partir de microrganismos que produzem cyFA e são expressados em uma célula hospedeira (por exemplo, *E. coli*).

Tabela 6: Genes para a síntese de CHC-CoA

Organismo	Gene	Número de acesso no GenBank
<i>Streptomyces collinus</i>	<i>ansJK</i> <i>ansL</i> <i>chcA</i> <i>ansM</i> <i>chcB</i>	U72144* AF268489
<i>Streptomyces</i> SP. HK803	<i>pmlJK</i> <i>pmlL</i> <i>chcA</i> <i>pmlM</i>	AAQ84158 AAQ84159 AAQ84160 AAQ84161
<i>Streptomyces</i>	<i>chcB/caiD</i>	NP_629292

<i>coelicolor</i>		
<i>Streptomyces avermitilis</i>	<i>chcB/caiD</i>	NP_629292

*Apenas *chcA* está registrado no GenBank na entrada U72144, *ansJKLM* estão de acordo com Chen *et al.* (*Eur. J. Biochem.* 261:98-107, 1999).

[223] Os genes listados na Tabela 5 (*fabH*, *ACP* e *fabF*) permitem a iniciação e o alongamento de derivados de ácido graxo ω -cíclico porque possuem ampla especificidade de substrato. Se a coexpressão de qualquer um desses genes com os genes listados na Tabela 6 não produzir *cyFA*, então os homólogos *fabH*, *ACP* e/ou *fabF* a partir de microrganismos que façam *cyFAs* (por exemplo, aqueles listados na Tabela 7) podem ser isolados (por exemplo, usando-se iniciadores de PCR degenerados ou provas de sequência de DNA heterólogas) e coexpressados.

Tabela 7: Exemplos não limitantes de microrganismos que contenham ácidos graxos ω -cíclicos

Organismo	Referência
<i>Curtobacterium pusillum</i>	ATCC19096
<i>Alicyclobacillus acidoterrestris</i>	ATCC49025
<i>Alicyclobacillus acidocaldarius</i>	ATCC27009
<i>Alicyclobacillus cycloheptanicus</i> *	MOORE, J. <i>Org. Chem.</i> , 62:pp.2173, 1997

* Utiliza ciclo-heptilcarbonil-CoA e não ciclo-hexilcarbonil-CoA como precursor para a biossíntese de *cyFA*.
Níveis de saturação de aldeído, álcool graxo e alceno

[224] O grau de saturação em derivados de ácido graxo pode ser controlado pela regulação do grau de saturação de derivados de ácido graxo intermediários. As famílias de gene *sfa*, *gns* e *fab* podem ser expressadas ou sobreexpressadas para controlar a saturação dos ácidos graxos. A Tabela A lista

exemplos não limitantes de genes nessas famílias de genes que podem ser usadas nos métodos de nas células hospedeiras descritos neste documento.

[225] Células hospedeiras podem ser projetadas para produzir ácidos graxos insaturados manipulando-se a célula hospedeira para sobreexpressar *fabB* ou pelo crescimento da célula hospedeira a baixas temperaturas (por exemplo, menos que 37°C). FabB possui predileção por cis- δ^3 decenoil-ACP e resulta na produção de ácido graxo insaturado em *E. coli*. A sobreexpressão de *fabB* resulta na produção de uma porcentagem significativa de ácidos graxos insaturados (DE MENDOZA *et al.*, *J. Biol. Chem.* 258:2098-2101, 1983. O gene *fabB* deve ser inserido e expressado em células hospedeiras que naturalmente não contenham o gene. Esse derivados insaturados de ácido graxo podem, então, se usados como intermediários em células hospedeiras que sejam projetadas para produzir derivados de ácido graxo, tais como aldeídos graxos, álcoois graxos ou alcenos.

[226] Em outras instâncias, um repressor de biossíntese de ácido graxo, por exemplo, *fabR* (acesso no GenBank NP_418398), pode ser suprimido, o que também irá resultar no aumento da produção de ácido graxo insaturado em *E. coli* (ZHANG *et al.*, *J. Biol. Chem.* 277:15558, 2002). Deleções semelhantes podem ser feitas em outras células hospedeiras. Outro aumento nos derivados insaturados de ácido graxo pode ser alcançado, por exemplo, pela sobreexpressão de *fabM* (trans-2, cis-3-decenoil-ACP isomerase, acesso no GenBank DAA05501) e pela expressão controlada de *fabK* (trans-2-enoil-ACP redutase II, acesso no GenBank NP_357969) a partir de *Streptococcus pneumoniae* (MARRAKCHI *et al.*, *J. Biol. Chem.*

277:44809, 202), enquanto se suprime *fabI* de *E. coli* (trans-2-enoil-ACP redutase, acesso no GenBank NP_415804). Em alguns exemplos, o gene endógeno *fabF* pode ser atenuado, assim aumentando a porcentagem de palmitoleato (C16:1) produzido.

Outros substratos

[227] Outros substratos que podem ser usados para produzir aldeído, álcoois graxos, alcanos e alcenos nos métodos descritos neste documento são acil-ACP, acil-CoA, um aldeído graxo ou um álcool graxo, o que está descrito, por exemplo, em PCT/US08/058788. Exemplos de genes que podem ser alterados para expressar ou sobreexpressar esses substratos em células hospedeiras estão listados na Tabela A. Outros exemplares de genes são descritos em PCT/US08/058788.

Manipulação genética de células hospedeiras para produzir aldeídos, álcoois graxos, alcano ou alcenos

[228] Diferentes células hospedeiras podem ser usadas para produzir aldeídos, álcoois graxos, alcanos e/ou alcenos, como descrito neste documento. Uma célula hospedeira pode ser uma célula procariótica ou eucariótica. Por exemplo, um polipeptídeo descrito neste documento pode ser expressado em células bacterianas (tais como *E. coli*), células de insetos, células de levedura ou mamíferas (tais como células de ovário de hamster chinês (CHO), células COS, células VERO, células BHK, células HeLa, células Cv1, células MDCK, células 293, células 3T3 ou células PC12). Outros exemplares de células hospedeiras incluem células a partir de membros dos gêneros *Escherichia*, *Bacillus*, *Lactobacillus*, *Rhodococcus*, *Pseudomonas*, *Aspergillus*, *Trichoderma*, *Neurospora*, *Fusarium*, *Humicola*, *Rhizomucor*, *Kluyveromyces*,

Pichia, *Mucor*, *Myceliophthora*, *Penicillium*, *Phanerochaete*, *Pleurotus*, *Trametes*, *Chrysosporium*, *Saccharomyces*, *Schizosaccharomyces*, *Yarrowia* ou *Streptomyces*. Ainda outros exemplares de células hospedeiras podem ser uma célula *Bacillus lentus*, uma célula *Bacillus brevis*, uma célula *Bacillus stearothermophilus*, uma célula *Bacillus licheniformis*, uma célula *Bacillus alkalophilus*, uma célula *Bacillus coagulans*, uma célula *Bacillus circulans*, uma célula *Bacillus pumilis*, uma célula *Bacillus thuringiensis*, uma célula *Bacillus clausii*, uma célula *Bacillus megaterium*, uma célula *Bacillus subtilis*, uma célula *Bacillus amyloliquefaciens*, uma célula *Trichoderma koningii*, uma célula *Trichoderma viride*, uma célula *Trichoderma reesei*, uma célula *Trichoderma longibrachiatum*, uma célula *Aspergillus awamori*, uma célula *Aspergillus fumigates*, uma célula *Aspergillus foetidus*, uma célula *Aspergillus nidulans*, uma célula *Aspergillus niger*, uma célula *Aspergillus oryzae*, uma célula *Humicola insolens*, uma célula *Humicola lanuginose*, uma célula *Rhizomucor miehei*, uma célula *Mucor michei*, uma célula *Streptomyces lividans*, uma célula *Streptomyces murinus*, ou uma célula *Actinomycetes*.

[229] Outros exemplos não limitantes de células hospedeiras são aqueles listados na Tabela 1.

[230] E uma concretização preferida, a célula hospedeira é uma célula de *E. coli*. Em uma concretização mais preferida, a célula hospedeira é *E. coli* de linhagens B, C, K ou W. Outras células hospedeiras adequadas são conhecidas daqueles hábeis no estado da técnica.

[231] Diferentes métodos bem conhecidos no estado da técnica podem ser usados para manipular geneticamente

células hospedeiras para produzir aldeídos, álcoois graxos, alcanos e/ou alcenos. Os métodos incluem o uso de vetores, preferencialmente vetores de expressão, contendo um ácido nucleico codificando um polipeptídeo biossintético de aldeído, álcool graxo, alcano e/ou alceno descrito neste documento, ou uma variante ou um fragmento de polipeptídeo deste. Como usado neste documento, o termo "vetor" refere-se a uma molécula de ácido nucleico capaz de transportar outro ácido nucleico ao qual ele tenha sido ligado. Um tipo de vetor é um "plasmídeo", que se refere a uma volta circular de DNA de dois filamentos dentro do qual segmentos adicionais de DNA possam ser ligados. Outro tipo de vetor é um vetor viral, no qual segmentos adicionais de DNA podem ser ligados dentro do genoma viral. Certos vetores são capazes de replicação autônoma em uma célula hospedeira na qual eles sejam introduzidos (por exemplo, vetores bacterianos contendo uma origem bacteriana de replicação e vetores mamíferos epissomáticos). Outros vetores (por exemplo, vetores mamíferos não epissomáticos) são integrados ao genoma de uma célula hospedeira até a introdução dentro da célula hospedeira e são, então, replicados juntos com o genoma hospedeiro. Além do mais, certos vetores, tais como vetores de expressão, são capazes de direcionar a expressão de genes aos quais eles sejam operacionalmente ligados. Em geral, vetores de expressão usados em técnicas de DNA recombinante estão frequentemente na forma de plasmídeos. Entretanto, outras formas de vetores de expressão, tais como vetores virais (por exemplo, retrovírus de replicação defeituosa, adenovírus e vírus adenoassociados), também podem ser usadas.

[232] Os vetores de expressão recombinante descritos neste documento incluem um ácido nucleico descrito neste documento em uma forma adequada para a expressão do ácido nucleico em uma célula hospedeira. Os vetores de expressão recombinante podem incluir uma ou mais sequências de controle, selecionadas na base da célula hospedeira a ser usada para a expressão. A sequência de controle é operacionalmente ligada à sequência de ácido nucleico para ser expressada. Tais sequências de controle são descritas, por exemplo, em Goeddel, *Gene Expression Technology: Methods in Enzymology 185*, Academic Press, San Diego, California (1990). Sequências de controle incluem aquelas que direcionam a expressão constitutiva de uma sequência de nucleotídeo apenas em determinadas células hospedeiras (por exemplo, sequências regulatórias específicas de tecido). Será apreciado por aqueles hábeis no estado da técnica que o projeto do vetor de expressão possa 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 descritos neste documento podem ser introduzidos dentro de células hospedeiras para produzir polipeptídeos, incluindo a fusão de polipeptídeos, codificados pelos ácidos nucleicos, como descrito neste documento.

[233] Vetores de expressão recombinante podem ser projetados para a expressão de um polipeptídeo biossintético de aldeído, álcool graxo, alceno e/ou alceno ou variantes em células procarióticas ou eucarióticas (por exemplo, células bacterianas, tais como *E. coli*, células de insetos (usando-se vetores de expressão de baculovírus), células de levedura

ou células de mamíferos). Células hospedeiras adequadas são ainda discutidas em Goeddel, *Gene Expression Technology: Methods in Enzymology 185*, Academic Press, San Diego, California (1990). Alternativamente, o vetor de expressão recombinante pode ser transcrito ou traduzido *in vitro*, por exemplo, usando-se sequências regulatórias de promotor T7 e polimerase T7.

[234] A expressão de polipeptídeos em procariontes, por exemplo, *E. coli*, é mais frequentemente realizada com vetores contendo promotores constitutivos ou induzíveis direcionando a expressão tanto de polipeptídeos de fusão quanto de não fusão. Vetores de fusão adicionam um número de aminoácidos a um polipeptídeo codificado neles, normalmente a um aminoterminal do polipeptídeo recombinante. Tais vetores de fusão tipicamente servem a três propósitos: (1) aumentar a expressão do polipeptídeo recombinante; (2) aumentar a solubilidade do polipeptídeo recombinante; e (3) auxiliar na purificação do polipeptídeo recombinante agindo como um ligante em purificação de afinidade. Frequentemente, em vetores de expressão de fusão, um local de clivagem proteolítica é introduzido na junção da metade de fusão e o polipeptídeo recombinante. Isso permite a separação do polipeptídeo recombinante da metade de fusão após a purificação do polipeptídeo de fusão. Exemplos de tais enzimas, e suas sequências cognatas de reconhecimento, incluem o Fator Xa, trombina e enteroquinase. Exemplares de vetores de expressão de fusão incluem pGEX (*Pharmacia Biotech Inc.*; SMITH *et al. Gene* (1988) 67:31-40), pMAL (*New England Biolabs*, Beverly, Mass.) e pRITS (*Pharmacia*, Piscataway, NJ), que fundem glutational S-transferase (GST), maltose E

ligando proteína ou proteína A, respectivamente, ao polipeptídeo recombinante alvo.

[235] Exemplos de vetores de expressão *E. coli* induzíveis de não fusão incluem pTrc (AMANN *et al.*, *Gene* (1988) 69:301-315) e pET 11d (STUDIER *et al.*, *Gene Expression Technology: Methods in Enzymology 185*, Academic Press, San Diego, California (1990) 60-89). A expressão do gene alvo a partir do vetor pTrc baseia-se na transcrição hospedeira de polimerase de RNA a partir de um promotor de fusão trp-lac híbrido. A expressão do gene alvo a partir do vetor pET 11d baseia-se na transcrição a partir do promotor de fusão T7 gn10-lac mediado por uma polimerase de RNA viral (T7 gn1) coexpressada. Essa polimerase viral é fornecida pelas linhagens hospedeiras BL21(DE3) ou HMS174(DE) a partir de um prófago λ residente nutrindo um gene T7 gn1 sob o controle transcricional do promotor lacUV 5.

[236] Uma estratégia para maximizar a expressão de polipeptídeo recombinante é expressar o polipeptídeo em uma célula hospedeira com uma capacidade debilitada para clivar proteoliticamente o polipeptídeo recombinante (ver GOTTESMAN, *Gene Expression Technology: Methods in Enzymology 185*, Academic Press, San Diego, California (1990) 119-128). Outra estratégia é alterar a sequência de ácido nucléico a ser inserida dentro de um vetor de expressão de forma que os códons individuais para cada aminoácido sejam aqueles preferencialmente utilizados na célula hospedeira (WADA *et al.* *Nucleic Acids Res.* (1992) 20:2111-2118). Tal alteração de sequências de ácido nucléico pode ser realizada por técnicas padrão de síntese de DNA.

[237] Em outra concretização, a célula hospedeira é uma

célula de levedura. Nesta concretização, o vetor de expressão é um vetor de expressão de levedura. Exemplos de vetores para a expressão em levedura *S. cerevisiae* incluem pYepSec1 (BALDARI *et al.* *EMBO J.* (1987) 6:229-234), pMFa (KURJAN *et al.* *Cell* (1982) 30:933-943), pJRY88 (SCHULTZ *et al.* *Gene* (1987) 54:113-123), pYES2 (Invitrogen Corporation, San Diego, California), e picZ (Invitrogen Corp. San Diego, California).

[238] Alternativamente, um polipeptídeo descrito neste documento pode ser expressado em células de insetos usando-se vetores de expressão de baculovírus. Vetores de baculovírus disponíveis para expressão de proteínas em cultura de células de insetos (por exemplo, células Sf9) incluem, por exemplo, as séries PAC (SMITH *et al.* *Mol. Cell Biol.* (1983) 3:2156-2165) e as séries pVL (LUCKLOW *et al.* *Virology* (1989) 170:31-39).

[239] Ainda em outra concretização, os ácidos nucleicos descritos neste documento podem ser expressados em células de mamíferos usando-se um vetor de expressão de mamífero. Exemplos de vetores de expressão de mamíferos incluem pCDM8 (SEED, *Nature* (1987) 329:840) e pMT2PC (KAUFMAN *et al.*, *EMBO J.* (1987) 6:187-195). Quando usado em células de mamíferos, as funções de controle do vetor de expressão podem ser fornecidas por elementos regulatórios virais. Por exemplo, promotores comumente usados são derivados a partir de polioma, Adenovírus 2, citomegalovírus e Simian Vírus 40. Outros sistemas de expressão adequados tanto para células procariotas quanto eucariotas estão descritos nos capítulos 16 e 17 de SAMBROOK *et al.*, eds. *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory,

Cold Spring Harbor Laboratory Press, Cold Spring Harbos, NY, 1989.

[240] Vetores podem ser introduzidos em células procariotas ou eucariotas através de técnicas convencionais de transformação ou transfecção. Como usado neste documento, os termos "transformação" e "transfecção" referem-se a uma variedade de técnicas reconhecidas no estado da técnica para a introdução de ácido nucleico estrangeiro (por exemplo, DNA) dentro de uma célula hospedeira, incluindo a coprecipitação de fosfato de cálcio ou de cloreto de cálcio, a transfecção de DEAE-dextran mediado, lipofecção ou eletroporação. Métodos adequados para a transformação ou a transfecção de células hospedeiras podem ser encontrados em, por exemplo, Sambrook *et al.* (*supra*).

[241] Para a transformação estável de células bacterianas, sabe-se que, dependendo do vetor de expressão e da técnica de transformação usados, apenas uma pequena fração de células será absorvida e replicará o vetor de expressão. A fim de identificar e selecionar esses transformantes, um gene que codifique um marcador selecionável (por exemplo, resistência a antibióticos) pode ser introduzido dentro da célula hospedeira junto com o gene de interesse. Marcadores selecionáveis incluem aqueles que conferem resistência a drogas, tais como ampicilina, canamicina, cloranfenicol ou tetraciclina. Ácidos nucleicos codificando um marcador selecionável podem ser introduzidos em uma célula hospedeira no mesmo vetor que aquele codificando um polipeptídeo descrito neste documento ou pode ser introduzido em um vetor separado. Células estavelmente transfectadas com o ácido nucleico introduzido podem ser

identificadas através de seleção por droga (por exemplo, células que incorporaram o gene do marcador selecionável irão sobreviver, enquanto as outras células morrerão).

[242] Para a transfecção estável de células de mamíferos, sabe-se que, dependendo do vetor de expressão e da técnica de transfecção usados, apenas uma pequena fração de células deve integrar o DNA estrangeiro dentro de seu genoma. A fim de identificar e selecionar esses integrantes, um gene que codifique um marcador selecionável (por exemplo, resistente a antibióticos) pode ser introduzido dentro das células hospedeiras junto com o gene de interesse. Marcadores selecionáveis preferidos incluem aqueles que conferem resistência a drogas, tais como G418, higromicina e metotrexato. Ácidos nucleicos codificando um marcador selecionável podem ser introduzidos em uma célula hospedeira no mesmo vetor que aquele codificando um polipeptídeo descrito neste documento ou podem ser introduzidos em um vetor separado. Células estávelmente transfectadas com o ácido nucleico introduzido podem ser identificadas através de seleção por droga (por exemplo, células que incorporaram o gene de marcador selecionável irão sobreviver, enquanto as outras células morrerão).

[243] Em certos métodos, um polipeptídeo biossintético de aldeído e um polipeptídeo biossintético de alceno ou alceno são coexpressados em uma única célula hospedeira. Em métodos alternativos, um polipeptídeo biossintético de aldeído e um polipeptídeo desidrogenase de álcool são coexpressados em uma única célula hospedeira.

Proteínas de transporte

[244] Proteínas de transporte podem exportar

polipeptídeos e hidrocarbonetos (por exemplo, aldeídos, alcanos e/ou alcenos) para fora da uma célula. Muitas proteínas de transporte e de efluxo servem para excretar uma ampla variedade de compostos e podem ser naturalmente modificadas para serem seletivas para tipos particulares de hidrocarbonetos.

[245] Exemplos não limitantes de proteínas de transporte adequadas são proteínas de transporte cassete de ligação a ATP (ABC), proteínas de efluxo e proteínas transportadoras de ácido graxo (FATP). Exemplos não limitantes adicionais de proteínas de transporte adequadas incluem as proteínas de transporte ABC a partir de organismos tais como *Caenorhabditis elegans*, *Arabidopsis thaliana*, *Alkaligenes eutrophus* e *Rhodococcus erythropolis*. Exemplos de proteína de transporte ABC que podem ser usadas estão listados na Tabela A (por exemplo, CER5, AtMRP5, Amis2 e AtPGP1). Células hospedeiras também podem ser escolhidas por sua habilidade endógena de secretar hidrocarbonetos. A eficiência de produção de hidrocarboneto e secreção dentro do ambiente da célula hospedeira (por exemplo, meio de cultura, caldo de fermentação) pode ser expressada como uma razão de produto intracelular para produto extracelular. Em alguns exemplos, a razão pode ser cerca de 5:1, 4:1, 3:1, 2:1, 1:1, 1:2, 1:3, 1:4 ou 1:5.

Fermentação

[246] A produção e o isolamento de aldeído, alcoóis graxos, alcanos e/ou alcenos podem ser melhorados empregando-se técnicas benéficas de fermentação. Um método para maximizar a produção enquanto se reduzem os gastos é aumentar a porcentagem da fonte de carbono que é convertida

em produtos de hidrocarboneto.

[247] Durante o ciclo de vida celular normal, o carbono é usado nas funções celulares, tais como produção de lipídios, sacarídeos, proteínas, ácidos orgânicos e ácidos nucleicos. A redução da quantidade de carbono necessária para as atividades relacionadas ao crescimento pode aumentar a eficiência da conversão de fonte de carbono em produto. Isso pode ser alcançado, por exemplo, primeiramente crescendo-se as células hospedeiras até uma densidade desejada (por exemplo, uma densidade alcançada no pico de fase log de crescimento). Em tal ponto, os genes *checkpoint* da replicação podem ser aproveitados para parar o crescimento das células. Especificamente, mecanismos *quorum sensing* (revisto em CAMILLI *et al. Science* 311:1113, 2006; VENTURI, *FEMS Microbio. Rev.* 30:274-291, 2006; e READING *et al. FEMS Microbiol. Lett.* 254:1-11, 2006) podem ser usados para ativar genes *checkpoint*, tais como *p53*, *p21* ou outros genes *checkpoint*.

[248] Genes que podem ser ativados para interromper a replicação e o crescimento da célula em *E. coli* incluem genes *umuDC*. A sobreexpressão de genes *umuDC* interrompe a progressão a partir da fase estacionária ao crescimento exponencial (MURLI *et al. J. of Bact.* 182:1127, 2000) O UmuC é uma polimerase de DNA que pode realizar síntese de translesão sobre lesões não codificantes - a mecânica básica da maioria das mutagêneses UV e químicas. Os produtos do gene *umuDC* estão envolvidos no processo de síntese de translesão e também servem como um ponto de checagem de perda da sequência de DNA. Os produtos do gene *umuDC* incluem UmuC, UmuD, *umuD'*, *UmuD'*₂C, *UmuD'*₂ e UmuD₂. Simultaneamente, genes de produção

de produtos podem ser ativados, assim minimizando a necessidade por vias de replicação e de manutenção para serem usadas enquanto um aldeído, um alcano e/ou um alceno estiver sendo feito. Células hospedeiras também podem ser projetadas para expressas *umuC* e *umuD* a partir de *E. coli* em pBAD24 sob o sistema promotor *prpBCDE* através da síntese *de novo* desse gene com genes apropriados de produção de produto final.

[249] A porcentagem de entrada de carbonos convertidos em aldeídos, alcoóis graxos, alcanos e/ou alcenos pode ser um direcionador de custos. Quanto mais eficiente for o processo (*i.e.*, quanto mais elevada a porcentagem de entrada de carbonos convertidos em aldeídos, alcoóis graxos, alcanos e/ou alcenos), menos caro o processo será. Para fontes de carbono contendo oxigênio (por exemplo, glicose e outras fontes baseadas em carboidratos), o oxigênio deve ser liberado na forma de dióxido de carbono. Para cada 2 átomos de oxigênio liberados, um átomo de carbono também é liberado levando a uma eficiência metabólica máxima teórica de aproximadamente 34% (p/p) (para produtos de derivados de ácido graxo). Essa figura, entretanto, muda para outros produtos de hidrocarbonetos e fontes de carbono. Eficiências típicas na literatura são de aproximadamente menos de 5%. Células hospedeiras projetadas para produzir aldeídos, alcanos e/ou alcenos podem ter eficiência superior a cerca de 1, 3, 5, 10, 15, 20, 25 e 30%. E um exemplo, células hospedeiras podem exibir uma eficiência de cerca de 10% a cerca de 25%. Em outros exemplos, tais células hospedeiras podem exibir uma eficiência de cerca de 25% a cerca de 30%. Em outros exemplos, células hospedeiras podem exibir eficiência superior a 30%.

[250] A célula hospedeira pode ser adicionalmente projetada para expressar celulosomas recombinantes, como aqueles descritos na aplicação PCT número PCT/US2007/003736. Esses celulosomas podem permitir que a célula hospedeira utilize material celulósico como fonte de carbono. Por exemplo, a célula hospedeira pode ser adicionalmente manipulada para expressar invertases (EC 3.2.1.26) de forma que a sacarose possa ser usada com uma fonte de carbono. Semelhantemente, a célula hospedeira pode ser projetada usando-se os ensinamentos descritos nas patentes norte-americanas números 5.000.000, 5.028.539, 5.424.202, 5.482.846 e 5.602.030, de forma que a célula hospedeira possa assimilar o carbono eficientemente e usar materiais celulósicos como fontes de carbono.

[251] Em um exemplo, a câmara de fermentação pode encerrar uma fermentação que esteja em redução contínua. Nessa instância, um ambiente redutivo estável pode ser criado. O equilíbrio de elétrons pode ser mantido pela liberação de dióxido de carbono (na forma gasosa). Esforços para ampliar o equilíbrio de NAD/H e NADP/H podem também facilitar a estabilização do equilíbrio de elétrons. A disponibilidade de NADPH intracelular pode também ser melhorada manipulando-se a célula hospedeira para expressar uma transidrogenase NADH:NADPH. A expressão de uma ou mais transidrogenases NADH:NADPH converte o NADH produzido em glicólise ao NADPH, o que pode melhorar a produção de aldeídos, alcanos e/ou alcenos.

[252] Para uma produção em pequena escala, as células hospedeiras manipuladas podem crescer em lotes de, por exemplo, cerca de 100mL, 500mL, 1L, 2L, 5L ou 10L;

fermentadas e induzidas a expressar os desejados aldeídos, alcoóis graxos, alcanos e/ou alcenos baseados nos genes específicos codificados nos plasmídeos apropriados. Por exemplo, células de *E. coli* BL21(DE3) alimentando pBAD24 (com resistência à ampicilina e via de síntese de aldeído, álcool graxo, alcano ou alceno) assim como pUMVC1 (com resistência à canamicina e o sistema de sobreexpressão de acetil CoA/malonil CoA) podem ser incubadas durante a noite em frascos de 2L a 37°C, agitados a > 200rpm em 500mL meio LB, suplementadas com 75µg/mL de ampicilina e 50µg/mL de canamicina até que as culturas atinjam um OD₆₀₀ de > 0,8. Ao alcançar um OD₆₀₀ de > 0,8, as células podem ser suplementadas com 25mM de proprionato de sódio (pH 8,0) para ativar os sistemas de genes manipulados para a produção e para interromper a proliferação celular por meio de ativação das proteínas UmuC e UmuD. A indução pode ser realizada por 6h a 30°C. Após a incubação, o meio pode ser examinado para aldeídos, alcoóis graxos, alcanos e/ou alcenos usando-se GC-MS.

[253] Para a produção em larga escala, as células hospedeiras manipuladas podem crescer em lotes de 10L, 100L, 1000L ou maiores; fermentadas e induzidas a expressar os desejados aldeídos, alcoóis graxos, alcanos e/ou alcenos baseados nos genes específicos codificados nos plasmídeos apropriados. Por exemplo, células de *E. coli* BL21(DE3) alimentando pBAD24 (com resistência à ampicilina e via de síntese de aldeído, álcool graxo, alcano ou alceno) assim como pUMVC1 (com resistência à canamicina e o sistema de sobreexpressão de acetil CoA/malonil CoA) podem ser incubadas a partir de uma cultura semente para fermentações de 10L (5L

para fermentações de 100L, etc.) e meio LB (livre de glicerol) com 50µg/mL de canamicina e 75µg/mL de ampicilina a 37°C, e agitados a > 200rpm até que as culturas atinjam um OD₆₀₀ de > 0,8 (tipicamente 16h). O meio pode ser continuamente suplementado para manter 25mM de proprionato de sódio (pH 8,0) para ativar os sistemas de genes manipulados para a produção e para interromper a proliferação celular por meio de ativação das proteínas UmuC e UmuD. O meio pode ser continuamente suplementado com glicose para manter a concentração de 25g/100mL.

[254] Após a primeira hora de indução, partes de não mais que 10% do volume total da célula podem ser removidas a cada hora e a elas é permitido sentar sem agitação para permitir aos aldeídos, alcanos e/ou alcenos ascender à superfície e iniciar uma fase de separação espontânea. O componente aldeído, álcoois graxos, alceno e/ou alceno pode ser, então, coletado, e a fase aquosa retornada à câmara de reação. A câmara de reação pode ser operada continuamente. Quando o OD₆₀₀ atinge abaixo de 0,6, as células podem ser substituídas com um novo lote crescido a partir de uma cultura semente.

Produção de aldeídos, álcoois graxos, alcanos e alcenos usando métodos livres de células

[255] Em alguns métodos descritos neste documento, um aldeído, álcoois graxos, alceno e/ou alceno podem ser produzidos usando-se um polipeptídeo purificado descrito neste documento e um substrato descrito neste documento. Por exemplo, uma célula hospedeira pode ser manipulada para expressar polipeptídeo biossintético de aldeído, álcoois graxos, alceno e/ou alceno ou variantes, como descrito neste documento. Extratos livres de células podem, então, ser

gerados usando-se métodos conhecidos. Por exemplo, as células hospedeiras podem ser lisadas usando-se detergentes ou por sonicação. Os polipeptídeos expressados podem ser purificados usando-se métodos conhecidos. Após a obtenção de extratos livres de células, os substratos descritos neste documento podem ser adicionados aos extratos livres de células e mantidos sob condições para permitir a conversão dos substratos em aldeídos, álcoois graxos, alcanos e/ou alcenos. Os aldeídos, álcoois graxos, alcanos e/ou alcenos podem, então, ser separados e purificados usando-se técnicas conhecidas.

Processamento de pós-produção

[256] Os aldeídos, álcoois graxos, alcanos e/ou alcenos produzidos durante a fermentação podem ser separados a partir do meio de fermentação. Qualquer técnica conhecida para a separação de aldeídos, álcoois graxos, alcanos e/ou alcenos a partir do meio aquoso pode ser usada. Um exemplar de processo de separação é um processo de separação de duas fases (bifásico). Esse processo envolve a fermentação das células hospedeiras geneticamente manipuladas sob condições suficientes para produzir um aldeído, álcoois graxos, alceno e/ou alceno, permitindo ao aldeído, álcoois graxos, alceno e/ou alceno serem coletados e uma fase orgânica, e separando a fase orgânica do caldo aquoso de fermentação. Esse método pode ser praticado tanto em um lote quanto em um conjunto de fermentação contínua.

[257] A separação bifásica utiliza a imiscibilidade relativa de aldeídos, álcoois graxos, alcanos e/ou alcenos para facilitar a separação. Imiscível refere-se à inabilidade relativa de um composto dissolver-se em água, e

é definida pelo coeficiente de partição do composto. Aquele de habilidade comum ao estado da arte apreciará que, escolhendo-se um caldo de fermentação e a fase orgânica, de modo que o aldeído, o alcano e/ou o alceno sendo produzidos possuam um elevado valor de logP, o aldeído, o alcano e/ou o alceno podem separar-se na fase orgânica, mesmo que em concentrações muito baixas no recipiente de fermentação.

[258] Os aldeídos, álcoois graxos, alcanos e/ou alcenos produzidos pelos métodos descritos neste documento podem ser relativamente imiscíveis no caldo de fermentação, assim como no citoplasma. No entanto, o aldeído, álcoois graxos, alcanos e/ou alcenos podem ser coletados em uma fase orgânica tanto intracelularmente quanto extracelularmente. A compilação dos produtos na fase orgânica pode diminuir o impacto do aldeído, álcoois graxos, alcanos e/ou alcenos na função celular e pode permitir à célula hospedeira produzir mais produtos.

[259] Os métodos descritos neste documento podem resultar na produção de compostos homogêneos nos quais pelo menos cerca de 60%, 70%, 80%, 90% ou 95% dos aldeídos, álcoois graxos, alcanos e/ou alcenos produzidos tenham comprimentos de cadeia de carbono que variem de menos que cerca de 6 carbonos, menos que cerca de 4 carbonos ou menos que cerca de 2 carbonos. Esses compostos podem também ser produzidos com um grau de saturação relativamente uniforme. Esses compostos podem ser usados diretamente como combustíveis, aditivos de combustível, especialidades químicas, matérias-primas para a produção de outros compostos químicos (por exemplo, polímeros, surfactantes, plásticos, têxteis, solventes, adesivos etc.), ou aditivos de produtos de cuidados pessoais. Esses compostos podem

igualmente ser usados como matérias-primas para reações subsequentes, por exemplo, hidrogenação, quebra catalítica (via hidrogenação, pirólise, ou ambas), para fazer outros produtos.

[260] Em algumas concretizações, os aldeídos, álcoois graxos, alcanos e/ou alcenos produzidos usando-se métodos descritos neste documento podem conter entre cerca de 50% a cerca de 90% de carbono, ou cerca de 5% e cerca de 25% de hidrogênio. Em outras concretizações, os aldeídos, álcoois graxos, alcanos e/ou alcenos produzidos usando-se métodos descritos neste documento podem conter entre cerca de 65% e cerca de 85% de carbono, ou entre cerca de 10% e cerca de 15% de hidrogênio.

Composições combustíveis e composições de especialidade química

[261] Os aldeídos, álcoois graxos, alcanos e/ou alcenos descritos neste documento podem ser usados como combustível ou convertido em combustível, ou como uma especialidade química. Aquele de habilidade comum ao estado da arte apreciará que, dependendo do propósito pretendido do combustível ou especialidade química, diferentes aldeídos, álcoois graxos, alcanos e/ou alcenos podem ser produzidos e usados. Por exemplo, um aldeído, álcool graxo, alcano e/ou alceno ramificado pode ser desejável para combustível de automóvel o qual se pretenda usar em climas frios. Além disso, quando os aldeídos, álcoois graxos, alcanos e/ou alcenos descritos neste documento são usados com uma matéria-prima para a produção de combustível ou especialidade química, aquele de habilidade comum ao estado da técnica apreciará que as características da matéria-prima de

aldeído, álcool graxo, alcano e/ou alceno afetem as características do combustível ou especialidade química produzido. Portanto, as características do produto combustível ou especialidade química podem ser selecionadas pela produção de aldeídos, álcoois graxos, alcanos e/ou alcenos especiais para uso como matéria-prima.

[262] Usando-se os métodos descritos neste documento, biocombustíveis possuindo qualidades combustíveis desejadas podem ser produzidos a partir de aldeídos, álcoois graxos, alcanos e/ou alcenos. Aldeídos, álcoois graxos, alcanos e/ou alcenos biologicamente produzidos representam uma nova fonte de biocombustíveis, que pode ser usada como combustível de jato, *diesel* ou gasolina. Alguns biocombustíveis feitos usando-se aldeídos, álcoois graxos, alcanos e/ou alcenos não foram produzidos a partir de fontes renováveis e são novas composições de matéria. Esses novos combustíveis ou especialidades químicas podem ser diferenciados dos combustíveis ou especialidades químicas derivadas a partir do carbono petroquímico com base na impressão digital isotópica de carbono duplo. Adicionalmente, a fonte específica de carbono de fonte biológica (por exemplo, glicose vs. glicerol) pode ser determinada pela impressão digital isotópica de carbono duplo (ver, por exemplo, patente norte-americana nº 7.169.588, em particular da col. 4, linha 31, à col. 6, linha 8).

[263] Os aldeídos, álcoois graxos, alcanos e/ou alcenos e os biocombustíveis associados, as especialidades químicas e as misturas podem ser diferenciadas a partir de seus homólogos derivados de petroquímico com base no ^{14}C (f_M) e na impressão digital isotópica de carbono duplo. Em alguns

exemplos, o aldeído, álcool graxo, alcano e/ou alceno na composição de biocombustível pode ter uma fração de carbono moderno (f_M^{14C}) de, por exemplo, pelo menos cerca de 1,003, 1,010 ou 1,5.

[264] Em alguns exemplos, uma composição de biocombustível pode ser feita para incluir aldeídos, álcoois graxos, alcanos e/ou alcenos contendo δ^{13C} de cerca de -15,4 a cerca de -10,9, onde os aldeídos, álcoois graxos, alcanos e/ou alcenos contam-se pelo menos cerca de 85% do material de biofonte (*i.e.*, derivado de uma fonte renovável, tal como biomassa, materiais celulósicos e açúcares) na composição.

[265] A habilidade para diferenciar esses produtos biologicamente derivados é benéfica no rastreamento desses materiais no comércio. Por exemplo, combustíveis ou especialidades químicas compreendendo tanto perfis isótopos de carbono biologicamente derivados quanto à base de petróleo podem ser diferenciados dos combustíveis e especialidades químicas feitas apenas de materiais à base de petróleo. Assim, os aldeídos, álcoois graxos, alcanos e/ou alcenos descritos neste documento podem ser seguidos no comércio ou identificados no comércio como um biocombustível com base em seus perfis únicos. Além disso, outros materiais concorrentes podem ser identificados como sendo biologicamente derivados ou derivados a partir de uma fonte petroquímica.

[266] Aditivos de combustíveis são usados para melhorar a performance de um combustível ou motor. Por exemplo, aditivos de combustível podem ser usados para alterar o ponto de resfriamento/gelificação, o ponto de nuvem, a lubrificidade, a viscosidade, a estabilidade oxidativa, a

qualidade da ignição, o nível de octano e/ou o ponto de centelha. Nos Estados Unidos, todos os aditivos de combustível devem ser registrados na Agência de Proteção Ambiental (EPA). Os nomes de aditivos de combustível e as companhias que os vendem estão publicamente disponíveis contactando-se a EPA ou vendo pelo *website* da agência. Aquele de habilidade comum ao estado da técnica apreciará que os biocombustíveis à base de aldeído e/ou alceno descritos neste documento possam ser misturados com um ou mais aditivos de combustíveis para conferir a qualidade desejada.

[267] Os biocombustíveis à base de aldeídos, álcoois graxos, alcanos e/ou alcenos descritos neste documento podem ser misturados a outros combustíveis, tais como diversos álcoois, tais como etanol e butanol, e produtos derivados de petróleo, tais como gasolina, *diesel* ou combustível de jato.

[268] Em alguns exemplos, a mistura pode incluir pelo menos cerca de 10%, 15%, 20%, 30%, 40%, 50% ou 60% por peso de aldeído, álcoois graxos, alceno ou alceno. Em outros exemplos, uma composição de biocombustível pode ser feita para incluir pelo menos cerca de 5%, 10%, 15%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou 95% de aldeído, álcoois graxos, alceno ou alceno que inclua uma cadeia de carbono que seja de 8, 10, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 ou 22 carbonos em comprimento. Tais composições de biocombustível podem opcionalmente incluir pelo menos um aditivo selecionado a partir de um aditivo de redução de ponto de nuvem que possa diminuir o ponto de nuvem para menos que cerca de 5°C ou 0°C; um surfactante; uma microemulsão; a pelo menos cerca de 5%, 10%, 15%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou 95% combustível *diesel* a partir de

triglicerídeos; gasolina derivada de petróleo, ou combustível *diesel* a partir de petróleo.

EXEMPLOS

[269] A invenção é ainda descrita nos exemplos seguintes, os quais não limitam o escopo da invenção descrito nas reivindicações.

Exemplo 1. Detecção e verificação de biossíntese de alcano na cianobactéria selecionada

[270] Sete cianobactérias, cujas sequências genômicas completas estão disponíveis ao público, foram selecionadas para verificação e / ou detecção de biossíntese de alcanos: *Synechococcus elongatus* PCC7942, *Synechococcus elongatus* PCC6301, *Anabaena variabilis* ATCC29413, *Synechocystis* sp. PCC6803, *Nostoc punctiforme* PCC73102, *Gloeobacter violaceus* ATCC 29082, e *Prochlorococcus marinus* CCMP 1986. Apenas as três primeiras linhagens de cianobactérias a partir dessa lista haviam sido relatadas previamente para conter alcanos (HAN *et al.*, *J. Am. Chem. Soc.* 91:5156-5159 (1969); FEHLER *et al.*, *Biochem.* 9:418-422 (1970)). As linhagens cresceram fotoautotroficamente em frascos de agitação em 100mL do meio apropriado (listado na Tabela 8) por 3-7 dias a 30°C a uma intensidade de luz de aproximadamente 3.500 lux. As células foram extraídas para a detecção de alcano da seguinte forma: as células de 1mL de volume de cultura foram centrifugadas por 1 min a 13.000 rpm, as pelotas de células foram ressuspensas em metanol, centrifugadas por 1 min e então sonicadas por 30 min. Após a centrifugação por 3 min a 13.000 rpm, os sobrenadantes foram transferidos para frascos frescos e analisados por GC-MS. As amostras foram analisadas tanto em coluna capilar DP-5 de 30 m (0,25 mm de diâmetro

interno) quanto em coluna capilar DP-5 de alta temperatura de 30 m (0,25 mm diâmetro interno) utilizando o seguinte método.

[271] Depois de uma injeção sem separação de 1 µL (temperatura de entrada realizada a 300°C) na coluna do CG/MS, a fornada foi realizada a 100°C por 3 minutos. A temperatura foi aumentada para 320°C a uma taxa de 20°C/min. A fornada foi realizada a 320°C por mais 5 min. A taxa de fluxo do transportador de gás hélio foi 1,3mL/min. O quadrupolo MS escaneado de 50 a 550 m/z. Os tempos de retenção e os padrões de fragmentação dos picos de produto foram comparadas com as referências para confirmar a identidade de pico.

[272] Das sete linhagens, seis produziram principalmente heptadecano e um produziu pentadecano (*P. marinus* CCMP1986); uma destas linhagens produziu metil-heptadecano além de heptadecano (*A. variabilis* ATCC29413) (ver Tabela 8). Assim, a biossíntese de alcanos em três cianobactérias anteriormente relatadas foi verificada, e biossíntese de alcanos foi detectado em quatro cianobactérias que não eram conhecidas anteriormente para produzir alcanos: *P. marinus* CCMP 1986 (ver Figura 1), *N. punctiforme* PCC73102 (ver Figura 2), *G. violaceus* ATCC 29082 (ver Figura 3) e *Synechocystis* sp. PCC6803 (ver Figura 4).

[273] A Figura 1A retrata o traço GC/MS de células *Prochlorococcus marinus* CCMP 1986 extraídas com metanol. O pico em 7,55 min tinha o mesmo tempo de retenção do pentadecano (Sigma). Na Figura 1B, o padrão de fragmentação de massas do pico de pentadecano é mostrado. O pico de 212 corresponde ao peso molecular de pentadecano.

[274] A Figura 2A mostra o traço GC/MS de células *Nostoc punctiforme* PCC73102 extraídas com metanol. O pico em 8,73 min tem o mesmo tempo de retenção do heptadecano (Sigma). Na Figura 2B, o padrão de fragmentação de massas do pico de heptadecano é mostrado. O pico de 240 corresponde ao peso molecular de heptadecano.

[275] A Figura 3A mostra o traço de GC/MS de células *Gloeobaceter violaceus* ATCC29082 extraídas com metanol. O pico em 8,72 min tem o mesmo tempo de retenção do heptadecano (Sigma). Na Figura 3B, o padrão de fragmentação de massas do pico de heptadecano é mostrado. O pico de 240 corresponde ao peso molecular de heptadecano.

[276] A Figura 4A mostra o traço GC/MS de células *Synechocystis* sp. PCC6803 extraídas com metanol. O pico em 7,36 min tem o mesmo tempo de retenção do heptadecano (Sigma). Na Figura 4B, o padrão de fragmentação de massas do pico de heptadecano é mostrado. O pico de 240 corresponde ao peso molecular de heptadecano.

Tabela 8: Hidrocarbonetos detectados na cianobactéria selecionada

Cianobactérias	ATCC#	Genoma	Meio	Alcanos	
				relatado	verificado ²
<i>Synechococcus elongatus</i> PCC7942	27144	2,7Mb	BG-11	C17:0	C17:0 , C15:0
<i>Synechococcus elongatus</i> PCC6301	33912	2,7Mb	BG-11	C17:0	C17:0 , C15:0
<i>Anabaena variabilis</i>	29413	6,4Mb	BG-11	C17:0, 7- ou 8-Me- C17:0	C17:0 , Me-C17:0
<i>Synechocystis</i> sp. PCC6803	27184	3,5Mb	BG-11		C17:0 , C15:0
<i>Prochlorococcus marinus</i> CCMP1986 ¹	-	1,7Mb	-		C15:0
<i>Nostoc punctiforme</i> PCC73102	29133	9,0Mb	ATCC819		C17:0
<i>Gloeobacter violaceus</i>	29082	4,6Mb	BG11		C17:0

¹células para a extração foram um presente de Jacob Waldbauer (MIT)

²maior hidrocarboneto em negrito

[277] Análise Genômica rendeu dois genes que estavam presentes nas linhagens de produção de alcano. O homólogo da *Synechococcus elongatus* PCC7942 desses genes estão representados na Tabela 9 e são Synpcc7942_1593 (SEQ ID NO: 1) e Synpcc7942_1594 (SEQ ID NO: 65).

Tabela 9: Genes cianobacteriais de produção de alcano

ID de Obejto do Gene	Etiqueta de Local	Adesão ao GenBank	Nome do Gene	Comprimento	COG	Pfam	InterPro	Notas
637800026	Synpcc7942_1593	YP_400610	proteína hipotética	231 aa	-	pfam02915	IPR009078 IPR003251	ferritina/ribonucleotídeo semelhante à reductase rubreritina
637800027	Synpcc7942_1594	YP_400611	proteína hipotética	341 aa	COG5322	pfam00106	IPR000408 IPR016040 IPR002198	deidrogenase prevista NAD(P)-ligante Deidrogenase de cadeia curta

Exemplo 2: Supressão dos genes sll0208 e sll0209 em *Synechocystis* sp. PCC6803 conduz à perda da biossíntese de alcano

[278] Os genes que codificam a descarboxilase suposta (sll0208; NP 442147) (SEQ ID NO: 3) e a enzima geradora de aldeídos (sll0209; NP 442146) (SEQ ID NO: 67) de *Synechocystis* sp. PCC6803 foram eliminados da seguinte forma. Cerca de 1 kb à montante e DNA flanqueado à jusante foram amplificados com o iniciador sll0208/9- K01 (CGCGGATCCCTTGATTCTACTGCGGCGAGT) com o iniciador sll0208/9- K02 (CACGCACCTAGGTTCACTCCCATGGTATAACAGGGGCGTTGGACTCCTGTG)

o iniciador s110208/9-KO3 (GTTATACCATGGGAGTGTGAACCTAGGTGCGTGGCCGACAGGATAGGGCGTGT) com o iniciador s110208/9-KO4 (CGCGGATCCAACGCATCCTCACTAGTCGGG), respectivamente. Os produtos de PCR foram usados em uma PCR cruzada com iniciadores s110208/9-KO1 e s110208/9-KO4 para amplificar o cassete de eliminação aproximadamente 2kb s110208/s110209, que foi clonado no local *Bam*HI do vetor de clonagem pUC19. Um cassete de resistência à canamicina (*aph*, KanR) foi, então, amplificado a partir do plasmídeo pRL27 (LARSEN *et al. Arch. Microbiol.* 178:193 (2002)), usando iniciadores Kan-*aph*-F (CATGCCATGGAAAGCCACGTTGTGTCTCAAAATCTCTG) e Kan-*aph*-R (CTAGTCTAGAGCGCTGAGGTCTGCCTCGTGAA), que foi então seccionado com *Nco*I e *Xba*I e clonado nos locais *Nco*I e *Avr*II do cassete de eliminação s110208/s110209, criando um cassete de inserção de KanR supressão de s110208/s110209 em pUC19. O vetor contendo cassete, que não se multiplica em cianobactérias, foi transformado em *Synechocystis* sp. PCC6803 (ZANG *et al.*, 2007, *J. Microbiol.* vol. 45, p. 241) e transformantes (por exemplo, integrantes cromossômicas por dupla recombinação homóloga) foram selecionados em placas de ágar BG-11 contendo 100µg/mL de canamicina em uma incubadora equipadas com luz a 30°C. As colônias resistentes à canamicina foram relistradas uma vez e, então, sujeitadas à análise genotípica usando-se PCR com iniciadores de diagnósticos.

[279] Os mutantes de inserção-supressão foram cultivados em 12mL de meio BG11 com 50 µg/mL de canamicina por 4 dias a 30°C em uma incubadora de agitação equipada com luz. 1 mL de caldo foi então centrifugado (1 min a 13.000 g) e os

aglomerados de células foram extraídos com 0,1 mL de metanol. Após a extração, as amostras foram novamente centrifugadas e os sobrenadantes foram submetidos à análise de GC-MS, conforme descrito no Exemplo 1.

[280] Como mostrado na Fig. 5, as linhagens de *Synechocystis* sp. PCC6803 das quais os genes *sll0208* e *sll0209* foram suprimidos perderam sua capacidade de produzir heptadeceno e octadecenal. Este resultado demonstra que os genes *sll0208* e *sll0209* em *Synechocystis* sp. PCC6803 e os genes ortólogos em outra cianobactéria (ver Tabela 1) são responsáveis pela biossíntese de alcanos e aldeídos graxos nestes organismos.

Exemplo 3: Produção de aldeídos graxos e álcoois graxos em *E. coli* através da expressão heteróloga de *Synechococcus elongatus* PCC7942 orf1594

[281] O DNA genômico codificando *Synechococcus elongatus* PCC7942 orf1594 (YP_400611; enzima suposta geradora de aldeídos) (SEQ ID NO: 65) foi amplificado e clonado nos locais *Nco*I e *Eco*RI do vetor OP-80 (pCL1920 derivados) sob o controle do promotor P_{trc} . A construção resultante ("OP80-PCC7942_1594") foi transformada em *E. coli* MG1655 e as células foram cultivadas a 37°C em meio M9 mínimo com 1% (p/v) de glicose como fonte de carbono e suplementado com 100µg/mL de espectinomicina. Quando a cultura atingiu OD₆₀₀ de 0,8-1,0, ela foi induzida com 1mM de IPTG e as células foram cultivadas por um adicional de 18-20 horas a 37°C. Células de 0,5mL de cultura foram extraídas com 0,5mL de acetato de etila. Após a sonicação por 60 min, a amostra foi centrifugada a 15.000 rpm por 5 min. A camada de solvente foi analisada por CG-MS, conforme descrito no Exemplo 1.

[282] Como mostrado na Fig. 6, células de *E. coli* transformadas com o vetor condutor de *Synechococcus elongatus* PCC7942 orf1594 produziram os seguintes aldeídos graxos e álcoois graxos: hexadecanal, octadecenal, tetradecenol, hexadecanol, hexadecenol e octadecenol. Esse resultado indica que PCC7942 orf1594 (i) gera aldeídos *in vivo* como possíveis substratos para descarbonilação e (ii) pode reduzir acil-ACPs como substratos, que são a forma mais abundante dos ácidos graxos ativados em células de *E. coli* do tipo selvagem. Portanto, a enzima foi denominada de acil-ACP redutase. *In vivo*, os aldeídos graxos aparentemente são ainda mais reduzidos para os álcoois graxos correspondentes por uma atividade endógena de *E. coli* aldeído redutase.

Exemplo 4: Produção de aldeídos graxos e álcoois graxos em *E. coli* através da expressão heteróloga de *Cyanothece* sp. ATCC51142 cce 1430

[283] O DNA genômico codificando *Cyanothece* sp. ATCC51142 cce_1430 (YP_001802846; enzima suposta geradora de aldeídos) (SEQ ID NO: 69) foi amplificado e clonado nos locais *Nco*I e *Eco*RI do vetor OP-80 (pCL1920 derivados), sob o controle do promotor P_{trc} . A construção resultante foi transformada em *E. coli* MG1655 e as células foram cultivadas a 37°C em meio M9 mínimo com 1% (p/v) de glicose como fonte de carbono e suplementado com 100 µg/mL de espectinomicina. As células foram cultivadas e extraídas como no Exemplo 3 e analisadas por GC-MS, conforme descrito no Exemplo 26.

[284] Como mostrado na Fig. 7, células de *E. coli* transformadas com o vetor transportador de *Cyanothece* sp. ATCC51142 cce_1430 produziram os seguintes aldeídos graxos e álcoois graxos: hexadecanal, octadecenal, tetradecenol,

hexadecanol, hexadecenol e octadecenol. Esse resultado indica que ATCC51142 *cce_1430* (i) gera aldeídos *in vivo* como possíveis substratos para descarbonilação e (ii) pode reduzir acil-ACPs como substratos, que são a forma mais abundante dos ácidos graxos ativados em células de *E. coli* do tipo selvagem. Portanto, esta enzima também é uma acil-ACP redutase.

Exemplo 5: Produção de alcanos e alcenos em *E. coli* através da expressão heteróloga de *Synechococcus elongatus* PCC7942 orf11594 e *Synechococcus elongatus* PCC7942 orf11593

[285] O DNA genômico codificando *Synechococcus elongatus* PCC7942 orf1593 (YP_400610; descarbonilase suposta) (SEQ ID NO: 1) foi amplificado e clonado nos locais *NdeI* e *XhoI* do vetor OP-183 (pACYC derivados) sob o controle do promotor P_{trc} . A construção resultante foi cotransformada com OP80-PCC7942_1594 em *E. coli* MG1655 e as células foram cultivadas a 37°C em meio M9 mínimo e suplementadas com 100 µg/mL de espectinomicina e 100 µg/mL de carbenicilina. As células foram cultivadas e extraídas como no Exemplo 3 e analisadas por GC-MS, conforme descrito no Exemplo 1.

[286] Como mostrado na Fig. 8, células de *E. coli* cotransformadas com os vetores de transporte *S. elongatus* PCC7942_1594 e *S. elongatus* PCC7942_1593 produziram os mesmos aldeídos graxos e álcoois graxos como no Exemplo 3, mas também pentadecano e heptadeceno. Esse resultado indica que PCC7942_1593 em *E. coli* converte hexadecanal e octadecenal para pentadecano e heptadeceno, respectivamente, e, portanto, é uma descarbonilase ativa de aldeído graxo.

Exemplo 6: Produção de alcanos e alcenos em *E. coli* através da expressão heteróloga de *Synechococcus elongatus* PCC7942

orf11594 e *Nostoc punctiforme* PCC73102 Npun02004178

[287] O DNA genômico codificando *Nostoc punctiforme* PCC73102 Npun02004178 (ZP_00108838; descarbonilase suposta) (SEQ ID NO: 5) foi amplificado e clonado nos locais *NdeI* e *XhoI* do vetor OP-183 (pACYC derivados), sob o controle do promotor P_{trc}. A construção resultante foi cotransformada com OP80-PCC7942_1594 em *E. coli* MG1655e as células foram cultivadas a 37°C em meio M9 mínimo e suplementadas com 100 µg/mL de espectinomicina e 100 µg/mL de carbenicilina. As células foram cultivadas e extraídas como no Exemplo 3 e analisadas por GC-MS, conforme descrito no Exemplo 1.

[288] Como mostrado na Fig. 9, células de *E. coli* cotransformados com os vetores de transporte *S. elongatus* PCC7942_1594 e *N. punctiforme* PCC73102 Npun02004178 produziram os mesmos aldeídos graxos e álcoois graxos como no Exemplo 3, mas também tridecano, pentadecano, pentadeceno e heptadeceno. Esse resultado indica que Npun02004178 em *E. coli* converte tetradecanal, hexadecenal, hexadecanal e octadecenal para tridecano, pentadecano pentadeceno e heptadeceno, respectivamente, e, portanto, é uma descarbonilase ativa de aldeído graxo.

Exemplo 7: Produção de alcanos e alcenos em *E. coli* através da expressão heteróloga de *Synechococcus elongatus* PCC7942 orf11594 e *Synechocystis* sp. PCC6803 sl10208

[289] O DNA genômico codificando *Synechocystis* sp. PCC6803 sl10208 (NP_442147; descarbonilase suposta) (SEQ ID NO: 3) foi amplificado e clonado nos locais *NdeI* e *XhoI* do vetor OP-183 (pACYC derivados), sob o controle do promotor P_{trc}. A construção resultante foi cotransformada com OP80-PCC7942_1594 em *E. coli* MG1655 e as células foram cultivadas

a 37°C em meio M9 mínimo suplementado com 100 µg/mL de espectinomicina e 100 µg/mL de carbenicilina. As células foram cultivadas e extraídas como no Exemplo 3 e analisadas por GC-MS, conforme descrito no Exemplo 1.

[290] Como mostrado na Fig. 10, células de *E. coli* cotransformadas com os vetores de transporte *S. elongatus* PCC7942_1594 e *Synechocystis* sp. PCC6803 sl10208 produziram os mesmos aldeídos graxos e álcoois graxos como no Exemplo 3, mas também pentadecano e heptadeceno. Esse resultado indica que Npun02004178 em *E. coli* converte hexadecanal e octadecenal para pentadecano e heptadeceno, respectivamente, e, portanto, é uma descarbonilase ativa de aldeído graxo.

Exemplo 8: Produção de alcanos e alcenos em *E. coli* através da expressão heteróloga de *Synechococcus elongatus* PCC7942 orf11594 e *Nostoc* sp. PCC7210 alr5283

[291] O DNA genômico codificando *Nostoc* sp. PCC7210 alr5283 (NP_489323; descarbonilase suposta) (SEQ ID NO: 7) foi amplificado e clonado nos locais *Nde*I e *Xho*I do vetor OP-183 (pACYC derivados), sob o controle do promotor P_{trc}. A construção resultante foi cotransformada com OP80-PCC7942_1594 em *E. coli* MG 1655 e as células foram cultivadas a 37°C em meio M9 mínimo suplementado com 100 µg/mL de espectinomicina e 100 µg/mL de carbenicilina. As células foram cultivadas e extraídas como no Exemplo 3 e analisadas por GC-MS, conforme descrito no Exemplo 1.

[292] Como mostrado na Fig. 11, células de *E. coli* cotransformadas com os vetores de transporte *S. elongatus* PCC7942_1594 e *Nostoc* sp. PCC7210 alr5283 produziram os mesmos aldeídos graxos e álcoois graxos como no Exemplo 3, mas também pentadecano e heptadeceno. Esse resultado indica

que alr5283 em *E. coli* converte hexadecanal e octadecenal para pentadecano e heptadeceno, respectivamente, e, portanto, é uma descarbonilase ativa de aldeído graxo.

Exemplo 9: Produção de alcanos e alcenos em *E. coli* através da expressão heteróloga de *Synechococcus elongatus* PCC7942 orf11594 e *Acaryochloris marina* MBIC11017 AM1 4041

[293] O DNA genômico codificando *Acaryochloris marina* MBIC11017 AM1_4041 (YP_001518340; descarbonilase suposta) (SEQ ID NO: 9) foi otimizado de códon para a expressão em *E. coli* (SEQ ID NO: 46), sintetizado e clonado nos locais *NdeI* e *XhoI* do vetor OP-183 (derivados pACYC) sob o controle do promotor P_{trc}. A construção resultante foi cotransformada com OP80-PCC7942_1594 em *E. coli* MG1655 e as células foram cultivadas a 37°C em meio M9 mínimo suplementado com 100 µg/mL de espectinomicina e 100 µg/mL de carbenicilina. As células foram cultivadas e extraídas como no Exemplo 3 e analisadas por GC-MS, conforme descrito no Exemplo 26.

[294] Como mostrado na Fig. 12, células de *E. coli* cotransformadas com os vetores de transporte *S. elongatus* PCC7942_1594 e *A.marina* MBIC11017 AM1_4041 produziram os mesmos aldeídos graxos e álcoois graxos como no Exemplo 3, mas também tridecano, pentadecano, pentadeceno e heptadeceno. Esse resultado indica que a AM1_4041 em *E. coli* converte tetradecanal, hexadecenal, hexadecanal e octadecenal para tridecano, pentadecano, pentadeceno e heptadeceno, respectivamente, e, portanto, é uma descarbonilase ativa de aldeído graxo.

Exemplo 10: Produção de alcanos e alcenos em *E. coli* através da expressão heteróloga de *Synechococcus elongatus* PCC7942 orf11594 e *Thermosynechococcus elongatus* BP-1 t111313

[295] O DNA genômico codificando *Thermosynechococcus elongatus* BP-1 tll1313 (NP_682103; descarbonilase suposta) (SEQ ID NO: 11) foi otimizado de códon para a expressão em *E. coli* (SEQ ID NO: 47), sintetizado e clonado nos locais *NdeI* e *XhoI* do vetor OP-183 (derivados pACYC) sob o controle do promotor P_{trc} . A construção resultante foi cotransformada com OP80-PCC7942_1594 em *E. coli* MG1655 e as células foram cultivadas a 37°C em meio M9 mínimo suplementado com 100 µg/mL de espectinomicina e 100 µg/mL de carbenicilina. As células foram cultivadas e extraídas como no Exemplo 3 e analisadas por GC-MS, conforme descrito no Exemplo 26.

[296] Como mostrado na Fig. 13, células de *E. coli* cotransformadas com os vetores de transporte *S. elongatus* PCC7942_1594 e *T. elongatus* BP-1 tll1313 produziram os mesmos aldeídos graxos e álcoois graxos como no Exemplo 3, mas também pentadecano e heptadeceno. Esse resultado indica que tll1313 em *E. coli* converte hexadecanal e octadecenal para pentadecano e heptadeceno, respectivamente, e, portanto, é uma descarbonilase ativa de aldeído graxo.

Exemplo 11: Produção de alcanos e alcenos em *E. coli* através da expressão heteróloga de *Synechococcus elongatus* PCC7942 orf11594 e *Synechococcus* sp. JA-3-3Ab CYA 0415

[297] O DNA genômico codificando *Synechococcus* sp. JA-3-3Ab CYA_0415 (YP_473897; descarbonilase suposta) (SEQ ID NO: 13) foi otimizado de códon para a expressão em *E. coli* (SEQ ID NO: 48), sintetizado e clonado nos locais *NdeI* e *XhoI* do vetor OP-183 (pACYC derivados) sob o controle do promotor P_{trc} . A construção resultante foi cotransformada com OP80-PCC7942_1594 em *E. coli* MG 1655 e as células foram cultivadas a 37°C em meio M9 mínimo suplementado com 100

µg/mL de espectinomicina e 100 µg/mL de carbenicilina. As células foram cultivadas e extraídas como no Exemplo 3 e analisadas por GC-MS, conforme descrito no Exemplo 26.

[298] Como mostrado na Fig. 14, células de *E. coli* cotransformadas com os vetores de transporte *S. elongatus* PCC7942_1594 e *Synechococcus* sp. JA-3-3Ab CYA_0415 produziram os mesmo aldeídos graxos e álcoois graxos como no Exemplo 3, mas também pentadecano e heptadeceno. Esse resultado indica que Npun02004178 em *E. coli* converte hexadecanal e octadecenal para pentadecano e heptadeceno, respectivamente, e, portanto, é uma descarbonilase ativa de aldeído graxo.

Exemplo 12: Produção de alcanos e alcenos em *E. coli* através da expressão heteróloga de *Synechococcus elongatus* PCC7942 orfl1594 e *Gloeobacter violaceus* PCC7421 gll3146

[299] O DNA genômico codificando *Gloeobacter violaceus* PCC7421 gll3146 (NP_926092; descarbonilase suposta) (SEQ ID NO: 15) foi amplificado e clonado nos locais *Nde*I e *Xho*I do vetor OP-183 (pACYC derivados), sob o controle do promotor *P_{trc}*. A construção resultante foi cotransformada com OP80-PCC7942_1594 em *E. coli* MG1655 e as células foram cultivadas a 37°C em meio M9 mínimo suplementado com 100 µg/mL de espectinomicina e 100 µg/mL de carbenicilina. As células foram cultivadas e extraídas como no Exemplo 3 e analisados por GC-MS, conforme descrito no Exemplo 1.

[300] Como mostrado na Fig. 15, células de *E. coli* cotransformadas com os vetores de transporte *S. elongatus* PCC7942_1594 e *G. violaceus* PCC7421 gll3146 produziram os mesmos aldeídos graxos e álcoois graxos como no Exemplo 3, mas também pentadecano e heptadeceno. Esse resultado indica

que gll13146 em *E. coli* converte hexadecanal e octadecenal para pentadecano e heptadeceno, respectivamente, e, portanto, é uma descarbonilase ativa de aldeído graxo.

Exemplo 13: Produção de alcanos e alcenos em *E. coli* através da expressão heteróloga de *Synechococcus elongatus* PCC7942 orfl1594 e *Prochlorococcus marinus* MIT9313 PMT1231

[301] O DNA genômico codificando *Prochlorococcus marinus* MIT9313 PMT1231 (NP_895059; descarbonilase suposta) (SEQ ID NO: 17) foi otimizado de códon para a expressão em *E. coli* (SEQ ID NO: 49), sintetizado e clonado nos locais *NdeI* e *XhoI* do vetor OP-183 (pACYC derivados) sob o controle do promotor P_{trc} . A construção resultante foi cotransformada com OP80-PCC7942_1594 em *E. coli* MG1655 e as células foram cultivadas a 37°C em meio M9 mínimo suplementado com 100 µg/mL de espectinomicina e 100 µg/mL de carbenicilina. As células foram cultivadas e extraídas como no Exemplo 3 e analisadas por GC-MS, conforme descrito no Exemplo 26.

[302] Como mostrado na Fig. 16, células de *E. coli* cotransformados com os vetores de transporte *S. elongatus* PCC7942_1594 e *P. marinus* MIT9313 PMT1231 produziram os mesmos aldeídos graxos e álcoois graxos como no Exemplo 3, mas também pentadecano e heptadeceno. Este resultado indica que a PMT1231 em *E. coli* converte hexadecanal e octadecenal para pentadecano e heptadeceno, respectivamente, e, portanto, é uma descarbonilase ativa de aldeído graxo.

Exemplo 14: Produção de alcanos e alcenos em *E. coli* através da expressão heteróloga de *Synechococcus elongatus* PCC7942 orfl1594 e *Prochlorococcus marinus* CCMP1986 PMM0532

[303] O DNA genômico codificando *Prochlorococcus marinus* CCMP1986 PMM0532 (NP_892650; descarbonilase suposta) (SEQ ID

NO: 19) foi amplificado e clonado nos locais *NdeI* e *XhoI* do vetor OP-183 (pACYC derivados) sob o controle do promotor P_{trc} . A construção resultante foi cotransformada com OP80-PCC7942_1594 em *E. coli* MG 1655 e as células foram cultivadas a 37°C em meio M9 mínimo suplementado com 100 µg/mL de espectinomicina e 100 µg/mL de carbenicilina. As células foram cultivadas e extraídas como no Exemplo 3 e analisadas por GC-MS, conforme descrito no Exemplo 1.

[304] Como mostrado na Fig. 17, células de *E. coli* cotransformadas com os vetores de transporte *S. elongatus* PCC7942_1594 e *P. marinus* CCMP1986 PMM0532 produziram os mesmos aldeídos graxos e álcoois graxos como no Exemplo 3, mas também pentadecano e heptadeceno. Esse resultado indica que PMM0532 em *E. coli* converte hexadecanal e octadecenal para pentadecano e heptadeceno, respectivamente, e, portanto, é uma descarbonilase ativa de aldeído graxo.

Exemplo 15: Produção de alcanos e alcenos em *E. coli* através da expressão heteróloga de *Synechococcus elongatus* PCC7942 orf11594 e *Prochlorococcus marinus* NATL2A PMN2A 1863

[305] O DNA genômico codificando *Prochlorococcus mariunus* NATL2A PMN2A_1863 (YP_293054; descarbonilase suposta) (SEQ ID NO: 21) foi otimizado de códon para a expressão em *E. coli* (SEQ ID NO: 51), sintetizado e clonado nos locais *NdeI* e *XhoI* do vetor OP-183 (pACYC derivados), sob o controle do promotor P_{trc} . A construção resultante foi cotransformada com OP80-PCC7942_1594 em *E. coli* MG1655 e as células foram cultivadas a 37°C em meio M9 mínimo suplementado com 100 µg/mL de espectinomicina e 100 µg/mL de carbenicilina. As células foram cultivadas e extraídas como no Exemplo 3 e analisadas por GC-MS, conforme descrito no

Exemplo 26.

[306] Como mostrado na Fig. 18, células de *E. coli* cotransformadas com os vetores de transporte *S. elongatus* PCC7942_1594 e *P. mariunus* NATL2A PMN2A_1863 produziram os mesmos aldeídos graxos e álcoois graxos como no Exemplo 3, mas também pentadecano e heptadeceno. Esse resultado indica que PMN2A_1863 em *E. coli* converte hexadecanal e octadecenal para pentadecano e heptadeceno, respectivamente, e, portanto, é uma descarbonilase ativa de aldeído graxo.

Exemplo 16: Produção de alcanos e alcenos em *E. coli* através da expressão heteróloga de *Synechococcus elongatus* PCC7942 orf11594 e *Synechococcus* sp. RS9917 RS9917_09941

[307] O DNA genômico codificando *Synechococcus* sp. RS9917 RS9917_09941 (ZP_O1079772; descarbonilase suposta) (SEQ ID NO: 23) foi otimizado de códon para a expressão em *E. coli* (SEQ ID NO: 52), sintetizado e clonado nos locais *NdeI* e *XhoI* do vetor OP-183 (pACYC derivados) sob o controle do promotor P_{trc} . A construção resultante foi cotransformada com OP80-PCC7942_1594 em *E. coli* MG 1655 e as células foram cultivadas a 37°C em meio M9 mínimo suplementado com 100 µg/mL de espectinomicina e 100 µg/mL de carbenicilina. As células foram cultivadas e extraídas como no Exemplo 3 e analisadas por GC-MS, conforme descrito no Exemplo 26.

[308] Como mostrado na Fig. 19, células de *E. coli* cotransformadas com os vetores de transporte *S. elongatus* PCC7942_1594 e *Synechococcus* sp. RS9917 RS9917_09941 produziram os mesmos aldeídos graxos e álcoois graxos como no Exemplo 3, mas também pentadecano e heptadeceno. Esse resultado indica que RS9917_09941 em *E. coli* converte hexadecanal e octadecenal para pentadecano e heptadeceno,

respectivamente, e, portanto, é uma descarbonilase ativa de aldeído graxo.

Exemplo 17: Produção de alcanos e alcenos em *E. coli* através da expressão heteróloga de *Synechococcus elongatus* PCC7942 orfl1594 e *Synechococcus* sp. RS9917 RS9917 12945

[309] O DNA genômico codificando *Synechococcus* sp. RS9917 RS9917_12945 (ZP_01080370; descarbonilase suposta) (SEQ ID NO: 25) foi otimizado de códon para a expressão em *E. coli* (SEQ ID NO: 53), sintetizado e clonado nos locais *NdeI* e *XhoI* do vetor OP-183 (pACYC derivados) sob o controle do promotor P_{trc} . A construção resultante foi cotransformada com OP80-PCC7942_1594 em *E. coli* MG 1655 e as células foram cultivadas a 37°C em meio M9 mínimo suplementado com 100 µg/mL de espectinomicina e 100 µg/mL de carbenicilina. As células foram cultivadas e extraídas como no Exemplo 3 e analisadas por GC-MS, conforme descrito no Exemplo 26.

[310] Como mostrado na Fig. 20, células de *E. coli* cotransformadas com os vetores de transporte *S. elongatus* PCC7942_1594 e *Synechococcus* sp. RS9917 RS9917_12945 produziram os mesmos aldeídos graxos e álcoois graxos como no Exemplo 3, mas também pentadecano e heptadeceno. Esse resultado indica que RS9917_12945 em *E. coli* converte hexadecanal e octadecenal para pentadecano e heptadeceno, respectivamente, e, portanto, é uma descarbonilase ativa de aldeído graxo.

Exemplo 18: Produção de alcanos e alcenos em *E. coli* através da expressão heteróloga de *Synechococcus elongatus* PCC7942 orfl1594 e *Cyanothece* sp. ATCC51142 cce 0778

[311] O DNA genômico codificando *Cyanothece* sp. ATCC51142 cce_0778 (YP_001802195; descarbonilase suposta)

(SEQ ID NO: 27) foi sintetizado e clonado nos locais *NdeI* e *XhoI* do vetor OP-183 (pACYC derivados) no âmbito do controle do promotor P_{trc} . A construção resultante foi cotransformada com OP80-PCC7942_1594 em *E. coli* MG1655 e as células foram cultivadas a 37°C em meio M9 mínimo suplementado com 100 µg/mL de espectinomicina e 100 µg/mL de carbenicilina. As células foram cultivadas e extraídas como no Exemplo 3 e analisadas por GC-MS, conforme descrito no Exemplo 26.

[312] Como mostrado na Fig. 21, células de *E. coli* cotransformados com os vetores de transporte *S. elongatus* PCC7942_1594 e *Cyanotheca* sp. ATCC51142 cce_0778 produziram os mesmos aldeídos graxos e álcoois graxos como no Exemplo 3, mas também tridecano, pentadecano, pentadeceno e heptadeceno. Esse resultado indica que ATCC51142 cce_0778 em *E. coli* converte tetradecanal, hexadecenal, hexadecanal e octadecenal para tridecano, pentadecano, pentadeceno e heptadeceno, respectivamente, e, portanto, é uma descarbonilase ativa de aldeído graxo.

Exemplo 19: Produção de alcanos e alcenos em *E. coli* através da expressão heteróloga de *Synechococcus elongatus* PCC7942 orf11594 e *Cyanotheca* sp. PCC7425 Cyan7425 0398

[313] O DNA genômico codificando *Cyanotheca* sp. PCC7425 Cyan7425_0398 (YP_002481151; descarbonilase suposta) (SEQ ID NO: 29) foi sintetizado e clonado nos locais *NdeI* e *XhoI* do vetor OP-183 (derivados pACYC) sob o controle do promotor P_{trc} . A construção resultante foi cotransformados com OP80-PCC7942_1594 em *E. coli* MG1655 e as células foram cultivadas a 37°C em meio M9 mínimo suplementado com 100 µg/mL de espectinomicina e 100 µg/mL de carbenicilina. As células foram cultivadas e extraídas como no Exemplo 3 e analisadas

por GC-MS, conforme descrito no Exemplo 26.

[314] Como mostrado na Fig. 22, células de *E. coli* cotransformados com os vetores de transporte *S. elongatus* PCC7942_1594 e *Cyanothece* sp. PCC7425 Cyan7425_0398 produziram os mesmos aldeídos graxos e álcoois graxos como no Exemplo 3, mas também tridecano, pentadecano, pentadeceno e heptadeceno. Esse resultado indica que Cyan7425_0398 em *E. coli* converte tetradecanal, hexadecenal, hexadecanal e octadecenal para tridecano, pentadecano, pentadeceno e heptadeceno, respectivamente, e, portanto, é uma descarboxilase ativa de aldeído graxo.

Exemplo 20: Produção de alcanos e alcenos em *E. coli* através da expressão heteróloga de *Synechococcus elongatus* PCC7942 orfl1594 e *Cyanothece* sp. PCC7425 Cyan7425 2986

[315] O DNA genômico codificando *Cyanothece* sp. PCC7425 Cyan7425_2986 (YP_002483683; descarboxilase suposta) (SEQ ID NO: 31) foi sintetizado e clonado nos locais *NdeI* e *XhoI* do vetor OP-183 (pACYC derivados) sob o controle do promotor P_{trc} . A construção resultante foi cotransformada com OP80-PCC7942_1594 em *E. coli* MG1655 e as células foram cultivadas a 37°C em meio M9 mínimo suplementado com 100 µg/mL de espectinomicina e 100 µg/mL de carbenicilina. As células foram cultivadas e extraídas como no Exemplo 3 e analisadas por GC-MS, conforme descrito no Exemplo 26.

[316] Como mostrado na Fig. 23, células de *E. coli* cotransformadas com os vetores de transporte *S. elongatus* PCC7942J594 e *Cyanothece* sp. PCC7425 Cyan7425_2986 produziram os mesmos aldeídos graxos e álcoois graxos como no Exemplo 3, mas também tridecano, pentadecano, pentadeceno e heptadeceno. Esse resultado indica que Cyan7425_2986 em *E.*

coli converte tetradecanal, hexadecenal, hexadecanal e octadecenal para tridecano, pentadecano, pentadeceno e heptadeceno, respectivamente, e, portanto, é uma descarbonilase ativa de aldeído graxo.

Exemplo 21: Produção de alcanos e alcenos em *E. coli* através da expressão heteróloga de *Prochlorococcus marinus* CCMP1986 PMM0533 e *Prochlorococcus mariunus* CCMP1986 PMM0532

[317] O DNA genômico codificando *P.mariunus* CCMP1986 PMM0533 (NP_892651; enzima suposta geradora de aldeído) (SEQ ID NO: 71) e *Prochlorococcus mariunus* CCMP1986 PMM0532 (NP_892650; descarbonilase suposta) (SEQ ID NO: 19) foram amplificados e clonados nos locais *NcoI* e *EcoRI* do vetor OP-80 e nos locais *NdeI* e *XhoI* do vetor OP-183, respectivamente. As construções resultantes foram transformadas em separado e cotransformadas em *E. coli* MG 1655 e as células foram cultivadas a 37°C em meio M9 mínimo suplementado com 100 µg/mL de espectinomicina e 100 µg/mL de carbenicilina. As células foram cultivadas e extraídas como no Exemplo 3 e analisadas por GC-MS, conforme descrito no Exemplo 26.

[318] Como mostrado na Fig. 24A, células de *E. coli* transformadas apenas o vetor de transporte *P.mariunus* CCMP1986 PMM0533 não produziram quaisquer aldeídos graxos ou álcoois graxos. No entanto, as células de *E. coli* cotransformados com vetores de transporte PMM0533 e PMM0532 produziram hexadecanol, pentadecano e heptadeceno (Fig. 24B). Esse resultado indica que PMM0533 só fornece substratos de aldeído graxo para a reação de descarbonilação quando interage com uma descarbonilase, como PMM0532.

Exemplo 22: Produção de alcanos e alcenos em uma linhagem *E. coli* produtora de acil-CoA graxo através da expressão heteróloga de *Synechococcus elongatus* PCC7942 orf1594 e

Acaryochloris marina MBIC11017 AM1 4041

[319] O DNA genômico codificando *Acaryochloris marina* MBIC11017 AM1_4041 (YP_001518340; descarbonilase suposta de aldeído graxo) (SEQ ID NO: 9) foi sintetizado e clonado nos locais *NdeI* e *XhoI* do vetor OP-183 (pACYC derivados) sob o controle do promotor P_{trc} . A construção resultante foi cotransformada com OP80-PCC7942_1594 em *E. coli* MG1655 $\Delta fadE lacZ::P_{trc}$ 'tesA-fadD. Esta linhagem expressa uma versão citoplasmática da tioesterase *E. coli*, 'TesA, e a sintetase acil-CoA *E. coli*, FADD, sob o controle do promotor P_{trc} e, portanto, produz acil-CoAs graxos. As células foram cultivadas a 37°C em meio M9 mínimo suplementado com 100 µg/mL de espectinomicina e 100 µg/mL de carbenicilina. As células foram cultivadas e extraídas como no Exemplo 3 e analisadas por GC-MS, conforme descrito no Exemplo 1.

[320] Como mostrado na Fig. 25, essas células *E. coli* cotransformadas com *S. elongatus* PCC7942_1594 e *A. marina* MBIC11017 AM1_4041 também produziram alcanos e álcoois graxos. Esse resultado indica que *S. elongatus* PCC7942_1594 é capaz de usar acil-CoA como substrato para a produção de hexadecanal, hexadecenal e octadecenal, que é então convertido em pentadecano, pentadeceno e heptadeceno, respectivamente, por *A. marina* MBIC11017 AM1_4041.

Exemplo 23: Produção de alcanos e alcenos em uma linhagem *E. coli* produtora de acil-CoA graxo através da expressão heteróloga de *Synechocystis* sp. PCC6803 s110209 e *Synechocystis* sp. PCC6803 s110208

[321] O DNA genômico codificando *Synechocystis* sp. PCC6803 s110208 (NP_442147; descarbonilase suposta de aldeído graxo) (SEQ ID NO: 3) foi sintetizado e clonado nos locais *NdeI* e *XhoI* do vetor OP-183 (pACYC derivados) sob o

controle do promotor P_{trc} . O DNA genômico codificando *Synechocystis* sp. PCC6803 sll0209 (NP_442146; acil-ACP redutase) (SEQ ID NO: 67) foi sintetizado e clonado nos locais *Nco*I e *Eco*RI do vetor OP-183 (pACYC derivados) sob o controle do promotor. P_{trc} . As construções resultantes foram cotransformadas com a *E. coli* MG1655 $\Delta fadE lacZ::P_{trc}$ 'tesA-*fadD*. Esta linhagem expressa uma versão citoplasmática da tioesterase *E. coli*, 'TesA, e a sintetase acil-CoA *E. coli*, *FADD*, sob o controle do promotor P_{trc} , portanto, produz acil-CoAs graxos. As células foram cultivadas a 37°C em meio M9 mínimo suplementado com 100 µg/mL de espectinomicina e 100 µg/mL de carbenicilina. As células foram cultivadas e extraídas como no Exemplo 3 e analisadas por GC-MS, conforme descrito no Exemplo 26.

[322] Como mostrado na Fig. 26, essas células de *E. coli* transformadas com *Synechocystis* sp. PCC6803 sll0209 não produziram quaisquer aldeídos graxos ou álcoois graxos. No entanto, quando cotransformadas com *Synechocystis* sp. PCC6803 sll0208 e sll0209, elas produziram alcanos, aldeídos graxos e álcoois graxos. Esse resultado indica que *Synechocystis* sp. PCC6803 sll0209 é capaz de usar acil-CoA como substrato para a produção de aldeídos graxos como hexadecanal, tetradecanal e octadecenal, mas só quando coexpressado com uma descarbonilase de aldeído graxo. Os aldeídos graxos, aparentemente, são ainda mais reduzidos para os álcoois graxos correspondentes, tetradecanol, hexadecanol e octadecenol, por atividade endógena de *E. coli* aldeído redutase. Neste experimento, octadecenal foi convertido em heptadeceno por *Synechocystis* sp. PCC6803 sll0208.

Exemplo 24: Produção de alcanos e alcenos em uma linhagem *E. coli* produtora de aldeído através da expressão heteróloga de *Nostoc punctiforme* PCC73102 Npun02004178 e vários de seus homólogos

[323] O DNA genômico codificando *Nostoc punctiforme* PCC73102 Npun02004178 (ZP_00108838; descarbonilase suposta de aldeído graxo) (SEQ ID NO: 5) foi amplificado e clonado nos locais *NdeI* e *XhoI* do vetor OP-183 (pACYC derivados) sob o controle do promotor P_{trc} . O DNA genômico codificando *Mycobacterium smegmatis* de linhagem MC2155 orf MSMEG_5739 (YP_889972, redutase suposta de ácido carboxílico) (SEQ ID NO: 85) foi amplificado e clonado nos locais *NcoI* e *EcoRI* do vetor OP-180 (pCL1920 derivados) sob o controle do promotor P_{trc} . As duas construções resultantes foram cotransformadas em *E. coli* MG 1655 $\Delta fadD lacZ::P_{trc}$ -*'tesA*. Nesta linhagem, aldeídos graxos foram fornecidos pela MSMEG_5739, que reduz ácidos graxos livres (formados pela ação de *'TesA*) para aldeídos graxos. As células foram cultivadas a 37°C em meio M9 mínimo suplementado com 100 µg/mL de espectinomicina e 100 µg/mL de carbenicilina. As células foram cultivadas e extraídas como no Exemplo 3 e analisadas por GC-MS, conforme descrito no Exemplo 1.

[324] Como mostrado na Fig. 27, essas células de *E. coli* cotransformadas com os vetores de transporte *N. punctiforme* PCC73102 Npun02004178 e *M. smegmatis* linhagem MC2 155 MSMEG_5739 produziram pentadeceno, tridecano e pentadecano. Esse resultado indica que Npun02004178 em *E. coli* converte tetradecanal, hexadecenal e hexadecanal fornecidos pela MSMEG_5739 redutase de ácido carboxílico para tridecano, pentadeceno e pentadecano. Como mostrado na figura. Como

Mostrado Na Fig. 28, no mesmo modelo experimental, as seguintes descarbonilases de aldeídos graxos também converteram os aldeídos graxos fornecidos por MSMEG_5739 para os alcanos correspondentes quando expressados em *E. coli* MG1655 $\Delta fadD$ *lacZ::Ptrc-tesA*: *Nostoc* sp. PCC7210 alr5283 (SEQ ID NO: 7), *P. mariunus* CCMP1986 PMM0532 (SEQ ID NO: 19), *G. violaceus* PCC7421 gll3146 (SEQ ID NO: 15), *Synechococcus* sp. RS9917_09941 (SEQ ID NO: 23), *Synechococcus* sp. RS9917_12945 (SEQ ID NO: 25), e *A. marina* MBIC11017 AM1_4041 (SEQ ID NO: 9).

Exemplo 25: Descarbonilases cianobacteriais de aldeídos graxos pertencem à classe das proteínas de diferro não heme. Mutagênese direcionada de local de histidinas conservadas para fenilalaninas em *Nostoc punctiforme* PCC73102 Npun02004178 não abole sua função catalítica

[325] Conforme discutido no Exemplo 13, a proteína hipotética PMT1231 a partir de *Prochlorococcus marinus* MIT9313 (SEQ ID NO: 18) é uma descarboxylase ativa de aldeído graxo. Com base na estrutura tridimensional da PMT1231, que está disponível na resolução 1,8 Å (pdb2OC5A) (ver Fig. 29B), descarbonilases cianobacteriais de aldeído graxo têm semelhança estrutural com proteínas diferro não heme, em especial com as proteínas de classe I ribonuclease redutase subunidade β , RNR β (RIGGS e STUBBE-GELASCO, TIBS 1998, vol 23, p. 438) (ver Fig. 29A). As classes Ia e Ib RNR β contêm um radical tirosil diférrico que medeia a atividade catalítica de RNR α (redução de ribonucleotídeos para desoxirribonucleotídeos). Em *E. coli* RNR β , esta tirosina está na posição 122 e está próximo a uma das moléculas de local ativo de ferro. O alinhamento estrutural mostrou que PMT1231 continha uma fenilalanina na mesma posição que RNR β

tyr122, sugerindo um mecanismo catalítico diferente para descarbonilases cianobacteriais de aldeídos graxos. No entanto, um alinhamento de todos as descarbonilases mostrou que dois resíduos de tirosina foram completamente conservados em todas as sequências, e tyr135 e tyr138 com respeito à PMT1231, com tyr135 estandado na proximidade (5,5 Å) para uma das moléculas de local ativo de ferro (ver Fig. 29C). Para verificar se um dos dois resíduos de tirosina conservado está envolvido no mecanismo catalítico de descarbonilases cianobacteriais de aldeídos graxos, estes resíduos foram substituídos por fenilalanina em Npun02004178 (tyr123 e tyr126), como segue.

[326] O DNA genômico codificando *S. elongatus* PCC7942 ORF1594 (SEQ ID NO: 65) foi clonado nos locais *NcoI* e *EcoRI* do vetor OP-80 (pCL1920 derivados) sob o controle do promotor P_{trc} . O DNA genômico codificando de *N. punctiforme* PCC73102 Npun02004178 (SEQ ID NO: 5) também foi clonado nos locais *NdeI* e *XhoI* do vetor OP-183 (pACYC177 derivados), sob o controle do promotor P_{trc} . Esta última construção foi utilizada como modelo para introduzir uma mutação nas posições 123 e 126 da proteína descarbonilase, mudando a tirosina para fenilalaninas utilizando os iniciadores gttttgcgatcgagcatttaacatttacatccccgttgccgacg e gttttgcgatcgagcatataacattttcatccccgttgccgacg, respectivamente. As construções resultantes foram, então, transformadas em *E. coli* MG1655. As células foram cultivadas a 37°C em meio M9 mínimo acrescido de glicose 1% (p/v) e 100 µg/mL de carbenicilina e espectinomicina. As células foram cultivadas e extraídas como no Exemplo 3.

[327] Como mostrado na Fig. 30, as duas variantes de

proteína Npun02004178 Tyr a Phe foram ativadas e produziram alcanos quando coexpressadas com *S. elongatus* PCC7942 ORF1594. Esse resultado indica que em contraste com as proteínas de classes Ia e Ib RNR β , o mecanismo catalítico de descarboxilases de aldeído graxo não envolve um radical tirosil.

Exemplo 26: Caracterização bioquímica de *Nostoc punctiforme* PCC73102 Npun02004178

[328] O DNA genômico codificando *N. punctiforme* PCC73102 Npun02004178 (SEQ ID NO: 5) foi clonado nos locais *NdeI* e *XhoI* do vetor pET-15b sob o controle do promotor T7. A proteína resultante Npun02004178 continha uma N-terminal His-tag. Uma linhagem de *E. coli* BL21 (DE3) (Invitrogen) foi transformada com o plasmídeo através de técnicas de transformação química de rotina. A expressão da proteína foi realizada pela primeira inoculação de uma colônia da linhagem de *E. coli* em 5 mL de meio LB suplementado com 100 mg/L de carbenicilina e agitada durante a noite a 37°C para produzir uma cultura *starter*. Essa cultura *starter* foi usada para inocular 0,5 L de LB meio suplementado com 100 mg/L de carbenecillina. A cultura foi agitada a 37°C até um valor de OD₆₀₀ de 0,8 ser alcançado e, em seguida, foi adicionado IPTG para uma concentração final de 1 mM. A cultura foi então agitada a 37°C por aproximadamente 3h adicionais. A cultura foi então centrifugada a 3.700 rpm por 20 min a 4°C. O sedimento foi então ressuspensão em 10 mL de tamponador contendo 100 mM de tamponador fosfato de sódio a pH 7,2 suplementado com bactérias ProteaseArrest (GBiosciences). As células foram então sonicadas a 12 W em gelo por 9 s com 1,5 s de sonicação seguido de 1,5 s de repouso. Este procedimento

foi repetido cinco vezes, com intervalos de um minuto entre cada ciclo de sonicação. O extrato de célula livre foi centrifugado a 10.000 rpm por 30 min a 4°C. 5 mL de Ni-NTA (Qiagen) foi adicionado ao sobrenadante e a mistura foi suavemente mexido a 4°C. A suspensão foi passada através de uma coluna para remover a resina a partir do lisado. A resina foi lavada com 30 mL de tamponador contendo 100 mM de tamponador fosfato de sódio a pH 7,2 acrescido de 30 mM de imidazol. Finalmente, a proteína foi eluída com 10 mL de 100 mM de tamponador fosfato de sódio em pH 7,2 acrescido de 250 mM de imidazol. A solução de proteína foi dialisada com 200 volumes de 100 mM de tamponador fosfato de sódio em pH 7,2 com 20% de glicerol. A concentração de proteínas foi determinada utilizando o ensaio de Bradford (BioRad). 5,6 mg / mL de proteína Npun02004178 foram obtidos.

[329] Para sintetizar octadecanal para a reação de descarboxilase, 500 mg de octadecanol (Sigma) foram dissolvidos em 25 mL de diclorometano. Em seguida, 200 mg de clorocromato de piridínio (TCI América) foram adicionados à solução e agitados durante a noite. A mistura de reação foi seca sob vácuo para remover o diclorometano. Os demais produtos foram ressuspensos em hexano e filtrados através de papel de filtro Whatman. O filtrado foi então seco sob vácuo e ressuspenso em 5 mL de hexano e purificado por cromatografia flash em sílica. A mistura foi carregada para a coluna de alimentação por gravidade em hexano e em seguida lavada com dois volumes da coluna de hexano. O octadecanal foi então eluído com uma mistura de 8:1 de hexano e acetato de etila. Frações contendo octadecanal foram agrupadas e analisadas através dos métodos CG/MS descritos abaixo. O

produto final foi 95% puro, como determinado por este método.

[330] Para testar a proteína Npun02004178 para a atividade de descarbonilação, os seguintes ensaios enzimáticos foram arranjados. Reações de 200 µL foram criadas em 100 mM de tamponador fosfato de sódio em pH 7.2 com os seguintes componentes em suas respectivas concentrações finais: 30 µM de proteína Npun02004178 purificada, 200 µM de octadecanal, 0,11 µg/mL de ferredoxina espinafre (Sigma), 0,05 unidades/mL de ferredoxina espinafre redutase (Sigma), e 1 mM NADPH (Sigma). Os controles negativos incluem a reação acima sem Npun02004178, a reação acima sem octadecanal, e a reação acima sem ferredoxina espinafre, ferredoxina redutase e NADPH. Cada reação foi incubada a 37°C por 2 horas antes de ser extraída com 100 µL de acetato de etila. As amostras foram analisadas por GC/MS utilizando-se os seguintes parâmetros: tempo de execução: 13,13 min; coluna: HP-5-MS Parte n° 19091S-433E (comprimento de 30 metros; ID: 0,25 milímetros de arco de abertura; filme: 0.25µM); injetar: 1 µl Agilent 6850 de entrada; entrada: 300 C sem quebra; gás transportador: hélio, fluxo: 1,3 mL/min, temperatura do forno: 75°C detido por 5 min, 320 a 40°C/min, 320 detido por 2 min; det: Agilent 5975B VL MSD; det. temp: 330°C; scan: 50-550 M/Z. O heptadecano da Sigma foi utilizado como uma referência autêntica para determinar o tempo de retenção do composto e o padrão de fragmentação.

[331] Como mostrado na Fig. 31, a conversão *in vitro* de octadecanal para heptadecano foi observada na presença de Npun02004178. O descarbonilação enzimática de octadecanal por Npun02004178 foi dependente da adição de ferredoxina espinafre reductase, ferredoxina e NADPH.

[332] Em seguida, foi determinado se ferredoxinas cianobacteriais e ferredoxinas redutases podem substituir as proteínas dos espinafres no ensaio *in vitro* de descarbonilase de aldeído graxo. Os seguintes quatro genes foram clonados separadamente nos locais *NdeI* e *XhoI* do pET-15b: *N. punctiforme* PCC73102 Npun02003626 (ZP_00109192, ferredoxina oxidoredutase petH sem o domínio vinculador n-terminal aloficocianina) (SEQ ID NO: 87), *N. punctiforme* PCC73102 Npun02001001 (ZP_00111633, ferredoxina 1) (SEQ ID NO: 89), *N. punctiforme* PCC73102 Npun02003530 (ZP_00109422, ferredoxina 2) (SEQ ID NO: 91) e *N. punctiforme* PCC73102 Npun02003123 (ZP_00109501, ferredoxina 3) (SEQ ID NO: 93). As quatro proteínas foram expressas e purificadas como descrito acima. 1 mg/mL de cada ferredoxina e 4 mg/mL da ferredoxina oxidoredutase foram obtidos. As três ferredoxinas cianobacteriais foram testadas com a ferredoxina oxidoredutase cianobacteriana usando-se o arranjo enzimático descrito anteriormente com as seguintes alterações. A concentração final da ferredoxina redutase foi de 60 µg/ml e as ferredoxinas estavam em 50 µg/mL. As reações enzimáticas foram extraídas por GC/MS utilizando-se os seguintes parâmetros: tempo de execução: 6,33 min; coluna: J&W 122-5711 DB-5ht (comprimento de 15 metros; ID: 0,25 milímetros arco de abertura; filme: 0,10µM); injetar : 1 µL Agilent 6850 de entrada; entrada: 300°C sem quebra; gás transportador: hélio; fluxo: 1,3 mL/min; temperatura do forno: 100°C mantida por 0,5 min, 260 a 30°C/min, 260 mantida por 0,5 min; det: Agilent 5975B VL MSD; det. temp: 230°C; scan: 50-550 M/Z.

[333] Como mostrado na Fig. 32, conversão *in vitro*

dependente de Npun02004178 de octadecanal para heptadecano foi observada na presença de NADPH e ferredoxina oxidorreductase cianobacterial e de qualquer uma das três ferredoxinas cianobacteriais.

Exemplo 27: Caracterização bioquímica de *Synechococcus elongatus* PCC7942 orf1594

[334] O DNA genômico codificando *S. elongatus* PCC7492 orf1594 (SEQ ID NO: 65) foi clonado nos locais *NcoI* e *XhoI* do vetor pET-28b sob o controle do promotor T7. A proteína resultante PCC7942_orf1594 continha um C-terminal His-tag. Uma linhagem de *E. coli* BL21 (DE3) (Invitrogen) foi transformada com o plasmídeo e a proteína PCC7942_orf1594 foi expressa e purificada como descrito no Exemplo 22. A solução de proteína foi armazenada no tamponador seguinte: fosfato de sódio 50 mM, pH 7,5, 100 mM de NaCl, 1 mM de THP, 10% de glicerol. A concentração de proteína foi determinada utilizando-se o ensaio de Bradford (BioRad). 2 mg/mL de proteína PCC7942_orf1594 foram obtidos.

[335] Para testar a proteína PCC7942_orf1594 para a atividade acil-ACP ou acil-CoA redutase, os seguintes ensaios de enzima foram arranjados. Reações de 100 µL foram colocados em 50 mM de tamponador Tris-HCl a pH 7,5 com os seguintes componentes em suas respectivas concentrações finais: 10 µM de proteína purificada PCC7942_orf1594, 0,01-1 mM de acil-CoA redutase ou acil-ACP, 2 mM de MgCl₂, 0,2-2 mM de NADPH. As reações foram incubadas por 1h a 37°C e foram interrompidas pela adição de 100 µL de acetato de etila (contendo 5 mg/L de 1-octadeceno como padrão interno). As amostras foram centrifugadas por 15 min e centrifugadas a velocidade máxima por 3 min para a separação de fase. 80 µL

da camada superior foram transferidos para frascos de vidro GC e analisados por GC/MS, conforme descrito no Exemplo 26. A quantidade de aldeído formado foi calculada com base no padrão interno.

[336] Como mostrado na Fig. 33, PCC7942_orf1594 foi capaz de reduzir octadecanoyl-CoA para octadecanal. A atividade redutase necessitou de cátions bivalentes tais como Mg^{2+} , Mn^{2+} ou Fe^{2+} e NADPH como doador de elétrons. NADH apoia a atividade redutase. PCC7942_orf1594 também foi capaz de reduzir octadecenoil-CoA e octadecenoil-ACP para octadecenal. Os valores de K_m para a redução da octadecanoil-CoA, octadecenoil-CoA e octadecenoil-ACP na presença de 2 mM de NADPH foram determinados como $45 \pm 20 \mu M$, $82 \pm 22 \mu M$ e $7,8 \pm 2 \mu M$, respectivamente. Esses resultados demonstram que PCC7942_orf1594, *in vitro*, reduz tanto acil-CoA quanto acil-ACP e que, aparentemente, a enzima tem maior afinidade para acil-ACPs em relação aos acil-CoAs. O valor de K_m para a NADPH na presença de 0,5 mM de octadecanoyl-CoA para PCC7942_orf1594 foi determinado como $400 \pm 80 \mu M$.

[337] Em seguida, a transferência de hidreto estereoespecífico de NADPH para um aldeído graxo catalisada por PCC7942_orf1594 foi examinada. Deutero-NADPH foi preparado de acordo com o protocolo a seguir. 5 mg de $NADP^+$ e 3,6 mg de D-glicose-1-d foram adicionados a 2,5 mL 50 mM de tamponador fosfato de sódio (pH 7,0). A produção enzimática de NADPH marcado foi iniciada pela adição de 5 unidades do desidrogenase de glicose a partir de qualquer *Bacillus megaterium* (USB Corporation) para a produção de R-(4- 2H)NADPH ou *Thermoplasma acidophilum* (Sigma) para a produção de S-(4- 2H)NADPH. A reação foi incubada por 15 min

a 37°C, centrifugada e filtrada usando uma centrífuga filtro 10 KDa MWCO Amicon Ultra (Millipore), congelada rapidamente em gelo seco e armazenada a -80°C.

[338] A reação no ensaio *in vitro* continha 50 mM de Tris-HCl (pH 7,5), 10 µM de proteína purificada PCC7942_orf1594, 1 mM de octadecanoyl-CoA redutase, 2 mM de MgCl₂, e 50 µL de deuterio-NADPH (preparado como descrito acima) em um volume total de 100 µL. Após uma incubação de 1 h, o produto da reação enzimática foi extraído e analisado conforme descrito acima. O aldeído graxo resultante detectado por GC/MS foi octadecanal (ver Fig. 34). Como a transferência de hidreto de NADPH é estereoespecífica, tanto *R*-(4-²H)NADPH e *S*-(4-²H)NADPH foram sintetizados. Octadecanal com uma unidade de massa a mais foi observado utilizando-se apenas o *S*-(4-²H)NADPH. O fato de que o aldeído graxo fora rotulado indica que o hidrogênio deuterado foi transferido do NADPH marcado para o aldeído graxo marcado. Isto demonstra que o NADPH é usado nesta reação enzimática e que a transferência de hidretos catalisadas por PCC7942_orf1594 é estereoespecífica.

Exemplo 28: Produção intracelular e extracelular de aldeídos graxos e álcoois graxos em *E. Coli* através de expressão heteróloga de *Synechococcus elongatus* PCC7942 orf1594

[339] O DNA genômico codificando *Synechococcus elongatus* PCC7942 orf1594 (YP_400611; acil-ACP redutase) (SEQ ID NO: 65) foi amplificado e clonado nos locais *Nco*I e *Eco*RI do vetor OP-80 (pCL1920 derivados) sob o controle do promotor P_{trc}. A construção resultante foi cotransformada em *E. coli* MG1655 Δ*fadE* e as células foram cultivadas a 37°C em 15 mL de meio Che-9 mínimo, com 3% (p/v) de glicose como fonte de

carbono e suplementado com 100 µg/mL de espectinomicina e carbenicilina, respectivamente. Quando a cultura atingiu OD₆₀₀ de 0,8-1,0, ela foi induzida com 1mM de IPTG e as células foram cultivadas por mais 24-48 h, a 37°C. O meio Che-9 mínimo é definido como: 6g/L de Na₂HPO₄, 3 g/L de KH₂PO₄, 0,5 g/L de NaCl, 2 g/L de NH₄Cl, 0,25 g/L de MgSO₄ x 7 H₂O, 11 mg/L de CaCl₂, 27 mg/L de Fe₃Cl x 6 H₂O, 2 mg/L de ZnCl x 4 H₂O, 2 mg/L de Na₂MoO₄ x 2 H₂O, 1,9 mg/L de CuSO₄ x 5 H₂O, 0,5 mg/L de H₃BO₃, 1 mg/L de tiamina, 200 mM Bis-Tris (pH 7,25) e 0,1% (v/v) de Triton-X100. Quando a cultura atingiu OD₆₀₀ de 1,0-1,2, ela foi induzida com 1 mM de IPTG e as células foram deixadas crescer por mais 40 horas a 37°C. Células de 0,5 mL de cultura foram extraídas com 0,5 mL de acetato de etila para a produção de hidrocarbonetos totais, conforme descrito no Exemplo 26. Além disso, as células e o sobrenadante foram separados por centrifugação (4.000 g, RT por 10 min) e extraídas em separado.

[340] A cultura produziu 620 mg/L de aldeídos graxos (tetradecanal, heptadecanal, heptadecenal e octadecenal) e 1.670 mg/L de álcoois graxos (dodecanol, tetradecenol, tetradecanol, heptadecenol, heptadecanol e octadecenol). A Fig. 35 mostra o cromatograma do sobrenadante extraído. Foi determinado que 73% dos aldeídos graxos e dos álcoois graxos estavam no sobrenadante livre de célula.

Exemplo 29: Produção intracelular e extracelular de alcanos e alcenos em *E. Coli* através de expressão heteróloga de *Synechococcus elongatus* PCC7942 orf1594 e *Nostoc punctiforme* PCC73102 Npun02004178

[341] O DNA genômico codificando *Synechococcus elongatus* PCC7942 orf1594 (YP_400611; acil-ACP redutase) (SEQ ID NO:

65) foi amplificado e clonado nos locais *NcoI* e *EcoRI* do vetor OP-80 (pCL1920 derivados) sob o controle do promotor P_{trc} . O DNA genômico codificando *Nostoc punctiforme* PCC73102 Npun02004178 (ZP_00108838; descarbonilase de aldeído graxo) (SEQ ID NO: 5) foi amplificado e clonado nos locais *NdeI* e *XhoI* do vetor OP-183 (pACYC derivados) sob o controle do promotor P_{trc} . As construções resultantes foram cotransformadas em *E. coli* MG1655 $\Delta fadE$ e as células foram cultivadas a 37°C em 15 mL de meio Che9 mínimo, com 3% (p/v) de glicose como fonte de carbono e suplementado com 100 µg/mL de espectinomicina e carbenicilina, respectivamente. As células foram cultivadas, separadas do caldo extraído, e analisadas como descrito no Exemplo 28.

[342] A cultura produziu 323 mg/L de alcanos e alcenos (tridecano, pentadecano, pentadeceno e heptadeceno), 367 mg/L de aldeídos graxos (tetradecanal, heptadecenal, heptadecanal e octadecenal) e 819 mg/L de álcoois graxos (tetradecanol, heptadecenol, heptadecanol e octadecenol). A Fig. 36 mostra o cromatograma do sobrenadante extraído. Foi determinado que 86% dos alcanos, alcenos, aldeídos graxos e álcoois graxos estavam no sobrenadante livre de células.

Exemplo 30: Produção alcanos e alcenos em *E. Coli* através de expressão heteróloga de *Nostoc sp.* PCC7210 alr5284 e *Nostoc sp.* PCC7210 alr5283

[343] O DNA genômico codificando *Nostoc sp.* PCC7210 alr5284 (NP_489324; enzima suposta geradora de aldeídos) (SEQ ID NO: 81) foi amplificado e clonado nos locais *NcoI* e *EcoRI* do vetor OP-80 (pCL1920 derivados) sob o controle do promotor P_{trc} . O DNA genômico codificando *Nostoc sp.* PCC7210 alr5283 (NP_489323; descarbonilase suposta) (SEQ ID NO: 7)

foi amplificado e clonado nos locais *NdeI* e *XhoI* do vetor OP-183 (pACYC derivados) sob o controle do promotor P_{trc} . As construções resultantes foram cotransformadas em *E. coli* MG1655 e as células foram cultivadas a 37°C em 15 mL de meio Che9 mínimo, com 3% (p/v) de glicose como fonte de carbono e suplementado com 100 µg/mL de espectinomicina e carbenicilina, respectivamente (como descrito no Exemplo 28). Células de 0,5 mL de cultura foram extraídas e analisadas conforme descrito no Exemplo 3 e analisadas por GC-MS, conforme descrito no Exemplo 26.

[344] Como mostrado na Fig. 37, as células de *E. coli* cotransformadas com os vetores de transporte *Nostoc* sp. PCC7210 alr5284 e *Nostoc* sp. PCC7210 alr5283 produziram tridecano, pentadeceno, tetradecanol, pentadecano e hexadecanol. Esse resultado indica que a coexpressão de *Nostoc* sp. PCC7210 alr5284 e alr5283 é suficiente para *E. coli* produzir álcoois graxos, alcanos e alcenos.

OUTRAS CONCRETIZAÇÕES

[345] É preciso entender que enquanto a invenção tem sido descrita em conjunto com a descrição detalhada da mesma, a descrição acima é destinado a ilustrar e não limitar o âmbito de aplicação da invenção, que é definida pelo alcance das reivindicações anexas. Outros aspectos, as vantagens e as modificações estão dentro do escopo das reivindicações a seguir.

REIVINDICAÇÕES

1. Microorganismo geneticamente modificado para produção de um hidrocarboneto **caracterizado** pelo fato de compreender uma sequência de polinucleotídeo exógena selecionada do grupo que consiste em SEQ ID NOs: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33 e 35 codificando um polipeptídeo biossintético de alcano ou alceno selecionado a partir do grupo que consiste em SEQ ID NOs: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34 e 36, em que o microorganismo geneticamente modificado é uma bactéria, levedura, fungo, fungo filamentosos ou alga.

2. Microorganismo geneticamente modificado, de acordo com a reivindicação 1, **caracterizado** pelo fato do referido microorganismo geneticamente modificado ser geneticamente modificado para produzir um hidrocarboneto selecionado do grupo que consiste em um alcano e um alceno, a partir de um aldeído.

3. Microorganismo geneticamente modificado, de acordo com a reivindicação 2, **caracterizado** pelo fato do referido aldeído ser produzido por um polipeptídeo biossintético de aldeído que é expresso no microorganismo geneticamente modificado.

4. Microorganismo geneticamente modificado, de acordo com a reivindicação 3, **caracterizado** pelo fato do referido polipeptídeo biossintético de aldeído ser uma acil-ACP redutase (AAR) que catalisa a conversão de uma acil-ACP em um aldeído graxo.

5. Microorganismo geneticamente modificado, de acordo com a reivindicação 4, **caracterizado** pelo fato da

referida AAR ser selecionada a partir do grupo que consiste nas SEQ ID NOs: 66, 68, 70, 72, 74, 76, 78, 80 e 82 e codificado por um polinucleotídeo selecionado do grupo que consiste em SEQ ID NOs: 65, 67, 69, 71, 73, 75, 77, 79 e 81.

6. Microorganismo geneticamente modificado, de acordo com a reivindicação 3, **caracterizado** pelo fato do referido polipeptídeo biossintético de aldeído ser uma redutase de ácido carboxílico (CAR) que catalisa a conversão de um ácido graxo em um aldeído graxo.

7. Microorganismo geneticamente modificado, de acordo com a reivindicação 6, **caracterizado** pelo fato da referida CAR compreender a SEQ ID NO. 86 e ser codificado pela SEQ ID NO. 85.

8. Microorganismo geneticamente modificado, de acordo com a reivindicação 1, **caracterizado** pelo fato do referido polipeptídeo biossintético de alceno ou alceno catalisar a conversão de um aldeído graxo em um alceno ou alceno.

9. Microorganismo geneticamente modificado, de acordo com a reivindicação 8, **caracterizado** pelo fato do referido microorganismo geneticamente modificado compreender ainda uma tioesterase que catalisa a conversão de um acil-ACP em um ácido graxo.

10. Microorganismo geneticamente modificado, de acordo com a reivindicação 1, **caracterizado** pelo fato do microorganismo ser uma bactéria.

11. Microorganismo geneticamente modificado, de acordo com a reivindicação 10, **caracterizado** pelo fato da bactéria ser *E. coli*.

12. Microorganismo geneticamente modificado, de acordo com a reivindicação 11, **caracterizado** pelo fato da bactéria ser cianobactéria.

13. Método de produção de um alcano ou alceno a partir de um microrganismo geneticamente modificado **caracterizado** pelo fato de compreender

transformar um microrganismo para expressar uma sequência polinucleotídica exógena selecionada a partir do grupo que consiste em SEQ ID NOs: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33 e 35 codificando um polipeptídeo biossintético de alcano ou alceno selecionado do grupo que consiste em SEQ ID NOs: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34 e 36, para produzir o microrganismo geneticamente modificado;

cultivar o microrganismo geneticamente modificado em um meio de cultura contendo uma fonte de carbono de carboidrato; e

isolar o alcano ou alceno do microrganismo geneticamente modificado;

em que o microrganismo geneticamente modificado é uma bactéria, levedura, fungo, fungo filamentoso ou alga.

14. Método, de acordo com a reivindicação 13, **caracterizado** pelo fato de que o alcano ou alceno é um alcano ou alcano C₁₃-C₂₁.

15. Método, de acordo com a reivindicação 13 ou 14, **caracterizado** pelo fato de que o alcano ou alceno é selecionado do grupo que consiste em tridecano, metiltridecano, nonadecano, metilnonadecano, heptadecano, metilheptadecano, pentadecano, metilpentadecano,

pentadeceno, heptadeceno, metilpentadeceno e metilheptadeceno.

16. Cultura de células **caracterizada** por compreender um microrganismo geneticamente modificado, em que

(b) um polinucleotídeo biossintético de alceno ou alceno selecionado a partir do grupo que consiste em SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33 e 35 codificando um polipeptídeo biossintético de alceno ou alceno selecionado do grupo que consiste em SEQ ID NOs: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34 e 36 são incorporados e superexpressos no referido microrganismo geneticamente modificado e um alceno ou alceno é produzido na cultura de células quando o referido microrganismo geneticamente modificado é cultivado em um meio de cultura contendo uma fonte de carbono de carboidrato para expressar o referido polipeptídeo biossintético de alceno ou alceno, e

em que o microrganismo geneticamente modificado é uma bactéria, levedura, fungo, fungo filamentoso ou alga.

17. Cultura de células, de acordo com a reivindicação 16, **caracterizada** pelo fato de que o alceno ou alceno é um alceno ou alceno C₁₃-C₂₁.

18. Cultura de células, de acordo com a reivindicação 16 ou 17, **caracterizada** pelo fato de que o alceno ou alceno é selecionado a partir do grupo que consiste em tridecano, metiltridecano, nonadecano, metilnonadecano, heptadecano, metilheptadecano, pentadecano, metilpentadecano, pentadeceno, heptadeceno, metilpentadeceno e metilheptadeceno.

19. Cultura de células, de acordo com qualquer uma das reivindicações 16 a 18, **caracterizada** pelo fato de compreender um alceno tendo um $\delta^{13}\text{C}$ de -15,4 ou maior.

20. Cultura de células, de acordo com qualquer uma das reivindicações 16 a 19, **caracterizada** pelo fato de compreender um alceno tendo um $f_M^{14}\text{C}$ de pelo menos 1,003.

FIGURA 1

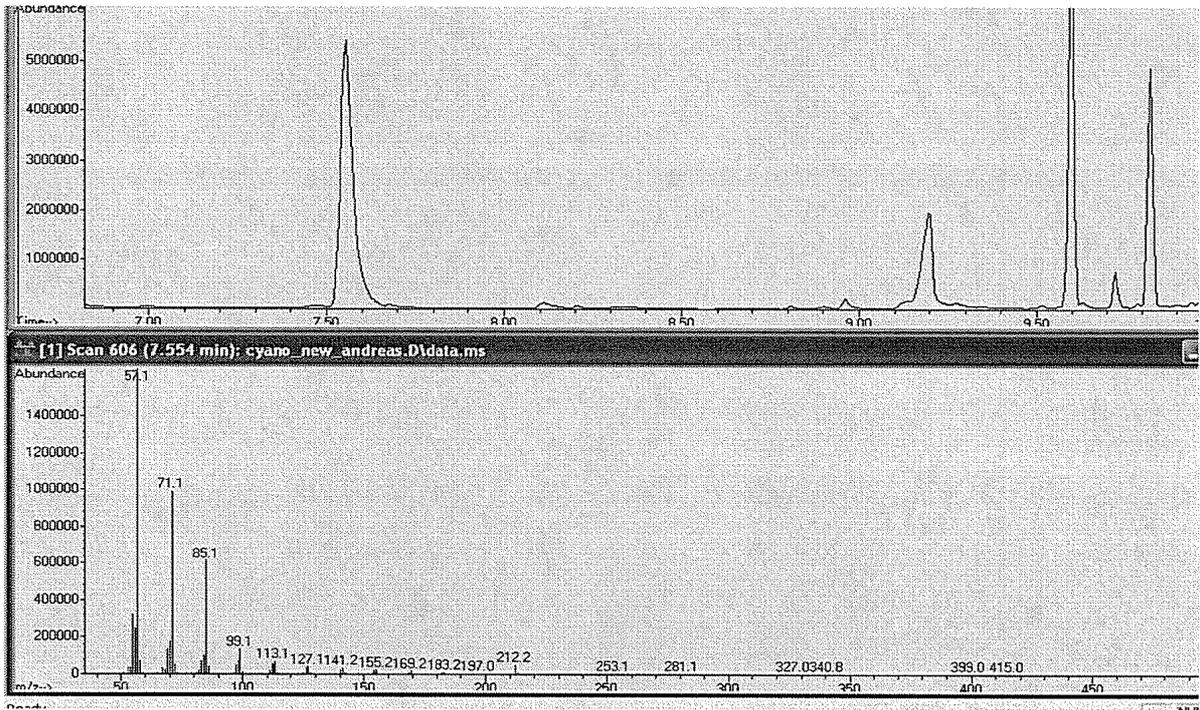


FIGURA 2

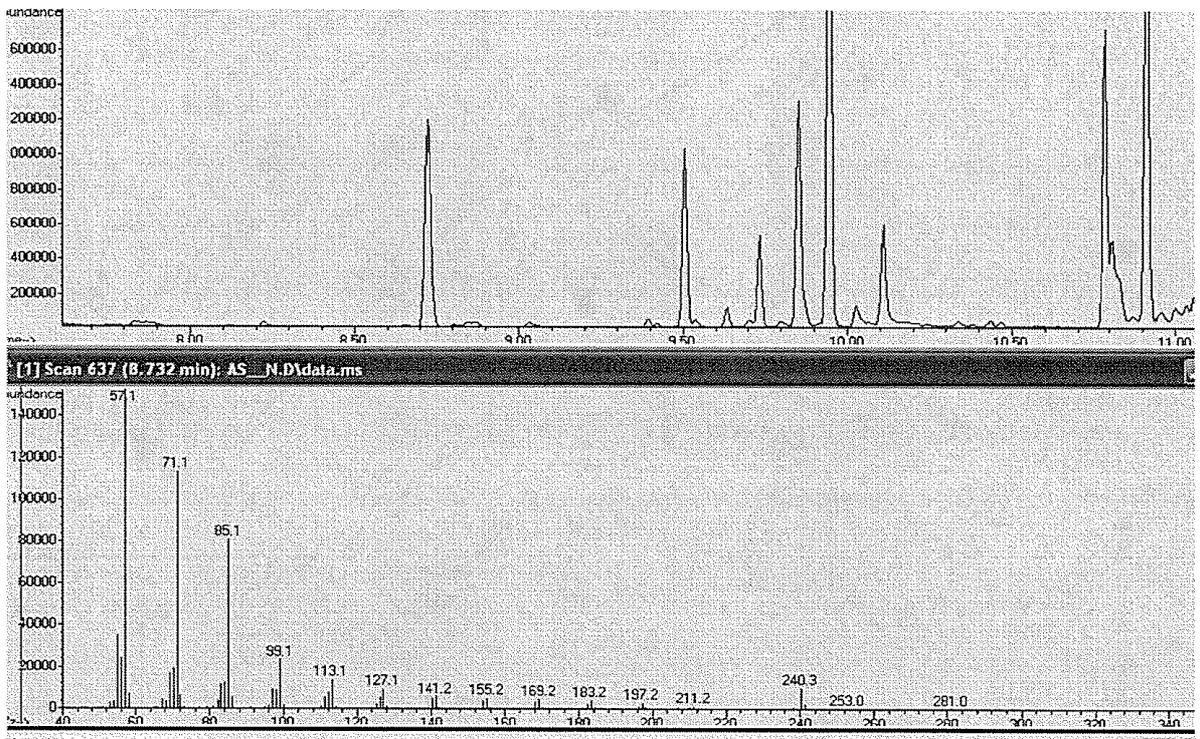


FIGURA 3

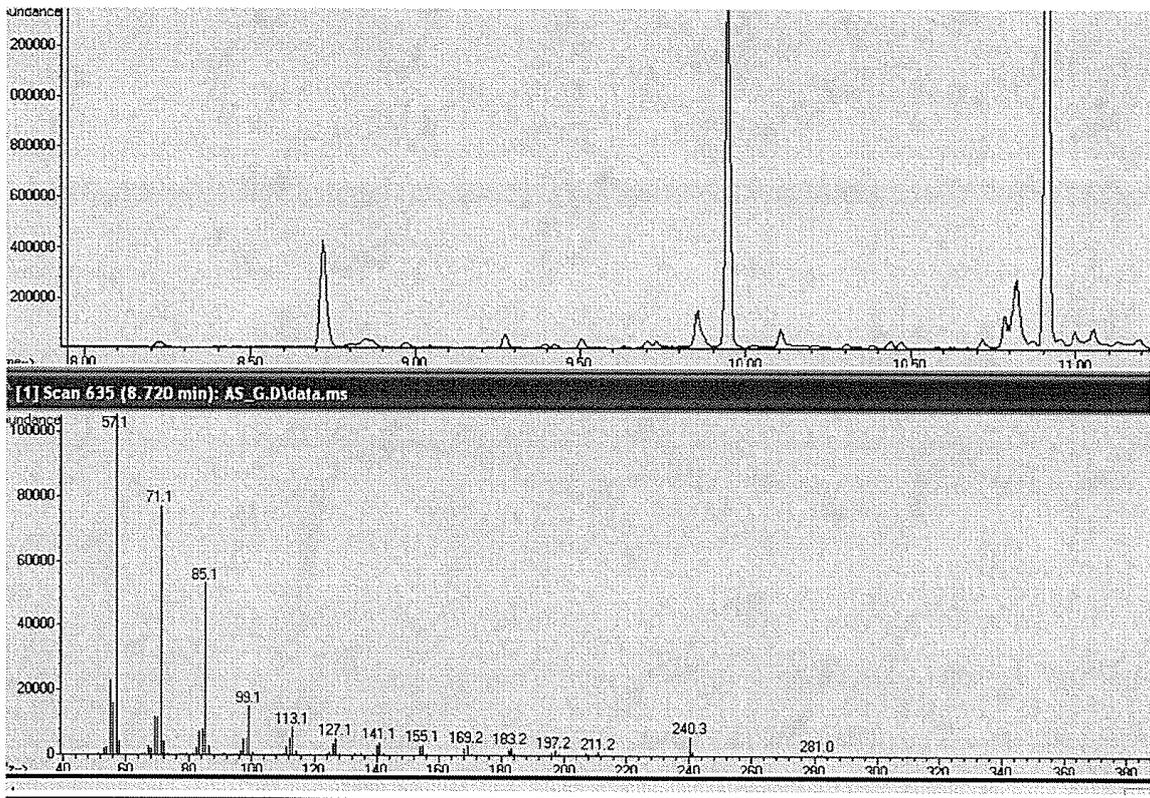


FIGURA 4

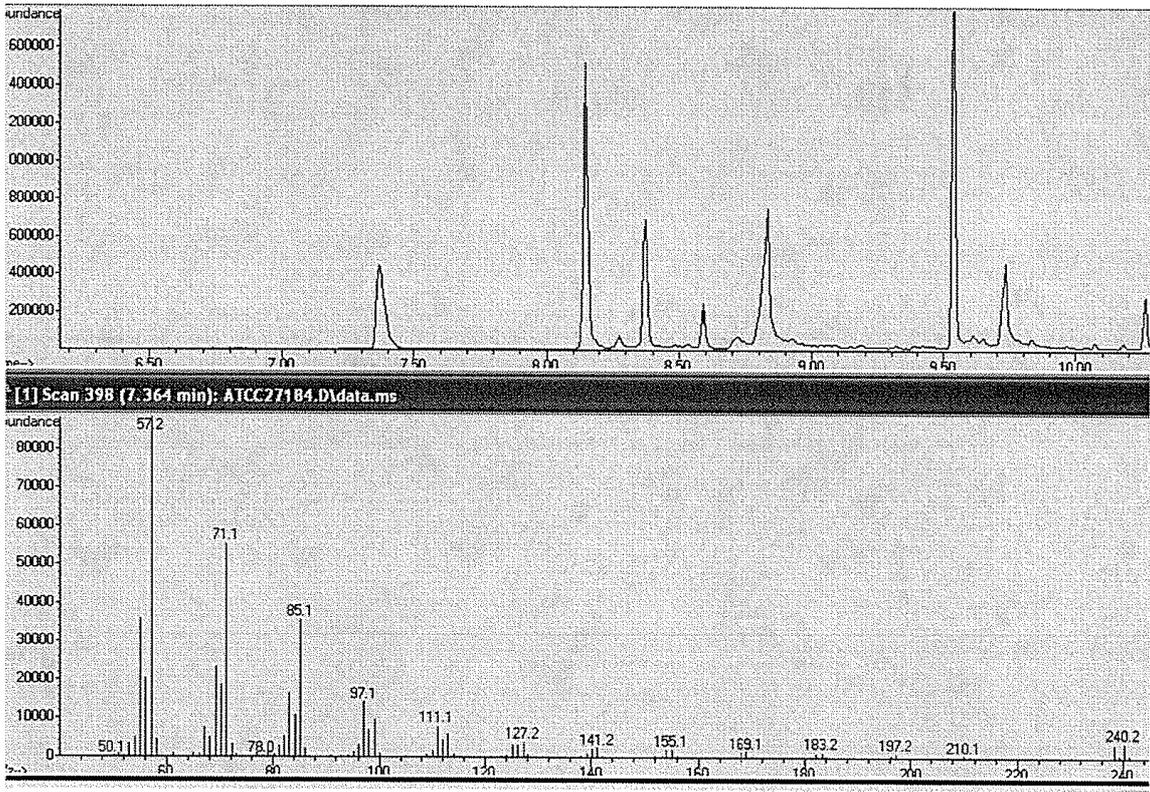


FIGURA 5

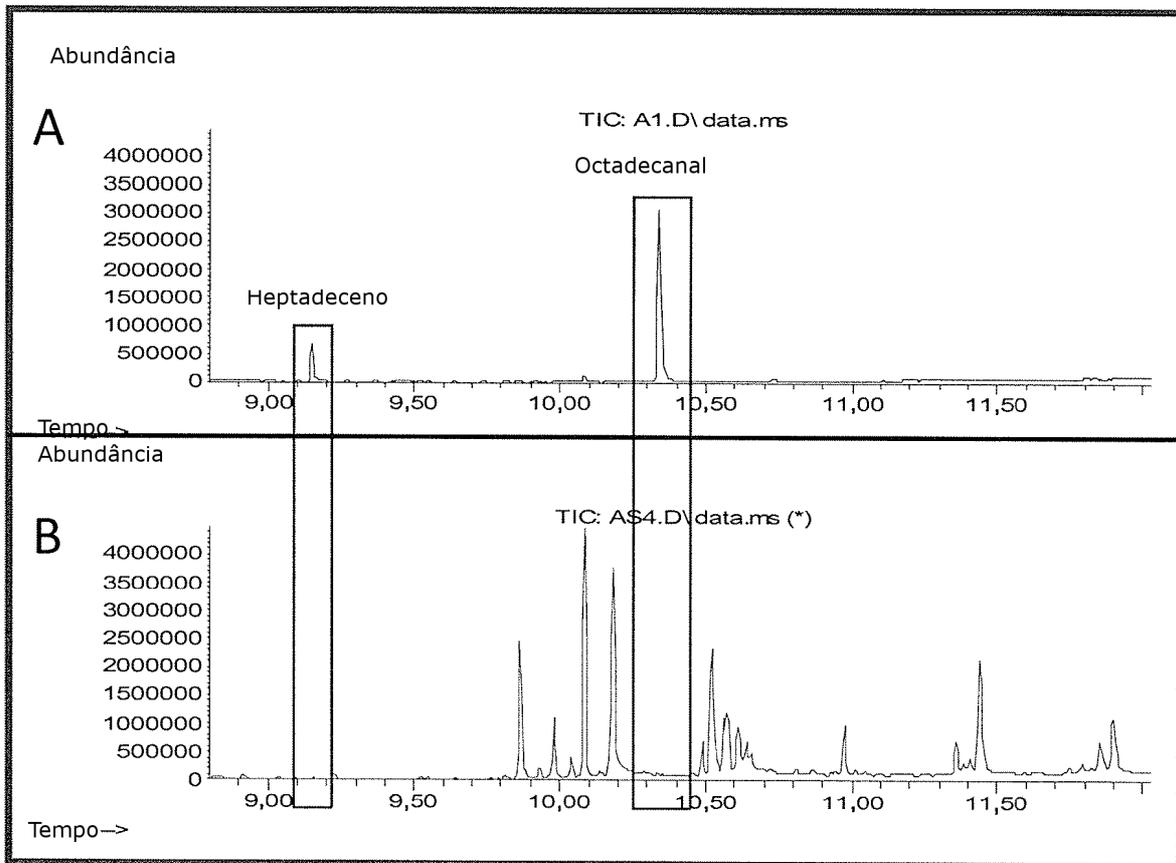


FIGURA 6

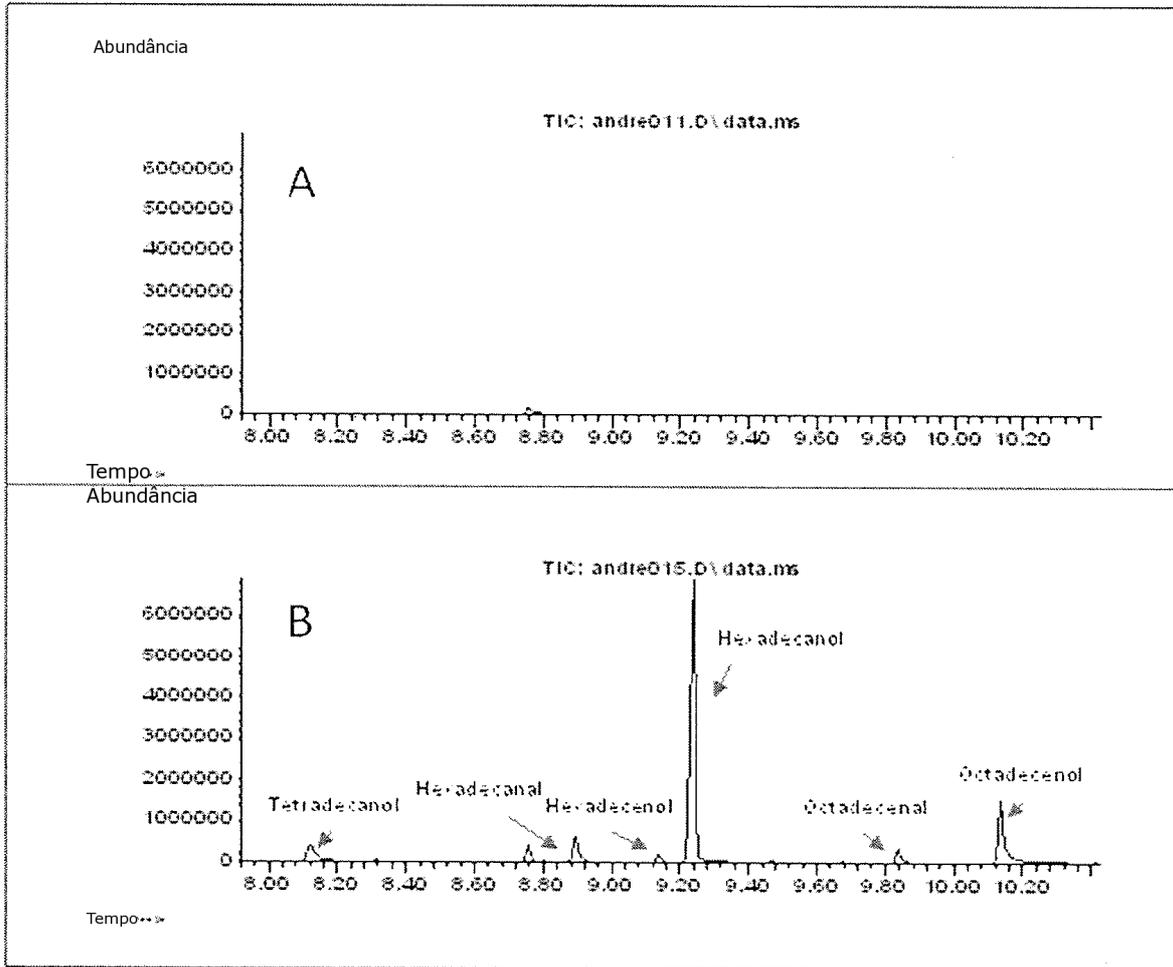


FIGURA 7

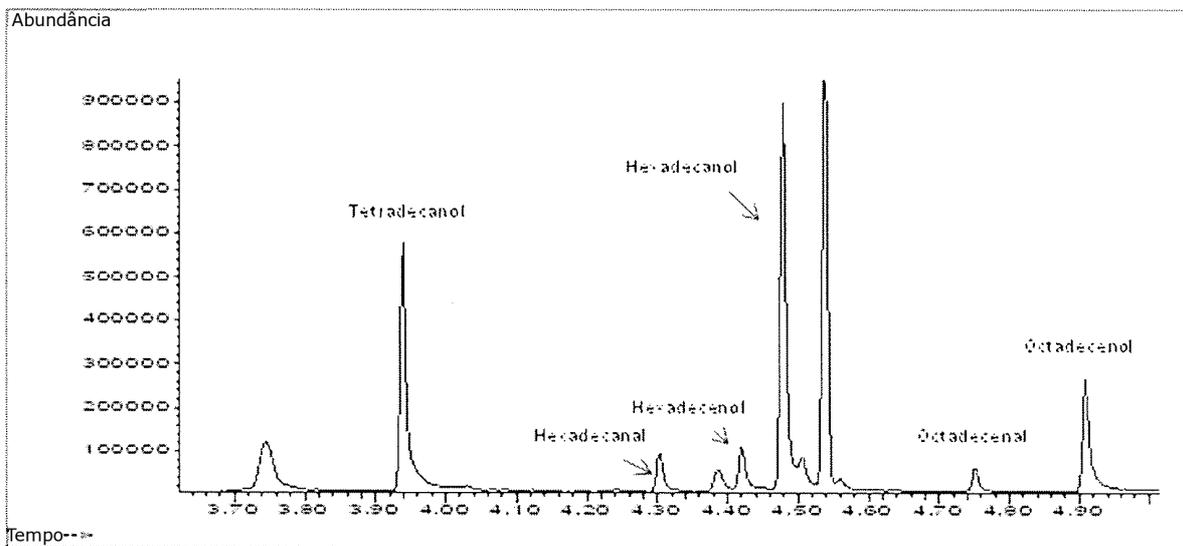


FIGURA 8

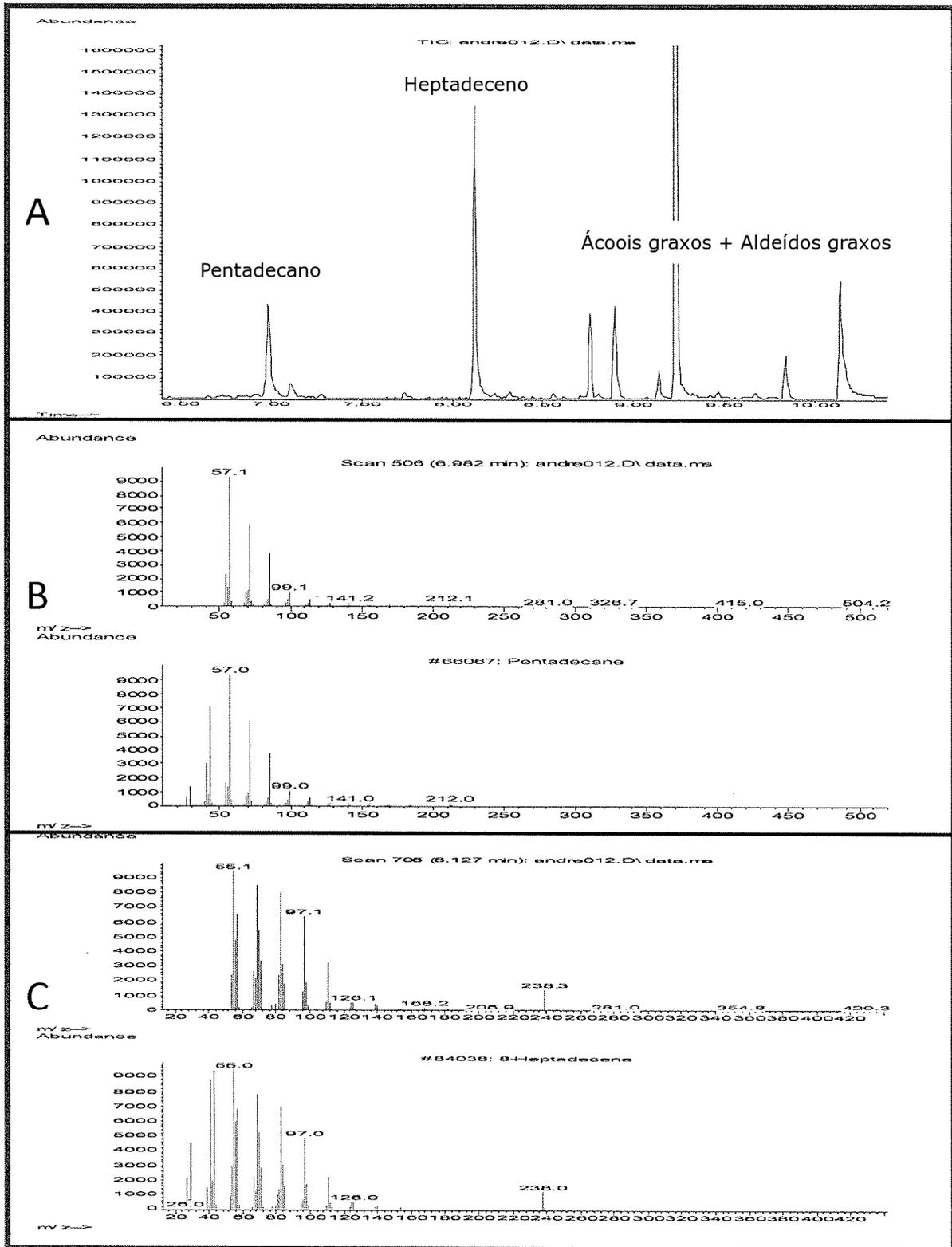


FIGURA 9

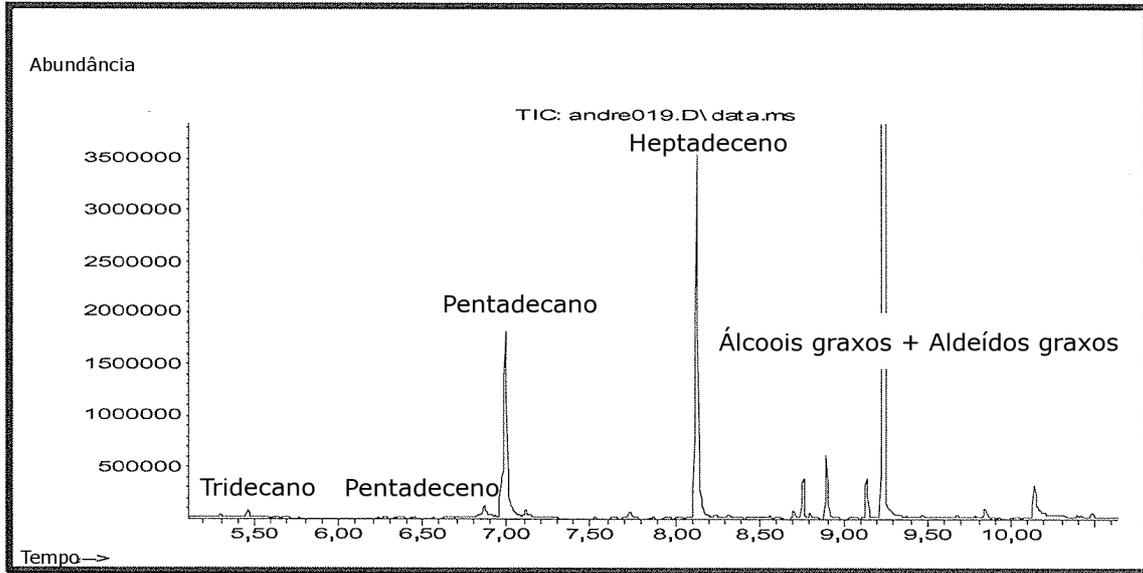


FIGURA 10

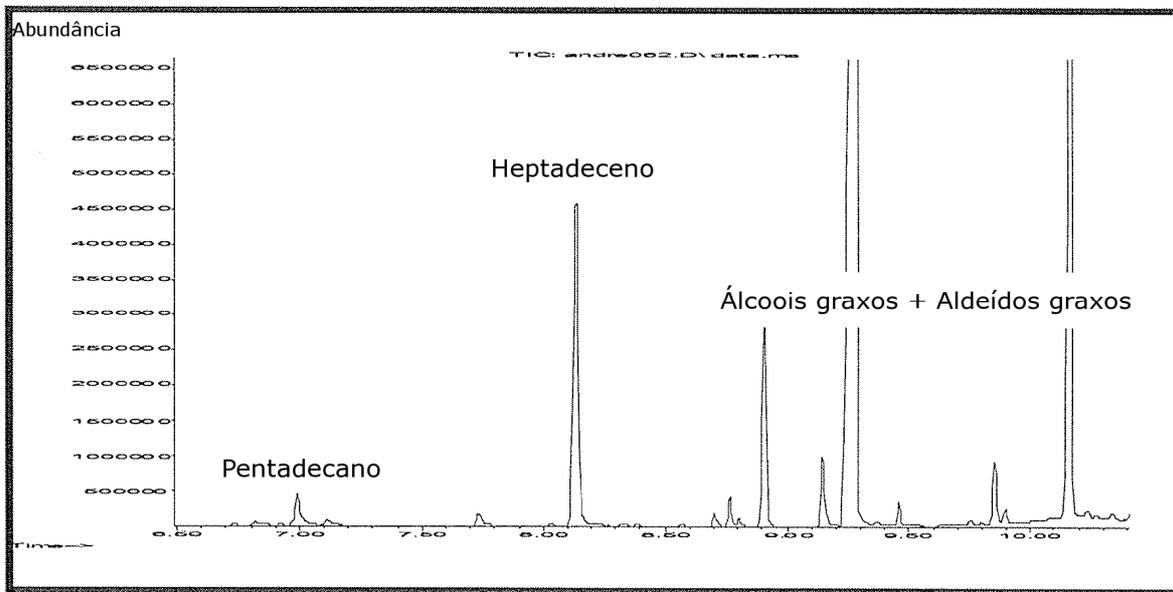


FIGURA 11

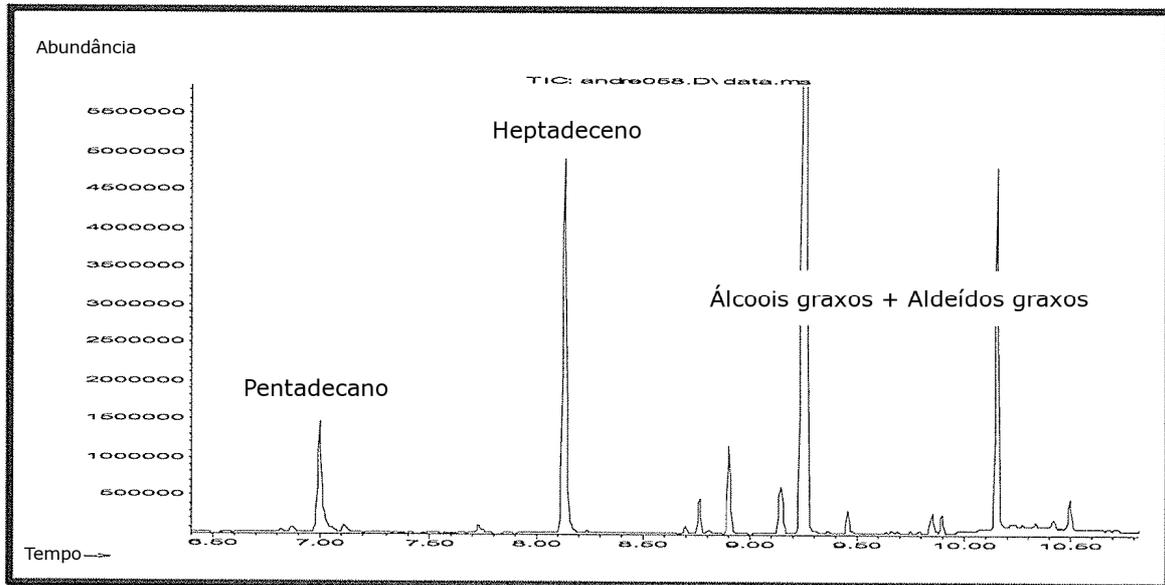


FIGURA 12

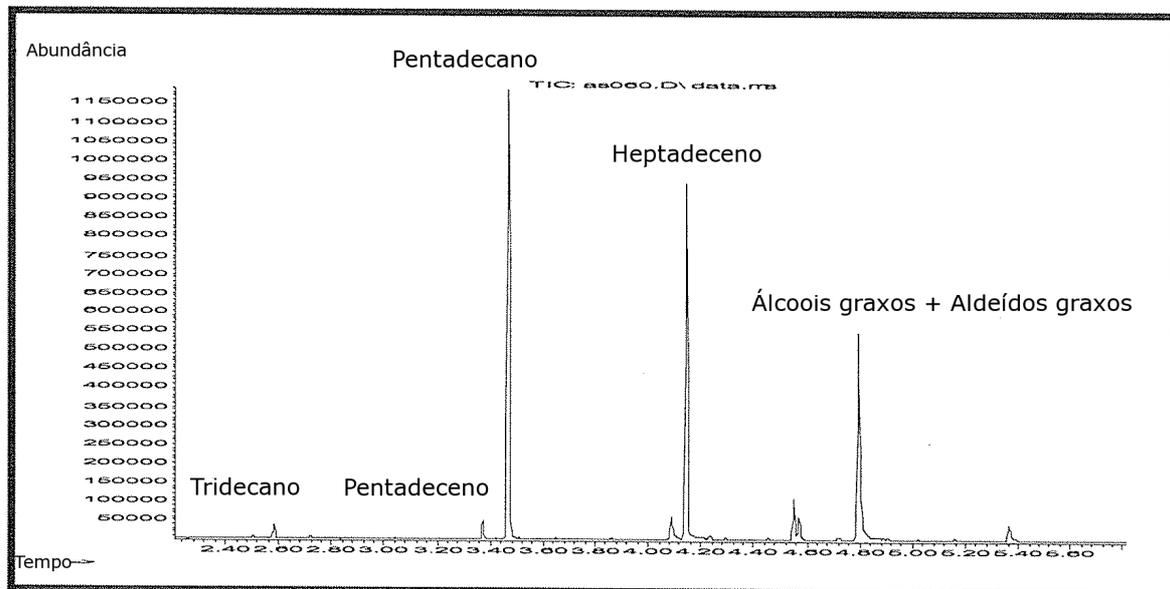


FIGURA 13

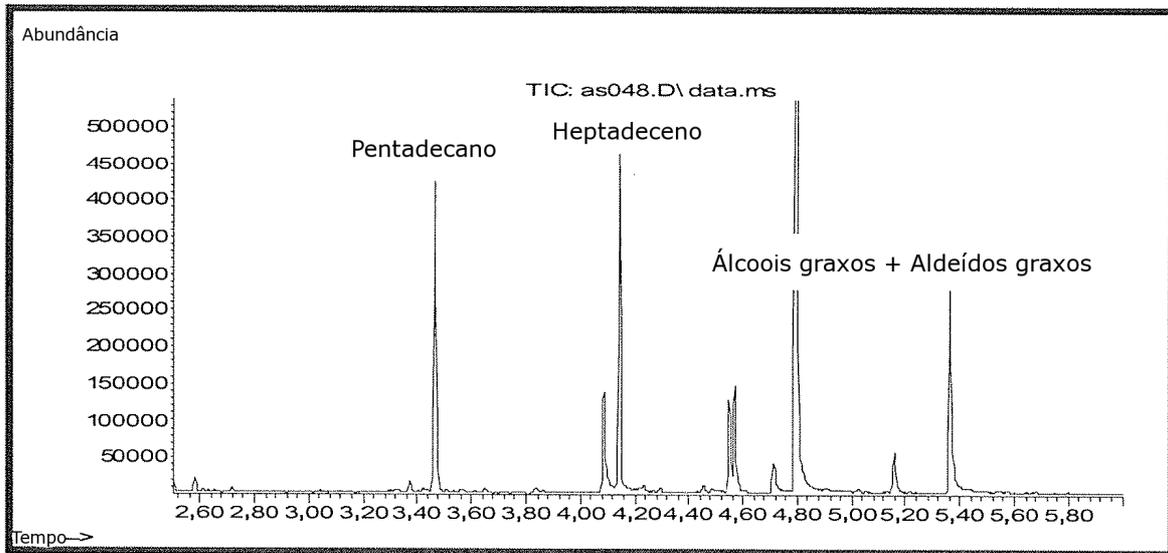


FIGURA 14

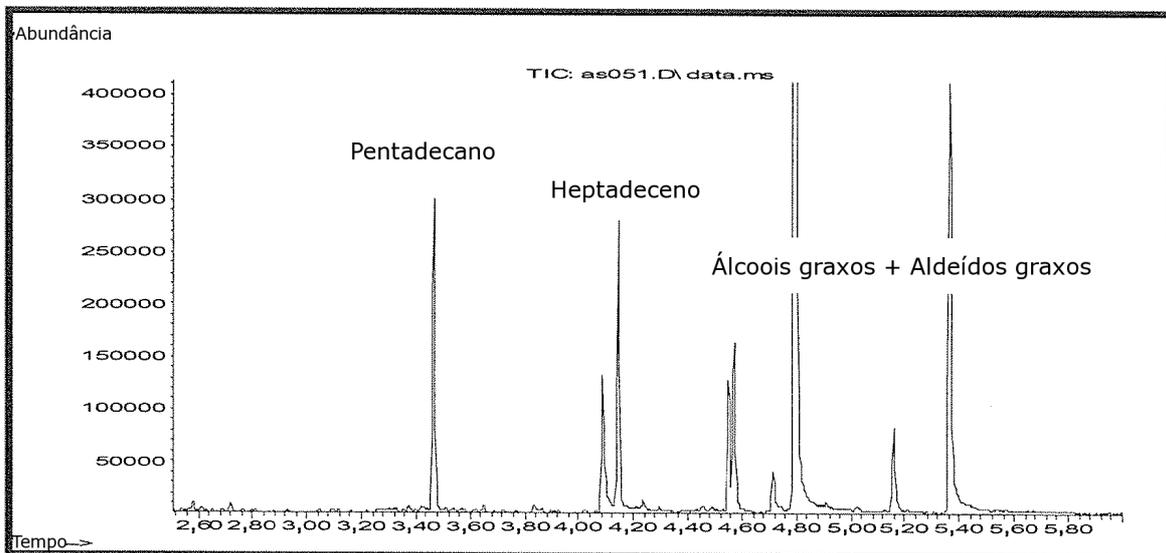


FIGURA 15

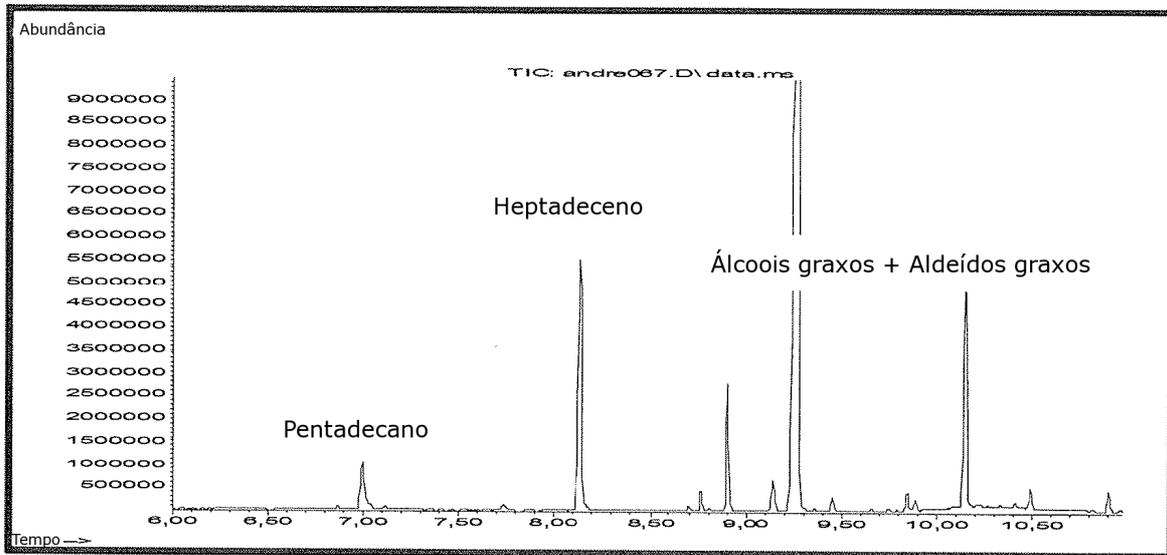


FIGURA 16

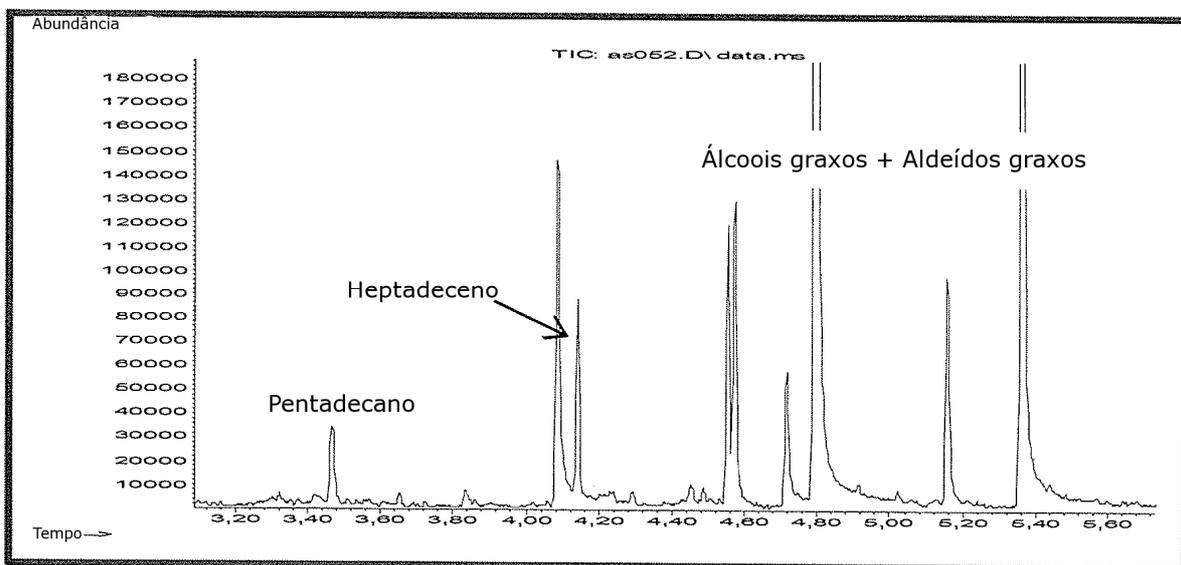


FIGURA 17

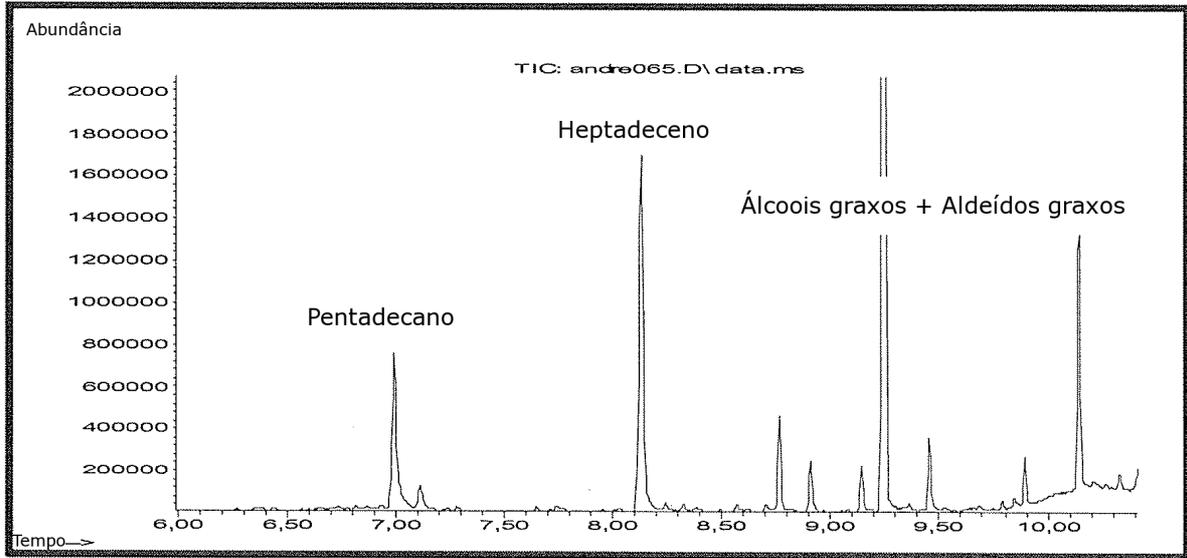


FIGURA 18

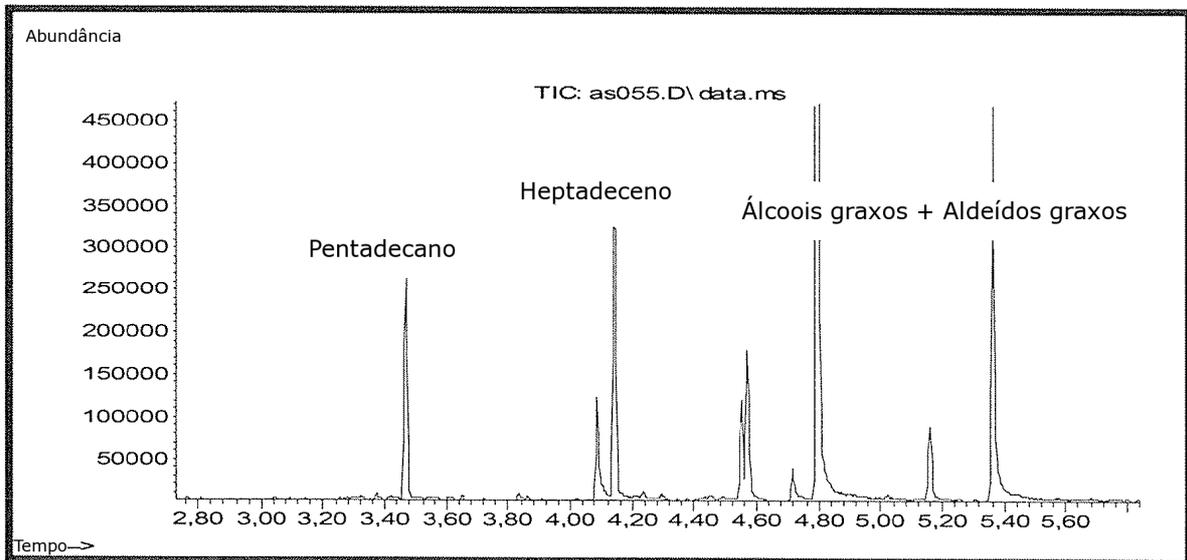


FIGURA 19

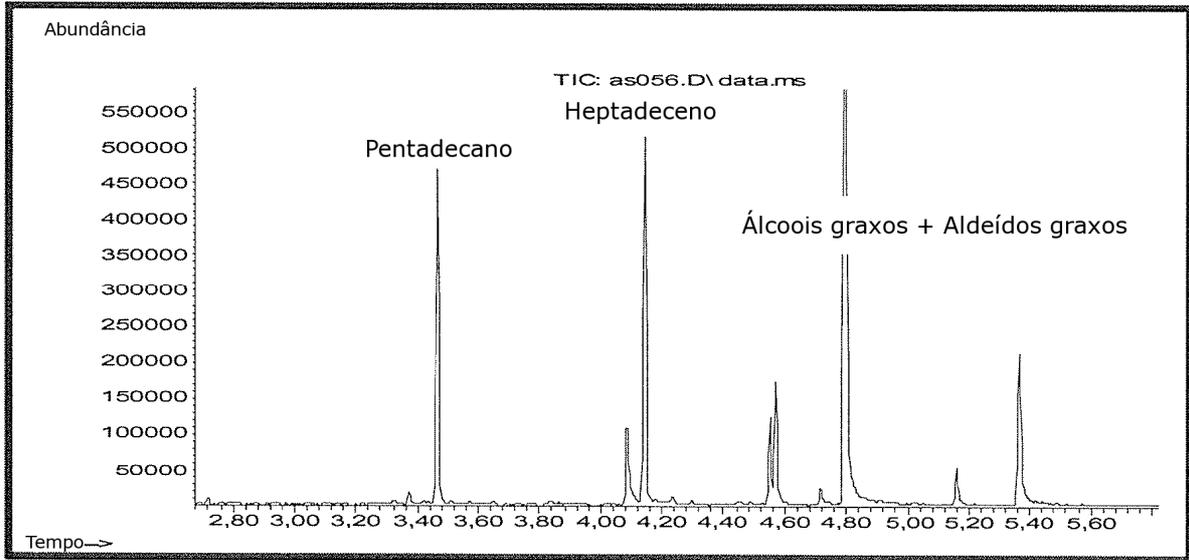


FIGURA 20

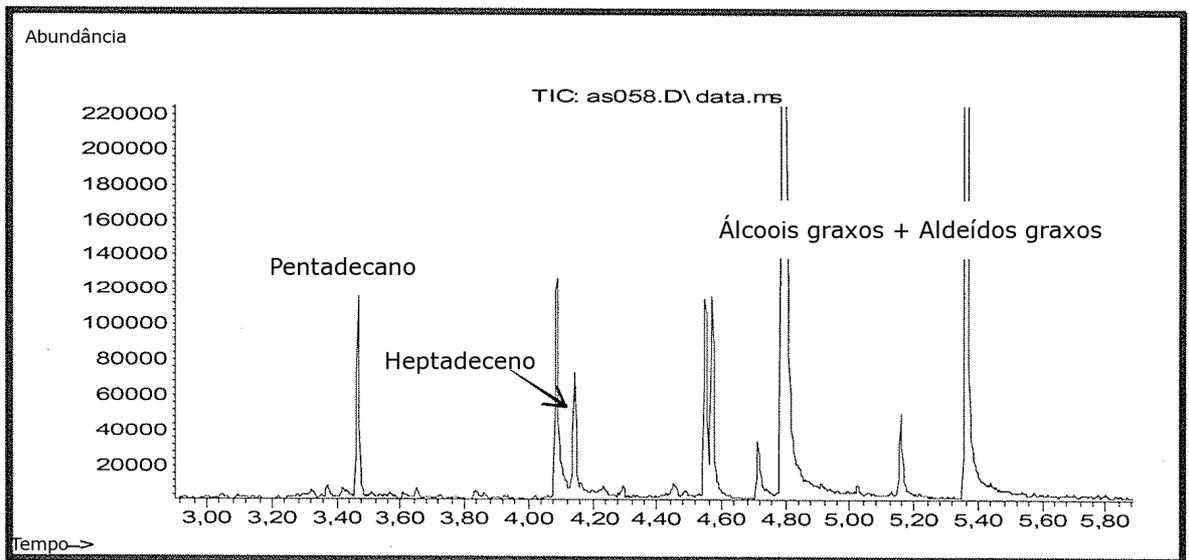


FIGURA 21

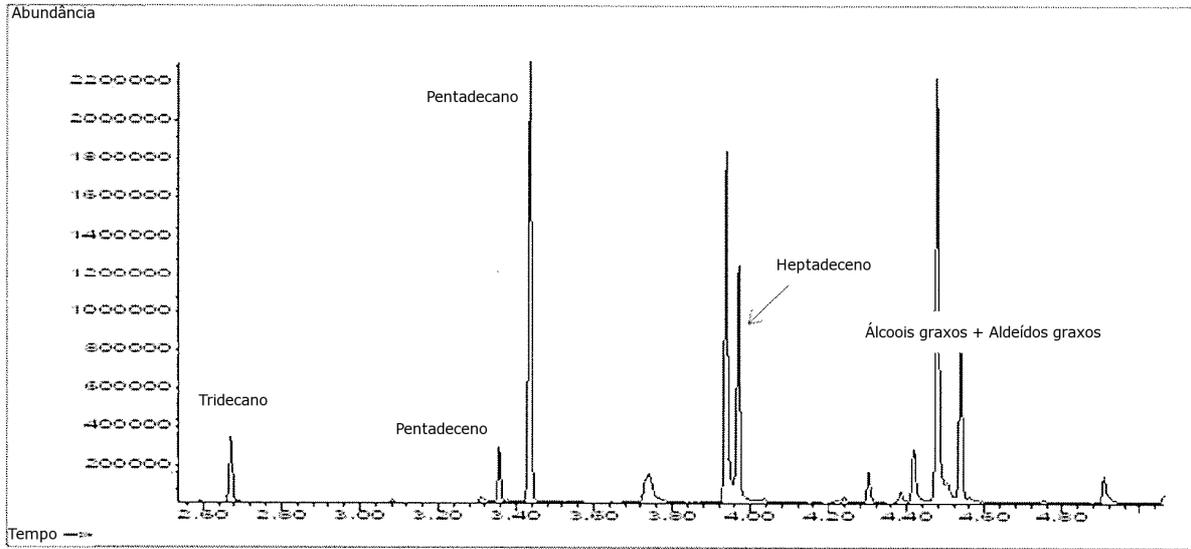


FIGURA 22

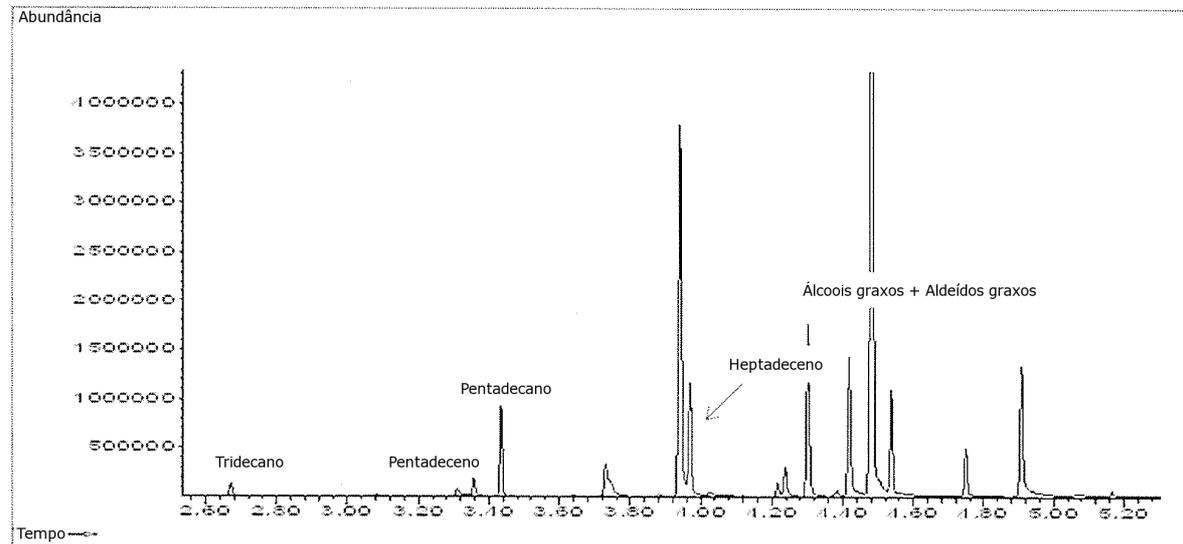


FIGURA 23

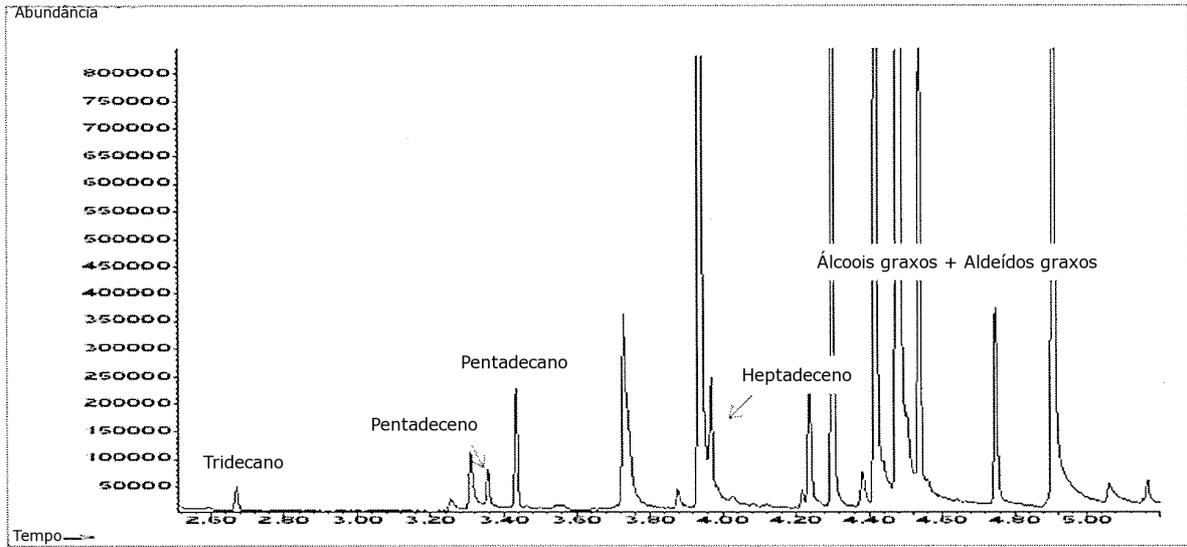


FIGURA 24

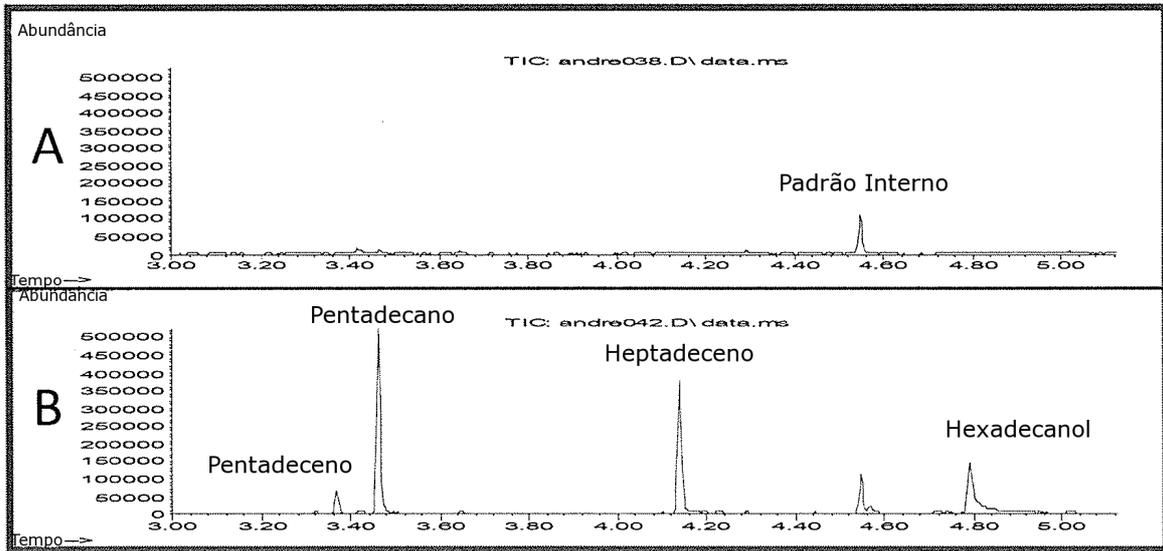


FIGURA 25

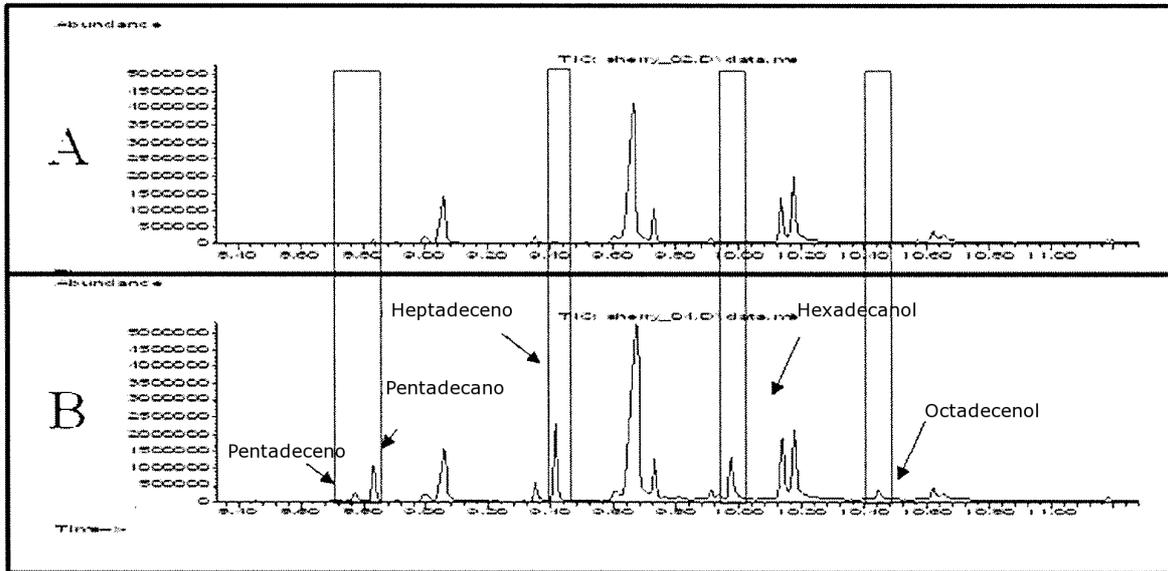


FIGURA 26

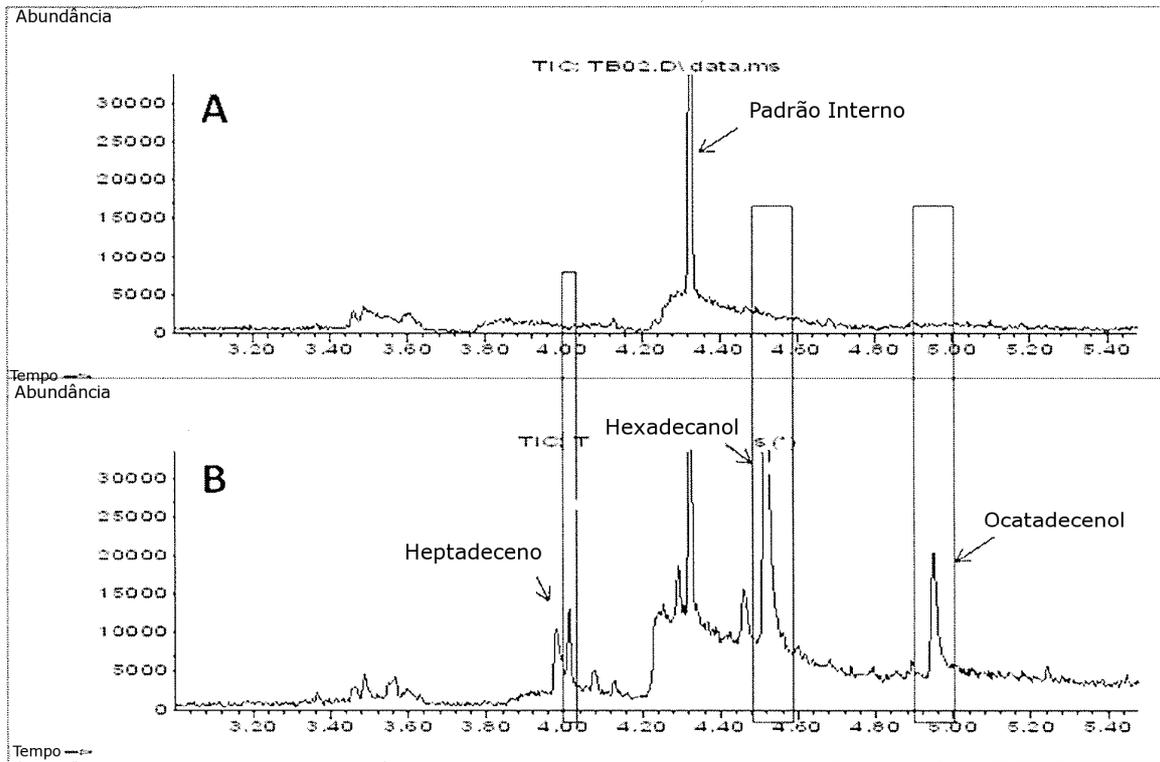


FIGURA 27

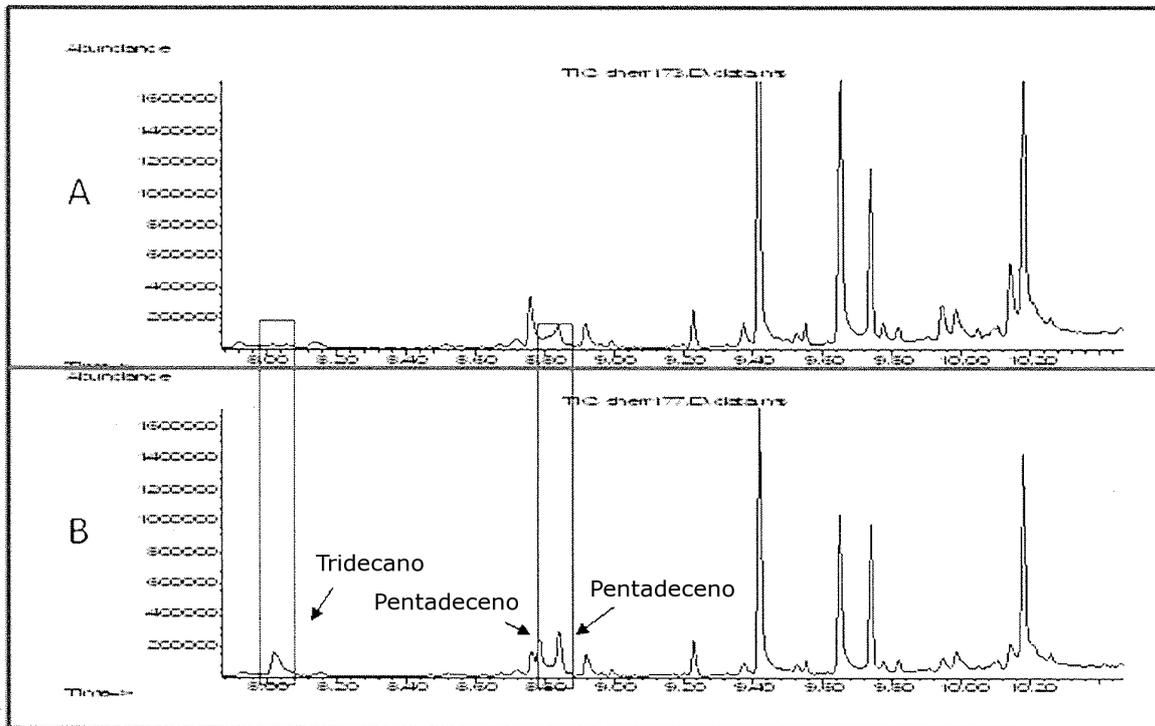


FIGURA 28

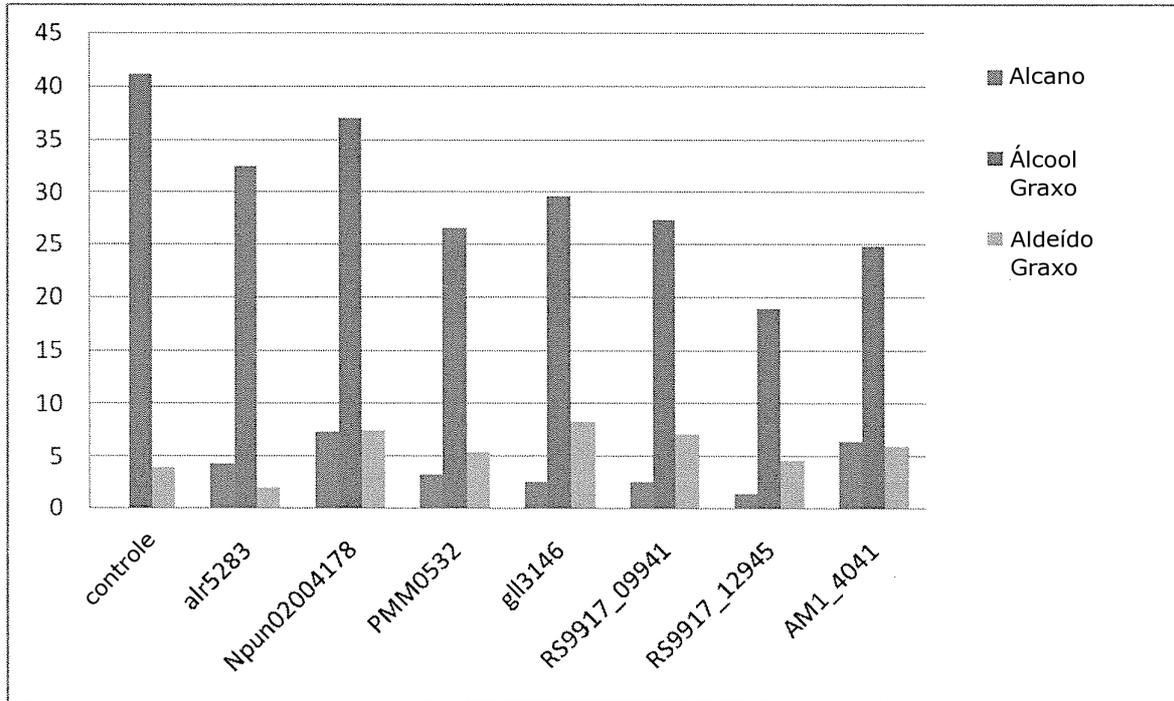


FIGURA 29

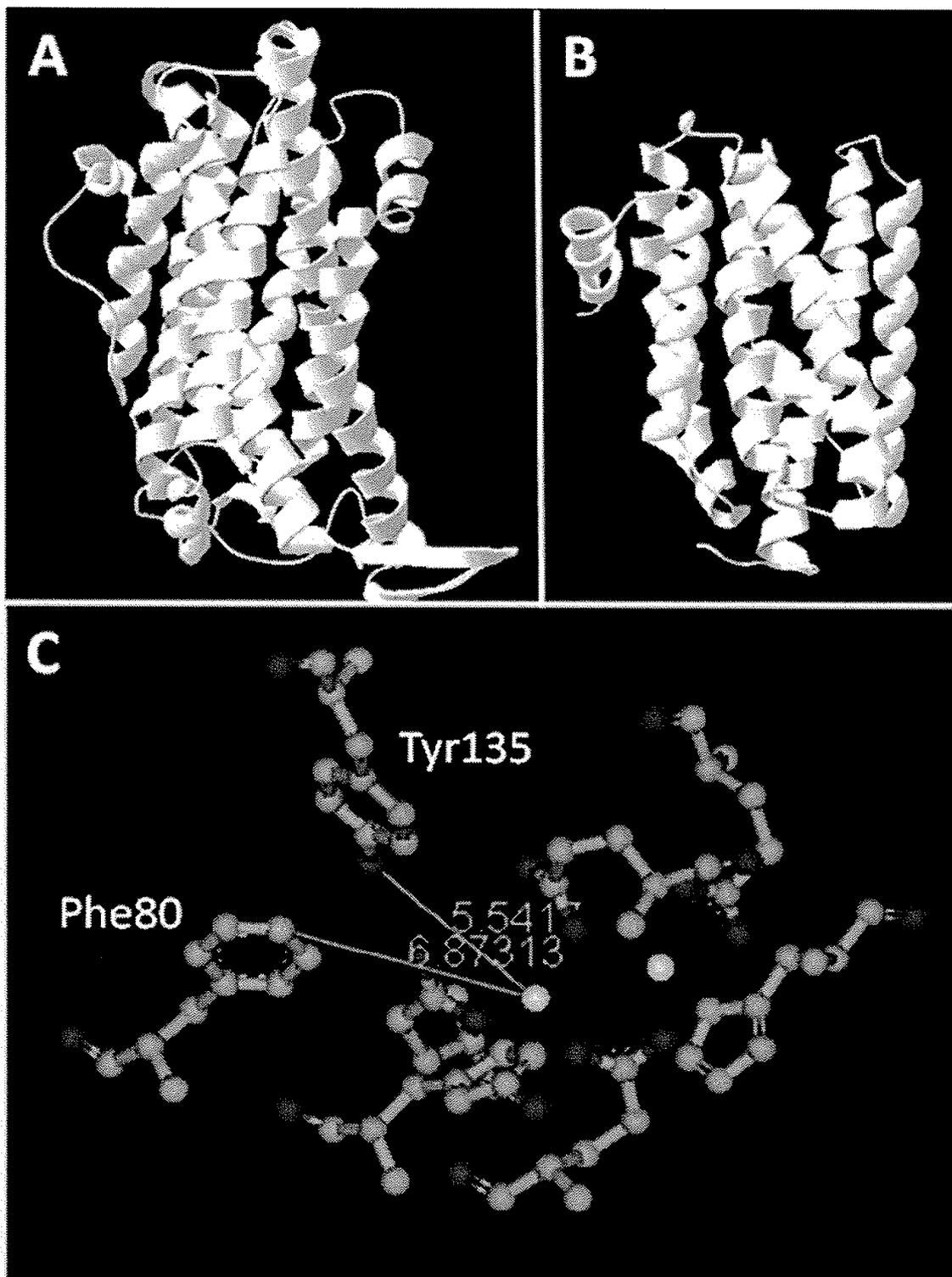


FIGURA 30

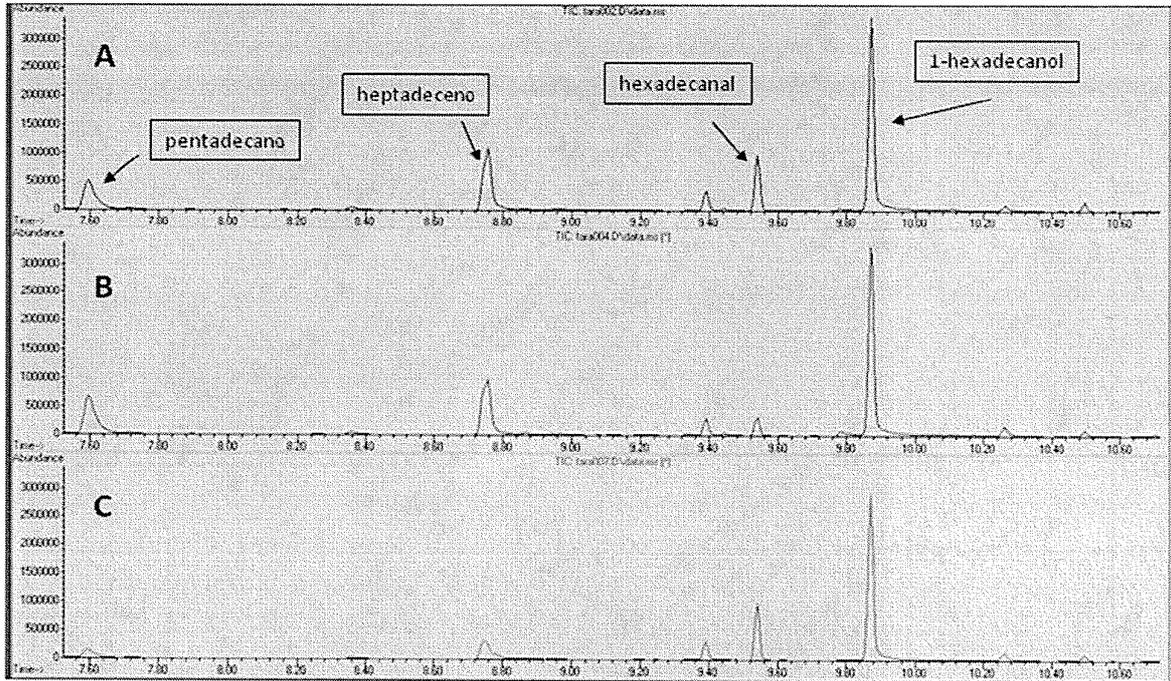


FIGURA 31

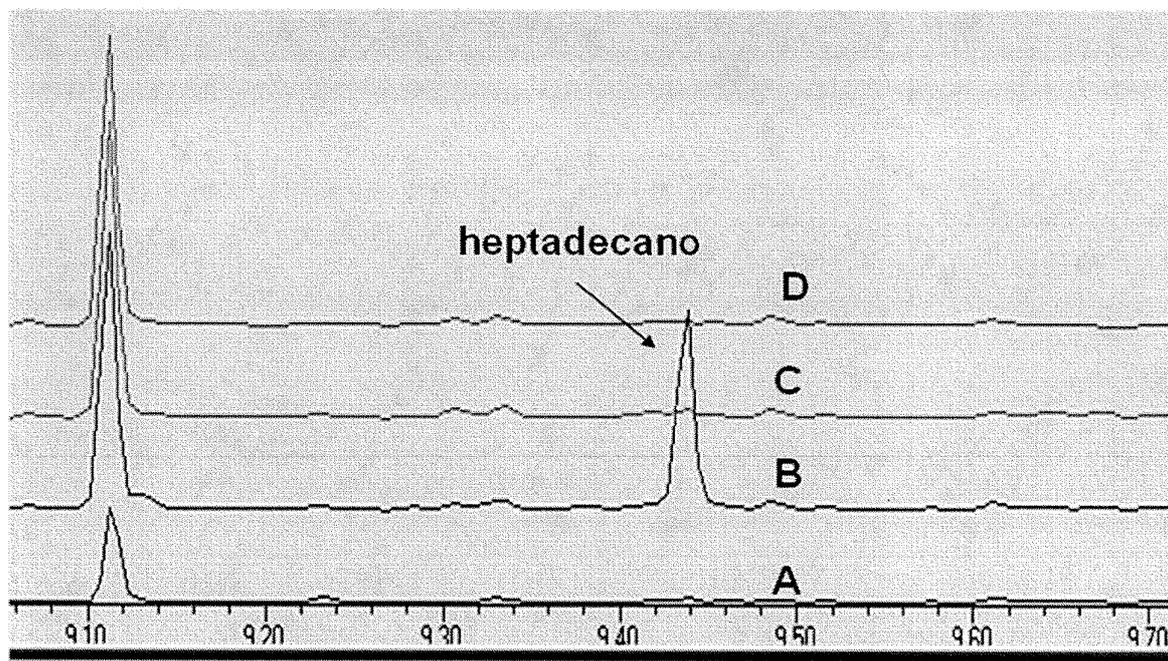


FIGURA 32

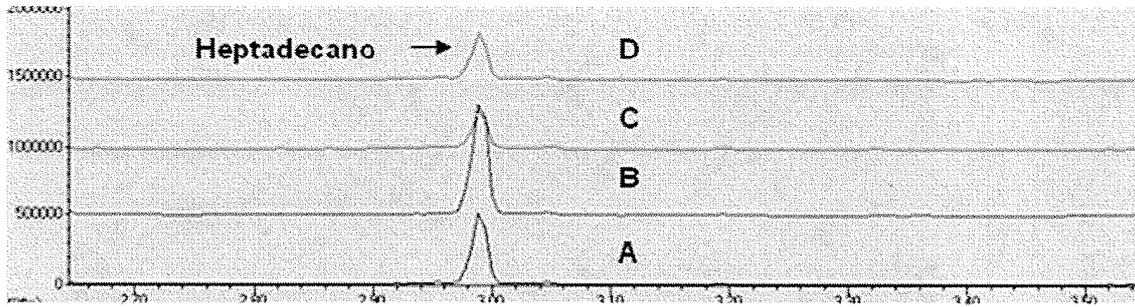


FIGURA 33

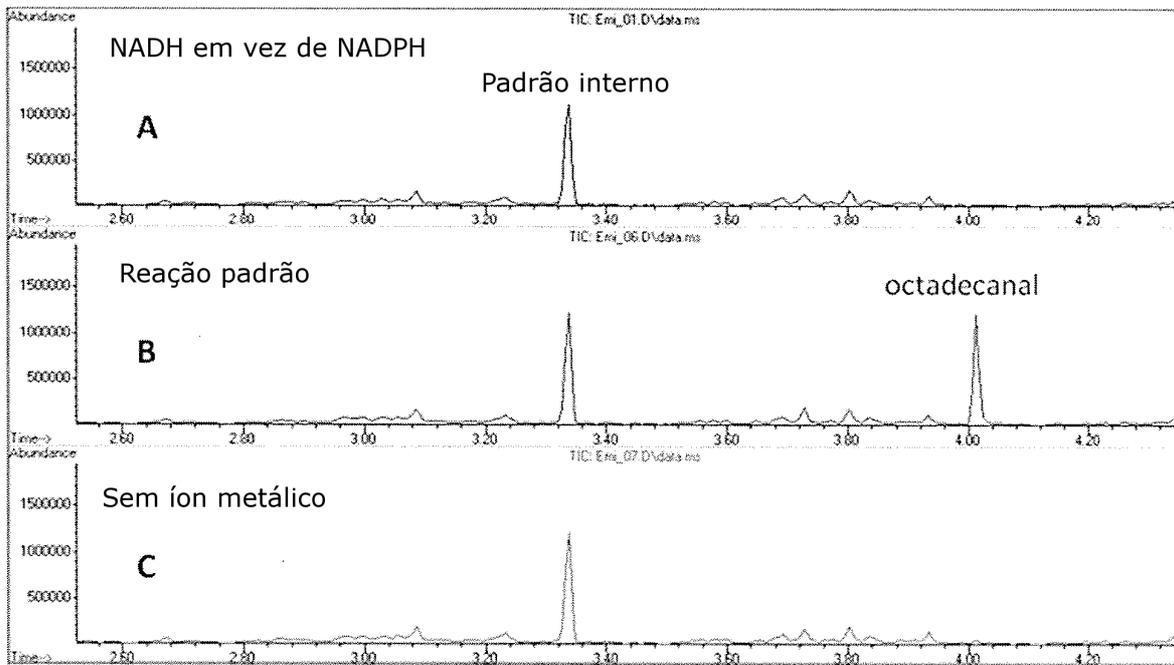


FIGURA 34

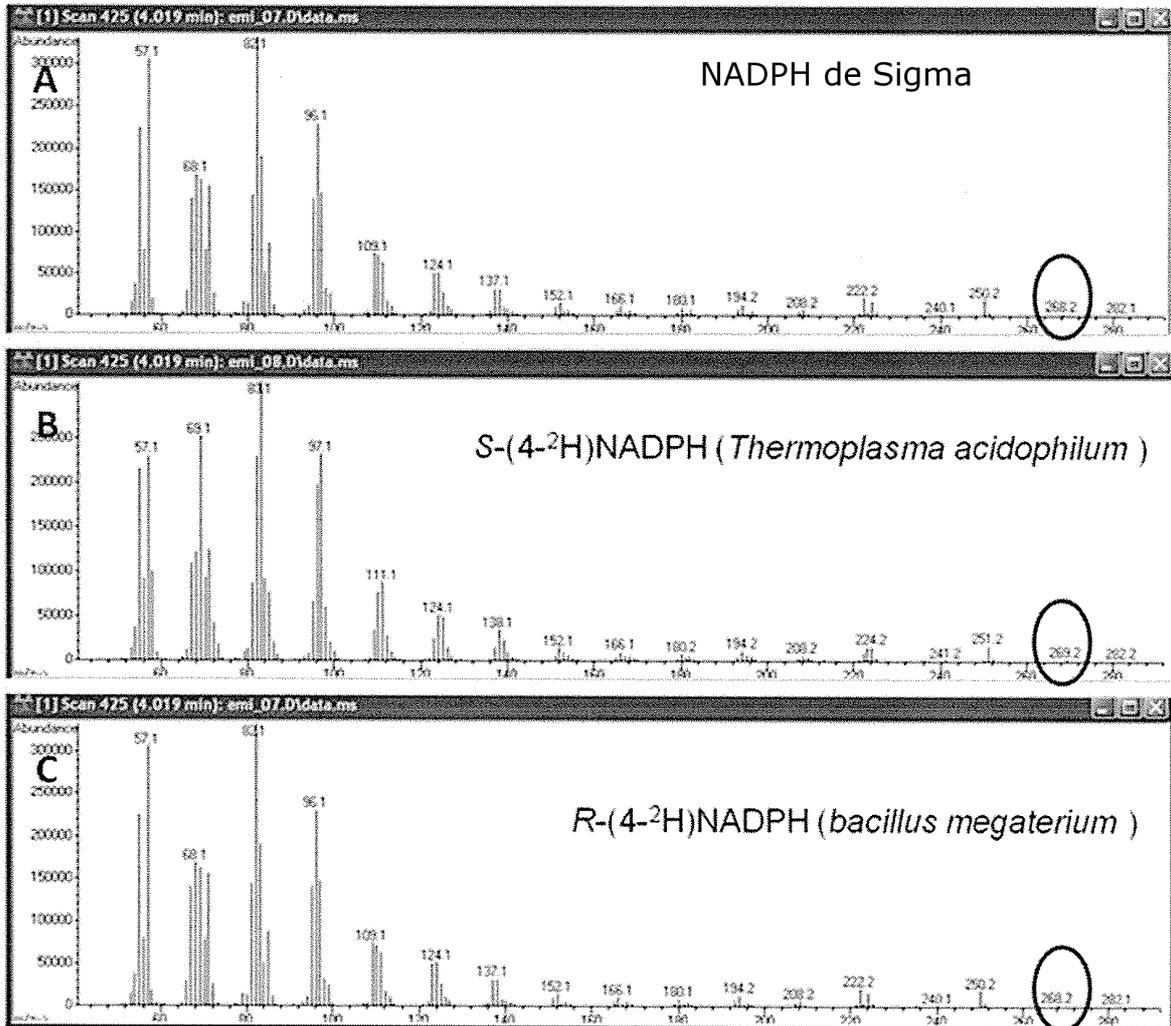


FIGURA 35

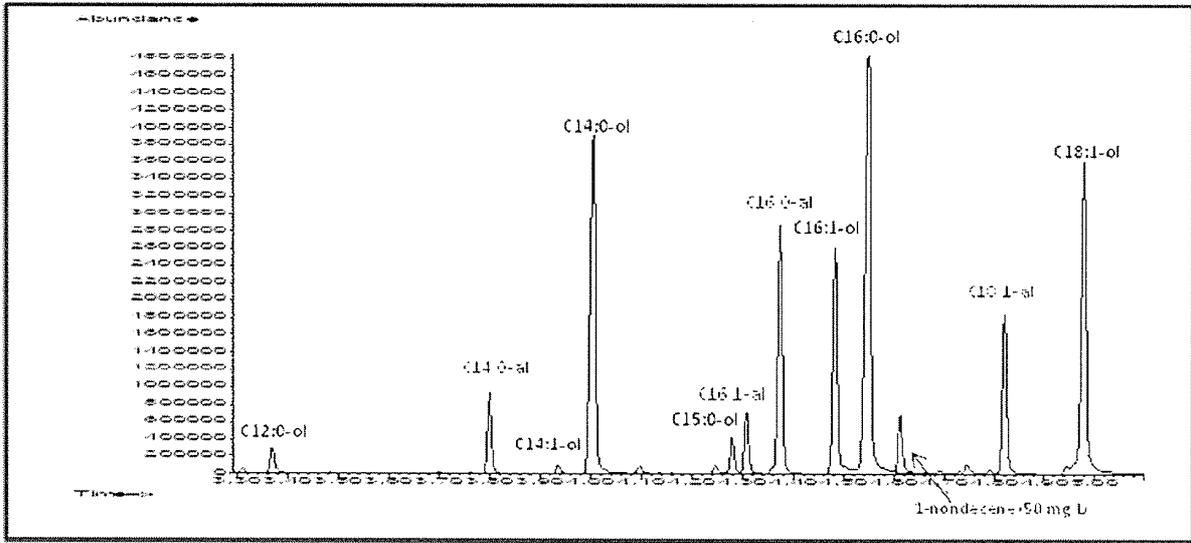


FIGURA 36

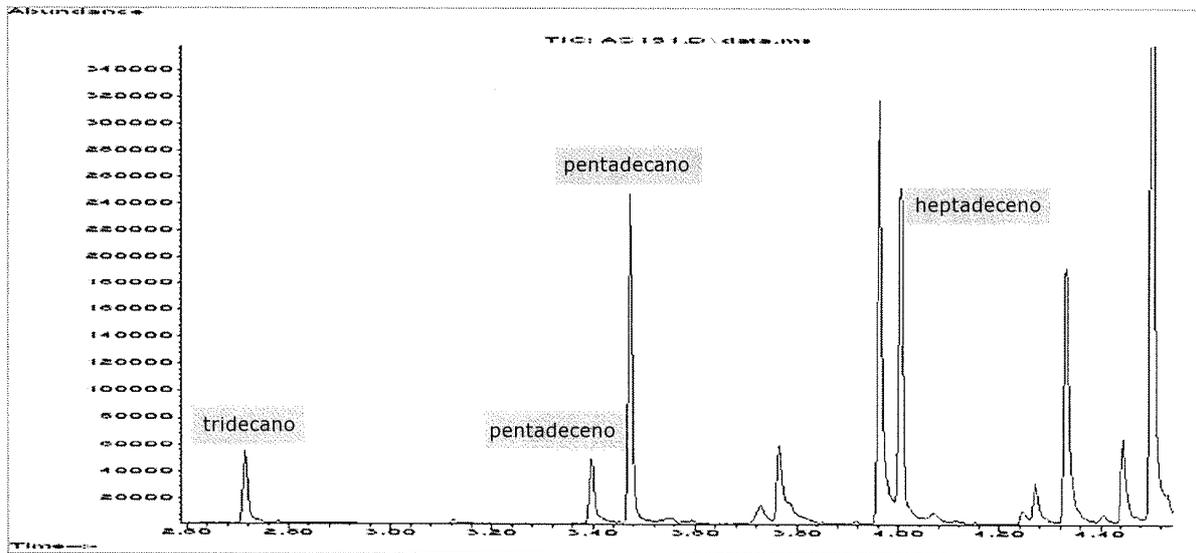


FIGURA 37

