

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

(21) **PI 0713112-7 A2**

(22) Data de Depósito: 29/06/2007
(43) Data da Publicação: 17/04/2012
(RPI 2154)



(51) *Int.Cl.:*
C12N 9/02

(54) Título: PRODUÇÃO DE HIDROCARBONETO VOLÁTIL DE CADEIA CURTA USANDO MICROALGAS, CIANOBACTÉRIAS OU BACTÉRIAS GENETICAMENTE MANIPULADAS

(57) Resumo: PRODUÇÃO DE HIDROCARBONETO VOLÁTIL DE CADEIA CURTA USANDO MICROALGAS, CIANOBACTÉRIAS OU BACTÉRIAS GENETICAMENTE MANIPULADAS. A presente invenção fornece métodos e composições para a produção de hidrocarbonetos de isopreno a partir de microalgas, cianobactérias e bactérias fotossintéticas e não-fotossintéticas.

(30) Prioridade Unionista: 28/06/2007 US 11/770,412, 29/06/2006 US 60/806,244

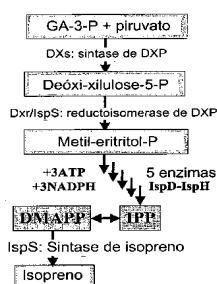
(73) Titular(es): The Regents Of The University Of California

(72) Inventor(es): Anastasios Melis

(74) Procurador(es): Dannemann, Siemsen, Bigler & Ipanema Moreira

(86) Pedido Internacional: PCT US2007072465 de 29/06/2007

(87) Publicação Internacional: WO 2008/003078 de 03/01/2008



Abreviações usadas:

GA-3-P = gliceraldeído-3-fosfato
IPP = pirofosfato de isopentenila
DMAPP = dimetilalil-pirofosfato

Relatório Descritivo da Patente de Invenção para "**PRODUÇÃO DE HIDROCARBONETO VOLÁTIL DE CADEIA CURTA USANDO MICROALGAS, CIANOBACTÉRIAS OU BACTÉRIAS GENETICAMENTE MANIPULADAS**".

5 REFERÊNCIA CRUZADA A PEDIDOS RELACIONADOS

O presente pedido reivindica o benefício do pedido provisório U.S. Nº 60/806.244, depositado em 29 de Junho de 2006, pedido o qual é aqui incorporado por referência.

ANTECEDENTES DA INVENÇÃO

10 Uma variedade de plantas coníferas, decíduas e herbáceas são conhecidas por possuir a capacidade genética e enzimática para a síntese e liberação de isoprenóides de cadeia curta (por exemplo, isopreno (C_5H_8) e metil-butenol ($C_5H_{10}O_1$)) no ambiente circundante. Tais isoprenóides de cadeia curta são derivados dos produtos do ciclo de Calvin precoce de fotosíntese e podem ser sintetizados no cloroplasto de plantas herbáceas, decí-
15 duas e coníferas através da assim denominada via de DXP-MEP em taxas substanciais sob determinadas condições de estresse ambiental. Estresse térmico do organismo é particularmente importante para a indução desse processo em plantas e a poluição por hidrocarboneto resultante da atmosfera tem sido o foco da técnica anterior nesse campo.
20

A emissão de isopreno de plantas herbáceas, decíduas e coníferas é em virtude da presença de um gene de isopreno sintase (*ispS*), um gene nuclear que codifica uma proteína localizada no cloroplasto que catalisa a conversão de dimetilalil difosfato (DMAPP) em isopreno. Conforme
25 mencionado acima, isoprenóides são sintetizados no cloroplasto a partir dos produtos precoces do ciclo de Calvin (fixação e redução de carbono, veja figura 1). Isoprenóides com 5 carbonos, por exemplo, isopreno (C_5H_8) e metil-butenol ($C_5H_{10}O_1$) são moléculas hidrofóbicas relativamente pequenas, sintetizadas diretamente a partir de DMAPP (figura 2). Esses isoprenóides
30 são moléculas voláteis que passam facilmente através das membranas celulares e, desse modo, são emitidas das folhas para a atmosfera. O processo de indução de estresse térmico e emissão de hidrocarbonetos de cadeia cur-

ta pelas plantas tem sido discutido como poluição indesejável da atmosfera na literatura. Não há descrição da geração em massa, coleta e captura desses hidrocarbonetos a partir de folhas de plantas herbáceas, decíduas e coníferas.

5 Há uma necessidade urgente para o desenvolvimento de biocombustíveis renováveis que ajudam a atingir as demandas globais por energia, mas sem contribuir para alterações climáticas. A presente invenção se dirige a essa necessidade proporcionando métodos e composições para gerar hidrocarbonetos voláteis de cadeia curta que são derivados totalmente da luz do sol, dióxido de carbono (CO_2) e água (H_2O). Esses hidrocarbonetos podem servir como biocombustível ou estoque de alimentação na indústria de química sintética.

BREVE SUMÁRIO DA INVENÇÃO

15 A invenção é baseada, em parte, na descoberta de que microalgas, cianobactérias e fotossíntese procariota podem ser empregadas, mediante modificação adequada, para produzir isoprenóides com 5 carbonos (por exemplo, figura 3). A via biossintética de isoprenóide DXP-MEP é absolutamente requerida em plantas e algas, uma vez que ela leva à síntese de muitos compostos celulares de cadeia mais longa essenciais. Algas verdes unicelulares expressam especificamente essa via em seu cloroplasto e utilizam as enzimas correspondentes para a biossíntese de uma grande variedade de moléculas (carotenóides, tocoferóis, fitol, esteróis, hormônios, dentre muitos outros). A presente invenção se refere a métodos e composições para o uso de microalgas, cianobactérias e bactérias fotossintéticas e não fotossintéticas geneticamente modificadas na produção e coleta de compostos de isoprenóide voláteis com 5 carbonos, por exemplo, isopreno e metil-butenol. Tais organismos geneticamente modificados podem ser usados comercialmente em um sistema de cultura em massa fechado, por exemplo, para proporcionar uma fonte de combustível renovável para motores de combustão interna ou, quando de reforma "on-board", em motores operados com célula de combustível; ou proporcionar uma fonte de isopreno para usos em outros processos químicos, tais como síntese química.

Microalgas, cianobactérias e bactérias fotossintéticas e não fotossintéticas não possuem um gene de isopreno sintase ou de metil-butenol sintase, o qual catalisa a última etapa obrigatória na biossíntese de isopreno (C_5H_8) e metil-butenol ($C_5H_{10}O_1$), respectivamente. A presente invenção, portanto, proporciona métodos e composições para modificar geneticamente microorganismos para expressar um gene de isopreno sintase, por exemplo, um gene de isopreno sintase de álamo códon-ajustado, de modo a conferir produção de isopreno (C_5H_8) ao organismo.

Em aspectos adicionais, a invenção também proporciona um método e composições para a modificação genética de microalgas, cianobactérias e bactérias fotossintéticas e não fotossintéticas para conferir a esses microorganismos superexpressão de genes endógenos e proteínas que codificam a primeira etapa obrigatória na biossíntese de isoprenóide. A invenção pode, assim, ainda compreender expressão crescente dos genes *Dxs* e *Dxr* nativos no microorganismo, por exemplo, algas verdes, tais como *Chlamydomonas reinhardtii*; cianobactérias, tais como *Synechocystis* sp.; ou bactérias fotossintéticas, tais como *Rhodospirillum rubrum* ou bactérias não fotossintéticas, tais como *Escherichia coli*. *Dxs* e *Dxr* codificam enzimas que catalisam as primeiras etapas obrigatórias na biossíntese de isoprenóide.

Em algumas modalidades, microalgas são empregadas. Microalgas são fábricas de fotossíntese, com o cloroplasto ocupando ~70% do volume da célula; o cloroplasto de algas verdes contém mais de 3 milhões de circuitos de transporte de elétrons, cada um sendo capaz de distribuir 100 elétrons por segundo ao ciclo de Calvin para conversão de CO_2 em GA-3-P; microalgas não têm raízes, caules, folhas ou flores sobre os quais investir recursos fotossintéticos, assim, uma maior fração do produto fotossintético pode ser dirigida à geração de isoprenóide volátil; microalgas crescem e se reproduzem mais rápido do que qualquer outra planta terrestre ou aquática, duplicando de biomassa por dia; e microalgas são não tóxicas e não poluentes, assim, ambientalmente seguras para a cultura e exploração comercial em massa. Consequentemente, em algumas modalidades, a invenção proporciona um processo para modificar o processo altamente eficiente de fotos-

síntese de microalgas para gerar, em alto volume, hidrocarbonetos de isoprenóide de cadeia curta (por exemplo, C_5H_8) a partir da luz do sol, CO_2 e H_2O . Tais microalgas modificadas podem ser cultivadas, por exemplo, em fotorreatores totalmente fechados de grande capacidade (por exemplo, 1.000-1.000.000 litros) para a produção e coleta de hidrocarbonetos de isoprenóide de cadeia curta voláteis.

A invenção ajudará a eliminar uma série de barreiras atuais na produção comercial, armazenamento e utilização de energia renovável incluindo, mas não limitado a: (a) diminuição do custo de produção e armazenamento de combustível, (b) melhora das proporções de peso/volume de combustível, (c) melhora da eficiência de produção/armazenamento de combustível, (d) aumento da durabilidade de armazenamento de combustível, (e) minimização do tempo de auto-reabastecimento, (f) oferta de armazenamento de combustível suficiente para uma gama de veículos aceitável, (g) produção de um combustível passível de um processo de regeneração, (h) o combustível não está sujeito à interferência pelo oxigênio, quer no estágio de produção ou armazenamento.

Em um aspecto, a invenção proporciona um método de produção de hidrocarbonetos de isopreno em um microorganismo selecionado do grupo consistindo em microalgas, cianobactérias ou bactérias fotossintéticas, o método compreendendo: introdução de um cassete de expressão que compreende uma sequência de ácido nucleico que codifica isopreno sintase no microorganismo; e cultura do microorganismo sob condições nas quais o ácido nucleico que codifica isopreno sintase é expresso. Em algumas modalidades, o microorganismo é uma microalga, tal como alga verde, por exemplo, *Chlamydomonas reinhardtii*, *Scenedesmus obliquus*, *Chlorella vulgaris* ou *Dunaliella salina*. Em modalidades alternativas, o microorganismo é uma cianobactéria, tal como uma *Synechocystis* sp. Em outras modalidades, o microorganismo é uma bactérias fotossintética, tal como *Rhodospirillum rubrum*. Alternativamente, em algumas modalidades, o microorganismo pode ser uma bactéria não fotossintética, tal como *Escherichia coli*.

Em algumas modalidades, o ácido nucleico introduzido no mi-

croorganismo compreende uma sequência que codifica um polipeptídeo de isopreno sintase que tem a sequência apresentada em SEQ ID NO: 2 ou tem a sequência apresentada em SEQ ID NO: 2, mas carece de uma região de peptídeo de trânsito. O polipeptídeo de isopreno sintase pode, por exemplo, compreender os resíduos de aminoácido 53 a 595 de SEQ ID NO: 2 ou resíduos 38 a 595 de SEQ ID NO: 2. Em algumas modalidades, o ácido nucleico compreende a sequência apresentada em SEQ ID NO: 1. Em outras modalidades, o ácido nucleico compreende a sequência de nucleotídeo de codificação de isopreno sintase apresentada em SEQ ID NO: 3 ou o ácido nucleico compreende a sequência de codificação de isopreno conforme apresentado em SEQ ID NO: 5.

Em outro aspecto, a invenção proporciona um microorganismo selecionado do grupo consistindo em uma célula de microalga, uma célula de cianobactéria e uma célula bacteriana fotossintética ou célula bacteriana não fotossintética, em que o microorganismo compreende um ácido nucleico heterólogo que codifica isopreno sintase e é operavelmente ligado a um promotor. O promotor pode ser um promotor constitutivo ou um promotor induzível. Em algumas modalidades, o microorganismo é uma alga verde, tal como *Chlamydomonas reinhardtii*, *Scenedesmus obliquus*, *Chlorella vulgaris* ou *Dunaliella salina*. Em outras modalidades, o microorganismo é uma bactéria fotossintética, tal como *Rhodospirillum rubrum*. Em algumas modalidades, o ácido nucleico heterólogo compreende uma sequência que codifica um gene de isopreno sintase que tem a sequência apresentada em SEQ ID NO: 2 ou tem a sequência apresentada em SEQ ID NO: 2, mas carece do peptídeo de trânsito. O polipeptídeo de isopreno sintase pode, por exemplo, compreender os resíduos de aminoácido 53 a 595 de SEQ ID NO: 2 ou os resíduos 38 a 595 de SEQ ID NO: 2. Em algumas modalidades, o ácido nucleico compreende a sequência apresentada em SEQ ID NO: 1. Em outras modalidades, o ácido nucleico compreende a sequência de nucleotídeo de codificação para isopreno sintase apresentada em SEQ ID NO: 3 ou o ácido nucleico compreende a sequência de codificação de isopreno conforme apresentado em SEQ ID NO: 5.

Em um outro aspecto, a invenção proporciona um método de produção de hidrocarbonetos de isopreno em um microorganismo que compreende um gene heterólogo que codifica isopreno sintase e que é selecionado do grupo consistindo em microalgas, cianobactérias, bactérias fotossintéticas e bactérias não fotossintéticas, o método compreendendo: cultura em massa do microorganismo em um bioreator fechado sob condições nas quais o gene de isopreno sintase é expresso; e coleta dos hidrocarbonetos de isopreno produzidos pelo microorganismo. Em algumas modalidades, o microorganismo é uma microalga que é uma microalga verde, tal como *Chlamydomonas reinhardtii*, *Scenedesmus obliquus*, *Chlorella vulgaris* ou *Dunaliella salina*. Alternativamente, o microorganismo pode ser uma cianobactéria, tal como *Synechocystis* sp. Em outras modalidades, o microorganismo é uma bactéria fotossintética, tal como *Rhodospirillum rubrum*. Em ainda outras modalidades, o microorganismo é uma bactéria não fotossintética, tal como *Escherichia coli*.

Em algumas modalidades dos métodos de cultura em massa da invenção, o gene heterólogo que codifica isopreno sintase compreende uma sequência que codifica um gene de isopreno sintase que tem a sequência apresentada em SEQ ID NO: 2 ou tem a sequência apresentada em SEQ ID NO: 2, mas carece do peptídeo de trânsito. O polipeptídeo de isopreno sintase pode, por exemplo, compreender os resíduos de aminoácido 53 a 595 de SEQ ID NO: 2 ou resíduos 38 a 595 de SEQ ID NO: 2. O ácido nucleico pode, por exemplo, compreender a sequência apresentada em SEQ ID NO: 1. Em outras modalidades, o ácido nucleico compreende a sequência de nucleotídeo de codificação para isopreno sintase apresentada em SEQ ID NO: 3 ou a sequência de codificação de isopreno conforme apresentado em SEQ ID NO: 5.

Em algumas modalidades dos métodos e composições da invenção, o ácido nucleico de IspS codifica uma proteína que compreende a sequência de aminoácido de SEQ ID NO: 8 ou que compreende os aminoácidos 46 a 608 de SEQ ID NO: 8.

BREVE DESCRIÇÃO DOS DESENHOS

Figura 1. Via esquemática de fixação e redução de dióxido de carbono no ciclo de Calvin de fotossíntese e de canalização de carbono orgânico a partir do gliceraldeído-3-fosfato (GA-3-P) abundante através da via biossintética de deóxi-xilulose/metil-erititol (DXP/MEP) em isoprenóides.

5 Figura 2. (Painel esquerdo). Reação enzimática em uma única etapa para a biossíntese de isopreno e metil-butenol no cloroplasto de árvores herbáceas/decíduas e pinheiros americanos, respectivamente. (Painel direito). Fórmula química do isopreno (C_5H_8) e metil-butenol ($C_5H_{10}O$).

10 Figura 3. A via biossintética de DXP/MEP que leva à formação de isoprenóides voláteis a partir dos metabólitos de cloroplasto GA-3-P (gliceraldeído-3-fosfato) e piruvato. Sete reações enzimáticas distintas são necessárias para sintetizar isopreno a partir de GA-3-P e piruvato. Algas verdes unicelulares, cianobactérias, bactérias fotossintéticas e determinadas bactérias não fotossintéticas possuem os primeiros seis desses genes, mas
15 carecem dos genes de isopreno sintase ou metil-butenol sintase.

 Figura 4. Co-transformação e recombinação homóloga de DNA de cloroplasto de algas verdes, por exemplo, *Chlamydomonas reinhardtii*, com o novo gene *Cr-IspS*. Este construto contém o promotor *atpA* (*PatpA*), fundido à extremidade 5' UTR de um DNA com sinalizador de epítipo de
20 hemaglutinina (HA) com três cópias códon-otimizado. A sequência de DNA é seguida pela região de codificação de *Cr-IspS*, seguido pela 3' UTR de *atpA*.

 Figura 5. Seleção de transformantes de *IspS* de *C. reinhardtii* (*Cr-IspS*) através de PCR do DNA genômico. Iniciadores N e C representam o conjunto de iniciadores usado para amplificação e seus locais de anela-
25 mento são mostrados na figura 4.

 Figura 6. A. Integridade do transgene *Cr-IspS* testado através de análise de Southern blot do DNA genômico. Filtros hibridizados com a sonda de DNA *Cr-IspS*. Hibridização com um fragmento de *NdeI/XbaI* radio-marcado da região de codificação de *Cr-IspS* identificou uma banda de 3,0 kbp
30 exclusivamente na linhagem Nº 9 de transformante *Cr-IspS*, enquanto que nenhuma banda detectável pôde ser observada na fileira da linhagem de controle Nº 7. B. Coloração com brometo de etídio para testar quantidades

iguais de aplicação de DNA em A.

Figura 7. A Representação esquemática do gene *Cr-IspS* sinalizado com 3X HA códon-otimizado. B. Validação de expressão do gene *Cr-IspS*. Frações de proteína solúvel, as quais correspondem a 10 ou 20 µg de
5 clorofila, foram submetidas a SDS-PAGE e análise por Western blot com anticorpos policlonais anti-HA específicos.

Figura 8. Componentes e estrutura do plasmídeo *plspS*. O novo gene de isopreno sintase (*Ss-IspS*) com uso de códon desenhado para expressão em cianobactérias, por exemplo, *Synechocystis*, o qual inclui um
10 gene de resistência à ampicilina. A nova sequência de DNA de *Ss-IspS* foi desenhada com base no molde da sequência de aminoácido da proteína de isopreno sintase de álamo, com critérios projetados para se conformar às preferências de códon de *Synechocystis*. Sítios de restrição foram introduzidos para facilitar clonagem. A nova sequência de DNA de *Ss-IspS* foi sintetizada e clonada no plasmídeo *plspS* para propagação em *E. coli*.
15

Figura 9. Construção do plasmídeo pAIGA para transformação de cianobactérias, por exemplo, *Synechocystis*. Sequências de flanqueamento do gene *psbA3* de *Synechocystis* foram usadas para recombinação homóloga do plasmídeo e subsequentemente direcionar a expressão do gene
20 *Ss-IspS* com um promotor forte. Um cassete de resistência à gentamicina foi introduzido no plasmídeo na extremidade 3' do gene *Ss-IspS* para servir como um marcador selecionável. O gene *Ss-IspS* foi clonado entre os sítios de restrição de *NcoI* e *PstI*.

Figura 10. Recombinação homóloga dupla. Esquema mostrando o princípio de transformação de *Synechocystis* sp. através de recombinação
25 homóloga dupla e substituição do gene *psbA3* nativo pela estrutura de resistência à Gm *Ss-IspS*.

Figura 11. Estrutura de um plasmídeo contendo *Ss-IspS* sinalizado com His para superexpressão de proteína recombinante em bactérias, por exemplo, *Escherichia coli*. O sinalizador de histidina N-terminal foi introduzido para facilitar purificação da proteína recombinante. Expressão em
30 *Escherichia coli* foi induzida pela adição de IPTG à cultura celular líquida.

Figura 12. Evidência de expressão da proteína recombinante *Ss-IspS* sinalizada com His em bactérias, por exemplo, *E. coli*. SDS-PAGE corado com Coomassie da proteína total separada eletroforéticamente de extratos de células de *E. coli* trazendo o plasmídeo pETIspS. Fileira 1: Cultura de controle não induzida. Fileiras 2 a 4 e 6 a 10: Culturas de *E. coli* induzidas. Fileira 5: Marcadores de tamanho de peso molecular de proteína.

Figura 13. Alinhamento pelo Clustal de quatro proteínas de isopreno sintase conhecidas.

DESCRIÇÃO DETALHADA DA INVENÇÃO

10 **Definições**

"Microalgas", "algas" ou semelhante se refere à plantas que pertencem ao subfilo Algae do filo Thallophyta. As algas são algas unicelulares, fotossintéticas, oxigênicas e são plantas não parasíticas sem raízes, caules ou folhas; elas contêm clorofila e têm uma grande variedade de tamanho, de microscópicas a grandes algas marinhas. Algas verdes, pertencendo ao Eukaryota--Viridiplantae--Chlorophyta--Chlorophyceae, podem ser usadas na invenção. Contudo, algas úteis na invenção também podem ser azuis-verdes, vermelhas ou marrons, contanto que as algas sejam capazes de desempenhar as etapas necessárias para fornecer um substrato para produzir isopreno.

Um "hidrocarboneto de isopreno volátil", no contexto da presente invenção, se refere a um isoprenóide com 5 carbonos de cadeia curta, por exemplo, isopreno ou metil-butenol.

Os termos "ácido nucleico" e "polinucleotídeo" são usados como sinônimos e se referem a um polímero fita simples ou dupla de bases de desoxirribonucleotídeo ou ribonucleotídeo lidas da extremidade 5' para 3'. Um ácido nucleico da presente invenção geralmente conterá ligações fosfodiéster embora, em alguns casos, análogos de ácido nucleico possam ser usados os quais podem ter estruturas alternativas compreendendo, por exemplo, ligações de fosforamidato, fosforotioato, fosforoditioato ou O-metilfosfoamidita (vide Eckstein, *Oligonucleotides and Analogues: A Practical Approach*, Oxford University Press); e estruturas e ligações de ácido peptídeo

nucleico. Outros ácidos nucleicos análogos incluem aqueles com estruturas positivas; estruturas não-iônicas e estruturas de não-ribose. Assim, ácidos nucleicos ou polinucleotídeos podem também incluir nucleotídeos modificados, que permitem leitura correta através de uma polimerase. "Sequência polinucleotídeo" ou "sequência de ácido nucleico" inclui as fitas senso e antissenso de um ácido nucleico, quer como fitas simples ou em uma dupla. Conforme será apreciado por aqueles na técnica, a representação de uma fita simples também define a sequência da fita complementar; assim, as sequências descritas aqui também proporcionam o complemento da sequência. A menos que de outro modo indicado, uma sequência de ácido nucleico particular também abrange, implicitamente, variantes da mesma (por exemplo, substituições de códon degenerado) e sequências complementares, bem como a sequência explicitamente indicada. O ácido nucleico pode ser DNA, genômico e cDNA, RNA ou um híbrido, onde o ácido nucleico pode conter combinações de desoxirribo- e ribonucleotídeos e combinações de bases, incluindo uracila, adenina, timina, citosina, guanina, inosina, xantina, hipoxantina, isocitosina, isoguanina, etc.

A expressão "uma sequência de ácido nucleico que codifica" se refere a um ácido nucleico o qual contém informação de sequência para um RNA estrutural, tal como rRNA, um tRNA ou a sequência de aminoácido primária de uma proteína ou peptídeo específico ou um sítio de ligação para um agente regulatório de trans-atuante. Essa expressão abrange especificamente códons degenerados (isto é, diferentes códons que codificam um único aminoácido) da sequência nativa ou sequências que podem ser introduzidas para se conformar à preferência de códon em uma célula hospedeira. No contexto da invenção, o termo "região de codificação de *IspS*", quando usada com referência a uma sequência de ácido nucleico de referência, tal como SEQ ID NOs: 3, 5 ou 7, se refere à região do ácido nucleico que codifica a proteína.

Um "gene" *IspS*, no contexto da presente invenção, se refere a um ácido nucleico que codifica uma proteína *IspS* ou fragmento da mesma. Assim, tal gene é, frequentemente, uma sequência de cDNA que codifica

IspS. Em outras modalidades, um gene *IspS* pode incluir sequências, tais como íntrons, que não estão presentes em um cDNA.

O termo "promotor" ou "elemento regulatório" se refere a uma região ou determinantes de sequência localizados à montante ou à jusante do início de transcrição que dirige a transcrição. Conforme usado aqui, um promotor inclui sequências de ácido nucleico necessárias próximas do sítio de início de transcrição tal como, no caso de um promotor do tipo polimerase II, um elemento TATA. Um promotor também inclui, opcionalmente, elementos distais, os quais podem estar localizados a centenas de milhares pares de base a partir do início de transcrição. Um promotor "constitutivo" é um promotor que é ativo sob a maioria das condições ambientais e de desenvolvimento. Um promotor "induzível" é um promotor que é ativo sob regulação ambiental ou de desenvolvimento. O termo "operavelmente ligado" se refere a uma ligação funcional entre uma sequência de controle de expressão de ácido nucleico (tal como um promotor) e uma segunda sequência de ácido nucleico, tal como um gene *IspS*, em que a sequência de controle de expressão dirige a transcrição do ácido nucleico correspondendo à segunda sequência. Um "promotor de alga" ou "promotor bacteriano" é um promotor capaz de iniciar a transcrição em células de alga e/ou bactérias, respectivamente. Tal promotor é, portanto, ativo em células de microalgas, cianobactérias ou bactérias, mas não precisa se originar desse organismo. Deve ser entendido que modificações limitadas podem ser feitas sem destruir a função biológica de um elemento regulatório e que tais modificações limitadas podem resultar em elementos regulatórios de algas que têm função substancialmente equivalente ou intensificada quando comparado com um elemento regulatório de alga do tipo selvagem. Essas modificações podem ser deliberadas, como através de mutagênese sítio-dirigida ou podem ser acidentais, tal como através de mutação em hospedeiros abrigando o elemento regulatório. Todas as tais sequências de nucleotídeo modificadas são incluídas na definição de um elemento regulatório de alga, contanto que a capacidade de conferir expressão em algas verdes unicelulares seja substancialmente retida.

Atividade ou expressão "aumentada" ou "intensificada" de um gene *Dxs* ou *Dxr* se refere a uma alteração na atividade de *Dxs* ou *Dxr*. Exemplos de tal atividade ou expressão aumentada incluem os seguintes. Atividade ou expressão de *Dxs* ou *Dxr* de um gene *Dxs* ou *Dxr* está aumentada
5 acima do nível daquele do microorganismo de controle não-transgênico do tipo selvagem (isto é, a quantidade de atividade de *Dxs* ou *Dxr* ou expressão do gene *Dxs* ou *Dxr* está aumentada). Atividade de *Dxs* ou *Dxr* ou expressão de um gene *Dxs* ou *Dxr* é em uma célula onde normalmente ela não é detectada em células não-transgênicas do tipo selvagem (isto é, expressão do
10 gene *Dxs* ou *Dxr* está aumentada). Atividade ou expressão de *Dxs* ou *Dxr* também está aumentada quando atividade de *Dxs* ou *Dxr* ou expressão do gene *Dxs* ou *Dxr* está presente em uma célula durante um período de tempo mais longo do que em um controle não-transgênico do tipo selvagem (isto é, duração de atividade de *Dxs* ou *Dxr* ou expressão do gene *Dxs* ou *Dxr* está
15 aumentada).

"Expressão" de um gene *lspS*, no contexto da presente invenção, se refere, tipicamente, à introdução de um gene *lspS* em uma célula, por exemplo, microalga, tal como microalga verde, cianobactérias ou bactérias fotossintéticas ou não-fotossintéticas, na qual normalmente ele não é
20 expresso. Consequentemente, um "aumento" na atividade ou expressão de *lspS* é geralmente determinado com relação à células do tipo selvagem, por exemplo, microalgas, cianobactérias ou bactérias fotossintéticas ou não-fotossintéticas, que não têm atividade de *lspS*.

Uma sequência de polinucleotídeo é "heteróloga" a uma segunda sequência de polinucleotídeo se ela se origina de uma espécie estranha ou, se da mesma espécie, é modificada pela ação humana com relação à sua forma original. Por exemplo, um promotor operavelmente ligado a uma sequência de codificação heteróloga se refere a uma sequência de codificação de uma espécie diferente daquela da qual o promotor foi derivado ou, se
25 da mesma espécie, uma sequência de codificação que é diferente de quaisquer variantes alélicas que ocorrem naturalmente.

Um "polinucleotídeo de *lspS*" é uma sequência de ácido nucleico

de SEQ ID NO: 1 ou SEQ ID NO: 7 ou as regiões de codificação de IspS de SEQ ID NO: 3 ou SEQ ID NO: 5; ou uma sequência de ácido nucleico que é substancialmente similar à SEQ ID NO: 1 ou as regiões de codificação de IspS de SEQ ID NO: 3 ou SEQ ID NO: 5; ou uma sequência de ácido nucleico que codifica um polipeptídeo de SEQ ID NO: 2 ou SEQ ID NO: 8 ou um polipeptídeo que é substancialmente similar à SEQ ID NO: 2 ou SEQ ID NO: 8 ou um fragmento ou domínio do mesmo. Assim, um polinucleotídeo de IspS: 1) compreende uma região de cerca de 15 a cerca de 50, 100, 150, 200, 300, 500, 1.000, 1.500 ou 2.000 ou mais nucleotídeos, algumas vezes de cerca de 20 ou cerca de 50 a cerca de 1800 nucleotídeos e, algumas vezes, de cerca de 200 a cerca de 600 ou cerca de 1500 nucleotídeos de SEQ ID NO: 1 ou SEQ ID NO: 7 ou a região de codificação de IspS de SEQ ID NOs: 3 ou 5; ou 2) hibridiza à SEQ ID NO: 1 ou SEQ ID NO: 7 ou à região de codificação de IspS de SEQ ID NO: 3 ou SEQ ID NO: 5 ou aos complementos das mesmas, sob condições estridentes ou 3) codifica um polipeptídeo ou fragmento de IspS de pelo menos 50 aminoácidos contínuos, tipicamente de pelo menos 100, 150, 200, 250, 300, 350, 400, 450, 500 ou 550 ou mais resíduos contínuos de um polipeptídeo de IspS, por exemplo, SEQ ID NO: 2 ou SEQ ID NO: 8; ou 4) codifica um polipeptídeo ou fragmento de IspS que tem pelo menos 55%, frequentemente pelo menos 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% ou mais de identidade à SEQ ID NO: 2 ou SEQ ID NO: 8 ou ao longo de uma janela de comparação de pelo menos 100, 200, 300, 400, 500 ou 550 resíduos de aminoácido de SEQ ID NO: 2 ou SEQ ID NO: 8; ou 5) tem uma sequência de ácido nucleico que tem mais do que cerca de 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% ou mais de identidade de sequência de nucleotídeo à SEQ ID NO: 1 ou SEQ ID NO: 7, pelo menos 80%, 85%, 90% ou pelo menos 95%, 96%, 97%, 98%, 99% ou mais de identidade ao longo de uma janela de comparação de pelo menos cerca de 50, 100, 200, 500, 1000 ou mais nucleotídeos de SEQ ID NO: 1 ou SEQ ID NO: 7 ou a região de codificação de IspS de SEQ ID NO: 3 ou SEQ ID NO: 5; ou 6) é amplificada por iniciadores para SEQ ID NO: 1 ou SEQ ID NO: 7 ou a região de codificação de IspS de SEQ ID NO: 3 ou SEQ

ID NO: 5. O termo "polinucleotídeo de *IspS*" se refere a ácidos nucleicos fita simples ou fita dupla. Os ácidos nucleicos de *IspS* para uso na invenção codificam uma *IspS* ativa que catalisa a conversão de um substrato dimetilalil difosfato em isopreno.

5 Um "polipeptídeo de *IspS*" é uma sequência de aminoácido que tem a sequência de aminoácido de SEQ ID NO: 2 ou SEQ ID NO: 8 ou é substancialmente similar à SEQ ID NO: 2 ou SEQ ID NO: 8 ou um fragmento ou domínio do mesmo. Assim, um polipeptídeo de *IspS* pode: 1) ter pelo menos 55% de identidade, tipicamente pelo menos 60%, 65%, 70%, 75%,
10 80%, 85%, 90%, 95% ou mais de identidade à SEQ ID NO: 2 ou SEQ ID NO: 8 ou ao longo de uma janela de comparação de pelo menos 100, tipicamente pelo menos 200, 250, 300, 350, 400, 450, 500, 550 ou mais aminoácidos de SEQ ID NO: 2 ou SEQ ID NO: 8; ou 2) compreender pelo menos 100, tipicamente pelo menos 200, 250, 300, 350, 400, 450, 500, 550 aminoácidos
15 contínuos de SEQ ID NOs: 2 ou 8; ou 3) se ligar a anticorpos criados contra um imunógeno compreendendo uma sequência de aminoácido de SEQ ID NOs: 2 ou 8 e variantes conservativamente modificadas das mesmas. Um polipeptídeo de *IspS*, no contexto da presente invenção, é uma proteína funcional que catalisa a conversão do substrato dimetilalil fosfato em isopreno.

20 Conforme usado aqui, um homólogo ou ortólogo de um gene *IspS* (por exemplo, SEQ ID NO: 1) é um segundo gene no mesmo tipo de planta ou em um tipo de planta diferente que é substancialmente idêntico (determinado conforme descrito abaixo) a uma sequência no primeiro gene.

Os termos ácidos nucleicos e polipeptídeo de "Dxs" e "Dxr" se referem a fragmentos, variantes e semelhantes. Sequências de Dxs e Dxr
25 exemplares incluem as sequências de ácido nucleico e polipeptídeo de Dxs e Dxr reveladas na Publicação de Pedido de Patente U.S. Nº 20030219798, por exemplo, sequências de *Chlamydomonas*. As sequências de Dxs e Dxr da Publicação de Pedido de Patente U.S. Nº 20030219798 são aqui incorpo-
30 radas por referência.

Um "cassete de expressão" se refere a um construto de ácido nucleico o qual, quando introduzido em uma célula hospedeira, resulta em

transcrição e/ou tradução de um RNA ou polipeptídeo, respectivamente.

No caso da expressão de transgenes, aqueles versados na técnica reconhecerão que a sequência de polinucleotídeo inserida pode não ser idêntica e pode ser "substancialmente idêntica" a uma sequência do gene do qual ela foi derivada. Conforme explicado abaixo, essas variantes são especificamente abrangidas pelo termo.

No caso onde a sequência de polinucleotídeo inserida é transcrita e traduzida para produzir um polipeptídeo funcional, aqueles versados na técnica reconhecerão que, em virtude de degeneração de códon, uma série de sequências de polinucleotídeo codificarão o mesmo polipeptídeo. Essas variantes são especificamente abrangidas pelo termo "sequência de polinucleotídeo de *lspS*" ou "gene *lspS*".

Duas sequências de ácido nucleico ou polipeptídeos são mencionados como sendo "idênticos" se a sequência de nucleotídeo ou resíduos de aminoácido, respectivamente, nas duas sequências for a mesma quando alinhada para correspondência máxima, conforme descrito abaixo. O termo "complementar a" é usado aqui para significar que a sequência é complementar a toda ou uma porção de uma sequência de polinucleotídeo de referência.

Alinhamento ótimo de sequências para comparação pode ser conduzido através do algoritmo de homologia local de Smith e Waterman, *Add. Apl. Math.* 2: 482 (1981), através do algoritmo de alinhamento por homologia de Needleman e Wunsch, *J. Mol. Biol.* 48: 443 (1970), através do método de busca por similaridade de Pearson e Lipman *Proc. Natl. Acad. Sci.* (U.S.A.) 85: 2444 (1988), através de implementações computadorizadas desses algoritmos (GAP, BESTFIT, BLAST, FASTA e TFASTA no Wisconsin Genetics Software Package, Genetics Computer Group (GCG), 575 Science Dr., Madison, WI) ou através de inspeção.

"Percentual de identidade de sequência" é determinado comparando duas sequências otimamente alinhadas ao longo uma janela de comparação, em que a porção da sequência de polinucleotídeo na janela de comparação pode compreender adições ou deleções (isto é, lacunas) quan-

do comparada com a sequência de referência (a qual não compreende adições ou deleções) para alinhamento ótimo das duas sequências. O percentual é calculado determinando o número de posições nas quais a base de ácido nucleico ou resíduo de aminoácido idêntico ocorre em ambas as sequências para proporcionar o número de posições pareadas, dividindo o número de posições pareadas pelo número total de posições na janela de comparação e multiplicando o resultado por 100 para proporcionar o percentual de identidade de sequência. Uma "janela de comparação", conforme usado aqui, inclui referência a um segmento de qualquer um da série de posições contíguas, por exemplo, 20 a 600, usualmente cerca de 50 a cerca de 200, mais usualmente cerca de 100 a cerca de 150, na qual uma sequência pode ser comparada com uma sequência de referência do mesmo número de posições contínuas após as duas sequências serem otimamente alinhadas.

O termo "identidade substancial", no contexto de sequências de polinucleotídeo ou aminoácido, significa que um polinucleotídeo ou polipeptídeo compreende uma sequência que tem pelo menos 50% de identidade de sequência a uma sequência de referência. Alternativamente, o percentual de identidade pode ser qualquer número inteiro de 50% a 100%. Modalidades exemplares incluem pelo menos: 55%, 57%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% ou 99% de identidade comparado com uma sequência de referência usando os programas descritos aqui; de preferência, BLAST usando parâmetros predeterminados, conforme descrito abaixo. Consequentemente, as sequências de IspS da invenção incluem sequências de ácido nucleico que têm identidade substancial à SEQ ID NO: 1 ou SEQ ID NO: 7 ou às regiões de codificação de IspS de SEQ ID NO: 3 ou SEQ ID NO: 5. Conforme mencionado acima, as sequências de polipeptídeo de IspS da invenção incluem sequências de polipeptídeo tendo identidade substancial à SEQ ID NO: 2 ou SEQ ID NO: 8.

Polipeptídeos que são "substancialmente similares" compartilham sequências conforme mencionado acima, exceto que posições de resíduo que não são idênticas podem diferir por alterações de aminoácido con-

servativas. Substituições de aminoácido conservativas se referem à capacidade de permuta de resíduos tendo cadeias laterais similares. Por exemplo, um grupo de aminoácidos tendo cadeias laterais alifáticas é glicina, alanina, valina, leucina e isoleucina; um grupo de aminoácidos tendo cadeias laterais de hidroxila alifática é serina e treonina; um grupo de aminoácidos tendo cadeias laterais contendo amida é asparagina e glutamina; um grupo de aminoácidos tendo cadeias laterais aromáticas é fenilalanina, tirosina e triptofano; um grupo de aminoácidos tendo cadeias laterais básicas é lisina, arginina e histidina; e um grupo de aminoácidos tendo cadeias laterais contendo enxofre é cisteína e metionina. Grupos de substituição de aminoácidos conservativos são: valina-leucina-isoleucina, fenilalanina-tirosina, lisina-arginina, alanina-valina, ácido aspártico-ácido glutâmico e asparagina-glutamina.

Outra indicação de que sequências de nucleotídeo são substancialmente idênticas é se duas moléculas hibridizam uma a outra ou a um terceiro ácido nucleico, sob condições estringentes. A expressão "condições de hibridização estringentes" se refere às condições sob as quais uma sonda hibridizará à sua sub-sequência alvo, tipicamente em uma mistura complexa de ácido nucleico, mas não a outras sequências. Condições estringentes são dependentes da sequência e serão diferentes em diferentes circunstâncias. Sequências mais longas hibridizam especificamente em temperaturas maiores. Um guia extensivo para a hibridização de ácidos nucleicos é encontrado em Tijssen, *Techniques in Biochemistry and Molecular Biology-Hybridization with Nucleic Probes*, "Overview of principles of hybridization and the strategy of nucleic acid assays" (1993). Geralmente, condições estringentes são selecionadas para ser cerca de 5 a 10°C menor do que o ponto de fusão térmica (T_m) para a sequência específica em um pH e força iônica definidos. A T_m é a temperatura (sob força iônica, pH e concentração de sal definidos) na qual 50% das sondas complementares ao alvo hibridizam à sequência alvo em equilíbrio (uma vez que as sequências alvo estão presentes em excesso, na T_m , 50% das sondas são ocupadas em equilíbrio). Condições estringentes serão aquelas nas quais a concentração de sal é menor do que cerca de 1,0 M de íon de sódio, tipicamente uma concentração de íons de

sódio de 0,01 a 1,0 M (ou outros sais) em um pH de 7,0 a 8,3 e a temperatura é pelo menos cerca de 30°C para sondas curtas (por exemplo, 10 a 50 nucleotídeos) e pelo menos cerca de 60°C para sondas longas (por exemplo, mais de 50 nucleotídeos). Condições estridentes também podem ser obtidas com a adição de agentes de desestabilização, tais como formamida. Para hibridização seletiva e específica, um sinal positivo é pelo menos duas vezes a hibridização de base, opcionalmente 10 vezes a hibridização de base. Condições de hibridização estridentes exemplares podem ser como segue: formamida a 50%, 5X SSC e SDS a 1%, incubação a 42°C ou 5X SSC, SDS a 1%, incubação a 65°C, com lavagem em 0,2X SSC e SDS a 0,1% a 55°C, 60°C ou 65°C. Tais lavagens podem ser realizadas durante 5, 15, 30, 60, 120 ou mais minutos.

Ácidos nucleicos que não hibridizam uns aos outros sob condições estridentes ainda são substancialmente idênticos se os polipeptídeos que eles codificam são substancialmente idênticos. Isso ocorre, por exemplo, quando uma cópia de um ácido nucleico é criada usando a degeneração de códon máxima permitida pelo código genético. Em tais casos, os ácidos nucleicos, tipicamente, hibridizam sob condições de hibridização moderadamente estridentes. Por exemplo, um polinucleotídeo de IspS também pode ser identificado por sua capacidade de hibridizar, sob condições estridentes (por exemplo, $T_m \sim 40^\circ\text{C}$) às sondas de ácido nucleico tendo a sequência de SEQ ID NO: 1 ou SEQ ID NO: 7. Tal sequência de ácido nucleico de IspS pode ter, por exemplo, cerca de 25 a 30% pareamentos errôneos de par de base ou menos com relação à sonda de ácido nucleico selecionada. SEQ ID NO: 1 é uma sequência de polinucleotídeo de IspS exemplar. "Condições de hibridização moderadamente estridentes" exemplares incluem uma hibridização em um tampão de formamida a 40%, NaCl a 1 M, SDS a 1% a 37°C e uma lavagem em SSC 1X a 45°C. Tais lavagens podem ser realizadas durante 5, 15, 30, 60, 120 ou mais minutos. Uma hibridização positiva é pelo menos duas vezes a de base. Aqueles versados na técnica reconhecerão que condições alternativas de hibridização e lavagem podem ser utilizadas para proporcionar condições de estridência similar.

O termo "isolado", quando aplicado a um ácido nucleico ou proteína, denota que o ácido nucleico ou proteína é essencialmente isento de outros componentes celulares com os quais ela está associada no estado natural. Ela está, de preferência, em um estado homogêneo e pode estar
5 seca ou em uma solução aquosa. A pureza e homogeneidade são, tipicamente, determinadas usando técnicas de química analítica, tais como eletroforese em gel de poliacrilamida ou cromatografia líquida de alta eficiência desempenho. Uma proteína que é a espécie predominante presente em um preparado é substancialmente purificada. Em particular, um gene isolado é
10 separado de fases abertas de leitura que flanqueiam o gene e codificam uma outra proteína que não o gene de interesse.

Conforme usado aqui, "cultura em massa" se refere ao crescimento de grandes quantidades de microalgas, cianobactérias ou bactérias fotossintéticas ou não-fotossintéticas que tenham sido modificadas para expressar um gene *IspS*. Uma "grande quantidade" está, geralmente, na faixa
15 de cerca de 100 litros a cerca de 1.500.000 litros ou mais. Em algumas modalidades, os organismos são cultivados em grandes quantidades em biorreatores modulares, cada um tendo uma capacidade de cerca de 1.000 a cerca de 1.000.000 litros.

Um "bioreator", no contexto da presente invenção, é qualquer recipiente de grande capacidade fechado no qual microalgas, cianobactérias ou bactérias fotossintéticas ou não-fotossintéticas são cultivadas. Um "recipiente de grande capacidade", no contexto da presente invenção, pode conter cerca de 100 litros, frequentemente cerca de 500 litros ou cerca de 1.000
25 litros a cerca de 1.000.000 litros ou mais.

Conforme usado aqui, "coleta" de hidrocarbonetos de isopreno voláteis se refere à captura e sequestro de tais hidrocarbonetos em um ambiente fechado ou contido.

Sequências de ácido nucleico de *IspS*, *Dxr* ou *Dxs*

30 A invenção emprega várias técnicas de ácido nucleico recombinante de rotina. Geralmente, a nomenclatura e os procedimentos de laboratório na tecnologia de DNA recombinante descritos abaixo são aqueles bem

conhecidos e comumente empregados na técnica. Muitos manuais que fornecem orientação para realização de manipulações de DNA recombinante estão disponíveis, por exemplo, Sambrook & Russell, *Molecular Cloning, A Laboratory Manual* (3ª Ed, 2001); e *Current Protocols in Molecular Biology* (Ausubel e colaboradores, eds., 1994-1999).

Sequências de polipeptídeo e ácido nucleico de IspS são conhecidas na técnica. Genes de IspS foram isolados e sequenciados de álamo e aspen (duas árvores relacionadas) e kudzu (uma uva). As espécies envolvidas e as sequências disponíveis no banco de dados do NCBI são fornecidas abaixo pelo número de acesso, cada uma das quais é incorporada por referência:

mRNA de IspS para isopreno sintase de *Populus alba* (álamo branco): ACESSO Nº AB198180;

isopreno sintase (IspS) de *Populus tremuloides* (quaking aspen): ACESSO Nº AY341431 (cds completa);

mRNA de IspS de *Populus alba* x *Populus tremula*: ACESSO Nº AJ294819;

mRNA para isopreno sintase (gene IspS) de *Populus nigra* (álamo Lombardy): ACESSO Nº AM410988;

isopreno sintase (IspS) de *Pueraria montana* var. lobata (uva kudzu): ACESSO Nº AY316691 (cds completa).

O exame dessas sequências de IspS revela um alto grau de idêntidades de sequência de nucleotídeo e aminoácido, por exemplo, sequências de cDNA de álamo e aspen híbridas são 98% idênticas a nível de polipeptídeo e nucleotídeo (vide, por exemplo, Sharkey et al., *Plant Physiol.* 137: 700-712, 1995). A sequência de nucleotídeo de codificação de isopreno sintase de aspen é 65% idêntica ao gene de kudzu, enquanto que as sequências de proteína (sem o peptídeo de trânsito de cloroplasto) são 57% idênticas.

A proteína de IspS de álamo tem uma alta densidade de aminoácidos cisteína e histidina na metade carboxi-terminal da proteína. Por exemplo, considerando a sequência de 591 aminoácidos da proteína Cr-IspS

(SEQ ID NO: 4), porções de cisteína são encontradas nas posições 34, 326, 378, 413, 484, 505 e 559, isto é, seis das sete cisteínas são encontradas nos 45% inferiores da proteína. Agrupamento adicional de histidinas em várias posições da metade C-terminal da proteína é também observado. Aminoácidos cisteína e histidina são conhecidos por participar no enovelamento apropriado e na estrutura do sítio catalítico de proteínas e podem ser componentes importantes para a atividade enzimática. Um alinhamento de quatro proteínas de IspS conhecidas mostrando alta conservação de Cys na parte C-terminal da molécula é fornecido na figura 13. Em um caso, a proteína de kudzu tem substituído uma sequência de Cys de outra forma conservada por Ser (Cys-509-Ser de *Alba* ou *nigra* ou *tremuloides*) no alinhamento pelo Clustal na figura 13. Serina é uma substituição altamente conservativa para cisteína, uma vez que a única diferença entre os dois aminoácidos é um grupo -OH no lugar do grupo -SH. O fato, o exame das quatro sequências de IspS revela a propriedade adicional de muitas Serinas conservadas na metade C-terminal da proteína. Consequentemente, em algumas modalidades, um ácido nucleico para uso na invenção codifica um polipeptídeo de IspS que compreende os 45% carbóxi-terminais de SEQ ID NO: 2 e retém a atividade catalítica na conversão de DMAPP em isopreno. Existem outros exemplos onde uma proteína relacionada em um microorganismo, tal como uma microalga verde, carece de uma porção substancial da porção N-terminal da proteína (com relação à forma da proteína presente em outro microorganismo, tal como bactérias) sem efeito adverso sobre a atividade (vide, por exemplo, Melis e Happe, *Plant Physiol.* 127: 740-748, 2001). Consequentemente, em algumas modalidades, um ácido nucleico de IspS para uso na invenção codifica um polipeptídeo que compreende de cerca do resíduo de aminoácido 330 até a terminação-C de SEQ ID NO: 2 ou SEQ ID NO: 8. Em algumas modalidades, o polipeptídeo de IspS codificado pelo ácido nucleico de IspS compreende de cerca do resíduo de aminoácido 300 à terminação-C de SEQ ID NO: 2 ou SEQ ID NO: 8. Em algumas modalidades, a sequência de IspS pode, adicionalmente, carecer dos últimos 10, 15 ou 20 resíduos de SEQ ID NO: 2 ou SEQ ID NO: 8.

- O peptídeo de trânsito da proteína de IspS inclui, minimamente, os aminoácidos 1 a 37 para álamo e aspen e 1 a 45 para kudzu. Com base nessa análise, a proteína madura começa com a sequência de aminoácido CSVSTEN...", etc. Sequências de ácido nucleico de IspS para uso na inven-
- 5 ção não precisam incluir sequências que codificam um polipeptídeo de trânsito e ainda omitem a sequência N-terminal adicional. Por exemplo, o constructo Ss-IspS apresentado na seção EXEMPLOS carece de 52 aminoácidos a partir do DNA do gene sintético de codificação. Isso não teve efeito sobre a síntese e acúmulo de proteína de IspS.
- 10 Em algumas modalidades da invenção, uma sequência de ácido nucleico que codifica um polipeptídeo de IspS de álamo ou aspen (por exemplo, SEQ ID NO: 2) é usada. Em outras modalidades, uma sequência de ácido nucleico que codifica um polipeptídeo de IspS de kudzu (por exemplo, SEQ ID NO: 8) é usada. Os polipeptídeos de IspS codificados pelos ácidos
- 15 nucleicos empregados nos métodos da invenção têm a atividade catalítica de conversão de DMAPP em isopreno. Tipicamente, o nível de atividade é equivalente à atividade exibida por um polipeptídeo de IspS de álamo ou aspen (por exemplo, codificado por SEQ ID NO: 1) ou um polipeptídeo de IspS de kudzu natural (por exemplo, codificado por SEQ ID NO: 7).
- 20 Sequências de Dxs e Dxr exemplares incluem as sequências de ácido nucleico e polipeptídeo de Dxs e Dxr reveladas na Publicação de Pedido de Patente U.S. Nº 20030219798, por exemplo, sequências de *C. hlamydomonas*. As sequências de Dxs e Dxr da Publicação de Pedido de Patente U.S. Nº 20030219798 são aqui incorporadas por referência.
- 25 Isolamento ou geração de sequências de polinucleotídeo de *IspS*, *Dxr* ou *Dxs* pode ser obtida através de uma série de técnicas. Clonagem e expressão de tal técnica será dirigida ao contexto de genes de IspS. Contudo, as mesmas técnicas podem ser usadas para isolar e expressar polinucleotídeos de Dxr ou Dxs. Por exemplo, sondas de oligonucleotídeo
- 30 baseadas nas sequências divulgadas aqui podem ser usadas para identificar o polinucleotídeo desejado em uma biblioteca de cDNA ou DNA genômico a partir de uma espécie de planta desejada. Talm biblioteca de cDNA ou ge-

nômica pode então ser pesquisada usando uma sonda baseada na sequência de um gene de IspS clonado, por exemplo, SEQ ID NO: 1 ou SEQ ID NO: 7. Sondas podem ser usadas para hibridizar com sequências de DNA genômico ou cDNA para isolar genes homólogos na mesma ou diferentes espécies de planta.

Alternativamente, os ácidos nucleicos de interesse podem ser amplificados a partir de amostras de ácido nucleico usando técnicas de amplificação. Por exemplo, PCR pode ser usada para amplificar as sequências dos genes diretamente a partir de mRNA, de cDNA, de bibliotecas genômicas ou bibliotecas de cDNA. PCR e outros métodos de amplificação *in vitro* também podem ser úteis, por exemplo, para clonar sequências de ácido nucleico que codificam proteínas a serem expressas, para fazer ácidos nucleicos para uso como sondas para detecção da presença do mRNA desejado em amostras, para sequenciamento de ácido nucleico ou para outras finalidades.

Iniciadores e sondas apropriadas para identificação de um gene de IspS de células de planta, por exemplo, álamo ou outra árvore decídua, podem ser gerados a partir de comparações das sequências fornecidas aqui. Para uma visão geral de PCR veja PCR Protocols: A Guide to Methods and Applications (Innis, M, Gelfand, D., Sninsky, J. e White, T., eds.), Academic Press, San Diego (1990). Uma PCR exemplar para amplificação de uma sequência de ácido nucleico de IspS é fornecida nos exemplos.

O gênero das sequências de ácido nucleico de IspS para uso na invenção inclui genes e produtos genéticos identificados e caracterizados através de técnicas tais como hibridização e/ou análise de sequência usando sequências de ácido nucleico exemplares, por exemplo, SEQ ID NO: 1 ou SEQ ID NO: 7 e sequências de proteína, por exemplo, SEQ ID NO: 2 ou SEQ ID NO: 8.

Preparo de vetores recombinantes

Para usar sequências isoladas nas técnicas acima, vetores de DNA recombinantes adequados para transformação de células de microalgas verdes, cianobactérias e bactérias fotossintéticas ou não fotossintéticas

são preparados. Técnicas para transformação são bem conhecidas e descritas na literatura técnica e científica. Por exemplo, uma sequência de DNA que codifica um gene *IspS* (descrito em maiores detalhes abaixo) pode ser combinada com sequências regulatórias de transcrição e outras as que dirigirão a transcrição da sequência a partir do gene nas células pretendidas das algas, cianobactérias ou bactérias transformadas. Em algumas modalidades, um vetor de expressão que compreende um cassete de expressão que compreende o gene *IspS* ainda compreende um promotor operavelmente ligado ao gene *IspS*. Em outras modalidades, um promotor e/ou outros elementos regulatórios que dirigem a transcrição do gene *IspS* são endógenos ao microorganismo e o cassete de expressão compreendendo o gene *IspS* é introduzido, por exemplo, através de recombinação homóloga, de modo que o gene *IspS* heterólogo esteja operavelmente ligado a um promotor endógeno e a expressão é dirigida pelo promotor endógeno.

Sequências regulatórias incluem promotores, os quais podem ser constitutivos ou induzíveis. Em algumas modalidades, um promotor pode ser usado para dirigir a expressão de ácidos nucleicos de *IspS* sob a influência de condições ambientais capazes de modificação. Exemplos de condições ambientais que podem realizar transcrição através de promotores induzíveis incluem condições anaeróbicas, temperatura elevada ou a presença de luz. Promotores que são induzíveis pela exposição a reagentes químicos também são usados para expressar ácidos nucleicos de *IspS*. Outros elementos regulatórios induzíveis incluem elementos regulatórios induzíveis por cobre (Mett et al., *Proc. Natl. Acad. Sci. USA* 90: 4567-4571 (1993); Furst et al., *Cell* 55: 705-717 (1988)); elementos regulatórios tetraciclina- e clortetraciclina-induzíveis (Gatz et al., *Plant J.* 2: 397-404 (1992); Roder et al., *Mol. Gen. Genet.* 243: 32-38 (1994); Gatz, *Meth. Cell Biol.* 50: 411-424(1995)); elementos regulatórios induzíveis por ecdisona (Christopherson et al., *Proc. Natl. Acad. Sci. USA* 89: 6314-6318 (1992); Kreutzweiser et al., *Ecotoxicol. Environ. Safety* 28: 14-24 (1994)); elementos regulatórios induzíveis por choque térmico (Takahashi et al., *Plant Physiol.* 99: 383-390 (1992); Yabe et al., *Plant Cell Physiol.* 35: 1207-1219 (1994); Ueda et al., *Mol. Gen.*

Genet. 250: 533-539 (1996)); e elementos do óperon lac, os quais são usados em combinação com um repressor lac constitutivamente expresso para conferir, por exemplo, expressão induzível por IPTG (Wilde e colaboradores, *EMBO J.* 11: 1251-1259 (1992)). Um elemento regulatório induzível também
 5 pode ser, por exemplo, um promotor induzível por nitrato, por exemplo, derivado do gene de nitrato redutase de espinafre (Back et al., *Plant Mol. Biol.* 17: 9 (1991)) ou um promotor induzível pela luz, tal como aquele associado à pequena subunidade da RuBP carboxilase ou famílias do gene de LHCP (Feinbaum et al., *Mol. Gen. Genet.* 226: 449 (1991); Lam e Chua, *Science*
 10 248: 471 (1990)) ou luz.

Em um exemplo, uma sequência promotora que é responsiva à luz pode ser usada para direcionar a expressão de um constructo de ácido nucleico de *IspS* que é introduzido em *Chlamydomonas* que é exposto à luz (por exemplo Hahn, *Curr Genet* 34: 459-66, 1999; Loppes, *Plant Mol Biol* 45:
 15 215-27, 2001; Villand, *Biochem J.*, 327: 51-7, 1997). Outros sistemas de promotor induzível pela luz também podem ser usados, tal como o sistema de fitocromo/PIF3 (Shimizu-Sato, *Nat Biotechnol* 20:1041-4, 2002). Ainda, pode ser usado um promotor que é também responsivo ao calor para direcionar a expressão em algas, tal como *Chlamydomonas* (Muller, *Gene* 111:
 20 165-73, 1992; von Gromoff, *Mol Cell Biol* 9: 3911-8, 1989). Promotores adicionais, por exemplo, para expressão em algas, tais como microalgas verdes, incluem os promotores de *RbcS2* e *PsaD* (veja, por exemplo, Stevens et al., *Mol. Gen. Genet.* 251: 23-30, 1996; Fischer & Rochaix, *Mol Genet Genomics* 265: 888-94, 2001).

25 Em algumas modalidades, o promotor pode ser de um gene associado à fotossíntese na espécie a ser transformada ou outra espécie. Por exemplo, tal promotor de uma espécie pode ser usado para dirigir a expressão de uma proteína em células de alga transformadas ou células de outro organismo marinho fotossintético. Promotores adequados podem ser isolados de ou sintetizados com base em sequências conhecidas de outros organismos fotossintéticos. Promotores preferidos são aqueles para genes de
 30 outras espécies fotossintéticas que são homólogos aos genes fotossintéticos

do hospedeiro de alga a ser transformado. Por exemplo, uma série de promotores de coleta leve de proteína de ligação à clorofila de fucoxantina foram identificados em *Phaeodactylum tricornutum* (veja, por exemplo, Apt et al., *Mol Gen. Genet.* 252: 572-579, 1996). Em outras modalidades, um promotor da proteína de ligação à clorofila de carotenóide, tal como aquele da proteína de ligação à clorofila de peridininina, pode ser usado.

Em algumas modalidades, um promotor usado para direcionar a expressão de um gene *lspS* heterólogo é um promotor constitutivo. Exemplos de promotores fortes constitutivos para uso em microalgas incluem, por exemplo, os promotores dos genes *atpA*, *atpB* e *rbcL*. Vários promotores que são ativos em cianobactérias também são conhecidos. Esses incluem promotores tais como o promotor (constitutivo) do gene *psbA3* em cianobactérias e promotores tais como aqueles apresentados na Publicação de Pedido de Patente U.S. Nº 20020164706, a qual é incorporada por referência. Outros promotores que são operativos em plantas, por exemplo, promotores derivados de vírus de planta, tais como os promotores CaMV35S, também podem ser empregados em algas.

Em algumas modalidades, promotores são identificados através de análise das sequências 5' de um clone genômico correspondendo a um gene *lspS*. Características de sequência das sequências do promotor podem ser usadas para identificar o promotor.

Um promotor pode ser avaliado, por exemplo, testando a capacidade do promotor de acionar expressão em células de planta, por exemplo, algas verdes, nas quais se deseja introduzir uma estrutura de expressão de *lspS*.

Um vetor compreendendo sequências de ácido nucleico de *lspS* compreenderão, tipicamente, um gene marcador que confere um fenotipo selecionável à algas ou células bacterianas. Tais marcadores são conhecidos. Por exemplo, o marcador pode codificar resistência a antibiótico, tal como resistência à canamicina, G418, bleomicina, higromicina e semelhantes. Em algumas modalidades, marcadores selecionáveis para uso em *Chlamydomonas* podem ser marcadores que proporcionam resistência à es-

pectinomomicina (Fargo, *Mol Cell Biol* 19: 6980-90, 1999), resistência à canamicina e amicacina (Bateman, *Mol Gen Genet* 263: 404-10, 2000), resistência à zeomicina e fleomicina (Stevens, *Mol Gen Genet* 251: 23-30, 1996) e resistência à paramomicina e neomicina (Sizova, *Gene* 277: 221-9, 2001).

5 Sequências de ácido nucleico de *IspS* da invenção são expressas recombinantemente em microorganismos, por exemplo, microalgas, cianobactérias ou bactérias fotossintéticas ou não fotossintéticas. Conforme apreciado por aqueles versados na técnica, estruturas de expressão podem ser desenhados levando-se em conta propriedades, tais como frequências
10 de uso de códon do organismo no qual o ácido nucleico de *IspS* tem de ser expresso. As frequências de uso de códon podem ser calculadas usando métodos conhecidos (vide, por exemplo, Nakamura e colaboradores, *Nucl. Acids Res.* 28: 292, 2000). Tabelas de frequência de uso de códon, incluindo aquelas para microalgas e cianobactéria, também estão disponíveis na técnica (por exemplo, em bancos de dados de uso de códon do Department of
15 Plant Genome Research, Kazusa DNA Research Institute (www.kazusa.or.jp/codon)).

Métodos de transformação de célula e marcadores selecionáveis para bactérias e cianobactérias são bem conhecidos na técnica (Wirth, *Mol Gen Genet*, Março de 1989; 216(1): 175-7; Koksharova, *Appl Microbiol Biotechnol*, Fevereiro de 2002; 58(2): 123-37; Thelwell). Métodos de transformação e marcadores selecionáveis para uso em bactérias são bem conhecidos (vide, por exemplo, Sambrook et al., *supra*).

Em microalgas, por exemplo, microalgas verdes, os genomas
25 nuclear, mitocondrial e de cloroplasto são transformados através de uma variedade de métodos conhecidos, incluindo através de bombardeamento de partícula ou usando um método com pérolas vidro (veja, por exemplo, Kindle, *J. Cell Biol* 109: 2589-601, 1989; Kindle, *Proc. Natl. Acad. Sci. USA* 87: 1228-32, 1990; Kindle, *Proc. Natl. Acad. Sci. USA* 88: 1721-5, 1991; Shimo-
30 gawara, *Genetics* 148: 1821-8, 1998; Boynton, *Science* 240: 1534-8, 1988; Boynton, *Methods Enzymol* 264: 279-96, 1996; Randolph-Anderson, *Mol Gen Genet* 236: 235-44, 1993). Em algumas modalidades, um gene *IspS* é

introduzido no cloroplasto de uma microalga. Em outras modalidades, um gene *IspS* é introduzido no núcleo.

As técnicas descritas aqui para obtenção e expressão de sequências de ácido nucleico de *IspS* em microalgas, cianobactérias ou bactérias fotossintéticas ou não fotossintéticas também podem ser empregadas para expressar sequências de ácido nucleico de *Dxr* ou *Dxs*.

Microorganismos que podem ser atingidos

IspS pode ser expresso em uma série de microalgas, por exemplo, algas verdes, ou cianobactérias ou bactérias fotossintéticas ou não fotossintéticas onde é desejável produzir isopreno. Microalgas, cianobactérias ou bactérias (bactérias fotossintéticas ou bactérias não fotossintéticas) que expressam um gene *IspS* heterólogo são crescidas sob condições de cultura em massa para a produção de hidrocarbonetos, por exemplo, a serem usados como uma fonte de combustível ou como estoque de alimentação em química sintética. Os organismos transformados são crescidos em bioreatores ou fermentadores que proporcionam um ambiente fechado para conter os hidrocarbonetos. Em modalidades típicas para cultura em massa, as microalgas, cianobactérias ou bactérias são crescidas em bioreatores fechados em quantidades de pelo menos cerca de 500 litros, frequentemente de pelo menos cerca de 1000 litros ou mais e, em algumas modalidades, em quantidades de cerca de 1.000.000 litros ou mais.

Em algumas modalidades, *IspS* é expressa em microalgas. Algas, algae ou semelhante se referem às plantas pertencendo ao subfilo Algae do filo Thallophyta. As algas são algas plantas unicelulares, fotossintéticas, oxigênicas e são não parasíticas sem raízes, caules ou folhas; elas contêm clorofila e têm uma grande variedade de tamanho, de microscópicas a grandes algas marinhas. Algas verdes, as quais são organismos eucariotas unicelulares de fotossíntese oxigênica dotada de clorofila a e clorofila b pertencendo a Eukaryota --Viridiplantae--Chlorophyta-Chlorophyceae, são frequentemente um alvo preferido. Por exemplo, *IspS* pode ser expresso em *C. reinhardtii*, a qual é classificada como Volvocales--Chlamydomonadaceae. Gêneros de algas que podem ser usados na presente invenção incluem, por

exemplo, *Chlamydomonas reinhardtii*, *Scenedesmus obliquus*, *Chlorella vulgaris*, *Botryococcus braunii*, *Botryococcus sudeticus*, *Dunaliella salina* e *Haematococcus pluvialis*.

Métodos de cultura em massa de algas são conhecidos. Por exemplo, algas podem ser crescidas em fotobioreatores de alta densidade (veja, por exemplo, Lee et al., *Biotech. Bioengineering* 44: 1161-1167, 1994; Chaumont, *J. Appl. Phycology* 5: 593-604, 1990), bioreatores tais como aqueles para tratamento de águas potáveis e água residual (por exemplo, Sawayama et al., *Appl. Micro. Biotech.*, 41: 729-731, 1994; Lincoln, *Bulletin De L'institut Oceangraphique (Mônaco)*, 12: 109-115, 1993), cultura em massa para eliminação de metais pesados de água contaminada (por exemplo, Wilkinson, *Biotech. Letters*, 11: 861-864, 1989), cultura em massa para a produção de β -caroteno (por exemplo, Yamaoka, *Seibutsu-Kogaku Kaishi*, 72: 111-114, 1994), hidrogênio (por exemplo, Publicação de Pedido de Patente U.S. Nº 20030162273) e compostos farmacêuticos (por exemplo, Cannell, 1990), bem como suplementos nutricionais para seres humanos e animais (Becker, 1993, *Bulletin De L'institut Oceanographique (Mônaco)*, 12, 141-155) e para a produção de outros compostos de valor nutricional.

Condições para crescimento de algas ou bactérias expressando IspS para as finalidades exemplificativas ilustradas acima são conhecidas na técnica (vide, por exemplo, referências exemplares citadas aqui). Hidrocarbonetos de isopreno voláteis produzidos pelos microorganismos modificados podem ser coletados usando técnicas conhecidas. Hidrocarbonetos de isopreno são não miscíveis em água e eles se elevam e flutuam na superfície do meio de crescimento de microorganismo. Eles são drenados da superfície e capturados em recipientes adequados. Além disso e, dependendo da temperatura que prevalece durante a cultura em massa dos microorganismos, isopreno pode existir na forma de vapor acima do meio aquoso no recipiente do bioreator (temperatura de ebulição do isopreno, $T = 34^{\circ}\text{C}$). Vapor de isopreno é canalizado do recipiente do bioreator e condensado na forma de combustível líquido por resfriamento ou compressão em baixo nível.

EXEMPLOS

Os exemplos descritos aqui são fornecidos como forma de ilustração apenas e não como forma de limitação. Aqueles versados na técnica reconhecerão prontamente uma variedade de parâmetros não críticos que poderiam ser alterados ou modificados para proporcionar resultados essencialmente similares.

Exemplo 1. Desenho e expressão do novo gene de Cr-IspS para produção de hidrocarboneto de isopreno em microalgas

Um constructo de DNA sintético códon-ajustado foi gerado baseado na sequência de proteína de isopreno sintase (IspS) nuclear-codificada conhecida de *Populus alba* (álamo). Essa sequência de aminoácido (SEQ ID NO: 2) foi usada como um molde para o desenho *de novo* de uma sequência de DNA de *IspS* para expressão do gene no cloroplasto da microalga modelo *Chlamydomonas reinhardtii*. Para fins da presente invenção, esse gene foi denominado *Cr-IspS*. Características desse gene incluíam: (1) Uso de códon era diferente daquele do álamo e especificamente selecionado para se adaptar ao uso de códon do cloroplasto de *Chlamydomonas reinhardtii*; (2) A sequência de direcionamento de cloroplasto de álamo da proteína foi omitida do desenho do novo gene *Cr-IspS*; (3) Três cópias de um gene códon-otimizado que codifica o sinalizador de epítipo de hemaglutinina (HA) foram fundidas a montante do gene *IspS*.

A sequência de DNA de *Cr-IspS* (SEQ ID NO:3) foi desenhada para codificar a proteína de isopreno sintase (SEQ ID NO: 4) especificamente no cloroplasto de microalgas, por exemplo, *Chlamydomonas reinhardtii*. Ajustes de uso de códon para expressão do gene no cloroplasto de *Chlamydomonas* foram feitos com base na tabela de uso de códon para o cloroplasto 6803 de *Chlamydomonas reinhardtii*, listada na seguinte URL: <http://www.bio.net/bionet/mm/chlamy/1997-March/000843.html>.

SEQ ID NO:4 também contém três cópias do sinalizador de hemaglutinina, as quais estão sublinhadas no lado N-terminal da sequência. Sítios de reconhecimento de enzima de restrição foram introduzidos nas extremidades da sequência de DNA de *Cr-IspS* para facilitar a clonagem do gene e a sequência toda foi sintetizada e clonada em um plasmídeo trans-

portador.

Um cloroplasto de *Chlamydomonas reinhardtii* transgênico foi gerado, o qual expressa o gene de isopreno sintase recombinante códon-otimizado (*Cr-IspS*). Isso foi obtido construindo um gene quimérico (figura 4 superior, *Cr-IspS*) contendo o promotor de *atpA* (PatpA), fundido à extremidade 5'UTR de um DNA sinalizado com epítipo de hemaglutinina (HA) códon-otimizado com três cópias (figura 4). Essa sequência de DNA foi seguida pela região de codificação de *Cr-IspS* (figura 4), seguido pela 3'UTR de *atpA* (figura 4, TatpA). Integração do gene quimérico construído no genoma de cloroplasto de *Chlamydomonas reinhardtii* foi obtida usando transformação biolística e recombinação homóloga, requerendo homologia de sequência entre o vetor de transformação e o genoma do cloroplasto (Boynton e colaboradores, *Science*, 240: 1534-1538, 1988). Para essa finalidade, o vetor p322 foi empregado, o qual contém um genoma de cloroplasto de *C. reinhardtii* parcial para o alvo de recombinação homóloga (Franklin et al., *Plant J.*, 30: 733-744, 2002). Conforme mostrado no diagrama da figura 4, o gene *Cr-IspS* quimérico foi ligado ao sítio de BamIII de p322 para gerar o plasmídeo pApISAt. O constructo pApISAt foi co-transformado no cloroplasto da cepa CC503 de *C. reinhardtii* por meio de bombardeamento de partícula (Boynton et al., *Science*, 240: 1534-1538, 1988), junto com o plasmídeo p228, contendo um gene ribossômico 16S modificado que confere resistência à espectinomicina (Franklin et al., *Plant J.*, 30: 733-744, 2002). Os iniciadores N e C na figura 4 marcam o local de anelamento dos iniciadores que foram usados para a subsequente seleção por PCR entre transformantes resistentes à espectinomicina isolados.

A figura 5 proporciona um exemplo da análise através de seleção por PCR genômica de transformantes primários que foram selecionados em meio contendo espectinomicina com relação à presença de regiões N- ou C-terminais do gene quimérico *Cr-IspS* de forma a selecionar transformantes de *Cr-IspS* de *C. reinhardtii*. Mais de cem colônias de transformantes resistentes à espectinomicina de *Chlamydomonas reinhardtii* foram isoladas e testadas, dentre as quais descobriu-se que duas linhagens independentes

(Nº 9 e Nº 20) contêm inequivocadamente o gene *Cr-lspS* estavelmente integrado em seu DNA de cloroplasto. Um transformante resistente à espectinomomicina (Nº 7, não mostrado) foi usado como controle negativo para a análise por PCR e o plasmídeo pAplSAT serviu como um molde para o controle positivo. Os iniciadores N e C representam o conjunto de primer usado para amplificação e seus locais de anelamento são mostrados na figura 4.

DNA genômico do controle *Chlamydomonas reinhardtii* (Nº 7) e a linhagem transformante de *Cr-lspS* putativa Nº 9 foram digeridos com BamHI, separados em um gel de Agarose e submetidos à análise por Southern blot de forma a testar a integridade do transgene *Cr-lspS*. Hibridização com um fragmento de NdeI/XbaI radio marcado da região de codificação de *Cr-lspS* identificou uma banda de ~ 3,0 kbp exclusivamente na linhagem de transformante Nº 9 de *Cr-lspS*, enquanto que nenhuma banda detectável pôde ser observada na fileira da linhagem de controle Nº 7 (figura 6A). Esses resultados validaram a integração estável de *Cr-lspS* no genoma de cloroplasto da linhagem de transformante Nº 9 de *Chlamydomonas reinhardtii* e são consistentes com os resultados da análise por PCR (figura 5). Coloração com brometo de etídio do gel de agarose (figura 6B) testou a quantidade igual de aplicação de DNA. Resultados similares foram obtidos com a linhagem transformante Nº 20 de *Cr-lspS* (não mostrado).

Acúmulo de proteína Cr-lspS na linhagem transgênica de *Chlamydomonas reinhardtii* Nº 9 foi verificado através de análise por Western blot (figura 7) de forma a demonstrar expressão do gene *Cr-lspS*. Anticorpo anti-sinalizador HA (a-HA) foi usado para avaliar a presença da proteína Cr-lspS recombinante e sua concentração celular. Três cópias de tag de hemaglutinina (HA) foram introduzidas em uma posição precedendo o gene *Cr-lspS* que codifica a proteína madura (figura 7A), para servir como um epítipo conveniente para a detecção de acúmulo de proteína Cr-lspS. Células de *Chlamydomonas reinhardtii* foram concentradas para 500×10^6 células/ml em tampão HEPES a 50 mM (pH de 7,0) e rompidas através de agitação com glóbulos de vidro durante 5 min para liberar a fração solúvel de cloroplasto. Frações de proteína solúvel, as quais correspondem a 10 ou 20 µg

de clorofila, foram submetidas a SDS-PAGE e análise por Western blot com anticorpos policlonais anti-HA específicos. Uma reação cruzada de anticorpo-proteína clara foi observada na banda de cerca de 67 kD nas fileiras aplicadas com amostra da linhagem N° 9 do transformante de *Chlamydomonas reinhardtii*, mas não no controle (C N° 7) (figura 7B). Além disso, reações cruzadas de anticorpo-proteína foram observadas em aproximadamente 38 kD, indicado por um asterisco na figura 5B. Acúmulo de proteína Cr-IspS como uma banda de 38 kD poderia indicar um término prematuro de tradução de mRNA de *Cr-IspS* ou uma atividade de degradação específica sobre a proteína recombinante. Não havia bandas de 67 kD ou 38 kD detectáveis na fileira de controle (C N° 7). A aparente reação cruzada correspondendo a uma proteína de 34 kD provavelmente é uma ligação inespecífica do anticorpo primário ou secundário a uma proteína de *Chlamydomonas reinhardtii*. Expressão da proteína Cr-IspS também foi detectada na linhagem de transformante N° 20 (não mostrado).

Exemplo 2. Desenho e expressão de um gene *Ss-IspS* para produção de hidrocarboneto isopreno em cianobactérias

De forma a expressar produção de hidrocarboneto isopreno em cianobactérias, uma estrutura de DNA sintética códon-ajustada foi gerada, com base na sequência de proteína de isopreno sintase IspS conhecida de *Populus alba* (álamo). Essa sequência de aminoácido foi usada como um molde para o desenho *de novo* de uma sequência de DNA de *IspS* para expressão do gene em cianobactérias, por exemplo, *Synechocystis* sp. Ajustes de uso de códon para expressão do gene em cianobactérias foram feitos com base na Tabela de uso de códon para *Synechocystis* PCC 6803, listada na seguinte URL: [http://gib.genes.nig.ac.jp/single/codon/main.php?spid= Syne_PCC6803](http://gib.genes.nig.ac.jp/single/codon/main.php?spid=Syne_PCC6803).

O gene códon-ajustado é referido aqui como *Ss-IspS*. Características desse gene incluem: (1) Uso de códon era diferente daquele do álamo e especificamente selecionado para se adaptar ao uso de códon de *Synechocystis*; (2) A sequência de direcionamento de cloroplasto de álamo da proteína foi omitida do desenho do novo gene *Ss-IspS*. A sequência de DNA

foi desenhada para codificar a proteína de isopreno sintase especificamente em cianobactérias, por exemplo, *Synechocystis*. A primeira sequência sublinhada de SEQ ID NO: 5 representa o gene de beta-lactamase (o complemento reverso), enquanto que a segunda sequência sublinhada é o DNA de *Ss-IspS*. Adicionalmente, as sequências em itálico são códons inicial e terminal e as sequências em negrito são sítios de restrição para clonagem. Sítios de reconhecimento de enzima de restrição foram introduzidos nas extremidades da sequência de DNA de *Ss-IspS* recentemente desenhada para facilitar a clonagem do gene e toda a sequência foi sintetizada e clonada em um plasmídeo transportador (figura 8).

O gene *Ss-IspS* códon-otimizado, com comprimento ajustado e quimicamente sintetizado foi clonado à jusante da região do promotor *psbA3* de *Synechocystis*, em fase com o códon inicial ATG do gene *psbA3*. O gene *Ss-IspS* foi seguido por um terminador de transcrição e um cassete de resistência à gentamicina e, após, pela sequência de *Synechocystis* imediatamente à jusante do gene *psbA3* (figura 8).

Essa novo constructo permitiu recombinação homóloga, isto é, inserção da sequência de DNA de *Ss-IspS* no genoma de *Synechocystis* através de substituição do gene *psbA3* endógeno via recombinação homóloga (figura 10). Seleção de transformantes de *Synechocystis* pôde ser feita usando gentamicina (Gm) como o marcador selecionável e o promotor forte *psbA3* direciona a expressão do gene *Ss-IspS*.

De forma a transformar *Synechocystis* com o constructo de *Ss-IspS*, células de *Synechocystis* sp. foram crescidas em um meio de culturas BG11 básico na presença de glicose a 5 mM, até a densidade celular atingir cerca de 50×10^6 células/ml⁻¹ (OD₇₃₀ = 0,5). As células foram, então, coletadas e concentradas para 10^{10} células/ml⁻¹, misturadas com o plasmídeo pAIGA para transformação e incubadas durante 4 a 6 h antes de dispersão da mistura sobre filtros por cima de lâminas de Ágar contendo BG11, também contendo 0,5 µg/ml de Gm, tiosulfato de sódio a 0,3% e TES-NaOH a 10 mM, pH de 8,0.

As placas de Petri foram mantidas sob baixa intensidade de luz

durante 1 a 2 dias e, após isso, movidas para condições de crescimento normais. Os filtros foram transferidos para placas contendo Gm fresco uma vez por semana e aquelas formadas na presença do marcador selecionável Gm foram isoladas e reestriadas sobre filtros frescos, seguido por transferência para meio de crescimento líquido BG11 sob condições seletivas contínuas na presença de Gm.

Exemplo 3. Expressão de *IspS* His-tagged em *Escherichia coli*

De forma a construir o vetor para expressão do gene *IspS* His-sinalizado em *E. coli*, DNA de *Ss-IspS*, códon-otimizado para expressão em cianobactérias, foi amplificado através de PCR usando os iniciadores: *IspS_F_NdeI*, 5'-CTGGGTCATATGGAAGCTCGACGAA-3' e *IspS_R_HindIII*, 5'-ATGGAAAACCTGAAGCTTTTAACGTTCAA-3'; introduzindo o sítio de *NdeI* e um sítio de *HindIII* nas extremidades 5' e 3' do gene, respectivamente. Esses sítios foram usados para clonar o gene no vetor de expressão pET1529, formando o vetor pET*IspS*, o qual traz um sinalizador de His na extremidade N-terminal da proteína (figura 11).

De forma a demonstrar expressão de *Ss-IspS* His sinalizado recombinante em *Escherichia coli*, bactérias *E. coli* foram transformadas com o plasmídeo pET*IspS*, o qual contém o gene *Ss-IspS* e um DNA codificante de sinalizador de His na extremidade 5' do gene *Ss-IspS*. Expressão com sucesso desse gene *His-Ss-IspS* em *E. coli* foi induzida pela da adição de IPTG a 0,1 mM à cultura celular líquida. As células foram coletadas e seu teor de proteína foi analisado através de SDS-PAGE e coloração com Coomassie (figura 12). Foi demonstrado que todos os clones trazendo o plasmídeo pET*IspS* estavam expressando a proteína His-*Ss-IspS* de ~65kD (figura 12, banda de ~65 kD).

Uma compreensão e demonstração similares de expressão do gene *IspS* e acúmulo de proteína de *IspS* recombinante em bactérias, por exemplo, *Escherichia coli*, foram conduzidas com sucesso com o gene *Cr-IspS*, códon-otimizado para expressão em algas verdes unicelulares, por exemplo, *Chlamydomonas reinhardtii* (resultados não mostrados).

Todas as publicações, números de acesso e pedidos de patente

citados no presente relatório descritivo são aqui incorporados por referência como se cada publicação ou pedido de patente individual fosse específica e individualmente indicado como sendo incorporado por referência.

Sequências de IspS Exemplares:

5 SEQ ID NO: 1 - cDNA de *Populus alba* para isopreno sintase, Nº Acesso AB198180

```

1 atggcaactg aattattgtg cttgcaccgt ccaatctcac tgacacacaa acttttcaga
61 aatcccttac ctaaagtcac ccaggccact cccttaactt tgaaactcag atgttctgta
121 agcacagaaa acgtcagctt cacagaaaca gaaacagaag ccagacggtc tgccaattat
181 gaaccaaata gctgggatta tgattatttg ctgtcttcag aactgacga atcgattgag
241 gtatacaaag acaaggccaa aaagctggag gctgagggtga gaagagagat taacaatgaa
301 aaggcagagt ttttgactct gcttgaactg atagataatg tccaaagggt aggattgggt
361 taccggttcg agagtgcacat aaggggagcc cttgatagat ttgtttcttc aggaggattt
421 gatgctgtta caaaaactag cttcatgggt actgctctta gcttcaggct tctcagacag
481 catgggtttt aggtctctca agaagcgttc agtggattca aggatcaaaa tggcaatttc
541 ttggaaaacc ttaaggagga catcaaggca atactaagcc tatatgaagc ttcatttctt
601 gcattagaag gagaaaatat cttggatgag gccaaagggt ttgcaatata acatctaaaa
661 gagctcagcg aagaaaagat tggaaaagag ctggccgaac aggtgaatca tgcattggag
721 cttccattgc atcgcaggac gcaaagacta gaagctgttt ggagcattga agcataccgt
781 aaaaaggaag atgcaaatca agtactgcta gaacttgcta tattggacta caacatgatt
841 caatcagtat accaaagaga tcttcgcgag acatcaaggt ggtggaggcg agtgggtctt
901 gcaacaaagt tgcattttgc tagagacagg ttaattgaaa gcttttactg ggcagttgga
961 gttgcgttcg agcctcaata cagtgattgc cgtaattcag tagcaaaaat gttttcattt

```

1021 gtaacaatca ttgatgatat ctatgatggt tatgggtactc tggacgagtt ggagctatatt
 1081 acagatgctg ttgagagatg ggatgttaat gccatcaatg atcttcgga ttatatgaag
 1141 ctctgcttcc tagctctcta caacactatc aatgagatag cttatgacaa tctgaaggac
 1201 aagggggaaa acattcttcc atacctaaca aaagcgtggg cagatttatg caatgcattc
 1261 ctacaagaag caaaatgggt gtacaataag tccacaccaa catttgatga ctatttcgga
 1321 aatgcatgga aatcatcctc agggcctctt caactagttt ttgcctactt tgccgtgggt
 1381 caaaacatca agaaagagga aattgaaaac ttacaaaagt atcatgatac catcagtagg
 1441 ccttcccaca tctttcgtct ttgcaacgac ctggcttcag catcggtga gatagcgaga
 1501 ggtgaaacag cgaattctgt atcatgctac atgcgtacaa aaggcatttc tgaggagctt
 1561 gctactgaat ccgtaatgaa cttgatcgac gaaacctgga aaaagatgaa caaagaaaag
 1621 cttgggtggct ctttgtttgc aaaacctttt gtcgaaacag ctattaacct tgcacggcaa
 1681 tcccattgca cttatcataa cggagatgcg catacttcac cagacgagct aactaggaaa
 1741 cgtgtcctgt cagtaatcac agagcctatt ctacctttg agagataa

SEQ ID NO: 2 - Sequência de polipeptídeo de *Populus alba* para isopreno sintase (do Nº de Acesso AB198180). A porção sublinhada da proteína denota um peptídeo de trânsito em cloroplasto.

MATELLCLHRPISLTHKLFRNPLPKVIQATPLTLKLRCSVSTENVSTETETEARRSAN
 YEPNSWDYDYLLSSDTDESIEVYKDKAKKLEAEVRREINNEKA EFLTLLELIDNVQR
 LGLGYRFESDIRGALDRFVSSGGFDAVTKTSLHGTALSFRLLRQHGFEVSQEAFSGFK
 DQNGNFLENLKEDIKAILSLYEASFLALEGENILDEAKVFAISHLKELSEEKIGKELAE
 QVNHALELPLHRRRTQRLEAVWSIEAYRKKEDANQVLLELAILDYNMIQSVYQRDLR
 ETSRWWRRVGLATKLHFARDRLIESFYWAVGVAFEPQYSDCRNSVAKMFSFVTIIDD
 IYDVYGTLDLELELFTDAVERWDVNAINDLDPDYMKLCFLALYNTINEIAYDNLKDKGE
 NILPYLTKAWADLCNAFLQEAKWLYNKSTPTFDDYFGNAWKSSSGPLQLVFAYFAV
 VQNIKKEBIENLQKYHDTISRPSHIFRLCNDLASASAEIARGETANSVSCYMRKGISE
 ELATESVMNLIDETWKKMNKEKLGGS�FAKPFVETAINLARQSHCTYHNGDAHTSP
 DELTRKRVLSVITEPILPER

SEQ ID NO: 3 - Gene *Cr-IspS* e sinalizador de hemaglutinina para transformação/expressão em algas verdes unicelulares. A sequência de *IspS* começa com o códon sublinhado "CCA".

ATGTATCCTTATGATGTTCCAGACTACGCAGGTTATCCTTATGATGTACCAGACTATGCA
 GGTTATCCTTACGATGTACCTGATTACGCTGGTCCATGGTGTCTGTTAGTACTGAAAATG
 TTTCATTTACTGAAACAGAAACAGAAGCACGTAGATCAGCAAATTATGAGCCAAATAGT
 TGGGATTATGACTATTTATTATCTAGTGATACAGATGAATCTATTGAAGTATATAAAGAT
 AAAGCAAAAAAATTAGAAGCAGAAGTACGTCGTGAAATTAATAACGAAAAAGCAGAAT
 TTCTTACTTTATTAGAATTAATTGATAATGTACAACGTTTAGGTTTAGGTTATCGTTTTGA
 ATCAGACATTCGTGGTGCATTAGATCGTTTTGTATCAAGTGGTGGTTTTGATGCTGTTACA
 AAAACTAGTTTACATGGTACTGCTTTAAGTTTTCTGTTTACTTCGTCAACATGGTTTTGAAG
 TAAGTCAAGAAGCTTTTTCTGGTTTTAAAGATCAAAATGGTAATTTCTTAGAAAAATTTAA
 AAGAAGATATTAAAGCTATTTTAAGTTTATACGAAGCATCATTTTTAGCTTTAGAAGGTG
 AAAATATTTTAGATGAAGCTAAAGTATTTGCTATTTCTCACTTAAAAGAATTATCAGAAG
 AAAAAATTGGTAAAGAATTAGCTGAACAAGTAAACCATGCATTAGAATTACCATTACAT
 CGTCGTACACAACGTTTAGAAGCAGTTTGGTCTATTGAAGCTTATCGTAAAAAAGAAGAT
 GCTAATCAAGTTTTATTAGAATTAGCAATTTAGATTATAATATGATTCAATCAGTATACC
 AACGTGATTTACGTGAAACAAGTCGTTGGTGGCGTCGTGTAGGTTTAGCTACTAAATTAC
 ATTTTGCTCGTGATCGTTTAATTGAAAGTTTTATTGGGCAGTTGGTGTAGCTTTTGAACC
 ACAATATTCAGATTGTCGTAATTCAGTTGCAAAAATGTTTCATTTGTAATATTATTGAT
 GATATTTATGATGTTTACGGTACATTAGATGAATTAGAATTATTCAGTATGCAGTAGAA
 CGTTGGGATGTTAATGCTATTAATGATTTACCAGATTATATGAAATTATGTTTTCTTGCTT
 TATATAACACTATTAATGAAATTGCTTATGATAACTTAAAAGATAAAGGTGAAAAATATT
 TACCATATTTAACAAAAGCTTGGGCTGATTTATGTAATGCTTTTTTACAAGAAGCTAAAT
 GGTTATATAATAAATCAACACCAACATTTGATGATTATTTGGTAATGCTTGGAAAAGTT
 CATCTGGTCCATTACAATTAGTTTTTGCTTATTTTGCTGTTGTTCAAAATATTAAAAAAGA
 AGAAATTGAAAAATTTACAAAAATATCATGATACAATTTACGTCCATCACATATTTTTCG
 TTTATGTAATGATTTAGCTTCAGCTTCAGCTGAAATTGCACGTGGTGAAACAGCAAATTC
 AGTTTCATGTTATATGCGTACAAAAGGTATTTCTGAAGAATTAGCTACAGAATCAGTTAT
 GAATTTAATTGATGAAACATGGAAAAAATGAATAAAGAAAAATTAGGTGGTCTTTAT
 TTGCTAAACCATTTGTTGAAACTGCTATTAATTTAGCACGTCAATCACATTGTACTTATCA
 TAATGGTGTATGCTCATACATCACCAGATGAATTAACACGTAAACGTGTTTATCAGTTAT
 TACAGAACCAATTTTACCATTGAAACGTAA

- SEQ ID NO: 4** - Sequência de polipeptídeo para o gene de isopreno sintase *Cr-IspS*. As três cópias do sinalizador de hemaglutinina HA estão sublinhadas. A sequência de isopreno sintase carece de uma sequência de direcionamento de cloroplasto da sequência de proteína de *IspS* de álamo. A sequência de *IspS* começa com "CS...", indicado pela alteração na fonte.

MYPYDVDPDYAGYPYDVDPDYAGYPYDVDPDYAGPWCSVSTENVSTETETETEARRSANYEPNSWDYDY
 LLSSDTDESIEVYKDKAKKLEAEVRREINNEKAEFLTLLLELIDNVQRLGLGYRFESDIRGALDR
 FVSSGGFDAVTKTSLHGTALSFRLLRQHGFVVSQEAQSGFKDQNGNLENLKEDIKAILSLYE
 ASFLALEGENILDEAKVFAISHLKELSEEKIGKELAEQVNHAELEPLHRRTQRLEAVWSIEAY
 RKKEDANQVLLLELAILDYNMIQSVYQRDLRETSRWRRVGLATKLHFARDRLIESFYWAVG
 VAFEPQYSDCRNSVAKMFSFVTIIDDYDVYGTLDLELELFTDAVERWDVNAINDLDPYMKLC
 FLALYNTINEIAYDNLKDKGENILPYLTKAWADLCNAFLQEAKWLYNKSTPTFDDYFGNAW
 KSSSGPLQLVFAYFAVVQNIKKEEIEENLQKYHDTISRPSHIFRLCNDLASASAEIARGETANSV
 SCYMRTKGISEELATESVMNLIDETWKKMNKEKLGGSFPAKPFVETAINLARQSHCTYHNGD
 AHTSPDELTRKRVLSVITEPILPFER

- SEQ ID NO: 5** - Sequência de nucleotídeo de DNA de *Ss-IspS* e plasmídeo plspS para cianobactéria. A primeira sequência sublinhada de SEQ ID NO: 5 representa o gene de beta-lactamase (complemento reverso), enquanto que a segunda sequência sublinhada é o DNA de *Ss-IspS*. Adicionalmente, as
- 5 sequências em *itálico* são os códons de início e parada e as sequências em **negrito** são sítios de restrição para clonagem

>pIspS

aaaaagcattgctcatcaatttgttgcaacgaacaggtcactatcagtcaaaataaaatcattatttaaaagggg
 cccgagcttaagactggcgctcgttttacaacacagaaagagttttagtagaaacgcaaaaaggccatccgtcaggg
 gccttctgcttagttgatgcctggcagttccctactctcgcttccgcttccctcgctcactgactcgctgcgct
 cggctcgttcggctgcggcgagcggtatcagctcactcaaaggcggttaatacggttatccacagaatcaggggata
 acgcaggaaagaacatgtgagcaaaaaggccagcaaaaaggccaggaaccgtaaaaaggccgcgttgctggcgttt
 tccataggctccgccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggac
 tataaagataaccaggcgtttccccctggaagctccctcgctgcgctctcctgttccgacctgcccgttacccgat
 acctgtccgcctttctcccttcgggaagcggtggcgctttctcatagctcacgctgtaggtatctcagttcggtgt
 aggtcgttcgctccaagctgggctgtgtgcacgaacccccgttcagcccgaccgctgcgccttatccggtaact
 atcgctcttgagtccaacccggtaagacacgacttatcgccactggcagcagccactggtaacaggattagcagag
 cgaggtatgtaggcggtgctacagagttcttgaagtgggtgggtaactacggctacactagaagaacagtatttg
 gtatctgcgctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaaacaaaccacg
 ctggtagcgggtgggttttttgtttgcaagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttga
 tcttttctacggggtctgacgctcagtggaaacgacgcgcgtaactcacgttaagggattttgggtcatgagctt
 gcgccgtcccgtaagtcagcgtaatgctctgcttaccgaatgcttaatacagtgaggcacctatctcagcgatctg
 tctatttctgctcatccatagttgcctgactccccgtcggttagataactacgatacgggaggggcttaccatctgg

cccagcgctgcatgataccgagagaaccacgctcaccgggtccggatttatcagcaataaaccagccagccgg
 aagggccgagcgcagaagtggctcctgcaactttatccgcctccatccagtctattaattgttgcgggaagctag
 agtaagtagttccgagttaatagtttgcgcaacggtgttgccatcgctacaggcatcggtgtcacgctcgtc
 gtttggtatggcttcattcagctccggttcccaacgatcaaggcgagttacatgatccccatggtgtgcaaaaa
 agcgggttagctccttcggtcctccgacgtgtgtcagaagtaagttggccgcagtggttatcactcatggttatggc
 agcactgcataattctcttactgtcatgccatccgtaagatgcttttctgtgactggtagtactcaaccaagtc
 attctgagaatagtgtatgcccgcaccgagttgctccttgcggcggtcaatacgggataataccgcccacatag
 cagaactttaaaagtgtcatcattggaataacggttcttcggggcgaaaaactctcaaggatcttaccgctgttgag
 atccagttcgatgtaacccactcgtgcacccaactgatcttcagcatcttttactttcaccagcgtttctgggtg
 agcaaaaacaggaaggcaaaaatgccgcaaaaaagggaataaggcgacacggaaatgttgaatactcatattctt
 cctttttcaatattattgaagcatttatcagggttattgtctcatgagcggatacatatttgaatgtatttagaa
 aaataaacaatataggggtcagtggtacaaccaattaaccaattctgaacattatcgcgagccattttatacctga
 atatggctcataacaccccttgacgtgcgactaacggcatgaagctcgtcggggaaataatgattttatatttgac
 tgatagtgcctgttcgttgcaacaaattgataagcaatgctttcttataatgccactttgtacaagaaagctg
 ggtccatggaagctcgacgaagcgctaattatgaaccaaatagttgggactacgattttctattatcctctgata
 cggatgagtcattgaagttttataaggataaaagctaagaaattggaagccgaagtcgcccgcgaaattaacaatg
 agaaagcggaaatttttgaccttattagaactcatcgataatgtgcaacgactgggattgggctatcgggttgaaa
 gtgacatccgcccgggactggatcgttttgatctagtggcggttgatggcgctcactaaaactagtttgacg
 cgaccgcactcagttttcggtattacgtcaacacgggttcgaagtgagtcagagcggttttagtggcttcaaag
 atcaaaatggcaattttctggaaaacttgaaagaagacacaaaagctatcctaagtttatagcaagctagtttct
 tcgcgctggaagggtgaaaatattctggatgagggtcgtgtatttgcaattttctcacctgaaagaattatctgaag
 aaaagattggcaagaactcgccgaacaggtaaatacagccttggaactgcccctccatcgctcgtaaccaacgat
 tggaaagctgtgtggagtatcgaaagcctatcgcaagaaagaagacgctaaccaagttttgttggaaactggccatct
 tggattataacatgattcaatccgtatatcagcgcatctacgtgaaacgtctcggtgggtggcggtgttgggc
 tcgctactaaattacattttgcaaaggatcgactcattgaatccttttattgggcccgtcggggtggcttttgaa
 ccagtcacagcgattgcccgaattctgtagcaaaaatgttttcttcgttacaattattgatgacatttatgacg
 tttacggcaccctcgacgaactggaattgttcaactgacgctgtggaacggtgggacgtaaatgccattaatgacc
 tgccagattacatgaagttgtgttttctcggttatataacaccattaatgaaattgcatacgacaattttaaagg
 ataaggagagagaatattctgccttatttgacgaaagcctgggcccgaatttgtgtaatgcctttttgcaggaagcta
 aatgggtatataacaaatccacccccacttttgatgactattttggcaatgcctggaagagcagcagcgggctc
 tccaactgatttttgcttattttgcggtagtacaaaacattaaagaaagaagagattgaaaatttgcaaaagtacc
 atgacattatttagtcggcccagtcataattttccgcttggtgcaacgacctggcatccgctagtgcgaaattgccc
 gtggcgaaacagctaatagtgtgagttgttacatgcccacaaaggcatttccgaagaactagctacggaaagtg
 tcatgaacctgattgacgagacttgcaagaaaatgaataaggaaaaattgggcccgtccctatttgccaaacct
 ttgtggaaaccgcgattaattttggctcgccaaagtcattgtacctatcacaatgggtgatgctcacaccagtcgcc
 atgaattaacccgtaaacgagttctgtctgtgattactgaacctattttgcccttgaaacgttaaaagtaacagg
 ttttccatgttgctgtctgcaagaacactgcagagcctgctttttgtacaaagttggcattata

SEQ ID NO: 6 - sequência de aminoácido da proteína de *Ss-IspS* de 65 kD
 esperada traduzida do plasmídeo de cianobactéria

1 MEARRSANYE PNSWDYDFLL SSDTDESIEV YKDKAKKLEA EVRREINNEK
 51 AEFLTLELELI DNVQRLGLGY RFESDIRRAL DRFVSSGGFD GVTKTSIHAT
 101 ALSFRLLRQH GFEVSQEAFS GFKDQNGNFL ENLKEDTKAI LSLYEASFLA
 151 LEGENILDEA RVFAISHLKE LSEEKIGKEL AEQVNHAELEL PLHRRTRQLE
 201 AVWSIEAYRK KEDANQVLE LAILDYNMIQ SVYQORDLRET SRWRRRVGLA
 251 TKLHFAKDRL IESFYWAVGV AFEPQYSDCR NSVAKMFSFV TIIDDIYDVY
 301 GTLDELELEFT DAVERWDVNA INDLPDYMKL CFLALYNTIN EIAYDNLKDK
 351 GENILPYLTK AWADLCNAFL QEAKWLYNKS TPTFDDYFGN AWKSSSGPLQ
 401 LIFAYFAVVQ NIKKEEIEENL QKYHDIISRP SHIFRLCNDL ASASAEIARG
 451 ETANSVSCYM RTKGISEELA TESVMNLIDE TCKKMNKEKL GGSFLAKPFV
 501 ETAINLARQS HCTYHNGDAH TSPDELTRKR VLSVITEPIL PFER*

SEQ ID NO: 7 - Isopreno sintase (IspS) de *Pueraria montana* var. lobata (uva kudzu); Nº de Acesso AY316691 (cds. completa). O códon de início atg está sublinhado e indica o início da região de codificação de proteína do cDNA

1 aatcaatata taatattttac ggaagatttg atgcctttcc tgatttttaat ttattttttat
 61 ccctgcataa aataattgtg gtcaccgtac actgttcttg tcacttggac aagaaatttg
 121 actagcaagc aaggtataat cattcatcta aacttatggg gatttatgtc cccacctcat
 181 caattttcgt gtgtttttatt ttagtgtcct tggatcctcg ttccaatata aaaggagAAC
 241 atggcatcgc aatttttagag catatcattg aaaagtcatt gcaaccaacc ttttatgctt
 301 gtctaataaa ttatcgtccc ccacaccaac accaagtact agatttccac aaagtaagaa
 361 cttcatcaca caaaaaacat ctcttgccaa tcccaaacct tggcgagtta tttgtgttac
 421 gagctctcaa tttacccaaa taacagaaca taatagtcgg cgttcagcta attaccagcc
 481 aaacctctgg aattttgaat ttctgcagtc tctggaaaat gaccttaagg tgattataca
 541 tatattccag ttaatttttc tttttttctt ttgtgatttt taaggaaatca tttagtttgg
 601 gaaagtattt tttttatttg cacttttaat tataaaaaatg ttatatcatt ttcacttttt
 661 tctattcatt ttcaaaattt tacatagaaa acagtaaatt ttttattttt tttattttct
 721 attttcatta tttctcaaat caaacgggat taaagcataa acaagaaat taatatgtt
 781 cttttaattt ttttttttta caataatggg aacgattata tattaggctg accttaataa
 841 gttatttttt ttttataata ttgttcttat tgtaacctaa cgacagggtg aaaaactaga
 901 agagaaggca acaaagctag aggaggagg acgatgcag atcaacagag tagacacaca
 961 accattaagc ttactagaat tgatcgacga tgtccagcgt ctaggattga cctacaagtt
 1021 tgagaaggac ataatacaag cccttgagaa tattgttttg ctggatgaga ataagaaaaa
 1081 taaaagtga ctcctatgta ctgctctcag ctcccgttta cttagacaac atggccttga
 1141 ggtttcccaa ggtatttatg tatatatatg ttaccactt agcaacatat atatatatat
 1201 atattatgat tcaactgacca tgcattgtgt gcagatgtgt ttgagagatt taaggacaag
 1261 gagggagggt tcagtgggtg acttaaagg gatgtgcaag ggttgctgag tctatatgaa
 1321 gcacccatc ttggccttga gggagaaaat ctcttgagg aggcaaggac attttcaata
 1381 acacatctca agaacaacct aaaagaagga ataaacacca aagtggcaga acaagttagt
 1441 catgcactgg aacttcccta tcatcaaaga ttgcatagac tagaagcacg atgggtcctt
 1501 gacaaatatg aaccaaagga accccaccat cagtactac tcgagcttgc aaagctagat
 1561 ttcaatatgg tgcaaacatt gcaccagaaa gaactgcaag acctgtcaag gttagaaatt
 1621 tcaattctca agtaattatt acctcataag aaattaaata acaataacaa tattgagtgt
 1681 agagatttcc aattaaaaat taacatacga gaggatcaat atatattctt aggtatgtgg
 1741 tactaatgaa atatatgcta ggtggtggac ggagatgggg ctagcaagca agctagactt
 1801 tgtccgagac agattaatgg aagtgtattt ttgggcgttg ggaatggcac ctgaccccca
 1861 attcgggtgaa tgtcgtaaag ctgtcactaa aatgttttga ttggtcacca tcatcgatga
 1921 tgtatatgac gtttatggta ctttgatga gctacaactc ttcactgatg ctggtgagag
 1981 gttcgttaatt gatttcagtc tgcattcagt tggaaattta ttattgctta attaatata
 2041 acttgcgtag atgcatacac acagatggga cgtgaatgcc ataaacacac ttccagacta
 2101 catgaagttg tgcttcctag cactttataa caccgtcaat gacacgtctt atagcatcct

```

2161 taaagaaaaa ggacacaaca acctttccta tttagacaaa tctgtacata tataactaatt
2221 atctccttgg ttgattaatt agtttagttt agtttagttg gtatgtcaac acaattaatt
2281 aatattatat atggatggtg acagtggcgt gagttatgca aagcattcct tcaagaagca
2341 aaatgggtcg acaacaaaat cattccagca tttagcaagt acctggaaaa tgcacgggtg
2401 tctctctccg gtgtggcttt gcttgctcct tctacttct cagtgtgcca acaacaagaa
2461 gatattctcag accatgctct tcgttcttta actgatttcc atggccttgt gcgctcctca
2521 tgcgtcattt tcagactctg caatgatttg gctacctcag cgggtgtgtaa ttaattacct
2581 taattaattt gtaacacttg ttagactaat atatataggt gtgtctgtta attactacag
2641 gctgagctag agaggggtga gacgacaaat tcaataatat cttatatgca tgagaatgac
2701 ggcacttctg aagagcaagc acgtgaggag ttgagaaaaat tgatcgatgc agagtgggaag
2761 aagatgaacc gagagcgagt ttcagattct acactactcc caaaagcttt tatggaaata
2821 gctgttaaca tggctcgagt ttgcattgc acataccaat atggagacgg acttggaagg
2881 ccagactacg ccacagagaa tagaatcaag ttgctactta tagaccctt tccaatcaat
2941 caactaatgt acgtgtaaca acacaatata aacacttttc tacaagtata tatttggtta
3001 atttcgggtg tgaattaggg gtcaacacag ctatatatac ttcaatggac caactcaacc
3061 aatctgataa gagaaaaaaa ataaaaataa ggtaggtta actttgtata aatccaagtt
3121 agatatcaag ttt

```

SEQ ID NO: 8 - Sequência de polipeptídeo de isopreno sintase de *Pueraria montana* var. lobata (uva kudzu)

```

MATNLLCLSNKLSSPTPTPSTRFPQSKNFITQKTSLANPKPWVICATSSQFTQITEHN
SRRSANYQPNLWNFEFLQSLNDLKVEKLEEKATKLEEEVRCMINRVDTQPLSLELI
DDVQRLGLTYKFEKDIIKALENIVLLDENKKNKSDLHATALSFRLLRQHGFVSQDV
FERFKDKEGGFSGELKGDVQGLLSLYEASYLGFEGENLLEEARTFSITHLKNNLKEGI
NTKVAEQVSHALELPYHQRLHRLEARWFLDKYEPKEPHHQLLLELAKLDFNMVQTL
HQKELQDLRWWTEMGLASKLDFVRDRLMEVYFWALGMAPDPQFGECKRAVTKM
FGLVTIIDDVYDVYGTLDLQLFTDAVERWDVNAINLTPDYMKLCFLALYNTVNDT
SYSILKEKGHNLSYLTKSWRELCKAFLQEAKWSNNKIIPAFSKYLENASVSSSGVAL
LAPSYFSVCQQQEDISDHALRSLTDFHGLVRSSCVIFRLCNDLATSAEELERGETTNSI
ISYMHENDGTSEEQAREELRKLIDAEWKKMNRERVSDSTLLPKAFMEIAVNMARVS
HCTYQYGDGLGRPDYATENRIKLLLIDPFPINQLMYV

```

LISTAGEM DE SEQUÊNCIA

- <110> Melis, Anastasios
The Regents of the University of Califórnia
- <120> Produção de Hidrocarboneto Volátil de Cadeia Curta Usando Microalgas, Cianobactérias ou Bactérias Geneticamente Manipuladas
- <130> 18941H-004010PC
- <140> WO PCT/US07/72465
- <141> 29/06/2007
- <150> US 60/806.244
- <151> 29/06/2006
- <150> US 11/770.412
- <151> 28/06/2007
- <160> 13
- <170> PatentIn Ver. 2.1
- <210> 1
- <211> 1788
- <212> DNA
- <213> Populus alba
- <220>
- <223> cDNA de isopreno sintase (IspS) de álamo branco

<400> 1

```

atggcaactg aattattgtg cttgcaccgt ccaatctcac tgacacacaa acttttcaga 60
aatcccttac ctaaagtcac ccaggccact cccttaactt tgaaactcag atgttctgta 120
agcacagaaa acgtcagctt cacagaaaca gaaacagaag ccagacggtc tgccaattat 180
gaaccaaata gctgggatta tgattatttg ctgtcttcag aactgacga atcgattgag 240
gtatacaaag acaaggccaa aaagctggag gctgaggtga gaagagagat taacaatgaa 300
aaggcagagt ttttgactct gcttgaactg atagataatg tccaaagggt aggattgggt 360
taccggttcg agagtgcacat aaggggagcc cttgatagat ttgtttcttc aggaggattt 420
gatgctgtta caaaaactag ccttcattgtt actgctotta gcttcaggct tctcagacag 480
catgggtttg aggtctctca agaagcgttc agtggattca aggatcaaaa tggcaatttc 540
ttggaaaacc ttaaggagga catcaaggca atactaagcc tatatgaagc ttcatttctt 600
gcattagaag gagaaaatat cttggatgag gccaaagggt ttgcaatatc acatctaaaa 660
gagctcagcg aagaaaagat tggaaaagag ctggccgaac aggtgaatca tgcattggag 720
cttccattgc atcgagaggac gcaaagacta gaagctgttt ggagcattga agcataccgt 780
aaaaaggaag atgcaaatca agtactgcta gaacttgcta tattggacta caacatgatt 840
caatcagtat accaaagaga tcttcgcgag acatcaagggt ggtggaggcg agtgggtctt 900
gcaacaaagt tgcattttgc tagagacagg ttaattgaaa gcttttactg ggcagttgga 960
gttgcgttcg agcctcaata cagtgattgc cgtaattcag tagcaaaaat gttttcattt 1020
gtaacaatca ttgatgatat ctatgatgtt tatggtactc tggacgagtt ggagctattt 1080
acagatgctg ttgagagatg ggatgttaat gccatcaatg atcttcogga ttatatgaag 1140
ctctgcttcc tagctctcta caacactatc aatgagatag cttatgacaa tctgaaggac 1200
aaggggggaaa acattcttcc atacctaaca aaagcgtggg cagatttatg caatgcattc 1260
ctacaagaag caaatgggtt gtacaataag tccacaccaa catttgatga ctatttcgga 1320
aatgcatgga aatcatcctc agggcctctt caactagtgt ttgcctactt tggcgtgggt 1380
caaaacatca agaaagagga aattgaaaac ttacaaaagt atcatgatac catcagtagg 1440
ccttcccaca tctttcgtct ttgcaacgac ctggcttcag catcggtgta gatagcgaga 1500
gggtgaaacag cgaattctgt atcatgctac atgctgacaa aaggcatttc tgaggagctt 1560
gctactgaat ccgtaatgaa cttgatcgac gaaacctgga aaaagatgaa caaagaaaaa 1620
cttggtggct ctttgtttgc aaaacctttt gtcgaaacag ctattaacct tgcacggcaa 1680
tcccattgca cttatcataa cggagatgcg catacttcac cagacgagct aactaggaaa 1740
cgtgtcctgt cagtaatcac agagcctatt ctaccctttg agagataa 1788

```

<210>	2
<211>	595
<212>	PRT
<213>	Populus alba
<220>	
<223>	Isopreno sintase (IspS) de álamo branco
<220>	
<221>	PEPTÍDEO
<222>	(1) . . (37)
<223>	peptídeo de trânsito de cloroplasto
<400>	2

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

Glu Lys Ile Gly Lys Glu Leu Ala Glu Gln Val Asn His Ala Leu Glu
 225 230 235 240
 Leu Pro Leu His Arg Arg Thr Gln Arg Leu Glu Ala Val Trp Ser Ile
 245 250 255
 Glu Ala Tyr Arg Lys Lys Glu Asp Ala Asn Gln Val Leu Leu Glu Leu
 260 265 270
 Ala Ile Leu Asp Tyr Asn Met Ile Gln Ser Val Tyr Gln Arg Asp Leu
 275 280 285
 Arg Glu Thr Ser Arg Trp Trp Arg Arg Val Gly Leu Ala Thr Lys Leu
 290 295 300
 His Phe Ala Arg Asp Arg Leu Ile Glu Ser Phe Tyr Trp Ala Val Gly
 305 310 315 320
 Val Ala Phe Glu Pro Gln Tyr Ser Asp Cys Arg Asn Ser Val Ala Lys
 325 330 335
 Met Phe Ser Phe Val Thr Ile Ile Asp Asp Ile Tyr Asp Val Tyr Gly
 340 345 350
 Thr Leu Asp Glu Leu Glu Leu Phe Thr Asp Ala Val Glu Arg Trp Asp
 355 360 365
 Val Asn Ala Ile Asn Asp Leu Pro Asp Tyr Met Lys Leu Cys Phe Leu
 370 375 380
 Ala Leu Tyr Asn Thr Ile Asn Glu Ile Ala Tyr Asp Asn Leu Lys Asp
 385 390 395 400
 Lys Gly Glu Asn Ile Leu Pro Tyr Leu Thr Lys Ala Trp Ala Asp Leu
 405 410 415
 Cys Asn Ala Phe Leu Gln Glu Ala Lys Trp Leu Tyr Asn Lys Ser Thr
 420 425 430
 Pro Thr Phe Asp Asp Tyr Phe Gly Asn Ala Trp Lys Ser Ser Ser Gly
 435 440 445
 Pro Leu Gln Leu Val Phe Ala Tyr Phe Ala Val Val Gln Asn Ile Lys
 450 455 460
 Lys Glu Glu Ile Glu Asn Leu Gln Lys Tyr His Asp Thr Ile Ser Arg
 465 470 475 480
 Pro Ser His Ile Phe Arg Leu Cys Asn Asp Leu Ala Ser Ala Ser Ala
 485 490 495
 Glu Ile Ala Arg Gly Glu Thr Ala Asn Ser Val Ser Cys Tyr Met Arg
 500 505 510
 Thr Lys Gly Ile Ser Glu Glu Leu Ala Thr Glu Ser Val Met Asn Leu
 515 520 525
 Ile Asp Glu Thr Trp Lys Lys Met Asn Lys Glu Lys Leu Gly Gly Ser
 530 535 540

Leu Phe Ala Lys Pro Phe Val Glu Thr Ala Ile Asn Leu Ala Arg Gln
545 550 555 560

Ser His Cys Thr Tyr His Asn Gly Asp Ala His Thr Ser Pro Asp Glu
565 570 575

Leu Thr Arg Lys Arg Val Leu Ser Val Ile Thr Glu Pro Ile Leu Pro
580 585 590

Phe Glu Arg
595

<210> 3

<211> 1776

<212> DNA

<213> sequência artificial

<220>

<223> Descrição de sequência artificial: Constructo de DNA sintético códon-ajustado para expressão do gene de IspS de *Populus alba* com sequência de direcionamento de cloroplasto omitida e sinalizador de epítipo de hemaglutinina (HA) em microalga *Chlamydomonas reinhardtii* (Cr-IspS)

<220>

<221> gene

<222> (100)..(1776)

<223> Gene de IspS de *Populus alba* com sequência de direcionamento de cloroplasto omitida códon-otimizado para expressão em cloroplasto de microalga *Chlamydomonas reinhardtii* (Cr-IspS)

<400> 3

```

atgtatcctt atgatgttcc agactacgca gggtatcctt atgatgtacc agactatgca 60
gggtatcctt acgatgtacc tgattacgct ggtccatggg gttctgttag tactgaaaat 120
gtttcattta ctgaaacaga aacagaagca cgtagatcag caaattatga gccaaatagt 180
tgggattatg actatttatt atctagtgat acagatgaat ctattgaagt atataaagat 240
aaagcaaaaa aattagaagc agaagtacgt cgtgaaatta ataacgaaaa agcagaattt 300
cttactttat tagaattaat tgataatgta caacgttttag gtttaggtta tcgttttgaa 360
tcagacattc gtggtgcatt agatcgtttt gtatcaagtg gtgggttttga tgctgttaca 420
aaaactagtt tacatggtac tgctttaagt tttcgtttac ttcgtcaaca tgggtttgaa 480
gtaagtcaag aagctttttc tgggttttaa gatcaaaatg gtaattttct agaaaattta 540
aaagaagata ttaaagctat ttttaagttta tacgaagcat catttttagc tttagaaggt 600
gaaaaatatt tagatgaagc taaagtattt gctattttct acttaaaaga attatcagaa 660
gaaaaaattg gtaaaagaatt agctgaacaa gtaaaaccatg cattagaatt accattacat 720
cgtcgtacac aacgttttaga agcagtttgg tctattgaag cttatcgtaa aaaagaagat 780
gctaatacaag ttttatttaga attagcaatt ttagattata atatgattca atcagtatac 840
caacgtgatt tacgtgaaac aagtcgttgg tggcgtcgtg taggttttagc tactaaatta 900
cattttgctc gtgatcgttt aattgaaagt ttttattggg cagttgggtg agcttttgaa 960
ccacaatatt cagattgtcg taattcagtt gcaaaaaatg tttcatttgt aactattatt 1020
gatgatattt atgatgttta cggtaacatta gatgaattag aattattcac attgacagta 1080
gaacgttggg atgttaatgc tattaatgat ttaccagatt atatgaaatt atgttttctt 1140
gctttatata acactattaa tgaaattgct tatgataact taaaagataa aggtgaaat 1200
attttaccat atttaacaaa agcttgggct gatttatgta atgctttttt acaagaagct 1260
aaatgggtat ataataaatc aacaccaaca tttgatgatt attttggtta tgcttggaaa 1320
agttcatctg gtccattaca attagttttt gcttattttg ctggtgttca aaatattaaa 1380
aaagaagaaa ttgaaaattt acaaaaaatat catgatataa tttcacgtcc atcacatatt 1440
tttcgtttat gtaatgattt agcttcagct tcagctgaaa ttgcacgtgg tgaacacagca 1500
aattcagttt catgttatat gcgtacaaaa ggtattttctg aagaatttagc tacagaatca 1560
gttatgaatt taattgatga aacatggaaa aaaatgaata aagaaaaatt aggtggttct 1620
ttatttgcta aaccatttgt tgaaactgct attaatntag cacgtcaatc acattgtact 1680

```

tatcataatg gtgatgctca tacatcacca gatgaattaa cacgtaaacg tgttttatca 1740
gttattacag aaccaatttt accatttgaa cgttaa 1776

<210> 4
<211> 591.
<212> PRT
<213> Sequência artificial
<220>
<223> Descrição de sequência artificial: Isopreno sintase (IspS) de *Populus alba* com sequência de direcionamento de cloroplasto omitida e proteína de fusão de sinalizador de epítipo de hemaglutinina (HA) para expressão em cloroplasto de microalga *Chlamydomonas reinhardtii*
<220>
<221> PEPTÍDEO
<222> (2) . . (10)
<223> peptídeo sinalizador de epítipo de hemaglutinina (HA)

<220>
<221> PEPTÍDEO
<222> (12). . (20)
<223> peptídeo sinalizador de epítipo de hemaglutinina (HA)
<220>
<221> PEPTÍDEO
<222> (22).. (30)
<223> peptídeo sinalizador de epítipo de hemaglutinina (HA)
<220>
<221> PEPTÍDEO
<222> (34)..(591)
<223> Isopreno sintase (IspS) de *Populus Alba* com sequência de direcionamento de cloroplasto omitida
<400> 4
Met Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Gly Tyr Pro Tyr Asp Val
1 5 10 15
Pro Asp Tyr Ala Gly Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Gly Pro
20 25 30
Trp Cys Ser Val Ser Thr Glu Asn Val Ser Phe Thr Glu Thr Glu Thr
35 40 45
Glu Ala Arg Arg Ser Ala Asn Tyr Glu Pro Asn Ser Trp Asp Tyr Asp
50 55 60
Tyr Leu Leu Ser Ser Asp Thr Asp Glu Ser Ile Glu Val Tyr Lys Asp
65 70 75 80
Lys Ala Lys Lys Leu Glu Ala Glu Val Arg Arg Glu Ile Asn Asn Glu
85 90 95

Lys Ala Glu Phe Leu Thr Leu Leu Glu Leu Ile Asp Asn Val Gln Arg
 100 105 110
 Leu Gly Leu Gly Tyr Arg Phe Glu Ser Asp Ile Arg Gly Ala Leu Asp
 115 120 125
 Arg Phe Val Ser Ser Gly Gly Phe Asp Ala Val Thr Lys Thr Ser Leu
 130 135 140
 His Gly Thr Ala Leu Ser Phe Arg Leu Leu Arg Gln His Gly Phe Glu
 145 150 155 160
 Val Ser Gln Glu Ala Phe Ser Gly Phe Lys Asp Gln Asn Gly Asn Phe
 165 170 175
 Leu Glu Asn Leu Lys Glu Asp Ile Lys Ala Ile Leu Ser Leu Tyr Glu
 180 185 190
 Ala Ser Phe Leu Ala Leu Glu Gly Glu Asn Ile Leu Asp Glu Ala Lys
 195 200 205
 Val Phe Ala Ile Ser His Leu Lys Glu Leu Ser Glu Glu Lys Ile Gly
 210 215 220
 Lys Glu Leu Ala Glu Gln Val Asn His Ala Leu Glu Leu Pro Leu His
 225 230 235 240
 Arg Arg Thr Gln Arg Leu Glu Ala Val Trp Ser Ile Glu Ala Tyr Arg
 245 250 255
 Lys Lys Glu Asp Ala Asn Gln Val Leu Leu Glu Leu Ala Ile Leu Asp
 260 265 270
 Tyr Asn Met Ile Gln Ser Val Tyr Gln Arg Asp Leu Arg Glu Thr Ser
 275 280 285
 Arg Trp Trp Arg Arg Val Gly Leu Ala Thr Lys Leu His Phe Ala Arg
 290 295 300
 Asp Arg Leu Ile Glu Ser Phe Tyr Trp Ala Val Gly Val Ala Phe Glu
 305 310 315 320
 Pro Gln Tyr Ser Asp Cys Arg Asn Ser Val Ala Lys Met Phe Ser Phe
 325 330 335
 Val Thr Ile Ile Asp Asp Ile Tyr Asp Val Tyr Gly Thr Leu Asp Glu
 340 345 350
 Leu Glu Leu Phe Thr Asp Ala Val Glu Arg Trp Asp Val Asn Ala Ile
 355 360 365
 Asn Asp Leu Pro Asp Tyr Met Lys Leu Cys Phe Leu Ala Leu Tyr Asn
 370 375 380
 Thr Ile Asn Glu Ile Ala Tyr Asp Asn Leu Lys Asp Lys Gly Glu Asn
 385 390 395 400
 Ile Leu Pro Tyr Leu Thr Lys Ala Trp Ala Asp Leu Cys Asn Ala Phe
 405 410 415
 Leu Gln Glu Ala Lys Trp Leu Tyr Asn Lys Ser Thr Pro Thr Phe Asp
 420 425 430
 Asp Tyr Phe Gly Asn Ala Trp Lys Ser Ser Ser Gly Pro Leu Gln Leu
 435 440 445

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

- <210> 5
 <211> 3965
 <212> DNA
 <213> Sequência artificial
 <220>
 <223> Descrição de sequência artificial: Constructo de DNA sintético códon-otimizado para expressão do gene de IspS de Populus alba com sequência de direcionamento de cloroplasto omitida em cianobactéria Synechocystis sp. (Ss-IspS) com gene de beta-lactamase (complemento reverso)
 <220>
 <221> gene
 <222> Complemento (1084)..(1944)
 <223> Gene de beta-lactamase (complemento reverso)
 <220>
 <221> gene
 <222> (2256)..(3890)
 <223> Gene de IspS de Populus Alba com sequência de direcionamento de cloroplasto omitida códon-otimizado para expressão em cianobactéria Synechocystis sp. (Ss-IspS)
 <400> 5

```

aaaaagcatt gctcatcaat ttgttgcaac gaacagggtca ctatcagtc aataaaatc 60
attattttaa agggggccga gcttaagact ggccgtcggt ttacaacaca gaaagagttt 120
gtagaacacg aaaaaggcca tccgtcaggg gccttctgct tagtttgatg cctggcagtt 180
ccctactctc gccttccgct tccctcgctca ctgactcgct gcgctcggtc gttcggctgc 240
ggcgagcggt atcagctcac tcaaaggcgg taatacgggt atccacagaa tcaggggata 300
acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc caggaaaccgt aaaaaggccg 360
cgttgctggc gtttttccat aggtcccgcc cccctgacga gcatacaaaa aatcgacgct 420
caagtcaagag gtggcgaaac ccgacaggac tataaagata ccaggcggtt cccctggaa 480
gctccctcgt gcgctctcct gttccgaacc tgccgcttac cggataacctg tccgcctttc 540

tcccttcggg aagcgtggcg ctttctcata gctcacgctg taggtatctc agttcgggtg 600
aggctcgttcg ctccaagctg ggctgtgtgc acgaaccccc cggtcagccc gaccgctgcg 660
ccttatccgg taactatcgt cttgagtcca acccggtatg acacgactta tccgactgg 720
cagcagccac tggtaacagg attagcagag cagggtatgt aggcgggtgct acagagttct 780
tgaagtgggt ggctaactac ggctacaacta gaagaacagt atttggtatc tgcgctctgc 840
tgaagccagt taccttcgga aaaaagagttg gtactctttg atccggcaaa caaaccaccg 900
ctggtagcgg tggttttttt gtttgcaagc agcagattac gcgcagaaaa aaaggatctc 960
aagaagatcc tttgatcttt tctacggggt ctgacgctca gtggaacgac gcgcgcgtaa 1020
ctcacgttaa gggatttttg tcatgagctt gcgccgtccc gtcaagtcag cgtaatgctc 1080
tgcttaccac tgcttaatca gtgaggcacc tatctcagcg atctgtctat ttcgttctac 1140
catagttgcc tgactccccg tegtgtagat aactacgata cgggagggct taccatctgg 1200
cccagcgtc gcgatgatac cgcgagaacc acgctcaccg gctccggatt tatcagcaat 1260
aaaccagcca gccggaaggg ccgagcgcag aagtggctct gcaactttat ccgcctccat 1320
ccagctctatt aattgttgcc gggaaagctag agtaagtagt tcgccagtta atagtttgcg 1380
caacgttggt gccatcgcta caggcatcgt gggtgcacgc tcgtcgtttg gtatggcttc 1440
attcagctcc ggttcccaac gatcaaggcg agttacatga tcccccatgt tgtgcaaaaa 1500
agcggtttagc tccctcggtc ctccgacgtg tgtcagaagt aagttggcgc cagtgttacc 1560
actcatggtt atggcagcac tgcataatct tcttactgtc atgccatccg taagatgctt 1620
ttctgtgact ggtgagtact caaccaagtc attctgagaa tagtgtatgc ggcgaccgag 1680
ttgtctttgc ccggcgtaaa tacgggataa taccgcgcca catagcagaa ctttaaaagt 1740
gctcatcatt ggaaaacgtt cttcggggcg aaaactctca aggatcttac cgctgttgag 1800
atccagttcg atgtaaccca ctctgcacc caactgatct tcagcatctt ttactttcac 1860
cagcgtttct gggtagcaa aaacaggaag gcaaaaagg gcaataaggc 1920
gacacgaaaa tgttgaatac tcatattctt cctttttcaa tattattgaa gcatttatca 1980
gggttattgt ctcatgagcg gatacatatt tgaatgtatt tagaaaaata aacaaatagg 2040
ggtcagtggt acaaccaatt aaccaattct gaacattatc gcgagcccat ttatacctga 2100
atatggctca taacaccctt tgcagtgcga ctaacggcat gaagctcgtc ggggaaataa 2160
tgattttatt ttgactgata gtgacctgtt cgttgcaaca aattgataag caatgcttcc 2220
ttataatgcc aactttgtac aagaaagctg ggtccatgga agctcgacga agcgctaatt 2280
atgaaccaa tagttgggac tacgattttc tattatctc tgatacggat gagtccattg 2340
aagtttataa ggataaagct aagaaattgg aagccgaagt gcgcgcgcaa attaacaatg 2400
agaaagcgga atttttgacc ttattagaac tcatcgataa tgtgcaacga ctgggatttg 2460
gctatcggtt tgaaagtac atccgcgggg cactggatcg ttttgatct agtggcggct 2520
ttgatggcgt cactaaaact agtttgacgc cgaccgcaat cagttttcgg ctattacgtc 2580
aacacgggtt cgaagtgaat caagaggcgt ttagtggctt caaagatcaa aatggcaatt 2640
ttctgaaaaa cttgaaagaa gacacaaaag ctatcctaag tttatacga gctagtttc 2700
tcgcgctgga aggtgaaaat attctggatg aggtcgtgt atttgcaatt tctcacctga 2760
aagaattatc tgaagaaaag attggcaag aactcgccga acaggtaaat cagcccttg 2820
aactgccccct ccactcgtcgt acccaacgat tggaaagctg gtggagtac gaagcctatc 2880
gcaagaaaga agacgctaac caagttttgt tggaaactgg catcttggtat tataacatga 2940
ttcaatccgt atatcagcgc gatctacgtg aaacgtctcg gtgggtggcg cgtgttgggc 3000
tcgctactaa attacatttt gcaaaggatc gactcattga atccttttat tgggcccgtc 3060
gggtggcttt tgaaccccag tacagcgatt gccgtaattc tgtagcaaaa atgttttctt 3120
tcgttaacaat tattgatgac atttatgacg tttacggcac cctcgacgaa ctggaattgt 3180
tcactgacgc tgtggaacgt tgggacgtaa atgccattaa tgacctgcca gattacatga 3240
agttgtgttt tctcgcgtta tataacacca ttaatgaaat tgcatacgac aatttaaagg 3300
ataagggaga gaattattct ccttatttga cgaaagcctg ggccgatttg tgtaatgcct 3360
ttttgcagga agctaaatgg ttatataaca aatccacccc cacttttgat gactattttg 3420
gcaatgcctg gaagagcagc agcgggcctc tccaactgat ttttgcttat tttcggtag 3480
tacaaaacat taagaaagaa gagattgaaa atttgcaaaa gtacctgac attattagtc 3540
ggcccagtc tattttccgc ttgtgcaacg acctggcatc cgctagtgc gaaattgcgc 3600
gtggcgaaac agctaatagt gtgagttgtt acatgcgcac aaagggcatt tccgaagaac 3660
tagctacgga aagtgtcatg aacctgattg acgagacttg caagaaaatg aataaggaaa 3720
aattgggcgg gtcctatttt gccaaacctt ttgtggaac cgcgattaat ttggctgcgc 3780
aaagtcatatg tacctatcac aatgggtgat ctcacaccag tcccgatgaa ttaaccgta 3840
aacagagttct gtctgtgatt actgaaccca ttttgccctt tgaacgttaa aagtaacagg 3900
ttttccatgt tgtcgtctgc aagaacactg cagagcctgc tttttgtac aaagtggca 3960
ttata-

```

<210> 6
 <211> 544
 <212> PRT
 <213> Sequência artificial
 <220>
 <223> Descrição de sequência artificial: Gene de isopreno sintase (IspS) de 65 kD de *Populus alba* com proteína de fusão da sequência de direcionamento de cloroplasto omitida a partir do plasmídeo de expressão plspS de cianobactéria *Synechocystis* sp. (Ss-IspS)
 <400> 6
 Met Glu Ala Arg Arg Ser Ala Asn Tyr Glu Pro Asn Ser Trp Asp Tyr
 1 5 10 15
 Asp Phe Leu Leu Ser Ser Asp Thr Asp Glu Ser Ile Glu Val Tyr Lys
 20 25 30
 Asp Lys Ala Lys Lys Leu Glu Ala Glu Val Arg Arg Glu Ile Asn Asn
 35 40 45
 Glu Lys Ala Glu Phe Leu Thr Leu Leu Glu Leu Ile Asp Asn Val Gln
 50 55 60
 Arg Leu Gly Leu Gly Tyr Arg Phe Glu Ser Asp Ile Arg Arg Ala Leu
 65 70 75 80
 Asp Arg Phe Val Ser Ser Gly Gly Phe Asp Gly Val Thr Lys Thr Ser
 85 90 95
 Leu His Ala Thr Ala Leu Ser Phe Arg Leu Leu Arg Gln His Gly Phe
 100 105 110
 Glu Val Ser Gln Glu Ala Phe Ser Gly Phe Lys Asp Gln Asn Gly Asn
 115 120 125
 Phe Leu Glu Asn Leu Lys Glu Asp Thr Lys Ala Ile Leu Ser Leu Tyr
 130 135 140
 Glu Ala Ser Phe Leu Ala Leu Glu Gly Glu Asn Ile Leu Asp Glu Ala
 145 150 155 160
 Arg Val Phe Ala Ile Ser His Leu Lys Glu Leu Ser Glu Glu Lys Ile
 165 170 175
 Gly Lys Glu Leu Ala Glu Gln Val Asn His Ala Leu Glu Leu Pro Leu
 180 185 190
 His Arg Arg Thr Gln Arg Leu Glu Ala Val Trp Ser Ile Glu Ala Tyr
 195 200 205
 Arg Lys Lys Glu Asp Ala Asn Gln Val Leu Leu Glu Leu Ala Ile Leu
 210 215 220
 Asp Tyr Asn Met Ile Gln Ser Val Tyr Gln Arg Asp Leu Arg Glu Thr
 225 230 235 240
 Ser Arg Trp Trp Arg Arg Val Gly Leu Ala Thr Lys Leu His Phe Ala
 245 250 255

Lys Asp Arg Leu Ile Glu Ser Phe Tyr Trp Ala Val Gly Val Ala Phe
 260 265 270
 Glu Pro Gln Tyr Ser Asp Cys Arg Asn Ser Val Ala Lys Met Phe Ser
 275 280 285
 Phe Val Thr Ile Ile Asp Asp Ile Tyr Asp Val Tyr Gly Thr Leu Asp
 290 295 300
 Glu Leu Glu Leu Phe Thr Asp Ala Val Glu Arg Trp Asp Val Asn Ala
 305 310 315 320
 Ile Asn Asp Leu Pro Asp Tyr Met Lys Leu Cys Phe Leu Ala Leu Tyr
 325 330 335
 Asn Thr Ile Asn Glu Ile Ala Tyr Asp Asn Leu Lys Asp Lys Gly Glu
 340 345 350
 Asn Ile Leu Pro Tyr Leu Thr Lys Ala Trp Ala Asp Leu Cys Asn Ala
 355 360 365
 Phe Leu Gln Glu Ala Lys Trp Leu Tyr Asn Lys Ser Thr Pro Thr Phe
 370 375 380
 Asp Asp Tyr Phe Gly Asn Ala Trp Lys Ser Ser Ser Gly Pro Leu Gln
 385 390 395 400
 Leu Ile Phe Ala Tyr Phe Ala Val Val Gln Asn Ile Lys Lys Glu Glu
 405 410 415
 Ile Glu Asn Leu Gln Lys Tyr His Asp Ile Ile Ser Arg Pro Ser His
 420 425 430
 Ile Phe Arg Leu Cys Asn Asp Leu Ala Ser Ala Ser Ala Glu Ile Ala
 435 440 445
 Arg Gly Glu Thr Ala Asn Ser Val Ser Cys Tyr Met Arg Thr Lys Gly
 450 455 460
 Ile Ser Glu Glu Leu Ala Thr Glu Ser Val Met Asn Leu Ile Asp Glu
 465 470 475 480
 Thr Cys Lys Lys Met Asn Lys Glu Lys Leu Gly Gly Ser Leu Phe Ala
 485 490 495
 Lys Pro Phe Val Glu Thr Ala Ile Asn Leu Ala Arg Gln Ser His Cys
 500 505 510
 Thr Tyr His Asn Gly Asp Ala His Thr Ser Pro Asp Glu Leu Thr Arg
 515 520 525
 Lys Arg Val Leu Ser Val Ile Thr Glu Pro Ile Leu Pro Phe Glu Arg
 530 535 540

<210> 7
 <211> 3133
 <212> DNA

<213> Pueraria montana var. lobata

<220>

<223> cDNA de Isopreno sintase (IspS) de uva kudzu

<400> 7

```

aatcaatata taatatttac ggaagatttg atgcctttcc tgatttttaatt ttattttttat 60
ccctgcataa aataattgtg gtcaccgtac actgttcttg tcaacttggaac aagaaatttg 120
actagcaagc aaggtataat cattcatcta aacttatggt gattttattgc cccacctcat 180
caattttcgt gtgttttatt ttagtgctct tggatcctcg ttccaatata aaaggagaac 240
atggcatcgc aatttttagag catatcattg aaaagtcacg gcaaccaacc ttttatgctt 300
gtctaataaa ttatcgtccc ccacaccaac accaagtact agattttccac aaagtaagaa 360
cttcatcaca caaaaaacat ctcttgccaa tcccaaacct tggcgagtta tttgtgctac 420
gagctctcaa tttacccaaa taacagaaca taatagtcgg cgctcagcta attaccagcc 480
aaacctctgg aattttgaat ttctgcagtc tctggaaaat gaccttaagg tgattatata 540
tatattccag ttaatttttc tttttttctt ttgtgatttt taaggaatca ttttagtttg 600
gaaagtattt tttttatttg cacttttaat tataaaaaatg ttatatcatt ttcacttttt 660
tctattcatt ttcaaaattt tacatagaaa acagtaaaatt ttttattttt tttattttct 720
attttcatta tttctcaaat caaacgggat taaagcataa acaaagaaat taatatttgt 780
cttttaattt tattttttta caataatggg aacgattata tattaggctg accttaataa 840
gttatttttt ttttataata ttgttcttat tghtaacctaa cgacaggtgg aacctataga 900
agagaaggca acaaagctag aggaggaggt acgatgcacg atcaacagag tagacacaca 960
accattaagc ttactagaat tgatcgacga tgtccagcgt ctaggattga cctacaagtt 1020
tgagaaggac ataatacaag cccttgagaa tattgttttg ctggatgaga ataagaaaaa 1080
taaaagtgac ctccatgcta ctgctctcag ctcccgttta cttagacaac atggccttga 1140
ggtttcccaa ggtatttatg tatatatatg ttacccaact agcaacatat atatatatat 1200
atattatgat tcaactgacca tgcattgtgt gcagatgtgt ttgagagatt taaggacaag 1260
gaggagggtt tcagtgggtg acttaaaagt gatgtgcaag ggttgctgag tctatatgaa 1320
gcatcctatc ttggctttga gggagaaaaa ctcttgaggg aggcaaggac attttcaata 1380
acacatctca agaacaacct aaaagaagga ataaacacca aagtggcaga acaagttagt 1440
catgcactgg aacttcctta tcatcaaaag ttgcatagac tagaagcacg atggttcctt 1500
gacaaatatg aaccaaagga accccaccat cagttactac tccagcttgc aaagctagat 1560
ttcaatatgg tgcaaacatt gcaccagaaa gaactgcaag acctgtcaag gttagaaatt 1620
tcaattctca agtaattatt acctataaga aaattaaata acaataacaa tattgagtgt 1680
agagatttcc aattaaaaat taacatacga gaggatcaat atatatctct aggtatgtgg 1740
tactaatgaa atatatgcta ggtggtggac ggagatgggg ctagcaagca agctagactt 1800
tgtccgagac agattaatgg aagtgtattt ttgggcgttg ggaatggcac ctgatcctca 1860
attcgggtgaa tgtcgtaaag ctgtcactaa aatgttttga ttggtcacca tcatcgatga 1920
tgtatatgac gtttatggta ctttgatgta gctacaactc ttcactgatg ctggtgagag 1980
gttcgtaatt gatttcagtc tcgattcagt tggaaattta ttattgctta attaataata 2040
acttgcgtac atgcatacac acagatggga cgtgaatgcc ataaacacac ttccagacta 2100
catgaagttg tgcttcctag cactttataa caccgtcaat gacacgtctt atagcatcct 2160
taaaagaaaaa ggacacaaca acctttccta tttgacaaaa tctgtacata tataactaatt 2220
atctccttgg ttgattaatt agtttagttt agtttagttg gtatgtcaac acaattaatt 2280
aatattatat atggatgttg acagtggcgt gagttatgca aagcattcct tcaagaagca 2340
aaatggctga acaacaaaaa cattccagca tttagcaagt acctggaaaa tgcctcgggtg 2400
tctcctcctg gtgtggcttt gcttgcctct tctacttct actgatttcc atggccttgt gcgtcctca 2520
gatattctag accatgctct tcgttcttta actgatttcc gctacctcag cgggtgtgtaa ttaattacct 2580
tgcgctcatt tcagactctg caatgatttg gctacctcag gtgtctgtaa attactacag 2640
taattaattt gtaacacttg ttagactaat atatataggt gtgtctgtaa tttatatgca tgagaatgac 2700
gctgagctag agaggggtga gacgacaaat ttagaaaaat tgatcgatgc agagtggag 2760
ggcacttctg aagagcaagc acgtgaggag ttgagaaaaa cactacttcc caaaaagctt tatggaaata 2820
aagatgaacc gagagcgagt ttcagattct acataccaat atggagacgg acttggaagg 2880
gctgttaaca tggctcgagt ttgcatttgc acataccaat atggagacgg acttggaagg 2880
ccagactaac ccacagagaa tagaatcaag ttgctactta tagaccctt tccaatcaat 2940
caactaatgt acgtgtaaca acacaatata aacacttttc tacaagtata tatttgttta 3000
atctcgggtg tgaattaggg gtcaacacag ctatatatac ttcaatggac caactcaacc 3060
aatctgataa gagaaaaaaa ataaaaataa ggtaggtta actttgtata aatccaagtt 3120
agatatcaag ttt 3133

```

<210> 8

<211> 608

<212> PRT

<213> Pueraria montana var. lobata

<220>

<223> Isopreno sintase (IspS) de uva kudzu

<400> 8

Met	Ala	Thr	Asn	Leu	Leu	Cys	Leu	Ser	Asn	Lys	Leu	Ser	Ser	Pro	Thr	1	5	10	15
Pro	Thr	Pro	Ser	Thr	Arg	Phe	Pro	Gln	Ser	Lys	Asn	Phe	Ile	Thr	Gln	20	25	30	
Lys	Thr	Ser	Leu	Ala	Asn	Pro	Lys	Pro	Trp	Arg	Val	Ile	Cys	Ala	Thr	35	40	45	
Ser	Ser	Gln	Phe	Thr	Gln	Ile	Thr	Glu	His	Asn	Ser	Arg	Arg	Ser	Ala	50	55	60	
Asn	Tyr	Gln	Pro	Asn	Leu	Trp	Asn	Phe	Glu	Phe	Leu	Gln	Ser	Leu	Glu	65	70	75	80
Asn	Asp	Leu	Lys	Val	Glu	Lys	Leu	Glu	Glu	Lys	Ala	Thr	Lys	Leu	Glu	85	90	95	
Glu	Glu	Val	Arg	Cys	Met	Ile	Asn	Arg	Val	Asp	Thr	Gln	Pro	Leu	Ser	100	105	110	
Leu	Leu	Glu	Leu	Ile	Asp	Asp	Val	Gln	Arg	Leu	Gly	Leu	Thr	Tyr	Lys	115	120	125	
Phe	Glu	Lys	Asp	Ile	Ile	Lys	Ala	Leu	Glu	Asn	Ile	Val	Leu	Leu	Asp	130	135	140	
Glu	Asn	Lys	Lys	Asn	Lys	Ser	Asp	Leu	His	Ala	Thr	Ala	Leu	Ser	Phe	145	150	155	160
Arg	Leu	Leu	Arg	Gln	His	Gly	Phe	Glu	Val	Ser	Gln	Asp	Val	Phe	Glu	165	170	175	
Arg	Phe	Lys	Asp	Lys	Glu	Gly	Gly	Phe	Ser	Gly	Glu	Leu	Lys	Gly	Asp	180	185	190	
Val	Gln	Gly	Leu	Leu	Ser	Leu	Tyr	Glu	Ala	Ser	Tyr	Leu	Gly	Phe	Glu	195	200	205	
Gly	Glu	Asn	Leu	Leu	Glu	Glu	Ala	Arg	Thr	Phe	Ser	Ile	Thr	His	Leu	210	215	220	
Lys	Asn	Asn	Leu	Lys	Glu	Gly	Ile	Asn	Thr	Lys	Val	Ala	Glu	Gln	Val	225	230	235	240
Ser	His	Ala	Leu	Glu	Leu	Pro	Tyr	His	Gln	Arg	Leu	His	Arg	Leu	Glu	245	250	255	
Ala	Arg	Trp	Phe	Leu	Asp	Lys	Tyr	Glu	Pro	Lys	Glu	Pro	His	His	Gln	260	265	270	

Leu Leu Leu Glu Leu Ala Lys Leu Asp Phe Asn Met Val Gln Thr Leu
 275 280 285
 His Gln Lys Glu Leu Gln Asp Leu Ser Arg Trp Trp Thr Glu Met Gly
 290 295 300
 Leu Ala Ser Lys Leu Asp Phe Val Arg Asp Arg Leu Met Glu Val Tyr
 305 310 315 320
 Phe Trp Ala Leu Gly Met Ala Pro Asp Pro Gln Phe Gly Glu Cys Arg
 325 330 335
 Lys Ala Val Thr Lys Met Phe Gly Leu Val Thr Ile Ile Asp Asp Val
 340 345 350
 Tyr Asp Val Tyr Gly Thr Leu Asp Glu Leu Gln Leu Phe Thr Asp Ala
 355 360 365
 Val Glu Arg Trp Asp Val Asn Ala Ile Asn Thr Leu Pro Asp Tyr Met
 370 375 380
 Lys Leu Cys Phe Leu Ala Leu Tyr Asn Thr Val Asn Asp Thr Ser Tyr
 385 390 395 400
 Ser Ile Leu Lys Glu Lys Gly His Asn Asn Leu Ser Tyr Leu Thr Lys
 405 410 415
 Ser Trp Arg Glu Leu Cys Lys Ala Phe Leu Gln Glu Ala Lys Trp Ser
 420 425 430
 Asn Asn Lys Ile Ile Pro Ala Phe Ser Lys Tyr Leu Glu Asn Ala Ser
 435 440 445
 Val Ser Ser Ser Gly Val Ala Leu Leu Ala Pro Ser Tyr Phe Ser Val
 450 455 460
 Cys Gln Gln Gln Glu Asp Ile Ser Asp His Ala Leu Arg Ser Leu Thr
 465 470 475 480
 Asp Phe His Gly Leu Val Arg Ser Ser Cys Val Ile Phe Arg Leu Cys
 485 490 495
 Asn Asp Leu Ala Thr Ser Ala Ala Glu Leu Glu Arg Gly Glu Thr Thr
 500 505 510
 Asn Ser Ile Ile Ser Tyr Met His Glu Asn Asp Gly Thr Ser Glu Glu
 515 520 525
 Gln Ala Arg Glu Glu Leu Arg Lys Leu Ile Asp Ala Glu Trp Lys Lys
 530 535 540
 Met Asn Arg Glu Arg Val Ser Asp Ser Thr Leu Leu Pro Lys Ala Phe
 545 550 555 560
 Met Glu Ile Ala Val Asn Met Ala Arg Val Ser His Cys Thr Tyr Gln
 565 570 575
 Tyr Gly Asp Gly Leu Gly Arg Pro Asp Tyr Ala Thr Glu Asn Arg Ile
 580 585 590

Lys Leu Leu Leu Ile Asp Pro Phe Pro Ile Asn Gln Leu Met Tyr Val
 595 600 605

<210> 9

<211> 595

<212> PRT

<213> Populus tremuloides

<220>

<223> Isopreno sintase de (IspS) quaking aspen

<400> 9

Met Ala Thr Glu Leu Leu Cys Leu His Arg Pro Ile Ser Leu Thr His
 1 5 10 15

Lys Leu Phe Arg Asn Pro Leu Pro Lys Val Ile Gln Ala Thr Pro Leu
 20 25 30

Thr Leu Lys Leu Arg Cys Ser Val Ser Thr Glu Asn Val Ser Phe Ser
 35 40 45

Glu Thr Glu Thr Glu Thr Arg Arg Ser Ala Asn Tyr Glu Pro Asn Ser
 50 55 60

Trp Asp Tyr Asp Tyr Leu Leu Ser Ser Asp Thr Asp Glu Ser Ile Glu
 65 70 75 80

Val His Lys Asp Lys Ala Lys Lys Leu Glu Ala Glu Val Arg Arg Glu
 85 90 95

Ile Asn Asn Glu Lys Ala Glu Phe Leu Thr Leu Leu Glu Leu Ile Asp
 100 105 110

Asn Val Gln Arg Leu Gly Leu Gly Tyr Arg Phe Glu Ser Asp Ile Arg
 115 120 125

Arg Ala Leu Asp Arg Phe Val Ser Ser Gly Gly Phe Asp Gly Val Thr
 130 135 140

Lys Thr Ser Leu His Gly Thr Ala Leu Ser Phe Arg Leu Leu Arg Gln
 145 150 155 160

His Gly Phe Glu Val Ser Gln Glu Ala Phe Ser Gly Phe Lys Asp Gln
 165 170 175

Asn Gly Asn Phe Leu Glu Asn Leu Lys Glu Asp Ile Lys Ala Ile Leu
 180 185 190

Ser Leu Tyr Glu Ala Ser Phe Leu Ala Leu Glu Gly Glu Asn Ile Leu
 195 200 205

Asp Glu Ala Lys Val Phe Ala Ile Ser His Leu Lys Glu Leu Ser Glu
 210 215 220

Glu Lys Ile Gly Lys Glu Leu Ala Glu Gln Val Ser His Ala Leu Glu
 225 230 235 240

Leu Pro Leu His Arg Arg Thr Gln Arg Leu Glu Ala Val Trp Ser Ile
 245 250 255

Glu Ala Tyr Arg Lys Lys Glu Asp Ala Asn Gln Val Leu Leu Glu Leu
 260 265 270
 Ala Ile Leu Asp Tyr Asn Met Ile Gln Ser Val Tyr Gln Arg Asp Leu
 275 280 285
 Arg Glu Thr Ser Arg Trp Trp Arg Arg Val Gly Leu Ala Thr Lys Leu
 290 295 300
 His Phe Ala Arg Asp Arg Leu Ile Glu Ser Phe Tyr Trp Ala Val Gly
 305 310 315 320
 Val Ala Phe Glu Pro Gln Tyr Ser Asp Cys Arg Asn Ser Val Ala Lys
 325 330 335
 Met Phe Ser Phe Val Thr Ile Ile Asp Asp Ile Tyr Asp Val Tyr Gly
 340 345 350
 Thr Leu Asp Glu Leu Glu Leu Phe Thr Asp Ala Val Glu Arg Trp Asp
 355 360 365
 Val Asn Ala Ile Asn Asp Leu Pro Asp Tyr Met Lys Leu Cys Phe Leu
 370 375 380
 Ala Leu Tyr Asn Thr Ile Asn Glu Ile Ala Tyr Asp Asn Leu Lys Asp
 385 390 395 400
 Lys Gly Glu Asn Ile Leu Pro Tyr Leu Thr Lys Ala Trp Ala Asp Leu
 405 410 415
 Cys Asn Ala Phe Leu Gln Glu Ala Lys Trp Leu Tyr Asn Lys Ser Thr
 420 425 430
 Pro Thr Phe Asp Asp Tyr Phe Gly Asn Ala Trp Lys Ser Ser Ser Gly
 435 440 445
 Pro Leu Gln Leu Ile Phe Ala Tyr Phe Ala Val Val Gln Asn Ile Lys
 450 455 460
 Lys Glu Glu Ile Glu Asn Leu Gln Lys Tyr His Asp Ile Ile Ser Arg
 465 470 475 480
 Pro Ser His Ile Phe Arg Leu Cys Asn Asp Leu Ala Ser Ala Ser Ala
 485 490 495
 Glu Ile Ala Arg Gly Glu Thr Ala Asn Ser Val Ser Cys Tyr Met Arg
 500 505 510
 Thr Lys Gly Ile Ser Glu Glu Leu Ala Thr Glu Ser Val Met Asn Leu
 515 520 525
 Ile Asp Glu Thr Trp Lys Lys Met Asn Lys Glu Lys Leu Gly Gly Ser
 530 535 540
 Leu Phe Ala Lys Pro Phe Val Glu Thr Ala Ile Asn Leu Ala Arg Gln
 545 550 555 560
 Ser His Cys Thr Tyr His Asn Gly Asp Ala His Thr Ser Pro Asp Glu
 565 570 575

Leu Thr Arg Lys Arg Val Leu Ser Val Ile Thr Glu Pro Ile Leu Pro
 580 585 590

Phe Glu Arg
 595

<210> 10

<211> 595

<212> PRT

<213> Populus nigra

<220>

<223> Isopreno sintase (IspS) de álamo Lombardy

<400> 10

Met Ala Thr Glu Leu Leu Cys Leu His Arg Pro Ile Ser Leu Thr His
 1 5 10 15

Lys Leu Phe Arg Asn Pro Leu Pro Lys Val Ile Gln Ala Thr Pro Leu
 20 25 30

Thr Leu Lys Leu Arg Cys Ser Val Ser Thr Glu Asn Val Ser Phe Thr
 35 40 45

Glu Thr Glu Thr Glu Thr Arg Arg Ser Ala Asn Tyr Glu Pro Asn Ser
 50 55 60

Trp Asp Tyr Asp Tyr Leu Leu Ser Ser Asp Thr Asp Glu Ser Ile Glu
 65 70 75 80

Val Tyr Lys Asp Lys Ala Lys Lys Leu Glu Ala Glu Val Arg Arg Glu
 85 90 95

Ile Asn Asn Glu Lys Ala Glu Phe Leu Thr Leu Pro Glu Leu Ile Asp
 100 105 110

Asn Val Gln Arg Leu Gly Leu Gly Tyr Arg Phe Glu Ser Asp Ile Arg
 115 120 125

Arg Ala Leu Asp Arg Phe Val Ser Ser Gly Gly Phe Asp Ala Val Thr
 130 135 140

Lys Thr Ser Leu His Ala Thr Ala Leu Ser Phe Arg Leu Leu Arg Gln
 145 150 155 160

His Gly Phe Glu Val Ser Gln Glu Ala Phe Ser Gly Phe Lys Asp Gln
 165 170 175

Asn Gly Asn Phe Leu Lys Asn Leu Lys Glu Asp Ile Lys Ala Ile Leu
 180 185 190

Ser Leu Tyr Glu Ala Ser Phe Leu Ala Leu Glu Gly Glu Asn Ile Leu
 195 200 205

Asp Glu Ala Lys Val Phe Ala Ile Ser His Leu Lys Glu Leu Ser Glu
 210 215 220

Glu Lys Ile Gly Lys Asp Leu Ala Glu Gln Val Asn His Ala Leu Glu
 225 230 235 240

Leu Pro Leu His Arg Arg Thr Gln Arg Leu Glu Ala Val Trp Ser Ile
 245 250 255
 Glu Ala Tyr Arg Lys Lys Glu Asp Ala Asp Gln Val Leu Leu Glu Leu
 260 265 270
 Ala Ile Leu Asp Tyr Asn Met Ile Gln Ser Val Tyr Gln Arg Asp Leu
 275 280 285
 Arg Glu Thr Ser Arg Trp Trp Arg Arg Val Gly Leu Ala Thr Lys Leu
 290 295 300
 His Phe Ala Arg Asp Arg Leu Ile Glu Ser Phe Tyr Trp Ala Val Gly
 305 310 315 320
 Val Ala Phe Glu Pro Gln Tyr Ser Asp Cys Arg Asn Ser Val Ala Lys
 325 330 335
 Met Phe Ser Phe Val Thr Ile Ile Asp Asp Ile Tyr Asp Val Tyr Gly
 340 345 350
 Thr Leu Asp Glu Leu Glu Leu Phe Thr Asp Ala Val Glu Arg Trp Asp
 355 360 365
 Val Asn Ala Ile Asp Asp Leu Pro Asp Tyr Met Lys Leu Cys Phe Leu
 370 375 380
 Ala Leu Tyr Asn Thr Ile Asn Glu Ile Ala Tyr Asp Asn Leu Lys Asp
 385 390 395 400
 Lys Gly Glu Asn Ile Leu Pro Tyr Leu Thr Lys Ala Trp Ala Asp Leu
 405 410 415
 Cys Asn Ala Phe Leu Gln Glu Ala Lys Trp Leu Tyr Asn Lys Ser Thr
 420 425 430
 Pro Thr Phe Asp Glu Tyr Phe Gly Asn Ala Trp Lys Ser Ser Ser Gly
 435 440 445
 Pro Leu Gln Leu Val Phe Ala Tyr Phe Ala Val Val Gln Asn Ile Lys
 450 455 460
 Lys Glu Glu Ile Asp Asn Leu Gln Lys Tyr His Asp Ile Ile Ser Arg
 465 470 475 480
 Pro Ser His Ile Phe Arg Leu Cys Asn Asp Leu Ala Ser Ala Ser Ala
 485 490 495
 Glu Ile Ala Arg Gly Glu Thr Ala Asn Ser Val Ser Cys Tyr Met Arg
 500 505 510
 Thr Lys Gly Ile Ser Glu Glu Leu Ala Thr Glu Ser Val Met Asn Leu
 515 520 525
 Ile Asp Glu Thr Trp Lys Lys Met Asn Lys Glu Lys Leu Gly Gly Ser
 530 535 540
 Leu Phe Ala Lys Pro Phe Val Glu Thr Ala Ile Asn Leu Ala Arg Gln
 545 550 555 560

Ser His Cys Thr Tyr His Asn Gly Asp Ala His Thr Ser Pro Asp Glu
565 570 575

Leu Thr Arg Lys Arg Val Leu Ser Val Ile Thr Glu Pro Ile Leu Pro
580 585 590

Phe Glu Arg
595

<210> 11

$\langle 211 \rangle$ 7

<212> PRT

<213> Sequência artificial

$\langle 220 \rangle$

<223> Descrição de sequência artificial: sequência inicial de proteína madura de isopreno sintase (IspS) de álamo branco com peptídeo de trânsito de cloroplasto omitido

<400> 11

Cys Ser Val Ser Thr Glu Asn
1 5

<210> 12

<211> 25

<212> DNA

<213> Sequência artificial

<220>

<223> Descrição de sequência artificial: Iniciador de amplificação por PCR IspS_F_NdeI

<400> 12

ctgggtcata tggaagctcg acgaa

25

<210> 13

<211> 29

<212> DNA

<213> Sequência artificial

<220>

<223> Descrição de sequência artificial: Iniciador de amplificação por PCR IspS_R_HindIII

<400> 13

atggaaaacc tgaagctttt aacgttcaa

29

REIVINDICAÇÕES

1. Método para a produção de hidrocarbonetos de isopreno em um microorganismo selecionado do grupo consistindo em microalgas, cianobactérias ou bactérias fotossintéticas, o método compreendendo:
 - 5 introdução de um cassete de expressão que compreende um ácido nucleico que codifica isopreno sintase, no microorganismo; e
 - cultura do microorganismo sob condições nas quais o ácido nucleico que codifica isopreno sintase é expresso.
2. Método de acordo com a reivindicação 1, em que o microor-
10 ganismo é uma microalga que é uma alga verde.
3. Método de acordo com a reivindicação 2, em que a alga verde é selecionada do grupo consistindo em *Chlamydomonas reinhardtii*, *Scenedesmus obliquus*, *Chlorella vulgaris* ou *Dunaliella salina*.
4. Método de acordo com a reivindicação 1, em que o microor-
15 ganismo é uma cianobactéria.
5. Método de acordo com a reivindicação 4, em que a cianobactéria é *Synechocystis* sp.
6. Método de acordo com a reivindicação 1, em que o microor-
ganismo é uma bactéria fotossintética.
- 20 7. Método de acordo com a reivindicação 6, em que a bactéria fotossintética é *Rhodospirillum rubrum*.
8. Método de acordo com a reivindicação 1, em que o ácido nucleico codifica uma isopreno sintase que compreende os resíduos de aminoácido 53 a 595 de SEQ ID NO: 2.
- 25 9. Microorganismo selecionado do grupo consistindo em uma célula de microalga, uma célula de cianobactéria e uma célula de bactéria fotossintética, em que o microorganismo compreende um ácido nucleico heterólogo que codifica isopreno sintase e é operavelmente ligado a um promotor.
- 30 10. Microorganismo de acordo com a reivindicação 9, em que o microorganismo é uma microalga que é uma alga verde.
11. Microorganismo de acordo com a reivindicação 10, em que a

alga verde é selecionada do grupo consistindo em *Chlamydomonas reinhardtii*, *Scenedesmus obliquus*, *Chlorella vulgaris* ou *Dunaliella salina*.

12. Microorganismo de acordo com a reivindicação 9, em que o microorganismo é uma cianobactéria.

5 13. Microorganismo de acordo com a reivindicação 12, em que a cianobactéria é *Synechocystis* sp.

14. Microorganismo de acordo com a reivindicação 9, em que o microorganismo é uma bactéria fotossintética.

10 15. Microorganismo de acordo com a reivindicação 14, em que a bactéria fotossintética é *Rhodospirillum rubrum*.

16. Microorganismo de acordo com a reivindicação 9, em que o ácido nucleico heterólogo codifica uma isopreno sintase que compreende os resíduos de aminoácido 53-595 de SEQ ID NO: 2.

15 17. Método de produção de hidrocarbonetos de isopreno em um microorganismo que compreende um ácido nucleico heterólogo que codifica isopreno sintase e que é selecionado do grupo consistindo em microalgas, cianobactérias, bactérias fotossintéticas e bactérias não-fotossintéticas, o método compreendendo:

20 cultura em massa do microorganismo em um bioreator vedado sob condições nas quais o gene de isopreno sintase é expresso; e coleta dos hidrocarbonetos de isopreno voláteis produzidos pelo microorganismo.

18. Método de acordo com a reivindicação 17, em que o microorganismo é uma microalga que é uma alga verde.

25 19. Método de acordo com a reivindicação 18, em que a alga verde é selecionada do grupo consistindo em *Chlamydomonas reinhardtii*, *Scenedesmus obliquus*, *Chlorella vulgaris* ou *Dunaliella salina*.

20. Método de acordo com a reivindicação 17, em que o microorganismo é uma cianobactéria.

30 21. Método de acordo com a reivindicação 20, em que a cianobactéria é *Synechocystis* sp.

22. Método de acordo com a reivindicação 17, em que o micro-

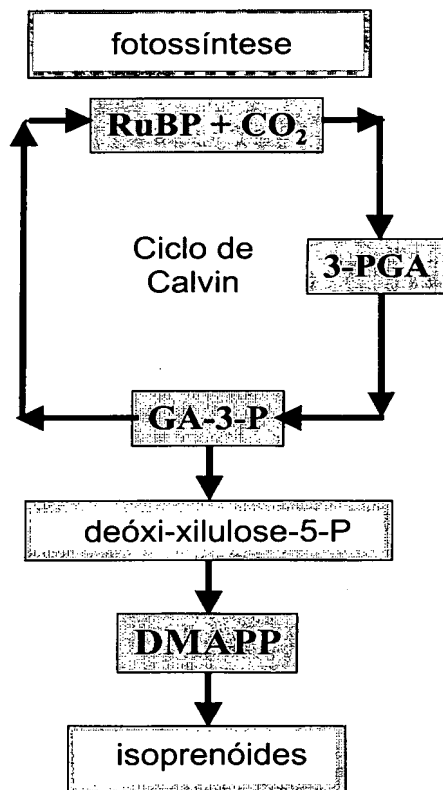
organismo é uma bactéria fotossintética.

23. Método de acordo com a reivindicação 22 em que a bactéria fotossintética é *Rhodospirillum rubrum*.

24. Método de acordo com a reivindicação 17 em que o microor-
5 ganismo é uma bactéria não-fotossintética.

25. Método de acordo com a reivindicação 24 em que a bactéria não-fotossintética é *Escherichia coli*.

26. Método de acordo com a reivindicação 17 em que o ácido nucleico heterólogo codifica uma isopreno sintase que compreende os resí-
10 duos de aminoácido 53 a 595 de SEQ ID NO: 2.



Abreviações usadas:

RuBP = bis-fosfato de ribulose
3-PGA = ácido 3-fosfoglicérico
GA-3-P = gliceraldeído-3-fosfato
DMAPP = dimetilalil-pirofosfato

FIG. 1

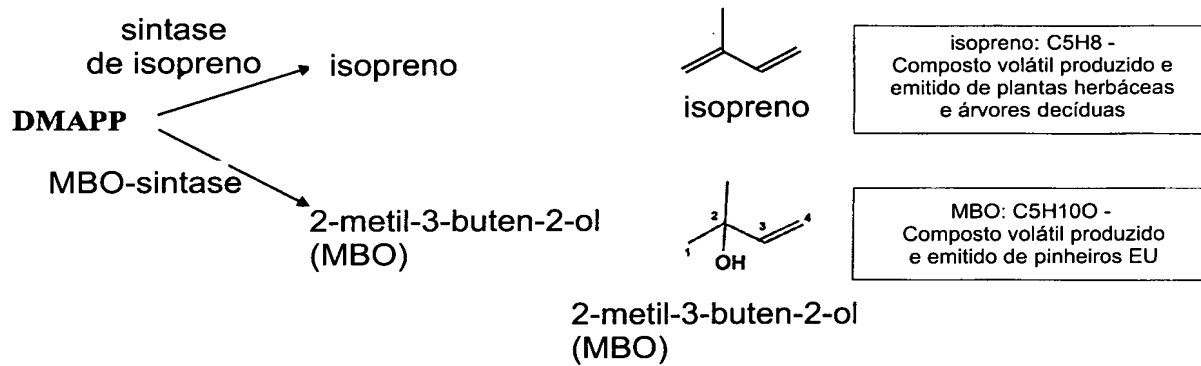
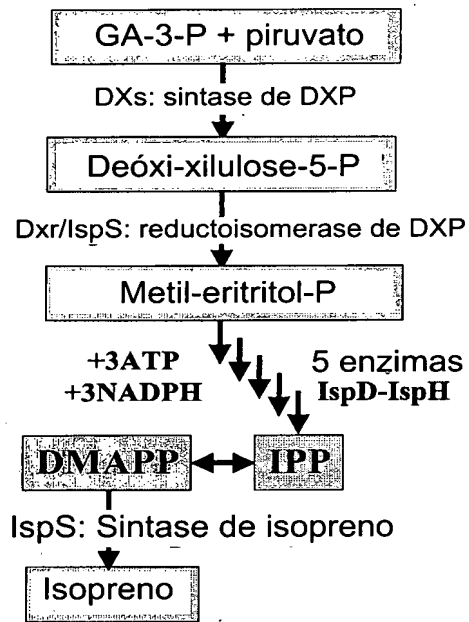


FIG. 2



Abreviações usadas:

GA-3-P = gliceraldeído-3-fosfato

IPP = pirofosfato de isopentenila

DMAPP = dimetilalil-pirofosfato

FIG. 3

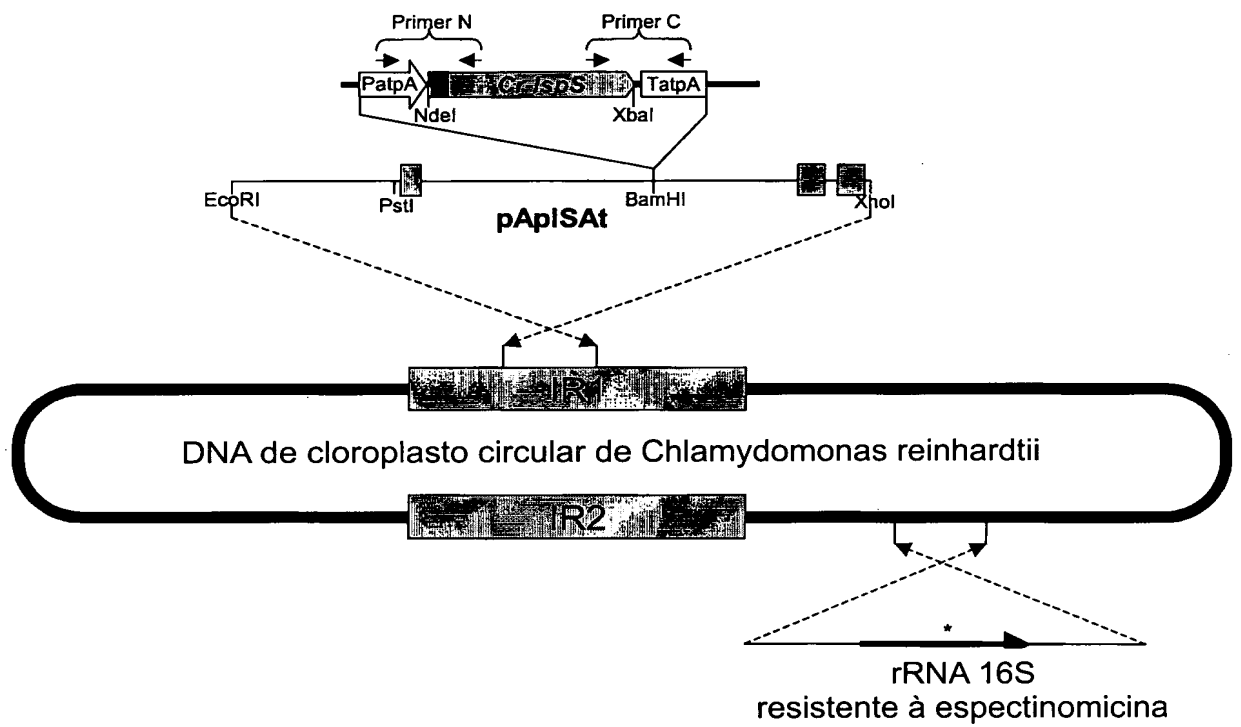


FIG. 4

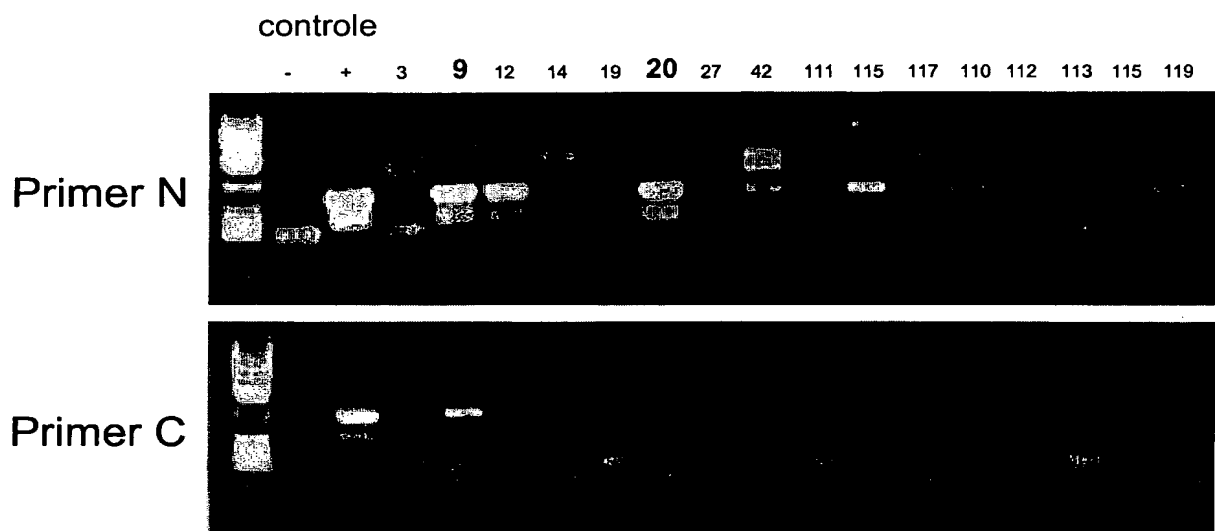


FIG. 5



FIG. 6a

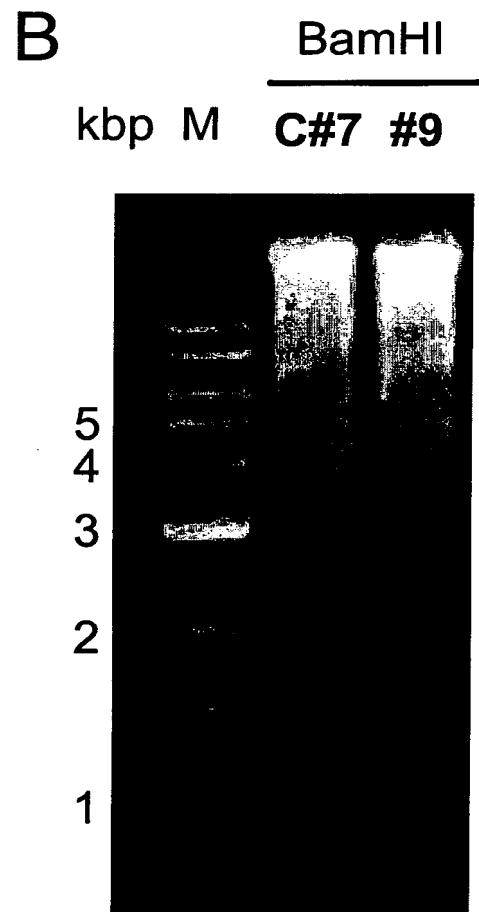


FIG. 6b

FIG. 7

A

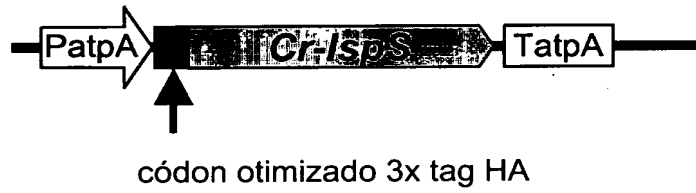
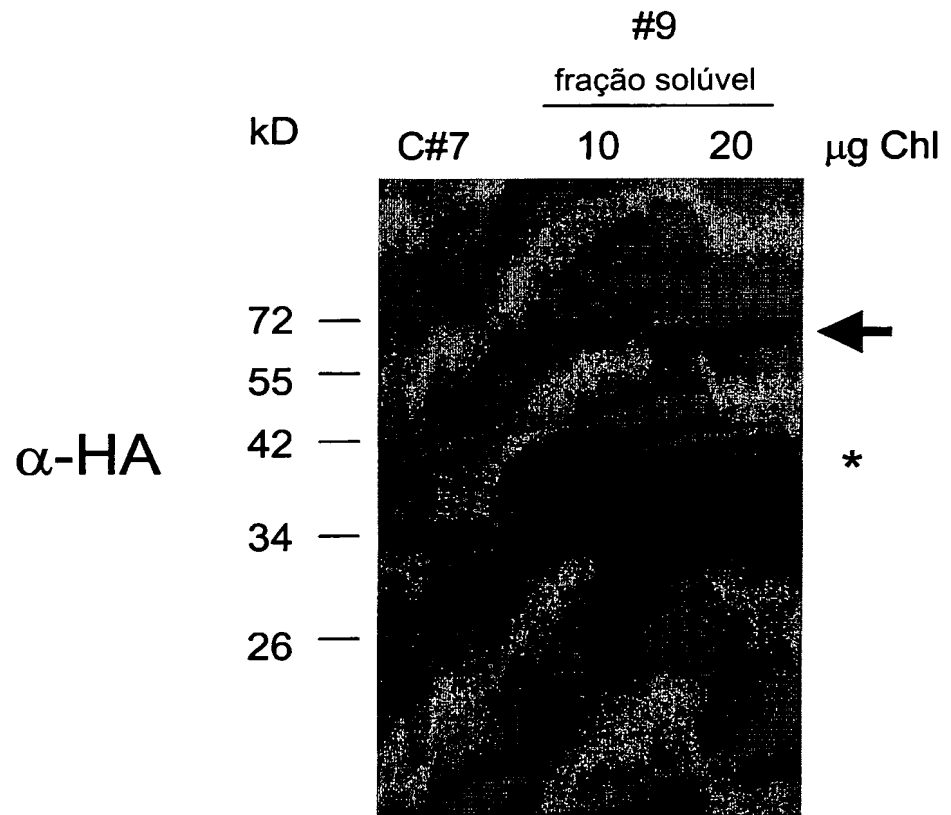


FIG. 7

B



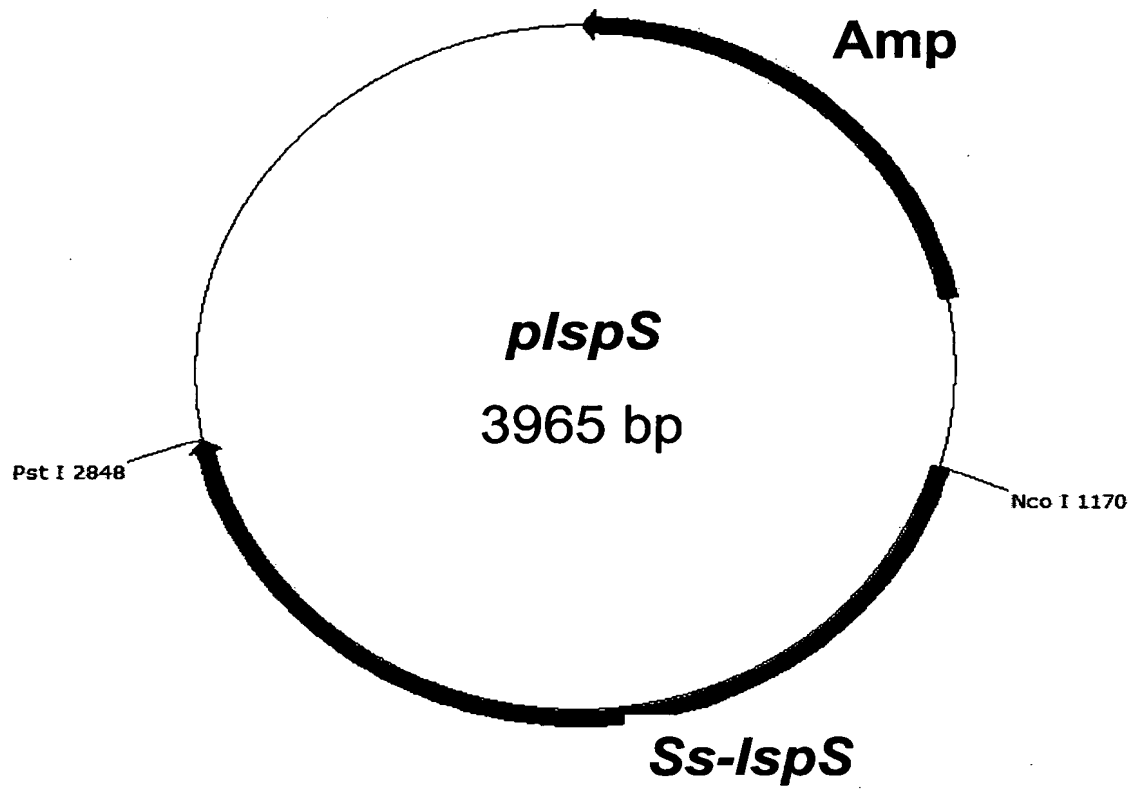


FIG. 8

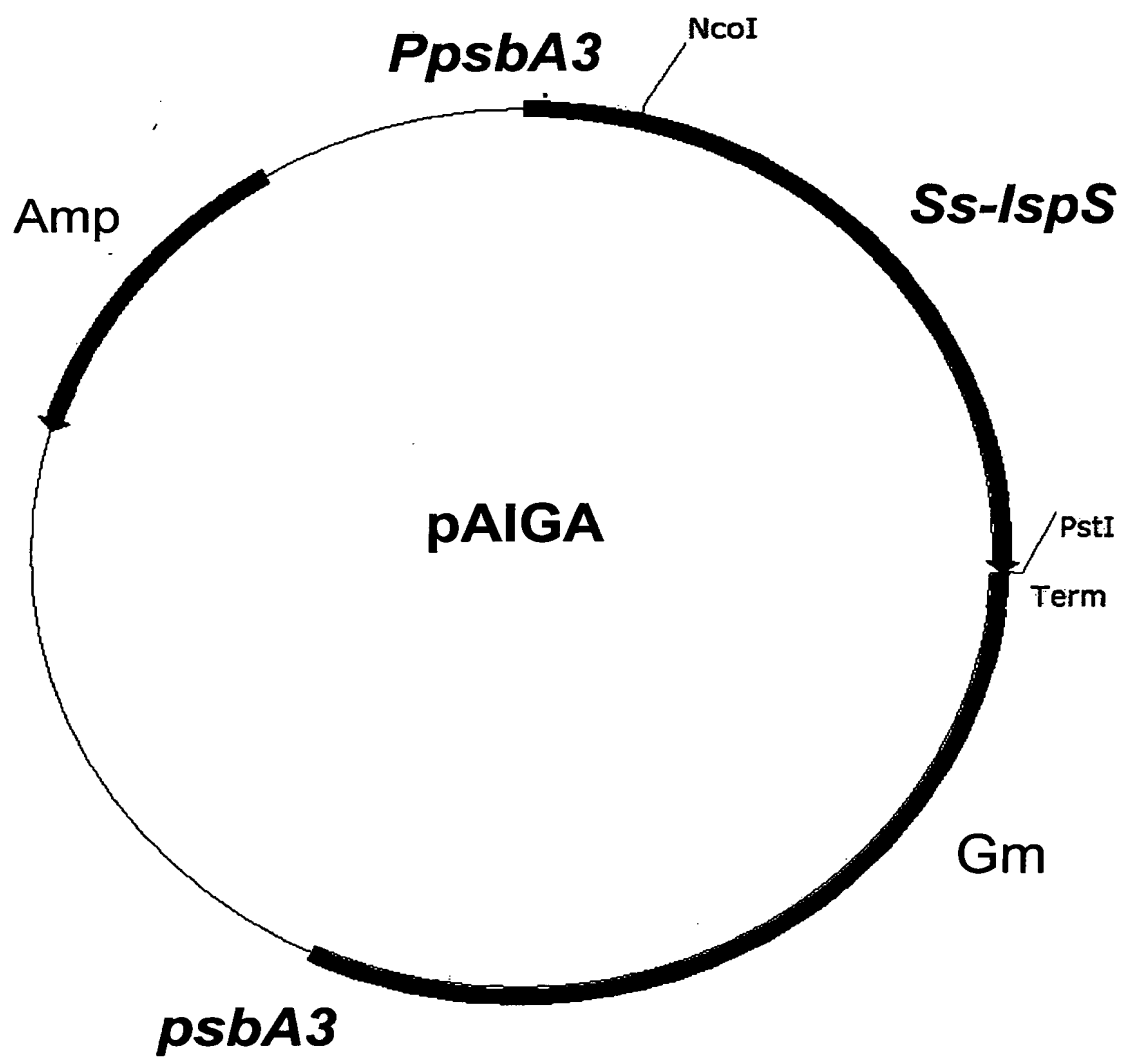


FIG. 9

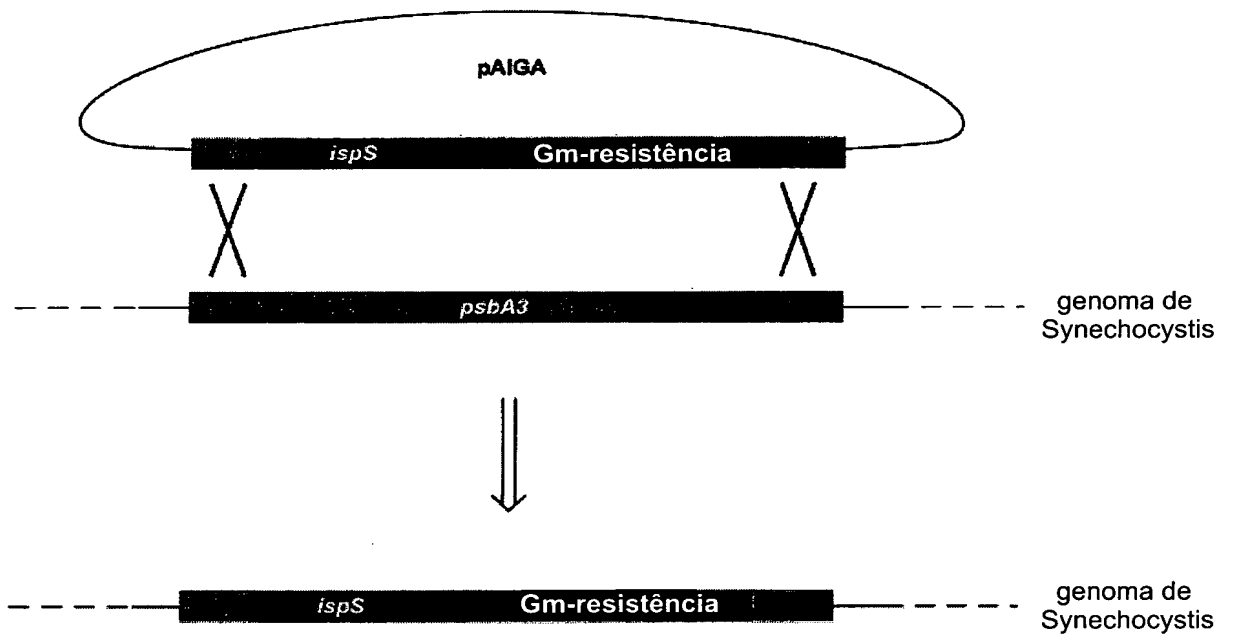


FIG. 10

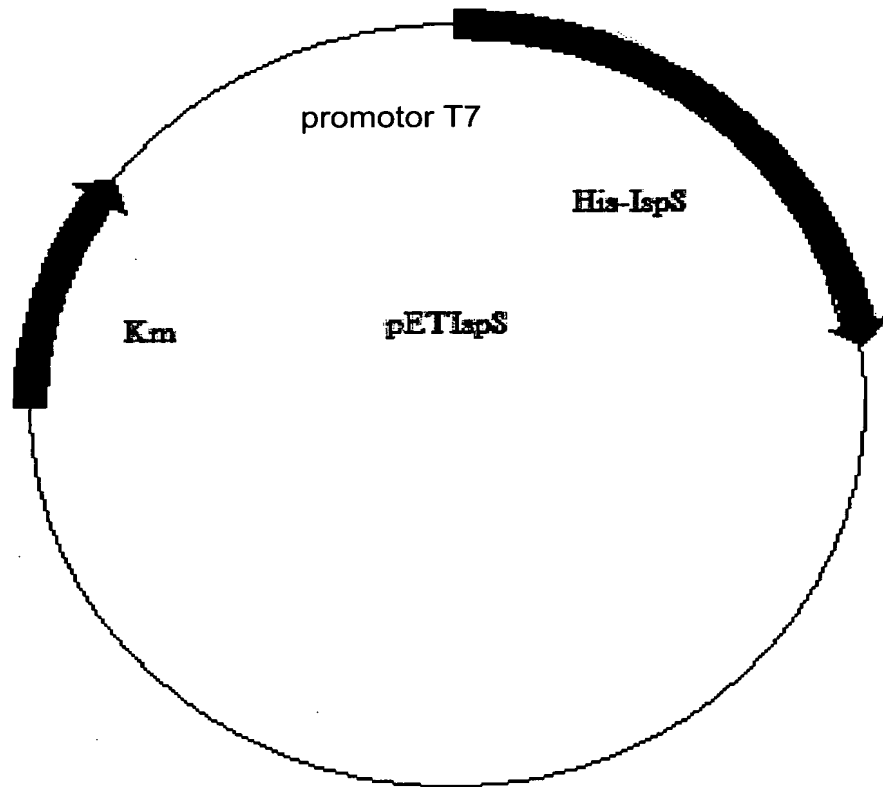
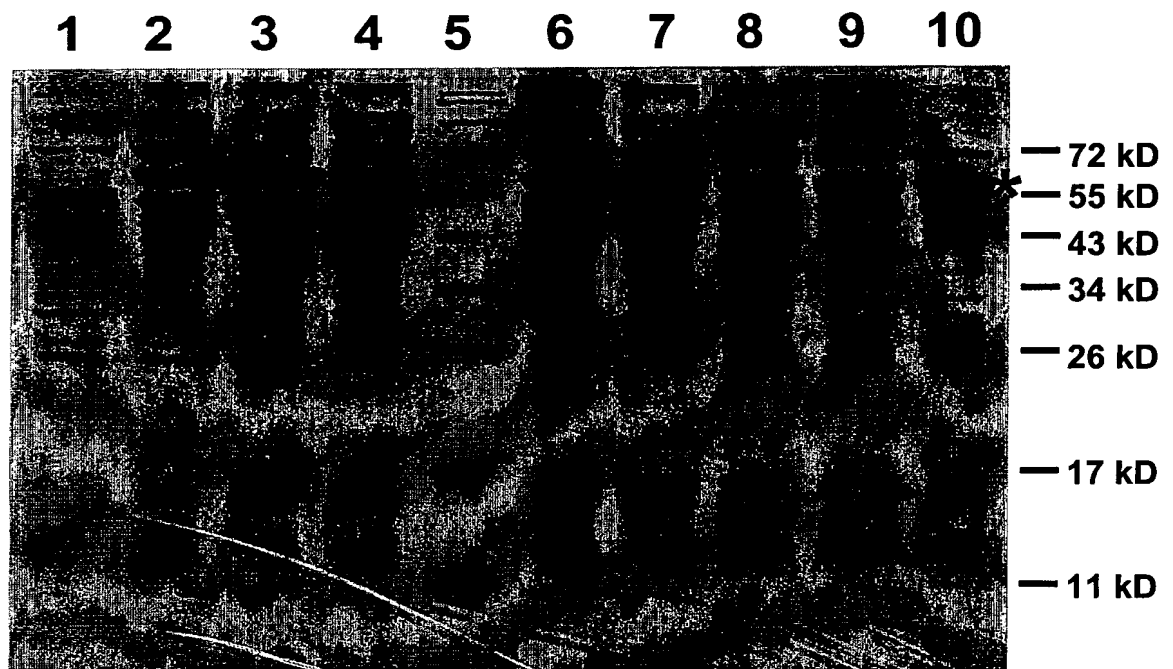


FIG. 11



Tamanho esperado de His-lspS é de 65 kF (asterisco)

FIG. 12

Múltiplos alinhamentos de sequência pelo CLUSTAL W (1.83) de proteínas de IspS conhecidas. O peptídeo de trânsito em cloroplasto putativo cpTP é mostrado pelas sequências sublinhadas; todos os aminoácidos Cys estão em fonte em realçada maior, incluindo substituições conservativas de Ser

```

alba      MATELLGLHRPISLTHKLFNRNLP-----KVIQATPLTLKLRCSVSTENVSTETET 52
tremuloides MATELLGLHRPISLTHKLFNRNLP-----KVIQATPLTLKLRCSVSTENVSTETET 52
nigra      MATELLGLHRPISLTHKLFNRNLP-----KVIQATPLTLKLRCSVSTENVSTETET 52
kudzu      MATNLLGLSNKLSSTPTPTSTRFQSKNFITQKTSLANPKPWRVICATSSQFTQITEHN- 59
          ***:*** . : * . . : * . * . * . : : * : * : . : : * :

alba      EARRSANYEPNSWDYDYLLSSDTDESIEVYKDKAKKLEAEVRREINNEKAEFLTLELID 112
tremuloides ETRRSANYEPNSWDYDYLLSSDTDESIEVHKDKAKKLEAEVRREINNEKAEFLTLELID 112
nigra      ETRRSANYEPNSWDYDYLLSSDTDESIEVYKDKAKKLEAEVRREINNEKAEFLTLELID 112
kudzu      -SRRSANYQPNLWNFEFLQSLNDLKVEKLEEKATKLEEEVRGMINRVDTQPLSLELID 118
          :*****:* *:::* * :.* :* :***** * * * . : : * * * * *

alba      NVQRLGLGYRFESDIRGALDRFVSSGGFDAVTKTSLHGTALSFRLLRQHGFVQSQAESG 172
tremuloides NVQRLGLGYRFESDIRRALDRFVSSGGFVGTVTKTSLHGTALSFRLLRQHGFVQSQAESG 172
nigra      NVQRLGLGYRFESDIRRALDRFVSSGGFDAVTKTSLHGTALSFRLLRQHGFVQSQAESG 172
kudzu      DVQRLGLTYKFEKDI I KALENIVLLD-ENKKNKSDLHGTALSFRLLRQHGFVQSQAESG 177
          :***** :*:*.* * :.* : . : . : * : * . * . * . * . * . * . : * .

alba      FKDQNGNFNLENLKEDIKAILSLYEASFLALEGENILDEAKVFAISHLKELSEEKIGKELA 232
tremuloides FKDQNGNFNLENLKEDIKAILSLYEASFLALEGENILDEAKVFAISHLKELSEEKIGKELA 232
nigra      FKDQNGNFNLENLKEDIKAILSLYEASFLALEGENILDEAKVFAISHLKELSEEKIGKELA 232
kudzu      FKDKGGFSGELKGDVQGLLSLYEASYLGFEGENLLEEARTFSITHLKNNLKEGINTKVA 237
          ***:.* * : * * : : * . * . * . * . * . * . * . * . * : * * . . : *

alba      EQVNHALELPLHRRRTQRLAEVWSIEAYRKEDANQVLELAILDYNMIQSVYQRDLRETS 292
tremuloides EQVSHALELPLHRRRTQRLAEVWSIEAYRKEDANQVLELAILDYNMIQSVYQRDLRETS 292
nigra      EQVNHALELPLHRRRTQRLAEVWSIEAYRKEDADQVLELAILDYNMIQSVYQRDLRETS 292
kudzu      EQVSHALELPHYHQLRLHRLAEARWFLDKYEPKEPHHQLLELAKLDFNMVQTLHQKELQDLS 297
          ***.***** *:* :***** * : : * . * * . :***** * : * : : : : : : *

alba      RWWRRVGLATKLHFARDRLIESFYWAVGVAFEPQYSDCRNSVAKMFSFVTIIDDYDVYG 352
tremuloides RWWRRVGLATKLHFARDRLIESFYWAVGVAFEPQYSDCRNSVAKMFSFVTIIDDYDVYG 352
nigra      RWWRRVGLATKLHFARDRLIESFYWAVGVAFEPQYSDCRNSVAKMFSFVTIIDDYDVYG 352
kudzu      RWWTEMGLASKLDFVRDRLMEVYFWALGMADPQPGECRKAVTKMFGVLVTIIDDYDVYG 357
          *** . :*****.* * . * . * : : * : * : * : : * : * : * : * : * : *

alba      TLDELELFTDAVERWDVNAINDLPDYMKLCLFLALYNTINEIAYDNLKDKGENILPYLTKA 412
tremuloides TLDELELFTDAVERWDVNAINDLPDYMKLCLFLALYNTINEIAYDNLKDKGENILPYLTKA 412
nigra      TLDELELFTDAVERWDVNAINDLPDYMKLCLFLALYNTINEIAYDNLKDKGENILPYLTKA 412
kudzu      TLDELQLFTDAVERWDVNAINTLPDYMKLCLFLALYNTVNDTSYSILKEKGHNLSYLTKS 417
          *****:*****: *****: : * . * : * . * . * : * : * : * :

alba      WADLCNAFLQEAKWLYNKSTPTFDDYFGNAWKSSSGPLQLVFAYFAVVQN---IKKEIE 469
tremuloides WADLCNAFLQEAKWLYNKSTPTFDDYFGNAWKSSSGPLQLVFAYFAVVQN---IKKEIE 469
nigra      WADLCNAFLQEAKWLYNKSTPTFDEYFGNAWKSSSGPLQLVFAYFAVVQN---IKKEID 469
kudzu      WRELCKAFLQEAKWSNNKI I PAFSKYLENASVSSSGVALLAPSYFSVCQQQEDISDHAR 477
          * : * : * : * : * * * * : * . : * * * * * : * : * : * : * : * :

alba      NLQKYHDTISRPCHIFRLCNDLASASAEIARGETANSVSCYMR-TKGISEELATESVMNL 528
tremuloides NLQKYHDTISRPCHIFRLCNDLASASAEIARGETANSVSCYMR-TKGISEELATESVMNL 528
nigra      NLQKYHDTISRPCHIFRLCNDLASASAEIARGETANSVSCYMR-TKGISEELATESVMNL 528
kudzu      SLTDFHGLVRSSCVIFRLCNDLATSAAELERGETTNSIIMYHENDGTSEEQAREELRKL 537
          . * . : * . : * : * : * : * : * : * : * : * : * : * : * : * : * :

```

FIG. 13

FIG. 13 - continuação

```
alba      IDETWKKMNKEKLGGS-LFAKPFVETA INLARQSHCTYHNGDAHTSPDELTRKRVLSVIT 587
tremuloi IDETWKKMNKEKLGGS-LFAKPFVETA INLARQSHCTYHNGDAHTSPDELTRKRVLSVIT 587
nigra     IDETWKKMNKEKLGGS-LFAKPFVETA INLARQSHCTYHNGDAHTSPDELTRKRVLSVIT 587
kudzu     IDAEWKMMNRERVSDSTLLPKAFMEIAVNMARVSHCTYQYGDLGRPDYATENRIKLILLI 597
**   *****:::. . * *:.*.* *:::** *****: **    ** *.:*:  ::
```

```
alba      EPILPFER--- 595
tremuloi  EPILPFER--- 595
nigra     EPILPFER--- 595
kudzu     DFFPINQLMYV 608
*:       :
```

RESUMO

Patente de Invenção: **"PRODUÇÃO DE HIDROCARBONETO VOLÁTIL DE CADEIA CURTA USANDO MICROALGAS, CIANOACTÉRIAS OU BACTÉRIAS GENETICAMENTE MANIPULADAS".**

5

A presente invenção fornece métodos e composições para a produção de hidrocarbonetos de isopreno a partir de microalgas, cianobactérias e bactérias fotossintéticas e não-fotossintéticas