



**(11) PI 0922490-4 B1**

**República Federativa do Brasil**  
Ministério do Desenvolvimento, Indústria,  
Comércio e Serviços

**(22) Data do Depósito:** 09/12/2009

**(45) Data de Concessão:** 12/12/2023

Instituto Nacional da Propriedade Industrial

---

**(54) Título:** MÉTODO PARA PRODUÇÃO DE BETA-SANTALENO

**(51) Int.Cl.:** C12P 5/00.

**(30) Prioridade Unionista:** 11/12/2008 EP 08171298.6; 17/08/2009 IB PCT/IB2009/053623.

**(73) Titular(es):** FIRMENICH SA.

**(72) Inventor(es):** MICHEL SCHALK.

**(86) Pedido PCT:** PCT IB2009055589 de 09/12/2009

**(87) Publicação PCT:** WO 2010/067309 de 17/06/2010

**(85) Data do Início da Fase Nacional:** 07/06/2011

**(57) Resumo:** MÉTODO PARA PRODUÇÃO DE BETA-SANTALENO. A presente invenção fornece um método de produção de f-santaleno, método este compreendendo o contato, no mínimo, de um polipeptídeo com pirofosfato de farnesila (FPP). Em particular, o referido método pode ser realizado in vitro ou in vivo para produzir Bsantaleno, um composto muito útil nos campos da perfumaria e aromatização. A presente invenção também fornece a sequência de aminoácido de um polipeptídeo, útil no método da invenção. Um ácido nucléico que codifica o polipeptídeo da invenção e um vetor de expressão contendo o referido ácido nucléico também são partes da presente invenção. Um organismo hospedeiro não-humano ou uma célula transformada para ser usada no método de produção de B-santaleno também é um objeto da presente invenção

## MÉTODO PARA PRODUÇÃO DE BETA-SANTALENO

### Campo Técnico

(001) A presente invenção fornece um método de produção de  $\beta$ -santaleno, método este compreendendo o contato, no mínimo, de um polipeptídeo com pirofosfato de farnesila (FPP). Em particular, o referido método pode ser realizado *in vitro* ou *in vivo* para produzir  $\beta$ -santaleno, um composto muito útil nos campos da perfumaria e aromatização. A presente invenção também fornece a sequência de aminoácido de um polipeptídeo, útil no método da invenção. Um ácido nucléico que codifica o polipeptídeo da invenção e um vetor de expressão contendo o referido ácido nucléico também são partes da presente invenção. Um organismo hospedeiro não-humano ou uma célula transformada para ser usada no método de produção de  $\beta$ -santaleno também é um objeto da presente invenção.

### Estado da Técnica

(002) Os terpenos são encontrados em muitos organismos (microorganismos, animais e plantas). Esses compostos são compostos de cinco unidades de carbono, chamadas unidades de isopreno e são classificados pela quantidade dessas unidades presentes em sua estrutura. Dessa forma, monoterpenos, sesquiterpenos e diterpenos são terpenos contendo 10, 15 e 20 átomos de carbono, respectivamente. Sesquiterpenos, por exemplo, são amplamente encontrados no reino vegetal. Muitas moléculas de sesquiterpeno são conhecidas por suas propriedades aromatizantes e de fragrância e seus efeitos cosméticos, medicinais e antimicrobianos. Mais de 300 hidrocarbonetos sesquiterpênicos e 3000 sesquiterpenóides foram identificados (Joulain, D., and König, W.A., *The Atlas of Spectral Data of Sesquiterpene Hydrocarbons*, EB Verlag, Hamburg, 1998; Connolly, J.D., Hill R.A., *Dictionary of Terpenoids*, Vol 1, Chapman and Hall (publicador),

1991), e muitas estruturas novas são identificadas todos os anos. Extratos vegetais obtidos por diferentes meios, como destilação de vapor ou extração de solvente, são usados como fonte de terpenos. Moléculas de terpeno são muitas vezes usadas como tal, mas em alguns casos reações químicas são usadas para transformar os terpenos em outras moléculas de valor alto.

(003) A produção biossintética de terpenos envolve enzimas chamadas terpeno sintases. Há praticamente uma infinidade de sesquiterpeno sintases presentes no reino vegetal, todas usando o mesmo substrato (pirofosfato de farnesila, FPP), mas tendo diferentes perfis de produto. Genes e cDNAs que codificam sesquiterpeno sintases foram clonados e as enzimas recombinantes correspondentes caracterizadas. A biossíntese de terpenos nas plantas e outros organismos foi extensivamente estudada e não é adicionalmente detalhada no presente, mas referência é feita a Dewick, *Nat. Prod. Rep.*, 2002, 19, 181-222, que revisa o estado da técnica das vias biossintéticas de terpeno.

(004) B-Santaleno é uma molécula de sesquiterpeno que ocorre naturalmente, que pode ser usada como material inicial para a síntese química ou a biossíntese de  $\beta$ -santalol (conforme representado na Figura 2B), que é um constituinte importante do óleo de sândalo. O óleo de sândalo é um importante componente de perfumaria, obtido pela destilação da cerne de espécies de Sândalo. O sândalo também é amplamente usado para incensos e medicina tradicional. O óleo contém 90% de álcoois sesquiterpênicos. Entre os diferentes isômeros de santalol, o  $\beta$ -santalol é o principal contribuinte para o odor amadeirado, doce e balsâmico típico do óleo de sândalo. Outros constituintes, como epi- $\beta$ -santalol e  $\alpha$ -santalol podem também contribuir com a nota de sândalo.

(005) Geralmente, o preço e a disponibilidade dos extratos naturais vegetais são

dependentes da abundância, do rendimento do óleo e da origem geográfica das plantas. Além disso, a disponibilidade e a qualidade dos extratos naturais são muito dependentes do clima e de outras condições locais, que levam à variabilidade de ano a ano, tornando o uso desses componentes na perfumaria de alta qualidade muito difícil ou mesmo impossível em alguns anos. Devido à exploração em excesso dos recursos naturais, dificuldades de cultivo, crescimento lento das plantas de *Santalum*, as disponibilidades da matéria-prima de sândalo diminuíram dramaticamente durante as últimas décadas. Portanto, seria uma vantagem fornecer uma fonte de  $\beta$ -santalol, que esteja menos sujeito a flutuações na disponibilidade e qualidade. Uma síntese química dos constituintes de sesquiterpeno de sândalo, até o momento, não está disponível. Uma via bioquímica que leva à síntese de  $\beta$ -santaleno, que poderia, então, ser usada para produzir  $\beta$ -santalol, portanto, seria de grande interesse.

(006) Sesquiterpeno do tipo santalano e, particularmente, sesquiterpenos com o esqueleto de  $\beta$ -santalano, foram identificados em diversas espécies de planta. Embora nenhuma sesquiterpeno sintase, capaz de produzir  $\beta$ -santaleno, tenha sido descrita até o momento.

(007) Uma sesquiterpeno sintase capaz de sintetizar, no mínimo, um sesquiterpeno bicíclico e/ou tricíclico, tendo um esqueleto de carbono santalano, o ácido nucléico correspondente e um método para produção desse composto, tendo um esqueleto de carbono santalano, são divulgados no pedido internacional de patente WO 2006/134523. Todavia, nenhum traço de  $\beta$ -santaleno foi detectado como produto das sesquiterpeno sintases divulgadas nos exemplos. O único produto com um esqueleto de santalano foi epi-beta-santaleno. As propriedades de epi-beta-santaleno são muito diferentes daqueles do  $\beta$ -santalano. Em particular, não tem

interesse na síntese de  $\beta$ -santalol. Além disso, a sesquiterpeno sintase divulgada na WO 2006/134523 compartilha apenas 27% da identificação com a sequência da invenção.

(008) A porcentagem de identificação entre sesquiterpeno sintases, conhecidas a partir de bancos de dados, e os polipeptídios da invenção é muito baixa. A sequência proteica mais próxima à  $\beta$ -santaleno sintase da invenção é uma monoterpeno sintase de *Santalum album* (acesso Nº ACF24767; Jones, C.G., Keeling, C.I., Ghisalberti, E.L., Barbour, E.L., Plummer, J.A. and Bohlmann, J. Arch. Biochem. Biophys., 2008, 477(1), 121-130), que compartilha 58% de identificação da sequência de aminoácido com a  $\beta$ -santaleno sintase da invenção. Quando em contato com FPP, essa enzima produz mais de 90% de  $\beta$ -bisaboleno e nenhum isômero santaleno é formado.

(009) Além da diferença entre as sequências por si só, também tem de ser apontado que a estrutura e as propriedades dos produtos sintetizados pela enzima acima mencionada são muito diferentes daquelas do sesquiterpeno  $\beta$ -santaleno. Em particular, os monoterpenos produzidos por essa enzima, isto é, alfa-terpineol, limoneno, geraniol, mirceno, linalol e alguns outros produtos não-significativos não são adequados como um material inicial para a síntese de  $\beta$ -santalol, que é um componente muito útil no campo da perfumaria.

(0010) Apesar de estudos extensivos de ciclização de terpeno, o isolamento e a caracterização das terpenos sintases ainda é difícil, particularmente nas plantas, devido a sua baixa abundância, seus padrões de expressão muitas vezes transitórios e a complexidade de purificá-las a partir das misturas de resinas e compostos fenólicos em tecidos onde elas são expressadas.

(0011) É um objetivo da presente invenção fornecer métodos para criação de  $\beta$ -

santaleno em uma maneira econômica, conforme indicado acima. Conformemente, a presente invenção tem o objetivo de produzir  $\beta$ -santaleno ao mesmo tempo em que tem pouco resíduo, um processo mais suficiente com relação à energia e recurso e ao mesmo tempo em que reduz a dependência de combustíveis fósseis. Um objetivo adicional é fornecer enzimas capazes de sintetizar  $\beta$ -santaleno, que é útil como componentes de perfumaria e/ou aroma.

#### Abreviações usadas

|      |                                          |
|------|------------------------------------------|
| ACC  | ácido 1-aminociclopropanecarboxílico     |
| bp   | par de base                              |
| kb   | quilobase                                |
| BSA  | albumina sérica bovina                   |
| 2,4D | ácido 2,4-diclorofenoxiacético           |
| DNA  | ácido desoxirribonucléico                |
| cDNA | DNA complementar                         |
| dNTP | trifosfato de desoxinucleotídeo          |
| DTT  | ditiotreitól                             |
| EDTA | ácido etilenodiaminotetracético          |
| FPP  | pirofosfato de farnesila                 |
| GC   | cromatografia gasosa                     |
| IPTG | isopropil-D-tiogalactopiranosídeo        |
| Kin  | cinetina                                 |
| LB   | caldo lisogênico                         |
| MS   | espectrômetro de massa                   |
| PCR  | reação em cadeia da polimerase           |
| RMCE | troca de cassete mediada por recombinase |

3’-/5’-RACE      amplificação rápida 3’ e 5’ das extremidades do cDNA

RNA              ácido ribonucléico

mRNA            ácido ribonucléico mensageiro

RNase            Ribonuclease

### Descrição da invenção

(0012)A presente invenção fornece um método para produzir biossinteticamente  $\beta$ -santaleno em uma maneira econômica, confiável e reproduzível, usando um polipeptídio tendo uma atividade de  $\beta$ -santaleno sintase. A presente invenção é particularmente útil porque nenhum desse polipeptídio era conhecido no estado da técnica e porque nenhuma biossíntese de  $\beta$ -santaleno foi descrita. Esta resolve o problema muito importante do fornecimento de  $\beta$ -santaleno, um composto que é muito útil para a indústria de perfumaria. A presente invenção também fornece uma sequência de ácido nucléico que codifica os polipeptídios usados no método da invenção, dessa forma, sendo intimamente ligado ao referido polipeptídio. O polipeptídio e o ácido nucléico são ferramentas muito importantes, que são necessárias para realizar o método da invenção. O mesmo é verdade para vetores e para organismos modificados com o ácido nucléico da invenção para expressarem heterologamente o polipeptídio da invenção.

(0013)Uma “sesquiterpeno sintase” ou um “polipeptídio tendo uma atividade de sesquiterpeno sintase” é proposto para o objetivo do presente pedido como um polipeptídio capaz de catalisar a síntese de uma molécula de sesquiterpeno ou de uma mistura de moléculas de sesquiterpeno do terpeno acíclico precursor FPP.

(0014)Como uma “ $\beta$ -santaleno sintase” ou como um “polipeptídio tendo uma atividade de  $\beta$ -santaleno sintase”, referimo-nos, no presente, a um polipeptídio capaz de catalisar a síntese de  $\beta$ -santaleno na forma de qualquer de seus

estereoisômeros ou uma mistura destes, começando do FPP. B-Santaleno pode ser o único produto ou pode ser parte de uma mistura de sesquiterpenos. B-Santaleno é definido pela maneira de sua estrutura, conforme representado na Figura 2A.

(0015) A capacidade de um polipeptídeo de catalisar a síntese de um sesquiterpeno particular (por exemplo,  $\beta$ -santaleno) pode ser simplesmente confirmada pela realização do ensaio enzimático, conforme detalhado no Exemplo 3.

(0016) De acordo com a presente invenção, os polipéptidos também devem incluir polipéptidos truncados, desde que mantenham a sua actividade de sesquiterpeno sintase como definido em qualquer uma das formas de realização anteriores e que partilham pelo menos a percentagem de identidade definida com o fragmento correspondente de SEQ ID NO: 15.

(0017) Como se pretende aqui abaixo, "uma sequência de nucleótidos obtida por modificação de SEQ ID NO: 14 ou 2" engloba qualquer sequência que tenha sido obtida alterando a sequência de SEQ ID NO: 14 ou de SEQ ID NO: 2 usando qualquer método conhecido na arte, por exemplo, introduzindo qualquer tipo de mutações, tais como mutações de exclusão, inserção ou substituição. Exemplos de tais métodos são citados na parte da descrição em relação aos polipéptidos variantes e os métodos para prepará-los.

(0018) De acordo com a presente invenção, polipeptídios também devem incluir polipeptídios truncados, considerando-se que eles mantenham sua atividade de sesquiterpeno sintase, conforme definido em qualquer uma das configurações acima e que eles compartilhem a porcentagem definida de identificação com o fragmento correspondente de SEQ ID NO:1.

(0019) Conforme proposto abaixo no presente, "uma sequência de nucleotídeo, obtida por modificação da SEQ ID NO:2, 4 ou o complemento desta" abrange

qualquer sequência que tenha sido obtida por alteração da sequência de SEQ ID NO:2, de SEQ ID NO:4 ou do complemento desta usando qualquer método conhecido na técnica, por exemplo, por introdução de qualquer tipo de mutações, como deleção, inserção ou mutações de substituição. Exemplos desses métodos são citados na parte da descrição, relativa aos polipeptídios variantes e aos métodos de preparação destes.

a. A porcentagem de identificação entre duas sequências peptídicas ou nucleotídeicas é uma função da quantidade de aminoácidos ou resíduos de nucleotídeo que são identificados nas duas sequências quando um alinhamento dessas duas sequências foi gerado. Resíduos idênticos são definidos como resíduos que são os mesmos nas duas sequências em uma dada posição do alinhamento. A porcentagem de identificação da sequência, conforme usada no presente, é calculada a partir do alinhamento favorável, pegando-se a quantidade de resíduos idênticos entre duas sequências, dividindo-a pela quantidade total de resíduos na sequência mais curta e multiplicando-se por 100. O alinhamento favorável é o alinhamento no qual a porcentagem de identificação é a mais alta possível. Lacunas podem ser introduzidas em um ou ambas as sequências em uma ou mais posições do alinhamento para obter o alinhamento favorável. Essas lacunas são, então, levadas em consideração como resíduos não-idênticos para o cálculo da porcentagem de identificação da sequência.

(0020)O alinhamento para o objetivo de determinação da porcentagem da sequência de aminoácido ou ácido nucléico pode ser atingido em várias maneiras, usando programas de computador e, por exemplo, programas de computador publicamente disponíveis na internet. Preferivelmente, o programa BLAST (Tatiana et al, *FEMS Microbiol Lett.*, 1999, 174:247-250, 1999) ajustado aos parâmetros

padrão, disponível a partir do Centro Nacional de Informações de Biotecnologia (NCBI) em <http://www.ncbi.nlm.nih.gov/BLAST/bl2seq/wblast2.cgi>, pode ser usado para obter um alinhamento favorável de sequências peptídicas ou nucleotídicas e calcular a porcentagem de identificação da sequência.

(0021) Um objeto da presente invenção é, portanto, um método para produção de  $\beta$ -santaleno compreendendo

a) o contato do FPP com, no mínimo, um polipeptídio tendo uma atividade de  $\beta$ -santaleno sintase e compreendendo uma sequência de aminoácido, no mínimo, 50% idêntica à SEQ ID NO:1 ou 3;

b) opcionalmente, o isolamento do  $\beta$ -santaleno produzido na etapa a).

(0022) O objeto da presente invenção é, portanto, um método para a produção de  $\beta$ -santaleno compreendendo

a) entrar em contato com FPP com pelo menos um polipeptídio possuindo uma atividade de  $\beta$ -santaleno sintase e compreendendo uma sequência de aminoácidos pelo menos 60% idêntica à SEQ ID NO: 15;

b) opcionalmente, isolando o  $\beta$ -santaleno produzido no passo a).

(0023) De acordo com uma configuração preferida, o  $\beta$ -santaleno representa, no mínimo, 20%, preferivelmente, no mínimo, 30%, preferivelmente, no mínimo, 35% dos produtos, produzidos pelo método da invenção.

(0024) O método pode ser realizado *in vitro*, bem como *in vivo*, considerando-se que os métodos, que envolvem apenas o metabolismo natural da planta, sem nenhuma transformação, não sejam abrangidos pelos métodos da presente invenção, conforme será explicado em detalhes adiante.

(0025) O polipeptídio a entrar em contato com FPP *in vitro* pode ser obtido por extração de qualquer organismo que o expresse, usando tecnologias padrão de

extração de proteína ou enzima. Se o organismo hospedeiro for um organismo unicelular ou célula que libera o polipeptídeo da invenção no meio de cultura, o polipeptídeo pode simplesmente ser coletado do meio de cultura, por exemplo, por centrifugação, opcionalmente seguida por etapas de lavagem e ressuspensão em soluções tampão adequadas. Se o organismo ou célula acumular o polipeptídeo dentro de suas células, o polipeptídeo pode ser obtido por ruptura ou lise das células e extração adicional do polipeptídeo do lisado celular.

(0026)O polipeptídeo tendo uma atividade de  $\beta$ -santaleno sintase, em uma forma isolado ou juntamente com outras proteínas, por exemplo, em um extrato de proteína crua, obtido das células ou microorganismos cultivados, pode, então, ser suspenso em uma solução tampão em pH favorável. Se adequado, sais, DTT, BSA e outros tipos de cofatores enzimáticos podem ser adicionados a fim de otimizar a atividade enzimática. A concentração desses cofatores pode ser ajustada a fim de atingir um rendimento otimizado. Por exemplo, a diminuição da concentração de íons  $Mg^{2+}$  na suspensão de polipeptídeo é particularmente vantajosa para aumentar o rendimento de  $\beta$ -santaleno. A concentração favorável de íons  $Mg^{2+}$  está compreendida entre 2 e 0,75mM. Condições adequadas são descritas em mais detalhes nos Exemplos adiante.

(0027)O precursor FPP é adicionado à suspensão de polipeptídeo, que é, então, incubada em temperatura favorável, por exemplo, entre 15 e 40°C, preferivelmente entre 25 e 35°C, mais preferivelmente a 30°C. Após incubação, o  $\beta$ -santaleno produzido pode ser isolado da solução incubada por procedimentos de isolamento padrão, como extração e destilação de solvente, opcionalmente após remoção de polipeptídeos da solução.

(0028)De acordo com outra configuração preferida, o método de qualquer uma das

configurações acima descritas é realizado *in vivo*. Nesse caso, a etapa a) compreende o cultivo de um organismo ou célula hospedeiros, não-humanos, capaz de produzir FPP e transformado para expressar, no mínimo, um polipeptídio compreendendo uma sequência de aminoácido, no mínimo, 50% idêntica à SEQ ID NO:1 ou 3 e tendo uma atividade de  $\beta$ -santaleno sintase, sob condições favoráveis à produção de  $\beta$ -santaleno.

(0029) De acordo com uma configuração mais preferida, o método ainda compreende, antes da etapa a), a transformação de um organismo ou célula não-humanos, capaz de produzir FPP com, no mínimo, um ácido nucléico, que codifica um polipeptídio compreendendo uma sequência de aminoácido, no mínimo, 50% idêntica à SEQ ID NO:1 ou 3 e tendo uma atividade de  $\beta$ -santaleno sintase, de maneira que o referido organismo expresse o referido polipeptídio.

(0030) De acordo com outra forma de realização preferida, o método de qualquer uma das formas de realização descritas acima é realizado *in vivo*. Neste caso, o passo a) compreende cultivar um organismo hospedeiro não-humano ou célula capaz de produzir FPP e transformado para expressar pelo menos um polipeptídio compreendendo uma sequência de aminoácidos pelo menos 60% idêntica à SEQ ID NO: 15 e tendo uma atividade de  $\beta$ -santaleno sintase, em condições propícias para a produção de  $\beta$ -santaleno.

(0031) De acordo com uma forma de realização mais preferida, o método compreende ainda, antes do passo a), transformar um organismo ou célula não humana capaz de produzir FPP com pelo menos um ácido nucleico que codifica um polipeptídio compreendendo uma sequência de aminoácidos pelo menos 60% idêntica à SEQ ID NO: 15 e tendo uma atividade de  $\beta$ -santaleno sintase, de modo que o referido organismo expresse o referido polipeptídio.

(0032)Essas configurações da invenção são particularmente vantajosas já que é possível realizar o método *in vivo* sem previamente isolar o polipeptídio. A reação ocorre diretamente dentro do organismo ou célula transformada para expressar o referido polipeptídio.

(0033)De acordo com uma configuração particular da invenção, no mínimo, um ácido nucléico, que codifica a  $\beta$ -santaleno sintase compreende uma sequência de nucleotídeo, no mínimo, 50%, preferivelmente, no mínimo, 55%, preferivelmente, no mínimo, 60%, preferivelmente, no mínimo, 65%, preferivelmente, no mínimo, 70%, preferivelmente, no mínimo, 75%, preferivelmente, no mínimo, 80%, preferivelmente, no mínimo, 85%, preferivelmente, no mínimo, 90%, mais preferivelmente, no mínimo, 95% e ainda mais preferivelmente, no mínimo, 98% idêntica à SEQ ID NO:2, 4 ou o complemento desta. De acordo com uma configuração mais preferida, o referido ácido nucléico compreende a sequência de nucleotídeo SEQ ID NO:2, 4 ou o complemento desta.

(0034)Em outra configuração particular, o ácido nucléico consiste de uma sequência de nucleotídeo, no mínimo, 50%, preferivelmente, no mínimo, 55%, preferivelmente, no mínimo, 60%, preferivelmente, no mínimo, 65%, preferivelmente, no mínimo, 70%, preferivelmente, no mínimo, 75%, preferivelmente, no mínimo, 80%, preferivelmente, no mínimo, 85%, preferivelmente, no mínimo, 90%, mais preferivelmente, no mínimo, 95% e ainda mais preferivelmente, no mínimo, 98% idêntica à SEQ ID NO:2, 4 ou o complemento desta. Em uma configuração ainda mais preferida, o referido ácido nucléico consiste de SEQ ID NO:2, 4 ou o complemento desta.

(0035)De acordo com uma configuração mais preferida, no mínimo, o ácido nucléico

usado em qualquer uma das configurações acima compreende a sequência de nucleotídeo que foi obtida pela modificação de SEQ ID NO:2, 3 ou o complemento desta. De acordo com uma configuração ainda mais preferida, no mínimo, o referido ácido nucléico consiste de uma sequência de nucleotídeo que foi obtida pela modificação de SEQ ID NO:2, 3 ou o complemento desta, preferivelmente SED ID NO:3 ou o complemento desta.

(0036) De acordo com uma concretização particular da invenção, o pelo menos um ácido nucleico que codifica a  $\beta$ -santaleno sintase compreende uma sequência de nucleotídeos pelo menos 60%, de preferência pelo menos 65%, de preferência pelo menos 70%, de preferência pelo menos 75%, de preferência a pelo menos 80%, de preferência pelo menos 85%, de preferência pelo menos 90%, mais preferencialmente pelo menos 95% e ainda mais preferencialmente pelo menos 98% idênticos à SEQ ID NO: 14, 2 ou o seu complemento, e de preferência a SEQ ID NO : 2 ou o seu complemento. De acordo com uma forma de realização mais preferida, o referido ácido nucleico compreende a sequência nucleotídica SEQ ID NO: 14, 2 ou o seu complemento, de preferência SEQ ID NO: 2 ou o seu complemento. Numa concretização ainda mais preferida, o referido ácido nucleico consiste em SEQ ID NO: 14, 2 ou o seu complemento, de preferência de SEQ ID NO: 2 ou o seu complemento.

(0037) De acordo com uma forma de realização mais preferida, o pelo menos um ácido nucleico utilizado em qualquer uma das formas de realização anteriores compreende uma sequência de nucleotídeos que foi obtida modificando SEQ ID NO: 14, 2 ou o seu complemento. De acordo com uma concretização ainda mais preferida, o referido pelo menos um ácido nucleico consiste numa sequência de nucleotídeos que foi obtida modificando SEQ ID NO: 14, 2 ou o seu complemento,

de preferência SEQ ID NO: 2 ou o seu complemento.

(0038) De acordo com outra configuração, no mínimo, um ácido nucléico é isolado de uma planta das espécies *Santalum*, preferivelmente de *Santalum album*.

(0039) O organismo ou célula deve “expressar” um polipeptídio, considerando-se que o organismo ou célula seja transformado para abrigar um ácido nucléico, que codifica o referido polipeptídio, esse ácido nucléico seja transcrito para mRNA e o polipeptídio seja encontrado no organismo ou célula hospedeiros. O termo “expressa” abrange “expressa heterologamente” ou “superexpressa”, o último referindo-se aos níveis de mRNA, polipeptídio e/ou atividade enzimática sobre e acima do que é medido em um organismo ou célula não-transformado. Uma descrição mais detalhada de métodos adequados para transformar um organismo ou célula hospedeiros, não-humanos será descrita posteriormente na parte da especificação que é dedicada a esses organismos ou células hospedeiros, não-humanos como objetos específicos da presente invenção e nos exemplos.

a. Um organismo ou célula particular deve ser “capaz de produzir FPP” quando produz FPP naturalmente ou quando não produz FPP naturalmente, mas é transformado para produzir FPP, ou antes da transformação com um ácido nucléico, conforme descrito no presente ou juntamente com o referido ácido nucléico. Organismos ou células transformadas para produzir uma quantidade maior de FPP do que o organismo ou célula que ocorre naturalmente também são abrangidos por “organismos ou células capazes de produzir FPP”. Métodos para transformar organismos, por exemplo, microorganismos, de maneira que eles produzam FPP já são conhecidos na técnica. Esses métodos podem, por exemplo, ser encontrados na literatura, por exemplo, nas seguintes publicações: Martin, V.J., Pitera, D.J., Withers, S.T., Newman, J.D., and Keasling, J.D. *Nat Biotechnol.*, 2003, 21(7), 796-

802 (transformação de *E. coli*); Wu, S., Schalk, M., Clark, A., Miles, R.B., Coates, R., and Chappell, J., *Nat Biotechnol.*, 2006, 24(11), 1441-1447 (transformação de plantas); Takahashi, S., Yeo, Y., Greenhagen, B. T., McMullin, T., Song, L., Maurina-Brunker, J., Rosson, R., Noel, J., Chappell, J., *Biotechnology and Bioengineering*, 2007, 97(1), 170-181 (transformação de levedura).

(0040) Para realizar a invenção *in vivo*, o organismo ou célula hospedeiro é cultivado sob condições favoráveis à produção de  $\beta$ -santaleno. Conformemente, se o hospedeiro for uma planta transgênica, condições de crescimento favoráveis são fornecidas, como condições de luz, água e nutriente favoráveis, por exemplo. Se o hospedeiro for um organismo unicelular, condições favoráveis à produção de  $\beta$ -santaleno podem compreender adição de cofatores adequados ao meio de cultura do hospedeiro. Além disso, um meio de cultura pode ser selecionado, de maneira a maximizar a síntese de  $\beta$ -santaleno. Condições de cultura favoráveis são descritas em uma maneira mais detalhada nos Exemplos a seguir.

(0041) Organismos hospedeiros não-humanos adequados para realizar o método da invenção *in vivo* podem ser quaisquer organismos multicelulares ou unicelulares não-humanos. Em uma configuração preferida, o organismo hospedeiro não-humano usado para realizar a invenção *in vivo* é uma planta, um procarionte ou um fungo. Qualquer planta, procarionte ou fungo pode ser usado. Plantas particularmente úteis são aquelas que naturalmente produzem altas quantidades de terpenos. Em uma configuração mais detalhada, a planta é selecionada da família de *Solanaceae*, *Poaceae*, *Brassicaceae*, *Fabaceae*, *Malvaceae*, *Asteraceae* ou *Lamiaceae*. Por exemplo, a planta é selecionada dos gêneros *Nicotiana*, *Solanum*, *Sorghum*, *Arabidopsis*, *Brassica* (colza), *Medicago* (alfalfa), *Gossypium* (algodão), *Artemisia*, *Salvia* e *Mentha*. Preferivelmente, a planta pertence às espécies de

*Nicotiana tabacum*.

(0042)Em uma configuração mais preferida, o organismo hospedeiro não-humano usado para realizar a invenção *in vivo* é um microorganismo. Qualquer microorganismo pode ser usado, mas de acordo com uma configuração ainda mais preferida, o referido microorganismo é uma bactéria ou levedura. Mais preferivelmente, a referida bactéria é *E. coli* e a referida levedura é *Saccharomyces cerevisiae*.

(0043)Alguns desses organismos não produzem FPP naturalmente. Para serem adequados para realizar o método da invenção, esses organismos têm de ser transformados para produzir o referido precursor. Eles podem ser então transformados antes da modificação com o ácido nucléico descrito de acordo com qualquer uma das configurações acima ou simultaneamente, conforme explicado acima.

(0044)Células eucarióticas maiores isoladas podem também ser usadas, ao invés de organismos completos, como hospedeiros para realizar o método da invenção *in vivo*. Células eucarióticas adequadas podem ser qualquer célula não-humana, mas são preferivelmente células vegetais ou fúngicas.

(0045)De acordo com uma configuração preferida, no mínimo, um polipeptídeo tendo uma atividade de  $\beta$ -santaleno sintase usado em qualquer uma das configurações acima descritas ou codificado pelo ácido nucléico usado em qualquer uma das configurações acima descritas compreende um amino, no mínimo, 55%, preferivelmente, no mínimo, 60%, preferivelmente, no mínimo, 65%, preferivelmente, no mínimo, 70%, preferivelmente, no mínimo, 75%, preferivelmente, no mínimo, 80%, preferivelmente, no mínimo, 85%, preferivelmente, no mínimo, 90%, mais preferivelmente, no mínimo, 95% e ainda mais preferivelmente, no mínimo, 98%

idêntico à SEQ ID NO:1 ou 3. De acordo com uma configuração mais preferida, o referido polipeptídio compreende a sequência de aminoácido SEQ ID NO:1 ou 3.

(0046) Em outra configuração preferida, o polipeptídio consiste de uma sequência de aminoácido, no mínimo, 50%, preferivelmente, no mínimo, 55%, preferivelmente, no mínimo, 60%, preferivelmente, no mínimo, 65%, preferivelmente, no mínimo, 70%, preferivelmente, no mínimo, 75%, preferivelmente, no mínimo, 80%, preferivelmente, no mínimo, 85%, preferivelmente, no mínimo, 90%, mais preferivelmente, no mínimo, 95% e ainda mais preferivelmente, no mínimo, 98% idêntico à SEQ ID NO:1 ou 3. Em uma configuração ainda mais preferida, o referido polipeptídio consiste da sequência de aminoácido SEQ ID NO:1 ou 3.

(0047) De acordo com outra configuração preferida, no mínimo, um polipeptídio tendo uma atividade de  $\beta$ -santaleno sintase usado em qualquer uma das configurações acima descritas ou codificado pelo ácido nucléico usado em qualquer uma das configurações acima descritas compreende uma sequência de aminoácido que é uma variante da SEQ ID NO:1 ou 3, obtida por engenharia genética, considerando-se que a referida variante mantenha sua atividade de  $\beta$ -santaleno sintase, conforme definido acima e tenha a porcentagem exigida de identificação à SEQ ID NO:1 ou 3. Em outros termos, o referido polipeptídio preferivelmente compreende uma sequência de aminoácido codificada por uma sequência de nucleotídeo que foi obtida por modificação da SEQ ID NO:2, 4 ou o complemento desta. De acordo com uma configuração mais preferida, no mínimo, um polipeptídio tendo uma atividade de  $\beta$ -santaleno sintase usado em qualquer uma das configurações acima descritas ou codificado pelo ácido nucléico usado em qualquer uma das configurações acima descritas consiste de uma sequência de aminoácido que é uma variante da SEQ ID NO:1 ou 3, obtida por engenharia genética, isto é,

uma sequência de aminoácido codificada por uma sequência de nucleotídeo que foi obtida por modificação da SEQ ID NO:2, 4 ou o complemento desta.

(0048) De acordo com outra configuração preferida, no mínimo, um polipeptídeo tendo uma atividade de  $\beta$ -santaleno sintase usado em qualquer uma das configurações acima descritas ou codificado pelo ácido nucléico usado em qualquer uma das configurações acima descritas é uma variante da SEQ ID NO:1 ou 3, que pode ser encontrada naturalmente em outros organismos, como outras espécies de planta, considerando-se mantenha sua atividade de  $\beta$ -santaleno sintase, conforme definido acima e tenha a porcentagem exigida de identificação à SEQ ID NO:1 ou 3.

(0049) Conforme usado no presente, o polipeptídeo é proposto como um polipeptídeo ou fragmento de peptídeo que abrange as sequências de aminoácido identificadas no presente, bem como polipeptídios truncados ou variantes, considerando-se que eles mantenham sua atividade de  $\beta$ -santaleno sintase conforme definido acima e que compartilhem a porcentagem definida de identificação com o fragmento correspondente de SEQ ID NO:1 ou 3.

(0050) De acordo com uma forma de realização preferida, o pelo menos um polipéptido possuindo uma actividade de  $\beta$ santaleno sintase utilizado em qualquer uma das formas de realização descritas acima ou codificado pelo ácido nucleico utilizado em qualquer uma das formas de realização descritas acima compreende uma sequência de aminoácidos pelo menos 65% de preferência pelo menos 70%, de preferência pelo menos 75%, de preferência pelo menos 80%, de preferência pelo menos 85%, de preferência pelo menos 90%, mais preferencialmente pelo menos 95% e ainda mais preferencialmente pelo menos 98% idênticos à SEQ ID NO: 15. De acordo com uma forma de realização mais preferida, o referido polipéptido compreende a sequência de aminoácidos SEQ ID NO: 15. Numa forma

de realização ainda mais preferida, o referido polipéptido consiste em SEQ ID NO: 15.

(0051) De acordo com outra forma de realização preferida, o pelo menos um polipéptido com uma actividade de  $\beta$ -santaleno sintase utilizado em qualquer uma das formas de realização descritas acima ou codificado pelo ácido nucleico utilizado em qualquer uma das formas de realização descritas acima compreende uma sequência de aminoácidos que é uma variante da SEQ ID NO: 15 obtida por engenharia genética, desde que a referida variante mantenha a sua actividade de  $\beta$ -santaleno sintase, como definido acima e tenha a percentagem de identidade requerida para SEQ ID NO: 15. Em outros termos, o referido polipéptido compreende uma sequência de aminoácidos codificada por uma sequência de nucleótidos que foi obtida modificando SEQ ID NO: 14, 2 ou o seu complemento, de preferência SEQ ID NO: 2 ou o seu complemento. De acordo com uma forma de realização mais preferida, o pelo menos um polipéptido possuindo uma actividade de  $\beta$ -santaleno sintase utilizado em qualquer uma das formas de realização descritas acima ou codificado pelo ácido nucleico utilizado em qualquer das formas de realização descritas acima consiste numa sequência de aminoácidos que é uma variante da SEQ ID NO: 15 obtida por engenharia genética, ou seja, uma sequência de aminoácidos codificada por uma sequência de nucleótidos que foi obtida modificando SEQ ID NO: 14, 2 ou o seu complemento, de preferência SEQ ID NO: 2 ou o complemento.

(0052) De acordo com outra forma de realização preferida, o pelo menos um polipéptido com uma actividade de  $\beta$ -santaleno sintase utilizado em qualquer uma das formas de realização descritas acima ou codificado pelo ácido nucleico utilizado em qualquer uma das formas de realização descritas acima é uma variante de SEQ

ID NO: 15 que podem ser encontrados naturalmente em outros organismos, tais como outras espécies de plantas, desde que mantenha a sua actividade de  $\beta$ -santaleno sintase como definido acima e tenha a percentagem de identidade requerida para SEQ ID NO: 15.

(0053) Tal como aqui utilizado, o polipéptido é pretendido como um fragmento de polipéptido ou péptido que engloba as sequências de aminoácidos aqui identificadas, bem como polipéptidos truncados ou variantes, desde que mantenham a sua actividade como definido acima e que partilham pelo menos a percentagem definida de identidade com o fragmento correspondente da SEQ ID NO: 15.

(0054) Exemplos de polipeptídios variantes são proteínas que ocorrem naturalmente, que resultam de eventos de splicing de mRNA alternativos ou de divisão proteolítica dos polipeptídios descritos no presente. Variações atribuíveis à proteólise incluem, por exemplo, diferenças nos terminais N ou C mediante expressão em diferentes tipos de células hospedeiras, devido à remoção proteolítica de um ou mais aminoácidos terminais dos polipeptídios da invenção. Os polipeptídios codificados por um ácido nucleico, obtidos por mutação natural ou artificial de um ácido nucleico da invenção, conforme descrito posteriormente, também são abrangidos pela invenção.

(0055) Variantes de polipeptídio resultando de uma fusão de sequências de peptídeo adicionais nas extremidades amino e carboxila-terminal também podem ser usadas nos métodos da invenção. Em particular, essa fusão pode aumentar a expressão dos polipeptídios, ser útil na purificação da proteína ou melhorar a atividade enzimática do polipeptídio em um ambiente desejado ou sistema de expressão. Essas sequências de peptídeo adicionais podem ser peptídeos de sinal, por exemplo. Conformemente, a presente invenção abrange métodos usando

polipeptídios variantes, como aqueles obtidos por fusão com outros oligo ou polipeptídios e/ou aqueles que são ligados a peptídeos de sinal. Polipeptídios resultantes de uma fusão com outra proteína funcional, como outra proteína da via de biossíntese de terpeno, também podem ser vantajosamente usados nos métodos da invenção.

(0056)De acordo com outra configuração, no mínimo, um polipeptídio tendo uma atividade de  $\beta$ -santaleno sintase usado em qualquer uma das configurações acima descritas ou codificado pelo ácido nucléico usado em qualquer uma das configurações acima descrita é isolado de uma planta das espécies de *Santalum*, preferivelmente de *Santalum album*.

(0057)Uma ferramenta importante para realizar o método da invenção é o próprio polipeptídio. Um polipeptídio tendo uma atividade de  $\beta$ -santaleno sintase e compreendendo uma sequência de aminoácido, no mínimo, 50% idêntica à SEQ ID NO:1 ou 3 é, portanto, outro objeto da presente invenção.

(0058)Uma ferramenta importante para realizar o método da invenção é o próprio polipeptídio. Um polipeptídio com uma atividade de  $\beta$ -santaleno sintase e que compreende uma sequência de aminoácidos com pelo menos 60% de identidade com SEQ ID NO: 15 é, portanto, outro objeto da presente invenção.

(0059)De acordo com uma configuração preferida, o polipeptídio é capaz de produzir uma mistura de sesquiterpenos onde  $\beta$ -santaleno representa, no mínimo, 20%, preferivelmente, no mínimo, 30%, preferivelmente, no mínimo, 35%, dos sesquiterpenos produzidos.

(0060)De acordo com uma configuração preferida, o polipeptídio compreende uma sequência de aminoácido, no mínimo, 55%, preferivelmente, no mínimo, 60%, preferivelmente, no mínimo, 65%, preferivelmente, no mínimo, 70%, preferivelmente,

no mínimo, 75%, preferivelmente, no mínimo, 80%, preferivelmente, no mínimo, 85%, preferivelmente, no mínimo, 90%, mais preferivelmente, no mínimo, 95% e ainda mais preferivelmente, no mínimo, 98% idêntico à SEQ ID NO:1 ou 3. De acordo com uma configuração mais preferida, o polipeptídio compreende a sequência de aminoácido SEQ ID NO:1 ou 3.

(0061)De acordo com outra configuração preferida, o polipeptídio consiste de uma sequência de aminoácido, no mínimo, 50%, preferivelmente, no mínimo, 55%, preferivelmente, no mínimo, 60%, preferivelmente, no mínimo, 65%, preferivelmente, no mínimo, 70%, preferivelmente, no mínimo, 75%, preferivelmente, no mínimo, 80%, preferivelmente, no mínimo, 85%, preferivelmente, no mínimo, 90%, mais preferivelmente, no mínimo, 95% e ainda mais preferivelmente, no mínimo, 98% idêntico à SEQ ID NO:1 ou 3. De acordo com uma configuração ainda mais preferida, o polipeptídio consiste de SEQ ID NO:1 ou 3.

(0062)No mínimo, um polipeptídio compreende uma sequência de aminoácido que é uma variante de SEQ ID NO:1 ou 3, obtida por engenharia genética ou encontrada naturalmente em plantas *Santalum* ou em outras espécies de planta.

(0063)Em outros termos, quando o polipeptídio variante é obtido por engenharia genética, o referido polipeptídio compreende uma sequência de aminoácido codificada por uma sequência de nucleotídeo que foi obtida por modificação da SEQ ID NO:2, 4 ou o complemento desta. De acordo com uma configuração mais preferida, no mínimo, um polipeptídio tendo uma atividade de  $\beta$ -santaleno sintase consiste de uma sequência de aminoácido que é uma variante da SEQ ID NO:1 ou 3, obtida por engenharia genética, isto é, uma sequência de aminoácido codificada por uma sequência de nucleotídeo que foi obtida por modificação da SEQ ID NO:2, 4 ou o complemento desta.

(0064) De acordo com uma forma de realização preferida, o polipeptídio compreende uma sequência de aminoácidos pelo menos 65%, de preferência pelo menos 70%, de preferência pelo menos 75%, de preferência pelo menos 80%, de preferência pelo menos 85%, de preferência pelo menos 90%, mais preferencialmente pelo menos 95% e ainda mais preferencialmente pelo menos 98% idênticos à SEQ ID NO: 15. De acordo com uma forma de realização mais preferida, o polipeptídio compreende a sequência de aminoácidos SEQ ID NO: 15. De acordo com uma concretização ainda mais preferida, o polipeptídio consiste na SEQ ID NO: 15.

(0065) O pelo menos um polipeptídio compreende uma sequência de aminoácidos que é uma variante de SEQ ID NO: 15, quer obtida por engenharia genética ou encontrada naturalmente em plantas de *Santalum* ou em outras espécies de plantas. Em outros termos, quando o polipeptídio variante é obtido por engenharia genética, o referido polipeptídio compreende uma sequência de aminoácidos codificada por uma sequência de nucleotídeos que foi obtida modificando SEQ ID NO: 14, 2 ou o seu complemento, de acordo com uma forma de realização mais preferida, pelo menos um polipeptídio com uma atividade de  $\beta$ -santaleno sintase consiste numa sequência de aminoácidos que é uma variante da SEQ ID NO: 15 obtida por engenharia genética, ou seja, uma sequência de aminoácidos codificada por uma sequência de nucleotídeos que foi obtida modificando SEQ ID NO: 14, 2 ou o seu complemento.

(0066) De acordo com outra configuração, o polipeptídio é isolado de uma planta das espécies *Santalum*, preferivelmente de *Santalum album*.

(0067) Conforme usado no presente, o polipeptídio é proposto como um polipeptídio ou fragmento de peptídeo que abrange as sequências de aminoácido

identificadas no presente, bem como polipeptídios truncados ou variantes, considerando-se que eles mantenham sua atividade conforme definido acima e que compartilhem, no mínimo, a porcentagem definida de identificação com o fragmento correspondente de SEQ ID NO:1 ou 3.

(0068) Tal como aqui utilizado, o polipeptídio pretende-se como um fragmento de polipeptídio ou péptico que abrange a sequência de aminoácidos aqui identificada, bem como polipeptídios truncados ou variantes, desde que mantenham a sua atividade como definido acima e que partilhem, pelo menos, a porcentagem definida de identidade com o fragmento correspondente da SEQ ID NO: 15.

(0069) Exemplos de polipeptídios variantes são proteínas que ocorrem naturalmente, que resultam de eventos de splicing de mRNA alternativos ou de divisão proteolítica dos polipeptídios descritos no presente. Variações atribuíveis à proteólise incluem, por exemplo, diferenças nos terminais N ou C mediante expressão em diferentes tipos de células hospedeiras, devido à remoção proteolítica de um ou mais aminoácidos terminais dos polipeptídios da invenção. Os polipeptídios codificados por um ácido nucléico, obtidos por mutação natural ou artificial de um ácido nucléico da invenção, conforme descrito posteriormente, também são abrangidos pela invenção.

(0070) Variantes de polipeptídio resultando de uma fusão de sequências de peptídeo adicionais nas extremidades amino e carboxila-terminal também são abrangidas pelos polipeptídios da invenção. Em particular, essa fusão pode aumentar a expressão dos polipeptídios, ser útil na purificação da proteína ou melhorar a atividade enzimática do polipeptídio em um ambiente desejado ou sistema de expressão. Essas sequências de peptídeo adicionais podem ser peptídeos de sinal, por exemplo. Conformemente, a presente invenção abrange variantes dos

polipeptídios da invenção, como aqueles obtidos por fusão com outros oligo ou polipeptídios e/ou aqueles que são ligados a peptídeos de sinal. Polipeptídios resultantes de uma fusão com outra proteína funcional, como outra proteína da via de biossíntese de terpeno, também são abrangidos pelos polipeptídios da invenção.

(0071)Conforme mencionado acima, o ácido nucléico que codifica o polipeptídio da invenção é uma ferramenta útil para modificar organismos ou células hospedeiros, não-humanos, propostos para serem usados quando o método for realizado *in vivo*.

(0072)Um ácido nucléico que codifica um polipeptídio, de acordo com qualquer uma das configurações acima descritas, portanto, também é um objeto da presente invenção.

(0073)De acordo com uma configuração preferida, o ácido nucléico compreende uma sequência de nucleotídeo, no mínimo, 50%, preferivelmente, no mínimo, 55%, preferivelmente, no mínimo, 60%, preferivelmente, no mínimo, 65%, preferivelmente, no mínimo, 70%, preferivelmente, no mínimo, 75%, preferivelmente, no mínimo, 80%, preferivelmente, no mínimo, 85%, preferivelmente, no mínimo, 90%, mais preferivelmente, no mínimo, 95% e ainda mais preferivelmente, no mínimo, 98% idêntica à SEQ ID NO:2, 4 ou o complemento desta. De acordo com uma configuração mais preferida, o ácido nucléico compreende a sequência de nucleotídeo SEQ ID NO:2, 4 ou o complemento desta.

(0074)De acordo com outra configuração preferida, o ácido nucléico consiste de uma sequência de nucleotídeo, no mínimo, 50%, preferivelmente, no mínimo, 55%, preferivelmente, no mínimo, 60%, preferivelmente, no mínimo, 65%, preferivelmente, no mínimo, 70%, preferivelmente, no mínimo, 75%, preferivelmente, no mínimo, 80%, preferivelmente, no mínimo, 85%, preferivelmente, no mínimo, 90%, mais preferivelmente, no mínimo, 95% e ainda mais preferivelmente, no mínimo, 98%

idêntica à SEQ ID NO:2, 4 ou o complemento desta. De acordo com uma configuração ainda mais preferida, o ácido nucléico consiste de SEQ ID NO:2, 4 ou o complemento desta.

(0075)De acordo com uma forma de realização preferida, o ácido nucleico compreende uma sequência de nucleotídeos pelo menos 60%, de preferência pelo menos 65%, de preferência pelo menos 70%, de preferência pelo menos 75%, de preferência pelo menos 80%, de preferência pelo menos 85%, de preferência a pelo menos 90%, mais preferencialmente pelo menos 95% e ainda mais preferencialmente pelo menos 98% idênticos à SEQ ID NO: 14, 2 ou o seu complemento, de preferência a SEQ ID NO: 2 ou o seu complemento. De acordo com uma forma de realização mais preferida, o ácido nucleico compreende a sequência nucleotídica SEQ ID NO: 14, 2 ou o seu complemento, de preferência SEQ ID NO: 2 ou o seu complemento. De acordo com uma concretização ainda mais preferida, o ácido nucleico consiste em SEQ ID NO: 14, 2 ou o seu complemento, de preferência de SEQ ID NO: 2 ou o seu complemento.

(0076)De acordo com uma forma de realização preferida, o ácido nucleico compreende uma sequência de nucleotídeos pelo menos 60%, de preferência pelo menos 65%, de preferência pelo menos 70%, de preferência pelo menos 75%, de preferência pelo menos 80%, de preferência pelo menos 85%, de preferência a pelo menos 90%, mais preferencialmente pelo menos 95% e ainda mais preferencialmente pelo menos 98% idênticos à SEQ ID NO: 14, 2 ou o seu complemento, de preferência a SEQ ID NO: 2 ou o seu complemento. De acordo com uma forma de realização mais preferida, o ácido nucleico compreende a sequência nucleotídica SEQ ID NO: 14, 2 ou o seu complemento, de preferência SEQ ID NO: 2 ou o seu complemento. De acordo com uma concretização ainda

mais preferida, o ácido nucleico consiste em SEQ ID NO: 14, 2 ou o seu complemento, de preferência de SEQ ID NO: 2 ou o seu complemento.

(0077) De acordo com outra configuração, o ácido nucléico é isolado de uma planta das espécies *Santalum*, preferivelmente de *Santalum album*.

(0078) O ácido nucléico da invenção pode ser definido como incluindo polímeros de desoxirribonucleotídeo ou ribonucleotídeo na forma de filamento simples ou duplo (DNA e/ou RNA). Os termos “sequência de nucleotídeo” devem também ser compreendidos como compreendendo uma molécula de polinucleotídeo ou uma molécula de oligonucleotídeo na forma de um fragmento separado ou como um componente de um ácido nucléico maior. Ácidos nucléicos da invenção também abrangem certas sequências de nucleotídeo isoladas, incluindo aquelas que são substancialmente livres de material endógeno contaminante. O ácido nucléico da invenção pode ser truncado, considerando-se que codifique um polipeptídeo abrangido pela presente invenção, conforme descrito acima.

(0079) O ácido nucléico da invenção pode estar presente naturalmente em plantas das espécies de sândalo ou outras espécies ou ser obtido por modificação da SEQ ID NO: 2, 4 ou o complemento desta. Preferivelmente, o referido ácido nucléico consiste de uma sequência de nucleotídeo que foi obtida pela modificação de SEQ ID NO: 2, 4 ou o complemento desta.

(0080) O ácido nucleico da invenção pode estar presente naturalmente em plantas da espécie *santalum* ou outras espécies, ou ser obtido modificando SEQ ID NO: 14, 2 ou o seu complemento, de preferência SEQ ID NO: 2 ou o seu complemento. De preferência, o referido ácido nucleico consiste numa sequência de nucleotídeos que foi obtida modificando SEQ ID NO: 14, 2 ou o seu complemento, de preferência SEQ ID NO: 2 ou o seu complemento.

(0081) Os ácidos nucleicos que compreendem uma sequência obtida por mutação de SEQ ID NO: 14, 2 ou o seu complemento são abrangidos pela invenção, desde que as sequências que compõem partilham pelo menos a percentagem definida de identidade com os fragmentos correspondentes de SEQ ID NO: 14, 2 ou o seu complemento e desde que codifiquem um polipeptídeo com uma atividade de  $\beta$ -santaleno sintase, como definido em qualquer uma das formas de realização acima. De preferência, a sequência é obtida por mutação de SEQ ID NO: 2 ou o seu complemento. As mutações podem ser qualquer tipo de mutações desses ácidos nucleicos, tais como mutações pontuais, mutações de deleção, mutações de inserção e / ou mutações de mudança de quadro. Um ácido nucleico variante pode ser preparado de modo a adaptar a sua sequência de nucleotídeos a um sistema de expressão específico. Por exemplo, os sistemas de expressão bacteriana são conhecidos por expressar polipeptídios de forma mais eficiente se os aminoácidos forem codificados por um códon preferido. Devido à degeneração do código genético, em que mais de um códon pode codificar o mesmo aminoácido, múltiplas sequências de DNA podem codificar para o mesmo polipeptídeo, estando abrangidas todas estas sequências de DNA pela invenção.

(0082) Os ácidos nucléicos que compreendem uma sequência obtida por mutação de SEQ ID NO: 2, 4 ou o complemento desta são abrangidos pela invenção, considerando-se que as sequências que eles compreendem compartilhem, no mínimo, a percentagem definida de identificação com os fragmentos correspondentes de SEQ ID NO: 2, 4 ou o complemento desta e considerando-se que codifiquem um polipeptídeo tendo uma atividade de  $\beta$ -santaleno sintase, conforme definido em qualquer uma das configurações acima. Mutações podem ser qualquer tipo de mutações desses ácidos nucléicos, como mutações pontuais,

mutações de deleção, mutações de inserção e/ou as mutações de estrutura de leitura. Um ácido nucléico variante pode ser preparado a fim de adaptar sua sequência de nucleotídeo a um sistema de expressão específico. Por exemplo, sistemas de expressão bacteriana são conhecidos por expressar mais eficientemente polipeptídios se aminoácidos forem codificados por um códon preferido. Devido à degeneração do código genético, em que mais de um códon pode codificar o mesmo aminoácido, múltiplas sequências de DNA podem codificar o mesmo polipeptídio, todas essas sequências de DNA sendo abrangidas pela invenção.

(0083)O ácido nucleico da invenção pode estar presente naturalmente em plantas da espécie *santalum* ou outras espécies, ou ser obtido modificando SEQ ID NO: 14, 2 ou o seu complemento, de preferência SEQ ID NO: 2 ou o seu complemento. De preferência, o referido ácido nucleico consiste numa sequência de nucleótidos que foi obtida modificando SEQ ID NO: 14, 2 ou o seu complemento, de preferência SEQ ID NO: 2 ou o seu complemento.

(0084)Os ácidos nucleicos que compreendem uma sequência obtida por mutação de SEQ ID NO: 14, 2 ou o seu complemento são abrangidos pela invenção, desde que as sequências que compõem partilham pelo menos a percentagem definida de identidade com os fragmentos correspondentes de SEQ ID NO: 14, 2 ou o seu complemento e desde que codifiquem um polipéptido com uma actividade de  $\beta$ -santaleno sintase, como definido em qualquer uma das formas de realização acima. De preferência, a sequência é obtida por mutação de SEQ ID NO: 2 ou o seu complemento. As mutações podem ser qualquer tipo de mutações desses ácidos nucleicos, tais como mutações pontuais, mutações de deleção, mutações de inserção e / ou mutações de mudança de quadro. Um ácido nucleico variante pode

ser preparado de modo a adaptar a sua sequência de nucleótidos a um sistema de expressão específico. Por exemplo, os sistemas de expressão bacteriana são conhecidos por expressar polipeptídeos de forma mais eficiente se os aminoácidos forem codificados por um codão preferido. Devido à degeneração do código genético, em que mais de um codão pode codificar o mesmo aminoácido, múltiplas sequências de DNA podem codificar para o mesmo polipéptido, estando abrangidas todas estas sequências de DNA pela invenção.

(0085) Outra ferramenta importante para transformação de organismos ou células hospedeiros, adequados para realizar o método da invenção *in vivo* é um vetor de expressão, compreendendo um ácido nucléico de acordo com qualquer configuração da invenção. Esse vetor, portanto, também é um objeto da presente invenção

(0086) Um “vetor de expressão”, conforme usado no presente, inclui qualquer vetor recombinante linear ou circular, incluindo, entre outros, vetores virais, bacteriófagos e plasmídeos. O perito é capaz de selecionar um vetor adequado, de acordo com o sistema de expressão. Em uma configuração, o vetor de expressão inclui o ácido nucléico da invenção operavelmente ligado a, no mínimo, uma sequência regulatória, que controla transcrição, tradução, iniciação e término, como um promotor, operador ou aumentador transcricionais ou um sítio de ligação ribossomal de mRNA e, opcionalmente, incluindo, no mínimo, um marcador de seleção. Sequências de nucleotídeo são “operavelmente ligadas” quando a sequência regulatória funcionalmente relaciona-se ao ácido nucléico da invenção.

(0087) Os vetores de expressão da presente invenção podem ser usados nos métodos para preparação de um organismo e/ou célula hospedeiros, geneticamente transformados, em organismos e/ou células hospedeiros, que abrigam os ácidos

nucléicos da invenção e nos métodos para criação de polipeptídios tendo uma atividade de  $\beta$ -santaleno sintase, conforme divulgada abaixo.

(0088) Organismos e células hospedeiros, não-humanos, recombinantes transformados para abrigar, no mínimo, um ácido nucléico da invenção de maneira que heterologicamente expresse ou superexpresse, no mínimo, um polipeptídio da invenção também são ferramentas muito úteis para realizar o método da invenção. Esses organismos e células hospedeiros, não-humanos são, portanto, outro objeto da presente invenção.

(0089) Um ácido nucléico, de acordo com qualquer uma das configurações acima descritas, pode ser usado para transformar os organismos e células hospedeiros, não-humanos e o polipeptídio expressado pode ser qualquer um dos polipeptídios acima descritos.

(0090) Organismos hospedeiros não-humanos da invenção podem ser quaisquer organismos multicelulares ou unicelulares não-humanos. Em uma configuração preferida, o organismo hospedeiro não-humano é uma planta, um procarionte ou um fungo. Qualquer planta, procarionte ou fungo é adequado para ser transformado, de acordo com a presente invenção. Plantas particularmente úteis são aquelas que naturalmente produzem altas quantidades de terpenos. Em uma configuração mais detalhada, a planta é selecionada da família de *Solanaceae*, *Poaceae*, *Brassicaceae*, *Fabaceae*, *Malvaceae*, *Asteraceae* ou *Lamiaceae*. Por exemplo, a planta é selecionada dos gêneros *Nicotiana*, *Solanum*, *Sorghum*, *Arabidopsis*, *Brassica* (colza), *Medicago* (alfalfa), *Gossypium* (algodão), *Artemisia*, *Salvia* e *Mentha*. Preferivelmente, a planta pertence às espécies de *Nicotiana tabacum*.

(0091) Em uma configuração mais preferida, o organismo hospedeiro não-humano é um microorganismo. Qualquer microorganismo é adequado para a presente

invenção, mas de acordo com uma configuração ainda mais preferida, o referido microorganismo é uma bactéria ou levedura. Mais preferivelmente, a referida bactéria é *E. coli* e a referida levedura é *Saccharomyces cerevisiae*.

(0092) Células eucarióticas maiores isoladas podem também ser transformadas, ao invés de organismos completos. Como células eucarióticas maiores, referimo-nos, no presente, a qualquer célula eucariótica não-humana, exceto células de levedura. Células eucarióticas maiores preferidas são células vegetais ou células fúngicas.

(0093) O termo “transformado” refere-se ao fato de que o hospedeiro foi submetido à engenharia genética para compreender uma, duas ou mais cópias de cada um dos ácidos nucléicos, exigidos em qualquer uma das configurações acima descritas. Preferivelmente, o termo “transformado” relaciona-se a hospedeiros heterologicamente expressando os polipeptídios codificados pelo ácido nucléico com os quais eles são transformados, bem como superexpressando os referidos polipeptídios. Conformemente, em uma configuração, a presente invenção fornece um organismo transformado, no qual os polipeptídios são expressados em maior quantidade do que no mesmo organismo não transformado.

(0094) Há diversos métodos conhecidos na técnica para a criação de organismos ou células hospedeiros transgênicos, como plantas, fungos, procariontes ou culturas de células eucarióticas maiores. Vetores de clonagem e expressão adequados para uso com hospedeiros celulares bacterianos, fúngicos, de levedura, vegetais e mamíferos são descritos, por exemplo, em Pouwels *et al.*, *Cloning Vector: A Laboratory Manual*, 1985, Elsevier, New York and Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, 2<sup>nd</sup> edition, 1989, Cold Spring Harbor Laboratory Press. Vetores de clonagem e expressão de plantas maiores e/ou células vegetais, em particular, são disponíveis ao perito. Vide, por exemplo, Schardl *et al.* *Gene* 61:

1-11, 1987.

(0095) Métodos de transformação de organismos ou células hospedeiros para abrigar ácidos nucleicos transgênicos são conhecidos pelo perito. Para a criação de plantas transgênicas, por exemplo, métodos atuais incluem: eletroporação de protoplastos vegetais, transformação mediada por lipossomo, transformação mediada por agrobactéria, transformação mediada por polietilenoglicol, bombardeio por partícula, microinjeção de células vegetais e transformação usando vírus.

(0096) Em uma configuração, DNA transformado é integrado em um cromossomo de um organismo e/ou célula hospedeiros, não-humanos de maneira que se resulte em um sistema recombinante estável. Qualquer método de integração cromossômica, conhecido na técnica, pode ser usado na prática da invenção, incluindo, entre outros, troca de cassete mediada por recombinase (RMCE), inserção cromossômica específica de sítio viral, injeção de adenovírus e pronuclear.

(0097) A fim de realizar o método de produção de  $\beta$ -santaleno *in vitro*, conforme exposto acima, é muito vantajoso fornecer um método de criação de, no mínimo, um polipeptídeo tendo uma atividade de  $\beta$ -santaleno sintase, conforme descrito em qualquer configuração da invenção. Portanto, a invenção fornece um método de produção de, no mínimo, um polipeptídeo, de acordo com qualquer configuração da invenção, compreendendo

(0098) o cultivo de um organismo ou célula hospedeiros, não-humanos, transformados com o vetor de expressão da invenção, de maneira que abrigue um ácido nucleico, de acordo com a invenção e expresse ou superexpresse um polipeptídeo da invenção;

(0099) o isolamento do polipeptídeo de organismo ou célula hospedeiros, não-humanos, cultivados na etapa a).

- a. De acordo com uma configuração preferida, o referido método ainda abrange, antes da etapa a), a transformação de um organismo ou célula hospedeiros, não-humanos com o vetor de expressão da invenção, de maneira que abrigue um ácido nucléico, de acordo com a invenção e expresse ou superexpresse um polipeptídio da invenção;
- b. Um ácido nucléico, de acordo com qualquer uma das configurações acima descritas, pode ser usado.
- c. A transformação e o cultivo do organismo ou célula hospedeiros, não-humanos podem ser realizados conforme descrito acima para o método de produção de  $\beta$ -santaleno *in vivo*. A etapa b) pode ser realizada usando qualquer técnica bem conhecida na técnica para isolar um polipeptídio particular de um organismo ou célula.
- d. Uma “variante de polipeptídio”, conforme referida no presente, significa um polipeptídio que tem uma atividade de  $\beta$ -santaleno sintase e é substancialmente homólogo ao polipeptídio, de acordo com qualquer uma das configurações acima, mas tendo uma sequência de aminoácido diferente daquela codificada por qualquer uma das sequências de ácido nucléico da invenção devido a uma ou mais deleções, inserções ou substituições.
- e. Variantes podem abranger sequências conservativamente substituídas, significando que um dado resíduo de aminoácido é substituído por um resíduo tendo características físicoquímicas semelhantes. Exemplos de substituições conservativas incluem substituição de um resíduo alifático por outro, como Ile, Val, Leu ou Ala por um outro, ou substituições de um resíduo polar por outro, como entre Lys e Arg; Glu e Asp; ou Gln e Asn. Vide Zubay, *Biochemistry*, 1983, Addison-Wesley Pub. Co. Os efeitos dessas substituições podem ser calculados usando

matrizes de classificação de substituição, como PAM-120, PAM-200 e PAM-250, conforme discutido em Altschul, J. *Mol. Biol.*, 1991, 219, 555-565. Outras substituições conservativas, por exemplo, substituições de regiões inteiras tendo características de hidrofobicidade semelhantes são bem conhecidas.

(00100) Variantes de peptídeo que ocorrem naturalmente também são abrangidas pela invenção. Exemplos dessas variantes são proteínas que resultam de eventos de splicing de mRNA alternativos ou de divisão proteolítica dos polipeptídios descritos no presente. Variações atribuíveis à proteólise incluem, por exemplo, diferenças nos terminais N ou C mediante expressão em diferentes tipos de células hospedeiras, devido à remoção proteolítica de um ou mais aminoácidos terminais dos polipeptídios codificados pelas sequências da invenção.

(00101) Variantes dos polipeptídios da invenção podem ser usados para atingir, por exemplo, atividade enzimática elevada ou reduzida desejada, regioquímica ou estereoquímica modificada ou utilização de substrato alterada ou distribuição de produto, afinidade elevada para o substrato, especificidade melhorada para a produção de um ou mais compostos desejados, velocidade elevada da reação enzimática, atividade ou estabilidade maior em um ambiente específico (pH, temperatura, solvente, etc.) ou nível de expressão melhorado em um sistema de expressão desejado. Um mutante direcionado por variante ou sítio pode ser criado por qualquer método conhecido na técnica. Variantes e derivados de polipeptídios nativos podem ser obtidos por isolamento de variantes que ocorrem naturalmente ou a sequência de nucleotídeo de variantes, de outras ou mesmas linhagens ou espécies de planta, por exemplo, plantas das espécies de *Santalum* ou por programação artificial de mutações de sequências de nucleotídeo que codificam os polipeptídios da invenção. Alterações da sequência de aminoácido nativa podem

ser realizadas por qualquer um de diversos métodos convencionais.

(00102) Variantes de polipeptídio resultantes de uma fusão de sequências de peptídeo adicional nas extremidades de amino e carboxila-terminal dos polipeptídios da invenção podem ser usadas para aumentar a expressão dos polipeptídios, ser úteis na purificação da proteína ou melhorar a atividade enzimática do polipeptídio em um ambiente ou sistema de expressão desejado. Essas sequências de peptídeo adicionais podem ser peptídeos de sinal, por exemplo. Conformemente, a presente invenção abrange variantes dos polipeptídios da invenção, como aqueles obtidos por fusão com outros oligo ou polipeptídios e/ou aqueles que são ligados a peptídeos de sinal. Polipeptídio de fusão, abrangidos pela invenção, também compreendem polipeptídios de fusão resultantes de uma fusão de outras proteínas funcionais, como outras proteínas da via de biossíntese de terpeno.

(00103) Portanto, em uma configuração, a presente invenção fornece um método de preparação de um polipeptídio variante tendo uma atividade de  $\beta$ -santaleno sintase, conforme descrito em qualquer uma das configurações acima e compreendendo as etapas de:

- (a) seleção de um ácido nucléico, de acordo com qualquer uma das configurações acima expostas;
- (b) modificação do ácido nucléico selecionado para obter, no mínimo, um ácido nucléico mutante;
- (c) transformação de células ou organismos unicelulares hospedeiros com a sequência de aminoácido mutante para expressar um polipeptídio codificado pela sequência de ácido nucléico mutante;
- (d) análise do polipeptídio com relação a, no mínimo, uma propriedade modificada; e,

(e) opcionalmente, se o polipeptídio não tiver atividade de  $\beta$ -santaleno sintase variante desejada, repetir as etapas (a) a (d) do processo até um polipeptídio com uma atividade de  $\beta$ -santaleno sintase variante desejada ser obtido;

(f) opcionalmente, se um polipeptídio tendo uma atividade de  $\beta$ -santaleno sintase variante desejada tiver sido identificado na etapa (d), isolar o ácido nucléico mutante correspondente obtido na etapa (c).

(00104) De acordo com uma configuração preferida, o polipeptídio variante preparado é capaz de produzir uma mistura de sesquiterpenos onde  $\beta$ -santaleno representa, no mínimo, 20%, preferivelmente, no mínimo, 30%, preferivelmente, no mínimo, 35%, dos sesquiterpenos produzidos.

(00105) Na etapa (b), uma grande quantidade de sequências de ácido nucléico pode ser criada, por exemplo, por mutagênese randômica, mutagênese específica de sítio ou recombinação homóloga do DNA. Os procedimentos detalhados de recombinação homóloga genética são encontrados em Stemmer, recombinação homóloga de DNA por fragmentação e remontagem randômicas: recombinação *in vitro* para evolução molecular. *Proc Natl Acad Sci USA.*, 1994, 91(22): 10747–1075. Resumidamente, a recombinação homóloga do DNA refere-se a um processo de recombinação randômica de sequências conhecidas *in vitro*, envolvendo, no mínimo, dois ácidos nucléicos selecionados para recombinação. Por exemplo, mutações podem ser introduzidas em locais particulares pela síntese de oligonucleotídeos contendo uma sequência mutante, flanqueada por sítios de restrição que permitem a ligação a fragmentos da sequência nativa. Após ligação, a sequência reconstruída resultante codifica um análogo tendo a inserção, substituição ou deleção de aminoácido desejada. Alternativamente, procedimentos de mutagênese de sítio específico, direcionada por oligonucleotídeo podem ser

empregados para fornecer um gene alterado em que códons pré-determinados podem ser alterações por substituição, deleção ou inserção.

(00106) Conformemente, o polipeptídio compreendendo SEQ ID NO:1 pode ser recombinado com quaisquer outros ácidos nucléicos que codificam sesquiterpeno sintase, por exemplo, isolados de um organismo diferente de *Santalum album*. Dessa forma, ácidos nucléicos mutantes podem ser obtidos e separados, que podem ser usados para transformação de uma célula hospedeira de acordo com os procedimentos padrão, por exemplo, conforme divulgado nos presentes exemplos.

(00107) Na etapa (d), o polipeptídio obtido na etapa (c) é analisado com relação a, no mínimo, uma propriedade modificada, por exemplo, uma atividade enzimática modificada desejada. Exemplos de atividades enzimáticas desejadas, para as quais um polipeptídio expressado pode ser analisado, incluem atividade enzimática elevada ou reduzida, conforme medido por valor de  $K_M$  ou  $V_{max}$ , regioquímica ou estereoquímica modificadas e utilização de substrato ou distribuição de produto alteradas. A análise da atividade enzimática pode ser realizada de acordo com procedimentos conhecidos pelo perito e aqueles divulgados nos presentes exemplos.

(00108) A etapa (e) fornece repetição das etapas (a)-(d) do processo, que podem preferivelmente ser realizadas em paralelo. Conformemente, através da criação de uma quantidade significativa de ácidos nucléicos mutantes, muitas células hospedeiras podem ser transformadas com diferentes ácidos nucléicos mutantes ao mesmo tempo, permitindo a análise subsequente de uma quantidade elevada de polipeptídios. As chances de obtenção de um polipeptídio variante desejado podem, dessa forma, ser elevadas a critério do perito.

(00109) Todas as publicações mencionadas neste pedido são incorporadas por

referência para divulgar e descrever os métodos e/ou materiais em conexão com os quais as publicações são citadas.

#### Descrição dos desenhos

(00110) Figura 1: Análise de GC/MS do sesquiterpeno produzido pela santaleno sintase recombinante de *Santalum album* (SaSantS).

(00111) Figura 1: Cromatograma de íon total. 1, *α*-santaleno; 2, *trans-α*-bergamoteno; 3, *epi-b*-santaleno; 4, *b*-santaleno; 5, *b*-farneseno

(00112) Figuras 2 a 4: Espectros de massa dos picos identificados como sesquiterpenos.

(00113) Figura 5: Estrutura molecular de  $\beta$ -santaleno e  $\beta$ -santalol.

#### Configurações específicas da invenção ou Exemplos

(00114) A invenção agora será descrita em detalhes adicionais através dos seguintes Exemplos.

#### Exemplo 1

(00115) Construção, sequenciamento e extração da biblioteca de DNA de sequências relacionadas com terpeno sintase

(00116) Jovens segmentos de hipocótilos, obtidos de sementes assepticamente germinadas de *Santalum album* L. (5 semanas de idade), foram usados para induzir formação de calos. As sementes de *S. album* foram obtidas da B&T World Seeds (Aigues-Vives, França) e da Sandeman Seeds (Lalongue, França). As sementes foram primeiramente esterilizadas na superfície em HClO a 2,5% durante 120 minutos e enxaguadas três vezes em água ultrapura estéril. As sementes foram, então, descascadas e colocadas em meio basal de MS (Murashige & Skoog, 1962, *Physiologia Plantarum* 15, 473-497), suplementado com 15g/L de sacarose e 7,8g/L de ágar, pH de 5,7. A germinação foi tipicamente observada após

9 a 18 dias com um rendimento de aproximadamente 40%. As plantas micropropagadas foram deixadas crescer *in vitro* durante 2 a 3 meses em uma sala de cultivo em uma temperatura de 27°C, com luz fluorescente, fria, branca e com um fotoperíodo de 16 horas. Para induzir a formação de calos verdes, os segmentos de hipocótilos foram cortados em segmentos transversais de 3-4mm, que foram colocados em meio basal de Gbg (Gamborg & al, 1968, *Exp Cell Res.* 50(1), 151-158) suplementado com 2,4D (ácido 2,4-Diclorofenoxiacético, Sigma-Aldrich Co.) a 0,5µM e Kin (Cinetina, Sigma-Aldrich Co.) a 10µM em placas de Petri. O crescimento do calo foi perpetuada por transferência do tecido a cada quatro semanas a meio fresco em placas de Petri. Todas as culturas de calo foram realizadas em uma câmara de crescimento nas mesmas condições conforme acima.

(00117) O calo obtido após um mês de cultura em meio de Gbg contendo Kin a 5µM e ACC a 2mM foram usados para a extração de RNA e construção da biblioteca de cDNA. RNA total foi extraído seguindo o protocolo descrito por Lefort e Douglas (Ann. For. Sci. 56 (1999), 259-263) exceto que o tratamento de RNase foi omitido. A pelota foi ressuspensa em água sem RNase a 200µl e centrifugada duas vezes durante 10 minutos a 20000g para remover os polissacarídeos. Aproximadamente 125µg de RNA total foram obtidos de 2,2g de células. Os mRNAs foram purificados usando o Kit de Isolamento de mRNA FastTrack® 2.0 (Invitrogen) e uma biblioteca de cDNA foi criada usando o Kit de Síntese de cDNA por PCR SMART® (Clontech Laboratories, Inc.) seguindo as instruções do fabricante.

(00118) A tecnologia de sequenciamento paralelo massivo de pequenas fragmentos de DNA desenvolvida pela Illumina (San Diego, Califórnia) foi usada para sequenciar a biblioteca de cDNA inteira. A preparação do DNA para sequenciamento, o sequenciamento e a montagem das sequencias foram realizadas

por Fasteris SA (Plan-les-Ouates, Suíça). A biblioteca de cDNA foi tratada seguindo o Kit de Preparo de Amostra Genômica (Illumina) e sequenciada no sistema Analisador de Genoma (Illumina). Um total de 4,03 milhões de sequências de 35bp (sequencias) foi obtido. Essas sequências foram montadas usando EDENA 2.1.1, um *software* que encontra sobreposições entre as leituras e monta contigs *de novo* (Hernandez *et al*, Sequenciamento de genoma bacteriano *de novo*: Milhões de sequencias muito curtas montadas em um computador. *Genome Res.* 2008;18:802–809). A montagem foi realizada com equivalências mínimas de 26 a 20 bases. Após eliminar os contigs menores do que 100 bases, 1983 a 3473 contigs únicos foram obtidos com um comprimento máximo de 1331 a 1914 dependendo dos parâmetros selecionados para a montagem. Outra montagem foi realizada usando o programa Velvet 1.0 (Zerbino and Birney (2008), Velvet: algoritmos para montagem de sequencia curta *de novo* usando gráficos de de Bruijn. *Genome Res.* 18(5), 821-829), fornecendo 5905 contigs únicos de comprimento entre 100 e 1616 bases.

(00119) Todos contigs gerados foram comparados contra um banco de dados de sequências de proteína (sequências de proteína não-redundantes, NCBI, <http://www.ncbi.nlm.nih.gov>) usando o algoritmo Blastx (Altschul *et al.*, *J. Mol. Biol.* 215, 403-410, 1990; <http://www.ncbi.nlm.nih.gov/blast/Blast.cgi>). Os contigs mostrando homologia de sequência significativa com sesquiterpeno sintases de planta foram retidos. Um total de 46 contigs com um comprimento de 100 a 621 bases foram, dessa forma, selecionados. Esses contigs foram, então, processados usando o programa CAP (Huang, *Genomics* 14(1), 18-25, 1992) para montá-los e gerar sequências maiores. Cinco contigs únicos de comprimento de 445 a 1064 foram, dessa forma, montados. As sequências de aminoácido deduzidas mostraram homologia significativa com terpeno sintases de planta e especialmente com

sequências descritas ou anotadas como monoterpeno sintases. O alinhamento dessas sequências de aminoácido mostrou que, no mínimo, dois cDNAs distintos foram presentes (duas sequências foram encontradas na maioria das posições entre o alinhamento). Esse alinhamento mostrou também que, no mínimo, uma sequência N-terminal e uma C-terminal estava presente. Para obter as sequências de comprimento total e determinar as sequências de extremidade 5' e extremidade 3' exatas para cada cDNA, uma amplificação rápida de experimento de extremidades de cDNA (RACE) foi empregada.

### Exemplo 2

(00120) Amplificação das sequências de comprimento total de um cDNA de terpeno sintase

(00121) Para os experimentos RACE, um conjunto de primers foi designado de um dos cinco contigs obtidos conforme descrito acima. Dessa forma, os primers diretos SCH5-Ct58-R1 (SEQ ID NO:6) e SCH5-Ct58-R2 (SEQ ID NO:7) e os primers reversos SCH5-Ct58-F3 (SEQ ID NO:8) e SCH5-Ct58-F4 (SEQ ID NO:9) foram deduzidos de SCH5-contig-5 (SEQ ID NO:5).

(00122) A PCR foi realizada com a Mistura Universal de Primer A (UPM) (Kit de Amplificação de cDNA por RACE SMART™, Clontech Laboratories, Inc.) em 50µl de volume final contendo 200µM de mistura de dNTPs, 5µl de biblioteca de cDNA (Exemplo 1), 0,2µM de primer específico de gene, 0,2µM de Mistura de Primer UPM (Clontech Laboratories, Inc.), 1µl de Mistura de Polimerase de Vantagem 2 (Clontech Laboratories, Inc.) e 5µl de Tampão de Reação por PCR de 10x cDNA (Clontech Laboratories, Inc.). As condições de ciclagem térmica foram as seguintes: 3 minutos a 94°C; ~5 ciclos de 30 segundos a 94°C e 3 minutos a 72°C; 5 ciclos de 30 segundos a 94°C e 3 minutos a 70°C; 5 ciclos de 30 segundos a 94°C e 3

minutos a 68°C; 3 minutos a 72°C. Com o 5'RACE, um fragmento de DNA de 610bp (SCH5-Ct58\_RR1, SEQ ID NO:10) incluindo a extremidade 5' do cDNA foi obtido. Com o 3'RACE, um fragmento de 1049bp (SCH5-Ct58-RF4, SEQ ID NO:11) foi obtido e a combinação dos dois produtos de RACE com a sequência SCH5-contig-5 (SEQ ID NO:5) permitiu a reconstituição de um novo cDNA de comprimento total (SCH5-Ct58, SEQ ID NO:12). O cDNA de SCH5-Ct58 de 2157bp, codificado para uma proteína de aminoácido 569 (SEQ ID NO:13) mostrando homologia com sequências de terpeno sintases de planta e contendo motivos característicos de terpeno sintases, como o motivo DDxxD presente em todas as monoterpeno e sesquiterpeno sintases. De forma interessante, a sequência de aminoácido mostrou maior similaridade às monoterpeno sintases do que às sesquiterpeno sintases. Entretanto, a presença de sinal de peptídeo do cloroplasto, uma característica comum nas monoterpeno sintases de planta, não foi prevista a partir da análise da sequência N-terminal (Emanuelsson, O., Nielsen, N., and von Heijne, G. 1999. ChloroP, um método baseado em rede neural para prever peptídeos de trânsito do cloroplasto e seus sítios de divisão Protein Science 8, 978-984).

### Exemplo 3

(00123) Expressão heteróloga e atividade enzimática *in vitro* de SCH5-Ct58

(00124) Decidimos modificar a sequência de DNA de SCH5-Ct58 (SEQ ID NO:12) e recriar a sequência para expressão heteróloga favorável em células de *E. coli*. Para começar com a sequência de aminoácido verdadeira, a sequência nucleotídica exata de SCH5-Ct58 na biblioteca de cDNA primeiramente teve de ser estabelecida. O *Software* Eland (Illumina) foi usado para recuperar todas as sequências equivalentes com a sequência SCH5-Ct58 (SEQ ID NO:12) com um máximo de 2 divergências. Um total de 5224 sequências foram recuperadas e foram

alinhas usando o programa CAP (Huang, Genomics 14(1), 18-25, 1992) com a sequência de DNA SCH5-Ct58 (SEQ ID NO:12) como referência. A cobertura média sobre a sequência inteira foi acima de 100X permitindo a dedução não ambígua da nova sequência de cDNA SCH5-Ct94 (SEQ ID NO:14). Nessa nova sequência, 5 bases foram corrigidas em comparação com a sequência SCH5-Ct58 (SEQ ID NO:12) deduzida dos resultados RACE e aquelas correções resultaram em uma nova sequência de aminoácido (SCH5-Ct94, SEQ ID NO:15) com uma diferença de dois resíduos. Para expressão heteróloga, a sequência de DNA de SCH5-Ct94 (SEQ ID NO:14) foi modificada para remover os primeiros 23 códons e substituir pela sequência ATGGCT e o uso de códon foi alterado para otimizar a sequência para expressão de *E. coli* (DNA 2.0, Menlo Park, CA, EUA). O cDNA, dessa forma, criado (SCH5-Ct94-opt, SEQ ID NO:2) foi sintetizado (DNA 2.0, Menlo Park, CA, EUA) e subclonado nos sítios *NdeI-KpnI* do plasmídeo pETDuet-1 fornecendo o plasmídeo Ct94-pETDuet. Essa sequência de cDNA otimizada codificou o polipeptídeo SCH5-Ct94-opt (SEQ ID NO:1).

(00125) Expressão heteróloga de Ct94 foi realizada em células BL21(DE3) de *E. coli* usando o plasmídeo Ct94-pETDuet. Ensaios enzimáticos *in vitro* foram realizados com FPP como substrato nas condições descritas acima e atividade de sesquiterpeno sintase foi observada com formação de uma mistura de cinco sesquiterpenos. A identificação desses sesquiterpenos foi confirmado por GC-MS como sendo o sesquiterpeno característico de *santalum album*: *a*-santaleno, *trans-a*-bergamoteno, *epi-b*-santaleno, *b*-santaleno e *b*-farneseno (Figura 1). Em pH de 7,0 e na presença de MgCl<sub>2</sub> a 15mM, a proporção relativa dos produtos de sesquiterpeno recombinantes foi 38,0% de *a*-santaleno, 18,2% de *trans-a*-bergamoteno, 5,7% de *epi-b*-santaleno, 36,7% de *b*-santaleno e 1,3% de *b*-

farneseno. Dessa forma, o cDNA de SCH-Ct98-opt codificou uma  $\beta$ -santaleno sintase. A razão dos produtos foi muito semelhante à proporção observada no óleo de *Santalum album* com relação aos produtos hidroxilados desses sesquiterpenos. Nenhuma atividade foi detectada quando  $MgCl_2$  foi omitido e o meio suplementado com EDTA a 2,5mM (para realização quelação de cátions residuais) mostrando a exigência estrita de cátions divalentes. A natureza e a concentração do cátion divalente presente no ensaio tiveram um efeito no perfil do produto (Tabela 1). Por exemplo, a diminuição da concentração de  $Mg^{2+}$  teve um efeito benéfico relacionado com  $\beta$ -santaleno, o último se tornando o produto principal da enzima. Além disso, a adição de  $Mn^{2+}$  teve um efeito negativo na formação de  $\beta$ -santaleno já que a proporção dos produtos de santaleno sesquiterpeno diminuiu e a proporção de *trans-a*-bergamoteno e *b*-farneseno aumentou, *trans-a*-bergamoteno sendo o principal produto da enzima na presença de  $MgCl_2$  a 1mM.

(00126) Tabela 1: Efeito da concentração de íons  $Mg^{2+}$  e  $Mn^{2+}$  na composição da mistura de sesquiterpenos obtidos pelo contato da SEQ ID NO:1 com FPP

|                             | Porcentagem, relativa à mistura de produto inteira |                |                   |                                    |
|-----------------------------|----------------------------------------------------|----------------|-------------------|------------------------------------|
|                             | $MgCl_2$ a 15mM                                    | $MgCl_2$ a 2mM | $MgCl_2$ a 0,75mM | $MgCl_2$ a 0,75mM + $MnCl_2$ a 1mM |
| <i>a</i> -santaleno         | 38,0                                               | 33,0           | 36,5              | 24,5                               |
| <i>trans-a</i> -bergamoteno | 18,2                                               | 11,8           | 12,6              | 35,4                               |
| <i>epi-b</i> -santaleno     | 5,7                                                | 6,4            | 5,6               | 4,1                                |
| <i>b</i> -santaleno         | 36,7                                               | 47,5           | 44,1              | 33,3                               |
| <i>b</i> -farneseno         | 1,3                                                | 1,3            | 1,1               | 2,75                               |

Exemplo 4

(00127) Produção *in vivo* de sesquiterpenos em *E coli* usando o Ct94 cDNA

(00128) O uso da santaleno sintase de *S. album* para a produção *in vivo* de sesquiterpenos em células de *E coli* foi avaliado pela coexpressão das enzimas de uma via biossintética de cinco etapas convertendo ácido mevalônico em FPP.

(00129) O gene de FPP sintase de levedura foi amplificado de DNA genômico de *S. cerevisiae* usando os primers FPPy\_NcoI (SEQ ID NO:16) E fppY-Eco (SEQ ID NO:17). O DNA amplificado foi ligado como fragmento NdeI-EcoRI no primeiro sítio de clonagem múltipla (MCS1) do plasmídeo pACYCDuet-1 fornecendo o plasmídeo FPPs-pACYCDuet abrigando o gene FPPs sob o controle do promotor T7. Um operon, incluindo os genes que codificam uma mevalonato quinase (mvaK1), uma fosfomevalonato quinase (mvaK2), uma mevalonato difosfato decarboxilase (MvaD) e uma isopentenil difosfato isomerase (idi), foi amplificado do DNA genômico de *Streptococcus pneumoniae* (ATCC BAA-334) com os primers MVA-up1-start (SEQ ID NO:18) e MVA-up2-stop (SEQ ID NO:19). A PCR foi realizada usando a polimerase de DNA de Fusão HS PfuUltra™ II (Stratagen). A composição da mistura de PCR foi de acordo com as instruções do fabricante. As condições de ciclagem térmica foram 2 minutos a 95°C; 30 ciclos de 20 segundos a 95°C, 20 segundos a 58°C e 90 segundos a 72°C; e 3 minutos a 72°C. O fragmento 3,8Kb foi purificado em um gel de agarose e ligado usando o Kit de Clonagem por PCR Dry-Down In-Fusion™ (clontech) no segundo MCS do plasmídeo FPPs-pACYCDuet, digerido com NdeI e XhoI, fornecendo o plasmídeo pACYCDuet-4506. As sequências das duas inserções foram totalmente sequenciadas para excluir qualquer mutação.

(00130) Células BL21 Star™(DE3) de *E. coli* (Invitrogen) foram

cotransformadas com os plasmídeos pACYDuet-4506 e Ct94-pETDuet e células transformadas foram selecionadas em placas de carbenicilina (50µg/ml), cloramfenicol (34µg/ml) de LB-agarose. Colônias únicas foram usadas para inocular meio de LB a 5mL com 50µg/ml de carbenicilina e 34µg/ml de cloramfenicol. A cultura foi incubada durante a noite a 37°C. No dia seguinte, 2mL de meio de TB, suplementado com os mesmos antibióticos foi inoculado com 0,2mL da cultura noturna. Após incubação de 6 horas a 37°C, a cultura foi resfriada para 28°C e IPTG a 1mM, 2mg/mL de mevalonato (preparada por dissolução de mevalonolactona (Sigma) em NaOH a 0,5N em uma concentração de 1g/mL e incubando a solução durante 30 minutos a 37°C) e decano a 0,2mL foram adicionados a cada tubo. As culturas foram incubadas durante 48 horas a 28°C. As culturas foram, então, extraídas duas vezes com 2 volumes de acetato de etila, a fase orgânica foi concentrada a 500µL e analisada por GC-MS conforme descrito acima no Exemplo 3. Nessas condições, produção de sesquiterpeno acima de 200mg/L foi rotineiramente atingida. Beta-santaleno foi produzido.

#### Exemplo 5

(00131) Produção *in vivo* de sesquiterpenos em *S. cerevisiae* usando o Ct94 cDNA

(00132) Para produção *in vivo* de sesquiterpeno em células de levedura, uma cepa de *saccharomyces cerevisiae* YNP5 em que o gene ERG9 (codificando a esqualeno sintase, a enzima que converte FPP em esqualeno) foi regulada de forma descendente pela substituição do promotor ERG9 nativo pelo promotor MET3 regulável. No trabalho anterior com sesquiterpeno sintases de planta, essa estratégia levou a uma biossíntese de ergosterol reduzida nas células e um acúmulo de FPP disponível para sesquiterpeno sintases (Asadollahi, *Biotechnology and*

*Bioengineering*, 99(3), 666-677, 2008).

(00133) O cDNA de SCH5-Ct94-opt (SEQ ID NO:2) foi amplificado do Ct94-pETDuet com os primers Ct94\_BamHI (SEQ ID NO:20) e T7term (SEQ ID NO:21). A PCR foi realizada com a Polimerase de DNA Pfu (Promega) usando as seguintes condições de ciclagem térmica: 90 segundos a 94°C; 35 ciclos de 45 segundos a 94°C, 45 segundos a 55°C; 4 minutos a 72°C; e 10 minutos a 72°C. O cDNA amplificado foi digerido com os sítios de restrição *Bam*HI e *Xho*I e ligado nos sítios correspondentes do plasmídeo pESC-URA (Stratagen) fornecendo o plasmídeo Ct94-pESC-ura. S.c. As células YNP5 foram transformadas usando Kit de Transformação S.c. EasyComp™ (Invitrogen).

(00134) Uma colônia única de cepas de levedura transformadas foi usada para inocular 20ml de meio de YNB (5g/L de (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>; 3g/L de KH<sub>2</sub>PO<sub>4</sub>; 0,5g/L de MgSO<sub>4</sub>.7 H<sub>2</sub>O; 1mL/L de solução de metal residual), suplementado com glicose a 2%. A cultura foi incubada durante 24 horas a 28°C. As células foram recuperadas por centrifugação e ressuspensas em 20mL de meio de YNB, suplementado com galactose a 2%. Após 1 hora em cultura, metionina em concentração final de 0,5mM e decano a 2mL foram adicionados à cultura. Após incubação de 24 horas a 28°C, as culturas foram extraídas com acetato de etila e analisadas por GC-MS conforme descrito no Exemplo 4. A quantidade total de sesquiterpenos produzidos pelas células de levedura nessas condições foi estimada em 50mg/L.

#### Exemplo 6

(00135) Isolamento de uma santaleno sintase de raízes de *santalum album*

(00136) Mudas de *Santalum album* obtidas de sementes assepticamente germinadas foram transferidas ao solo 5 a 10 semanas após germinação. Já que espécies de *santalum* são hemiparasitas de raiz, a adaptação ao solo foi feita em

contato próximo com plantas cítricas de 6 meses a 1 ano de idade (*Citrus sinensis*). As raízes das plantas de sândalo foram colhidas, 2-3 anos após a transferência aos solos e separadas das raízes da planta hospedeira. Análise de GC-MS de um extrato dessas raízes mostrou a presença do óleo de sândalo, característico dos sesquiterpenos. O RNA total foi extraído das raízes usando o Reagente de RNA de Planta Concert™ (Invitrogen). De 12g de tecido, 640µg de RNA total foram isolados. Os mRNAs foram purificados usando o Kit de Isolamento de mRNA FastTrack® 2.0 (Invitrogen) e uma biblioteca de cDNA foi criada usando o Kit de Amplificação de cDNA Marathon™ (Clontech Laboratories, Inc.) seguindo as instruções do fabricante.

(00137) Uma quantidade de 1µg de cDNA foi usada para sequenciamento usando o Sistema Analisador de Genoma (Illumina). Um total de 10,3 milhões de sequências de comprimento de 35bp foi obtido. Essas sequências foram montadas usando em paralelo os *software* de montagem Edena (Hernandez *et al*, 2008, *Genome Res.* 18, 802–809) e o Velvet (Zerbino and Birney, 2008, *Genome Res.* 18: 821-829) resultando em contigs únicos 18'937 e 22'414 com uma variação média de 242 e 211bp. As sequências foram pesquisadas usando o programa tBlastn (Altschul *et al*, 1990, *J. Mol. Biol.* 215, 403-410) com a sequência de aminoácido SCH5-CT94 (SEQ ID NO:15) como sequência de busca. Quinze contigs foram selecionados, mostrando homologia significativa de suas sequências de aminoácido deduzidas com sesquiterpeno sintases de planta. Esses contigs selecionados foram remontados em duas sequências distintas, das quais SCH10-Ct8201 (SEQ ID NO:22) foi 383bp no comprimento e mostrou a homologia mais alta com a sequência de DNA SCH5-CT94 (SEQ ID NO:14). O primer direto SCH10-Ctg8201-F2 (SEQ ID NO:23) foi criado da sequência SCH10-Ct8201 e usado de

forma bem-sucedida para 3'RACE usando o Kit de Amplificação de cDNA Marathon™ (Clontech Laboratories, Inc.). Da sequência do produto 3'RACE, dessa forma, obtido, dois primers reversos (SCH10-Ct19779-R3 (SEQ ID NO:24) e SCH10-Ct19779-R4 (SEQ ID NO:25)) foram criados e usados de forma bem-sucedida para a amplificação por 5'RACE da extremidade 5' do cDNA correspondente. Das sequências do 3'RACE e 5'RACE, uma sequência de comprimento total de uma nova terpeno sintase foi, dessa forma, reconstituída. A fim de verificar a sequência, o programa MAQ (Li *et al*, 2008, *Genome Res.* 18(11), 1851–1858) foi usado para pesquisar e mapear todas as sequências com um máximo de 2 divergências. Essa abordagem forneceu uma sequência de DNA de comprimento 1725bp (SEQ ID NO:26) codificando uma proteína de comprimento de 570 aminoácidos (SEQ ID NO:27) tendo 91,9% de identificação com a sequência de aminoácido de SCH5-Ct94 (SEQ ID NO:15).

(00138) Para expressão heteróloga em *E coli*, um cDNA otimizado foi criado pela deleção dos primeiros 21 códons, adicionando a sequência ATGGCTACC como os primeiros 3 códons e otimizando o uso de códon para *E coli*. Essa sequência otimizada (SCH10-Ct8201-opt, SEQ ID NO:4) que codifica proteína modificada de N-terminal SCH10-Tps8201-opt (SEQ ID NO:3) foi sintetizada (DNA 2.0; Menlo Park, CA, EUA) e subclonada nos sítios NdeI-KpnI do plasmídeo de expressão pETDuet-1 (Novagen). Expressão heteróloga e caracterização enzimática de SCH10-Tps8201-opt (SEQ ID NO:3) foram realizadas conforme descrito no Exemplo 3. A proteína recombinante mostrou atividade de sesquiterpeno sintase e produziu de FPP a mesma mistura de sesquiterpenos como a proteína recombinante SCH5-CT94 (SEQ ID NO:1, Exemplo 3) com as mesmas proporções relativas.

## REIVINDICAÇÕES

### 1. MÉTODO PARA PRODUÇÃO DE B-SANTALENO, CARACTERIZADO

por compreender

- a) o contato do polipeptídeo com pirofosfato de farnesila (FPP) com, no mínimo, um polipeptídeo tendo uma atividade de  $\beta$ -santaleno sintase e compreendendo uma sequência de aminoácido idêntica à SEQ ID NO:15 ou SEQ ID NO: 27; ou codificada por uma sequência de nucleotídeo idêntica a SEQ ID NO: 2 ou SEQ ID NO: 14;
- b) opcionalmente, o isolamento do  $\beta$ -santaleno produzido na etapa a).

### 2. MÉTODO, de acordo com a reivindicação 1, **CARACTERIZADO** pela etapa

- a) compreender o cultivo de um organismo hospedeiro ou célula hospedeira não-humana, capazes de produzir um polipeptídeo com pirofosfato de farnesila (FPP) e transformado para expressar, no mínimo, um polipeptídeo compreendendo uma sequência de aminoácido idêntica à SEQ ID NO:15 ou SEQ ID NO: 27 e tendo uma atividade de  $\beta$ -santaleno sintase ou codificada por uma sequência de nucleotídeo idêntica a SEQ ID NO: 2 ou SEQ ID NO: 14, em que o organismo hospedeiro não humano ou células são *E. coli* ou *Saccharomyces cerevisiae*.

- 3. **MÉTODO**, de acordo com a reivindicação 2, **CARACTERIZADO** pelo método ainda compreender, antes da etapa a), a transformação de um organismo ou célula não-humana, capaz de produzir um polipeptídeo com pirofosfato de farnesila (FPP) com, no mínimo, um ácido nucléico, que codifica um polipeptídeo compreendendo uma sequência de aminoácido idêntica à SEQ ID NO:15 ou SEQ ID NO: 27 e tendo uma atividade de  $\beta$ -santaleno sintase ou codificada por uma sequência de nucleotídeo idêntica a SEQ ID NO: 2 ou SEQ ID NO: 14, de maneira que o referido organismo expresse o referido polipeptídeo, em que o organismo hospedeiro não humano ou células são *E. coli* ou *Saccharomyces cerevisiae*.

4. **ÁCIDO NUCLEÍCO ISOLADO RECOMBINANTE, CARACTERIZADO** por compreender uma sequência de nucleotídeo idêntica à SEQ ID NO:2, ou SEQ ID NO:14.

5. **VETOR DE EXPRESSÃO RECOMBINANTE, CARACTERIZADO** por compreender as sequências definidas pelas SEQ ID NO:2, ou SEQ ID NO:14 ligadas de modo operativo às sequências de nucleotídeos.

6. **ORGANISMO HOSPEDEIRO NÃO HUMANO OU CÉLULA HOSPEDEIRA NÃO-HUMANA**, transformados para abrigar, no mínimo, um ácido nucléico, de acordo com a reivindicação 4, por serem, de maneira que heterologicamente expresse ou superexpresse, no mínimo, um polipeptídeo, **CARACTERIZADO** pelo organismo hospedeiro não humano ou célula hospedeira ser *E. coli* ou *Saccharomyces cerevisiae*.

7. **MÉTODO DE PRODUÇÃO**, de, no mínimo, um polipeptídeo, compreendendo:

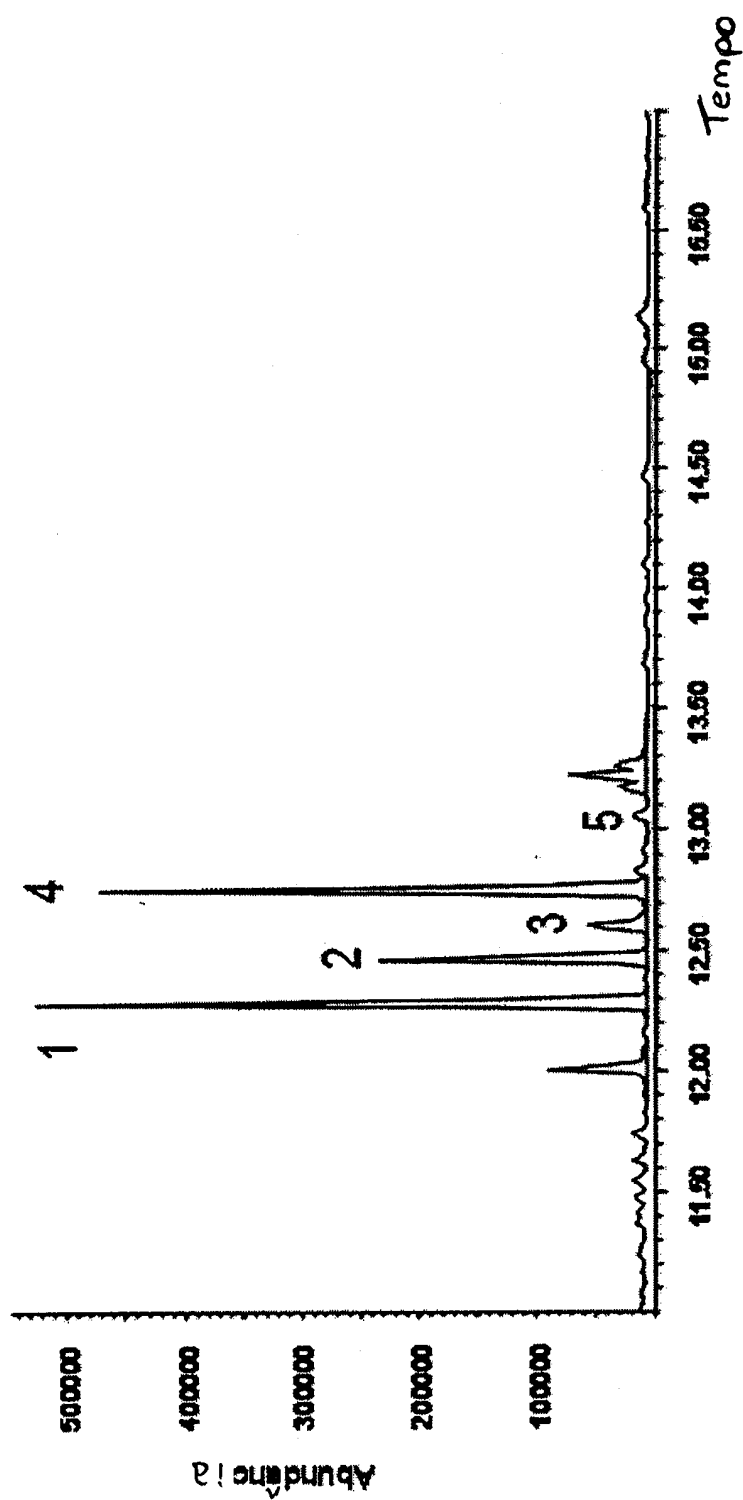
a) o cultivo de um organismo hospedeiro não humano ou célula hospedeiros, transformados com o vetor de expressão da reivindicação 5, de maneira que abrigue um ácido nucléico, de acordo com a reivindicação 4 e expresse ou superexpresse um polipeptídeo;

b) o isolamento do polipeptídeo de organismo ou célula hospedeiros, não-humanos, cultivados na etapa a),

**CARACTERIZADO** pelo fato de que o organismo hospedeiro não humano é um *E. coli* ou *Saccharomyces cerevisiae*.

8. **MÉTODO**, de acordo com a reivindicação 7, ainda compreender, antes da etapa a), a transformação de um organismo hospedeiro não humano ou célula hospedeiro, com o vetor de expressão da reivindicação 5, de maneira que abrigue um

ácido nucléico, de acordo com a reivindicação 4 e expresse ou superexpresse um polipeptídeo, **CARACTERIZADO** pelo fato de que o organismo hospedeiro não humano ser um *E. coli* ou *Saccharomyces cerevisiae*.



Figură 1, parte 4

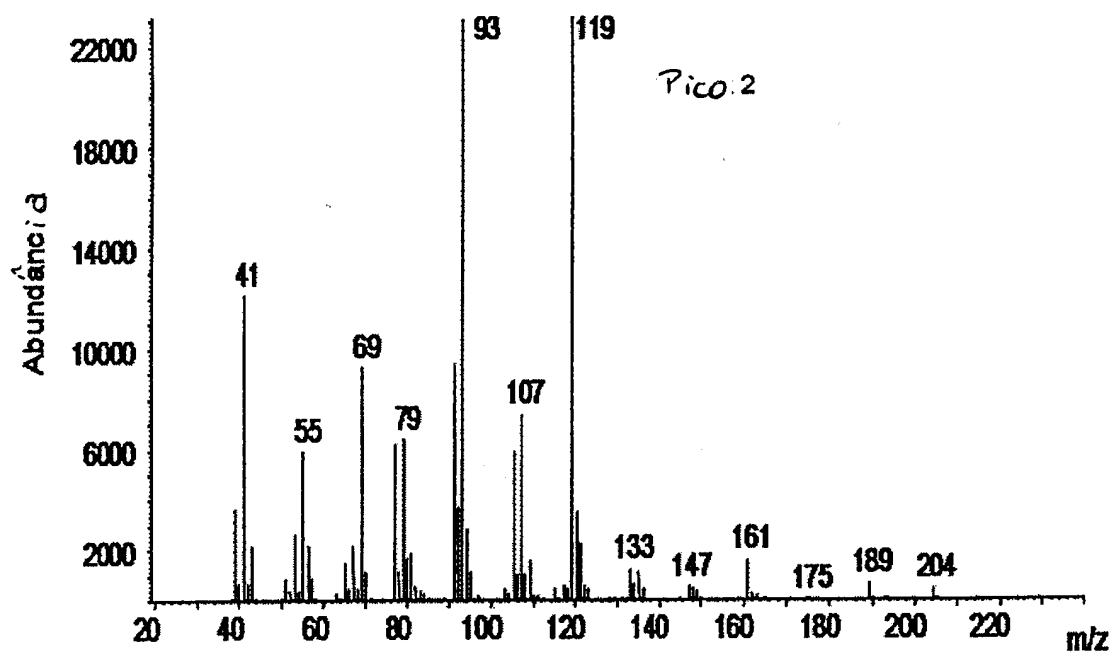
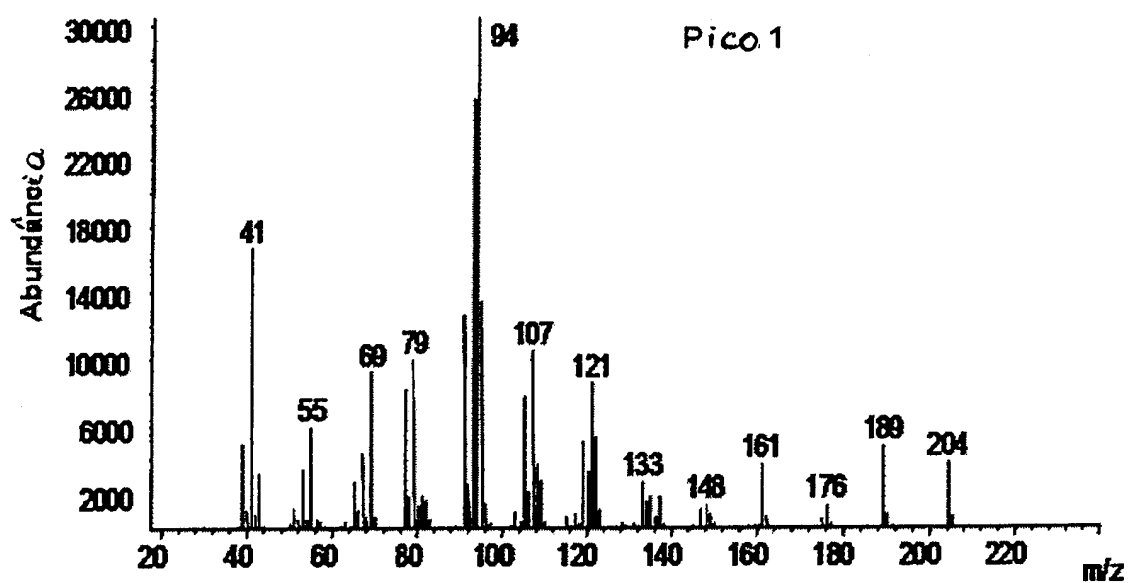


Figura 1, parte 4

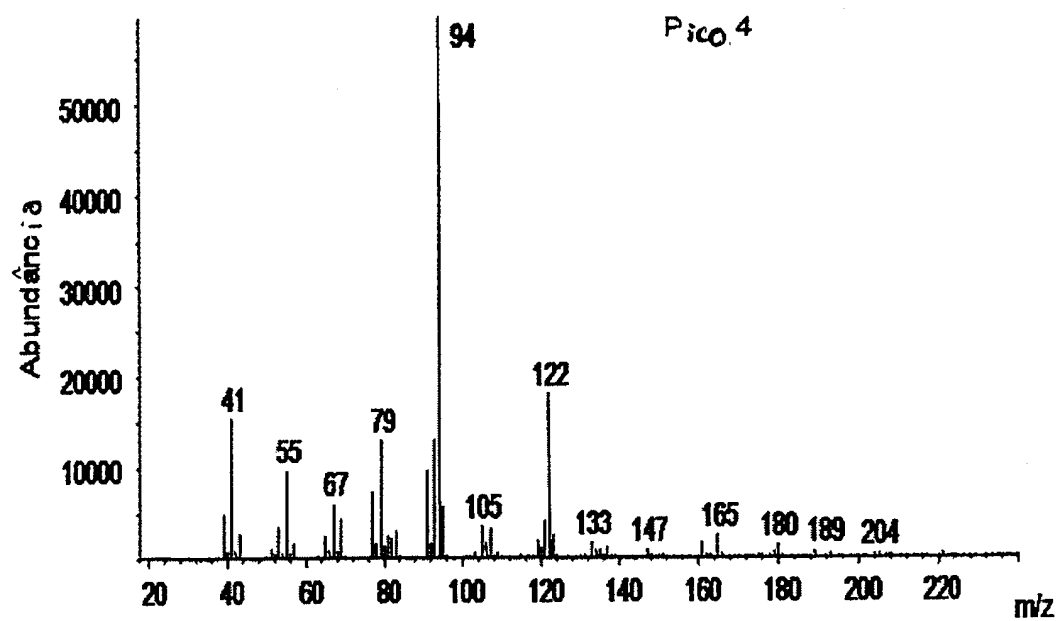
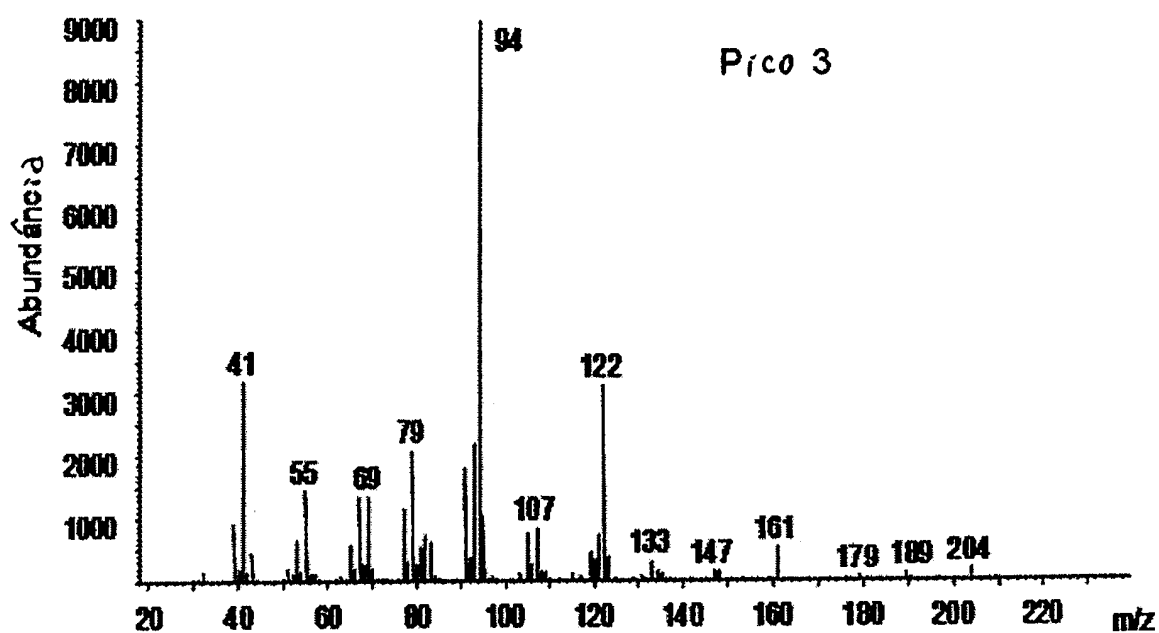


Figura 1, parte 4

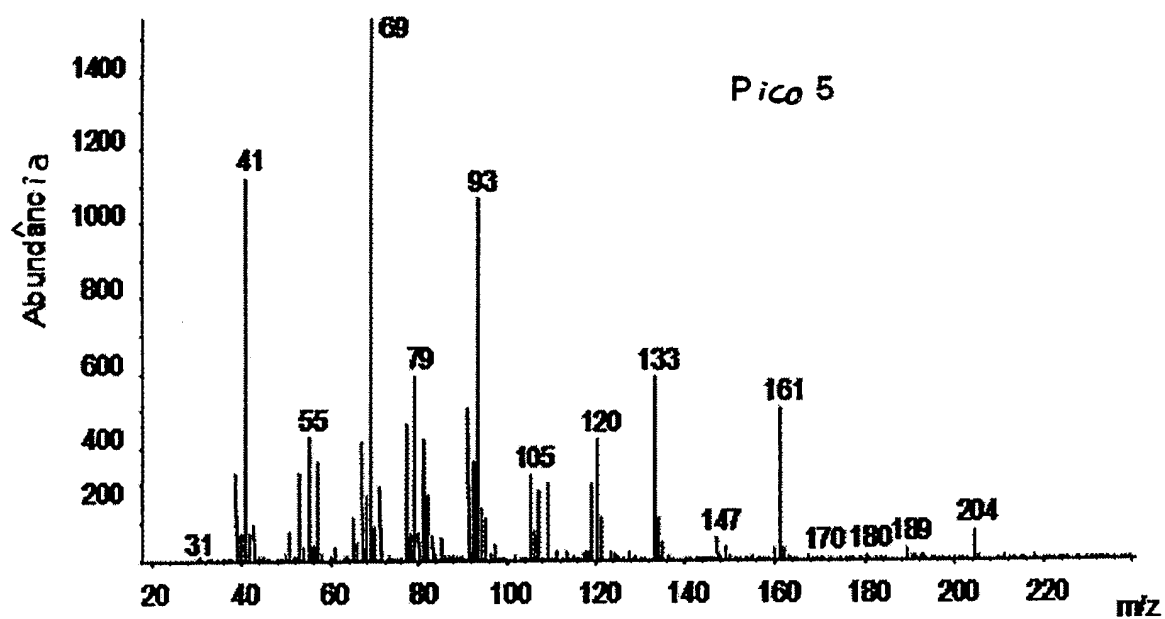
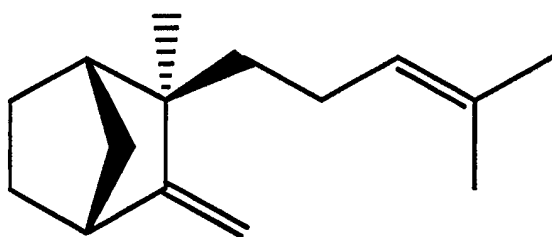
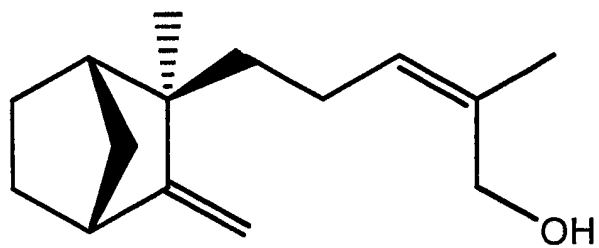


Figura 1, parte 4



Beta Santaleno

Figura 2A



Beta Santalol

Figura 2B