

(22) Data de pedido: **2005.02.25**

(30) Prioridade(s): **2004.02.25 JP 2004049123**

(43) Data de publicação do pedido: **2006.11.15**

(45) Data e BPI da concessão: **2011.01.05**  
**042/2011**

(73) Titular(es):

**MEIJI SEIKA KAISHA LTD.**  
**4-16, KYOBASHI 2-CHOME, CHUO-KU TOKYO**  
**104-8002 JP**

(72) Inventor(es):

**HIROAKI HAYASHI JP**  
**KENICHIRO INOUE JP**  
**MASATERU HOSHINO JP**  
**SHIBUYA MASHIBUYA JP**  
**EBIZUKA YUTAKA JP**

(74) Mandatário:

**JOSÉ EDUARDO LOPES VIEIRA DE SAMPAIO**  
**R DO SALITRE 195 RC DTO 1250-199 LISBOA PT**

(54) Epígrafe: **HIDROXILASE DE TRITERPENO**

(57) Resumo:

O SAPOGENOL B DE SOJA É BIOSINTETIZADO, ATRAVÉS DE DOIS PASSOS DA REACÇÃO DE HIDROXILAÇÃO DO SEU PRECURSOR A &#914;- AMIRINA. CONTUDO, O GENE DA HIDROXILASE ENVOLVIDO NESTA REACÇÃO NÃO TINHA SIDO REVELADO. POR ISSO, ERA IMPOSSÍVEL APLICAR UMA TÉCNICA DE ENGENHARIA GENÉTICA À HIDROXILASE. OS PRESENTES INVENTORES REVELAM QUE UMA SEQUÊNCIA QUE CORRESPONDE AO GENE CYP93E1 DO CITOCROMO P-450 DERIVADO DA SOJA CODIFICA PARA UMA PROTEÍNA ENZIMA QUE EFECTUA A HIDROXILAÇÃO NA POSIÇÃO 24 DE UM TRITERPENO DO TIPO OLEANANO, E FORNECE TAMBÉM UM PROCESSO DE APLICAÇÃO DO DITO GENE UTILIZANDO UMA TÉCNICA DE ENGENHARIA GENÉTICA.

## Descrição

### "Hidroxilase de triterpeno"

#### Campo da invenção

Esta invenção diz respeito a uma célula transformada por uma técnica de engenharia genética com um gene de enzima que está envolvido na biossíntese do sapogenol B de soja derivado de planta, e a um processo para produzir sapogenol B de soja utilizando a célula.

#### Antecedentes da Invenção

O sapogenol B de soja (12-oleaneno-3,22,24-triol) é um triterpeno que possui um esqueleto de oleanano que foi isolado a partir de uma semente de soja e a sua estrutura foi determinada (Chem. Pharm. Bull., 24 p. 121-129, 1976 Chem. Pharm. Bull., 30, p. 2294-2297, 1982) (referências que não são patentes 1 e 2) e o seu glicósido saponina de soja encontra-se largamente distribuído nas plantas leguminosas.

No que diz respeito ao sapogenol B de soja foi até agora reportada a sua actividade anti-complemento e acção inibidora da aglutinação das plaquetas (JP-A-61-37749) (referência 1 das patentes), actividade anti-tumoral (JP-A-10-234396) (referência 2 das patentes) e actividade hepatoprotectora (Bioorg. Med. Chem. Lett., 7, 85-88, 1997) (referência 3 documento que não é uma patente) e semelhantes, e é esperada a sua utilidade como uma preparação farmacêutica, ou um seu material. No que diz respeito ao processo de produção do sapogenol B de soja é conhecido um processo em que cadeias de açúcar de saponina

contidas em sementes de soja são hidrolizadas, e depois o sapogenol B de soja é purificado, mas uma vez que a proporção de saponina contida nas sementes de soja é cerca de 0,2% que é muito pequena (Yakugaku Zasshi (Journal of Pharmaceutical Sciences), 104, 162-168, 1984) (referência 4 do documento que não é patente) é necessário um processo de produção mais eficiente.

Considera-se que a  $\beta$ -amirina como precursor da biossíntese do sapogenol B de soja é biosintetizada, através do fecho do anel de 2,3-oxidoesqualeno que é formado, através da via do mevalonato e o sapogenol B de soja é sintetizado a seguir, através de dois passos de reacções de hidroxilação.

Soforadiol (12-oleaneno-3,22-diol) estruturalmente próximo do sapogenol B de soja é uma substância que foi reportada como sendo um componente da Kaika (Sofora japónica) (Yakugaku Zasshi, 78, 1090-1094, 1958) (referência 5 que não é patente), e é possível produzir sapogenol B de soja por hidroxilação da posição 24 do soforadiol.

Actualmente um processo para produzir sapogenol B de soja por hidroxilação da posição 24 do soforadiol utilizando uma hidroxilase derivada de uma célula cultivada de *Glycyrrhiza glabra* foi revelado (WO 02/086142) (referência 3 das patentes).

Referência 1 das patentes: JP-A-61-37749

Referência 2 das patentes: JP-A-10-234396

Referência 3 patentes: publicação internacional WO 02/086142

Referência 1 que não é patente: Chem. Pharm. Bull., 24 p. 121-129, 1976

Referência 2 que não é patente: Chem. Pharm. Bull., 30 p. 2294-2297, 1982

Referência 3 que não é patente: Bioorg. Med. Chem. Lett., 7, 85-88, 1997

Referência 4 que não é patente: Yakugaku Zasshi (Journal of Pharmaceutical Science), 104, 162-168, 1984

Referência 5 que não é patente: Yakugaku Zasshi, 78, 1090-1094, 1958

## **Detalhes da Invenção**

### **Problemas solucionados pela invenção**

A  $\beta$ -amirina é um precursor biossintético do sapogenol B de soja. É biosintetizada, através do fecho do anel de 2,3-oxidoesqualeno que é formado, através da via do mevalonato e o sapogenol B de soja é sintetizado a seguir através de dois passos da reacção de hidroxilação. Contudo, o gene da hidroxilase envolvido nesta reacção não foi revelado. Por isso, foi impossível aplicar uma técnica de engenharia genética à hidroxilase.

Steele C. L. et al. ("Molecular Characterization of the Aryl Migration Reaction of Isoflavoid Biosynthesis in Soybean"; Archives of Biochemistry and Biophysics, Vol. 367, N° 1, 1999, páginas 146-150) revelam o gene do citocromo P450 clonado num vector de baculovírus e a expressão de P450 em células de insectos e de plantas. O nucleótido CYP93E1 e sequência da proteína foram submetidos à base de dados EMBL/GenBank/DDBJ (N° de acesso à Genbank AF135485).

A US2004/031072A1 revela construções de ADN recombinante compreendendo polinucleótidos (de acordo com as sequências SEQ ID N°:1 à SEQ ID N°:142,842) úteis para a melhoria das plantas. As sequências 100510 e 243352 reveladas na US2004/031072A1 correspondem respectivamente ao gene e polipéptido do citocromo P450. O material

genético exógeno pode ser transferido para uma célula de planta, através da utilização de um vector de ADN que pode ser desenhado para se replicar em *E. coli* e *A. lumefaciens*.

A WO 02086142A1 revela a produção do sapogenol B de soja, através do tratamento do soforadiol com uma hidroxilase originária de uma planta para hidroxilar o composto na sua posição 24.

Hayashi H. et al. (Glycyrrhetic acid 24-hydroxylase activity in microsomes of cultured licorice cells"; Photochemistry vol. 34, N°. 5, 1993, páginas 1303-1307) divulgam a existência de uma monoxigenase dependente do citocromo P-450 que catalisa a hidroxilação do grupo metilo C-24 do ácido glicirretínico.

#### Processos para Resolver os Problemas

Com base no pressuposto de que o gene de uma enzima do tipo do citocromo P-450 envolvido na biossíntese do sapogenol B de soja da  $\beta$ -amirina está contido num clone EST (Etiquetas da sequência de expressão -Expression Sequence Tags) da soja que produz saponina de soja a uma taxa elevada de produção, ou num clone cuja função não está identificada, os presentes inventores efectuaram análises de funções destes clones de soja utilizando uma estirpe mutante de levedura deficiente em lanosterol. Entre os clones analisados, a actividade de hidroxilação da posição 24 de um triterpeno do tipo oleanano que não pode ser detectado inicialmente, foi detectado numa estirpe de levedura em que o polinucleótido representado pela SEQ ID N°:8 foi transcrito e traduzido. A sequência polinucleótídica da actividade de hidroxilação da enzima detectada é a SEQ ID N°:8, e a sequência polipeptídica deduzida é a SEQ ID N°:9. Como uma sequência que é semelhante à SEQ ID N° 8, é conhecido o gene CYP93E1 do citocromo P-450 (N° de acesso à

GenBank AF135485, SEQ ID N°:10, a sequência polipeptídica deduzida é a SEQ ID N°:11). O polinucleótido representado pela SEQ ID N°:8 e o polinucleótido representado pela SEQ ID N°:10 são diferentes um do outro em termos de 3 posições a posição 121<sup>a</sup>, a posição 171<sup>a</sup> e a 108<sup>a</sup> posição (daqui para a frente a sequência representada pela SEQ ID N° 8 é também designada nalguns casos gene CYP93E1 do citocromo P-450). Além disso, o polipéptido representado pela SEQ ID N°:9 e o polinucleótido representado pela SEQ ID N°:11 são diferentes um do outro em termos dos aminoácidos na posição 41<sup>a</sup> e 61<sup>a</sup>. Os presentes inventores revelaram que o polinucleótido representado pela SEQ ID N° 8 codifica para uma proteína de enzima que efectua a hidroxilação da posição 24 de um triterpeno do tipo oleanano. Com base neste facto, os inventores efectuaram esforços intensos e realizaram a invenção.

Assim, a invenção diz respeito a um processo para produzir um triterpeno do tipo oleano em que a posição 24 é hidroxilada, compreendendo:

a cultura de um transformante compreendendo um vector de expressão que possui um polinucleótido que se hibridiza com uma cadeia complementar do polinucleótido representado pela SEQ ID N°:8 em condições severas e também codifica para um polipéptido que possui a actividade de hidroxilar a posição 24 de um triterpeno do tipo oleano de modo a produzir um polipéptido da SEQ ID N°:9; e permitindo o transformante actuar sobre um triterpeno do tipo oleano, em que o transformante é um microrganismo.

Além disso, a presente invenção diz respeito a um processo como descrito acima, em que o polinucleótido é o polinucleótido representado pela SEQ ID N°:8.

Além disso, a presente invenção diz respeito a um processo como especificado acima, em que o transformante é uma levedura.

## Vantagens da Invenção

Através da invenção, foi possível revelar a sequência nucleotídica do gene da hidroxilase da posição 24 de um triterpeno do tipo oleanano e uma sua sequência de aminoácidos. Também a utilização do dito gene através de técnicas de engenharia genética torna possível a produção em grande escala da proteína da enzima como um produto génico.

Além disso, tornou-se possível efectuar a hidroxilação da posição 24 de um triterpeno, através da utilização da proteína enzimática produzida, ou de um transformante contendo a dita proteína enzimática. Também se tornou possível produzir um triterpeno em que a posição 24 é hidroxilada por cultura directa utilizando um transformante transformado com o dito gene e um gene da sintase da  $\beta$ -amirina.

## Melhor Modo de Efectuar a Invenção

Como o triterpeno do tipo oleanano da invenção,  $\beta$ -amirina, soforadiol, sapogenois A e B de soja e semelhantes são conhecidos, mas o triterpeno do tipo oleanano da invenção não se encontra limitado ao caso acima.

A  $\beta$ -amirina e o soforadiol podem ser tomados como exemplos de um triterpeno do tipo oleanano em que a posição 24 vai ser hidroxilada, mas isto não se encontra limitado ao caso acima com a condição de que seja um composto em que a posição 24 vai ser hidroxilada, através do processo da invenção. Exemplos do triterpeno em que a posição 24 se encontra hidroxilada podem ser os sapogenois de soja A e B, mas não se encontram limitados aos sapogenois de soja A e

B, com a condição de que são produtos de oxidação através da invenção.

De acordo com a invenção um triterpeno do tipo oleanano em que a posição 24 vai ser hidroxilada pode ser produzido utilizando os produtos de transcrição e tradução do polinucleótido do gene CYP93E1 do citocromo P-450, e suas formas equivalentes. Neste contexto formas equivalentes significam aquelas sequências que possuem a mesma função e hibridizam-se sob condições severas com uma cadeia complementar da sequência descrita no gene CYP93E1 do citocromo P-450.

No que diz respeito a "hibridizarem sob condições severas" a hibridização de um nucleótido pode ser verificada, através da utilização de um processo (por ex. hibridização de colônia, hibridização em placa, hibridização por transferência de Southern ou semelhante), em que a hibridização é efectuada utilizando como sonda, uma parte ou uma porção completa (ou uma sua cadeia complementar) de um ADN que possui a sequência nucleotídica representada pelo gene CYP39E1 do citocromo P-450. Pode ser dado como exemplo de forma ilustrativa, um caso em que a hibridização é efectuada a 55°C na presença de 0,5 mol/l de cloreto de sódio, e depois 2x uma solução SSC (composição de uma solução 1xSSC consiste em 150 mM de NaCl, 15 mM de citrato de sódio, pH 7,0).

A hibridização pode ser efectuada de acordo com o processo descrito em *Molecular Cloning, A Laboratory Manual*, editado por T. Maniatis et al., Cold Spring Harbor Laboratory, 1989, ou semelhantes. Pode ser exemplificado de forma ilustrativa um ADN que pode ser hibridizado, um ADN que apresenta uma homologia de pelo menos 80% ou mais, preferencialmente 90% ou mais, mais preferencialmente 95% ou mais, com a sequência nucleotídica representada pelo gene CYP93E1 do citocromo P-450, quando calculado utilizando

BLAST (National Center for Biotechnology Information). Neste contexto, a homologia de acordo com a invenção significa o valor numérico, quando os parâmetros de BLAST são colocados para o tamanho de palavras 3, Matriz: BLOSUM 62, Custos de lacunas: Existência:11, Extensão:1.

Para além disso, o “polinucleótido que se hibridiza com uma cadeia complementar do polinucleótido do gene CYP93E1 do citocromo P-450 sob condições severas e codifica também para um polipéptido que possui a actividade de hidroxilar a posição 24 de um triterpeno do tipo oleanano” inclui de forma ilustrativa um polinucleótido em que um ou dois ou mais nucleótidos na sequência polinucleotídica do gene CYP93E1 do citocromo P-450 são suprimidas, substituídas, inseridas, ou adicionadas, e que codifica para um polipéptido que possui a actividade de hidroxilar a posição 24 de um triterpeno do tipo oleanano. O número de nucleótidos pode ser substituído na sequência polinucleotídica do gene CYP93E1 do citocromo P-450 não se encontra particularmente limitado com a condição de que seja um número que satisfaz a homologia anteriormente mencionada e codifica para um polipéptido que possui a actividade de hidroxilar na posição 24 um triterpeno do tipo oleanano.

Uma mutação no polinucleótido do gene CYP93E1 do citocromo P450 inclui uma mutação artificial para além de mutações que foram geradas no mundo natural. Um exemplo dos meios para provocar uma mutação artificial é um processo para obter um polinucleótido em que pelo menos uma supressão, substituição, inserção e adição é efectuada num em dois, ou em vários nucleótidos, através de uma técnica de engenharia genética introduzindo uma mutação aleatória, ou uma mutação dirigida para o local utilizando o polinucleótido do gene CYP93E1 do gene CYP93E1 do citocromo P-450. Utilizando o polinucleótido mutante obtido desta

maneira torna-se possível obter um polipéptido com uma temperatura óptima diferente, estabilidade ao calor, pH óptimo, estabilidade ao pH, especificidade para o substrato e propriedades semelhantes da actividade desta enzima.

Além disso, um polinucleótido que se hibridiza com uma cadeia complementar do polinucleótido do gene CYP93E1 do citocromo P-450 sob condições severas e codifica também para um polipéptido que possui a actividade de hidroxilar a posição 24 de um triterpeno do tipo oleanano pode também ser obtido através de um processo em que uma hibridização (por exemplo hibridização de colónia, hibridização em placa, hibridização de transferência de Southern ou semelhantes) é aplicada a outros microrganismos, plantas, ou animais capazes de produzir um triterpeno do tipo oleanano em que a posição 24 está hidroxilada (preferencialmente uma planta, mais preferencialmente uma planta leguminosa, mais preferencialmente soja, que pode produzir um triterpeno do tipo oleanano), utilizando como sonda, uma parte, ou uma porção completa (ou uma sua cadeia complementar) de um nucleótido que possui a sequência nucleotídica do gene CYP93E1 do citocromo P-450, ou através de um método em que é efectuada uma PCR utilizando como sonda, uma parte, ou a porção completa (ou uma sua porção complementar) de um nucleótido que possui a sequência nucleotídica do gene CYP93E1 do citocromo P-450, ou semelhantes.

Além disso, o polinucleótido anteriormente mencionado pode também ser obtido por síntese química com base na informação da sequência nucleotídica. Este processo pode ser efectuado de acordo com as descrições em Gene, vol. 60(1), p. 115-127(1987).

Além disso a invenção diz respeito a um transformante preparado transformando um hospedeiro com um vector que se

pode replicar de forma autónoma (preferencialmente um vector de expressão) para albergar e/ou expressar o polinucleótido do gene CYP93E1 do citocromo P-450 e seus equivalentes. O dito vector pode ainda conter um gene da sintase de  $\beta$ -amirina para além do gene CYP93E1 do citocromo P-450 e dos seus equivalentes.

É utilizado um microrganismo como hospedeiro. Como exemplo de microrganismo pode ser referida uma levedura, *Escherichia coli* e semelhantes e é utilizada preferencialmente uma levedura. É possível disponibilizar uma planta que possui um teor acrescido de um triterpeno do tipo oleanano em que a posição 24 é hidroxilada transferindo o vector da invenção para a planta.

Como exemplo de uma levedura a ser transformada é a levedura deficiente em lanosterol sintase GIL 77 (Kushiro T. et al., Eur. J. Biochem., 256, 238-244, 1998). Torna-se possível produzir em cultura um triterpeno em que a posição 24 se encontra hidroxilada, integrando um ADNc correspondente ao gene CYP93E1 do citocromo P-450 anteriormente mencionado e um gene da sintase da  $\beta$ -amirina derivado da ervilha num vector de expressão da levedura pESC-ERA (fab. por Stratagene) e transformando com ele uma levedura deficiente em lanosterol sintase GIL 77, e efectuando a co-expressão dos dois genes.

Como vector de expressão, são utilizados aqueles que se podem replicar de forma autónoma na célula hospedeira ou que podem ser integrados no cromossoma, e possuem um promotor na posição em que o polinucleótido da invenção pode ser transcrito.

Como exemplo de um vector de expressão, quando a célula hospedeira é um microrganismo, é por exemplo pBluescript (fabr. por Stratagene), pUC18 (fabr. por Takara Bio), pUC118 (fabr. por Takara Bio), pUC19 (fabr. por Takara Bio), pUC119 (fabr. por Takara Bio) e semelhantes.

No que diz respeito ao promotor pode ser qualquer promotor que pode efectuar a expressão em *Escherichia coli*, um fungo e células hospedeiras semelhantes. Por exemplo, podem ser citados um promotor *trp* ( $P_{trp}$ ), um promotor *lac* ( $P_{lac}$ ) e promotores semelhantes derivados de *Escherichia coli*, um fago e semelhantes, e um promotor do gene da Taka-amilase, um promotor do gene *TEF1* e promotores semelhantes derivados de uma estirpe de *Aspergillus*.

Além disso, um promotor concebido artificialmente e modificado, ou semelhantes pode ser utilizado.

No que diz respeito ao processo de transferir um vector recombinante, pode ser utilizado qualquer processo para transferir um polinucleótido nas células hospedeiras anteriormente mencionadas e pode por exemplo ser citado um processo que utiliza um íão cálcio [Proc. Natl. Acad. Sci. USA, 69, 2110(1972)] e semelhantes.

Exemplos de vectores de expressão, quando uma estirpe de levedura é a célula hospedeira são por exemplo, pAUR101 (fabr. por Takara Bio), pAUR112 (fabr. por Takara Bio), pI-REDI (fabr. por Toyobo) e semelhantes.

Como promotor pode ser qualquer promotor que pode efectuar a expressão na estirpe de levedura. Por exemplo, podem ser citados um promotor do gene da enzima do metabolismo glicolítico, um promotor *Gal* e promotores semelhantes.

No que diz respeito a um processo para transferir um vector recombinante pode ser utilizado qualquer processo para transferir um polinucleótido para uma estirpe de levedura, e por exemplo, pode ser citado a electroporação [Methods. Enzymol., 194, 182 (1990)], o método do esferoplasto [Proc. Natl. Acad. Sci. USA, 84, 1929 (1978)], o processo do acetato de lítio [Journal of Bacteriology (J. Bacteriol.), 153, 163(1983)] e semelhantes.

No que diz respeito ao meio e condições de cultura das células hospedeiras é possível seleccioná-los opcionalmente de acordo com processos convencionalmente conhecidos. Quando um microrganismo é utilizado como célula hospedeira, o meio a ser utilizado para cultivar o transformante obtido pode ser um meio natural, ou um meio sintético com a condição de ser um meio que contém uma fonte de carbono, uma fonte de azoto, sais inorgânicos e semelhantes que podem ser assimilados pelo dito microrganismo e que pode efectuar a cultura do transformante de forma eficiente.

Como fonte de carbono pode ser utilizada, dextrose de batata, glucose, sacarose, amido solúvel, glicerol, dextrina, melaços, ácidos orgânicos e análogos. Como fonte de azoto pode ser utilizado sulfato de amónio, carbonato de amónio, fosfato de amónio, acetato de amónio e sais orgânicos semelhantes ou sais de amónio de ácidos orgânicos, outros compostos de azoto, peptona, extracto de levedura, água de maceração do milho, hidrolisados de caseína e extracto de carne. Como sais inorgânicos podem ser utilizados fosfato de potássio primário, fosfato de potássio secundário, fosfato de magnésio, sulfato de magnésio, cloreto de sódio, sulfato ferroso, sulfato de manganês, sulfato de cobre, carbonato de cálcio e semelhantes.

Quando a célula hospedeira é um bicho da seda pode ser expresso o polipéptido da invenção, por exemplo, através de um processo conhecido convencionalmente que utiliza um sistema de expressão de baculovírus [Appl. Microbiol. Biotechnol., 62, 1-20(2003)]. Também por exemplo, quando é obtida uma planta transformada com o polipéptido da invenção utilizando uma célula de planta como hospedeiro é eficaz uma transferência directa do gene que utiliza um sistema de plasmídeo Ti ou um plasmídeo binário de *Agrobacterium tumefaciens*, um plasmídeo Ri de *Agrobacterium*

*rhizogenes* ou polietileno glicol ou a electroporação [Methods in Molecular Biology, 267, Recombinant Gene Expression, 329-50(2004)].

Além disso, quando um transformante transferido com um vector de expressão que possui um promotor indutível é cultivado pode ser adicionado um indutor ao meio conforme necessário. Por exemplo, isopropil- $\beta$ -D-tiogalactopiranósido ou semelhante pode ser adicionado ao meio, quando o promotor lac é utilizado, ou ácido indolacrílico ou semelhante, quando o promotor trp é utilizado.

Neste contexto, no que diz respeito a um processo para expressar o polipéptido da invenção que não a sua expressão directa este pode ser efectuado de acordo com o processo descrito na segunda edição de Molecular Cloning ou semelhantes.

Pode ser produzido um triterpeno do tipo oleanano em que a posição 24 se encontra hidroxilada utilizando o transformante descrito acima. O transformante é cultivado no meio adicionando um triterpeno do tipo oleanano como substrato. O composto hidroxilado na posição 24 obtido deste modo é extraído com acetato de etilo, éter, ou solvente orgânico semelhante e purificado utilizando sílica gel ou ODS.

Além disso, um triterpeno do tipo oleanano em que a posição 24 é hidroxilada pode também ser produzido preparando um extracto isento de células do meio de cultura do transformante. Naquele caso, as células recolhidas são suspensas num líquido de suspensão, as células resultantes são destruídas utilizando um homogeneizador, um desintegrador sónico ou uma prensa francesa, ou semelhantes e depois centrifugadas para se obter um extracto isento de células. De modo a evitar a inactivação do polipéptido pode ser adicionado ao líquido tampão um anti-oxidante, um estabilizador enzimático, um agente adsorvente

polifenólico, um ligando metálico e semelhantes. É eficaz purificar o polipéptido para um maior aumento da actividade específica, e centrifugação com uma centrífuga, adição de sal ("salting out") com sulfato de amónio, ou semelhante, filtração em gel, cromatografia de permuta iónica, cromatografia de afinidade, electroforese e técnicas semelhantes podem ser utilizadas, isoladamente, ou em combinação.

O triterpeno do tipo oleanano pode ser utilizado como substrato e uma co-enzima pode ser adicionada ao polipeptido obtido deste modo contendo tampão e incubado a uma temperatura de 15 a 40°C, preferencialmente de 20 a 37°C. Como coenzima pode ser utilizada NADH ou NADPH, e um sistema de reconstrução de NADPH que utiliza glucose 6-fosfato e glucose 6-fosfato desidrogenase pode também ser utilizado conjuntamente. Além disso, é também possível efectuar a reacção de hidroxilação adicionando uma redutase NADPH-P-450 diferente da NADPH-P-450 redutase produzida pelas células de transformação.

Quando são utilizados os transformantes descritos no parágrafo [0012] que contêm ainda um gene da sintase da  $\beta$ -amirina é produzido um triterpeno do tipo oleanano utilizando o 2,3-oxidoesqualeno produzido pelas próprias células transformantes de modo que o triterpeno do tipo oleanano hidroxilado na posição 24 pode ser produzido sem adicionar o triterpeno do tipo oleanano do resíduo exterior. O composto hidroxilado na posição 24 assim obtido é extraído com acetato de etilo, éter ou um solvente orgânico semelhante e depois purificado utilizando sílica gel ou ODS.

Resumos dos exemplos da invenção são descritos a seguir.

Foram seleccionadas sete espécies de EST derivada da soja e clones do citocromo P-450, cujas funções não estão

identificadas, mas foram divulgadas as sequências nucleótidas completas (Nº de acesso à GenBank: AF135485, Y10491, Y10982, Y10983, Y10493 e AF022459, e Nº de acesso TIGRA: TC 100921). Entre elas, CYP93E1 (N<sup>os</sup> de acesso à GenBank AF 135485) apresentou actividade e a sequência polinucleótídica SEQ ID Nº:8 que apresentou uma homologia elevada em relação a ela é descrita a seguir. Foi amplificado um ADNc que corresponde a CYP93E1 (SEQ ID Nº:8), através do método RT-PCR a partir do ARN preparado a partir de rebentos de soja e integrados num vector de expressão da levedura pESC-ERA (fabr. por Stratagene), e uma levedura deficiente em lanosterol sintase GIL 77 (Kushiro, T. et al., Eur. J. Biochem., 256, 238-244, 1998) foi transformada com ele para efectuar análise funcional. Deixou-se reagir extractos isentos de células dos transformantes de levedura com  $\beta$ -amirina e os produtos foram acetilados e analisados por GC-MS (cromatografia gasosa-espectrometria de massa). Foi detectado como resultado 3,24-diacetóxi-12-oleaneno.

Da mesma maneira, extractos isentos de células dos transformantes de levedura foram deixados reagir com soforadiol, e os produtos foram acetilados e analisados por GC-MS. Como resultado foi detectado triacetilsapogenol de soja B.

Foi adicionada  $\beta$ -amirina ao meio de cultura de transformantes de levedura e deixados reagir e depois as células foram recolhidas. Fracções de gordura solúvel foram extraídas e acetiladas e depois analisadas por GC-MS. Foi detectado como resultado 3,24-diacetóxi-12-oleaneno.

O ADNc anteriormente referido da SEQ ID Nº:8 e o gene da sintase da  $\beta$ -amirina derivado da ervilha foi integrado num vector de expressão da levedura pESC-ERA (fabr. por Stratagene), e a levedura deficiente em lanosterol sintase GIL 77 foi com ele transformada, afectando deste modo a co-

expressão dos dois genes. Esta levedura transformada denominada GIL77/pESC.PSY.CYP93E1 foi depositada em 6 de Fevereiro, 2004, como FERM P-19675 (transferido para FERMBP-10201 em Janeiro de 2006, 2005) no Depositário de Organismos das Patentes de Invenção Internacionais, National Institute of Advanced Industrial Science and Technology (código postal 305-8566; Central 6, 1-1-1 Higashi, Tsukuba, Ibaraki, Japão).

A levedura transformada foi cultivada e as células foram recolhidas. A fracção solúvel em gordura foi extraída e acetilada e depois analisada por GC-MS.

Foi detectado como resultado 3,24-diacetóxi-12-oleaneno.

Da mesma maneira o ADNC anteriormente referido da SEQ ID N°:8 e o gene da sintase de triterpenos misturados derivado de *Arabidopsis thaliana* foi integrado num vector de expressão da levedura pESC-ERA (fabr. por Stratagene), e a levedura deficiente em lanosterol sintase GIL77 foi com ele transformada, afectando deste modo a co-expressão dos dois genes.

A levedura transformada foi cultivada e as células foram recolhidas. A fracção solúvel em gordura foi extraída e acetilada e depois analisada por GC-MS.

Foi detectado como resultado 3,24-diacetóxi-12-oleaneno. Não foram detectados outros diacetóxitriterpenos.

Com base nos resultados acima a actividade de hidroxilação na posição 24 do sofradiol e da  $\beta$ -amirina que não foi detectada inicialmente foi confirmada de modo que foi capaz de revelar que a SEQ ID N:8 é um gene que codifica para uma enzima que faz a hidroxilação na posição 24 de um triterpeno do tipo oleanano. Por outro lado, esta actividade não foi detectada nos outros genes 6P-450 examinados da mesma maneira.

Seguidamente descreve-se a invenção mais em detalhe com base em exemplos, mas a invenção não se encontra limitada a estes exemplos.

#### Exemplo 1

##### (1) Preparação de ADNc de rebentos de soja

O ARN total foi extraído, através do método de fenol/clorofórmio de folhas jovens de soja (soja verde de amadurecimento precoce, Atariya Noen)) após 14 dias de impregnação em água. Utilizando isto como molde foi preparado ADNc utilizando um Superscript II de transcriptase reversa (fabr. por GIBCO BRL) e o iniciador indicado na SEQ ID N°:1.

##### (2) Amplificação do polinucleótido da SEQ ID N:8

Utilizando o ADNc preparado em (1) anteriormente referido como molde, e os oligo fragmentos de ADN que correspondem ao terminal N e terminal C do polipéptido apresentado na SEQ ID N°: 2 e 3 como iniciadores, foi efectuada uma PCR (30 ciclos, Ex Taq ADN polimerase fabr. por Takara Shuzo) a uma temperatura de emparelhamento a 65°C para obter um clone de CYP93E1 de comprimento completo (SEQ ID N°:8).

##### (3) Construção de pESC-CYP93E1 e preparação de levedura transformada

O clone de comprimento completo em (2) acima foi tratado com Spel e ClaI e integrado no local Spel-ClaI de um vector de expressão de pESC-URA (fabr. por Stratagene). Este foi designado pESC-CYP93E. pESC-CYP93E foi transferido para uma

estirpe de levedura INVSC 2 (fabr. por Invitrogen) utilizando transformação II de levedura EZ congelada (fabr. por Zymo Research).

(4) Teste da actividade da enzima *in vitro*

A levedura transformada foi inoculada em 20 ml de meio SC-U (Methods in Yeast Genetics, A Laboratory Course Manual, Cold Spring Harbor Laboratory Press, 1990) contendo de 2% de rafinose em vez de glucose e cultivada a 30°C e a 220 rpm durante 18 horas. Adicionou-se Hemin (13 µg/ml na concentração final) e galactose (2% na concentração final) para efectuar a cultura durante 20 horas nas mesmas condições. As células foram recolhidas por centrifugação, transferidas para um frasco roscado de 2 ml e novamente suspensas adicionando 100 µl de um tampão de extracção (preparado adicionando 10% de sacarose, 1 mM de EDTA e 14 mM de 2-mercaptoetanol a 50 mM de tampão de fosfato de potássio de pH 7,5). Esferas de vidro possuindo um diâmetro de 0,4 a 0,6 mm (fabr. por Luchi Seieido) foram lavadas com ácido clorídrico diluído e adicionadas. Arrefecendo a 4°C foi efectuada a quebra das células utilizando MINI-BEADBREAD-ER (fabr. por BIOSPEC). Isto foi ainda misturado com 400 µl do tampão de extracção e bem agitado, e depois centrifugado a 3500 g durante 5 minutos enquanto que se arrefece a 4°C para recuperar cerca de 400 µL do sobrenadante como um líquido de enzima em bruto. A isto adicionou-se 100 µl de um tampão de reacção concentrado (preparado adicionando 10 mM de NADPH, 75 mM de glucose-6-fosfato (C6P) e 2,5 U/ml de glucose-6-fosfato desidrogenase (G6PDH) ao tampão de extracção) e 5 µl de 10 mM de β-amirina em solução de metanol. Deixou-se reagir a 30°C durante 6 horas. Após se ter adicionado 10 µl de ácido

clorídrico 12 N, isto foi concentrado efectuando a extracção de componentes solúveis em gordura duas vezes utilizando 500  $\mu\text{l}$  de acetato de etilo. A acetilação do extracto foi efectuada adicionando 20  $\mu\text{l}$  de piridina e de anidrido acético e deixando a repousar de um dia para o outro. A reacção foi parada adicionando 200  $\mu\text{l}$  de solução aquosa de metanol a 50%, e isto foi concentrado efectuando a extracção duas vezes utilizando 200  $\mu\text{l}$  de hexano (1)). Como teste de controlo, as amostras foram preparadas, através do mesmo processo em 2) um caso em que um líquido de enzima em bruto derivou de um transformante preparado utilizando pESC-URA, 3) um caso em que não foi adicionado  $\beta$ -amirina como substrato, 4) um caso em que a reacção foi efectuada utilizando um líquido de enzima em bruto tratada por calor a 100°C durante 5 minutos, e 5) um caso em que a mesma quantidade de glucose foi adicionada em vez de galactose com o objectivo de inibir o promotor GAL 1 de pESC-CYP93E1. Cada um deles foi dissolvido em 20  $\mu\text{l}$  de hexano e uma sua porção de 1  $\mu\text{l}$  foi submetida a análise por GC-MS (um espectrómetro de massa de cromatografia gasosa GCMS-QP 2010 fabr. pela empresa Shimadzu, coluna: Rtx-5MS fabr, por RESTEK, diâmetro interno 0,25 mm, espessura da membrana 0,25  $\mu\text{m}$ , comprimento 30 m, programa de rampa de temperatura: 3 minutos para manter uma temperatura a 230°C, subida de temperatura a 10°C/min, 8 minutos para manter a temperatura a 330°C). Foi efectuada uma monitorização de iões totais (TIM), e a presença, ou ausência de produto foi analisada por um cromatograma de massa de  $m/z=218$  (pico de base de  $3\beta$ , 24-diacetóxi-12-oleaneno). A formação de  $3\beta$ , 24-diacetóxi-12-oleaneno não foi encontrada por TIM (os resultados não são descritos). Contudo, foi observado  $3\beta$ , 24-diacetóxi-12-oleaneno no cromatograma de massas  $m/z=218$  nas condições de 1) (Fig. 1). Com base neste resultado foi

confirmado que o produto de tradução CYP93E1 possui a actividade para efectuar a hidroxilação da posição 24 da  $\beta$ -amirina.

A seguir, para examinar a sua reactividade para o soforadiol foi efectuada a reacção enzimática da mesma maneira utilizando o soforadiol (5  $\mu$ l, 10 mM) como substrato, e a análise por GC-MS foi efectuada da mesma maneira que acima. O pico de triacetil sapogenol de soja B foi detectado por análise por cromatografia de massa (m/z=216 pico de base do triacetil sapogenol de soja B) do produto sob as condições reaccionais do acima mencionado 1) (Fig. 2). Com base neste resultado foi revelado que o produto de tradução CYP93E1 possui a actividade para efectuar a hidroxilação da posição 24, não apenas da  $\beta$ -amirina, mas também do soforadiol.

#### (5) Teste da actividade da enzima *in vivo*

A levedura transformada foi inoculada em 20 ml do meio SC-U contendo 2% de rafinose em vez de glucose e cultivada a 30°C e a 220 rpm durante 18 horas. Adicionou-se Hemin (13  $\mu$ g/ml na concentração final) e galactose (2% na concentração final) e 10  $\mu$ l de 10 mM de  $\beta$ -amirina em solução de metanol foi ainda adicionada como substracto. De modo a fornecer oxigénio, a parte de cima do tubo Falcon foi selada com um tampão de algodão e isto foi depois cultivado durante 24 horas aseticamente e aerobicamente. As células foram recolhidas por centrifugação e transferidas para um frasco roscado de 2ml de capacidade. Adicionou-se uma porção de 250  $\mu$ l de uma solução aquosa de hidróxido de potássio a 40% e 250  $\mu$ l de metanol e a mistura foi vigorosamente agitada e tratada por calor a 100°C durante 5 minutos. Isto foi concentrado efectuando duas

vezes a extracção dos componentes solúveis em gordura utilizando 500  $\mu$ l de hexano. Foi efectuada a acetilação do extracto adicionando 20  $\mu$ l de piridina e anidrido acético e deixando a repousar de um dia para o outro. A reacção foi parada adicionando 20  $\mu$ l de uma solução aquosa de metanol a 50% e concentrou-se efectuando duas vezes uma extracção utilizando 200  $\mu$ l de hexano (1)). Como teste de controlo, as amostras foram preparadas pelo mesmo processo de 2) em casos em que foram utilizados transformantes preparados utilizando pESC-URA, 3) casos em que a  $\beta$ -amirina não foi adicionada como substrato, e 4) em casos em que a mesma quantidade de glucose foi adicionada em vez de galactose com o objectivo de inibir o promotor GAL 1 de pESC-CYP93E1. Cada um deles foi dissolvido em 10  $\mu$ l de hexano, e uma porção de 1  $\mu$ l foi submetida a uma análise por GC-MS na mesma condição que em (4) (as condições são as mesmas que o teste de (4)). O pico de 3 $\beta$ ,24-diacetóxi-12-oleaneno foi observado por TIM sob as condições de 1) (Fig. 3), e o padrão de fragmentação por MS (espectrometria de massa) coincidia com a preparação padrão (Fig. 4). Com base nos resultados da determinação, através da área do pico de TIM, pode-se considerar que vários  $\mu$ g de 3 $\beta$ ,24-dihidróxi-12-oleaneno são obtidos, quando cultivados utilizando 1 litro de meio nas mesmas condições do caso acima (adição de cerca de 2 mg de  $\beta$ -amirina).

#### (6) Construção do plasmídeo de expressão pESC-PSY

Utilizando um plasmídeo integrado com um gene PSY da  $\beta$ -amirina sintase derivado da ervilha (AB034802, Eur. J. Biochem., 267, 3543-3460, 2000) como molde, e a utilização de oligo fragmentos de ADN apresentados na SEQ ID N°s 4 e 5 que correspondem ao terminal N e terminal C do polipéptido

como iniciadores, PCR (30 ciclos, Ex Taq ADN polimerase fabricado por Takara Shuzo) foi efectuada a uma temperatura de emparelhamento de 58 °C para obter um fragmento PSY em que os locais *SalI* e *NheI* foram respectivamente introduzidos no terminal N e terminal C. O pESC-PSY foi preparado introduzindo isto nos locais *SalI* e *NheI* de pESC-URA e a actividade da  $\beta$ -amirina sintase foi confirmada, através de um processo convencional conhecido (Eur. J. Biochem., 267, 3543-3460, 2000).

(7) Construção do plasmídeo de expressão pESC-PSY-CYP93E1 e preparação da levedura transformada.

pESC-PSY e pESC-CYP93E1 foi digerida com *SalI* e *Clal*. Ligando o fragmento assim obtido contendo PSY e o fragmento contendo CYP93E1, foi construído um plasmídeo pESC-PSY-CYP93E1 que co-expressa PSY e CYP93E1. Foi obtido um transformante transferindo isto para uma estirpe de levedura GIL 77 utilizando levedura congelada EZ Transformação II.

(8) Teste de co-expressão de PSY e CYP93E1

A estirpe de levedura transformada foi inoculada num meio preparado suplementando 20 ml do meio SC-U contendo 2% de glucose como fonte de carbono com hemin (13  $\mu$ g/ml na concentração final), ergosterol (20  $\mu$ g/ml na concentração final) e Tween 80 (5 mg/ml na concentração final), e cultivada a 30°C e 220 rpm durante 1,5 dias. O meio foi substituído por um meio preparado adicionando hemin (13  $\mu$ g/ml na concentração final), ergosterol (20  $\mu$ g/ml na concentração final) e Tween 80 (5 mg/ml na concentração final) a 20 ml do meio SC-U contendo 2% de galactose como fonte de carbono, e depois a cultura foi continuada a 30°C

e 220 rpm durante 1 dia. As células foram transferidas para um tampão de fosfato de potássio 50 mM de pH 7,5, misturado com hemin (13 µg/ml na concentração final) e glucose (3% na concentração final) e ainda incubada a 30°C e a 220 rpm durante 1 dia. Uma amostra acetilada para análise por GC-MS foi preparada da mesma maneira que o método de teste em (4). Como processo de controlo, as amostras foram preparadas pelo mesmo processo num transformante preparado utilizando pESC-URA e um transformante preparado utilizando pESC-PSY e pESC-CYP93E1. Cada uma destas amostras foi dissolvida em 1000 µl de hexano, e uma porção de 1 µl foi sujeita a análise por GC-MS (as condições são as mesmas que no teste em (4)). Como indicado na Fig. 5, foi observado um pico correspondendo ao 3β,24-diacetóxi-12-oleanano por TIM, apenas quando foi utilizada a estirpe de levedura transformada com pESC-PSY-CYP93E1. Como resultado da análise determinativa pela proporção da área do pico pode-se considerar que várias centenas de µg de 3β,24-diacetóxi-12-oleanano são obtidas quando cultivadas utilizando 1 litro do meio sob esta condição.

(9) Cultura de GIL 77/pESC-PSY-CYP93E1 em grande escala  
(11)

A estirpe de levedura transformada foi inoculada num frasco cónico de 500 ml de capacidade carregado com 250 ml de um meio preparado adicionando hemin (13 µg/ml na concentração final), ergosterol (20 µg/ml na concentração final) e Tween 80 (5 mg/ml na concentração final) a 250 ml de meio SC-U contendo 2% de rafinose como fonte de carbono. Isto foi preparado em 4 frascos, e foi efectuada uma cultura com um volume total de 1 litro. Após cultura a 30°C e a 220 rpm durante 20 horas suplementou-se com galactose

(2% de concentração final) e continuou-se a cultura a 30°C e a 220 rpm durante 20 horas. As células completas foram transferidas para 100 ml de tampão de fosfato de potássio a 50 mM de pH 7,5 misturada com hemin (13 µg/ml na concentração final) e glucose (3% na concentração final) e incubou-se a 30°C a 220 rpm durante 1 dia.

(10) Isolamento do produto a partir de 1 litro de cultura de GIL 77/pESC-PSY-CYP93E1

As células foram recolhidas a partir da mistura de cultura obtida em (9), misturadas com 50 ml de solução aquosa de hidróxido de potássio a 40% e 50 ml de metanol e aquecidas em refluxo durante 1 hora. Extração da fracção solúvel em gordura foi efectuada utilizando 50 ml de hexano. A fracção do hexano foi exaustivamente lavada 3 vezes com 50 ml de solução aquosa saturada de bicarbonato de sódio. A extração foi efectuada repetindo a operação 3 vezes para obter 23 mg da fracção solúvel em gordura.

Esta foi purificada em dois passos em cromatografia em coluna flash em sílica gel. Em primeiro lugar, a anteriormente mencionada fracção solúvel em gordura foi dissolvida em benzeno e purificada utilizando sílica gel FC-40 (4 g, fabr. por Wako Pure Chemical Industries) e um sistema de solvente hexano-acetato de etilo. Posteriormente, a dita fracção foi purificada, utilizando sílica gel FC-40 (2 g) e um sistema de solvente de benzeno acetato de etilo para obter 0,55 mg de 3β,24-dihidróxi-12-oleaneno.

(11) Construção de um plasmídeo de expressão pESC-YUP43

Um plasmídeo integrado com um gene YUP43 de sintase de triterpeno multifuncional derivado de *Arabidopsis thaliana*

(Tetrahedron Lett., 41, 7705-7710, 2000) que disponibiliza 9 espécies de triterpeno incluindo  $\beta$ -amirina foi utilizado como molde, e utilizando os fragmentos de ADN apresentados nas SEQ ID N<sup>o</sup>s 6 e 7 que correspondem ao terminal N e terminal C do polipéptido como iniciadores, PCR (temperatura de emparelhamento 58°C, 30 ciclos, Ex Taq ADN polimerase fabr. por Takara Shuzo) foi efectuada para obter um fragmento YUP43 em que os locais *SalI* e *NheI* foram respectivamente introduzidos no terminal N e no terminal C. O pESC-YUP43 foi preparado introduzindo-o nos locais *SalI* e *NheI* de pESC-URA, e a actividade da sintase de triterpenos multifuncionais foi confirmada, através de um processo convencionalmente conhecido (Tetrahedron Lett., 41, 7705-7710, 2000).

(12) Construção do plasmídeo de expressão pESC-YUP43-CYP93E1 e preparação da levedura transformada

Os plasmídeos pESC-YUP43 e pESC-CYP93E1 foram digeridos com *SalI* e *ClaI*. Ligando o fragmento contendo YUP43 e o fragmento contendo CYP93E1 foi construído um plasmídeo pESC-YUP43-CYP93E1 que co-expressa YUP43 e CYP93E1. Foi obtido um transformante transferindo isto para uma estirpe de levedura GIL 77 utilizando Transformação II de levedura EZ congelada.

(13) Teste de co-expressão de polipéptidos de produtos de transcrição e de tradução de YUP43 e polinucleótido da SEQ ID N<sup>o</sup>:8

Através do mesmo processo de (8), foram preparadas amostras de transformantes obtidos respectivamente utilizando pESC-URA, pESC-YUP43, pESC-CYP93E1 e pESC-YUP43-CYP93E1. Cada um deles foi dissolvido em 1000  $\mu$ l de hexano, e uma sua porção

de 1 µl foi sujeita a uma análise por GC-MS (as condições são as mesmas que no teste de (4)). Como indicado na Fig. 7, foi observado um pico de 3β,24-diacetóxi-12-oleaneno por TIM, quando foi utilizada apenas a estirpe de levedura transformada com pESC-YUP43-CYP93E1. Uma vez que a quantidade produzida de β-amirina por YUP43 era mais baixo do que PSY, a quantidade de 3β,24-diacetóxi-12-oleaneno produzida era menor em comparação com o transformante pESC-PSY-CYP93E1. Por outro lado, os picos considerados por serem formas de hidroxilação de outros triterpenos (lupeol, butirospermol, tirucaladienol, taraxasterol, pseudotaraxasterol, baureelenol, α-amirina e multiflorenol) foram iguais ou inferiores ao limite de detecção. Enquanto que a invenção foi descrita em detalhe e fazendo referência às suas concretizações específicas será aparente a um perito no estado da técnica que várias alterações e modificações podem aí ser feitas sem se afastarem do espírito e âmbito da invenção.

O pedido de patente de invenção encontra-se baseado num pedido de patente de invenção japonesa registado em 25 de Fevereiro de 2004 (pedido de patente de invenção N° 2004-049123),

#### Aplicabilidade Industrial

Através da invenção torna-se possível manipular através de uma técnica de engenharia genética uma enzima que efectua a hidroxilação na posição 24 de um triterpeno do tipo oleanano. Neste contexto, a utilização de uma célula, no qual o dito gene de hidroxilase foi integrado torna possível a sua utilização, por exemplo, na produção da hidroxilase, através de uma estirpe de levedura, aplicação de uma reacção de hidroxilação, produção microbiana de um triterpeno de uma planta e semelhantes. Além disso,

integrando o gene da hidroxilase numa planta torna-se possível aplicá-la no campo da agricultura, tal como o aumento da produção de sapogenol de soja ou triterpenos semelhantes.

#### Breve Descrição das Figuras

A Fig. 1 apresenta a actividade de hidroxilação na posição 24 da  $\beta$ -amirina (*in vitro*) do produto de tradução da SEQ ID N° 8. De forma mais ilustrativa mostra um cromatograma de espectrometria de massa que monitoriza  $m/Z=218$  que foi analisado por GC-MS após acetilação do produto. A a F apresenta os seguintes resultados. A: amostra padrão de  $3\beta,24$ -diacetóxi-12-oleaneno. B: um produto obtido deixando um líquido de enzima bruta preparado a partir de uma estirpe de levedura transformada com pESC-CYP93E1 reagir com  $\beta$ -amirina na coexistência de um sistema de regeneração da NADPH. C: um produto da reacção B efectuada removendo a  $\beta$ -amirina do sistema reaccional. D: um produto da reacção B efectuada utilizando líquido de enzima em bruto desnaturado pelo calor. E: estirpe de levedura transformada com pESC-CYP93E1 foi cultivada adicionando glucose, as outras condições são as mesmas que em B. F: um líquido de enzima em bruto preparado a partir de uma estirpe de levedura transformada com um plasmídeo vazio pESC-URA, as outras condições são as mesmas que em B.

A Fig. 2 apresenta a actividade de hidroxilação na posição 24 do soforodiol (*in vitro*) do produto de tradução da SEQ ID N°8. De forma mais ilustrativa apresenta um cromatograma de espectrometria de massa que monitoriza  $m/Z=216$  que foi analisado por GC-MS após acetilação do produto. A a F apresenta os resultados seguintes. A: amostra padrão de triacetil sapogenol B de soja. B: um produto obtido deixando um líquido de enzima em bruto preparado a partir

de uma estirpe de levedura transformada com pESC-CYP93E1 reagir com soforodiol na coexistência de um sistema de regeneração da NADPH. C: um produto da reacção B efectuada removendo soforodiol do sistema reaccional. D: um produto da reacção B efectuado utilizando líquido de enzima em bruto desnaturado pelo calor. E: estirpe de levedura transformada com pESC-CYP93E1 foi cultivado adicionando glucose, outras condições são as mesmas que em B. F: foi utilizado líquido de enzima em bruto preparado a partir de uma estirpe de levedura transformada com um plasmídeo vazio pESC-URA, as outras condições são as mesmas que em B.

A Fig. 3 apresenta a actividade de hidroxilação da posição 24 da  $\beta$ -amirina (*in vivo*) do produto de tradução da SEQ ID N°8. De forma mais ilustrativa mostra um cromatograma de GC-MS TIM após acetilação do produto. A a E apresenta os resultados seguintes. A: amostra padrão de 3 $\beta$ ,24-diacetóxi-12-oleaneno (20 pmol). B: um produto obtido adicionando  $\beta$ -amirina a uma estirpe de levedura transformada com pESC-CYP93E1. C: um caso em que a  $\beta$ -amirina não foi adicionada, as outras condições são as mesmas que em B. D: a levedura transformada foi cultivada adicionando glucose, outras condições são as mesmas que em B. E: a  $\beta$ -amirina foi adicionada a uma estirpe de levedura transformada com um plasmídeo vazio pESC-URA, as outras condições são as mesmas que em B.

A Fig. 4 apresenta a actividade de hidroxilação da posição 24 da  $\beta$ -amirina (*in vivo*) do produto de tradução da SEQ ID N°8. A e B apresentam os resultados seguintes. A: o espectro de massa do pico a um tempo de retenção de 15,35 minutos detectado em B da Fig. 3. B: espectro de massa de A da Fig.3 (amostra padrão 3 $\beta$ ,24-diacetóxi-12-oleaneno).

A Fig. 5 apresenta a produção de 3 $\beta$ ,24-diacetóxi-12-oleaneno por co-expressão de CYP93E1 e sintase de  $\beta$ -amirina

(PSY). De forma mais ilustrativa apresenta um cromatograma por TIM, através de análise por GC-MS após acetilação da fracção solúvel em gordura obtida a partir da respectiva levedura transformada (GIL 77). A a E apresentam os resultados seguintes. A: amostra padrão de  $3\beta,24$ -diacetóxi-12-oleaneno (20 pmol). B: um extracto de estirpe de levedura transformada com pESC-PSY-CYP93E1. C: um extracto de estirpe de levedura transformada com pESC-PSY. D: um extracto de estirpe de levedura transformada com pESC-CYP93E1. E: um extracto de estirpe de levedura transformado com pESC-URA.

A Fig. 6 apresenta um RMN  $H^1$  do produto obtido a partir de 1 litro de meio de cultura de GIL 77/pESC-PSY-CYP93E1.

A Fig. 7 apresenta a produção de  $3\beta,24$ -diacetóxi-12-oleaneno por co-expressão do produto de tradução da SEQ ID N°:8 e uma sintase de triterpeno multifuncional (YUP43). Mais ilustrativamente apresenta um cromatograma por TIM, através de análise por GC-MS após acetilação da fracção solúvel em gordura obtida a partir da respectiva levedura transformada (GIL 77). A a E apresenta os resultados seguintes. A: amostra padrão de  $3\beta,24$ -diacetóxi-12-oleaneno (20 pmol). B: um extracto de estirpe de levedura transformado com pESC-YUP43-CYP93E1. C: um extracto de estirpe de levedura transformado com pESC-YUP43. D: um extracto de estirpe de levedura transformado com pESC-CYP93E1. E: um extracto de estirpe de levedura transformado com pESC-Ura.

### Lista de Sequências

<110> Meiji Seika Kaisha LTD.  
 <130> Triterpene hydroxylase  
 <130> P05160903  
 <150> JP 2004-049123  
 <151> 2004-02-25

<160> 11

<170> Patentn version 3.1

<210> 1

<211> 32

<212> DNA

<213> Artificial

<220>

<223> primer for cDNA

<400> 1

gactcagctc gacaacgati tttttttt tt 32

<210> 2

<211> 31

<212> DNA

<213> Artificial

<220>

<223> Forward primer sequence for CYP93E1

<400> 2

aaacacatagt atgctagaca taaaaggcta c 31

<210> 3

<211> 31

<212> DNA

<213> Artificial

<220>

<223> Reverse primer sequence for CYP93E1

<400> 3

ttaacatgat tcaggcagcg aacggagctg a 31

<210> 4

<211> 31

<212> DNA

<213> Artificial

<220>

<223> Forward primer for PSY

<400> 4

ctctctctgac aagatgtgga ggtgacagat a 31

<210> 5

<211> 31

<212> DNA

<213> Artificial

<220>

<223> Reverse primer for PSY

<400> 5

gtctctctgac tcaaggcaaa ggaactcttc t 31

<210> 8  
 <211> 31  
 <212> DNA  
 <213> Artificial

<220>  
 <223> Forward primer for YUP43

<220>  
 <221> misc\_feature  
 <223> Forward primer for YUP43

<400> 8  
 taaggtagac atatgtgga agtgaagat a 31

<210> 7  
 <211> 31  
 <212> DNA  
 <213> Artificial

<220>  
 <223> Reverse primer for YUP43

<400> 7  
 taaggtagac ctaagatct tgaatgag c 31

<210> 8  
 <211> 1542  
 <212> DNA  
 <213> Soybean

<220>  
 <221> CDS  
 <222> (1)..(1542)  
 <223>

<400> 8

atg cta gac atc aaa ggc tac ctc gta ctc ttc ttc cta tgg ttc ata 48  
 Met Leu Asp Ile Lys Gly Tyr Leu Val Leu Phe Phe Leu Trp Phe Ile  
 1 5 10 15

tca acc att ctg ata cgt tcc atc ttc aag aaa cca cag cgt cta aga 96  
 Ser Thr Ile Leu Ile Arg Ser Ile Phe Lys Lys Pro Gln Arg Leu Arg  
 20 25 30

ctc cca ccg ggt cct cca act tca gta ccc ttg ctg gga cac ggc cca	184
Leu Pro Pro Gly Pro Pro Ile Ser Val Pro Leu Leu Gly His Ala Pro	
35 40 45	
cat ctc cgt tca ctg etc cac caa gcc ttg tac aag cta tca ctg cgc	192
Tyr Leu Arg Ser Leu Leu His Gln Ala Leu Tyr Lys Leu Ser Leu Arg	
50 55 60	
tat gga ccc ttg etc cac gtc atg etc ggt tgg aag cac ctg gtc gtc	240
Tyr Gly Pro Leu Ile His Val Met Ile Gly Ser Lys His Val Val Val	
65 70 75 80	
gag tgg tgg gag gag acg gcc aag cag ata ctc aaa cac tgg gag gag	288
Ala Ser Ser Ala Glu Thr Ala Lys Gln Ile Leu Lys Thr Ser Glu Glu	
85 90 95	
gaa ttc tgc aac cgt ccc tta atg ata gag agc gag agc cta acc tac	336
Ala Phe Cys Asn Arg Pro Leu Met Ile Ala Ser Glu Ser Leu Thr Tyr	
100 105 110	
ggc gag ggc gar tac ttc ttc atc ccc tac ggc acc tac tgg cgg ttc	384
Gly Ala Ala Asp Tyr Phe Phe Ile Pro Tyr Gly Thr Tyr Trp Arg Phe	
115 120 125	
ctg aag aag ctc tgc atg acg gag ctt ctg agc ggg aag acc ctg gag	432
Leu Lys Lys Leu Cys Met Thr Gln Leu Leu Ser Gly Lys Thr Leu Glu	
130 135 140	
cat ttc gtc aga etc cgc gag agc gag gtc gag ggc ttc ctg aag aga	480
His Phe Val Arg Ile Arg Glu Ser Glu Val Glu Ala Phe Leu Lys Arg	
145 150 155 160	
atg atg gag att tca ggc aat gga aat tac gag gtc gtc atg agg aag	528
Met Met Glu Ile Ser Gly Asn Gly Asn Tyr Glu Val Val Met Arg Lys	
165 170 175	
gag ctg ata aag cac acg aat aac atc atc acg agg atg ata atg ggg	576
Glu Leu Ile Thr His Thr Asn Asn Ile Ile Thr Arg Met Ile Met Gly	
180 185 190	
aag aag agt aat gcc gaa aac gat gag gtc gcc agg ttg aag aag gtc	624
Lys Lys Ser Asn Ala Glu Asn Asp Glu Val Ala Arg Leu Arg Lys Val	
195 200 205	
gtg agg gag gtc ggg gag ttg ctt ggg gcy tta aac ctg ggg gat gtc	672
Val Arg Glu Val Gly Glu Leu Leu Gly Ala Phe Asn Leu Gly Asp Val	
210 215 220	
att ggg ttc atg agg cct ctg gat ctg caa ggg ttt ggg aag aag aac	720
Ile Gly Phe Met Arg Pro Leu Asp Leu Gln Gly Phe Gly Lys Lys Asn	
225 230 235 240	
atg gaa act cac cac aag gtc gat gcg atg atg gcy aag gtc ttg agg	768
Met Glu Thr His His Lys Val Asp Ala Met Met Glu Lys Val Leu Arg	
245 250 255	
gag cat gag gag gct agg gct aag gaa gat gct gac tct gat agg aag	816
Glu His Glu Glu Ala Arg Ala Lys Glu Asp Ala Asp Ser Asp Arg Lys	
260 265 270	

aag gat ctt ttt gat att ttg ttg aac ctc att gaa gct gat ggt gct	884
Lys Asp Leu Phe Asp Ile Leu Leu Asn Leu Ile Glu Ala Asp Gly Ala	
275 288 285	
gat aat aag ctc act aga gag agt gcc aac gcc ttt gct ctg aac atg	912
Asp Asn Lys Leu Thr Arg Gln Ser Ala Lys Ala Phe Ala Leu Asp Met	
290 295 300	
ttc ctg gcc ggc aca aac ggc acc gaa agc gtc cta gag tgg aca ctg	940
Phe Ile Ala Gly Thr Asn Gly Pro Ala Ser Val Leu Glu Trp Ser Leu	
305 310 315 320	
ggg gag ctg gtg aga aac acc acc gtt ttc aag aag gca aga gaa gag	1008
Ala Glu Leu Val Arg Asn Pro His Val Phe Lys Lys Ala Arg Glu Glu	
325 330 335 335	
act gag aca gtg gta ggc aaa gaa agg ctg gtc aac aca tca gac att	1056
Ile Glu Ser Val Val Gly Lys Glu Arg Leu Val Lys Glu Ser Asp Ile	
340 345 350	
ccc aac cta aca tac cta caa gca ttg ccg aag gaa acc cta agg ctg	1104
Pro Asn Leu Pro Tyr Leu Glu Ala Leu Leu Lys Glu Thr Leu Arg Leu	
355 360 365	
cac ccg cca aac cca aca ttc gaa aga gaa gcc atg cca aca tgc cag	1152
His Pro Pro Thr Pro Ile Phe Ala Arg Glu Ala Met Arg Thr Cys Glu	
370 375 380	
ggt gaa ggc tac gat att ccg gaa aat tcc act att ttg atc ayc aca	1200
Val Glu Gly Tyr Asp Ile Pro Glu Asn Ser Thr Ile Leu Ile Ser Thr	
385 390 395 400	
tgg gcc att ggt agg gat cca aat tac tgg gat gac gca ctg gag tac	1248
Trp Ala Ile Gly Arg Asp Pro Asn Tyr Trp Asp Asp Ala Leu Glu Tyr	
405 410 415	
aag ccg gag agg ttc ttg ttc tcc gat gac ccg gcc aag agc aag att	1296
Lys Pro Glu Arg Phe Leu Phe Ser Asp Asp Pro Gly Lys Ser Lys Ile	
420 425 430	
gac gtc agg ggg cag tac tat aag ctc ctg acc ttt ggg acc ggg aga	1344
Asp Val Arg Gly Glu Tyr Tyr Glu Leu Leu Pro Phe Gly Ser Gly Arg	
435 440 445	
aga agc tgc acc gga gac tgg cta gaa ttg ctt gtc atg caa gca aag	1392
Arg Ser Cys Pro Gly Ala Ser Leu Ala Leu Leu Val Met Glu Ala Thr	
450 455 460	
cta gcg agt ttg atc aag tgc ttc gac tgg atc gtt aat gat ggt aaa	1440
Leu Ala Ser Leu Ile Glu Cys Phe Asp Trp Ile Val Asn Asp Gly Lys	
465 470 475 480	
aac cat cat gtt gac atg tct gag gaa ggg agg gtc act gtc tcc ttg	1488
Asn His His Val Asp Met Ser Glu Glu Gly Arg Val Thr Val Phe Leu	
485 490 495	
gcc aag aca ctg aag tgc aag cat gtt ccg agt ttc act ccg ttc gct	1536
Ala Lys Pro Leu Lys Cys Lys Pro Val Pro Arg Phe Thr Pro Phe Ala	
500 505 510	
gcc tgg	1542

Ala

<210>0  
 <211>513  
 <212>PRT  
 <213>Soybean  
 <400>0

Met Leu Asp Ile Lys Gly Tyr Leu Val Leu Phe Phe Leu Trp Phe Ile  
 1 5 10 15

Ser Thr Ile Leu Ile Arg Ser Ile Phe Lys Lys Pro Glu Arg Leu Arg  
 20 25 30

Leu Pro Pro Gly Pro Pro Ile Ser Val Pro Leu Leu Gly His Ala Pro  
 35 40 45

Tyr Leu Arg Ser Leu Leu His Glu Ala Leu Tyr Lys Leu Ser Leu Arg  
 50 55 60

Tyr Gly Pro Leu Ile His Val Met Ile Gly Ser Lys His Val Val Val  
 65 70 75 80

Ala Ser Ser Ala Glu Thr Ala Lys Glu Ile Leu Lys Thr Ser Glu Glu  
 85 90 95

Ala Phe Cys Asn Arg Pro Leu Met Ile Ala Ser Glu Ser Leu Thr Tyr  
 100 105 110

Gly Ala Ala Asp Tyr Phe Phe Ile Pro Tyr Gly Thr Tyr Trp Arg Phe  
 115 120 125

Leu Lys Lys Leu Cys Met Thr Glu Leu Leu Ser Gly Lys Thr Leu Glu  
 130 135 140

His Phe Val Arg Ile Arg Glu Ser Glu Val Glu Ala Phe Leu Lys Arg  
 145 150 155 160

Met Met Glu Ile Ser Gly Asn Gly Asn Tyr Glu Val Val Met Arg Lys  
 165 170 175

Glu Leu Ile Thr His Thr Asn Asn Ile Ile Thr Arg Met Ile Met Gly  
 180 185 190

Lys Lys Ser Asn Ala Glu Asn Asp Glu Val Ala Arg Leu Arg Lys Val  
 195 200 205

Val Arg Glu Val Gly Glu Leu Leu Gly Ala Phe Asn Leu Gly Asp Val  
 210 215 220

Ile Gly Phe Met Arg Pro Leu Asp Leu Gln Gly Phe Gly Lys Lys Asn  
 225 230 235 240

Met Glu Thr His His Lys Val Asp Ala Met Met Glu Lys Val Leu Arg  
 245 250 255

Glu His Glu Glu Ala Arg Ala Lys Glu Asp Ala Asp Ser Asp Arg Lys  
 260 265 270

Lys Asp Leu Phe Asp Ile Leu Leu Asn Leu Ile Glu Ala Asp Gly Ala  
 275 280 285

Asp Asn Lys Leu Thr Arg Glu Ser Ala Lys Ala Phe Ala Leu Asp Met  
 290 295 300

Phe Ile Ala Gly Thr Asn Gly Pro Ala Ser Val Leu Glu Trp Ser Leu  
 305 310 315 320

Ala Glu Leu Val Arg Asn Pro His Val Phe Lys Lys Ala Arg Glu Glu  
 325 330 335

Ile Glu Ser Val Val Gly Lys Glu Arg Leu Val Lys Glu Ser Asp Ile  
 340 345 350

Pro Asn Leu Pro Tyr Leu Gln Ala Leu Leu Lys Glu Thr Leu Arg Leu  
 355 360 365

His Pro Pro Thr Pro Ile Phe Ala Arg Glu Ala Met Arg Thr Cys Glu  
 370 375 380

Val Glu Gly Tyr Asp Ile Pro Glu Asn Ser Thr Ile Leu Ile Ser Thr  
 385 390 395 400

Trp Ala Ile Gly Arg Asp Pro Asn Tyr Trp Asp Asp Ala Leu Glu Tyr  
 405 410 415

Lys Pro Glu Arg Phe Leu Phe Ser Asp Asp Pro Gly Lys Ser Lys Ile  
 420 425 430

Asp Val Arg Gly Glu Tyr Tyr Gln Leu Leu Pro Phe Gly Ser Gly Arg  
 435 440 445

Arg Ser Cys Pro Gly Ala Ser Leu Ala Leu Leu Val Met Gln Ala Thr  
 450 455 460

Leu Ala Ser Leu Ile Gln Cys Phe Asp Trp Ile Val Asn Asp Gly Lys  
 465 470 475 480

Asn His His Val Asp Met Ser Gln Gln Gly Arg Val Thr Val Phe Leu  
 485 490 495

Ala Lys Pro Leu Lys Cys Lys Pro Val Pro Arg Phe Thr Pro Phe Ala  
 500 505 510

Ala

<10> 10  
 <11> 1542  
 <12> DNA  
 <13> Soybean

<20>  
 <21> CDS  
 <22> (1), (1542)  
 <23>

<400> 10

atg cta gac atc aaa ggc tac ctc gta ctc ttc ttc cta tgg ttc ata	48
Met Leu Asp Ile Lys Gly Tyr Leu Val Leu Phe Phe Leu Trp Phe Ile	
1 5 10 15	
tca acc att ctg ata cgt tcc acc ttc sag aas cca cag cgt cca aga	96
Ser Thr Ile Leu Ile Arg Ser Ile Phe Lys Lys Pro Gln Arg Leu Arg	
20 25 30	
ctc aca ccg ggt cct cca att tca ata ccc ttg ctg gaa ccc gcg cca	144
Leu Pro Pro Gly Phe Phe Ile Ser Ile Pro Leu Leu Gly His Ala Pro	
35 40 45	
tat ctc cgt tca ctg ctg ccc caa gca ttg tac aag cta tca ctg cgc	192
Tyr Leu Arg Ser Leu Leu His Gln Ala Leu Tyr Lys Leu Ser Leu Arg	
50 55 60	
tat gga ccc ttg atc ccc gtc atg atc ggt tgg aag cac gta gta gta	240
Tyr Gly Pro Leu Ile His Val Met Ile Gly Ser Lys His Val Val Val	
65 70 75 80	
gcg tcg tcg gcg gag ccg gcc aag cag atc ctg aas acc tcg gag gag	288
Ala Ser Ser Ala Gln Thr Ala Lys Gln Ile Leu Lys Thr Ser Gln Gln	
85 90 95	
gca ttc tgc aac cgt ccc tta atg ata gag aac gag aac cta acc tac	336
Ala Phe Cys Asn Arg Phe Leu Met Ile Ala Ser Gln Ser Leu Thr Tyr	
100 105 110	
ggc gcg gcg gca tcc ttc ttc atc ccc tac ggc acc tac tgg cgg ttc	384

Gly	Ala	Ala	Asp	Tyr	Phe	Phe	Ile	Pro	Tyr	Gly	Thr	Tyr	Trp	Arg	Phe		
		115					120						125				
ctg	aag	aag	ctc	tgc	atg	acg	gag	ctt	ctg	agc	ggg	aag	acc	ctg	gag		432
Leu	Lys	Lys	Leu	Cys	Met	Thr	Glu	Leu	Leu	Ser	Gly	Lys	Thr	Leu	Glu		
		130				135						140					
cat	tcc	gtg	aga	atc	agc	gag	acc	gag	gtc	gag	gag	ttc	ctc	aag	aga		480
His	Phe	Val	Arg	Ile	Arg	Glu	Ser	Glu	Val	Glu	Ala	Phe	Leu	Lys	Arg		
					150						155				160		
atg	atg	ggg	att	tca	ggc	aat	gaa	aat	tac	gag	gtg	gtg	atg	agg	aag		528
Met	Met	Glu	Ile	Ser	Gly	Asn	Gly	Asn	Tyr	Glu	Val	Val	Met	Arg	Lys		
				165					170						175		
gag	ctc	ata	acg	cac	acg	aat	aac	atc	atc	acg	agg	atg	ata	atg	ggg		576
Glu	Leu	Ile	Thr	His	Thr	Asn	Asn	Ile	Ile	Thr	Arg	Met	Ile	Met	Gly		
		180						185						190			
aag	aag	agt	aat	ggc	gaa	aac	gat	gag	gag	gac	agg	ttg	agg	aag	gtg		624
Lys	Lys	Ser	Asn	Ala	Glu	Asn	Asp	Glu	Val	Ala	Arg	Leu	Arg	Lys	Val		
		195					200					205					
gtg	agg	gag	gtc	ggg	gag	tgg	ctt	ggg	ggg	ttt	aac	tgg	ggg	gat	glt		672
Val	Arg	Glu	Val	Gly	Glu	Leu	Leu	Gly	Ala	Phe	Asn	Leu	Gly	Asp	Val		
		210				215						220					
att	ggg	ttc	atg	agg	cat	ttg	gat	ctg	aaa	ggg	ttt	ggg	aag	aag	aac		720
Ile	Gly	Phe	Met	Arg	Pro	Leu	Asp	Leu	Gln	Gly	Phe	Gly	Lys	Lys	Asn		
					230						235				240		
atg	gaa	act	cac	cac	aag	gtg	gat	ggc	atg	atg	gag	aag	gtg	ttg	agg		768
Met	Glu	Thr	His	His	Lys	Val	Asp	Ala	Met	Met	Glu	Lys	Val	Leu	Arg		
				245				250						255			
gag	cat	gag	gag	gct	agg	gct	aag	gaa	gat	gct	gac	ctt	gat	agg	agg		816
Glu	His	Glu	Glu	Ala	Arg	Ala	Lys	Glu	Asp	Ala	Asp	Ser	Asp	Arg	Lys		
			260					265						270			
aag	gat	ctt	ttt	gat	att	ttg	ttg	aac	ctc	att	gaa	gct	gat	ggt	gct		864
Lys	Asp	Leu	Phe	Asp	Ile	Leu	Leu	Asn	Leu	Ile	Glu	Ala	Asp	Gly	Ala		
		275					280						285				
gac	aat	aag	ctc	act	aga	gag	agt	gtc	aaa	gtc	ttt	gct	ctg	gac	ctg		912
Asp	Asn	Lys	Leu	Thr	Arg	Glu	Ser	Ala	Lys	Ala	Phe	Ala	Leu	Asp	Met		
		290				295					300						
ttc	atc	ggc	ggc	aca	aac	ggc	ctc	gca	agc	gtc	ata	gag	tgg	tca	ctg		960
Phe	Ile	Ala	Gly	Thr	Asn	Gly	Pro	Ala	Ser	Val	Leu	Glu	Trp	Ser	Leu		
				310							315				320		
ggg	gag	ctg	gtg	aga	aac	ccc	cac	gtt	ttc	aag	aag	gca	aga	gaa	ggg		1008
Ala	Glu	Leu	Val	Arg	Asn	Pro	His	Val	Phe	Lys	Lys	Ala	Arg	Glu	Glu		
				325						330					335		
att	gag	tca	gtg	gta	ggc	aaa	gaa	agg	ctg	gtc	aaa	gaa	tca	gat	att		1056
Ile	Glu	Ser	Val	Val	Gly	Lys	Glu	Arg	Leu	Val	Lys	Glu	Ser	Asp	Ile		
			340					345						350			
ccc	aac	cta	cca	tac	cta	caa	gaa	gtg	ctg	aag	gaa	acc	cta	agg	ctg		1104
Pro	Asn	Leu	Pro	Tyr	Leu	Gln	Ala	Val	Leu	Lys	Glu	Thr	Leu	Arg	Leu		

355	360	365	
cac ccg cca acc cca cca tcc gca aga gaa gcc atg cga acc tgc cag			3152
His Pro Pro Thr Pro Ile Phe Ala Arg Glu Ala Met Arg Thr Cys Glu			
370	375	380	
ggt gaa ggc tac gac att ccg gaa aat tcc act att ttg atc agc aca			3200
Val Glu Gly Tyr Asp Ile Pro Glu Asn Ser Thr Ile Leu Ile Ser Thr			
385	390	395	400
tgg gcc att ggt agg gat cca aat tac tgg gat gaa cca ctc gag tac			3248
Trp Ala Ile Gly Arg Asp Pro Asn Tyr Trp Asp Asp Ala Leu Glu Tyr			
	405	410	415
gag ccg gag agg ttc ttg ttc tcc gac gac ccg ggc aag agc aag att			3296
Lys Pro Glu Arg Phe Leu Phe Ser Asp Asp Pro Gly Lys Ser Lys Ile			
	420	425	430
gac ggg agg ggg cag tac tat ccg ctc ctg ccc ttt ggg agc ggg aga			3344
Asp Val Arg Gly Glu Tyr Tyr Glu Leu Leu Pro Phe Gly Ser Gly Arg			
	435	440	445
aga agc tgc ccc gga gcc tcc cta ccg ttg ctt gtc atg caa gca acg			3392
Arg Ser Cys Pro Gly Ala Ser Leu Ala Leu Leu Val Met Glu Ala Thr			
	450	455	460
cta ccg agt ttg atc cag tgc ttc gac tgg atc gtt aat gat ggt aas			3440
Leu Ala Ser Leu Ile Glu Cys Phe Asp Trp Ile Val Asn Asp Gly Lys			
	465	470	475
aac cat cat gtt gac atg tct gag gaa ggg agg gtg act gtg ttt ttg			3488
Asn His His Val Asp Met Ser Glu Glu Gly Arg Val Thr Val Phe Leu			
	485	490	495
gcc aag cca ctc aag tgc aag cct gtt ccg agt ttc act ccg ttc gct			3536
Ala Lys Pro Leu Lys Cys Lys Pro Val Pro Arg Phe Thr Pro Phe Ala			
	500	505	510
gac tga			3584
Ala			

<210> 11  
 <211> 613  
 <212> PRT  
 <213> Soybean

<400> 11

Met Leu Asp Ile Lys Gly Tyr Leu Val Leu Phe Phe Leu Trp Phe Ile  
 1 5 10 15

Ser Thr Ile Leu Ile Arg Ser Ile Phe Lys Lys Pro Glu Arg Leu Arg  
 20 25 30

Leu Pro Pro Gly Pro Pro Ile Ser Ile Pro Leu Leu Gly His Ala Pro  
 35 40 45

Tyr Leu Arg Ser Leu Leu His Glu Ala Leu Tyr Lys Leu Ser Leu Arg  
 59 55 60

Tyr Gly Pro Leu Ile His Val Met Ile Gly Ser Lys His Val Val Val  
 65 70 75 80

Ala Ser Ser Ala Glu Thr Ala Lys Glu Ile Leu Lys Thr Ser Glu Glu  
 85 90 95

Ala Phe Cys Asn Arg Pro Leu Met Ile Ala Ser Glu Ser Leu Thr Tyr  
 100 105 110

Gly Ala Ala Asp Tyr Phe Phe Ile Pro Tyr Gly Thr Tyr Trp Arg Phe  
 115 120 125

Leu Lys Lys Leu Cys Met Thr Glu Leu Leu Ser Gly Lys Thr Leu Glu  
 130 135 140

His Phe Val Arg Ile Arg Glu Ser Glu Val Glu Ala Phe Leu Lys Arg  
 145 150 155 160

Met Met Glu Ile Ser Gly Asn Gly Asn Tyr Glu Val Val Met Arg Lys  
 165 170 175

Glu Leu Ile Thr His Thr Asn Asn Ile Ile Thr Arg Met Ile Met Gly  
 180 185 190

Lys Lys Ser Asn Ala Glu Asn Asp Glu Val Ala Arg Leu Arg Lys Val  
 195 200 205

Val Arg Glu Val Gly Glu Leu Leu Gly Ala Phe Asn Leu Gly Asp Val  
 210 215 220

Ile Gly Phe Met Arg Pro Leu Asp Leu Glu Gly Phe Gly Lys Lys Asn  
 225 230 235 240

Met Glu Thr His His Lys Val Asp Ala Met Met Glu Lys Val Leu Arg  
 245 250 255

Glu His Glu Glu Ala Arg Ala Lys Glu Asp Ala Asp Ser Asp Arg Lys  
 260 265 270

Lys Asp Leu Phe Asp Ile Leu Leu Asn Leu Ile Glu Ala Asp Gly Ala  
 275 280 285

Asp Asn Lys Leu Thr Arg Glu Ser Ala Lys Ala Phe Ala Leu Asp Met  
 280 295 300

Phe Ile Ala Gly Thr Asn Gly Pro Ala Ser Val Leu Glu Trp Ser Leu  
 305 310 315 320

Ala Glu Leu Val Arg Asn Pro His Val Phe Lys Lys Ala Arg Glu Glu  
 325 330 335

Ile Glu Ser Val Val Gly Lys Glu Arg Leu Val Lys Glu Ser Asp Ile  
 340 345 350

Pro Asn Leu Pro Tyr Leu Glu Ala Val Leu Lys Glu Thr Leu Arg Leu  
 355 360 365

His Pro Pro Thr Pro Ile Phe Ala Arg Glu Ala Met Arg Thr Cys Glu  
 370 375 380

Val Glu Gly Tyr Asp Ile Pro Glu Asn Ser Thr Ile Leu Ile Ser Thr  
 385 390 395 400

Trp Ala Ile Gly Arg Asp Pro Asn Tyr Trp Asp Asp Ala Leu Glu Tyr  
 405 410 415

Lys Pro Glu Arg Phe Leu Phe Ser Asp Asp Pro Gly Lys Ser Lys Ile  
 420 425 430

Asp Val Arg Gly Glu Tyr Tyr Glu Leu Leu Pro Phe Gly Ser Gly Arg  
 435 440 445

Arg Ser Cys Pro Gly Ala Ser Leu Ala Leu Leu Val Met Glu Ala Thr  
 450 455 460

Leu Ala Ser Leu Ile Glu Cys Phe Asp Trp Ile Val Asn Asp Gly Lys  
 465 470 475 480

Asn His His Val Asp Met Ser Glu Glu Gly Arg Val Thr Val Phe Leu  
 485 490 495

Ala Lys Pro Leu Lys Cys Lys Pro Val Pro Arg Phe Thr Pro Phe Ala  
 500 505 510

Ala

Lisboa, 22 de Fevereiro de 2011.

## **Reivindicações**

- 1.** Processo para produzir um triterpeno do tipo oleano em que a posição 24 é hidroxilada compreendendo:  
cultivar um transformante compreendendo um vector de expressão que possui um polinucleótido que hibridiza com uma cadeia complementar do polinucleótido representado pela SEQ ID N°:8 em condições severas e que também codifica para um polipéptido que possui a actividade de hidroxilar a posição 24 de um triterpeno do tipo oleano de modo a produzir um polipéptido da SEQ ID N°9; e deixar o transformante actuar num triterpeno do tipo oleano, em que o transformante é um microrganismo.
  
- 2.** Processo de acordo com a reivindicação 1, em que o polinucleótido é o polinucleótido representado pela SEQ ID N°:8.
  
- 3.** Processo de acordo com a reivindicação 1, em que o transformante é uma levedura.

Lisboa, 22 de Fevereiro de 2011.

## Resumo

### "Hidroxilase de triterpeno"

O sapogenol B de soja é biosintetizado, através de dois passos da reacção de hidroxilação do seu precursor a  $\beta$ -amirina. Contudo, o gene da hidroxilase envolvido nesta reacção não tinha sido revelado. Por isso, era impossível aplicar uma técnica de engenharia genética à hidroxilase. Os presentes inventores revelam que uma sequência que corresponde ao gene CYP93E1 do citocromo P-450 derivado da soja codifica para uma proteína enzima que efectua a hidroxilação na posição 24 de um triterpeno do tipo oleanano, e fornece também um processo de aplicação do dito gene utilizando uma técnica de engenharia genética.

FIG. 1

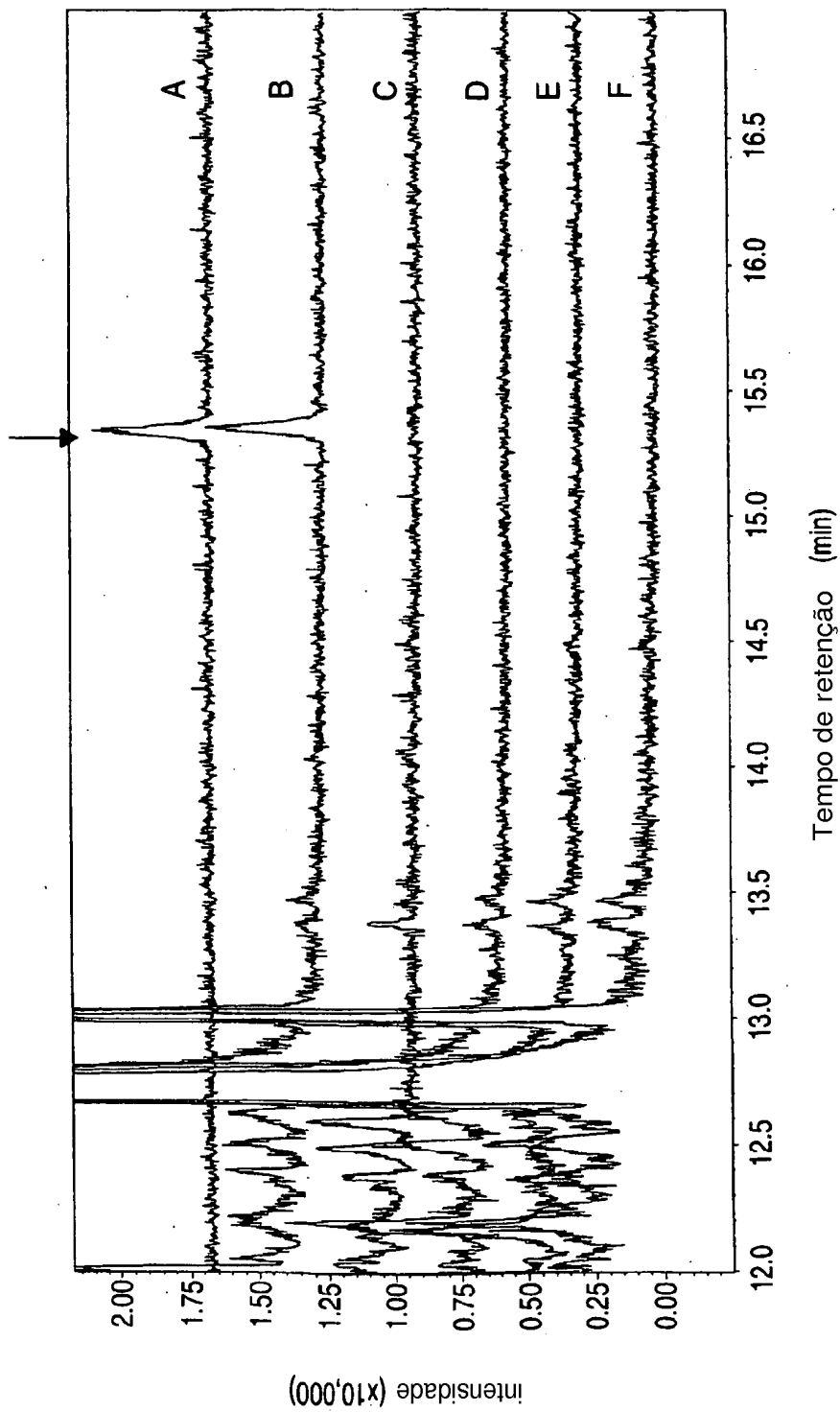


FIG. 2

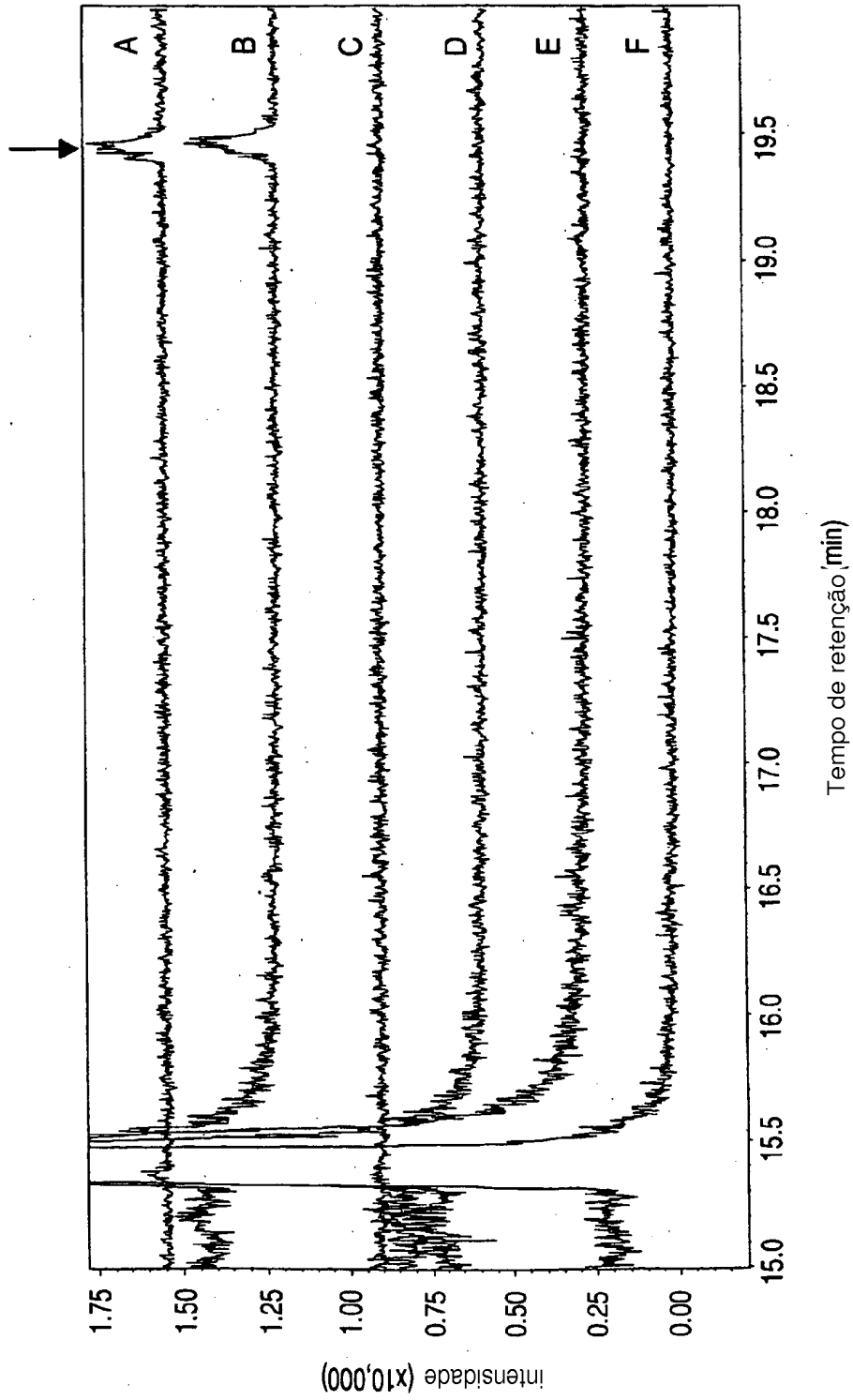


FIG. 3

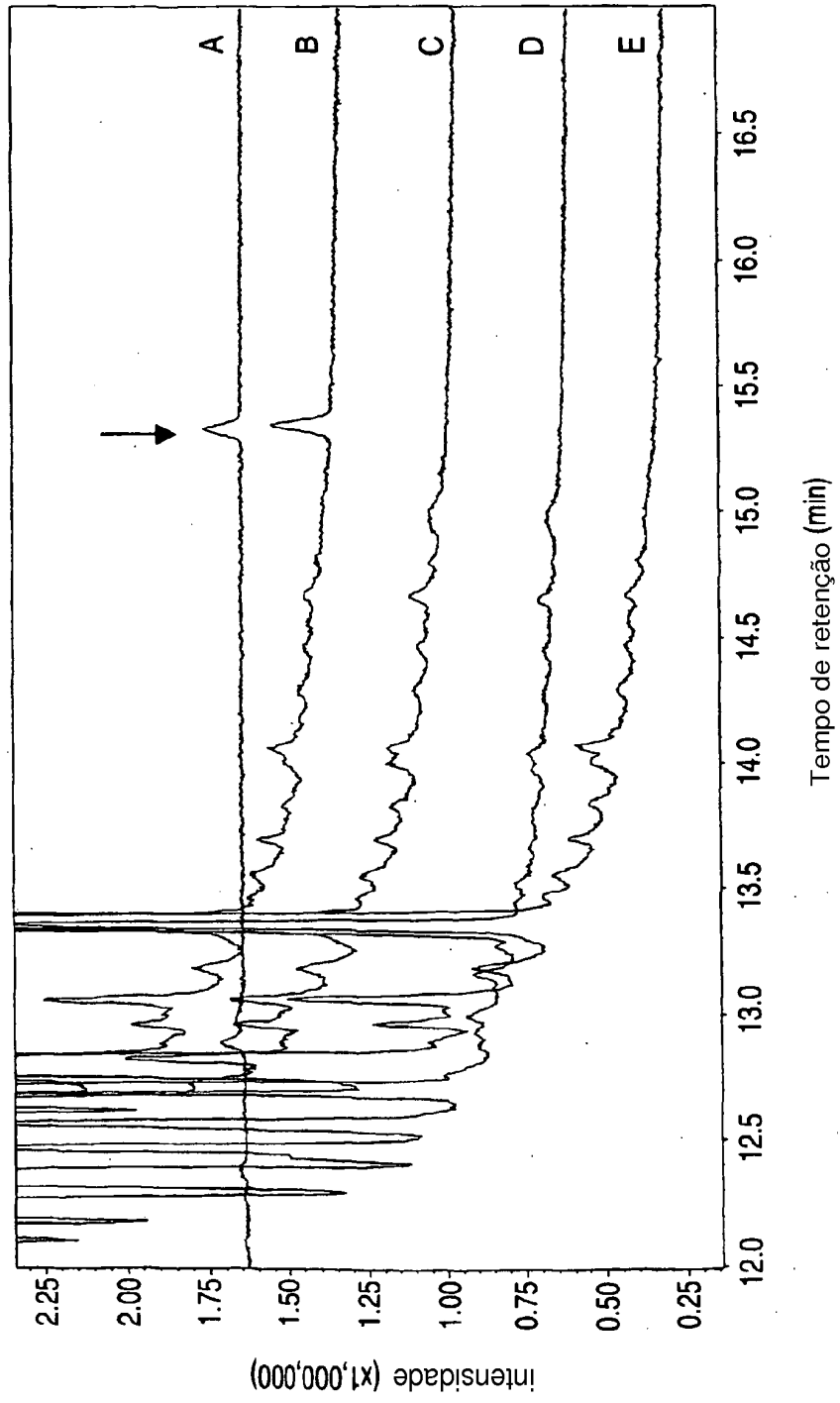


FIG. 4

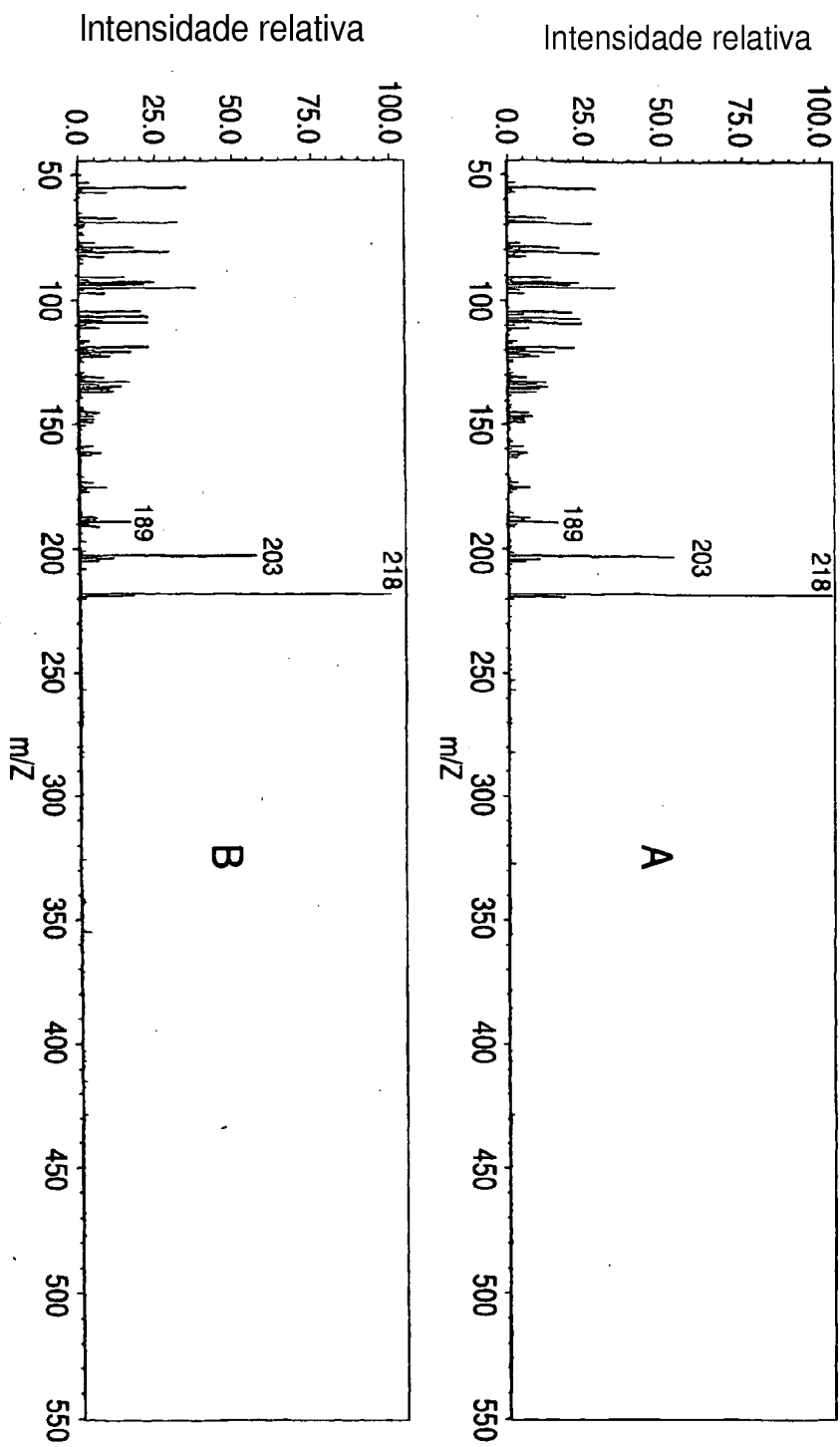


FIG. 5

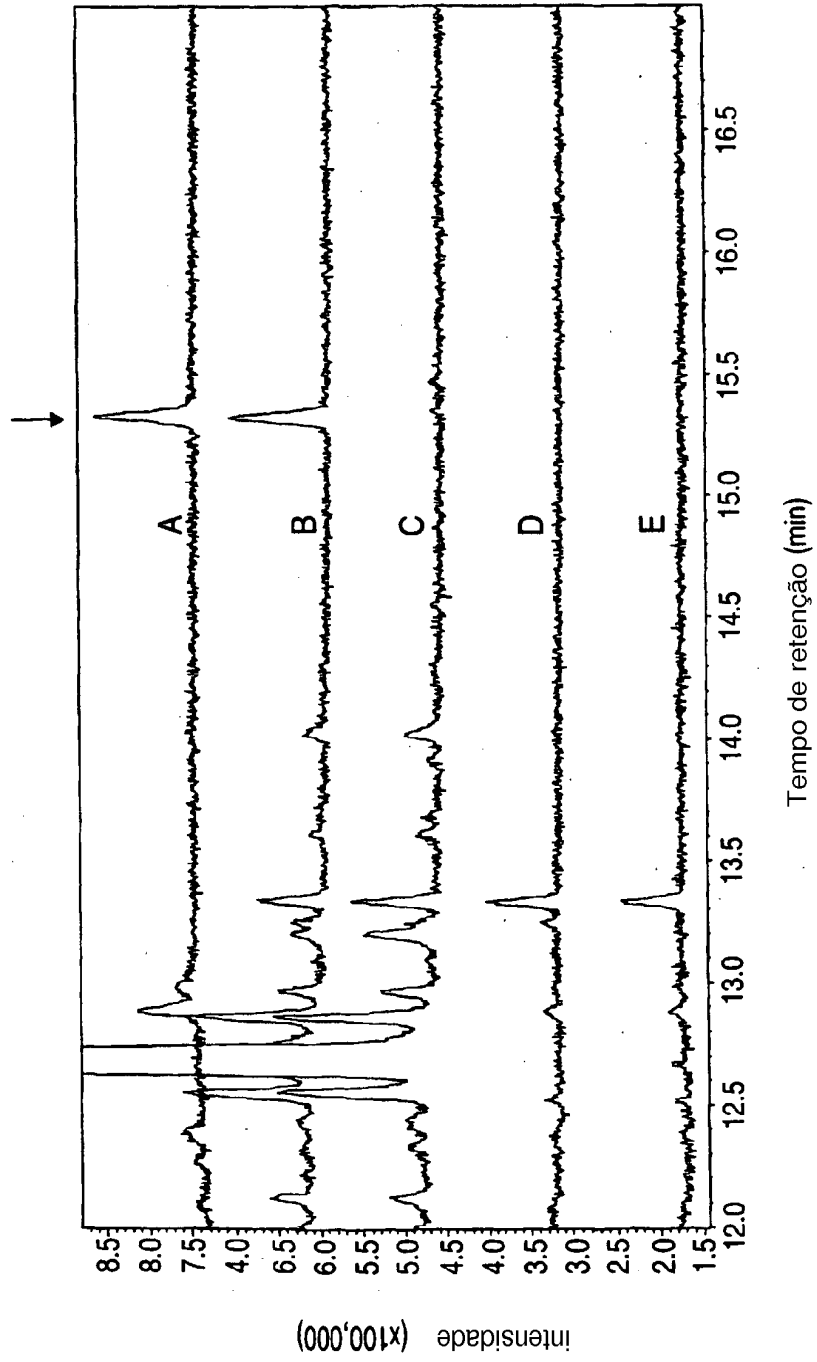


FIG. 6

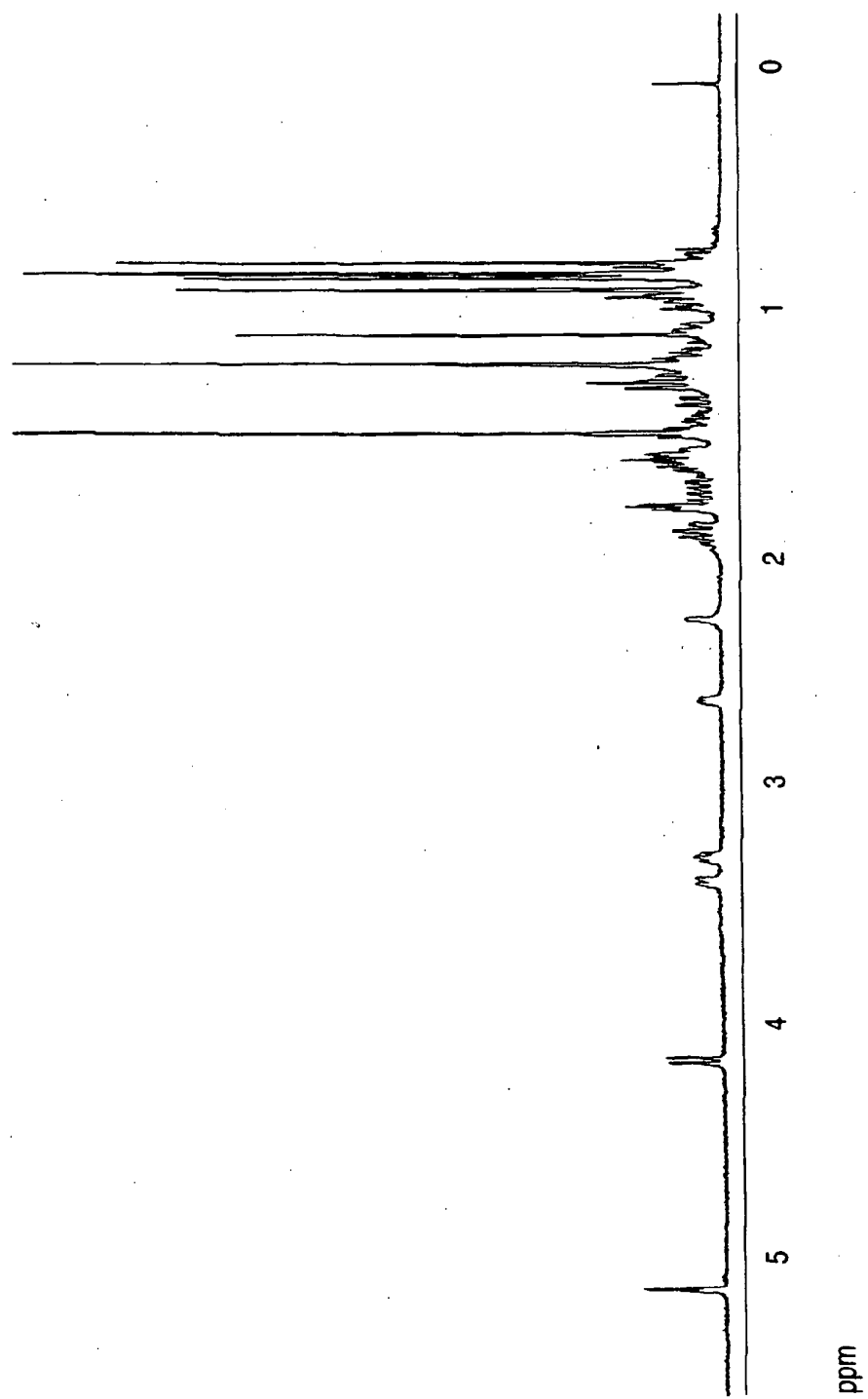


FIG. 7

