Methods and materials for production of terpenoids

Hamberger, Björn; Ranberg, Johan Andersen; Pateraki, Eirini; Møller, Birger Lindberg; Nielsen, Morten Thrane; Hansen, Nikolaj Lervad

Publication date:
2015

Document Version
Publisher's PDF, also known as Version of record

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
30 December 2015 (30.12.2015)

WIPO | PCT

(10) International Publication Number
WO 2015/197075 A1

(51) International Patent Classification:
C07K 14/415 (2006.01) *C12P 7/00* (2006.01)
C12P 5/00 (2006.01)

(21) International Application Number:
PCT/DK2015/050181

(22) International Filing Date:
23 June 2015 (23.06.2015)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
PA 2014 70381 23 June 2014 (23.06.2014) DK

(71) Applicants: **UNIVERSITY OF COPENHAGEN** [DK/DK]; Nørregade 10, DK-1165 Copenhagen K (DK). **DANMARKS TEKNISKE UNIVERSITET** [DK/DK]; Anker Engelundsvej 1, Bygning 101 A, DK-2800 Kgs Lyngby (DK).

(72) Inventors: **HAMBERGER, Björn**; Borgbygård Allee 11, DK-2770 Kastrup (DK). **ANDERSEN-RANBERG, Johan**; Hothers Plads 13, 3.th., DK-2200 Copenhagen N (DK). **PATERAKI, Eirini**; Gråspurvevej 71, 2.-2, DK-2400 Copenhagen NV (DK). **MØLLER, Birger Lindberg**; Kongstedvej 5, DK-2700 Brønshøj (DK). **NIELSEN, Morten Thrane**; Heinesgade 10, 1. sal, DK-2200 Copenhagen N (DK). **HANSEN, Nikolaj Lervad**; Vestbanevej 6, 3. th., DK-2500 Valby (DK).

(74) Agent: **HØIBERG A/S**; St. Kongensgade 59 A, 1264 Copenhagen K (DK).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

— *of inventorship (Rule 4.17(iv))*

Published:

— *with international search report (Art. 21(3))*

— *with sequence listing part of description (Rule 5.2(a))*

WO 2015/197075 A1

(54) **Title**: METHODS AND MATERIALS FOR PRODUCTION OF TERPENOIDS

(57) **Abstract**: The invention provide novel methods for production of terpenoids. The methods take advantage of combinations of enzymes, which are not present in nature. Thus the methods involves use of host organisms expressing a diTPS of class II, diTPS of class I, and one or more CYPs.

Methods and materials for production of terpenoids

Field of invention

The present invention relates to the field of biosynthetic methods for producing diterpenoids.

Background of invention

Terpenes constitute a large and diverse class of organic compounds produced by a variety of plants as well as other species. Terpenes modified by oxidation or rearrangements are generally referred to as terpenoids.

Terpenes and terpenoids find multiple uses, for example as flavor compounds, additives for food, as fragrances and in medical treatment

Terpenes are derived biosynthetically from units of isoprene, which has the molecular formula C_5H_8. Diterpenes are composed of *four isoprene* units and in nature they are produced from geranylgeranyl pyrophosphate.

Summary of invention

In nature diterpenes are produced with the aid of specific pairs of diterpene synthases (diTPS) derived from two classes, class I and class II. Diterpenoids may be produced from diterpenes with the action of various enzymes, including CYP enzymes.

Most CYP enzymes have a narrow substrate specificity, which has been a major hurdle in broadening P450 applications in synthetic chemistry (Podust et al., 2012, Nat. Prod. Rep., 29, 1251). In particular, plant CYPs are known to be recruited to highly specialised functions, and show a high degree of substrate specificity.

The present invention discloses that by combining diTPS enzymes of class I and class II, and CYPs from different species, many different diterpenoids may be produced including diterpenoids not identified in nature. Surprisingly, it is revealed that a diTPS enzyme of class I, a diTPS enzyme of class II may be combined with one or more

CYPs, wherein the diTPSs and the CYP(s) are from different species, resulting in a high diversity of diterpenoids, which can be produced.

Thus, the invention features an inventory of functional class II and class I diTPS and CYPs from a range of plants, which are useful for accumulating high-value and bioactive diterpenoids. When these diTPS and CYPs are paired into specific modules consisting of new-to-nature combinations, such as using enzymes from different plant species, both the structure and the stereochemistry of the formed diterpenoids can be controlled. This strategy gives access to a novel structural diversity of highly complex diterpenoids, many representing valuable compounds, for example flavours, fragrances, pharmaceuticals and fine chemicals. The diterpenoids may also be potentially bioactive molecules, starting materials for chemical synthesis, and intermediates for further functionalization to flavours, fragrances, pharmaceuticals and fine chemicals.

The invention thus in one aspect provides methods of producing a terpenoid, said methods comprising the steps of:

a) providing one or more host organisms, wherein
   I. at least one of the host organisms comprises a heterologous nucleic acid encoding a diTPS of class II,
   II. at least one of the host organisms comprises a heterologous nucleic acid encoding a diTPS of class I,
   III. at least one of the host organisms comprises a heterologous nucleic acid encoding a cytochrome P450 (CYP)

   with the proviso that said diTPS of class II, said diTPS of class I and said CYP are not all from the same species;

b) Incubating said one or more host organisms either simultaneously or sequentially under conditions allowing growth of said host organisms, wherein at least the host organism comprising a heterologous nucleic acid encoding a diTPS of class II is incubated in the presence of geranylgeranyl pyrophosphate (GGPP);

c) Optionally isolating diterpenoid from the host organism(s) and/or from cultivation medium in which said host organism has been incubated.

The invention also provides a host organism comprising

I. a heterologous nucleic acid encoding a diTPS of class II,
II. a heterologous nucleic acid encoding a diTPS of class I,
III. a heterologous nucleic acid encoding a cytochrome P450 (CYP)

with the proviso that said diTPS of class II, said diTPS of class I and said CYP are not all from the same species.

The invention furthermore provides a collection of host organisms comprising at least two different host organisms, wherein

I. at least one of the host organisms comprises a heterologous nucleic acid encoding a diTPS of class II,
II. at least one of the host organisms comprises a heterologous nucleic acid encoding a diTPS of class I,
III. at least one of the host organisms comprises a heterologous nucleic acid encoding a cytochrome P450 (CYP)

with the proviso that said diTPS of class II, said diTPS of class I and said CYP are not all from the same species.

The term "said diTPS of class II, said diTPS of class I and said CYP are not all from the same species" means that the diTPS of class II, the diTPS of class I and the CYP are derived from at least two different species. For example the diTPS of class II and the diTPS of class I may be from the same species, whereas the CYP is from a different species. In another example the diTPS of class II and the CYP may be from the same species, whereas the diTPS of class I is from a different species. In another example the diTPS of class I and the CYP may be from the same species, whereas the diTPS of classII is from a different species. In another example the diTPS of class II and the diTPS of class I and the CYP are from three different species.

Description of Drawings

Figure 1 provides an example of biosynthesis pathways to diterpenes of different stereochemistry. The figure shows biosynthesis of three different isomers of manool by

using diTPS enzymes from four different species: *Oryza sativa* (rice), *Zea maize* (maize), *Coleus forskohlii* (Indian coleus, medicinal plant) and *Salvia sclarea* (Clary sage, medicinal plant). The diTPS from *Oryza sativa* may for example be the enzyme of SEQ ID NO:1. The diTPS from *Zea maize* may for example be the enzyme of SEQ ID NO:3. The diTPS from *Coleus forskohlii* may for example be the enzyme of SEQ ID NO:5. The diTPS from *Salvia sclarea* may for example be the enzyme of SEQ ID NO:11.

Figures 2A and 2B show "Combinatorial wheels" showing examples of compounds, which can be made by combining different diTPS enzymes. The universal precursor, GGPP is shown in the middle. The next ring shows various examples of diTPS class II enzymes. The next ring shows various examples of diTPS class I enzymes. The outer ring shows the diterpenes produced by the indicated combinations of diTPS class II and diTPS class I enzymes. Each diterpene has been assigned a compound number used to identify said diterpene herein. The sequences of all of diTPS class II and diTPS class I enzymes are provided herein in the sequence listing. Table 1 provides a list and structures of the diterpenes.

Figures 3A and 3B shows the reactions catalysed by various class II diTPS enzymes as well as the diterpene pyrophosphate intermediates generated by the reactions.

Figure 4 shows an alignment of the amino acid sequences of selected diTPS enzymes of class I.

Figure 5 shows an alignment of the amino acid sequences of selected diTPS enzymes of class II.

Figure 6 shows an alignment of amino acid sequences of selected CYPs. 6 motifs named SRS1, SRS2, SRS3, SRS4, SRS5 and SRS6 are marked above the alignment with black squares. 4 highly conserved motifs are marked by numbers 1 to 4 and shown in black boxes. Other highly conserved residues are also marked in black boxes. Somewhat conserved residues are marked with grey boxes.

Figure 7 shows CYP catalysed bioconversion of Dehydroabietadiene (1.1) and Miltiradiene (1.2). Cytochrome P450 enzymes CfCYP76AH11 (top) and CfCYP71D381 (bottom) catalyzes oxidation of Dehydroabietadiene (1.1) and Miltiradiene (1.2) into

novel compounds 1.1.1-1.1.4 and 1.2.1 corresponding to hydroxylated or ketonated derivatives of Dehydroabietadiene (1.1) and Miltiradiene (1.2).

Figure 8A shows GC-MS chromatogram of *N. benthamiana* expressing CfTPS1+CfTPS3+CYP76AH11 and mass spectra of compound 1.1.1 & 1.2.1

Figure 8B shows GC-MS chromatogram of *N. benthamiana* expressing CfTPS1 + CfTPS3 + CYP71D381 and mass spectra of compound 1.1.2-1.1.4.

Figure 9 shows CYP catalysed bioconversion of Manool (1.3). Cytochrome P450 enzymes CfCYP76AH8 (top), CfCYP76AH11 (middle) and PsCYP720B4 (bottom) catalyzes oxidation of Manool (1.3) into novel compounds 1.3.1-1.3.10 corresponding to hydroxylated or dihydroxylated derivatives of Manool (1.3).

Figure 10A shows GC-MS chromatogram of *N. benthamiana* expressing CfTPS1+SsSCS+CYP76AH8 and mass spectra of compound 1.3.1 & 1.3.5.

Figure 10B shows GC-MS chromatogram of *N. benthamiana* expressing CfTPS1+SsSCS+CYP76AH11 and mass spectra of compound 1.3.6 & 1.3.9.

Figure 10C shows GC-MS chromatogram of *N. benthamiana* expressing CfTPS1+SsSCS+CYP720B4 and mass spectra of compound 1.3.10.

Figure 11 shows CYP catalysed bioconversion of syn-manool (1.4). Cytochrome P450 enzymes CfCYP71D381 catalyzes oxidation of syn-manool (1.4) into novel compounds 1.4.1-1.4.3 where 1.4.2 and 1.4.3 correspond to hydroxylated derivatives of syn-manool (1.4).

Figure 12 shows GC-MS chromatogram of *N. benthamiana* expressing sCPSsyn+SsSCS+CYP71D381 and mass spectra of compound 1.4.1-1.4.3.

Figure 13 shows production of compound 1.2.1 by a yeast strain expressing the diterpene synthases CfTPS1 and CfTPS3 as well as the cytochrome P450 PsCYP720B4, Panel A shows extracted ion chromatogram recorded by LC-MS analysis of ethanol extracts generated from media as well as cell material of strain IT-7-

T13. The black trace corresponds to a yeast strain expressing the diterpene synthases CfTPS1 and CfTPS3 as well as the cytochrome P450 PsCYP720B4, whereas the grey trace corresponds to a yeast strain expressing only the diterpene synthases CfTPS1 and CfTPS3. Panel B, UV absorption (top), primary mass spectra (middle) as well as MS-n fragmentation mass spectra (bottom). Panel C shows a schematic representation offormation of miltiradiene (1.2) and dehydroabietadiene by CfTPS1 and CfTPS3 and proposed cytochrome P450 catalysed oxidation of miltiradiene into the corresponding carboxylic acid, based on the parental mass.

Figure 14 shows production of oxidized manool. Panel A, shows a schematic presentation of formation of manool (1.3) and proposed cytochrome P450 catalyzed oxidation of manool in both the plant *Nicotiana benthamiana* and in yeast after heterologous expression of CfTPS1, SsSCS and the cytochrome P450 enzymes CYP76AH8, CYP76AH11 and CYP720B4, which catalyse oxidation of manool into novel compounds 1.3.1-1.3.16. Panel B shows extracted ion chromatograms of three strains recorded by LC-MS analysis of ethanol extracts generated from media as well as cell material. Extracted ion trace combine the ions corresponding to the mass of manool (273 m/z), monohydroxy-manool (289 m/z, hydrogen adduct of water loss ion), dihydroxymanool (347 m/z) and keto-hydroxy-manool (345 m/z). Result from yeast strains expressing CfTPS1, SsSCS and CfCYP76AH8 (strain IT7-T50)(black trace, top), CfTPS1 and SsSCS (dark grey trace, middle) or an untransformed reference strain (light grey, bottom) are shown. Panels C, D and E display, UV (top), primary mass spectra (middle) as well as MS-n fragmentation mass spectra (bottom) of putative keto-hydroxy mannol (1.3.11), dihydroxy manool (1.3.12) and hydroxyl manool (1.3.13) in yeast cell expressing CfTPS1, SsSCS, CfCYP76AH8, strain name IT7-T50.

Figure 15 shows production of oxidized manool. Panel A shows extracted ion chromatograms of two strains recorded by LC-MS analysis of ethanol extracts generated from media as well as cell material. Extracted ion trace combine the ions corresponding to the mass of monohydroxy-manool (329 m/z, sodium adduct), dihydroxymanool (345 m/z, sodium adduct) and keto-hydroxy-manool or a carboxylic acid of manool (343 m/z, sodium adduct). Result from yeast strains expressing CfTPS1, SsSCS and PsCYP720B4 (black trace, top), CfTPS1 and SsSCS (dark grey trace, bottom) are shown. Panels B display, UV (top), primary mass spectra (bottom) of

the oxidized compounds 1.3.12-1.3.16 in yeast cells expressing CfTPS1, SsSCS, PsCYP720B4, strain name IT7-T37.

Figure 16 shows detection of oxidised isopimaradiene. Panel A shows a schematic representation of formation of syn-or iso-pimaradienes (1.5-group) and proposed cytochrome P450 catalyzed oxidation of syn-or iso-pimaradienes into novel compounds 1.5.1-1.5.6 tentatively identified as hydroxylated, dihydroxylated or ketonated derivatives of either syn- or iso-pimaradiene based on their parental masses. Panel B shows extracted ion chromatogram recorded by LC-MS analysis of ethanol extracts generated from media as well as cell material. Extracted ion correspond to the mass of syn-pimara-9,(11),15-diene or similar C20H32 diterpene oxidized to contain either one keto- and one hydroxyl group or a single carboxylic acid group (303 m/z) for yeast cells expressing OssynCPP, CfTPS3 and PsCYP720B4 (IT7-T58)(black trace), untransformed reference strain (grey trace). Panel C, UV (top), primary mass spectra (middle) as well as MS-n fragmentation mass spectra (bottom) of the five oxidized diterpene derivatives. The 257 m/z fragment observed in all compounds 1.5.1-1.5.5 is a signature ion of diterpenes. This ion is observed as 253 m/z in 1.5.6 possibly indicative of introduction of two double bonds in the structure. (OssynCPP, CfTPS3, PsCYP720B4, strain name IT7-T58)

Figure 17 shows detection of higher oxidised manool through pairs of cytochrome P450s. Panel A shows a schematic representation. Cytochrome P450 enzyme pairs CYP76AH8, CYP71D381; CYP76AH8, CYP76AH11 and CYP76AH8, CYP720B4 catalyse oxidation of Manool into novel compounds 1.3.14-1.3.18. Panel B shows extracted ion chromatograms of three strains recorded by LC-MS analysis of ethanol extracts generated from media as well as cell material. Extracted ion trace combine the ions corresponding to the mass of monohydroxy-manool (289 m/z = hydrogen adduct of water loss ion) and dihydroxymanool (345 m/z = sodium adduct, 305 = hydrogen adduct of water loss ion). Yeast cells expressing CfTPS1, SsSCS and CfCYP76AH8 and CfCYP71D381 (strain IT7-T32)(black trace, top), yeast cells expressing CfTPS1, SsSCS and CfCYP76AH8 (strain IT7-T50)(dark grey trace, middle) or an untransformed control strain (light grey, bottom). Panel C display, UV (top), primary mass spectra (middle) as well as MS-n fragmentation mass spectra (bottom) of oxidized diterpenes.

Figure 18 shows detection of oxidized manool. Panel A shows extracted ion chromatograms of three strains recorded by LC-MS analysis of ethanol extracts generated from media as well as cell material. Extracted ion trace combine the ions corresponding to the mass of monohydroxy-manool (289 m/z = hydrogen adduct of water loss ion) and dihydroxymanool (345 m/z = sodium adduct, 305 = hydrogen adduct of water loss ion). CfTPS1, SsSCS and CfCYP76AH8 and CfCYP76AH11 (strain IT7-T31)(black trace, top), CfTPS1, SsSCS and CfCYP76AH8 (dark grey trace, middle), untransformed control strain (light grey, bottom). Panel B display, UV (top), primary mass spectra (middle) as well as MS-n fragmentation mass spectra (bottom) of oxidized diterpenes.

Figure 19 shows detection of oxidized manool. Panel A shows extracted ion chromatograms of three strains recorded by LC-MS analysis of ethanol extracts generated from media as well as cell material. Yeast cells expressing CfTPS1, SsSCS and CfCYP76AH8 and PsCYP720B4 (strain IT7-T9)(black trace, top), yeast cells expressing CfTPS1, SsSCS and CfCYP76AH8 (dark grey trace, middle), yeast cells expressing CfTPS1, SsSCS and PsCYP720B4 (light grey, bottom). Panel B displays, UV (top), primary mass spectra (middle) as well as MS-n fragmentation mass spectra (bottom) of oxidized diterpenes.

Figure 20 shows a schematic pathway for production of ferruginol. Thus, co-expression of TwTPS7, CfTPS3 and CYP76AH4 resulted in the formation of ferruginol.

Detailed description of the invention

Method for producing diterpenoids

The present invention relates to a biosynthetic method for producing diterpenoids. The methods typically involves the steps of

a) providing one or more host organisms, wherein
   I. at least one of the host organisms comprises a heterologous nucleic acid encoding a diTPS of class II,
   II. at least one of the host organisms comprises a heterologous nucleic acid encoding a diTPS of class I,

III. at least one of the host organisms comprises a heterologous nucleic acid encoding a cytochrome P450 (CYP)

with the proviso that said diTPS of class II, said diTPS of class I and said CYP are not all from the same species;

b) Incubating said one or more host organisms either simultaneously or sequentially under conditions allowing growth of said host organisms, wherein at least the host organism comprising a heterologous nucleic acid encoding a diTPS of class II is incubated in the presence of geranylgeranyl pyrophosphate (GGPP);

c) Optionally isolating diterpenoid(s) from the host organism(s) and/or from its surroundings.

The host organism may be cultivated in a cultivation medium. For example, in embodiments of the invention, wherein the host organism is a microorganism, then typically said host organism is cultivated in a cultivation medium under conditions allowing growth of the particular microorganism. When the host organism is cultivated in a cultivation medium, then the diterpenoid may be isolated from the host organism(s) and/or from the cultivation medium. Similarly, intermediates may also be obtained and/or isolated from the host organism(s) and/or the cultivation medium.

In embodiments of the invention, wherein the host organism is a plant, said host organism may be cultivated in soil. In such embodiments, then the diterpenoid may typically be isolated from the host organism or parts thereof. Similarly, intermediates may also be obtained and/or isolated from the host organism or part thereof.

In one preferred embodiment of the invention step a) comprises providing one host organism comprising:

I. a heterologous nucleic acid encoding a diTPS of class II,
II. a heterologous nucleic acid encoding a diTPS of class I,
III. a heterologous nucleic acid encoding a cytochrome P450 (CYP).

In addition to the aforementioned heterologous nucleic acids said host organism may contain one or more additional heterologous nucleic acids. For example the host organism may comprise at least two heterologous nucleic acids encoding different

CYPs. Said CYPs may be from the same or from different species. Thus, the host organism may comprise:

I. a heterologous nucleic acid encoding a diTPS of class II,
II. a heterologous nucleic acid encoding a diTPS of class I,
III. at least two heterologous nucleic acids each encoding a different cytochrome P450 (CYP).

Even though it is preferred that one host organism expresses both diTPS of class II, diTPS of class I and one or more CYPs, it is also comprised with the invention that the methods may employ different host organisms expressing one or more of said enzymes.

In embodiments of the invention employing use of more than one host organism, said host organisms may be incubated in cultivation medium together or sequentially. When incubated sequentially one or more host organisms may be incubated together followed by incubation of one or more other host organisms. Alternatively, each host organism may be incubated separately.

When the host organisms are incubated sequentially, the order of incubation is of importance. Thus, for example:

I. a host organism comprising a heterologous nucleic acid encoding a diTPS of class II may beincubated in cultivation medium before or simultaneously with any host organism comprising a heterologous nucleic acid encoding a diTPS of class I
II. a host organism comprising a heterologous nucleic acid encoding a diTPS of class I may beincubated in cultivation medium before or simultaneously with any host organism comprising a heterologous nucleic acid encoding a CYP.

Thus, in one embodiment the methods of the invention may comprise the steps of

a) providing
   I. a host organism comprising a heterologous nucleic acid encoding a diTPS of class II,
   II. a host organism comprising a heterologous nucleic acid encoding a diTPS of class I,

III. a host organism comprising one or more heterologous nucleic acids encoding a cytochrome P450 (CYP)

with the proviso that said diTPS of class II, said diTPS of class I and said CYP are not all from the same species;

b) Incubating said host organisms under conditions allowing growth of said host organisms, wherein

i. all host organisms are incubated together in the presence of geranylgeranyl pyrophosphate (GGPP); or

ii. the host organisms under I. and II are incubated together in the presence of GGPP, thereby producing a diterpene, followed by incubation of the host organism(s) under III in the presence of said diterpene; or

iii. the host organism under I. is incubated in the presence of GGPP thereby producing a diterpene pyrophosphate intermediate, followed by incubation of the host organisms under II. and III. in the presence of said diterpene pyrophosphate intermediate; or

iv. the host organism under I. is incubated in the presence of GGPP thereby producing a diterpene pyrophosphate intermediate, followed by incubation of the host organism under II. in the presence of said diterpene pyrophosphate intermediate there by producing a diterpene, followed by incubation of the host organism(s) under III. in the presence of said diterpene,

c) Optionally isolating diterpenoid from the host organism(s) and/or from the cultivation medium if the host organism has been cultivated in a cultivation medium.

It is also comprised within the invention that the order of incubation may be in the reverse order, i.e. starting with incubation of host organism(s) containing downstream enzymes, such as CYP or diTPS of class II, followed by incubation of host organism(s) containing upstream enzymes, such as diTPS of class I. This may be done in order to prevent accumulation of intermediates and/or undesired side products.

In another embodiment the methods of the invention may comprise the steps of

a) providing two host organisms, whereon the first host organism comprises
   I. a heterologous nucleic acid encoding a diTPS of class II,
   II. a heterologous nucleic acid encoding a diTPS of class I,

   and the second host organism comprises:
   III. a heterologous nucleic acid encoding a cytochrome P450 (CYP)

   with the proviso that said diTPS of class II, said diTPS of class I and said CYP are not all from the same species;

b) Incubating said host organisms under conditions allowing growth of said host organisms, wherein
   i. both host organisms are incubated simultaneously in the presence of geranylgeranyl pyrophosphate (GGPP); or
   ii. the first host organism is incubated in the presence of GGPP thereby producing a diterpene, followed by incubation of the second host organism in the presence of said diterpene;

c) Optionally isolating diterpenoid from the host organism(s) and/or from the cultivation medium if the host organism has been cultivated in a cultivation medium.

In yet another embodiment the methods of the invention may comprise the steps of

a) providing two host organisms, wherein the first host organism comprises
   I. a heterologous nucleic acid encoding a diTPS of class II,
   II. a heterologous nucleic acid encoding a diTPS of class I,

   and the second host organism comprises
   III. at least two heterologous nucleic acids each encoding a different cytochrome P450 (CYP)

   with the proviso that said diTPS of class II, said diTPS of class I and said CYP are not all from the same species;

b) Incubating said host organisms in cultivation medium under conditions allowing growth of said host organisms, wherein
   i. both host organisms are incubated simultaneously in the presence of geranylgeranyl pyrophosphate (GGPP); or

ii. the first host organism is incubated in the presence of GGPP thereby producing a diterpene, followed by incubation of the second host organism in the presence of said diterpene;

c) Optionally isolating diterpenoid from the host organism(s) and/or from the cultivation medium if the host organism has been cultivated in a cultivation medium.

In yet another embodiment the methods of the invention may comprise the steps of

a) providing at least three host organisms, wherein the first host organism comprises
   I. a heterologous nucleic acid encoding a diTPS of class II,
   II. a heterologous nucleic acid encoding a diTPS of class I,
   and the second host organism comprises
   III. a heterologous nucleic acids encoding a cytochrome P450 (CYP)
   and the third and optionally further host organisms each comprises
   IV. a heterologous nucleic acids encoding a cytochrome P450 (CYP) different to the CYP under III.,

   with the proviso that said diTPS of class II, said diTPS of class I and said CYP are not all from the same species;
b) Incubating said host organisms in cultivation medium under conditions allowing growth of said host organisms, wherein
   i. all host organisms are incubated simultaneously in the presence of geranylgeranyl pyrophosphate (GGPP); or
   ii. the first host organism is incubated in the presence of GGPP thereby producing a diterpene, followed by incubation of the second host organism in the presence of said diterpene, thereby producing an intermediate diterpenoid, followed by incubation of the third host organism in the presence of said intermediate diterpenoid; or
   iii. the first host organism is incubated in the presence of GGPP thereby producing a diterpene, followed by incubation of the second host organism in the presence of said diterpene, thereby producing an intermediate diterpenoid, followed by incubation of the third host organism in the presence of said intermediate

diterpenoid, thereby producing a second intermediate diterpenoid; followed by simultaneous or sequential incubation of said further host organisms in the presence of intermediate diterpenoids; or

iv. the first and the second host organism are incubated together in the presence of GGPP thereby producing an intermediate diterpenoid, followed by incubation of the third host organism in the presence of said intermediate diterpenoid; or

v. the first and the second host organism are incubated together in the presence of GGPP thereby producing an intermediate diterpenoid, followed by incubation of the third host organism in the presence of said intermediate diterpenoid, thereby producing a second intermediate diterpenoid; followed by simultaneous or sequential incubation of said further host organisms in the presence of intermediate diterpenoids; or

Optionally isolating diterpenoid from the host organism(s) and/or from the cultivation medium if the host organism has been cultivated in a cultivation medium.

In addition to the methods specifically outlined herein above, additional combinations of host organisms comprising diTPS of class I, diTPS of class II and various CYPs may be envisaged. Such combinations may also be employed with the present invention.

The method may be performed in vitro or in vivo. It is generally preferred, that the methods of the invention are performed in vivo. The term “in vivo” as used herein refers to that the method is performed within a host organism, which for example may be any of the host organisms described herein below in the section “Host organism”.

The diterpene pyrophosphate intermediate and the diterpene may for example be any of the compounds described herein below in the sections “Diterpene pyrophosphate intermediates” and “Diterpenes”.

The diterpenoid intermediates and the diterpenoids may be any of the compounds described herein below in the section “Diterpenoids”.

In preferred embodiments of the invention the host organism is capable of producing GGPP. In particular, it is preferred that the host organism comprising a heterologous nucleic acid encoding a diTPS of class II is capable of producing GGPP. Thus, the step of incubating the host organism(s) in the presence of GGPP may simply be performed by cultivating said host organism. Many host organisms produce GGPP endogenously. Thus, the host organism may be a host organism, which endogenously produce GGPP. Such host organisms for example include plants and yeast. Even if the host organism produces GGPP endogenously, the host organism may be recombinantly modulated to upregulate production of GGPP. This may be achieved by expressing one or more enzymes involved in the production of GGPP in said host organism. In addition or alternatively, the expression of one or more enzymes involved in reducing the level of GGPP may be reduced or even abolished.

For example, the host organism may comprise a heterologous nucleic acid encoding a 1-deoxy-D-xylulose-5-phosphate synthase. The host organism may comprise a heterologous nucleic encoding DXS, such as DXS *Coleus forskohlii*. DXS may be a polypeptide encoded by SEQ ID NO:25 or a functional homologue thereof sharing at least 70%, such as at least 80%, for example at least 75%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 91%, such as at least 92%, such as at least 93%, such as at least 94%, such as at least 95%, such as at least 96%, such as at least 97%, such as at least 98%, such as at least 99%, such as 100% sequence identity therewith.

In one aspect the invention relates to DXS encoded by SEQ ID NO:25 or a functional homologue, as well as to host organisms containing a nucleic acid encoding same.

The host organism may comprise a heterologous nucleic acid encoding a GGPP synthase (GGPPS), for example a GGPPS1, such as GGPPS1 of *Coleus forskohlii*. The GGPPS1 may be a polypeptide encoded by SEQ ID NO:26 or a functional homologue thereof sharing at least 70%, such as at least 80%, for example at least 75%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 91%, such as at least 92%, such as at least 93%, such as at least 94%, such as at least 95%, such as at least 96%, such as at least 97%, such as at least 98%, such as at least 99%, such as 100% sequence identity therewith.

In one aspect the invention relates to GGPPS1 encoded by SEQ ID NO:26 or a functional homologue, as well as to host organisms containing a nucleic acid encoding same.

It is also comprised within the invention that GGPP is introduced to the host organism. If the host organism is a microorganism, then GGPP may be added to the cultivation medium of said microorganism. If the host organism is a plant, then GGPP may be added to the growing soil of the plant or it may be introduced into the plant by infiltration. Thus, if the heterologous nucleic(s) are introduced into the plant by infiltration, then GGPP may be co-infiltrated together with the heterologous nucleic acid(s).

When the methods of producing a diterpernoid are performed in vitro, the methods typically comprises the steps of:

a) providing a host organism or a collection of host organism, wherein
   I. at least one of the host organisms comprises a heterologous nucleic acid encoding a diTPS of class II,
   II. at least one of the host organisms comprises a heterologous nucleic acid encoding a diTPS of class I,
   III. at least one of the host organisms comprises a heterologous nucleic acid encoding a cytochrome P450 (CYP)

b) preparing an extract of said host organism(s) or preparing individual extracts or each host organism;

c) providing GGPP

d) incubating one or more of said extracts with GGPP, wherein at least the extract of the host organism comprising a heterologous nucleic acid encoding a diTPS of class II is incubated in the presence of geranylgeranyl pyrophosphate (GGPP);

e) optionally, incubating the product of the incubation under d) with one or more of said extracts,

f) optionally, incubating the product of the incubation under e) with one or more of said extracts, wherein step f) may optionally be repeated;

thereby producing a diterpenoid.

Thus, when the methods are performed in vitro, an extract containing all enzymes may be incubated with GGPP thereby producing a diterpenoid, or extracts containing different enzymes may be used sequentially. If the extracts are used sequentially the order is of importance.

Thus, steps d) to f) may for example be performed in either of the following ways:

i. incubating an extract containing diTPS of class I, diTPS of class II and one or more CYPs in the presence of geranylgeranyl pyrophosphate (GGPP); or
ii. incubating an extract containing diTPS of class I and diTPS of class II in the presence of GGPP, thereby producing a diterpene, followed by incubation an extract containing CYP in the presence of said diterpene; or
iii. incubating an extract containing diTPS of class II in the presence of GGPP thereby producing a diterpene pyrophosphate intermediate, followed by incubation of an extract containing diTPS of class I and CYP in the presence of said diterpene pyrophosphate intermediate; or
iv. incubating an extract containing diTPS of class II in the presence of GGPP thereby producing a diterpene pyrophosphate intermediate, followed by incubation of an extract containing diTPS of class I in the presence of said diterpene pyrophosphate intermediate there by producing a diterpene, followed by incubation of an extract containing a CYP in the presence of said diterpene,
v. incubating an extract containing diTPS of class I and diTPS of class II in the presence of GGPP, thereby producing a diterpene, followed by incubation of an extract comprising a CYP in the presence of said diterpene, thereby producing an intermediate diterpenoid, followed by incubation of an extract comprising at least one other CYP in the presence of said intermediate diterpenoid,
vi. incubating an extract containing diTPS of class I and diTPS of class II in the presence of GGPP, thereby producing a diterpene, followed by incubation of an extract comprising a CYP in the presence of said diterpene, thereby producing an intermediate diterpenoid, followed by incubation of an extract

comprising at least one other CYP in the presence of said intermediate diterpenoid, thereby producing a second intermediate diterpenoid; followed by simultaneous or sequential incubation of extract(s) containing additional CYP(s) in the presence of intermediate diterpenoids; or

vii. incubating an extract containing diTPS of class I and diTPS of class II and CYP(s) in the presence of GGPP, thereby producing an intermediate diterpenoid, followed by incubation of an extract comprising at least one other CYP in the presence of said intermediate diterpenoid; or

viii. incubating an extract containing diTPS of class I and diTPS of class II and CYP(s) in the presence of GGPP, thereby producing an intermediate diterpenoid, followed by incubation of an extract comprising at least one other CYP in the presence of said intermediate diterpenoid, followed by incubation of an extract comprising at least one other CYP in the presence of said intermediate diterpenoid, thereby producing a second intermediate diterpenoid; followed by simultaneous or sequential incubation of extract(s) containing additional CYP(s) in the presence of intermediate diterpenoids.

Alternative combinations of extracts comprising diTPSs and CYPS may also be envisaged and are comprised within the scope of the present invention.

When the methods are performed sequentially, the intermediate products, for example the diterpene, the diterpene pyrophosphate intermediate or any of the intermediate diterpenoids, may be purified or partly purified before performing the next step of the method. However, it is also comprised within the invention that crude extract or cultivation medium is used for the subsequent steps.

In order to produce a specific diterpenoid according to the present invention, a useful combination of a diTPS of class II and a diTPS of class I and one or more CYPs must be employed. Examples of specific combinations of a diTPS of class II and a diTPS of class I, which leads to production of specific diterpenes, are shown in figure 2. Thus, for production of a diterpenoid having any of the diterpenes shown in figure 2 as intermediate, then the specific combination of a diTPS of class II and a diTPS of class I shown in figure 2 may be employed in combination with one or more CYP.

Other combinations of diTPS of class II and diTPS of class I may be used. In general, the diTPS of class II is selected so that it produces a diterpene pyrophosphate intermediate containing a decalin core having the desired stereochemistry at the 9 and 10 substitutions. Useful diTPS of class II are described below and also specific diTPS of class II catalysing formation of diterpene pyrophosphate intermediates with a specific stereochemistry are described.

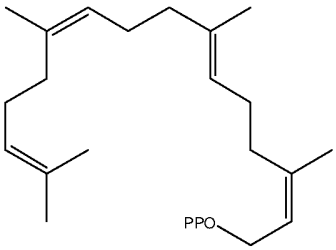
The diTPS of class I is selected so that is catalyses the conversion of the diterpene pyrophosphate intermediate to a diterpene, which can serve as an intermediate in the production of the final diterpenoid. Since a diterpenoid according to the present invention is a diterpene substituted with one or more hydroxyl, keto and/or carboxyl groups, for simplicity the unsubstitued diterpene may be referred to as the "precursor diterpene" of the diterpenoid. By way of example, dehydroabietadiene may be referred to as the precursor diterpene of hydroxy dehydroabietadiene. Thus, diTPS of class I is selected so that is catalyses the conversion of the diterpene pyrophosphate intermediate to a precursor diterpene of the diterpenoid of interest. Useful diTPS of class I are described below. Also specific reactions catalysed by various diTPS of class I are described, enabling the skilled person to select a useful diTPS of class I for production of a precursor diterpene of a desired diterpenoid.

The CYP(s) are collectively selected so that they catalyse oxidation of the precursor diterpene to obtain the desired diterpenoid. If the diterpenoid is a hydroxylated diterpene it is generally sufficient to use one CYP. Said one CYP is selected so that it can catalyse hydroxylation of the precursor diterpene. If the diterpenoid contains ketone and/or carboxyl and/or several hydroxyl groups, then one or more CYP may be employed. Details regarding how to select useful CYPs are given herein below in the section "CYP".

Once a useful diTPS of class II and diTPS of class I and one or more CYPs have been selected, nucleic acids encoding same may be expressed in the host organism(s) allowing production of the diterpenoid in a host organism. Putative useful combinations of a diTPS of class II and a diTPS of class I and CYP(s) for production of a given diterpenoid may be tested by expressing said diTPS of class II and said diTPS of class I and said CYP(s) in one or more host organism followed by testing for production of the diterpenoid, e.g. by GC-MS analysis and/or NMR analysis. Putative useful

combinations of a diTPS of class II and a diTPS of class I and CYP for production of a given diterpenoid may in particular be tested as described in Example 1 herein below. Methods for expression of enzymes in host organisms are well known to skilled person, and may for example include the methods described herein below in the section "Heterologous nucleic acids".

The term GGPP as used herein refers to geranylgeranyl diphosphate and is a compound of the following structure:



wherein PPO- is diphosphate. PPO- and –OPP may be used interchangeably herein.

As outlined above the methods of the invention involves use of diTPS of class II or of heterologous nucleic acids encoding diTPS of class II. The diTPS of class II may be any of the diTPS of class II described herein below in the section "diTPS of class II".

As outlined above the methods of the invention involves use of diTPS of class I or of heterologous nucleic acids encoding diTPS of class I. The diTPS of class I may be any of the diTPS of class I described herein below in the section "diTPS of class I".

As outlined above the methods of the invention involves use of CYP or of heterologous nucleic acids encoding CYP. The CYP may be any of the CYPs described herein below in the section "CYP".

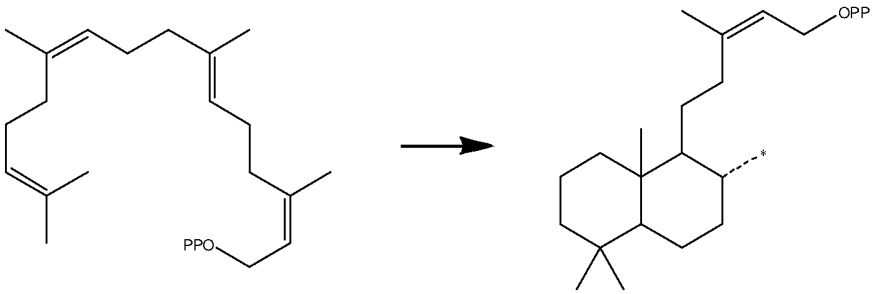
diTPS of class II

The methods of the invention use a diTPS of class II. The invention also features host organisms comprising a heterologous nucleic acid encoding a diTPS of class II.

Said diTPS of class II is an enzyme capable of catalysing protonation-initiated cationic cycloisomerization of GGPP to form a diterpene pyrophosphate intermediate. The class

II diTPS reaction, may be terminated either by deprotonation or by water capture of the diphosphate carbocation.

In particular the diTPS of class II may be an enzyme capable of catalysing the reaction I:



wherein PPO- is diphosphate and the *---- indicates either a double bond or two single bonds, wherein one is substituted with –OH and the other with $–CH_3$.

Thus, *---- may be 

or 

.

When no stereochemistry is indicated, the bond may be in any conformation. By selecting appropriate diTPS of class II the stereochemistry of the diterpene produced may be controlled. Accordingly, by following the description of the present invention, the skilled person may be able to design the production of a given precursor diterpene by selecting appropriate diTPS enzymes of class II and class I as described herein.

The diTPS of class II is generally a polypeptide sharing at least some sequence similarity to at least one of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7 or SEQ ID NO:8. In particular, it is preferred that the diTPS of class II shares at least 30%, preferably at least 40% sequence identity with at least one of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7 and SEQ ID NO:8. In particular, it is preferred that the diTPS of class II shares at least 30%, such as at least 35% sequence identity to the sequence of SsLPPS (SEQ ID NO:6) or to the sequence of AtCPS (see figure 5). Furthermore, it is preferred that the diTPS of class II in addition to above mentioned sequence identity also contains the following motif of four amino acids:

D/E-X-D-D,
wherein X may be any amino acid, such as any naturally occurring amino acids. In particular, X may be an amino acid with a hydrophobic side chain, and thus X may for example be selected from the group consisting of A, I, L, M, F, W, Y and V. Even more preferably X is an amino acid with a small hydrophobic side chain, and thus X may be selected from the group consisting of A, I, L and V.

In one embodiment of the invention said motif of four amino acids is:

D/E-I/V-D-D

D/E indicates that said amino acid may be D or E and I/V indicates that said amino acid may be I or V.

Amino acids are herein named using the IUPAC nomenclature for amino acids.

In particular, it is preferred that the diTPS of class II contains above described motif in a position corresponding to position aa 372 to 375 of SsLPPS of SEQ ID NO:6. A position corresponding to position aa 372 to 375 of SsLPPS of SEQ ID NO:6 is identified by aligning the sequence of a diTPS of class II of interest to SEQ ID NO:6 and optionally to additional sequences of diTPS of class II as e.g. shown in figure 5 and identifying the amino acids of said diTPS of class II aligning with aa 372 to 375 of SsLPPS of SEQ ID NO:6.

It is furthermore preferred that in addition to sharing above mentioned sequence identity and containing said motif, then as many as possible of the amino acids marked with a black box in figure 5 are retained. Thus, when aligned to the sequence of ScLPPS (SEQ ID NO:6), then preferably the diTPS of class II also contains at least 80%, more preferably at least 90%, for example at least 95%, such as all of the amino acids marked by a black box in figure 5. Alternatively, when aligned to the sequence of sequence of AtCPS (see figure 5), then preferably the diTPS of class II also contains at least 80%, more preferably at least 90%, for example at least 95%, such as all of the amino acids marked by a black box in figure 5.

Thus, the diTPS of class II may for example be selected from the group consisting of diTPS of class II of the following types:

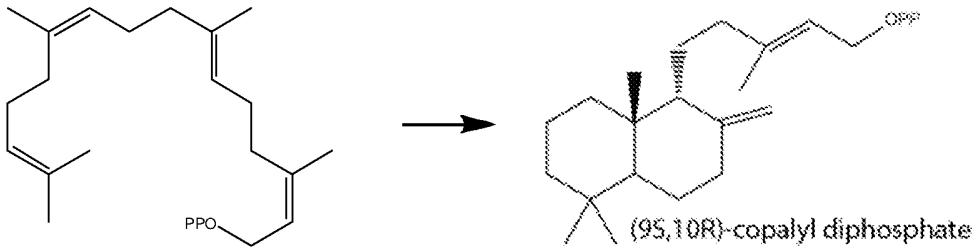
i. syn-CPP type, such as any of the enzymes described herein below in the section "syn-CPP type diTPS"
ii. ent-CPP type, such as any of the enzymes described herein below in the section "ent-CPP type diTPS"
iii. (+)-CPP type, such as any of the enzymes described herein below in the section "(+)-CPP type diTPS"
iv. LPP type, such as any of the such as any of the enzymes described herein below in the section "LPP type diTPS"
v. LPP like type, such as any of the enzymes described herein below in the section "LPP like type diTPS"

Certain diTPS enzymes are bifunctional in the sense that they may be classified as both class II and class I diTPS enzymes. Such bifunctional diTPS enzymes in general contain both the four amino acids motif: D/E-X-D-D, described herein above, as well as the five amino acid motif: D-D-X-X-D/E, described herein below. It is preferred that the diTPS of class II is not a bifunctional enzyme of both class II and class I. It is also preferred that the diTPS of class I is not a bifunctional enzyme of both class II and class I.

syn-CPP type diTPS

The methods of the invention involve use of a diTPS of class II. The invention also features host organisms comprising a heterologous nucleic acid encoding a diTPS of class II. In one embodiment said diTPS of class II is a syn-CPP type diTPS. Such diTPS of class II are in particular useful in embodiments of the inventions, wherein the diterpenoid to be produced contains a 9S,10R decalin core.

As used herein the term "syn-CPP type diTPS" refers to any enzyme capable of catalysing the reaction II:



wherein PPO- refers to diphosphate.

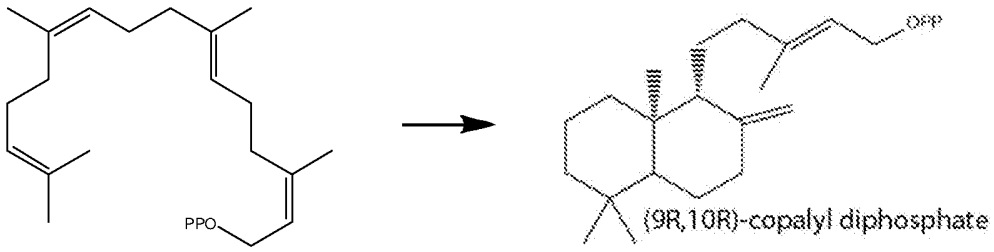
In one embodiment the syn-CPP type diTPS may be syn-copalyl pyrophosphate synthase (syn-CPP), such as syn-CPP from Oryza sativa. Syn-CPP may also be referred to as CPSsyn. In particular, said syn-CPP type diTPS may be a polypeptide of SEQ ID NO:1 or a functional homologue thereof sharing at least 70%, such as at least 80%, for example at least 75%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 91%, such as at least 92%, such as at least 93%, such as at least 94%, such as at least 95%, such as at least 96%, such as at least 97%, such as at least 98%, such as at least 99%, such as 100% sequence identity therewith. The sequence identity is preferably calculated as described herein below in the section "Sequence identity". A functional homologue of a syn-CPP is a polypeptide, which is also capable of catalysing reaction II described above.

The heterologous nucleic acid encoding said syn-CPP type diTPS may for example be any nucleic acid encoding any of the syn-CPP type diTPS described herein. Thus, the heterologous nucleic acid may be any nucleic acid encoding the polypeptide of SEQ ID NO:1 or any of functional homologue thereof described in the previous paragraph. For example, said heterologous nucleic acid may comprise or consist of SEQ ID NO:31.

ent-CPP type

The methods of the invention involve use of a diTPS of class II. The invention also features host organisms comprising a heterologous nucleic acid encoding a diTPS of class II. In one embodiment said diTPS of class II is an ent-CPP type diTPS. Such diTPS of class II are in particular useful in embodiments of the inventions, wherein the diterpenoid to be produced contains a 9R,10R decalin core.

As used herein the term “ent-CPP type diTPS” refers to any enzyme capable of catalysing the reaction III:



wherein PPO- refers to diphosphate.

In one embodiment the ent-CPP type diTPS may be EpTPS7. In particular, said ent-CPP type diTPS may be a polypeptide of SEQ ID NO:2 or a functional homologue thereof sharing at least 70%, such as at least 80%, for example at least 75%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 91%, such as at least 92%, such as at least 93%, such as at least 94%, such as at least 95%, such as at least 96%, such as at least 97%, such as at least 98%, such as at least 99%, such as 100% sequence identity therewith.

The heterologous nucleic acid encoding said ent-CPP type diTPS, may for example be any nucleic acid encoding any of the ent-CPP type diTPS described herein. Thus, the heterologous nucleic acid may be any nucleic acid encoding the polypeptide of SEQ ID NO:2 or any of functional homologue thereof described in the previous paragraph.

In another embodiment the ent-CPP type diTPS may be ZmAN2. In particular, said ent-CPP type diTPS may be a polypeptide of SEQ ID NO:3 or a functional homologue thereof sharing at least 70%, such as at least 80%, for example at least 75%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 91%, such as at least 92%, such as at least 93%, such as at least 94%, such as at least 95%, such as at least 96%, such as at least 97%, such as at least 98%, such as at least 99%, such as 100% sequence identity therewith.

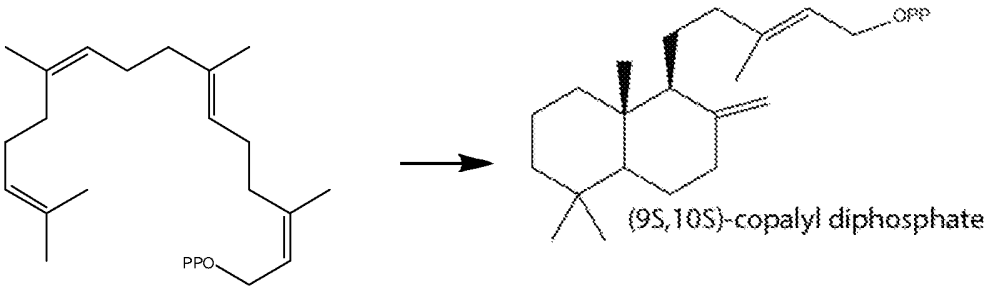
The heterologous nucleic acid may for example be any nucleic acid encoding the polypeptide of SEQ ID NO:3 or any of functional homologue thereof described in the previous paragraph.

The sequence identity is preferably calculated as described herein below in the section "Sequence identity". A functional homologue of an ent-CPP is a polypeptide, which is also capable of catalysing reaction III described above.

(+)-CPP type diTPS

The methods of the invention involve use of a diTPS of class II. The invention also features host organisms comprising a heterologous nucleic acid encoding a diTPS of class II. In one embodiment said diTPS of class II is a (+)-CPP type diTPS. Such diTPS of class II are in particular useful in embodiments of the inventions, wherein the diterpenoid to be produced contains a 9S,10S decalin core.

As used herein the term "(+)-CPP type diTPS" refers to any enzyme capable of catalysing the reaction IV:



wherein PPO- refers to diphosphate.

In one embodiment the (+)-CPP type diTPS may be TwTPS7. In particular, said (+)-CPP type diTPS may be a polypeptide of SEQ ID NO:4 or a functional homologue thereof sharing at least 70%, such as at least 80%, for example at least 75%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 91%, such as at least 92%, such as at least 93%, such as at least 94%, such as at least 95%, such as at least 96%, such as at least 97%, such as at least 98%, such as at least 99%, such as 100% sequence identity therewith.

The heterologous nucleic acid encoding said (+)-CPP type diTPS may for example be any nucleic acid encoding any of the (+)-CPP type diTPS described herein. Thus, the heterologous nucleic acid may be any nucleic acid encoding the polypeptide of SEQ ID NO:4 or any of functional homologue thereof described in the previous paragraph.

In another embodiment the (+)-CPP type diTPS may be CfTPS1. In particular, said (+)-CPP type diTPS may be a polypeptide of SEQ ID NO:5 or a functional homologue thereof sharing at least 70%, such as at least 80%, for example at least 75%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 91%, such as at least 92%, such as at least 93%, such as at least 94%, such as at least 95%, such as at least 96%, such as at least 97%, such as at least 98%, such as at least 99%, such as 100% sequence identity therewith.

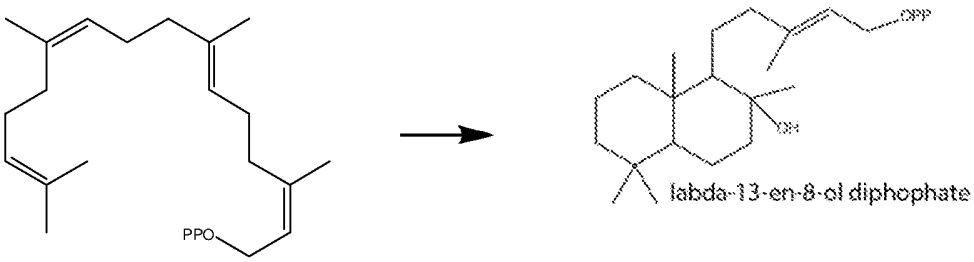
The heterologous nucleic acid may for example be any nucleic acid encoding the polypeptide of SEQ ID NO:5 or any of functional homologue thereof described in the previous paragraph. For example, said heterologous nucleic acid may comprise or consist of SEQ ID NO:27. Said heterologous nucleic acid may also comprise or consist of SEQ ID NO:28. This may in particular be the case in embodiments of the invention wherein the host organism is yeast, e.g. *S. cerevisiae*.

The sequence identity is preferably calculated as described herein below in the section "Sequence identity". A functional homologue of a (+)-CPP is a polypeptide, which is also capable of catalysing reaction IV described above.

LPP type diTPS

The methods of the invention involve use of a diTPS of class II. The invention also features host organisms comprising a heterologous nucleic acid encoding a diTPS of class II. In one embodiment said diTPS of class II is a LPP type diTPS. Such diTPS of class II are in particular useful in embodiments of the inventions, wherein the diterpenoid to be produced contains a 8-hydroxy-decalin core. However, LPP type diTPS may also be useful in other embodiments of the invention.

As used herein the term "LPP type diTPS" refers to any enzyme capable of catalysing the reaction V:



wherein PPO- refers to diphosphate.

In one embodiment the LPP type diTPS may be labda-13-en-8-ol pyrophosphate synthase, such as SsLPPS. In particular, said LPP type diTPS may be a polypeptide of SEQ ID NO:6 or a functional homologue thereof sharing at least 70%, such as at least 80%, for example at least 75%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 91%, such as at least 92%, such as at least 93%, such as at least 94%, such as at least 95%, such as at least 96%, such as at least 97%, such as at least 98%, such as at least 99%, such as 100% sequence identity therewith. In embodiments of the invention, wherein the diTPS of class II is SsLPPS or a functional homologue thereof sharing above mentioned sequence identity, then it may be preferred that the diTPS of class I is not SsSCS, CfTPS3, CfTPS4 or EpTPS8 or a functional homologue of any of the aforementioned sharing at least 70% sequence identity therewith. Thus, in embodiments of the invention, wherein the diTPS of class II is SsLPPS, then it is preferred that the diTPS of class I is not SsSCS, CfTPS3, CfTPS4 or EpTPS8.

The heterologous nucleic acid encoding said LPP type diTPS may for example be any nucleic acid encoding any of the LPP type diTPS described herein. Thus, the heterologous nucleic acid may be any nucleic acid encoding the polypeptide of SEQ ID NO:6 or any of functional homologue thereof described in the previous paragraph.

It is also preferred that if the diTPS of class II is SsCPSL, then it is preferred that the diTPS of class I is not SsKSL1 or SsKSL2.

In another embodiment the LPP type diTPS may be TwTPS21. In particular, said LPP type diTPS may be a polypeptide of SEQ ID NO:7 or a functional homologue thereof

sharing at least 70%, such as at least 80%, for example at least 75%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 91%, such as at least 92%, such as at least 93%, such as at least 94%, such as at least 95%, such as at least 96%, such as at least 97%, such as at least 98%, such as at least 99%, such as 100% sequence identity therewith.

The heterologous nucleic acid may for example be any nucleic acid encoding the polypeptide of SEQ ID NO:7 or any of functional homologue thereof described in the previous paragraph.

In another embodiment the LPP type diTPS may be CfTPS2. In particular, said LPP type diTPS may be a polypeptide of SEQ ID NO:17 or a functional homologue thereof sharing at least 70%, such as at least 80%, for example at least 75%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 91%, such as at least 92%, such as at least 93%, such as at least 94%, such as at least 95%, such as at least 96%, such as at least 97%, such as at least 98%, such as at least 99%, such as 100% sequence identity therewith. In embodiments of the invention, wherein the diTPS of class II is CfTPS2 or a functional homologue thereof sharing above mentioned sequence identity, then it is preferred that the diTPS of class I is not CfTPS3 or CfTPS4 or EpTPS8 or a functional homologue of the aforementioned sharing at least 70% sequence identity therewith. Thus, in embodiments of the invention, wherein the diTPS of class II is CfTPS2, then it is preferred that the diTPS of class I is not CfTPS3 or CfTPS4 or EpTPS8.

The heterologous nucleic acid may for example be any nucleic acid encoding the polypeptide of SEQ ID NO:17 or any of functional homologue thereof described in the previous paragraph.

The sequence identity is preferably calculated as described herein below in the section “Sequence identity”. A functional homologue of a LPP is a polypeptide, which is also capable of catalysing reaction V described above.

LPP like type diTPS

The methods of the invention involve use of a diTPS of class II. The invention also features host organisms comprising a heterologous nucleic acid encoding a diTPS of class II. In one embodiment said diTPS of class II is a LPP like type diTPS.

In one embodiment the LPP like type diTPS may be TwTPS14/28. In particular, said LPP like type diTPS may be a polypeptide of SEQ ID NO:8 or a functional homologue thereof sharing at least 70%, such as at least 80%, for example at least 75%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 91%, such as at least 92%, such as at least 93%, such as at least 94%, such as at least 95%, such as at least 96%, such as at least 97%, such as at least 98%, such as at least 99%, such as 100% sequence identity therewith.
The heterologous nucleic acid encoding said LPP like type diTPS may for exampe be any nucleic acid encoding any of the LPP like type diTPS described herein. Thus, the heterologous nucleic acid may be any nucleic acid encoding the polypeptide of SEQ ID NO:8 or any of functional homologue thereof described in the previous paragraph.

The sequence identity is preferably calculated as described herein below in the section “Sequence identity”.

diTPS of class I

The methods of the invention involve use of a diTPS of class I. The invention also features host organisms comprising a heterologous nucleic acid encoding a diTPS of class I.

Said diTPS of class I is an enzyme capable of catalyzing cleavage of the diphosphate group of the diterpene pyrophosphate intermediate and additionally preferably also is capable of catalysing cyclization and/or rearrangement reactions on the resulting carbocation. As with the class II diTPSs, deprotonation or water capture may terminate the class I diTPS reaction leading to hydroxylation of the diterpene pyrophosphate intermediate.

The diTPS of class I is generally a polypeptide sharing at least some sequence similarity to at least one of SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16 or SEQ ID NO:17. In particular, it is preferred

that the diTPS of class I shares at least 30%, preferably at least 40%, more preferably at least 45% sequence identity with at least one of SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16 and SEQ ID NO:17. In particular, it is preferred that the diTPS of class I shares at least 30%, such as at least 35% sequence identity to the sequence of ScSCS (SEQ ID NO:11) or to the sequence of AtEKS (see figure 4). Furthermore, it is preferred that the diTPS of class I in addition to above mentioned sequence identity also contains the following motif of five amino acids:

D-D-X-X-D/E,

wherein X may be any amino acid, such as any naturally occurring amino acids. In particular, X may be an amino acid with a hydrophobic side chain, and thus X may for example be selected from the group consisting of A, I, L, M, F, W, Y and V. Even more preferably X is an amino acid with a small hydrophobic side chain, and thus X may be selected from the group consisting of A, I, L and V.

In one embodiment of the invention said motif of five amino acids is:

D-D-F-F-D/E

D/E indicates that said amino acid may be D or E.

In particular, it is preferred that the diTPS of class I contains said motif in a position corresponding to position aa 329-333 of SsSCS of SEQ ID NO:11. A position corresponding to position aa 329-333 of SsSCS of SEQ ID NO:11 is identified by aligning the sequence of a diTPS of class I of interest to SEQ ID NO:11 and optionally to additional sequences of diTPS of class I as e.g. shown in figure 4, and identifying the amino acids of said diTPS of class I aligned with aa 329-333 of SsSCS of SEQ ID NO:11.

It is furthermore preferred that in addition to sharing above mentioned sequence identity and containing said motif, then as many as possible of the amino acids marked with a black box in figure 4 are retained. Thus, when aligned to the sequence of ScSCS (SEQ ID NO:11), then preferably the diTPS of class I also contains at least 80%, more preferably at least 90%, for example at least 95%, such as all of the amino acids marked by a black box in figure 4. Alternatively, when aligned to the sequence of sequence of AtEKS (see figure 4), then preferably the diTPS of class I also contains at

least 80%, more preferably at least 90%, for example at least 95%, such as all of the amino acids marked by a black box in figure 4.

Thus, the diTPS of class I may for example be selected from the group consisting of diTPS of class I of the following types:

i. EpTPS8 like diTPS, such as any of the enzymes described herein below in the section "EpTPS8"
ii. EpTPS23 like diTPS, such as any of the enzymes described herein below in the section "EpTPS23"
iii. SsSCS like diTPS, such as any of the enzymes described herein below in the section "SsSCS"
iv. CfTPS3 like diTPS, such as any of the enzymes described herein below in the section "CfTPS3"
v. CfTPS4 like diTPS, such as any of the enzymes described herein below in the section "CfTPS4"
vi. TwTPS2 like diTPS, such as any of the enzymes described herein below in the section "TwTPS2"
vii. EpTPS1 like diTPS, such as any of the enzymes described herein below in the section "TwTPS1"
viii. CfTPS14 like diTPS, such as any of the enzymes described herein below in the section "CfTPS14"

The diTPS of class I may in one embodiment also be MvTPS5 like diTPS, such as any of the enzymes described herein below in the section "MvTPS5".

In one embodiment of the invention the diTPS of class I may be a TPS from *Tripterygium Wilfordii.* For example, the diTPS of class I may be the polypeptide of SEQ ID NO:41 or a functional homologue thereof sharing at least 70%, such as at least 80%, for example at least 75%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 91%, such as at least 92%, such as at least 93%, such as at least 94%, such as at least 95%, such as at least 96%, such as at least 97%, such as at least 98%, such as at least 99%, such as 100% sequence identity therewith.

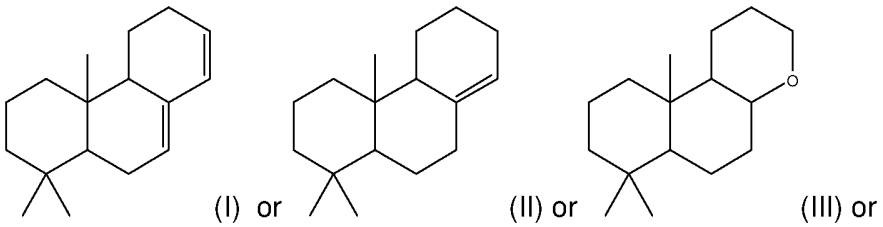
The heterologous nucleic acid encoding said diTPS of *Tripterygium Wilfordii*, may for example be any nucleic acid encoding any of the diTPS of *Tripterygium Wilfordii*

described herein. Thus, the heterologous nucleic acid may be any nucleic acid encoding the polypeptide of SEQ ID NO:41 or any of functional homologue thereof described in the previous paragraph. For example the heterologous nucleic acid may be a nucleic acid of SEQ ID NO:45.

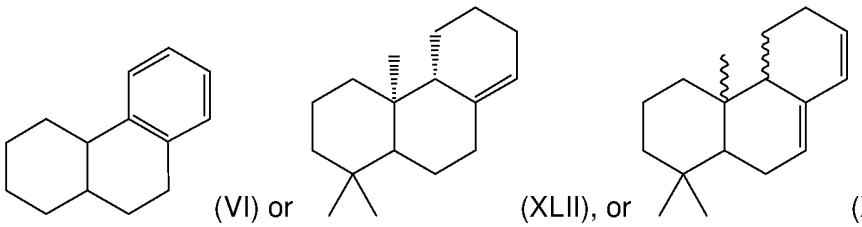
EpTPS8

The invention involves use of a diTPS of class I. In one embodiment said diTPS of class I may be an EpTPS8 like diTPS. In embodiments of the invention, wherein the diTPS of class I is a EpTPS8 like diTPS, then it may be preferred that the diTPS of class II is not CfTPS2 or SsLPPS or a functional homologue of any of the aforementioned sharing at least 70% sequence identity therewith. Thus, in embodiments of the invention, wherein the diTPS of class I is EpTPS8, then it may be preferred that the diTPS of class II is not CfTPS2 or SsLPPS.

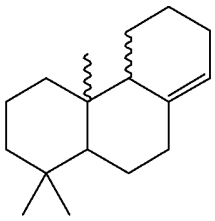
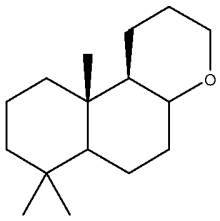
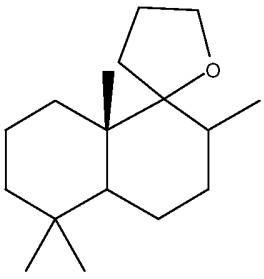
In particular, said diTPS of class I may be an EpTPS8 like diTPS in embodiments of the invention, wherein the diterpene to be produced contains a tricyclic ring structure. For example said diTPS of class I may be and EpTPS8 like diTPS in embodiments of the invention, wherein the diterpene to be produced contains a core of any of the formulas I, II, III, XLI, XLII, XLIII, XLIV or XLV:



(I) or (II) or (III) or



(VI) or (XLII), or (XLIII), or


(XLIV), or 
(XLV) or 
(XLI)..

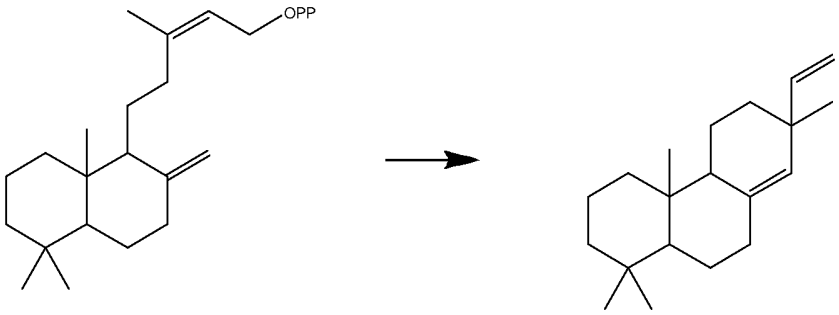
The waved line " ≀ " as used herein indicates a bond of undefined stereochemistry, i.e. the bond may be either a " | " or " ⁞ ".

Dependent on the structure of the diterpene pyrophosphate intermediate then the diterpene containing a core of formula I or II may have different stereochemistry. In general the stereochemistry of the decalin core present in the diterpene pyrophosphate intermediate is maintained after the reaction catalysed by a EpTPS8 like diTPS.

The EpTPS8 like diTPS may be any enzyme capable of catalysing the reaction VII:

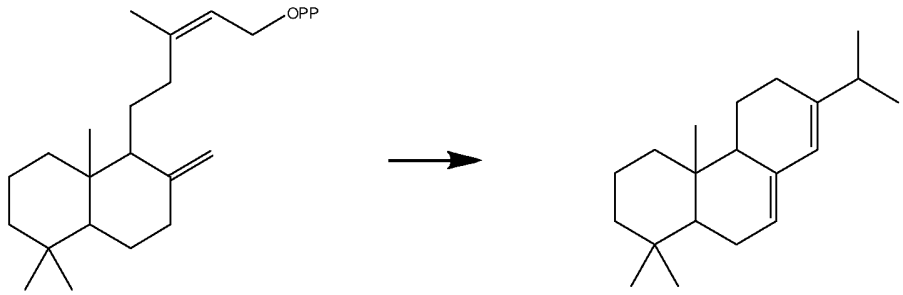
Diterpene pyrophosphate intermediate containing a decalin core structure ⟶ Diterpene containing a core structure of formula I or formula II or formula III or formula VI.

In particular EpTPS8 like diTPS may be an enzyme catalysing the reaction VIII:



wherein -OPP indicates diphosphate. During reaction VIII the produced diterpene will in general maintain the stereochemistry around the decalin core found in the starting diterpene pyrophosphate intermediate.

The EpTPS8 like diTPS may also be an enzyme catalysing the reaction IX:

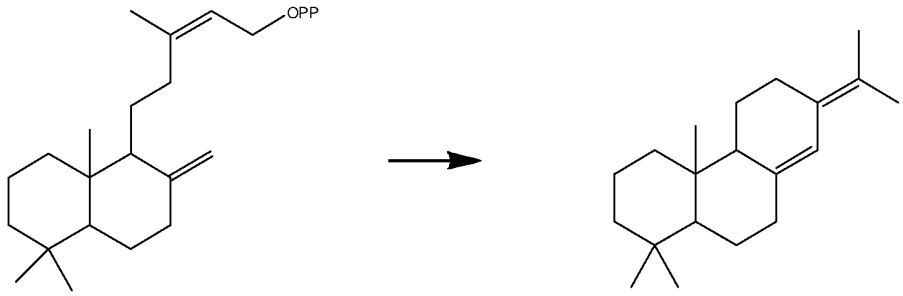


wherein OPP indicated diphosphate. During reaction IX the produced diterpene will in general maintain the stereochemistry around the decalin core found in the starting diterpene pyrophosphate intermediate.

The EpTPS8 like diTPS may also be an enzyme catalysing the reaction X:

wherein -OPP indicated diphosphate. During reaction X the produced diterpene will in general maintain the stereochemistry around the decalin core found in the starting diterpene pyrophosphate intermediate.

In particular, the EpTPS8 like diTPS may be an enzyme catalysing the reaction XXV:



wherein -OPP indicates diphosphate. During reaction XXV the produced diterpene will in general maintain the stereochemistry around the decalin core found in the starting diterpene pyrophosphate intermediate.

In one embodiment EpTPS8 like diTPS may be a terpene synthase from *Euphobia peplus*, and in particular it may be TPS8 from *Euphobia peplus*. TPS8 from *Euphobia peplus* is also referred to as EpTPS herein. In particular, said EpTPS8 like diTPS may be a polypeptide of SEQ ID NO:9 or a functional homologue thereof sharing at least 70%, such as at least 80%, for example at least 75%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 91%, such as at least 92%, such as at least 93%, such as at least 94%, such as at least 95%, such as at least 96%, such as at least 97%, such as at least 98%, such as at least 99%, such as 100% sequence identity therewith.

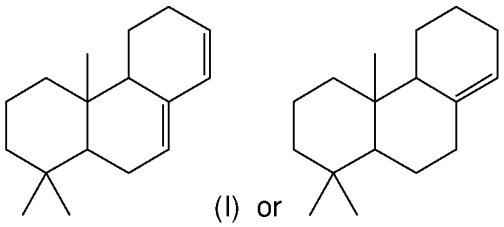
The heterologous nucleic acid encoding said EpTPS8 like diTPS may for example be any nucleic acid encoding any of the EpTPS8 like diTPS described herein. Thus, the heterologous nucleic acid may for example be any nucleic acid encoding the polypeptide of SEQ ID NO:9 or any of functional homologue thereof described in the previous paragraph.

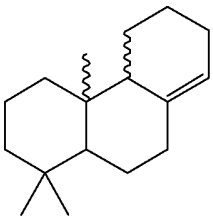
The sequence identity is preferably calculated as described herein below in the section "Sequence identity". A functional homologue of EpTPS8 is a polypeptide, which is also capable of catalysing at least one of reactions VII, VIII, IX, X and XXV described above.

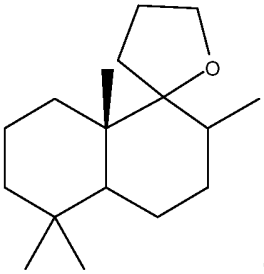
EpTPS23

The invention involves use of a diTPS of class I. In one embodiment said diTPS of class I may be an EpTPS23 like diTPS.

In particular, said diTPS of class I may be an EpTPS23 like diTPS in embodiments of the invention, wherein the diterpene to be produced contains a tricyclic ring structure. For example said diTPS of class I may be an EpTPS23 like diTPS in embodiments of the invention, wherein the diterpene to be produced contains a core of any of the formulas I, II, XLIV or XLI:


(I) or

(II). or

(XLIV) or



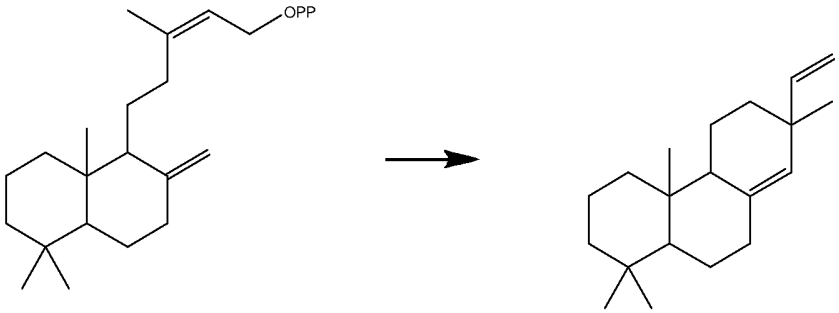
(XLI).

Dependent on the structure of the diterpene pyrophosphate intermediate then the diterpene containing a core of formula I or II may have different stereochemistry. In general the stereochemistry of the decalin core present in the diterpene pyrophosphate intermediate is maintained after the reaction catalysed by an EpTPS23 like diTPS.

The EpTPS23 like diTPS may in particular be an enzyme capable of catalysing the reaction XI:

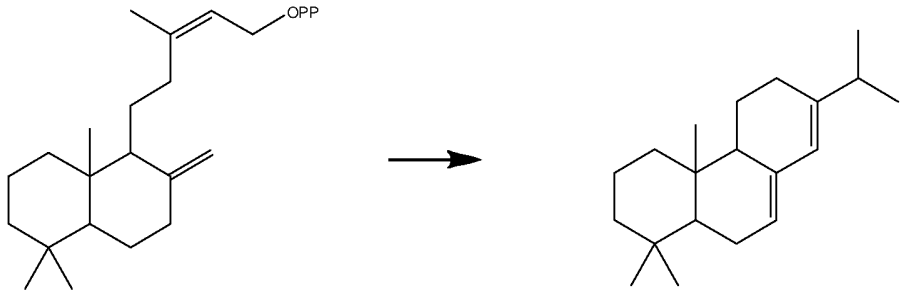
Diterpene pyrophosphate intermediate containing a decalin core structure ⟶ Diterpene containing a core structure of formula I or formula II

In particular an EpTPS23 like diTPS may be an enzyme catalysing the reaction VIII:



wherein -OPP indicated diphosphate. During reaction VIII the produced diterpene will in general maintain the stereochemistry around the decalin core found in the starting diterpene pyrophosphate intermediate.

The EpTPS23 like diTPS may also be an enzyme catalysing the reaction IX:



wherein -OPP indicated diphosphate. During reaction IX the produced diterpene will in general maintain the stereochemistry around the decalin core found in the starting diterpene pyrophosphate intermediate.

In one embodiment an EpTPS23 like diTPS may be a diterpene synthase from *Euphobia peplus.* In particular, the EpTPS23 like diTPS may be TPS23 of *Euphobia peplus*. TPS23 of *Euphobia peplus* may also be referred to as EpTPS23 herein. In particular, said EpTPS23 like diTPS may be a polypeptide of SEQ ID NO:10 or a functional homologue thereof sharing at least 70%, such as at least 80%, for example at least 75%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 91%, such as at least 92%, such as at least 93%, such as at least 94%, such as at least 95%, such as at least 96%, such as at least 97%, such as at least 98%, such as at least 99%, such as 100% sequence identity therewith.

The heterologous nucleic acid encoding said EpTPS23 like diTPS may for example be any nucleic acid encoding any of the EpTPS23 like diTPS described herein. Thus, the heterologous nucleic acid may be any nucleic acid encoding the polypeptide of SEQ ID NO:10 or any of functional homologue thereof described in the previous paragraph.

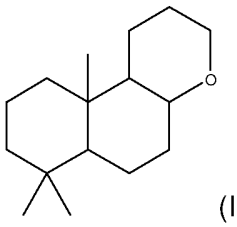
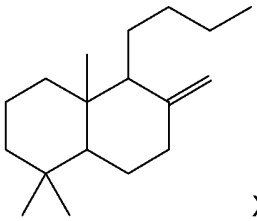
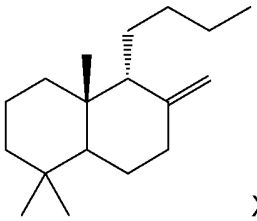
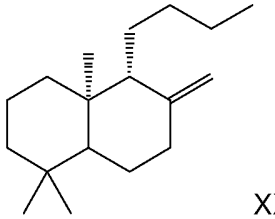
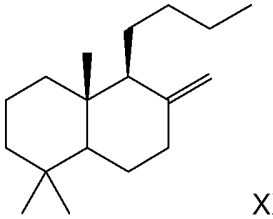
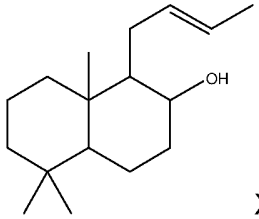
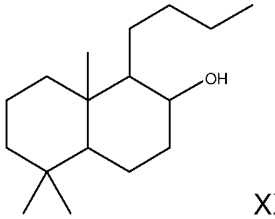
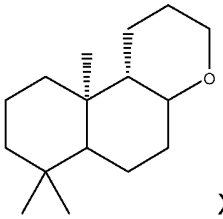
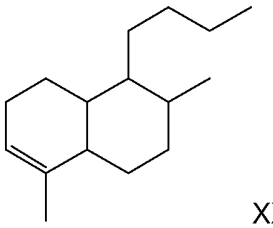
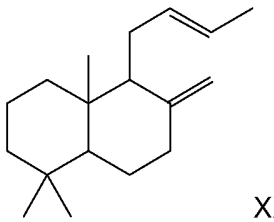
The sequence identity is preferably calculated as described herein below in the section “Sequence identity”. A functional homologue of EpTPS23 is a polypeptide, which is also capable of catalysing at least one of reactions VIII or IX described above.

SsSCS

The invention involves use of a diTPS of class I. In one embodiment said diTPS of class I may be a SsSCS like diTPS.

In particular, said diTPS of class I may be a SsSCS like diTPS in embodiments of the invention, wherein the diterpene to be produced contains a decalin substituted at the 10 position with C_5-alkenyl chain, which optionally may be substituted with a hydroxyl and/or a methyl group and/or =C.

Furthermore, said diTPS of class I may be a SsSCS like diTPS in embodiments of the invention, wherein the diterpene to be produced contains a core of formula III, XLVI, XXVII, XXVIII, XXIX, XXX, XXXI, XXXII, XXXIII, or XXXIV:


(III), or

XLVI, or

XXVII, or

XXVIII, or

XXIX, or

XXX, or

XXXI, or

XXXII, or

XXXIII, or

XXXIV

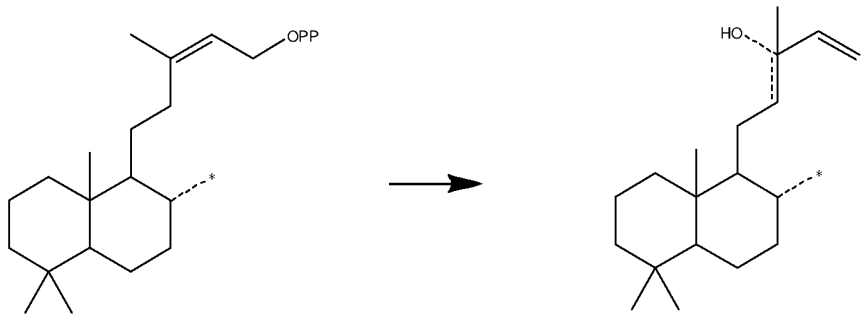
Dependent on the structure of the diterpene pyrophosphate intermediate then the diterpene containing a decalin substituted at the 10 position with said C_5-alkenyl chain,

or the diterpene containing a core of formula III may have different stereochemistry. In general the stereochemistry of the decalin core present in the diterpene pyrophosphate intermediate is maintained after the reaction catalysed by a SsSCS like diTPS.

The SsSCS like diTPS may be any enzyme capable of catalysing the following reaction XII:

Diterpene pyrophosphate intermediate containing a decalin core structure ⟶ Diterpene containing a decalin core substituted at the 10 position with C_5-alkenyl chain, which optionally may be substituted with a hydroxyl and/or a methyl group and/or =C OR diterpene containing a core structure of formula III.

The SsSCS like diTPS may in particular be an enzyme capable of catalysing the reaction XVI:



wherein -OPP is diphosphate; and

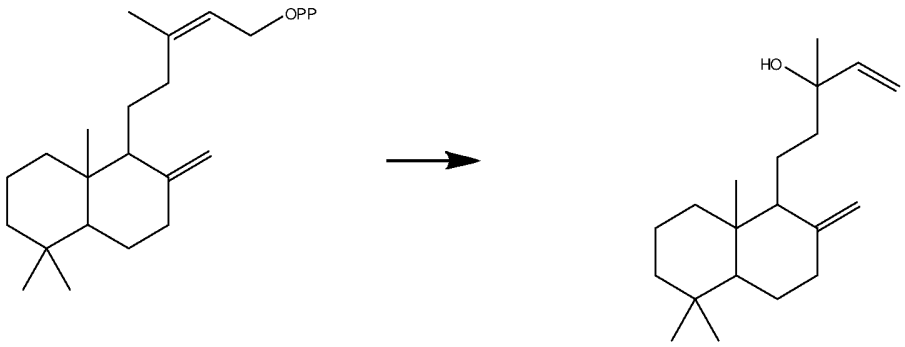
 indicates either a double bond or two single bonds, wherein one is substituted with –OH and the other with $–CH_3$; and

the dotted lines without star indicates a bond, which optionally is present.

Thus, may be  or  .

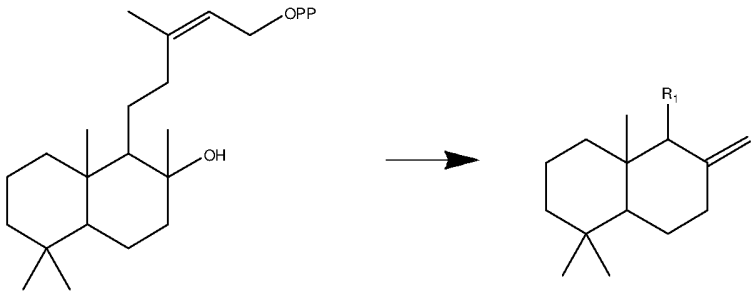
It is to be understood that in embodiments of the invention, wherein the dotted line shown as  is not present, then also the hydroxyl group is not present. It is preferred that one and only one of the dotted lines without star indicates a bond.

A SsSCS like diTPS may in particular be an enzyme capable of catalysing the reaction XVII:



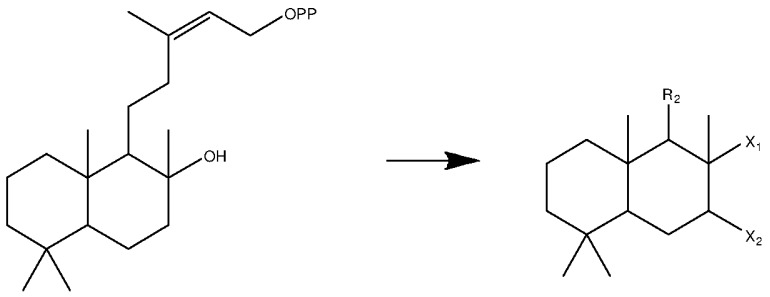
wherein OPP indicated diphosphate. During reaction XVII the produced diterpene will in general maintain the stereochemistry around the decalin core found in the starting diterpene pyrophosphate intermediate. Thus, the SsSCS like diTPS may be an enzyme catalysing any of the reactions XIII, XIV and XV shown in figure 1.

The SsSCS like diTPS may also be an enzyme catalysing the following reaction XXVIII:



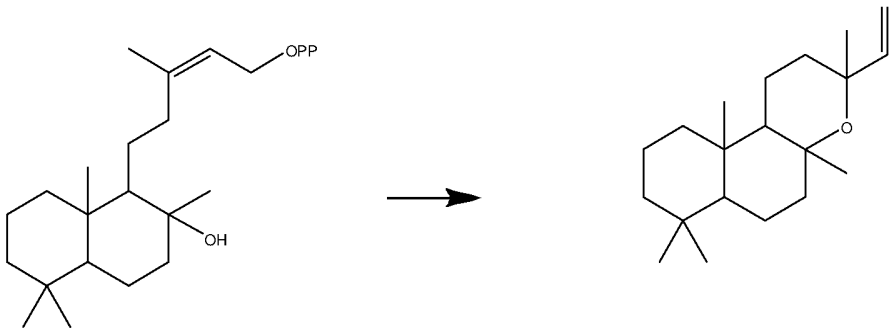
wherein OPP is diphosphate and R_1 is a C_5-alkenyl substituted with methyl and/or hydroxyl. Preferably, R_1 is C_5-alkenyl containing one or two double bonds. When R_1 is alkenyl containing one double bond, said alkenyl is preferably substituted with hydroxyl and methyl. When R_1 is alkenyl containing two double bonds, said alkenyl is preferably substituted with methyl.

The SsSCS like diTPS may also be an enzyme catalysing the following reaction XXIX:



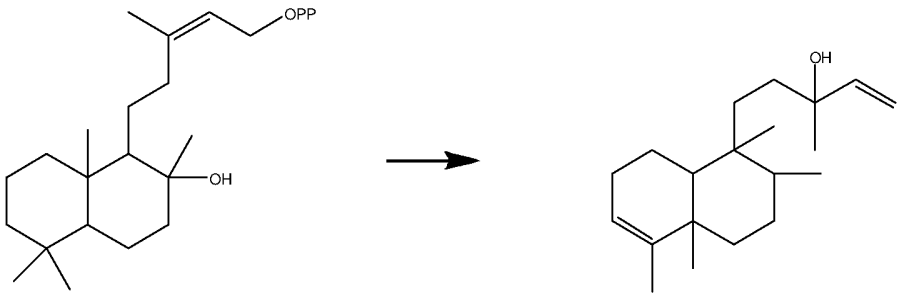
wherein -OPP is diphosphate and R_2 is a C_5-alkenyl substituted with methyl and/or hydroxyl or with =C, and X_1 is either –OH or methyl, and X_2 is either –H or –OH, wherein one and only one of X_1 and X_2 is -OH. Preferably, R_2 is C_5-alkenyl containing one or two double bonds. When R_2 is alkenyl containing one double bond, said alkenyl is preferably substituted with hydroxyl and methyl or with =C. When R_2 is alkenyl containing two double bonds, said alkenyl is preferably substituted with methyl.

The SsSCS like diTPS may also be an enzyme catalysing the reaction X:



wherein OPP indicates diphosphate. During reaction X the produced diterpene will in general maintain the stereochemistry around the decalin core found in the starting diterpene pyrophosphate intermediate.

The SsSCS like diTPS may also be an enzyme catalysing the reaction LII:



wherein OPP indicates diphosphate.

In one embodiment a SsSCS like diTPS may be SClareol Synthase (SCS) from *Salvia Sclarea*. SCS from *Salvia Sclarea* may also be referred to as SsSCS herein. In particular, said SsSCS like diTPS may be a polypeptide of SEQ ID NO:11 or a functional homologue thereof sharing at least 70%, such as at least 80%, for example

at least 75%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 91%, such as at least 92%, such as at least 93%, such as at least 94%, such as at least 95%, such as at least 96%, such as at least 97%, such as at least 98%, such as at least 99%, such as 100% sequence identity therewith.

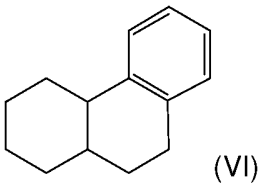
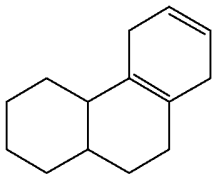
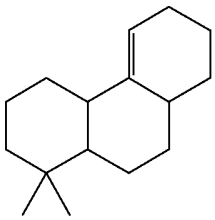
The heterologous nucleic acid encoding said SsSCS like diTPS may for example be any nucleic acid encoding any of the SsSCS like diTPS described herein. Thus, the heterologous nucleic acid may for example be any nucleic acid encoding the polypeptide of SEQ ID NO:11 or any of functional homologue thereof described in the previous paragraph. For example, said heterologous nucleic acid may comprise or consist of SEQ ID NO:32. Said heterologous nucleic acid may also comprise or consist of SEQ ID NO:33. This may in particular be the case in embodiments of the invention wherein the host organism is yeast, e.g. *S. cerevisiae*.

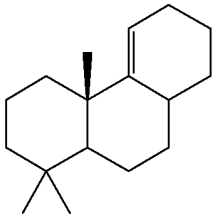
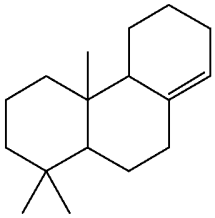
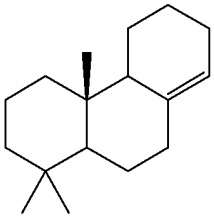
The sequence identity is preferably calculated as described herein below in the section "Sequence identity". A functional homologue of SsSCS is a polypeptide, which is also capable of catalysing at least one of reactions XII, XIII, XIV, XV, XVI, XVII, XXVIII or XXIX or LII described above.

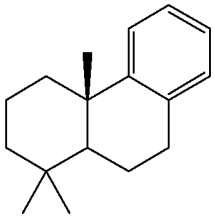
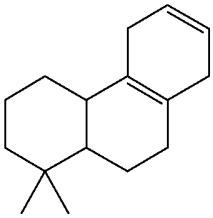
CfTPS3

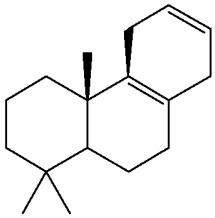
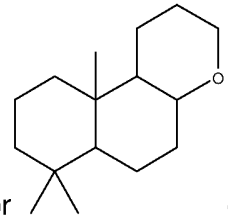
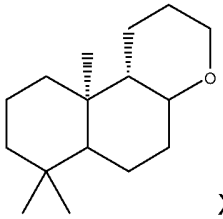
The invention involves use of a diTPS of class I. In one embodiment said diTPS of class I may be a CfTPS3 like diTPS. In embodiments of the invention, wherein the diTPS of class I is a CfTPS3 like diTPS, then it may be preferred that the diTPS of class II is not CfTPS2 or SsLPPS or a functional homologue thereof sharing at least 70% sequence identity therewith. Thus, in embodiments of the invention, wherein the diTPS of class I is CfTPS3, then it may be preferred that the diTPS of class II is not CfTPS2 or SsLPPS.

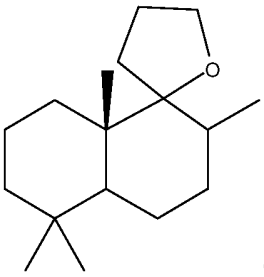
In particular, said diTPS of class I may be a CfTPS3 like diTPS in embodiments of the invention, wherein the diterpene to be produced contains a tricyclic ring structure. For example said diTPS of class I may be a CFTPS3 like diTPS in embodiments of the invention, wherein the diterpene to be produced contains a core of any of the formulas VI, IX XXXV, XXXVI, II, XXXVII, XXXVIII, XXXIX, XL, III or XXXII:


(VI) or

(IX), or

(XXXV) or


(XXXVI) or

(II) or

(XXXVII) or


(XXXVIII), or

(XXXIX) or


(XL) or

(III) or

XXXII or


(XLI).

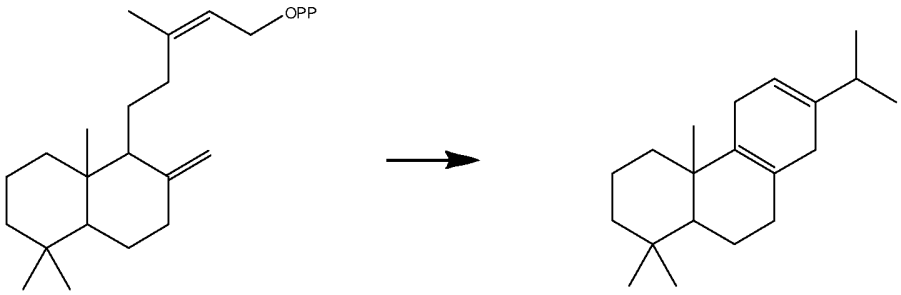
Dependent on the structure of the diterpene pyrophosphate intermediate then the diterpene containing a core of formula VI, IX, XXXV, II or XXXIX may have different stereochemistry. In general the stereochemistry of the decalin core present in the diterpene pyrophosphate intermediate is maintained after the reaction catalysed by the CfTPS3 like diTPS.

The CfTPS3 like diTPS may be any enzyme capable of catalysing the reaction XXIII:

Diterpene pyrophosphate intermediate containing a decalin core structure ⟶

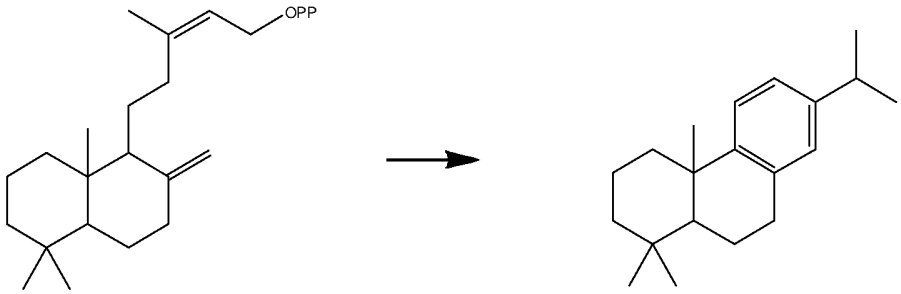
Diterpene containing a core structure of formula VI or formula IX, XXXV, XXXVI, II,XXXVII, XXXVIII, XXXIX, XL, III or XXXII.

The CfTPS3 like diTPS may in particular be an enzyme capable of catalysing the reaction XXIV:



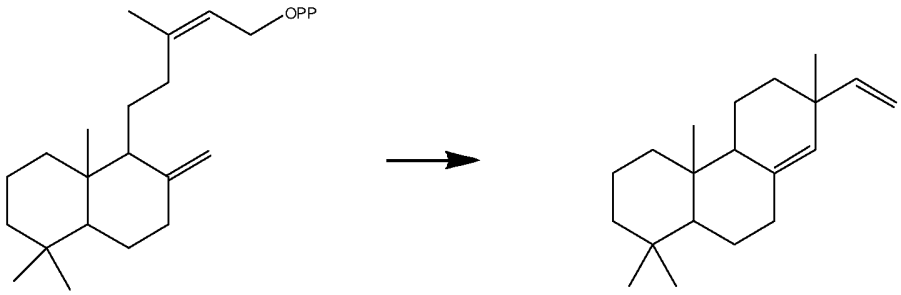
wherein OPP indicates diphosphate. During reaction XXIV the produced diterpene will in general maintain the stereochemistry around the decalin core found in the starting diterpene pyrophosphate intermediate.

The CfTPS3 like diTPS may in particular be an enzyme capable of catalysing the reaction XXII:



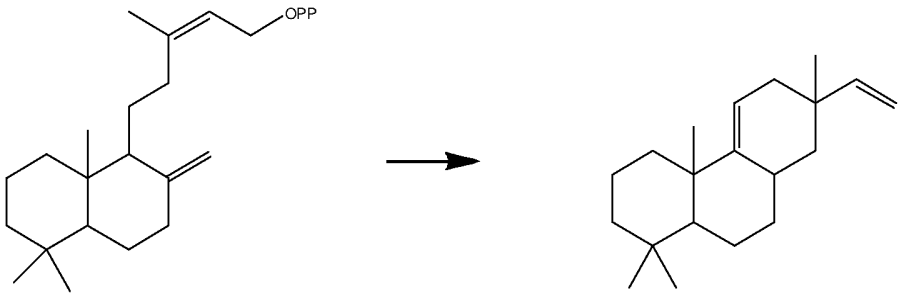
wherein OPP is diphosphate. During reaction XXII the produced diterpene will in general maintain the stereochemistry around the decalin core found in the starting diterpene pyrophosphate intermediate.

The CfTPS3 like diTPS may in particular be an enzyme capable of catalysing the reaction LIII:



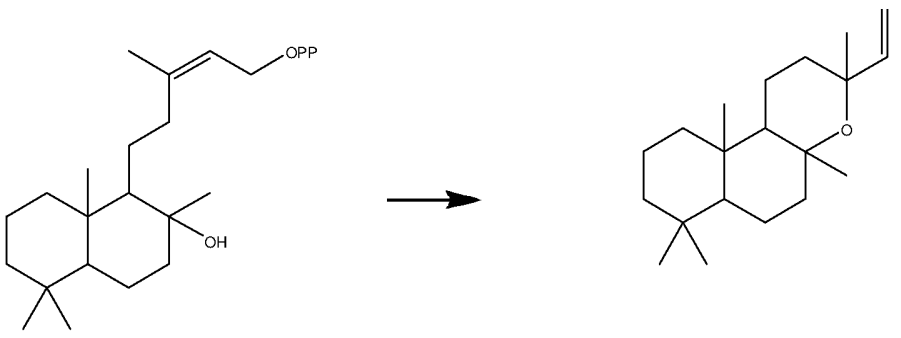
wherein OPP is diphosphate. During reaction LIII the produced diterpene will in general maintain the stereochemistry around the decalin core found in the starting diterpene pyrophosphate intermediate.

The CfTPS3 like diTPS may in particular be an enzyme capable of catalysing the reaction LIV:



wherein OPP is diphosphate. During reaction LIV the produced diterpene will in general maintain the stereochemistry around the decalin core found in the starting diterpene pyrophosphate intermediate.

The CfTPS3 like diTPS may also be an enzyme catalysing the reaction X:



wherein OPP indicates diphosphate. During reaction X the produced diterpene will in general maintain the stereochemistry around the decalin core found in the starting diterpene pyrophosphate intermediate.

In one embodiment the CfTPS3 like diTPS may be a diterpene synthase from *Coleus forskohlii*. In particular, the CfTPS3 like diTPS may be a TPS3 from *Coleus forskohlii*. TPS3 from *Coleus forskohlii* may also be referred to as CfTPS3. In particular, said CfTPS3 like diTPS may be a polypeptide of SEQ ID NO:12 or a functional homologue thereof sharing at least 70%, such as at least 80%, for example at least 75%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 91%, such as at least 92%, such as at least 93%, such as at least 94%, such as at least 95%, such as at least 96%, such as at least 97%, such as at least 98%, such as at least 99%, such as 100% sequence identity therewith.

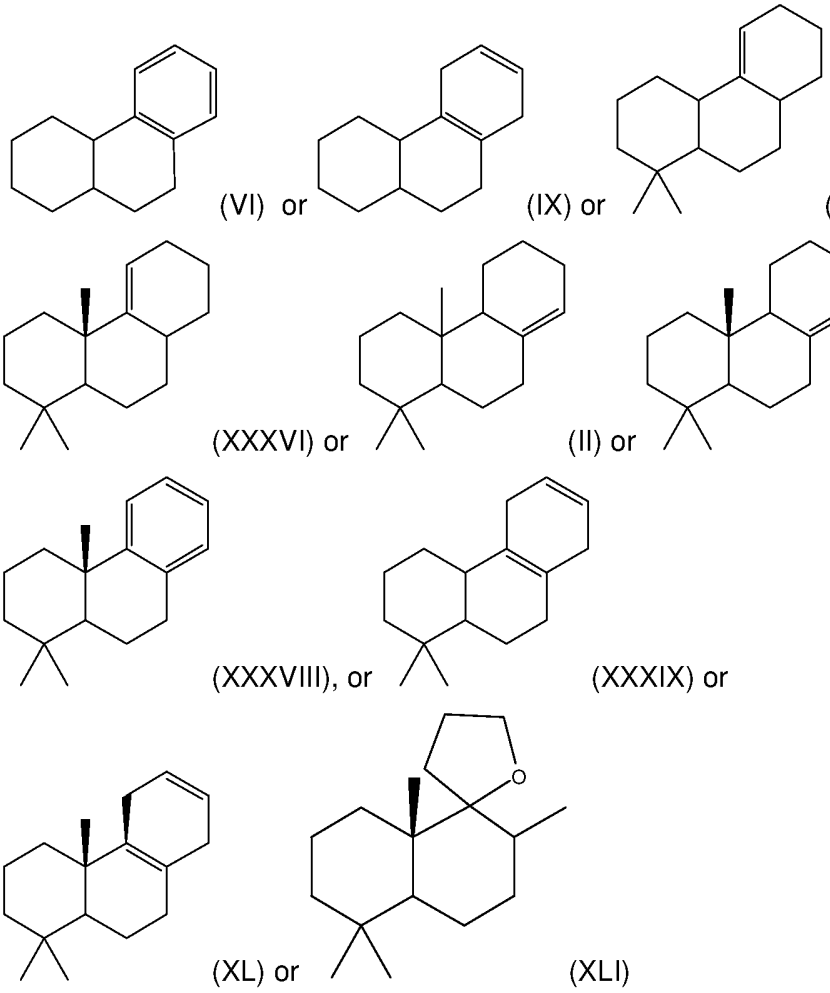
The heterologous nucleic acid encoding said CfTPS3 like diTPS may for example be any nucleic acid encoding any of the CfTPS3 like diTPS described herein. Thus, the heterologous nucleic acid may for example be any nucleic acid encoding the polypeptide of SEQ ID NO:12 or any of functional homologue thereof described in the previous paragraph. For example, said heterologous nucleic acid may comprise or consist of SEQ ID NO:29. Said heterologous nucleic acid may also comprise or consist of SEQ ID NO:30. This may in particular be the case in embodiments of the invention wherein the host organism is yeast, e.g. *S. cerevisiae*.

The sequence identity is preferably calculated as described herein below in the section "Sequence identity". A functional homologue of CfTPS3 is a polypeptide, which is also capable of catalysing at least one of reactions XXII, XXIII or XXIV described above.

CfTPS4

The invention involves use of a diTPS of class I. In one embodiment said diTPS of class I may be a CfTPS4 like diTPS. In embodiments of the invention, wherein the diTPS of class I is a CfTPS4 like diTPS, then it may be preferred that the diTPS of class II is not CfTPS2 or SsLPPSor a functional homologue of any of the aforementioned sharing at least 70% sequence identity therewith. Thus, in embodiments of the invention, wherein the diTPS of class I is CfTPS4, then it is preferred that the diTPS of class II is not CfTPS2 or SsLPPS.

In particular, said diTPS of class I may be a CfTPS4 like diTPS in embodiments of the invention, wherein the diterpene to be produced contains a tricyclic ring structure. For example said diTPS of class I may be a CfTPS4 like diTPS in embodiments of the invention, wherein the diterpene to be produced contains a core of any of the formulas VI, IX, XXXV, XXXVI, II, XXXVII, XXXVIII, XXXIX or XL:

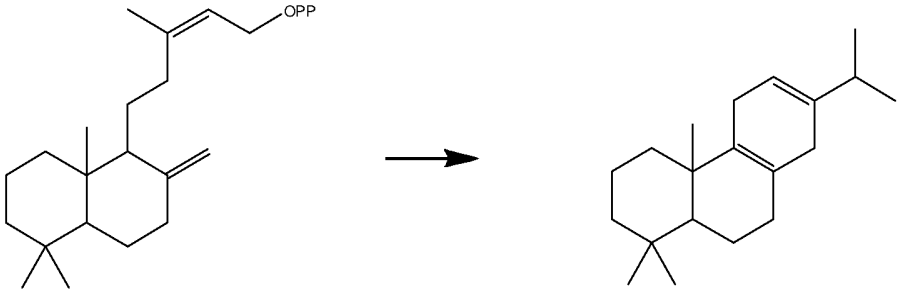


Dependent on the structure of the diterpene pyrophosphate intermediate then the diterpene containing a core of formula VI, IX, XXXV, II, or XXXIX may have different stereochemistry. In general the stereochemistry of the decalin core present in the diterpene pyrophosphate intermediate is maintained after the reaction catalysed by the CfTPS4 like diTPS.

The CfTPS4 like diTPS may be any enzyme capable of catalysing the reaction XXIII:

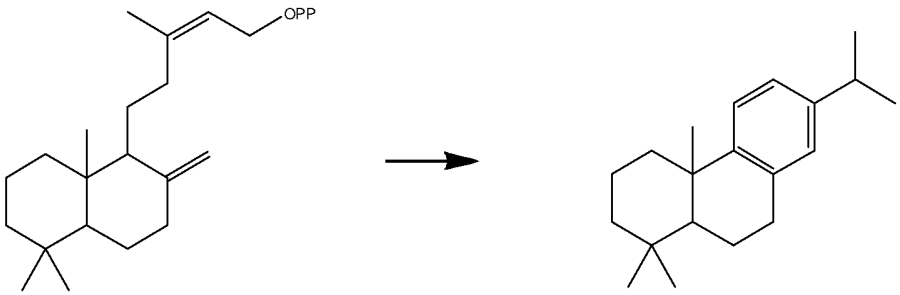
Diterpene pyrophosphate intermediate containing a decalin core structure ⟶ Diterpene containing a core structure of formula VI or formula IX, XXXV, XXXVI, II, XXXVII, XXXVIII, XXXIX or XL.

The CfTPS4 like diTPS may in particular be an enzyme capable of catalysing the reaction XXIV:



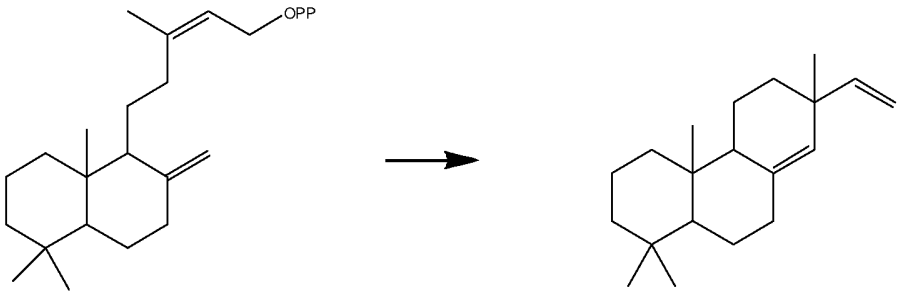
wherein OPP indicates diphosphate. During reaction XXIV the produced diterpene will in general maintain the stereochemistry around the decalin core found in the starting diterpene pyrophosphate intermediate.

The CfTPS4 like diTPS may in particular be an enzyme capable of catalysing the reaction XXII:



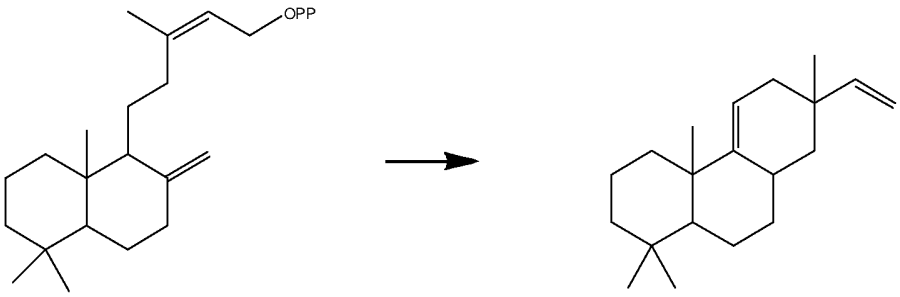
wherein OPP is diphosphate. During reaction XXII the produced diterpene will in general maintain the stereochemistry around the decalin core found in the starting diterpene pyrophosphate intermediate.

The CfTPS4 like diTPS may in particular be an enzyme capable of catalysing the reaction LIII:



wherein OPP is diphosphate. During reaction LIII the produced diterpene will in general maintain the stereochemistry around the decalin core found in the starting diterpene pyrophosphate intermediate.

The CfTPS4 like diTPS may in particular be an enzyme capable of catalysing the reaction LIV:



wherein OPP is diphosphate. During reaction LIV the produced diterpene will in general maintain the stereochemistry around the decalin core found in the starting diterpene pyrophosphate intermediate.

In one embodiment the CfTPS4 like diTPS may be a diterpene synthase from *Coleus forskohlii*. In particular, the CfTPS4 like diTPS may be a TPS4 from *Coleus forskohlii*. TPS4 from *Coleus forskohlii* may also be referred to as CfTPS4. In particular, said CfTPS4 like diTPS may be a polypeptide of SEQ ID NO:13 or a functional homologue thereof sharing at least 70%, such as at least 80%, for example at least 75%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 91%, such as at least 92%, such as at least 93%, such as at least 94%, such as at least 95%, such as at least 96%, such as at least 97%, such as at least 98%, such as at least 99%, such as 100% sequence identity therewith.

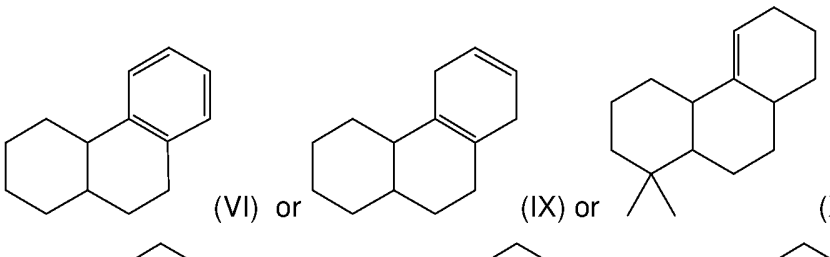
The heterologous nucleic acid encoding said CfTPS4 like diTPS may for example be any nucleic acid encoding any of the CfTPS4 like diTPS described herein. Thus, the heterologous nucleic acid may for example be any nucleic acid encoding the polypeptide of SEQ ID NO:13 or any of functional homologue thereof described in the previous paragraph.

The sequence identity is preferably calculated as described herein below in the section "Sequence identity". A functional homologue of CfTPS4 is a polypeptide, which is also capable of catalysing at least one of reactions XXII, XXIII or XXIV described above.

MvTPS5

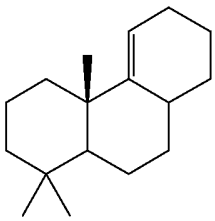
The invention involves use of a diTPS of class I. In one embodiment said diTPS of class I may be a MvTPS5 like diTPS.

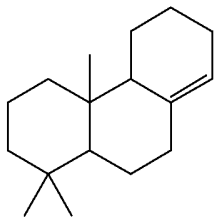
In particular, said diTPS of class I may be a MvTPS5 like diTPS in embodiments of the invention, wherein the diterpene to be produced contains a tricyclic ring structure. For example said diTPS of class I may be a MvTPS5 like diTPS in embodiments of the invention, wherein the diterpene to be produced contains a core of any of the formulas VI, IX, XXXV, XXXVI, II,XXXVII, XXXVIII, XXXIX, XL, III or XXXII:

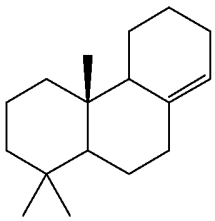

(VI) or

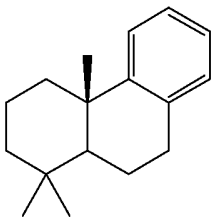
(IX) or

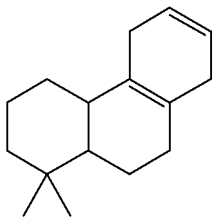
(XXXV) or

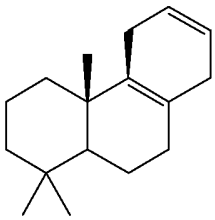
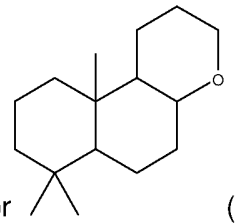
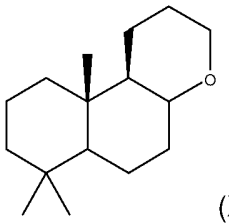
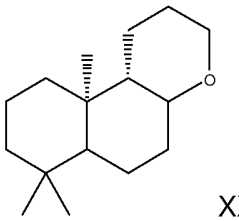

(XXXVI) or


(II) or


(XXXVII) or


(XXXVIII), or


(XXXIX) or

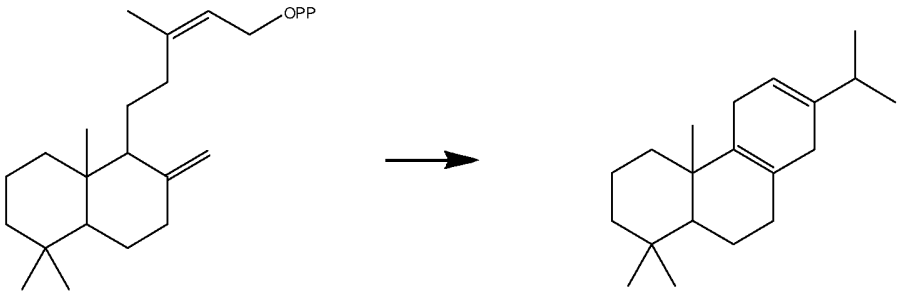

(XL) or

(III) or

(XXV), or

XXXII.

Dependent on the structure of the diterpene pyrophosphate intermediate then the diterpene containing a core of formula VI, IX, XXXV, II, XXXIX or III, may have different stereochemistry. In general the stereochemistry of the decalin core present in the diterpene pyrophosphate intermediate is maintained after the reaction catalysed by the MvTPS5 like diTPS.

The MvTPS5 like diTPS may be any enzyme capable of catalysing the reaction XXIII:

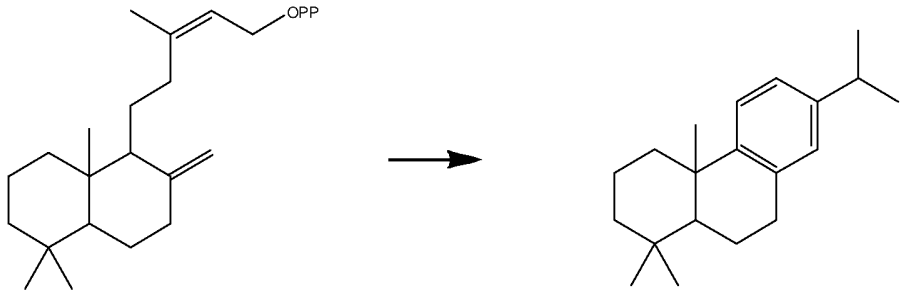
Diterpene pyrophosphate intermediate containing a decalin core structure ⟶ Diterpene containing a core structure of formula VI, IX, XXXV, XXXVI, II,XXXVII, XXXVIII, XXXIX, XL, III or XXXII.

The MvTPS5 like diTPS may in particular be an enzyme capable of catalysing the reaction LV:



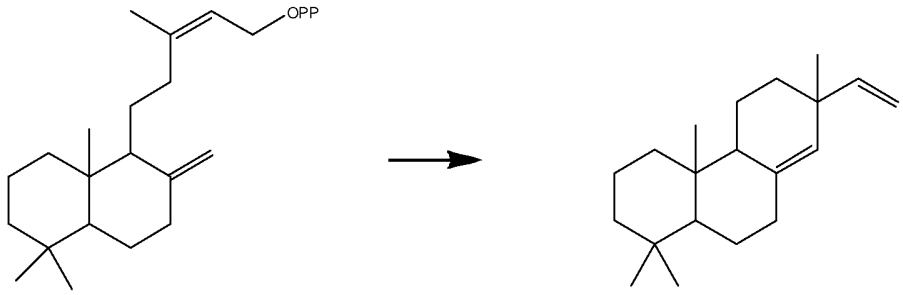
wherein OPP indicates diphosphate. During reaction LV the produced diterpene will in general maintain the stereochemistry around the decalin core found in the starting diterpene pyrophosphate intermediate.

The MvTPS5 like diTPS may in particular be an enzyme capable of catalysing the reaction LVI:



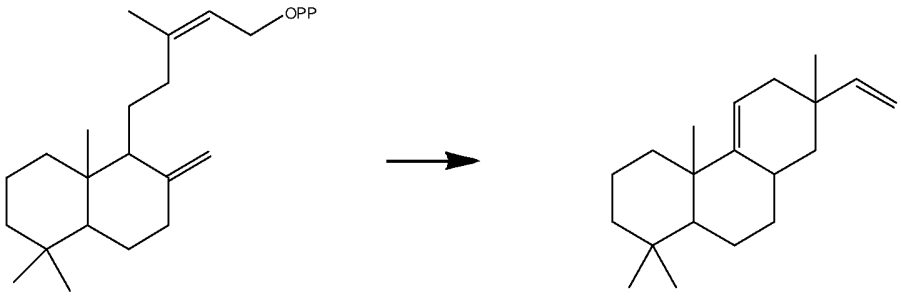
wherein OPP is diphosphate. During reaction LVI the produced diterpene will in general maintain the stereochemistry around the decalin core found in the starting diterpene pyrophosphate intermediate.

The MvTPS5 like diTPS may in particular be an enzyme capable of catalysing the reaction LIII:



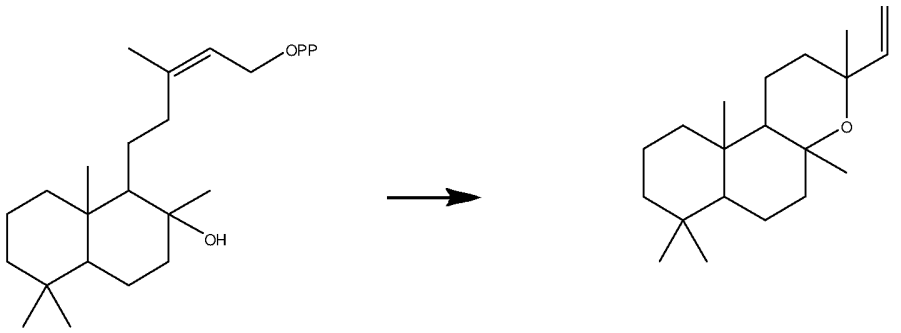
wherein OPP is diphosphate. During reaction LIII the produced diterpene will in general maintain the stereochemistry around the decalin core found in the starting diterpene pyrophosphate intermediate.

The MvTPS5 like diTPS may in particular be an enzyme capable of catalysing the reaction LIV:



wherein OPP is diphosphate. During reaction LIV the produced diterpene will in general maintain the stereochemistry around the decalin core found in the starting diterpene pyrophosphate intermediate.

The MvTPS5 like diTPS may also be an enzyme catalysing the reaction X:



wherein OPP indicates diphosphate. During reaction X the produced diterpene will in general maintain the stereochemistry around the decalin core found in the starting diterpene pyrophosphate intermediate.

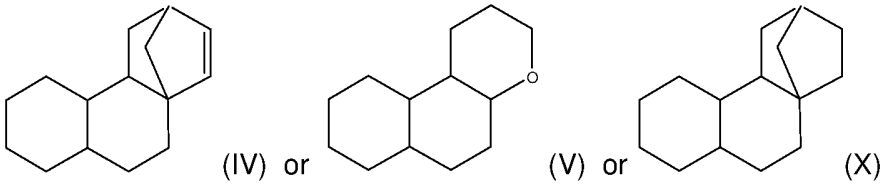
In one embodiment the MvTPS5 like diTPS may be a diterpene synthase from *Marrubium vulgare*. In particular, the MvTPS5 like diTPS may be a TPS5 from *Marrubium vulgare*. TPS5 from *Marrubium vulgare* may also be referred to as MvTPS5. In particular, said MvTPS5 like diTPS may be a polypeptide of SEQ ID NO:46 or a functional homologue thereof sharing at least 70%, such as at least 80%, for example at least 75%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 91%, such as at least 92%, such as at least 93%, such as at least 94%, such as at least 95%, such as at least 96%, such as at least 97%, such as at least 98%, such as at least 99%, such as 100% sequence identity therewith.

The sequence identity is preferably calculated as described herein below in the section "Sequence identity". A functional homologue of MvTPS5 is a polypeptide, which is also capable of catalysing at least one of reactions LIII, LIV, LV or LVI described above.

TwTPS2

The invention involves use of a diTPS of class I. In one embodiment said diTPS of class I may be a TwTPS2 like diTPS.

In particular, said diTPS of class I may be a TwTPS2 like diTPS in embodiments of the invention, wherein the diterpene to be produced contains a tricyclic ring structure. For example said diTPS of class I may be a TwTPS2 like diTPS in embodiments of the invention, wherein the diterpene to be produced contains a core of any of the formulas IV, V or X:

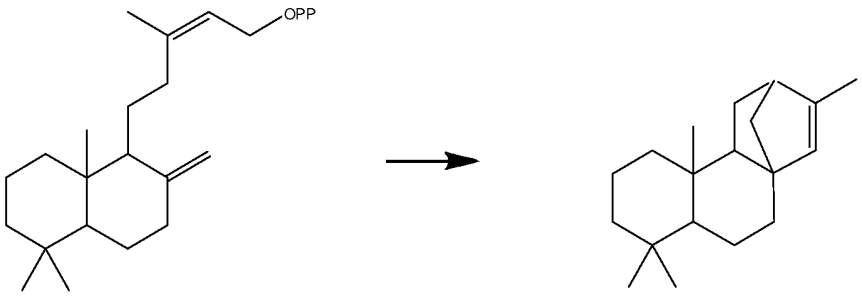


Dependent on the structure of the diterpene pyrophosphate intermediate then the diterpene containing a core of formula IV and V, may have different stereochemistry. In general the stereochemistry of the decalin core present in the diterpene pyrophosphate intermediate is maintained after the reaction catalysed by the TwTPS2 like diTPS.

The TwTPS2 like diTPS may be any enzyme capable of catalysing the reaction XXVI:

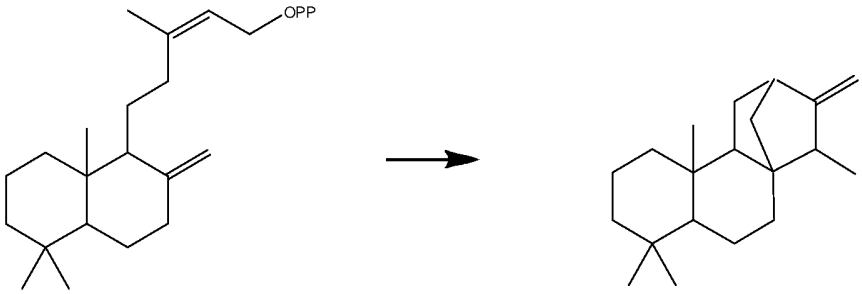
Diterpene pyrophosphate intermediate containing a decalin core structure ⟶ Diterpene containing a core structure of formula IV or formula V or formula X

The TwTPS2 like diTPS may be any enzyme capable of catalysing conversion of a diterpene pyrophosphate intermediate to a diterpene containing a core of either formula IV or V. The TwTPS2 like diTPS may in particular be an enzyme capable of catalysing the reaction XIX:



wherein OPP is diphosphate. During reaction XIX the produced diterpene will in general maintain the stereochemistry around the decalin core found in the starting diterpene pyrophosphate intermediate.

The TwTPS2 like diTPS may in particular be an enzyme capable of catalysing the reaction XXVII:



wherein OPP is diphosphate. During reaction XIX the produced diterpene will in general maintain the stereochemistry around the decalin core found in the starting diterpene pyrophosphate intermediate.

The TwTPS2 like diTPS may in particular be an enzyme capable of catalysing the reaction XX:

wherein OPP indicated diphosphate. During reaction XX the produced diterpene will in general maintain the stereochemistry around the decalin core found in the starting diterpene pyrophosphate intermediate.

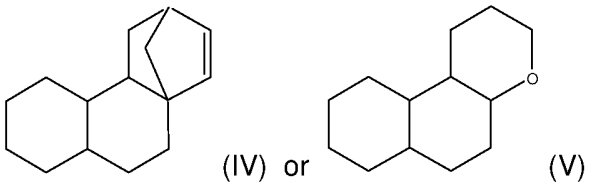
In one embodiment the TwTPS2 like diTPS may be a diterpene synthase from *Tripterygium wilfordii*. In particular, the TwTPS2 like diTPS may be a TPS2 from *Tripterygium wilfordii*. TPS2 from *Tripterygium wilfordii* may also be referred to as TwTPS2. In particular, said TwTPS2 like diTPS may be a polypeptide of SEQ ID NO:14 or a functional homologue thereof sharing at least 70%, such as at least 80%, for example at least 75%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 91%, such as at least 92%, such as at least 93%, such as at least 94%, such as at least 95%, such as at least 96%, such as at least 97%, such as at least 98%, such as at least 99%, such as 100% sequence identity therewith. The heterologous nucleic acid encoding said TwTPS2 like diTPS may for example be any nucleic acid encoding any of the TwTPS2 like diTPS described herein. Thus, the heterologous nucleic acid may for example be any nucleic acid encoding the polypeptide of SEQ ID NO:14 or any of functional homologue thereof described in the previous paragraph.

The sequence identity is preferably calculated as described herein below in the section “Sequence identity”. A functional homologue of TwTPS2 is a polypeptide, which is also capable of catalysing at least one of reactions, XIX, XX, XXVI or XXVII described above.

EpTPS1

The invention involves use of a diTPS of class I. In one embodiment said diTPS of class I may be an EpTPS1 like diTPS.

In particular, said diTPS of class I may be an EpTPS1 like diTPS in embodiments of the invention, wherein the diterpene to be produced contains a tricyclic ring structure. For example said diTPS of class I may be an EpTPS1 like diTPS in embodiments of the invention, wherein the diterpene to be produced contains a core of any of the formulas IV or V:



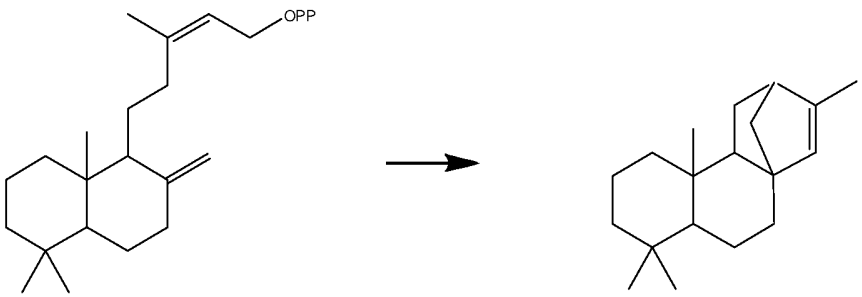
(IV) or (V)

Dependent on the structure of the diterpene pyrophosphate intermediate then the diterpene containing a core of formula IV and V, may have different stereochemistry. In general the stereochemistry of the decalin core present in the diterpene pyrophosphate intermediate is maintained after the reaction catalysed by the EpTPS1 like diTPS.

The EpTPS1 like diTPS may be any enzyme capable of catalysing the reaction XVIII:

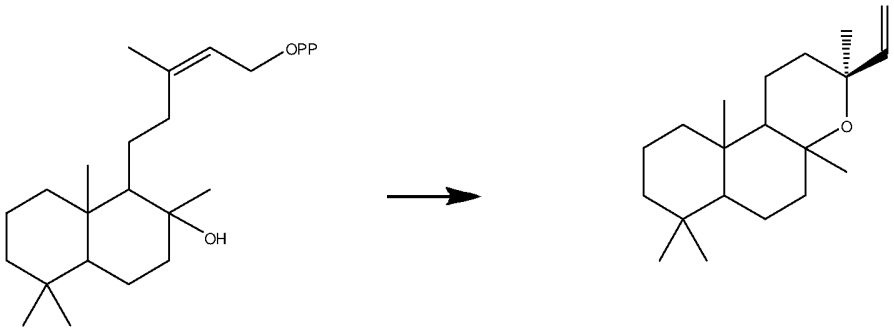
Diterpene pyrophosphate intermediate containing a decalin core structure ⟶
Diterpene containing a core structure of formula IV or formula V

The EpTPS1 like diTPS may be any enzyme capable of catalysing conversion of a diterpene pyrophosphate intermediate to a diterpene containing a core of either formula IV or V. The EpTPS1 like diTPS may in particular be an enzyme capable of catalysing the reaction XIX:



wherein OPP is diphosphate. During reaction XIX the produced diterpene will in general maintain the stereochemistry around the decalin core found in the starting diterpene pyrophosphate intermediate.

The EpTPS1 like diTPS may in particular be an enzyme capable of catalysing the reaction XX:



wherein OPP indicated diphosphate. During reaction XX the produced diterpene will in general maintain the stereochemistry around the decalin core found in the starting diterpene pyrophosphate intermediate.

In one embodiment the EpTPS1 like diTPS may be a diterpene synthase from *Euphobia peplus*. In particular, the EpTPS1 like diTPS may be a TPS1 from *Euphobia peplus*. TPS1 from *Euphobia peplus* may also be referred to as EpTPS1. In particular, said EpTPS1 like diTPS may be a polypeptide of SEQ ID NO:15 or a functional homologue thereof sharing at least 70%, such as at least 80%, for example at least 75%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 91%, such as at least 92%, such as at least 93%, such as at least 94%, such as at least 95%, such as at least 96%, such as at least 97%, such as at least 98%, such as at least 99%, such as 100% sequence identity therewith.

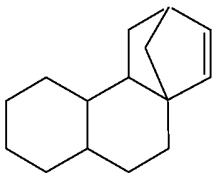
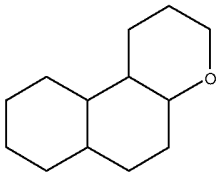
The heterologous nucleic acid encoding said EpTPS1 like diTPS may for example be any nucleic acid encoding any of the EpTPS1 like diTPS described herein. Thus, the heterologous nucleic acid may for example be any nucleic acid encoding the polypeptide of SEQ ID NO:15 or any of functional homologue thereof described in the previous paragraph.

The sequence identity is preferably calculated as described herein below in the section "Sequence identity". A functional homologue of EpTPS1 is a polypeptide, which is also capable of catalysing at least one of reactions XVIII, XIX or XX described above.

CfTPS14

The invention involves use of a diTPS of class I. In one embodiment said diTPS of class I may be an CfTPS14 like diTPS.

In particular, said diTPS of class I may be an CfTPS14 like diTPS in embodiments of the invention, wherein the diterpene to be produced contains a tricyclic ring structure. For example said diTPS of class I may be an CfTPS14 like diTPS in embodiments of the invention, wherein the diterpene to be produced contains a core of any of the formulas IV or V:

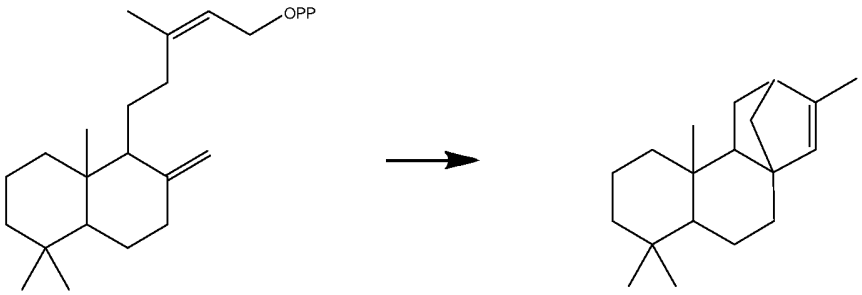

(IV) or

(V)

Dependent on the structure of the diterpene pyrophosphate intermediate then the diterpene containing a core of formula IV and V, may have different stereochemistry. In general the stereochemistry of the decalin core present in the diterpene pyrophosphate intermediate is maintained after the reaction catalysed by the CfTPS14 like diTPS.

The CfTPS14 like diTPS may be any enzyme capable of catalysing the reaction XVIII:

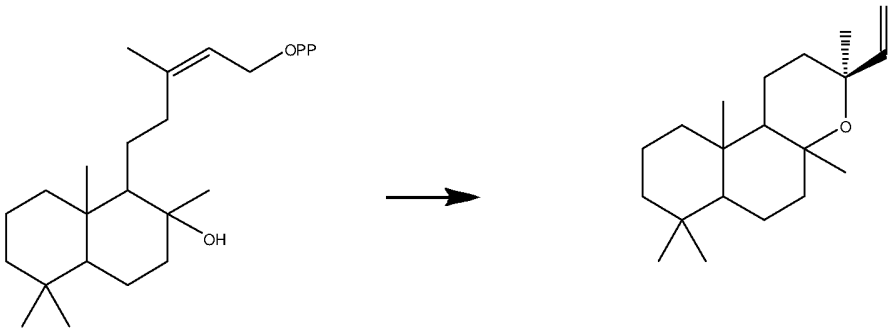
Diterpene pyrophosphate intermediate containing a decalin core structure ⟶ Diterpene containing a core structure of formula IV or formula V

The CfTPS14 like diTPS may be any enzyme capable of catalysing conversion of a diterpene pyrophosphate intermediate to a diterpene containing a core of either formula IV or V. The CfTPS14 like diTPS may in particular be an enzyme capable of catalysing the reaction XIX:



wherein OPP is diphosphate. During reaction XIX the produced diterpene will in general maintain the stereochemistry around the decalin core found in the starting diterpene pyrophosphate intermediate.

The CfTPS14 like diTPS may in particular be an enzyme capable of catalysing the reaction XX:



wherein OPP indicated diphosphate. During reaction XX the produced diterpene will in general maintain the stereochemistry around the decalin core found in the starting diterpene pyrophosphate intermediate.

In one embodiment the CfTPS14 like diTPS may be a diterpene synthase from *Coleus forskohlii*. In particular, the CfTPS14 like diTPS may be a TPS14 from *Coleus forskohlii*. TPS14 from *Coleus forskohlii* may also be referred to as CfTPS14. In particular, said CfTPS14 like diTPS may be a polypeptide of SEQ ID NO:16 or a functional homologue thereof sharing at least 70%, such as at least 80%, for example at least 75%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 91%, such as at least 92%, such as at least 93%, such as at least 94%, such as at least 95%, such as at least 96%, such as at least 97%, such as at least 98%, such as at least 99%, such as 100% sequence identity therewith.

The heterologous nucleic acid encoding said CfTPS14 like diTPS may for example be any nucleic acid encoding any of the CfTPS14 like diTPS described herein. Thus, the heterologous nucleic acid may for example be any nucleic acid encoding the polypeptide of SEQ ID NO:16 or any of functional homologue thereof described in the previous paragraph.

The sequence identity is preferably calculated as described herein below in the section "Sequence identity". A functional homologue of CfTPS14 is a polypeptide, which is also capable of catalysing at least one of reactions XVIII, XIX or XX described above.

CYP

The methods of the invention use one of more CYPs. The invention also features host organisms comprising one or more heterologous nucleic acid encoding a CYP.

The term "Cytochrome P450" as herein refers to an enzyme of the Cytochrome P450 family. Cytochrome P450s may also be referred to as "CYPs" and one cytochrome P450 may be referred to as a CYP. CYPs according to the present invention are enzymes capable of catalyzing oxidation reactions using NAD(P)H as electron donor.

CYPs may for example be enzymes capable of catalyzing one or more of the following reactions:

Hydroxylation;
Epoxidation;
Dealkylation;
Isomerization;
Dimerization;
Dehydration;
Carbon-carbon cleavage;
Decarboxylations;
Nitrogen and sulfur oxidation
Dehalogenations
Deaminations
Alkyl- and aryl-shifts, ,
Heteroatom oxidation
Heteroatom release

Ring coupling formation and contraction; and
Dehalogenation.

Preferred CYPs according to the present invention are hemoproteins capable of catalyzing oxidation reactions that utilize NADPH and/or NADH to reductively cleave atmospheric dioxygen to produce a functionalized organic substrate and a molecule of water.

Thus, a preferred CYP according to the invention is an enzyme capable of catalyzing one or more of the following reactions:

Hydroxylation
Oxidation leading to a formation of a carbonyl grops
Oxidation leading to formation of a carboxyl group.

As stated herein above the methods of the invention may involve use of one or more CYPs. However, at least one CYP is preferably and enzyme capable of catalyzing the reaction XXX:

Diterpene ⟶ hydroxyl-diterpene,

wherein the diterpene for example may be any of the diterpenes described in the section “Diterpenes” herein below. In particular, at least one CYP is an enzyme capable of catalyzing hydroxylation of a diterpene containing a core structure of one of following formulas I, II, III, IV, V, VI, IX, X, XXII, XXIII or XXIV, for example any one of the diterpenes described in Table I.

Thus in embodiments of the invention, wherein the diterpenoid is a hydroxyl-diterpene, then the CYP preferably is an enzyme capable of catalysing reaction XXX.

The CYP may also be an enzyme capable of catalysing hydroxylation of a hydroxyl-diterpene to form dihydroxy-diterpene and/or diterpene ketone and/or diterpene aldehyde. Thus, the CYP may be an enzyme capable of catalysing:

Reaction XXXI: Hydroxyl-diterpene ⟶ dihydroxy-diterpene
Reaction XXXII: Hydroxyl-diterpene ⟶ diterpene ketone

Reaction XXXIII: Hydroxyl-diterpene ⟶ diterpene aldehyde

Said hydroxyl-diterpene may in particular be a hydroxyl-diterpene containing a core structure of one of following formulas I, II, III, IV, V, VI, IX, X, XXII, XXIII or XXIV, for example it may be hydroxyl-diterpene, wherein the diterpene is any one of the diterpenes described in Table I. Preferably, the hydroxyl groups of said dihydroxy-diterpene are attached to different carbon atoms of the diterpene.

In reaction XXXII, then preferably, the ketone group of the diterpene ketone is positioned on the same carbon atom as the hydroxyl group of the hydroxyl-diterpene. Furthermore, it is preferred that said carbon atom is a carbon atom attached to exactly two hydrogen atoms in the precursor diterpene.

In reaction XXXIII, then preferably, the aldehyde group of the diterpene aldehyde is positioned on the same carbon atom as the hydroxyl group of the hydroxyl-diterpene. Furthermore, it is preferred that said carbon atom, is a carbon atom attached to exactly 3 hydrogen atoms in the precursor diterpene.

In embodiments of the invention wherein the diterpenoid is dihydroxy-diterpene, then the methods preferably involve use of a CYP, which is capable of catalysing reaction XXX and a CYP, which is capable of catalysing reaction XXXI. Said CYP may be one CYP capable of catalysing both reaction or it may be two different CYPs.

In embodiments of the invention wherein the diterpenoid is diterpene ketone, then the methods preferably involve use of one or more CYP(s), which can catalyse reactions XXX and XXXII.

In embodiments of the invention wherein the diterpenoid is diterpene aldehyde, then the methods involve use of one or more CYP(s), which can catalyse reactions XXX and XXXIII.

The CYP may also be an enzyme capable of catalysing hydroxylation of a dihydroxy-diterpene to form trihydroxy-diterpene and/or diterpene carboxylic acid. Thus, the CYP may be an enzyme capable of catalysing:

Reaction XXXIV: Dihydroxy-diterpene ⟶ trihydroxy-diterpene
Reaction XXXVa: Dihydroxy-diterpene ⟶ Hydroxy-diterpene ketone
Reaction XXXVb: Dihydroxy-diterpene ⟶ Hydroxy-diterpene carboxylic acid

Said dihydroxy-diterpene may in particular be a dihydroxy-diterpene containing a core structure of one of following formulas I, II, III, IV, V, VI, IX, X, XXII, XXIII or XXIV, for example it may be dihydroxy-diterpene, wherein the diterpene is any one of the diterpenes described in Table I. Preferably, the hydroxyl groups of said dihydroxy-diterpene are attached to different carbon atoms of the diterpene. Similarly, it is preferred that the hydroxyl groups of the trihydroxy-diterpenene are attached to different carbon atoms.

In reaction XXXVa, then preferably, the ketone group of the hydroxy-diterpene ketone is positioned on the same carbon atom as one of the hydroxyl group of the dihydroxy-diterpene. Furthermore, it is preferred that said carbon atom, is a carbon atom attached to exactly 2 hydrogen atoms in the precursor diterpene.

In reaction XXXVb, then preferably, the carboylix acid group of the hydroxy-diterpene carboxylic acid is positioned on the same carbon atom as one of the hydroxyl group of the dihydroxy-diterpene. Furthermore, it is preferred that said carbon atom, is a carbon atom attached to exactly 3 hydrogen atoms in the precursor diterpene.

In embodiments of the invention, wherein the diterpenoid is trihydroxy-diterpene, then the methods may involve use of a CYP capable of catalysing reaction XXX and of a CYP capable of catalysing reaction XXXI and of a CYP capable of catalysing reaction XXXIV. This may be different CYPs, or one CYP can be capable of catalysing two or all three of said reactions.

In embodiments of the invention wherein the diterpenoid is a hydroxyl-diterpene carboxylic acid, then the methods may involve use of a CYP capable of catalysing reaction XXX and of a CYP capable of catalysing reaction XXXI and of a CYP capable of catalysing reaction XXXV. This may be different CYPs, or one CYP can be capable of catalysing two or all three of said reactions.

The CYP may also be an enzyme capable of catalysing hydroxylation of diterpene-ketone to form hydroxyl-diterpene ketone and/or diterpene carboxylic acid. Thus, the CYP may be an enzyme capable of catalysing:

Reaction XXXVI: diterpene ketone ⟶ hydroxyl-diterpene ketone
Reaction XXXVII: diterpene ketone ⟶ diterpene carboxylic acid

Said diterpene ketone may in particular be a diterpene substituted with a ketone group, wherein the diterpene contains a core structure of one of following formulas I, II, III, IV, V, VI, IX, X, XXII, XXIII or XXIV, for example it may be any one of the diterpenes described in Table I. It is preferred that the ketone group and the hydroxyl group of the hydroxyl-diterpene ketone are attached to different carbon atoms.

In reaction XXXVII, then preferably, the ketone group of the diterpene ketone is positioned on the same carbon atom as the carboxylic acid of the diterpene carboxylic acid. Furthermore, it is preferred that said carbon atom, is a carbon atom attached to exactly 3 hydrogen atoms in the precursor diterpene.

In embodiments of the invention wherein the diterpenoid is hydroxy-diterpene ketone, then the methods may involve use of a CYP capable of catalysing reaction XXX and of a CYP capable of catalysing reaction XXXII and of a CYP capable of catalysing reaction XXXVI. This may be different CYPs, or one CYP can be capable of catalysing two or all three of said reactions.

In embodiments of the invention wherein the diterpenoid is diterpene carboxylic acid, then the methods may involve use of a CYP capable of catalysing reaction XXX and of a CYP capable of catalysing reaction XXXII and of a CYP capable of catalysing reaction XXXVII. This may be different CYPs, or one CYP can be capable of catalysing two or all three of said reactions.

The CYP may also be an enzyme capable of catalysing hydroxylation of diterpene-carboxylic acid to form hydroxyl-diterpene carboxylic acid. Thus, the CYP may be an enzyme capable of catalysing:

Reaction XXXVIII: diterpene carboxylic acid ⟶ hydroxyl-diterpene carboxylic acid

Said diterpene carboxylic acid may in particular be a diterpene substituted with a carboxyl group, wherein the diterpene contains a core structure of one of following formulas I, II, III, IV, V, VI, IX, X, XXII, XXIII or XXIV, for example it may be any one of the diterpenes described in Table I.The carboxylic acid group will generally be attached to a different carbon atom as the hydroxyl group.

In embodiments of the invention wherein the diterpenoid is hydroxyl-diterpene carboxylic acid, then the methods may involve use of a CYP capable of catalysing reaction XXX and of a CYP capable of catalysing reaction XXXI and of a CYP capable of catalysing reaction XXXV and of a CYP capable of catalysing reaction XXXVIII. This may be different CYPs, or one CYP can be capable of catalysing two, three or even all of said reactions.

In embodiments of the invention wherein the diterpenoid is hydroxyl-diterpene carboxylic acid, then the methods may involve use of a CYP capable of catalysing reaction XXX and of a CYP capable of catalysing reaction XXXII and of a CYP capable of catalysing reaction XXXVII and of a CYP capable of catalysing reaction XXXVIII. This may be different CYPs, or one CYP can be capable of catalysing two, three or even all of said reactions.

In addition to above mentioned reactions the methods may also employ use of one or more CYPs capable of catalysing one or more of the following reactions:

Trihydroxy-diterpene ⟶ tetrahydroxy-diterpene
Trihydroxy-diterpene ⟶ dihydroxy-diterpene ketone
Trihydroxy-diterpene ⟶ dihydroxy-diterpene carboxylic acid
hydroxyl-diterpene ketone ⟶ dihydroxy –diterpene ketone
hydroxyl-diterpene ketone ⟶ diterpene-di-ketone
hydroxyl-diterpene carboxylic acid ⟶ dihydroxy-diterpene carboxylic acid
hydroxyl-diterpene carboxylic acid ⟶ diterpene-ketone-carboxylic acid
tetrahydroxy-diterpene ⟶ pentahydroxy-diterpene
tetrahydroxy-diterpene ⟶ trihydroxy-diterpene ketone
tetrahydroxy-diterpene ⟶ tri-hydroxy-diterpene carboxylic acid
dihydroxy-diterpene ketone ⟶ trihydroxy-diterpene ketone

dihydroxy-diterpene ketone ⟶ hydroxy-diterpene di-ketone
dihydroxy-diterpene ketone ⟶ hydroxyl-diterpene ketone carboxylic acid

CYPs are encoded by gene superfamily, which is divided into families sharing at least 40% sequence identity. The families are divided into subfamilies sharing at least 55% sequence identity. The CYP families is has a number, which generally is written after "CYP". Thus, by way of example CYPs of family 74 are named CYP74. The subfamilies are indicated by a capital letter after the family number. Thus by way of example a CYP of family 74 and subfamily A is named CYP74A. Additional description of CYPs, the structural characteristics and the nomenclature thereof may for example be found in Schuler et al., Annu Rev. Plant Biol., 2003, 54:629-67 and in Podust et al., 2012, Nat. Prod. Rep., 29:1251-1266. Thus, the CYP to be used with the present invention may be a CYP as described in any of these references.

In one embodiment of the invention the CYP is a polypeptide sharing at least some sequence similarity to at least one of SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22 or SEQ ID NO:23. In particular, it is preferred that the CYP shares at least 30%, preferably at least 40%, more preferably at least 45% sequence identity with at least one of SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, or SEQ ID NO:23.

Furthermore, it is preferred that the CYP contains the following motif of five amino acids:

A/G-G-X-X-T/S,

wherein X may be any amino acid, such as any naturally occurring amino acids. In particular, one of the X may be an amino acid with a charged side chain, and in particular an acidic side chain, such as glutamic acid (E).

A/G indicates that the amino acid may be A or G. Similarly, T/S indicates that the amino acid may be T or S.

In one embodiment of the invention said motif of five amino acids is:
A/G-X-E-T/S
wherein X may be any amino acid, such as any naturally occurring amino acids.

Said motif of five amino acids may for example be

A/G-G-H/S/G-E-T/S.

H/S/G indicates that the amino acid may be H or S or G.

In particular, it is preferred that the CYP contains said motif in a position corresponding to of position aa 290-294 of PsCYP720B4 SEQ ID NO:23. A position corresponding to position aa 290-294 of PsCYP720B4 SEQ ID NO:23 is identified by aligning the sequence of a CYP of interest to SEQ ID NO:23 and optionally to additional sequences of CYPs as e.g. shown in figure 6, and identifying the amino acids of said CYP aligned with aa 290-294of PsCYP720B4 SEQ ID NO:23. This motif is shown as "1" in figure 6.

Furthermore, it is preferred that the CYP contains the following motif 4 amino acids:

E-X-X-R,

wherein X may be any amino acid, such as any naturally occurring amino acids. In particular, X may be an amino acid with an uncharged side chain, such as an hydrophobic side chain.

In one embodiment of the invention said motif of four amino acids is:
E-X-L-R
wherein X may be any amino acid, such as any naturally occurring amino acids.

Said motif of five amino acids may for example be
E-T/V/M/A-L-R.

T/V/M/A indicates that the amino acid may be T or V or M or A.

In particular, it is preferred that the CYP contains said motif in a position corresponding to of position aa 350-353 of PsCYP720B4 SEQ ID NO:23. A position corresponding to position aa 350-353 of PsCYP720B4 SEQ ID NO:23 is identified by aligning the sequence of a CYP of interest to SEQ ID NO:23 and optionally to additional sequences of CYPs as e.g. shown in figure 6, and identifying the amino acids of said CYP aligned with aa 350-353 of PsCYP720B4 SEQ ID NO:23. This motif is shown as "2" in figure 6.

Furthermore, it is preferred that the CYP contains the following motif following motif of 10 amino acids:

F-X-X-G-X-X-X-C-X-G,

wherein X may be any amino acid, such as any naturally occurring amino acids.

In one embodiment of the invention said motif of 10 amino acids is:

F-G-G/A-G-A/R-R-X-C-P-G,

wherein X may be any amino acid, such as any naturally occurring amino acids.

G/A indicates that the amino acid may be G or A and A/R indicates that the amino acid may be A or R.

In particular, it is preferred that the CYP contains said motif in a position corresponding to of position aa 422-431 of PsCYP720B4 SEQ ID NO:23. A position corresponding to position aa 422-431of PsCYP720B4 SEQ ID NO:23 is identified by aligning the sequence of a CYP of interest to SEQ ID NO:23 and optionally to additional sequences of CYPs as e.g. shown in figure 6, and identifying the amino acids of said CYP aligned with aa 422-431of PsCYP720B4 SEQ ID NO:23. This motif is shown as “3” in figure 6.

Furthermore, it is preferred that the CYP contains the following motif following motif of 3 amino acids:

P-F-G

In particular, it is preferred that the CYP contains said motif in a position corresponding to of position aa 421-423 of PsCYP720B4 SEQ ID NO:23. A position corresponding to position aa 421-423of PsCYP720B4 SEQ ID NO:23 is identified by aligning the sequence of a CYP of interest to SEQ ID NO:23 and optionally to additional sequences of CYPs as e.g. shown in figure 6, and identifying the amino acids of said CYP aligned with aa 421-423of PsCYP720B4 SEQ ID NO:23. This motif may be overlapping with the motif of 10 amino acids described above. The motif is shown as “4” in figure 6.
In addition to the motifs described above, the CYP may contain one or more of the following motifs:

SRS1: P/V-X-X-X-K-$(X)_{7-9}$-F

In particular, it is preferred that the CYP contains SRS1 in a position corresponding to of position aa 100-114 of PsCYP720B4 SEQ ID NO:23. The motif is shown as "SRS1" in figure 6.

SRS5: N/P-F/P-G/L-P-X-V/L-X-X-R

In particular, it is preferred that the CYP contains SRS5 in a position corresponding to of position aa 356-360 of PsCYP720B4 SEQ ID NO:23. The motif is shown as "SRS5" in figure 6.

The CYP may also contain one or more of the motifs SRS2, SRS3 and/or SRS6 shown in figure 6.

It is furthermore preferred that in addition to sharing above mentioned sequence identity and/or containing one or more said motifs (preferably containing all of said motifs), then as many as possible of the amino acids shown in a black box in figure 6 are retained. Thus, when aligned to the sequence of of PsCYP720B4 (SEQ ID NO:23), then preferably the CYP also contains at least 80%, more preferably at least 90%, for example at least 95%, such as all of the amino acids shown in a black box in figure 6. It is also preferred that in addition to sharing above mentioned sequence identity and/or containing one or more said motifs, then as many as possible of the amino acids shown in a grey box in figure 6 are retained. Thus, when aligned to the sequence of of PsCYP720B4 (SEQ ID NO:23), then preferably the CYP also contains at least 80%, more preferably at least 90%, for example at least 95%, such as at least 98% of the amino acids shown in a grey box in figure 6.

In one embodiment of the invention the CYP may be one of the following CYPs:

i. A CYP of family 76, such as any of the CYPs described herein below in the section "CYP76"
ii. A CYP of family 71, such as any of the CYPs described herein below in the section "CYP71"

iii. A CYP of family 720, such as any of the CYPs described herein below in the section "CYP720"

CYP76

In one embodiment of the invention at least one CYP may be a CYP of family 76. A CYP of family 76 may preferably be capable of catalysing reaction XXX outlined above. In particular a CYP of family 76 may be an enzyme capable of catalysing one or more of the following reactions (preferably said CYP of family 76 is capable of catalysing all of the following reactions):

i. Reaction XXXIX: hydroxylation of dehydroabietadiene to form hydroxyl dehydroabientadiene
ii. Reaction XL: Rehydroxylation of miltiradiene to form hydroxyl mitradiene
iii. Reaction XLI: hydroxylation of manool to form hydroxyl manool

The CYP of family 76 may also be caoable of catalysing one or more of the following reactions:
Reaction XLV: manool -> keto-hydroxy manool
Reaction XLVI: manool -> dihydroxy manool
Reaction XLVII: manool -> carboboxy manool
Reaction LI: hydroxylation of hydroxy manool to form dihydroxymanool

In one embodiment the CYP of family 76 may be a CYP from *Coleus forskohlii*. In particular, said CYP of family 76 may be a polypeptide sharing at least 40%, such as at least 45%, for example at least 60%, such as at least 55% sequence identity with SEQ ID NO:20 and/or SEQ ID NO:21.

In particular the CYP of family 76 may be a CYP of subfamily AH. Thus said CYP may be a CYP76AH. Said CYP of family 76 and subfamily AH may be a polypeptide sharing at least 55%, such as at least 60%, for example at least 65%, such as at least 70% sequence identity with SEQ ID NO:20 and/or SEQ ID NO:21.

In a preferred embodiment the CYP of family 76 is CYP76AH8 of SEQ ID NO:20, CYP76AH11 of SEQ ID NO:21 or a functional homologue thereof sharing at least 70%, such as at least 80%, for example at least 75%, such as at least 80%, such as at least

85%, such as at least 90%, such as at least 91%, such as at least 92%, such as at least 93%, such as at least 94%, such as at least 95%, such as at least 96%, such as at least 97%, such as at least 98%, such as at least 99%, such as 100% sequence identity therewith.

In another embodiment the CYP of family 76 is CYP76AH15 of SEQ ID NO:40, CYP76AH17 of SEQ ID NO:41 or a functional homologue thereof sharing at least 70%, such as at least 80%, for example at least 75%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 91%, such as at least 92%, such as at least 93%, such as at least 94%, such as at least 95%, such as at least 96%, such as at least 97%, such as at least 98%, such as at least 99%, such as 100% sequence identity therewith. CYP76AH15 of SEQ ID NO:40 and CYP76AH17 of SEQ ID NO:41 have an activity similar to CYP76AH8 of SEQ ID NO:20, and thus CYP76AH15 of SEQ ID NO:40 or CYP76AH17 of SEQ ID NO:41 or functional homologues thereof may be employed in place of CYP76AH8 of SEQ ID NO:20 or functional homologues thereof.

The heterologous nucleic acid encoding said CYP of family 76 may for example be any nucleic acid encoding any of the CYPs of family 76 described herein. Thus, the heterologous nucleic acid may for example be any nucleic acid encoding the polypeptide of SEQ ID NO:20, or the polypeptide of SEQ ID NO:21 or any of the functional homologue of the aforementioned described in the previous paragraph. For example, said heterologous nucleic acid may comprise or consist of SEQ ID NO:35. Said heterologous nucleic acid may also comprise or consist of SEQ ID NO:36. This may in particular be the case in embodiments of the invention wherein the host organism is yeast, e.g. *S. cerevisiae*.

In another embodiment the CYP of family 76 is CYP76AH9 of SEQ ID NO:24 or a functional homologue thereof sharing at least 70%, such as at least 80%, for example at least 75%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 91%, such as at least 92%, such as at least 93%, such as at least 94%such as at least 95%, such as at least 96%, such as at least 97%, such as at least 98%, such as at least 99%, such as 100% sequence identity therewith.

The heterologous nucleic acid may for example be any nucleic acid encoding the polypeptide of SEQ ID NO:24 or any of functional homologue thereof described in the previous paragraph.

In another embodiment the CYP of family 76 is CYP76AH1, such as CYP76AH1 of SEQ ID NO:42 or a functional homologue thereof sharing at least 70%, such as at least 80%, for example at least 75%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 91%, such as at least 92%, such as at least 93%, such as at least 94%, such as at least 95%, such as at least 96%, such as at least 97%, such as at least 98%, such as at least 99%, such as 100% sequence identity therewith.

The heterologous nucleic acid may for example be any nucleic acid encoding the polypeptide of SEQ ID NO:42 or any of functional homologue thereof described in the previous paragraph. For example the heterologous nucleic acid may be the nucleic acid of sequence with the Genbank accession number KP337687.1.

In another embodiment the CYP of family 76 is CYP76AH4, such as CYP76AH4 of SEQ ID NO:43 or a functional homologue thereof sharing at least 70%, such as at least 80%, for example at least 75%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 91%, such as at least 92%, such as at least 93%, such as at least 94%, such as at least 95%, such as at least 96%, such as at least 97%, such as at least 98%, such as at least 99%, such as 100% sequence identity therewith. Said functional homologue may for example be the protein with Genbank accession number AJQ30187.1 or the protein with Genbank accession number AJQ30188.1.

The heterologous nucleic acid may for example be any nucleic acid encoding the polypeptide of SEQ ID NO:43 or any of functional homologue thereof described in the previous paragraph. For example the heterologous nucleic acid may be the nucleic acid of sequence with the Genbank accession number KP091843.1 of the Genbank accession number KP091844.1.

The sequence identity is preferably calculated as described herein below in the section “Sequence identity”. A functional homologue of CYP76AH8 or CYP76AH11 or CYP76AH9 is a polypeptide, which is also capable of catalysing at least one of reactions XXXIX, XL, XLI, XLV, XLVI, XLVII, LI described above.

It is furthermore preferred that in addition to sharing above mentioned sequence identity, then as many as possible of the amino acids shown in a black box in figure 6 are retained. Thus, when aligned to the sequence of CYP76AH8 (SEQ ID NO:20) and/or CYP76AH11 (SEQ ID NO:21), then preferably the CYP of family 76 also contains at least 80%, more preferably at least 90%, for example at least 95%, such as all of the amino acids shown in a black box in figure 6. Furthermore, when aligned to the sequence of CYP76AH8 (SEQ ID NO:20) and/or CYP76AH11 (SEQ ID NO:21), then preferably the CYP of family 76 also contains at least 80%, more preferably at least 90%, for example at least 95%, such as at least 98% of the amino acids shown in a grey box in figure 6.

CYP71

In one embodiment of the invention at least one CYP may be a CYP of family 71. A CYP of family 71 may preferably be capable of catalysing reaction XXX outlined above. In particular a CYP of family 71 may be an enzyme capable of catalysing one or more of the following reactions (preferably said CYP of family 71 is capable of catalysing all of the following reactions):

i. Reaction XXXIX: Oxidation of dehydroabietadiene and/or mitradiene
ii. Reaction XLII: hydroxylation of syn-manool to form hydroxy syn-manool
Reaction LI: hydroxylation of hydroxy manool to form dihydroxymanool
iii.

In one embodiment the CYP of family 71 may be a CYP from *Coleus forskohlii*. In particular, said CYP of family 71 may be a polypeptide sharing at least 40%, such as at least 45%, for example at least 60%, such as at least 55% sequence identity with SEQ ID NO:22.

In particular the CYP of family 71 may be a CYP of subfamily D. Thus said CYP may be a CYP71D. Said CYP of family 71 and subfamily D may be a polypeptide sharing at least 55%, such as at least 60%, for example at least 65%, such as at least 70% sequence identity with SEQ ID NO:22.

In a preferred embodiment the CYP of family 71 is CYP71D381 of SEQ ID NO:22 or a functional homologue thereof sharing at least 70%, such as at least 80%, for example at least 75%, such as at least 80%, such as at least 85%, such as at least 90%, such

as at least 91%, such as at least 92%, such as at least 93%, such as at least 94%, such as at least 95%, such as at least 96%, such as at least 97%, such as at least 98%, such as at least 99%, such as 100% sequence identity therewith.

The heterologous nucleic acid encoding said CYP of family 71 may for example be any nucleic acid encoding any of the CYPs of family 71 described herein. Thus, the heterologous nucleic acid may for example be any nucleic acid encoding the polypeptide of SEQ ID NO:22 or any of functional homologue thereof described in the previous paragraph. For example, said heterologous nucleic acid may comprise or consist of SEQ ID NO:34.

The sequence identity is preferably calculated as described herein below in the section "Sequence identity". A functional homologue of CYP71D381 is a polypeptide, which is also capable of catalysing at least one of reactions XXXIX, XL, XLII or LI described above.

It is furthermore preferred that in addition to sharing above mentioned sequence identity, then as many as possible of the amino acids shown in a black box in figure 6 are retained. Thus, when aligned to the sequence of CYP71D381 (SEQ ID NO:22), then preferably the CYP of family 71 also contains at least 80%, more preferably at least 90%, for example at least 95%, such as all of the amino acids shown in a black box in figure 6. Furthermore, when aligned to the sequence of CYP71D381 (SEQ ID NO:22), then preferably the CYP of family 71 also contains at least 80%, more preferably at least 90%, for example at least 95%, such as at least 98% of the amino acids shown in a grey box in figure 6.

CYP720

In one embodiment of the invention at least one CYP may be a CYP of family 720. A CYP of family 720 may preferably be capable of catalysing reaction XXX outlined above. In particular a CYP of family 720 may be an enzyme capable of catalysing the following reaction:

i. Reaction XLI: hydroxylation of manool to form hydroxyl manool

The CYP of family 720 may also be an enzyme capable of catalysing one of both of the following reactions:

Reaction XLIII:	dehydroabietadiene -> dehydroabietic acid
Reaction XLIV:	miltradiene -> miltiradienic acid
Reaction XLVIII:	hydroxylation of syn-pimara-9,(11),15-diene OR syn-isoprimara-7,15-diene OR debydroabietadiene to form hydroxy syn-pimara-9,(11),15-diene OR hydroxy syn-isoprimara-7,15-diene OR hydroxy debydroabietadiene, respectively
Reaction XLIX:	addition of keto group to syn-pimara-9,(11),15-diene OR syn-isoprimara-7,15-diene OR debydroabietadiene to form keto syn-pimara-9,(11),15-diene OR keto syn-isoprimara-7,15-diene OR keto debydroabietadiene, respectively
Reaction L:	carboxylation of syn-pimara-9,(11),15-diene OR syn-isoprimara-7,15-diene OR debydroabietadiene to form carboxy syn-pimara-9,(11),15-diene OR carboxy syn-isoprimara-7,15-diene OR carboxy debydroabietadiene, respectively
Reaction LI:	hydroxylation of hydroxy manool to form dihydroxymanool
Reaction LII:	miltiradiene OR abietatriene -> dehydroabietic acid

In one embodiment the CYP of family 720 may be a CYP from *Picea sitchensis*. In particular, said CYP of family 720 may be a polypeptide sharing at least 40%, such as at least 45%, for example at least 60%, such as at least 55% sequence identity with SEQ ID NO:23.

In particular the CYP of family 720 may be a CYP of subfamily B. Thus said CYP may be a CYP720B. Said CYP of family 720 and subfamily B may be a polypeptide sharing at least 55%, such as at least 60%, for example at least 65%, such as at least 70% sequence identity with SEQ ID NO:23.

In a preferred embodiment the CYP of family 720 is CYP720B4 of SEQ ID NO:23 or a functional homologue thereof sharing at least 70%, such as at least 80%, for example at least 75%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 91%, such as at least 92%, such as at least 93%, such as at least 94%,

such as at least 95%, such as at least 96%, such as at least 97%, such as at least 98%, such as at least 99%, such as 100% sequence identity therewith.

The heterologous nucleic acid encoding said CYP of family 720 may for example be any nucleic acid encoding any of the CYPs of family 720 described herein. Thus, the heterologous nucleic acid may for example be any nucleic acid encoding the polypeptide of SEQ ID NO:23 or any of functional homologue thereof described in the previous paragraph. For example, said heterologous nucleic acid may comprise or consist of SEQ ID NO:39.

The sequence identity is preferably calculated as described herein below in the section "Sequence identity". A functional homologue of CYP720B4 is a polypeptide, which is also capable of catalysing one, two or all of reactions XLI, XLIII, XLIV, XLVIII, XLIX, L, LI or LII described above.

It is furthermore preferred that in addition to sharing above mentioned sequence identity, then as many as possible of the amino acids marked with a black box in figure 6 are retained. Thus, when aligned to the sequence of CYP720B4 (SEQ ID NO:23), then preferably the CYP of family 720 also contains at least 80%, more preferably at least 90%, for example at least 95%, such as all of the amino acids marked by a black box in figure 6. Furthermore, when aligned to the sequence of CYP720B4 (SEQ ID NO:23), then preferably the CYP of family 720 also contains at least 80%, more preferably at least 90%, for example at least 95%, such as at least 98% of the amino acids shown in a grey box in figure 6.

Additional recombinant modifications

The host organisms according to the present invention may also be recombinantly modified in addition to comprising the heterologous nucleic acids encoding a diTPS of class I and a diTPS of class II and one or more CYPs as described herein.

For example the host organism may be modified to increase the pool of GGPP. As described herein elsewhere, GGPP is the starting compound for production of diterpenes. Thus, if the host organism is modified to increase the pool of GGPP, then

frequently, the host organism will be capable of producing increased amounts of diterpene.

Various methods for increasing the pool of GGPP are well known in the art. These includes methods of reducing the activity of enzymes reducing the level of GGPP.

In one embodiment the pool of GGPP is increased by expression of one or more enzymes involved in synthesis of GGPP.

Thus, it may be preferred that the host organism comprises a heterologous nucleic acid encoding GGPP synthase (GGPPS). Said GGPPS may be any GGPPS, e.g. BTS1 of *S. cerevisiae*.

In particular, the GGPPS may be the GGPPS described by Zhou, Y. J., W. Gao, Q. Rong, G. Jin, H. Chu, W. Liu, W. Yang, Z. Zhu, G. Li, G. Zhu, L. Huang and Z. K. Zhao (2012). "Modular Pathway Engineering of Diterpenoid Synthases and the Mevalonic Acid Pathway for Miltiradiene Production." Journal of the American Chemical Society **134**(6): 3234-3241.

Accordingly, the host organism may express a fusion of SmCPS and SmKSL, and/or a fusion of BTS1 (GGPP synthase) and ERG20 (farnesyl diphosphate synthase) as described in Zhou et al., 2012.

The host organism may also comprise a heterologous nucleic acid encoding a GGPPS from a plant, e.g. from *Coleus forskohlii.* Thus, in one embodiment the host organism comprises:

a) a heterologous nucleic acid encoding *Coleus forskohlii* deoxyxylulose 5-phosphate synthase (CfDXS) of SEQ ID NO:25 or a functional homologue of any of the aforementioned sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith and/or

b) a heterologous nucleic acid encoding *Coleus forskohlii* geranylgeranylpyrophosphate synthase (CfGGPPs) of SEQ ID NO:26 or a functional homologue of any of the aforementioned sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as

at least 95%, such as at least 98%, such as at least 99% sequence identity therewith.

The host organism may also comprise a heterologous nucleic acid encoding *Coleus forskohlii* geranylgeranylpyrophosphate synthase (CfGGPPs) of SEQ ID NO:26 or a functional homologue of any of the aforementioned sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith. The host organism may also comprise a a heterologous nucleic acid encoding GGPP7 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith. The sequence of *GGPPS7* of *Synechococcus sp*. is available under Gene ID:86553638.

The host organism may also comprise a heterologous nucleic acid encoding a cytochrome P450 oxidoreductases (POR). PORs are known in the art. Said POR may for example be ATR1 from *A. thaliana* or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith. The sequence of ATR1 from A. thaliana is available under Gene ID:3150037. Said POR may for example be POR from *Stevia rebaudiana* or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith. The sequence of POR from S. rebaudiana is available under Gene ID: 93211213.

Said POR may also be a CPR, such as CPR from *Coleus forskohlii.* Thus, the host organism may in addition to the heterologous nucleic acids encoding diTPS and CYPs also contain a heterologous nucleic acic encoding CfCPR of SEQ ID NO:44 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith. Said heterologous nucleic acid may be any heterologous nucleic acid encoding CPR, for example the nucleic acid may comprise or consist of SEQ ID NO:45.

Sequence identity

A high level of sequence identity indicates likelihood that the first sequence is derived from the second sequence. Amino acid sequence identity requires identical amino acid sequences between two aligned sequences. Thus, a candidate sequence sharing 80% amino acid identity with a reference sequence, requires that, following alignment, 80% of the amino acids in the candidate sequence are identical to the corresponding amino acids in the reference sequence. Identity according to the present invention is determined by aid of computer analysis, such as, without limitations, the ClustalW computer alignment program (Higgins D., Thompson J., Gibson T., Thompson J.D., Higgins D.G., Gibson T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22:4673-4680), and the default parameters suggested therein. The ClustalW software is available from as a ClustalW WWW Service at the European Bioinformatics Institute http://www.ebi.ac.uk/clustalw or via the software BioEdit. Using this program with its default settings, the mature (bioactive) part of a query and a reference polypeptide are aligned. The number of fully conserved residues are counted and divided by the length of the reference polypeptide. Thus if nothing else is indicated a sequence identity is given over the entire length of the reference polypeptide.

Thus, sequence identity is calculated over the entire length of the reference polypeptide.

The ClustalW algorithm may similarly be used to align nucleotide sequences. Sequence identities may be calculated in a similar way as indicated for amino acid sequences

Heterologous nucleic acid

The term “heterologous nucleic acid” as used herein refers to a nucleic acid sequence, which has been introduced into the host organism, wherein said host does not endogenously comprise said nucleic acid. For example, said heterologous nucleic acid may be introduced into the host organism by recombinant methods. Thus, the genome of the host organism has been augmented by at least one incorporated heterologous

nucleic acid sequence. It will be appreciated that typically the genome of a recombinant host described herein is augmented through the stable introduction of one or more heterologous nucleic acids encoding one or more diTPS's and/or CYPs.

Suitable host organisms include microorganisms, plant cells, and plants, and may for example be any of the host organisms described herein below in the section "Host organism".

In general the heterologous nucleic acid encoding a polypeptide (also referred to as "coding sequence" in the following) is operably linked in sense orientation to one or more regulatory regions suitable for expressing the polypeptide. Because many microorganisms are capable of expressing multiple gene products from a polycistronic mRNA, multiple polypeptides can be expressed under the control of a single regulatory region for those microorganisms, if desired. A coding sequence and a regulatory region are considered to be operably linked when the regulatory region and coding sequence are positioned so that the regulatory region is effective for regulating transcription or translation of the sequence. Typically, the translation initiation site of the translational reading frame of the coding sequence is positioned between one and about fifty nucleotides downstream of the regulatory region for a monocistronic gene.

"Regulatory region" refers to a nucleic acid having nucleotide sequences that influence transcription or translation initiation and rate, and stability and/or mobility of a transcription or translation product. Regulatory regions include, without limitation, promoter sequences, enhancer sequences, response elements, protein recognition sites, inducible elements, protein binding sequences, 5´ and 3´ untranslated regions (UTRs), transcriptional start sites, termination sequences, polyadenylation sequences, introns, and combinations thereof. A regulatory region typically comprises at least a core (basal) promoter. A regulatory region also may include at least one control element, such as an enhancer sequence, an upstream element or an upstream activation region (UAR). A regulatory region is operably linked to a coding sequence by positioning the regulatory region and the coding sequence so that the regulatory region is effective for regulating transcription or translation of the sequence. For example, to operably link a coding sequence and a promoter sequence, the translation initiation site of the translational reading frame of the coding sequence is typically positioned between one and about fifty nucleotides downstream of the promoter. A regulatory

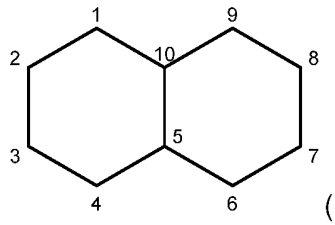
region can, however, be positioned at further distance, for example as much as about 5,000 nucleotides upstream of the translation initiation site, or about 2,000 nucleotides upstream of the transcription start site.

The choice of regulatory regions to be included depends upon several factors, including the type of host organism. It is a routine matter for one of skill in the art to modulate the expression of a coding sequence by appropriately selecting and positioning regulatory regions relative to the coding sequence. It will be understood that more than one regulatory region may be present, *e.g.*, introns, enhancers, upstream activation regions, transcription terminators, and inducible elements.

It will be appreciated that because of the degeneracy of the genetic code, a number of nucleic acids can encode a particular polypeptide; *i.e.*, for many amino acids, there is more than one nucleotide triplet that serves as the codon for the amino acid. Thus, codons in the coding sequence for a given polypeptide can be modified such that optimal expression in a particular host organisms obtained, using appropriate codon bias tables for that host (e.g., microorganism). Nucleic acids may also be optimized to a GC-content preferable to a particular host, and/or to reduce the number of repeat sequences. Such nucleic acids may be referred to as “codon optimized”. As isolated nucleic acids, these modified sequences can exist as purified molecules and can be incorporated into a vector or a virus for use in constructing modules for recombinant nucleic acid constructs. Software for preparing nucleic acid sequences codon optimized for expression in a given host organism is commercial available. For example codon optimization may be performed using the Geneart service from LifeTechnologies or OptimumGene™ Codon Optimization available from GenScript, United States, or GeneArt® Invitrogen.

Diterpene pyrophosphate intermediate

The term ”decalin” as used herein refers to a compound of the formula VII:



(VII)

The numbering of carbon atoms provided in formula VII is adhered to throughout this description.

A compound containing or comprising a " decalin core" as used herein refers to a compound comprising above mentioned structure of formula VII, wherein each of the carbon atoms numbered 1 to 10 may be substituted with one or two substituents. It is possible that two of said substituents are fused to form a ring, and thus compound containing or comprising decalin may contain 3 or more rings.

The term "diterpene pyrophosphate intermediate" as used herein refers to a compound, which is the product of bicyclisation of GGPP in a reaction catalysed by a diTPS class II enzyme. The diterpene pyrophosphate intermediate according to the invention contains a decalin core, and comprises a pyrophosphate group.

It is preferred that the diterpene pyrophosphate intermediate of the invention is a compound containing a decalin core, which is substituted at one of more positions with substituents selected from the group consisting of alkyl, alkenyl and hydroxyl, wherein one of said alkyl or alkenyl is substituted with O-pyrophosphate.

The terms "diphosphate" and "pyrophosphate" are used interchangeably herein. The abbreviation "OPP", "-OPP" or "PPO-" as used herein refers to diphosphate.

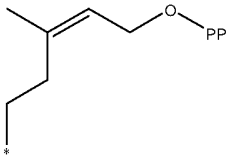
The term "alkyl" as used herein refers to a saturated, straight or branched hydrocarbon chain. The hydrocarbon chain preferably contains of from one to eighteen carbon atoms (C_{1-18}-alkyl), more preferred of from one to six carbon atoms (C_{1-6}-alkyl), including methyl, ethyl, propyl, isopropyl, butyl, isobutyl, secondary butyl, tertiary butyl, pentyl, isopentyl, neopentyl, tertiary pentyl, hexyl and isohexyl.

The term "alkenyl" as used herein refers to a saturated, straight or branched hydrocarbon chain containing at least one double bond. Alkenyl may preferably be any of the alkyls described above containing one or more double bonds.

In particular, the diterpene pyrophosphate intermediate of the invention is a compound containing a decalin core, wherein said decalin is

i. substituted at the 4 position with one or two alkyl, such as with two alkyl, wherein said alkyl for example may be C_{1-3}, alkyl, for example said alkyl may be methyl;
ii. substituted at the 8 position with one or two substituents individually selected from the group consisting of alkyl, hydroxyl and alkenyl, wherein said alkyl for example may be C_{1-3} alkyl, for example said alkyl may be methyl, and said alkenyl may be C_{1-3} alkenyl, for example said alkenyl may be =C;
iii. substituted at the 9 position with alkenyl-O-PP, wherein said alkenyl for example may be branched C4-8-alkenyl, such as branched C5-7-alkenyl, for example branched C6-alkenyl; and
iv. substituted at the 10 position with alkyl, wherein said alkyl for example may be C_{1-3}, alkyl, for example said alkyl may be methyl.

In particular, the substituent at the 9 position may be alkenyl of formula VIII:



wherein the asterisk indicates the point of attachment to the decalin core.

It is also preferred that the stereochemistry around substituents 9 and 10 is predetermined. Thus, said diterpene pyrophosphate intermediate may contain a decalin core substituted as indicated above, wherein the substitutions at the 9 and 10 positions are (9R, 10R), (9S,10S), (9S, 10R) or (9R, 10S), for example the substitutions at the 9 and 10 positions are (9R, 10R), (9S,10S) or (9S, 10R).

In preferred embodiments, the diterpene pyrophosphate intermediate may be any of the diterpene pyrophosphate intermediates shown in figure 3, i.e. the diterpene pyrophosphate intermediate may be selected from the group consisting of (9R,10R)-

copalyl diphosphate, (9S,10S)-copalyl diphosphate, labda-13-en-8-ol diphosphate and (9S, 10R)-copalyl diphosphate.

Diterpenes

The term "diterpene" as used herein refers to a compound derived or prepared from four isoprene units. A diterpene according to the invention is a C_{20}- molecule consisting of 20 carbon atoms, up to three oxygen atoms and hydrogen atoms.

The term "precursor diterpene" as used herein refers to a diterpene, which may serve as the precursor for formation of a diterpenoid. Thus, a precursor diterpene may be a diterpene, which after one or more oxidation reactions becomes a diterpenoid. The precursor diterpene may for example be any of the diterpenes described herein in this section.

The diterpene typically contains one or more ring structures, such as one or more monocyclic, bicyclic, tricyclic or tetracyclic ring structure(s). The diterpene may contain one or more double bonds. Frequently, a diterpene according to the invention contains at least one double bond and often they contain in the range of 1 to 3 double bonds.

The diterpene may comprise up to three oxygen atom, although it is also possible that the diterpene contains no oxygen and consists solely of carbon and hydrogen atoms. The oxygen atom are generally present in the form of hydroxyl groups, or part of a ring structure.

In principle, the methods of the invention can be used to produce any diterpenoid by selecting an appropriate combination of diTPS of class II and diTPS of class I and CYP. A suitable precursor diterpene may be produced by selecting an appropriate combination of diTPS of class II and diTPS of class I.

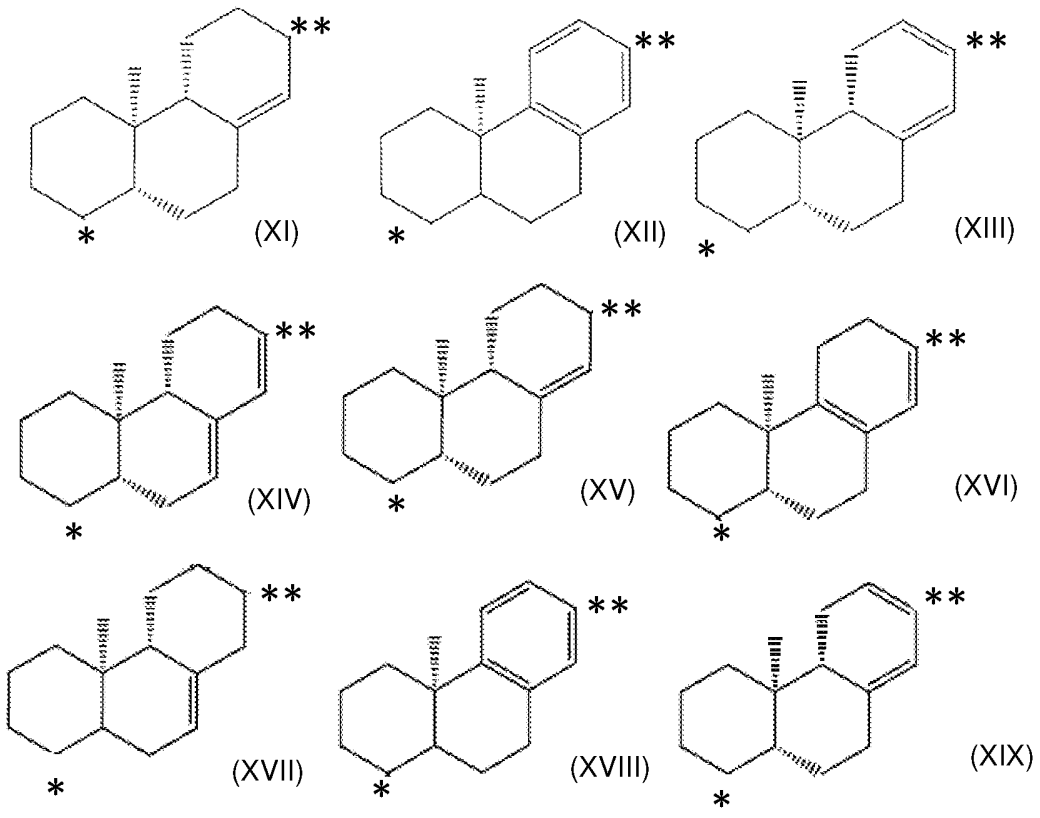
In one preferred embodiment a diterpene according to the present invention is a C_{20}-molecule containing a decalin core structure.

As used herein the term "containing a core structure of formula " or the term "containing a core of formula" refers to a molecule containing a structure of the indicated formula,

wherein said structure may be substituted at one or more positions. The term "substituted" as used herein in relation to organic compounds refer to one hydrogen being substituted with another group or atom.

Said decalin may be substituted at one or more positions, and it is also contained within the invention that two substituents are fused, thus leading to a tricyclic or higher cyclic structure.

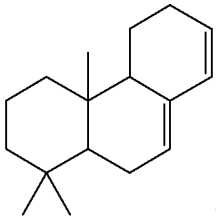
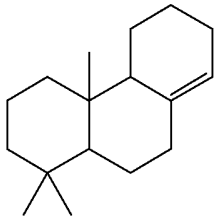
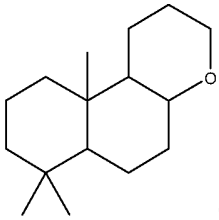
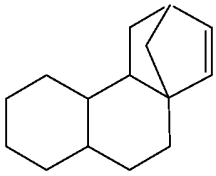
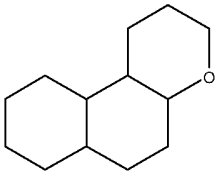
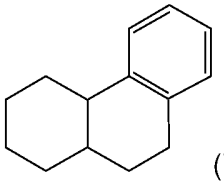
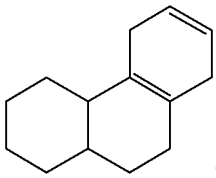
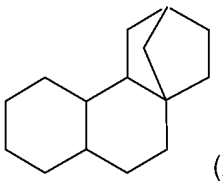
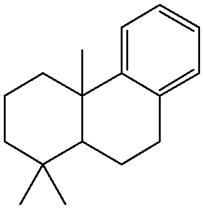
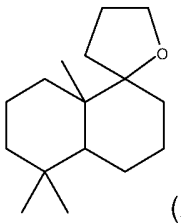
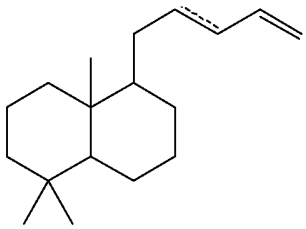
In particular, the diterpene may be a C_{20}-molecule containing a core structure of one of following formulas XI, XII, XIII, XIV, XV, XVI, XVII, XVIII or XIX:



The diterpene containing a core structure of any of formulas XI, XII, XIII, XIV, XV, XVI, XVII, XVIII or XIX, may be a C_{20}-molecule consisting of the formulas XI, XII, XIII, XIV, XV, XVI, XVII, XVIII or XIX substituted at one or more positions. In particular, said diterpene may be a C_{20}-molecule substituted at the position marked by * with one or two alkyl, such as one or two C_{1-3}-alkyl, such as with one or two methyl groups. In

addition said diterpene may be substituted at the position marked by ** with one or two groups individually selected from alkyl and alkenyl. Said alkyl may for example be C_{1-6}-alkyl, such as C_{1-3}-alkyl, for example isopropyl or methyl. Said alkenyl may me C_{1-6} alkenyl, such as C_{2-4}-alkenyl, such as C_{2-3}-alkenyl.

In preferred embodiments of the invention the diterpene to be produced may be a C_{20}-molecule containing a core structure of one of following formulas I, II, III, IV, V, VI, IX, X, XXII, XXIII or XXIV:


(I) or

(II) or

(III) or

(IV) or

(V) or

(VI) or

(IX) or

(X) or

(XXII) or

(XXIII) or

(XXIV).

The diterpene containing a core structure of any of formulas I, II, III, IV, V, VI, IX, X, XXII, XXIII or XXIV, may be a C_{20}-molecule consisting of the formulas I, II, III, IV, V, VI, IX, X, XXII, XXIII or XXIV substituted at one or more positions, for example by one or more groups selected from the group consisting of:

a) alkyl, such as C_{1-6}-alkyl, for example C_{1-3}, wherein said alkyl may be linear or branched, for example alkyl may be isopropyl or methyl

b) alkenyl, such as C_{1-6} alkenyl, such as C_{2-4}-alkenyl, such as C_{2-3}-alkenyl
c) hydroxyl

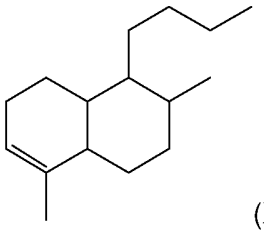
In particular said diterpene containing a core structure of any of formulas formulas I, II, III, IV, V, VI, IX or X, may be a C_{20}-molecule substituted

a) at the position corresponding to the 4 position of decalin with one or two alkyl, such as one or two C_{1-3}-alkyl, such as with one or two methyl groups, for example with two methyl; and/or
b) at the position corresponding to the 10 position of decalin with alkyl, such as with C_{1-3}-alkyl, such as with methyl; and/or
c) at the position corresponding to the position marked by ** in relations to formulas XI-XIX, with one or two groups individually selected from alkyl and alkenyl. Said alkyl may for example be C_{1-6}-alkyl, such as C_{1-3}-alkyl, for example isopropyl or methyl. Said alkenyl may me C_{1-6} alkenyl, such as C_{2-4}-alkenyl, such as C_{2-3}-alkenyl; and/or
d) hydroxyl.

The diterpene according to the invention may also be a C_{20}-molecule consisting of 20 carbon atoms, up to three oxygen atoms and hydrogen atoms, and which contains a core structure of any of formulas I, II, III, IV, VI, X, XII, XLII, XLIII, XLIV, XLV, XLVI, XXVII, XXVIII, XXIX, XXX, XXXI, XXXII, XXXIII, XXXIV, XXXV, XXXVI, XXXVIII, XXXIX, XL and/or XLI.

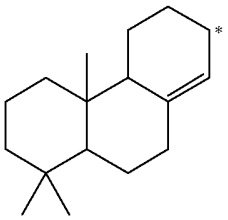
The structure of the formulas I, II, III, IV, VI, X, XII, XLII, XLIII, XLIV, XLV, XLVI, XXVII, XXVIII, XXIX, XXX, XXXI, XXXII, XXXIII, XXXIV, XXXV, XXXVI, XXXVII, XXXVIII, XXXIX, XL and XLI are as indicated herein above.

In one embodiment the diterpene is a C_{20}-molecule containing a core of formula XXXIII:

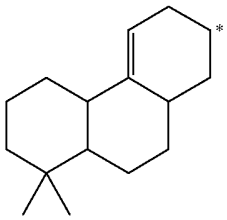


(XXXIII). Said diterpene may in particular contain a core of formula XXXIII substituted with alkyl, alkenyl and/or hydroxyl, preferably substituted with methyl, $=CH_2$ and hydroxyl.

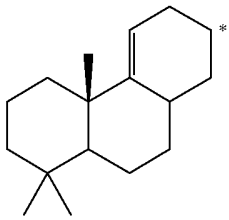
In another embodiment the diterpene is a C_{20}-molecule containing a core of any of formulas II, XXXV, XXXVI and/or XXXVII:



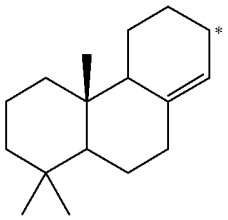
(II), or



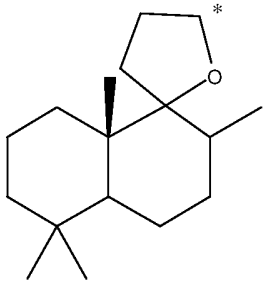
(XXXV) or



(XXXVI) or

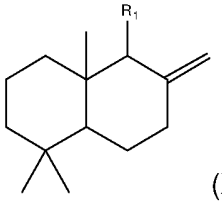


(XXXVII) or or



(XLI), wherein said core may be substituted with one or more alkyl or alkenyl . In particular, the position marked by asterisk may be substituted with one or two substituents selected from the group consisting of C_{1-2}-alkyl and C_{1-2}-alkenyl, preferably the position marked by asterisk may be substituted with one methyl group and ethenyl group.

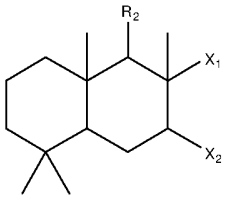
In one embodiment, said diterpene to be produced is a C_{20}-molecule containing a decalin substituted at the 10 position with C_5-alkenyl chain, which optionally may be substituted with a hydroxyl and/or a methyl group and/or =C. For example, said diterpene may be a C_{20}-molecule of the formula XX:



(XX),

wherein R_1 is a C_5-alkenyl substituted with methyl and/or hydroxyl. Preferably, R_1 is C_5-alkenyl containing one or two double bonds. When R_1 is alkenyl containing one double bond, said alkenyl is preferably substituted with hydroxyl and methyl. When R_1 is alkenyl containing two double bonds, said alkenyl is preferably substituted with methyl.

For example, said diterpene may be a C_{20}-molecule of the formula XXI:



(XXI)

wherein R_2 is a C_5-alkenyl substituted with methyl and/or hydroxyl or with =C, and X_1 is either –OH or methyl, and X_2 is either –H or –OH, wherein one and only one of X_1 and X_2 is -OH. Preferably, R_2 is C_5-alkenyl containing one or two double bonds. When R_2 is alkenyl containing one double bond, said alkenyl is preferably substituted with hydroxyl and methyl or with =C. When R_2 is alkenyl containing two double bonds, said alkenyl is preferably substituted with methyl.

It is also comprised within the invention that the diterpene is the product of any of the reactions VII to XIX described herein above.

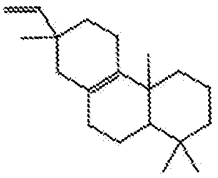
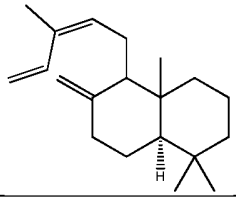
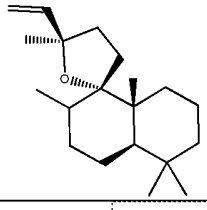
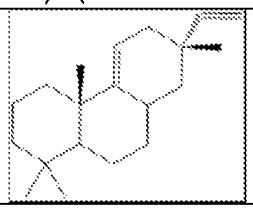
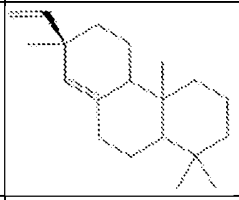
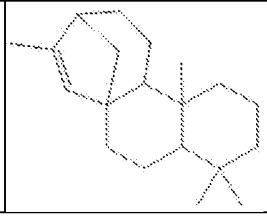
In particular, the diterpene may be any of the compounds 1 to 47 shown in figure 2 and/or Table 1.

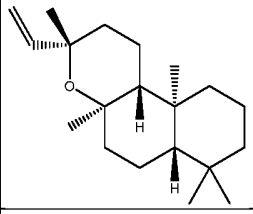
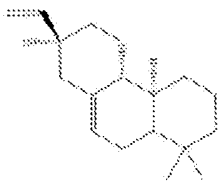
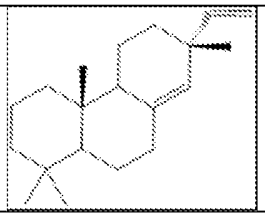
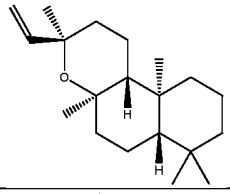
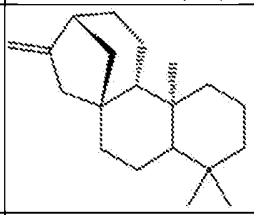
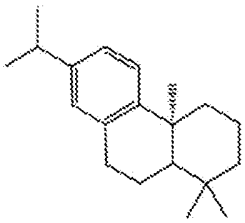
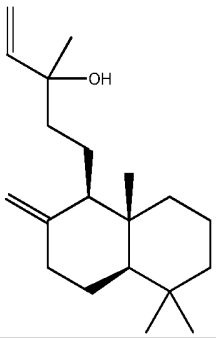
It is preferred that the diterpene is not 13R-manoyl oxide.

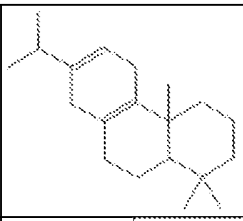
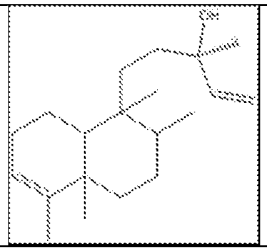
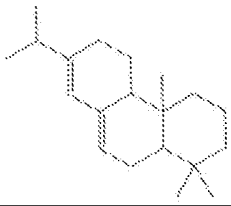
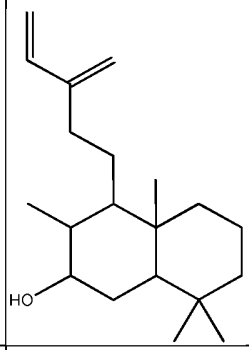
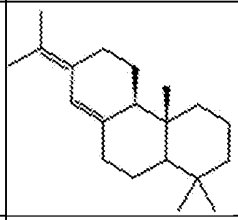
Table 1 shows preferred diterpenes, which may be precursor diterpenes according to the present invention. The left hand column provides a number of the diterpene used herein, the middle column provides the retention index (RI) and the right hand column the structure. The retention index (RI) is determined by GC-MS and may also be referred to as Kovat's retention index.

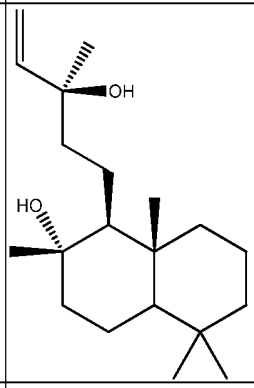
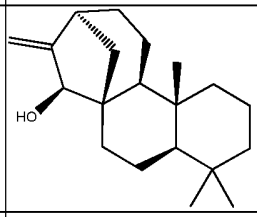
Table 1

Compound	RI	structure

(1)	1906	
(2)	1923	
(3)	1938	
(4)	1938	
(5)	1952	
(6)	1953	
(7)	1957	
(8)	1967	
(9)	1972	
(10)	1975	
(11)	1981	
(12)	1989	
(13)	1994	
(14)	1997	
(15)	2000	

(16)	2002	
(17)	2005	
(19)	2014	
(18)	2020	
(20)	2026	
(21)	2055	
(22)	2062	
(23)	2065	
(24)	2072	

(25)	2074	
(26)	2074	
(27)	2091	
(28)	2095	
(29)	2098	
(30)	2115	
(31)	2120	
(32)	2128	
(33)	2129	
(34)	2158	
(35)	2167	

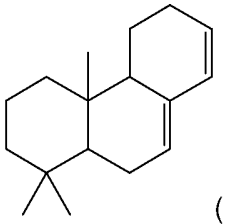
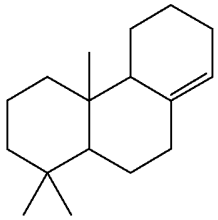
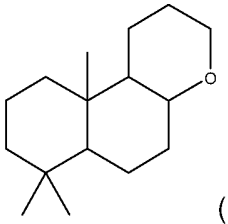
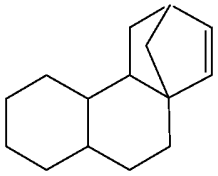
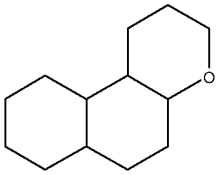
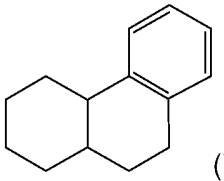
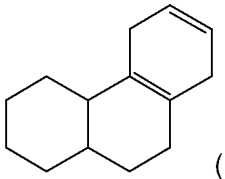
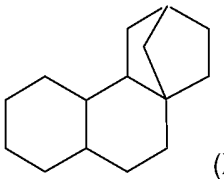
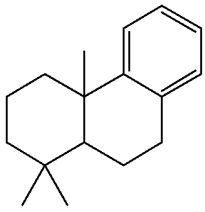
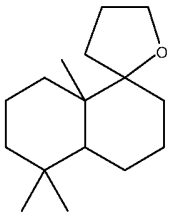
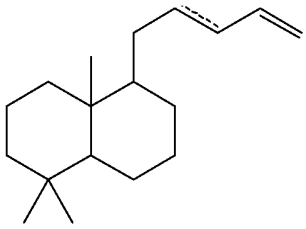
(36)	2186	
(37)	2192	
(38)	2194	
(39)	2199	
(40)	2202	
(41)	2215	
(42)	2227	
(43)	2230	
(44)	2236	
(45)	2244	
(46)	2246	
(47)	2268	

Diterpenoid

The term “diterpenoid” refers to a diterpene, which has been functionalised by addition of one or more functional groups. Said diterpene may be any of the diterpenes described herein above in the section “Diterpene”. Preferably, the diterpenoid according to the present invention is any of the diterpenes described herein above in the section “Diterpene” substituted at one or more positions with substituent(s) selected from the group consisting of –OH, =O and –COOH

Thus, the diterpenoid may be a C_{20}-molecule substituted at one or more positions with substituent(s) selected from the group consisting of –OH, =O and –COOH. Preferably, the diterpenoid is a C_{20}-molecule containing a decalin core and one or more substitutents selected from the group consisting of –OH, =O and/or –COOH.

In a preferred embodiment of the invention the diterpenoid is a C_{20}-molecule containing a core structure of one of following formulas I, II, III, IV, V, VI, IX, X, XXII, XXIII or XXIV:


(I) or

(II) or

(III) or

(IV) or

(V) or

(VI) or

(IX) or

(X) or

(XXII) or

(XXIII) or

(XXIV).

In particular, the diterpenoid may be a C_{20}-molecule containing a cores structure of formulas I, II, III, IV, V, VI, IX, X, XXII, XXIII or XXIV substituted at one or more positions by one or more groups selected from the group consisting of:

a) alkyl, such as C_{1-6}-alkyl, for example C_{1-3}, wherein said alkyl may be linear or branched, for example alkyl may be isopropyl or methyl, wherein said alkyl may be substituted with one or more groups selected from the group consisting of –OH, =O and –COOH;

b) alkenyl, such as C_{1-6} alkenyl, such as C_{2-4}-alkenyl, such as C_{2-3}-alkenyl, wherein said alkyl may be substituted with one or more groups selected from the group consisting of –OH, =O and –COOH;
c) –OH;
d) =O; and
e) –COOH.

The diterpenoid according to the invention may also be a C_{20}-molecule, which contains a core structure of any of formulas I, II, III, IV, VI, X, XII, XLII, XLIII, XLIV, XLV, XLVI, XXVII, XXVIII, XXIX, XXX, XXXI, XXXII, XXXIII, XXXIV, XXXV, XXXVI, XXXVIII, XXXIX, XL and/or XLI, and substituted at one or more positions by one or more groups selected from the group consisting of:

a) alkyl, such as C_{1-6}-alkyl, for example C_{1-3}, wherein said alkyl may be linear or branched, for example alkyl may be isopropyl or methyl, wherein said alkyl may be substituted with one or more groups selected from the group consisting of –OH, =O and –COOH;
b) alkenyl, such as C_{1-6} alkenyl, such as C_{2-4}-alkenyl, such as C_{2-3}-alkenyl, wherein said alkyl may be substituted with one or more groups selected from the group consisting of –OH, =O and –COOH;
c) –OH;
d) =O; and
e) –COOH.

The diterpenoid according to the invention may also be a C_{20}-molecule which contains a core structure of any of formulas I, II, IV, VI, X, XII, XLII, XLIII, XLIV, XLVI, XXVII, XXVIII, XXIX, XXX, XXXI, XXXIII, XXXIV, XXXV, XXXVI, XXXVII, XXXVIII, XXXIX, XL and/or XLI, and which is substituted at one or more positions by one or more groups selected from the group consisting of:

a) alkyl, such as C_{1-6}-alkyl, for example C_{1-3}, wherein said alkyl may be linear or branched, for example alkyl may be isopropyl or methyl, wherein said alkyl may be substituted with one or more groups selected from the group consisting of –OH, =O and –COOH;

b) alkenyl, such as C_{1-6} alkenyl, such as C_{2-4}-alkenyl, such as C_{2-3}-alkenyl, wherein said alkyl may be substituted with one or more groups selected from the group consisting of –OH, =O and –COOH;
c) –OH;
d) =O; and
e) –COOH.

In one embodiment of the invention the diterpenoid is a diterpene, which has been substituted at one or more positions with substituent(s) selected from the group consisting of –OH, =O and –COOH, wherein said diterpene is the product of any of the reactions VII to XIX described herein above.

In another embodiment of the invention the diterpenoid is any of the dierpenes describes herein above in the section “Diterpenes”, which has been substituted at one or more positions with substituent(s) selected from the group consisting of –OH, =O and –COOH .

In a preferred embodiment of the invention the diterpenoid is a diterpene, which has been substituted at one or more positions with substituent(s) selected from the group consisting of –OH, =O and –COOH, wherein the diterpene is any of the compounds 1 to 47 of Table 1.

In one embodiment the diterpenoid is dehydroabietadiene substituted at one or more positions with substituent(s) selected from the group consisting of –OH, =O and –COOH.

In another embodiment of the invention the diterpenoid is miltiradiene substituted at one or more positions with substituent(s) selected from the group consisting of –OH, =O and –COOH.

In another embodiment of the invention the diterpenoid is manool substituted at one or more positions with substituent(s) selected from the group consisting of –OH, =O and –COOH.

In another embodiment of the invention the diterpenoid is syn-manool substituted at one or more positions with substituent(s) selected from the group consisting of –OH, =O and –COOH.

In another embodiment of the invention the diterpenoid is syn-pimara-9,(11),15-diene substituted at one or more positions with substituent(s) selected from the group consisting of –OH, =O and –COOH.

In another embodiment of the invention the diterpenoid is syn-isopimara-7,15-diene substituted at one or more positions with substituent(s) selected from the group consisting of –OH, =O and –COOH.

In one embodiment of the invention the diterpenoid is selected from the group consisting of hydroxy dehydroabietadiene, hydroxy miltiradiene and hydroxy syn-manool..

In one embodiment of the invention the diterpenoid is carboxyl miltradiene, for example miltiradienic acid.

In one embodiment of the invention the diterpenoid is selected from the group consisting of monohydroxy-manool, dihydroxymanool and keto-hydroxy-manool.

In one embodiment of the invention, the diterpenoid is selected from the group consisting of syn-pimara-9,(11),15-diene oxidized to contain either one keto- and one hydroxyl group or a single carboxylic acid group.

In one embodiment the diterpenoid is ferruginol.

A diterpenoid intermediate according to the present invention is a diterpene, which has been substituted with one or more groups selected from the group consisting of –OH, =O and –COOH, wherein oxidation of said diterpenoid intermediate leads to formation of a desired diterpenoid. Accordingly, the diterpenoid intermediate may be any of the diterpenoids described herein in this section.

In one embodiment, the diterpenoid is not forskolin.

Host organism

The host organism to be used with the methods of the invention, may be any suitable host organism containing
a heterologous nucleic acid encoding a diTPS of class II , which may be any of diTPS of class II described herein in any of the sections “diTPS of class II”, “syn-CPP type diTPS”, “ent-CPP type diTPS”, “(+)-CPP type diTPS”, “LPP type diTPS”, and “LPP like type diTPS”; and/or
a heterologous nucleic acid encoding a diTPS of class I, which may be any of diTPS of class I described herein in any of the sections “diTPS of class I”, “EpTPS8”, “EpTPS23”, “SsSCS”, “CfTPS3”, “CfTPS4”, “MvTPS5”, “TwTPS2”, “EpTPS1” , and “CfTPS14” and/or
a heterologous nucleic acid encoding a CYP, which may be any of CYPs described herein in any of the sections “CYP”, “CYP76”, “CYP71” and “CYP720”.

Suitable host organisms include microorganisms, plant cells, and plants.

The microorganism can be any microorganism suitable for expression of heterologous nucleic acids. In one embodiment the host organism of the invention is a eukaryotic cell. In another embodiment the host organism is a prokaryotic cell.
In a preferred embodiment, the host organism is a fungal cell such as a yeast or filamentous fungus. In particular the host organism may be a yeast cell.

In a further embodiment the yeast cell is selected from the group consisting of Saccharomyces cerevisiae, Schizosaccharomyces pombe, Yarrowia lipolytica, Candida glabrata, Ashbya gossypii, Cyberlindnera jadinii, and Candida albicans.
In general, yeasts and fungi are excellent microorganism to be used with the present invention. They offer a desired ease of genetic manipulation and rapid growth to high cell densities on inexpensive media. For instance yeasts grow on a wide range of carbon sources and are not restricted to glucose. Thus, the microorganism to be used with the present invention may be selected from the group of yeasts described below:

Arxula adeninivorans (Blastobotrys adeninivorans) is a dimorphic yeast (it grows as a budding yeast like the baker’s yeast up to a temperature of 42 °C, above this threshold it grows in a filamentous form) with unusual biochemical characteristics. It can grow on

a wide range of substrates and can assimilate nitrate. It has successfully been applied to the generation of strains that can produce natural plastics or the development of a biosensor for estrogens in environmental samples.

Candida boidinii is a methylotrophic yeast (it can grow on methanol). Like other methylotrophic species such as Hansenula polymorpha and Pichia pastoris, it provides an excellent platform for the production of heterologous proteins. Yields in a multigram range of a secreted foreign protein have been reported. A computational method, IPRO, recently predicted mutations that experimentally switched the cofactor specificity of Candida boidinii xylose reductase from NADPH to NADH. Details on how to download the software implemented in Python and experimental testing of predictions are outlined in the following paper.

Hansenula polymorpha (Pichia angusta) is another methylotrophic yeast (see Candida boidinii). It can furthermore grow on a wide range of other substrates; it is thermotolerant and can assimilate nitrate (see also Kluyveromyces lactis). It has been applied to the production of hepatitis B vaccines, insulin and interferon alpha-2a for the treatment of hepatitis C, furthermore to a range of technical enzymes.

Kluyveromyces lactis is a yeast regularly applied to the production of kefir. It can grow on several sugars, most importantly on lactose which is present in milk and whey. It has successfully been applied among others to the production of chymosin (an enzyme that is usually present in the stomach of calves) for the production of cheese. Production takes place in fermenters on a 40,000 L scale.

Pichia pastoris is a methylotrophic yeast (see Candida boidinii and Hansenula polymorpha). It provides an efficient platform for the production of foreign proteins. Platform elements are available as a kit and it is worldwide used in academia for the production of proteins. Strains have been engineered that can produce complex human N-glycan (yeast glycans are similar but not identical to those found in humans).

Saccharomyces cerevisiae is the traditional baker's yeast known for its use in brewing and baking and for the production of alcohol. As protein factory it has successfully been applied to the production of technical enzymes and of pharmaceuticals like insulin and hepatitis B vaccines. Also it has been useful for production of terpenoids.

Yarrowia lipolytica is a dimorphic yeast (see Arxula adeninivorans) that can grow on a wide range of substrates. It has a high potential for industrial applications.

In another embodiment the host organism is a microalgae such as Chlorella and Prototheca.

In another embodiment of the invention the host organism is a filamentous fungus, for example Aspergillus.

In further yet another embodiment the host organism is a plant cell. The host organism may be a cell of a higher plant, but the host organism may also be cells from organisms not belonging to higher plants for example cells from the moss Physcomitrella patens.

In another embodiment the host organism is a mammalian cell, such as a human, feline, porcine, simian, canine, murine, rat, mouse or rabbit cell.

As mentioned, the host organism can also be a prokaryotic cell such as a bacterial cell. If the host organism is a prokaryotic cell the cell may be selected from, but not limited to E. coli, Corynebacterium, Bacillus, Pseudomonas and Streptomyces cells.

The host organism may also be a plant.

A plant or plant cell can be transformed by having a heterologous nucleic acid integrated into its genome, *i.e.*, it can be stably transformed. Stably transformed cells typically retain the introduced nucleic acid with each cell division. A plant or plant cell can also be transiently transformed such that the recombinant gene is not integrated into its genome. Transiently transformed cells typically lose all or some portion of the introduced nucleic acid with each cell division such that the introduced nucleic acid cannot be detected in daughter cells after a certain number of cell divisions. Both transiently transformed and stably transformed transgenic plants and plant cells can be useful in the methods described herein.

Plant cells comprising a heterologous nucleic acid used in methods described herein can constitute part or all of a whole plant. Such plants can be grown in a manner

suitable for the species under consideration, either in a growth chamber, a greenhouse, or in a field. Plants may also be progeny of an initial plant comprising a heterologous nucleic acid provided the progeny inherits the heterologous nucleic acid. Seeds produced by a transgenic plant can be grown and then selfed (or outcrossed and selfed) to obtain seeds homozygous for the nucleic acid construct.

The plants to be used with the invention can be grown in suspension culture, or tissue or organ culture. For the purposes of this invention, solid and/or liquid tissue culture techniques can be used. When using solid medium, plant cells can be placed directly onto the medium or can be placed onto a filter that is then placed in contact with the medium. When using liquid medium, transgenic plant cells can be placed onto a flotation device, *e.g.*, a porous membrane that contacts the liquid medium.

When transiently transformed plant cells are used, a reporter sequence encoding a reporter polypeptide having a reporter activity can be included in the transformation procedure and an assay for reporter activity or expression can be performed at a suitable time after transformation. A suitable time for conducting the assay typically is about 1-21 days after transformation, *e.g.*, about 1-14 days, about 1-7 days, or about 1-3 days. The use of transient assays is particularly convenient for rapid analysis in different species, or to confirm expression of a heterologous polypeptide whose expression has not previously been confirmed in particular recipient cells.

Techniques for introducing nucleic acids into monocotyledonous and dicotyledonous plants are known in the art, and include, without limitation, *Agrobacterium*-mediated transformation, viral vector-mediated transformation, electroporation and particle gun transformation, U.S. Patent Nos 5,538,880; 5,204,253; 6,329,571; and 6,013,863. If a cell or cultured tissue is used as the recipient tissue for transformation, plants can be regenerated from transformed cultures if desired, by techniques known to those skilled in the art.

The plant comprising a heterologous nucleic acid to be used with the present invention may for example be selected from: corn (Zea. mays), canola (Brassica napus, Brassica rapa ssp.), alfalfa (Medicago sativa), rice (Oryza sativa), rye (Secale cerale), sorghum (Sorghum bicolor, Sorghum vulgare), sunflower (Helianthus annuas), wheat (Tritium aestivum and other species), Triticale, Rye (Secale) soybean (Glycine max), tobacco

(Nicotiana tabacum or Nicothiana Benthamiana), potato (Solanum tuberosum), peanuts (Arachis hypogaea), cotton (Gossypium hirsutum), sweet potato (Impomoea batatus), cassava (Manihot esculenta), coffee (Cofea spp.), coconut (Cocos nucifera), pineapple (Anana comosus), citrus (Citrus spp.) cocoa (Theobroma cacao), tea (Camellia senensis), banana (Musa spp.), avacado (Persea americana), fig (Ficus casica), guava (Psidium guajava), mango (Mangifer indica), olive (Olea europaea), papaya (Carica papaya), cashew (Anacardium occidentale), macadamia (Macadamia intergrifolia), almond (Primus amygdalus), apple (Malus spp), Pear (Pyrus spp), plum and cherry tree (Prunus spp), Ribes (currant etc.), Vitis, Jerusalem artichoke (Helianthemum spp), non-cereal grasses (Grass family), sugar and fodder beets (Beta vulgaris), chicory, oats, barley, vegetables, and ornamentals.

For example, plants of the present invention are crop plants (for example, cereals and pulses, maize, wheat, potatoes, tapioca, rice, sorghum, millet, cassava, barley, pea, sugar beets, sugar cane, soybean, oilseed rape, sunflower and other root, tuber or seed crops. Other important plants maybe fruit trees, crop trees, forest trees or plants grown for their use as spices or pharmaceutical products (Mentha spp, clove, Artemesia spp, Thymus spp, Lavendula spp, Allium spp., Hypericum, Catharanthus spp, Vinca spp, Papaver spp., Digitalis spp, Rawolfia spp., Vanilla spp., Petrusilium spp., Eucalyptus, tea tree, Picea spp, Pinus spp, Abies spp, Juniperus spp,. Horticultural plants which may be used with the present invention may include lettuce, endive, and vegetable brassicas including cabbage, broccoli, and cauliflower, carrots, and carnations and geraniums.

The plant may also be selected from the group consisting of tobacco, cucurbits, carrot, strawberry, sunflower, tomato, pepper and Chrysanthemum.

The plant may also be a grain plants for example oil-seed plants or leguminous plants. Seeds of interest include grain seeds, such as corn, wheat, barley, sorghum, rye, etc. Oil-seed plants include cotton soybean, safflower, sunflower, Brassica, maize, alfalfa, palm, coconut, etc. Leguminous plants include beans and peas. Beans include guar, locust bean, fenugreek, soybean, garden beans, cowpea, mung bean, lima bean, fava bean, lentils, chickpea.

In a further embodiment of the invention said plant is selected from the following group: maize, rice, wheat, sugar beet, sugar cane, tobacco, oil seed rape, potato and soybean. Thus, the plant may for example be rice.

The whole genome of *Arabidopsis thaliana* plant has been sequenced (The Arabidopsis Genome Initiative (2000). "Analysis of the genome sequence of the flowering plant Arabidopsis thaliana". *Nature* **408** (6814): 796–815. doi:10.1038/35048692. PMID 11130711). Consequently, very detailed knowledge is available for this plant and it may therefore be a useful plant to work with.

Accordingly, one plant, which may be used with the present invention is an *Arabidopsis* and in particular an *Arabidopsis thaliana*.

In embodiments of the invention, wherein one or more host organism(s) are plants, then the diterpenoid may be isolated from said plant by conventional methods. Thus the diterpenoid may be isolated from the entire plant or from parts thereof. Said parts may for example be selected from the group consisting of leaves, fruits, flowers, stems, seeds and roots of plants. Intermediates may also be isolated from said plants or parts thereof.

It may be preferred that the host organism does not naturally produce the diterpenoid to be produced by the methods of the invention.

In one embodiment of the invention, the host organism may comprise at least the following heterologous nucleic acids:

a) a heterologous nucleic acid encoding CfTPS1 of SEQ ID NO:5 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith; and

b) a heterologous nucleic acid encoding CfTPS3 of SEQ ID NO:12 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith;

c) a heterologous nucleic acid encoding CYP76AH11 of SEQ ID NO:21 or a functional homologue thereof sharing at least 70%, such as at least 80%, such

as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith.

Such a host organism is in particular useful for production of diterpenoids having a core of formulas VI, IX, XXXVIII, XXXIX or XL, for example for production of oxidised dehydroabietadien and/or production of oxidised miltiradiene, e.g hydroxyl dehydroabietadiene or hydroxyl miltiradiene.

In one embodiment of the invention, the host organism may comprise at least the following heterologous nucleic acids:

a) a heterologous nucleic acid encoding TwTPS7 of SEQ ID NO:4 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith; and
b) a heterologous nucleic acid encoding CfTPS3 of SEQ ID NO:12 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith;
c) a heterologous nucleic acid encoding CYP76AH11 of SEQ ID NO:21 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith.

Such a host organism is in particular useful for production of diterpenoids having a core of formulas VI, IX, XXXVIII, XXXIX or XL, for example for production of oxidised dehydroabietadiene and/or production of oxidised miltiradiene, e.g hydroxyl dehydroabietadiene or hydroxyl miltiradiene.

In one embodiment of the invention, the host organism may comprise at least the following heterologous nucleic acids:

a) a heterologous nucleic acid encoding CfTPS1 of SEQ ID NO:5 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith; and
b) a heterologous nucleic acid encoding CfTPS3 of SEQ ID NO:12 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least

85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith;

c) a heterologous nucleic acid encoding CYP71D381 of SEQ ID NO:22 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith.

Such a host organism is in particular useful for production of diterpenoids having a core of formulas VI, IX, XXXVIII, XXXIX or XL, for example for production of oxidised dehydroabietadiene and/or production of oxidised miltiradiene, e.g hydroxyl dehydroabietadiene or hydroxyl miltiradiene.

In one embodiment of the invention, the host organism may comprise at least the following heterologous nucleic acids:

a) a heterologous nucleic acid encoding TwTPS7 of SEQ ID NO:4 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith; and

b) a heterologous nucleic acid encoding CfTPS3 of SEQ ID NO:12 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith;

c) a heterologous nucleic acid encoding CYP71D381 of SEQ ID NO:22 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith.

Such a host organism is in particular useful for production of diterpenoids having a core of formulas VI, IX, XXXVIII, XXXIX or XL, for example for production of oxidised dehydroabietadiene and/or production of oxidised miltiradiene, e.g hydroxyl dehydroabietadiene or hydroxyl miltiradiene.

In one embodiment of the invention, the host organism may comprise at least the following heterologous nucleic acids:

a) a heterologous nucleic acid encoding CfTPS1 of SEQ ID NO:5 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least

85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith; and

b) a heterologous nucleic acid encoding CfTPS3 of SEQ ID NO:12 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith;

c) a heterologous nucleic acid encoding CYP720B4 of SEQ ID NO:23 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith.

Such a host organism is in particular useful for production of diterpenoids having a core of formulas VI, IX, XXXVIII, XXXIX or XL, for example for production of oxidised dehydroabietadien and/or production of oxidised miltiradiene, e.g hydroxyl dehydroabietadiene or hydroxyl miltiradiene.

In one embodiment of the invention, the host organism may comprise at least the following heterologous nucleic acids:

a) a heterologous nucleic acid encoding Ossyn-CPP of SEQ ID NO:1 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith; and

b) a heterologous nucleic acid encoding CfTPS3 of SEQ ID NO:12 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith;

c) a heterologous nucleic acid encoding PsCYP720B4 of SEQ ID NO:23 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith.

Such a host organism is in particular useful for production of diterpenoids having a core of formulas II, VI, XXXV, XXXVI, XXXVII or XXXVIII, and in particular for for production of diterpenoids having a core of formulas II, VI, XXXV, XXXVI, XXXVII or XXXVIII and comprising at least one keto group, hydroxyl group and/or carboxylic acid group. Thus, for example said host organism may be useful for production of oxidised syn-pimara-9,(11),15-diene, oxidised syn-isopimara-7,15-diene and/or oxidised

dehydroabietadiene, e.g keto syn-pimara-9,(11),15-diene, keto syn-isopimara-7,15-diene, keto dehydroabietadiene, hydroxyl-syn-pimara-9,(11),15-diene, hydroxyl-syn-isopimara-7,15-diene, hydroxy-dehydroabietadiene, carboxyl-syn-pimara-9,(11),15-diene, carboxyl-syn-isopimara-7,15-diene and/or carboxyl-dehydroabietadiene.

In one embodiment of the invention, the host organism may comprise at least the following heterologous nucleic acids:

a) a heterologous nucleic acid encoding CfTPS1 of SEQ ID NO:5 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith; and
b) a heterologous nucleic acid encoding SsSCS of SEQ ID NO:11 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith;
c) a heterologous nucleic acid encoding CYP76AH8 of SEQ ID NO:20 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith.

Such a host organism is in particular useful for production of diterpenoids having a core of formulas XXIX or XLVI, for example for production of oxidised manool, such as hydroxy manool or dihydroxy manool or keto-hydroxy-manool.

In one embodiment of the invention, the host organism may comprise at least the following heterologous nucleic acids:

a) a heterologous nucleic acid encoding CfTPS1 of SEQ ID NO:5 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith; and
b) a heterologous nucleic acid encoding SsSCS of SEQ ID NO:11 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith;
c) a heterologous nucleic acid encoding CYP76AH11 of SEQ ID NO:21 or a functional homologue thereof sharing at least 70%, such as at least 80%, such

as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith.

Such a host organism is in particular useful for production of diterpenoids having a core of formulas XXIX or XLVI, for example for production of oxidised manool, such as hydroxy manool or dihydroxy manool or keto-hydroxy-manool.

In one embodiment of the invention, the host organism may comprise at least the following heterologous nucleic acids:

a) a heterologous nucleic acid encoding CfTPS1 of SEQ ID NO:5 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith; and
b) a heterologous nucleic acid encoding SsSCS of SEQ ID NO:11 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith;
c) a heterologous nucleic acid encoding CYP720B4 of SEQ ID NO:23 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith.

Such a host organism is in particular useful for production of diterpenoids having a core of formulas XXIX or XLVI, for example for production of oxidised manool, such as hydroxy manool or dihydroxy manool or keto-hydroxy-manool or a carboxylic acid of manool.

In one embodiment of the invention, the host organism may comprise at least the following heterologous nucleic acids:

a) a heterologous nucleic acid encoding Ossyn-CPP of SEQ ID NO:1 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith; and
b) a heterologous nucleic acid encoding SsSCS of SEQ ID NO:11 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith;

c) a heterologous nucleic acid encoding CYP71D381 of SEQ ID NO:22 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith.

Such a host organism is in particular useful for production of diterpenoids having a core of formulas XXVII or XLVI, for example for production of oxidised syn-manool, such as hydroxylated syn-manool.

In one embodiment of the invention, the host organism may comprise at least the following heterologous nucleic acids:

a) a heterologous nucleic acid encoding CfTPS1 of SEQ ID NO:5 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith; and

b) a heterologous nucleic acid encoding SsSCS of SEQ ID NO:11 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith;

c) a heterologous nucleic acid encoding CYP76AH8 of SEQ ID NO:20 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith; and

d) a heterologous nucleic acid encoding CYP71D381 of SEQ ID NO:22 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith.

Such a host organism is in particular useful for production of diterpenoids having a core of formulas XXIX or XLVI, for example for production of oxidised manool, and in particular for production of highly oxidised manool. Thus, such host organism may be particularly useful for production of dihydroxy manool.

In one embodiment of the invention, the host organism may comprise at least the following heterologous nucleic acids:

a) a heterologous nucleic acid encoding CfTPS1 of SEQ ID NO:5 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least

85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith; and

b) a heterologous nucleic acid encoding SsSCS of SEQ ID NO:11 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith;

c) a heterologous nucleic acid encoding CYP76AH8 of SEQ ID NO:20 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith; and

d) a heterologous nucleic acid encoding CYP76AH11 of SEQ ID NO:21 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith.

Such a host organism is in particular useful for production of diterpenoids having a core of formulas XXIX or XLVI, for example for production of oxidised manool, and in particular for production of highly oxidised manool. Thus, such host organism may be particularly useful for production of dihydroxy manool.

In one embodiment of the invention, the host organism may comprise at least the following heterologous nucleic acids:

a) a heterologous nucleic acid encoding CfTPS1 of SEQ ID NO:5 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith; and

b) a heterologous nucleic acid encoding SsSCS of SEQ ID NO:11 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith;

c) a heterologous nucleic acid encoding CYP76AH8 of SEQ ID NO:20 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith; and

d) a heterologous nucleic acid encoding CYP720B4 of SEQ ID NO:23 or a functional homologue thereof sharing at least 70%, such as at least 80%, such

as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith.

Such a host organism is in particular useful for production of diterpenoids having a core of formulas XXIX or XLVI, for example for production of oxidised manool, and in particular for production of highly oxidised manool. Thus, such host organism may be particularly useful for production of dihydroxy manool or a dioxidized manool derivative with three additional degrees of desaturation, i.e. double bonds.

In one embodiment of the invention, the host organism may comprise at least the following heterologous nucleic acids:

a) a heterologous nucleic acid encoding TwTPS7 of SEQ ID NO:4 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith; and
b) a heterologous nucleic acid encoding CfTPS3 of SEQ ID NO:12 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith;
c) a heterologous nucleic acid encoding CYP76AH1 of SEQ ID NO:42 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith; and

Such a host organism is in particular useful for production of diterpenoids having a core of formulas VI or XII, for example for production of oxidised miltradiene, for example for production of ferruginol.

In one embodiment of the invention, the host organism may comprise at least the following heterologous nucleic acids:

a) a heterologous nucleic acid encoding TwTPS7 of SEQ ID NO:4 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith; and
b) a heterologous nucleic acid encoding CfTPS3 of SEQ ID NO:12 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least

85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith;

c) a heterologous nucleic acid encoding CYP76AH4 of SEQ ID NO:43 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith; and

Such a host organism is in particular useful for production of diterpenoids having a core of formulas VI or XII, for example for production of oxidised miltradiene, for example for production of ferruginol.

Sequence listing

SEQ ID NO:1	Amino acid sequence of syn-CPP from *Oryza sativa*
SEQ ID NO:2	Amino acid sequence of TPS7 from *Euphobia peplus*
SEQ ID NO:3	Amino acid sequence of AN2 from *Zea Maiz*
SEQ ID NO:4	Amino acid sequence of TPS7 from *Tripterygium Wilfordii*
SEQ ID NO:5	Amino acid sequence of TPS1 from *Coleus forskohlii*
SEQ ID NO:6	Amino acid sequence of LPPS from *Salvia scarea*
SEQ ID NO:7	Amino acid sequence of TPS21 from *Tripterygium Wilfordii*
SEQ ID NO:8	Amino acid sequence of TPS14/28 from *Tripterygium Wilfordii*
SEQ ID NO:9	Amino acid sequence of TPS8 of *Euphobia peplus*
SEQ ID NO:10	Amino acid sequence of TPS23 of *Euphobia peplus*
SEQ ID NO:11	Amino acid sequence of SCS of *Salvia scarea*
SEQ ID NO:12	Amino acid sequence of TPS3 of *Coleus forskohlii*
SEQ ID NO:13	Amino acid sequence of TPS4 of *Coleus forskohlii*
SEQ ID NO:14	Amino acid sequence of TPS2 of *Tripterygium Wilfordii*
SEQ ID NO:15	Amino acid sequence of TPS1 of *Euphobia peplus*
SEQ ID NO:16	Amino acid sequence of TPS14 of *Coleus forskohlii*
SEQ ID NO:17	Amino acid sequence of TPS2 of *Coleus forskohlii*
SEQ ID NO:18	Amino acid sequence of AtCPS of *Arabidopsis thaliana*
SEQ ID NO: 19	Amino acid sequence of AtEKS of *Arabidopsis thaliana*
SEQ ID NO: 20	Amino acid sequence of CYP76AH8 of *Coleus forskohlii*
SEQ ID NO: 21	Amino acid sequence of CYP76AH11 of *Coleus forskohlii*
SEQ ID NO: 22	Amino acid sequence of CYP71D381 of *Coleus forskohlii*

SEQ ID NO: 23	Amino acid sequence of CYP720B4 of *Picea sitchensis* (Sitka spruce)
SEQ ID NO:24	Amino acid sequence of CYP76AH9 of *Coleus forskohlii*
SEQ ID NO:25	cDNA sequence encoding CfDXS of *Coleus forskohlii*
SEQ ID NO:26	cDNA sequence encoding CfGGPPS of *Coleus forskohlii*
SEQ ID NO:27	cDNA encoding TPS1 of *Coleus forskohlii*
SEQ ID NO:28	DNA encoding TPS1 of *Coleus forskohlii* codon optimised for expression in *S. cerevisiae*
SEQ ID NO:29	cDNA encoding TPS3 of *Coleus forskohlii*
SEQ ID NO:30	DNA encoding TPS3 of *Coleus forskohlii* codon optimised for expression in *S. cerevisiae*
SEQ ID NO:31	cDNA encoding syn-CPP of *Oryza sativa*
SEQ ID NO:32	cDNA encoding SCS of *Salvia scarea*
SEQ ID NO:33	DNA encoding SCS of *Salvia scarea* codon optimised for expression in *S. cerevisiae*
SEQ ID NO:34	cDNA encoding CYP71D381 of *Coleus forskohlii*
SEQ ID NO:35	cDNA encoding CYP76AH8 of *Coleus forskohlii*
SEQ ID NO:36	DNA encoding CYP76AH8 of *Coleus forskohlii* codon optimised for expression in *S. cerevisiae*
SEQ ID NO:37	cDNA encoding CYP76AH11 of *Coleus forskohlii*
SEQ ID NO:38	DNA encoding CYP76AH11 of *Coleus forskohlii* codon optimised for expression in *S. cerevisiae*
SEQ ID NO:39	cDNA encoding CYP720B4 of *Picea sitchensis* (Sitka spruce)
SEQ ID NO:40	Amino acid sequence of CYP76AH15 of *Coleus forskohlii*
SEQ ID NO:41	Amino acid sequence of CYP76AH17 of *Coleus forskohlii*
SEQ ID NO:42	Amino acid sequence of CYP76AH1 from Salvia miltiorrhiza
SEQ ID NO:43	Amino acid sequence of CYP76AH4 from Rosmarinus officinalis
SEQ ID NO:44	Amino acid sequence of CPR of *Coleus forskohlii*
SEQ ID NO:45	cDNA encoding CPR of *Coleus forskohlii*
SEQ ID NO:46	Amino acid sequence of TPS5 of *Marrubium vulgare*

SEQ ID NO:1- Os syn-CPP
MPVFTASFQCVTLFGQPASAADAQPLLQGQRPFLHLHARRRRPCGPMLISKSPPYPASEETREW
EAEGQHEHTDELRETTTTMIDGIRTALRSIGEGEISISAYDTSLVALLKRLDGGDGPQFPSTID
WIVQNQLPDGSWGDASFFMMGDRIMSTLACVVALKSWNIHTDKCERGLLFIQENMWRLAHEEED
WMLVGFEIALPSLLDMAKDLDLDIPYDEPALKAIYAERERKLAKIPRDVLHAMPTTLLHSLEGM
VDLDWEKLLKLRCLDGSFHCSPASTATAFQQTGDQKCFEYLDGIVKKFNGGVPCIYPLDVYERL
WAVDRLTRLGISRHFTSEIEDCLDYIFRNWTPDGLAHTKNCPVKDIDDTAMGFRLLRLYGYQVD
PCVLKKFEKDGKFFCLHGESNPSSVTPMYNTYRASQLKFPGDDGVLGRAEVFCRSFLQDRRGSN

RMKDKWAIAKDIPGEVEYAMDYPWKASLPRIETRLYLDQYGGSGDVWIGKVLHRMTLFCNDLYL
KAAKADFSNFQKECRVELNGLRRWYLRSNLERFGGTDPQTTLMTSYFLASANIFEPNRAAERLG
WARVALLADAVSSHFRRIGGPKNLTSNLEELISLVPFDDAYSGSLREAWKQWLMAWTAKESSQE
SIEGDTAILLVRAIEIFGGRHVLTGQRPDLWEYSQLEQLTSSICRKLYRRVLAQENGKSTEKVE
EIDQQLDLEMQELTRRVLQGCSAINRLTRETFLHVVKSFCYVAYCSPETIDNHIDKVIFQDVI*

SEQ ID NO:2 – EpTPS7
MAAAANPSNSILNHHLLSSAAARSVSTSQLLFHSRPLVLSGAKDKRDSFVFRIKCSAVSN
PRIQEQTDVFQKNGLPVIKWHEFVETDIDHEQVSKVSVSNEIKKRVESIKAILESMEDGD
ITISAYDTAWVALVEDINGSGAPQFPASLQWIANNQLPDGSWGDAEIFTAHDRILNTLSC
VVALKSWNIHPDMCERGMKYFRENLCKLEDENIEHMPIGFEVAFPSLLELAKKLEIQVPE
DSPVLKDVYDSRNLKLKKIPKDIMHKVPTTLLHSLEGMPGLEWEKLLKLQSKDGSFLFSP
SSTAYALMQTKDQNCLEYLTKIVHKFNGGVPNVYPVDLFEHIWAVDRLQRLGISRYFQPQ
LKDSVDYVARYWEEDGICWARNSSVHDVDDTAMGFRVLRSFGHHVSADVFKHFKKGDTFF
CFAGQSTQAVTGMYNLLRASQLMFPGEKILEEAKQFSSAFLKVKQDANEVLDKWIITKDL
PGEVKYALDIPWYASLPRVESRFYIEQYGGSDDVWIGKTLYRMPIVNNDEYLKLAKLDYN
NCQAVHRSEWDNIQKWYEESDLAEFGVSRREILMAYYLAAASIFEPEKSRERIAWAKTSV
LLNTIQAYFHENNSTIHEKAAFVQLFKSGFAINARKLEGKTMEKLGRIIVGTLNDVSLDT
AMAYGKDISRDLRHAWDICLQKWEESGDMHQGEAQLIVNTINLTSDAWNFNDLSSHYHQF
FQLVNEICYKLRKYKKNKVNDKKKTTTPEIESHMQELVKLVLESSDDLDSNLKQIFLTVA
RSFYYPAVCDAGTINYHIARVLFERVY*

SEQ ID NO:3 – ZmAN2
MVLSSSCTTVPHLSSLAVVQLGPWSSRIKKKTDTVAVPAAAGRWRRALARAQHTSESAAV
AKGSSLTPIVRTDAESRRTRWPTDDDDAEPLVDEIRAMLTSMSDGDISVSAYDTAWVGLV
PRLDGGEGPQFPAAVRWIRNNQLPDGSWGDAALFSAYDRLINTLACVVTLTRWSLEPEMR
GRGLSFLGRNMWKLATEDEESMPIGFELAFPSLIELAKSLGVHDFPYDHQALQGIYSSRE
IKMKRIPKEVMHTVPTSILHSLEGMPGLDWAKLLKLQSSDGSFLFSPAATAYALMNTGDD
RCFSYIDRTVKKFNGGVPNVYPVDLFEHIWAVDRLERLGISRYFQKEIEQCMDYVNRHWT
EDGICWARNSDVKEVDDTAMAFRLLRLHGYSVSPDVFKNFEKDGEFFAFVGQSNQAVTGM
YNLNRASQISFPGEDVLHRAGAFSYEFLRRKEAEGALRDKWIISKDLPGEVVYTLDFPWY
GNLPRVEARDYLEQYGGGDDVWIGKTLYRMPLVNNDVYLELARMDFNHCQALHQLEWQGL
KRWYTENRLMDFGVAQEDALRAYFLAAASVYEPCRAAERLAWARAAILANAVSTHLRNSP
SFRERLEHSLRCRPSEETDGSWFNSSSGSDAVLVKAVLRLTDSLAREAQPIHGGDPEDII
HKLLRSAWAEWVREKADAADSVCNGSSAVEQEGSRMVHDKQTCLLLARMIEISAGRAAGE
AASEDGDRRIIQLTGSICDSLKQKMLVSQDPEKNEEMMSHVDDELKLRIREFVQYLLRLG
EKKTGSSETRQTFLSIVKSCYYAAHCPPHVVDRHISRVIFEPVSAAK*

SEQ ID NO:4 – TwTPS7
MHSLLMKKVIMYSSQTTHVFPSPLHCTIPKSSSFFLDAPVVRLHCLSGHGAKKKRLHFDI
QQGRNAISKTHTPEDLYAKQEYSVPEIVKDDDKEEEVVKIKEHVDIIKSMLSSMEDGEIS
ISAYDTAWVALIQDIHNNGAPQFPSSLLWIAENQLPDGSWGDSRVFLAFDRIINTLACVV
ALKSWNVHPDKCERGISFLKENISMLEKDDSEHMLVGFEFGFPVLLDMARRLGIDVPDDS
PFLQEIYVQRDLKLKRIPKDILHNAPTTLLHSLEAIPDLDWTKLLKLQCQDGSLLFSPSS
TAMAFINTKDENCLRYLNYVVQRFNGGAPTVYPYDLFEHNWAVDRLQRLGISRFFQPEIR
ECMSYVYRYWTKDGIFCTRNSRVHDVDDTAMGFRLLRLHGYEVHPDAFRQFKKGCEFICY
EGQSHPTVTVMYNLYRASQLMFPEEKILDEAKQFTEKFLGEKRSANKLLDKWIITKDLPG
EVGFALDVPWYASLPRVEARFFIQHYGGEDDVWLDKALYRMPYVNNNVYLELAKLDYNYC
QALHRTEWGHIQKWYEECKPRDFGISRECLLRAYFMAAASIFEPERSMERLAWAKTAILL
EIIVSYFNEVGNSTEQRIAFTTEFSIRASPMGGYINGRKLDKIGTTQELIQMLLATIDQF
SQDAFAAYGHDITRHLHNSWKMWLLKWQEEGDRWLGEAELLIQTINLMADHKIAEKLFMG
HTNYEQLFSLTNKVCYSLGHHELQNNKELEHDMQRLVQLVLTNSSDGIDSDIKKTFLAVA
KRFYYTAFVDPETVNVHIAKVLFERVD*

SEQ ID NO:5 – CfTPS1
MGSLSTMNLNHSPMSYSGILPSSSAKAKLLLPGCFSISAWMNNGKNLNCQLTHKKISKVA
EIRVATVNAPPVHDQDDSTENQCHDAVNNIEDPIEYIRTLLRTTGDGRISVSPYDTAWVA
LIKDLQGRDAPEFPSSLEWIIQNQLADGSWGDAKFFCVYDRLVNTIACVVALRSWDVHAE
KVERGVRYINENVEKLRDGNEEHMTCGFEVVFPALLQRAKSLGIQDLPYDAPVIQEIYHS
REQKSKRIPLEMMHKVPTSLLFSLEGLENLEWDKLLKLQSADGSFLTSPSSTAFAFMQTR
DPKCYQFIKNTIQTFNGGAPHTYPVDVFGRLWAIDRLQRLGISRFFESEIADCIAHIHRF
WTEKGVFSGRESEFCDIDDTSMGVRLMRMHGYDVDPNVLKNFKKDDKFSCYGGQMIESPS
PIYNLYRASQLRFPGEQILEDANKFAYDFLQEKLAHNQILDKWVISKHLPDEIKLGLEMP
WYATLPRVEARYYIQYYAGSGDVWIGKTLYRMPEISNDTYHELAKTDFKRCQAQHQFEWI
YMQEWYESCNMEEFGISRKELLVAYFLATASIFELERANERIAWAKSQIISTIIASFFNN
QNTSPEDKLAFLTDFKNGNSTNMALVTLTQFLEGFDRYTSHQLKNAWSVWLRKLQQGEGN
GGADAELLVNTLNICAGHIAFREEILAHNDYKTLSNLTSKICRQLSQIQNEKELETEGQK
TSIKNKELEEDMQRLVKLVLEKSRVGINRDMKKTFLAVVKTYYYKAYHSAQAIDNHMFKV
LFEPVA*

SEQ ID NO:6 – SsLPPS
MTSVNLSRAPAAITRRRLQLQPEFHAECSWLKSSSKHAPLTLSCQIRPKQLSQIAELRVT
SLDASQASEKDISLVQTPHKVEVNEKIEESIEYVQNLLMTSGDGRISVSPYDTAVIALIK
DLKGRDAPQFPSCLEWIAHHQLADGSWGDEFFCIYDRILNTLACVVALKSWNLHSDIIEK
GVTYIKENVHKLKGANVEHRTAGFELVVPTFMQMATDLGIQDLPYDHPLIKEIADTKQQR
LKEIPKDLVYQMPTNLLYSLEGLGDLEWERLLKLQSGNGSFLTSPSSTAAVLMHTKDEKC
LKYIENALKNCDGGAPHTYPVDIFSRLWAIDRLQRLGISRFFQHEIKYFLDHIESVWEET
GVFSGRYTKFSDIDDTSMGVRLLKMHGYDVDPNVLKHFKQQDGKFSCYIGQSVESASPMY
NLYRAAQLRFPGEEVLEEATKFAFNFLQEMLVKDRLQERWVISDHLFDEIKLGLKMPWYA
TLPRVEAAYYLDHYAGSGDVWIGKSFYRMPEISNDTYKELAILDFNRCQTQHQLEWIHMQ
EWYDRCSLSEFGISKRELLRSYFLAAATIFEPERTQERLLWAKTRILSKMITSFVNISGT
TLSLDYNFNGLDEIISSANEDQGLAGTLLATFHQLLDGFDIYTLHQLKHVWSQWFMKVQQ
GEGSGGEDAVLLANTLNICAGLNEDVLSNNEYTALSTLTNKICNRLAQIQDNKILQVVDG
SIKDKELEQDMQALVKLVLQENGGAVDRNIRHTFLSVSKTFYYDAYHDDETTDLHIFKVL
FRPVV*

SEQ ID NO:7 – TwTPS21
MFMSSSSSSHARRPQLSSFSYLHPPLPFPGLSFFNTRDKRVNFDSTRIICIAKSKPARTT
PEYSDVLQTGLPLIVEDDIQEQEEPLEVSLENQIRQGVDIVKSMLGSMEDGETSISAYDT
AWVALVENIHHPGSPQFPSSLQWIANNQLPDGSWGDPDVFLAHDRLINTLACVIALKKWN
IHPHKCKRGLSFVKENISKLEKENEEHMLIGFEIAFPSLLEMAKKLGIEIPDDSPALQDI
YTKRDLKLTRIPKDKMHNVPTTLLHSLEGLPDLDWEKLVKLQFQNGSFLFSPSSTAFAFM
HTKDGNCLSYLNDLVHKFNGGVPTAYPVDLFEHIWSVDRLQRLGISRFFHPEIKECLGYV
HRYWTKDGICWARNSRVQDIDDTAMGFRLLRLHGYEVSPDVFKQFRKGDEFVCFMGQSNQ
AITGIYNLYRASQMMFPEETILEEAKKFSVNFLREKRAASELLDKWIITKDLPNEVGFAL
DVPWYACLPRVETRLYIEQYGGQDDVWIGKTLYRMPYVNNNVYLELAKLDYNNCQSLHRI
EWDNIQKWYEGYNLGGFGVNKRSLLRTYFLATSNIFEPERSVERLTWAKTAILVQAIASY
FENSREERIEFANEFQKFPNTRGYINGRRLDVKQATKGLIEMVFATLNQFSLDALVVHGE
DITHHLYQSWEKWVLTWQEGGDRREGEAELLVQTINLMAGHTHSQEEELYERLFKLTNTV
CHQLGHYHHLNKDKQPQQVEDNGGYNNSNPESISKLQIESDMRELVQLVLNSSDGMDSNI
KQTFLAVTKSFYYTAFTHPGTVNYHIAKVLFERVV*

SEQ ID NO:8 – TwTPS14/28
MFMSSSSSSHARRPQLSSFSYLHPPLPFPGLSFFNTRDKRVNFDSTRIICIAKSKPARTT
PEYSDVLQTGLPLIVEDDIQEQEEPLEVSLENQIRQGVDIVKSMLGSMEDGETSISAYDT
AWVALVENIHHPGSPQFPSSLQWIANNQLPDGSWGDPDVFLAHDRLINTLACVIALKKWN

IHPHKCKRGLSFVKENISKLEKENEEHMLIGFEIAFPSLLEMAKKLGIEIPDDSPALQDI
YTKRDLKLTRIPKDIMHNVPTTLLYSLEGLPSLDWEKLVKLQCTDGSFLFSPSSTACALM
HTKDGNCFSYINNLVHKFNGGVPTVYPVDLFEHIWCVDRLQRLGISRFFHPEIKECLGYV
HRYWTKDGICWARNSRVQDIDDTAMGFRLLRLHGYEVSPDVFKQFRKGDEFVCFMGQSNQ
AITGIYNLYRASQMMFPEETILEEAKKFSVNFLREKRAASELLDKWIITKDLPNEVGFAL
DVPWYACLPRVETRLYIEQYGGQDDVWIGKTLYRMPYVNNNVYLELAKLDYNNCQSLHRI
EWDNIQKWYEGYNLGGFGVNKRSLLRTYFLATSNIFEPERSVERLTWAKTAILVQAIASY
FENSREERIEFANEFQKFPNTRGYINGRRLDVKQATKGLIEMVFATLNQFSLDALVVHGE
DITHHLYQSWEKWVLTWQEGGDRREGEAELLVQTINLMAGHTHSQEEELYERLFKLTNTV
CHQLGHYHHLNKDKQPQQVEDNGGYNNSNPESISKLQIESDMRELVQLVLNSSDGMDSNI
KQTFLAVTKSFYYTAFTHPGTVNYHIAKVLFERVV*

SEQ ID NO:9 – EpTPS8
MQVSLSLTTGSEPCITRIHAPSDAPLKQRNNEREKGTLELNGKVSLKKMGEMLRTIENVP
IVGSTSSYDTAWVGMVPCSSNSSKPLFPESLKWIMENQNPEGNWAVDHAHHPLLLKDSLS
STLACVLALHKWNLAPQLVHSGLDFIGSNLWAAMDFRQRSPLGFDVIFPGMIHQAIDLGI
NLPFNNSSIENMLTNPLLDIQSFEAGKTSHIAYFAEGLGSRLKDWEQLLQYQTSNGSLFN
SPSTTAAAAIHLRDEKCLNYLHSLTKQFDNGAVPTLYPLDARTRISIIDSLEKFGIHSHF
IQEMTILLDQIYSFWKEGNEEIFKDPGCCATAFRLLRKHGYDVSSDSLAEFEKKEIFYHS
SAASAHEIDTKSILELFRASQMKILQNEPILDRIYDWTSIFLRDQLVKGLIENKSLYEEV
NFALGHPFANLDRLEARSYIDNYDPYDVPLLKTSYRSSNIDNKDLWTIAFQDFNKCQALH
RVELDYLEKWVKEYKLDTLKWARQKTEYALFTIGAILSEPEYADARISWSQNTVFVTIVD
DFFDYGGSLDECRNLINLMHKWDDHLTVGFLSEKVEIVFYSMYGTLNDLAAKAEVRQGRC
VRSHLVNLWIWVMENMLKEREWADYNLVPTFYEYVAAGHITIGLGPVLLIALYFMGYPLS
EDVVQSQEYKGVYLNVSIIARLLNDRVTVKRESAQGKLNGVSLFVEHGRGAVDEETSMKE
VERLVESHKRELLRLIVQKTEGSVVPQSCKDLAWRVSKVLHLLYMDDDGFTCPVKMLNAT
NAIVNEPLLLTS*

SEQ ID NO:10 – EpTPS23
MLLASSTSSRFFTKEWEPSNKTFSGSVRAQLSQRVKNIVVTPDQVKESESSGTSLRLKEM
LKKVEMPISSYDTAWVAMVPSMEHSRNKPLFPNSLKWVMENQQPDGSWCFDDSNHPWLIK
DSLSSTLASVLALKKWNVGQQLIDKGLEYIGSNMWAATDMHQYSPIGFNIIFPSMVEHAN
KLGLSLSLDHSLFQSMLRNRDMETKSLNGRNMAYVAEGLNGSNNWKEVMKYQRRNGSILN
SPATTAAALIHLNDVKCFEYLDSLLTKFQHAVPTLYPFDIYARLCILDELEKLGVDRFVE
IEKMLLLDYIYRCWLEGSEEILEDPTCCAMAFRFLRMNGYVVSPDVLQGFEEEEKLFHVK
DTKSVLELLKASQLKVSEKEGILDRIYSWATSYLKHQLFNASISDKSLQNEVDYVVKHPH
AILRRIENRNYIENYNTKNVSLRKTSFRFVNVDKRSDLLAHSRQDFNKCQIQFKKELAYL
SRWEKKYGLDKLKYARQRLEVVYFSIASNLFEPEFSDARLAWTQYAILTTVVDDFFEYAA
SMDELVNLTNLIERWDEHGSEEFKSKEVEILFYAIYDLVNEDAEKAKKYQGRCIKSHLVH
IWIDILKAMLKESEYVRYNIVPTLDEYISNGCTSISFGAILLIPLYFLGKMSEEVVTSKE
YQKLYMHISMLGRLLNDRVTSQKDMAQGKLNSVSLRVLHSNGTLTEEEAKEEVDKIIEKH
RRELLRMVVQTEGSVVPKACKKLFWMTSKELHLFYMTEDCFTCPTKLLSAVNSTLKDPLL
MP*

SEQ ID NO:11 SsSCS
MSLAFNVGVTPFSGQRVGSRKEKFPVQGFPVTTPNRSRLIVNCSLTTIDFMAKMKENFKREDDK
FPTTTTLRSEDIPSNLCIIDTLQRLGVDQFFQYEINTILDNTFRLWQEKHKVIYGNVTTHAMAF
RLLRVKGYEVSSEELAPYGNQEAVSQQTNDLPMIIELYRAANERIYEEERSLEKILAWTTIFLN
KQVQDNSIPDKKLHKLVEFYLRNYKGITIRLGARRNLELYDMTYYQALKSTNRFSNLCNEDFLV
FAKQDFDIHEAQNQKGLQQLQRWYADCRLDTLNFGRDVVIIANYLASLIIGDHAFDYVRLAFAK
TSVLVTIMDDFFDCHGSSQECDKIIELVKEWKENPDAEYGSEELEILFMALYNTVNELAERARV
EQGRSVKEFLVKLWVEILSAFKIELDTWSNGTQQSFDEYISSSWLSNGSRLTGLLTMQFVGVKL

SDEMLMSEECTDLARHVCMVGRLLNDVCSSEREREENIAGKSYSILLATEKDGRKVSEDEAIAE
INEMVEYHWRKVLQIVYKKESILPRRCKDVFLEMAKGTFYAYGINDELTSPQQSKEDMKSFVF*

SEQ ID NO:12 CfTPS3
MSSLAGNLRVIPFSGNRVQTRTGILPVHQTPMITSKSSAAVKCSLTTPTDLMGKIKEVFN
REVDTSPAAMTTHSTDIPSNLCIIDTLQRLGIDQYFQSEIDAVLHDTYRLWQLKKKDIFS
DITTHAMAFRLLRVKGYEVASDELAPYADQERINLQTIDVPTVVELYRAAQERLTEEDST
LEKLYVWTSAFLKQQLLTDAIPDKKLHKQVEYYLKNYHGILDRMGVRRNLDLYDISHYKS
LKAAHRFYNLSNEDILAFARQDFNISQAQHQKELQQLQRWYADCRLDTLKFGRDVVRIGN
FLTSAMIGDPELSDLRLAFAKHIVLVTRIDDFFDHGGPKEESYEILELVKEWKEKPAGEY
VSEEVEILFTAVYNTVNELAEMAHIEQGRSVKDLLVKLWVEILSVFRIELDTWTNDTALT
LEEYLSQSWVSIGCRICILISMQFQGVKLSDEMLQSEECTDLCRYVSMVDRLLNDVQTFE
KERKENTGNSVSLLQAAHKDERVINEEEACIKVKELAEYNRRKLMQIVYKTGTIFPRKCK
DLFLKACRIGCYLYSSGDEFTSPQQMMEDMKSLVYEPLPISPPEANNASGEKMSCVSN*

SEQ ID NO:13 – CfTPS4
MSITINLRVIAFPGHGVQSRQGIFAVMEFPRNKNTFKSSFAVKCSLSTPTDLMGKIKEKL
SEKVDNSVAAMATDSADMPTNLCIVDSLQRLGVEKYFQSEIDTVLDDAYRLWQLKQK
DIFSDITTHAMAFRLLRVKGYDVSSEELAPYADQEGMNLQTIDLAAVIELYRAAQERVA
EEDSTLEKLYVWTSTFLKQQLLAGAIPDQKLHKQVEYYLKNYHGILDRMGVRKGLDLYD
AGYYKALKAADRLVDLCNEDLLAFARQDFNINQAQHRKELEQLQRWYADCRLDKLEF
GRDVVRVSNFLTSAILGDPELSEVRLVFAKHIVLVTRIDDFFDHGGPREESHKILELIKEW
KEKPAGEYVSKEVEILYTAVYNTVNELAERANVEQGRNVEPFLRTLWVQILSIFKIELDT
WSDDTALTLDDYLNNSWVSIGCRICILMSMQFIGMKLPEEMLLSEECVDLCRHVSMV
DRLLNDVQTFEKERKENTGNAVSLLLAAHKGERAFSEEEAIAKAKYLADCNRRSLMQI
VYKTGTIFPRKCKDMFLKVCRIGCYLYASGDEFTSPQQMMEDMKSLVYEPLQIHPPAAA*

SEQ ID NO:14 – TwTPS2
MFDKTQLSVSAYDTAWVAMVSSPNSRQAPWFPECVNWLLDNQLSDGSWGLPPHHPSLVKD
ALSSTLACLLALKRWGLGEQQMTKGLQFIESNFTSINDEEQHTPIGFNIIFPGMIETAID
MNLNLPLRSEDINVMLHNRDLELRRNKLEGREAYLAYVSEGMGKLQDWEMVMKYQRKNGS
LFNSPSTTAAALSHLGNAGCFHYINSLVAKFGNAVPTVYPSDKYALLCMIESLERLGIDR
HFSKEIRDVLEETYRCWLQGDEEIFSDADTCAMAFRILRVHGYEVSSDPLTQCAEHHFSR
SFGGHLKDFSTALELFKASQFVIFPEESGLEKQMSWTNQFLKQEFSNGTTRADRFSKYFS
IEVHDTLKFPFHANVERLAHRRNIEHHHVDNTRILKTSYCFSNISNADFLQLAVEDFNRC
QSIHREELKHLERWVVETKLDRLKFARQKMAYCYFSAAGTCFSPELSDARISWAKNSVLT
TVADDFFDIVGSEEELANLVHLLENWDANGSPHYCSEPVEIIFSALRSTICEIGDKALAW
QGRSVTHHVIEMWLDLLKSALREAEWARNKVVPTFDEYVENGYVSMALGPIVLPAVYLIG
PKVSEEVVRSPEFHNLFKLMSICGRLINDTRTFKRESEAGKLNSVLLHMIHSGSGTTEEE
AVEKIRGMIADGRRELLRLVLQEKDSVVPRACKDLFWKMVQVLHLFYMDGDGFSSPDMML
NAVNALIREPISL*

SEQ ID NO: 15 – EpTPS1
MSATPNSFFTSPISAKLGHPKSQSVAESNTRIQQLDGTREKIKKMFDKVELSVSPYDTAW
VAMVPSPNSLEAPYFPECSKWIVDNQLNDGSWGVYHRDPLLVKDSISSTLACVLALKRWG
IGEKQVNKGLEFIELNSASLNDLKQYKPVGFDITFPRMLEHAKDFGLNLPLDPKYVEAVI
FSRDLDLKSGCDSTTEGRKAYLAYISEGIGNLQDWNMVMKYQRRNGSIFDSPSATAAASI
HLHDASCLRYLRCALKKFGNAVPTIYPFNIYVRLSMVDAIESLGIARHFQEEIKTVLDET
YRYWLQGNEEIFQDCTTCAMAFRILRANGYNVSSEKLNQFTEDHFSNSLGGYLEDMRPVL
ELYKASQLIFPDELFLEKQFSWTSQCLKQKISSGLRHTDGINKHITEEVNDVLKFASYAD
LERLTNWRRIAVYRANETKMLKTSYRCSNIANEHFLELAVEDFNVCQSMHREELKHLGRW

VVEKRLDKLKFARQKLGYCYFSSAASLFAPEMSDARISWAKNAVLTTVVDDFFDVGGSEE
ELINLVQLIERWDVDGSSHFCSEHVEIVFSALHSTICEIGEKAFAYQGRRMTSHVIKIWL
DLLKSMLTETLWSKSKATPTLNEYMTNGNTSFALGPIVLPALFFVGPKLTDEDLKSHELH
DLFKTMSTCGRLLNDWRSYERESEEGKLNAVSLHMIYGNGSVAATEEEATQKIKGLIESE
RRELMRLVLQEKDSKIPRPCKDLFWKMLKVLHMFYLKDDGFTSNQMMKTANSLINQPISL
HER*

SEQ ID NO:16 – CfTPS14
MSLPLSTCVLFVPKGSQFWSSRFSYASASLEVGFQRATSAQIAPLSKSFEETKGRIAKLF
HKDELSISTYDTAWVAMVPSPTSSEEPCFPACLNWLLENQCLDGSWARPHHHPMLKKDVL
SSTLACILALKKWGVGEEQINRGLHFIELNFASATEKCQITPMGFDIVFPAMLDRARALS
LNIRLEPTTLNDLMNKRDLELNRCYQSSSTEREVYRAYIAEGMGKLQNWESVMKYQRKNG
TLFNCPSTTAAAFTALRNSDCLNYLHLALNKFGDAVPAVFPLDIYSQLCIVDNLERVGIS
RHFLTEIQSVLDGTYRSWLQGDEQIFMDASTCALAFRTLRMNGYNVSSDPITKLIQEGSF
SRNTMDINTTLELYRASELILYPDERDLEEHNLRLKTILDQELSGGGFILSRQLGRNINA
EVKQALESPFYAIMDRMAKRRSIEHYHIDNTRILKTSYCSPNFGNEDFLSLSVEDFNRCQ
VIHREELRELERWVIENRLDELKFARSKSAYCYFSAAATIFSPELSDARMSWAKNGVLTT
VVDDFFDVGGSVEELKNLIQLVELWDVDVSRECISPSVQIIFSALKHTIREIGDKGFKLQ
GRSITDHIIAIWLDLLYSMMKESEWGREKAVPTIDEYISNAYVSFALGPIVLPALYLVGP
KLSEEMVNHADYHNLFKSMSTCGRLLNDIRGYERELKDGKLNTLSLYMVNNEGEISWEAA
ILEVKSWIERERRELLRSVLEEEKSVVPKACKELFWHMCTVVHLFYSKDDGFTSQDLLSA
VNAIIYQPLVLE*

SEQ ID NO:17 – CfTPS2
MKMLMIKSQFRVHSIVSAWANNSNKRQSLGHQIRRKQRSQVTECRVASLDALNGIQKVGP
ATIGTPEEENKKIEDSIEYVKELLKTMGDGRISVSPYDTAIVALIKDLEGGDGPEFPSCL
EWIAQNQLADGSWGDHFFCIYDRVVNTAACVVALKSWNVHADKIEKGAVYLKENVHKLKD
GKIEHMPAGFEFVVPATLERAKALGIKGLPYDDPFIREIYSAKQTRLTKIPKGMIYESPT
SLLYSLDGLEGLEWDKILKLQSADGSFITSVSSTAFVFMHTNDLKCHAFIKNALTNCNGG
VPHTYPVDIFARLWAVDRLQRLGISRFFEPEIKYLMDHINNVWREKGVFSSRHSQFADID
DTSMGIRLLKMHGYNVNPNALEHFKQKDGKFTCYADQHIESPSPMYNLYRAAQLRFPGEE
ILQQALQFAYNFLHENLASNHFQEKWVISDHLIDEVRIGLKMPWYATLPRVEASYYLQHY
GGSSDVWIGKTLYRMPEISNDTYKILAQLDFNKCQAQHQLEWMSMKEWYQSNNVKEFGIS
KKELLLAYFLAAATMFEPERTQERIMWAKTQVVSRMITSFLNKENTMSFDLKIALLTQPQ
HQINGSEMKNGLAQTLPAAFRQLLKEFDKYTRHQLRNTWNKWLMKLKQGDDNGGADAELL
ANTLNICAGHNEDILSHYEYTALSSLTNKICQRLSQIQDKKMLEIEEGSIKDKEMELEIQ
TLVKLVLQETSGGIDRNIKQTFLSVFKTFYYRAYHDAKTIDAHIFQVLFEPVV*

SEQ ID NO:18 – AtCPS

MSLQYHVLNSIPSTTFLSSTKTTISSSFLTISGSPLNVARDKSRSGSIHCSKLRTQEYINSQEV
QHDLPLIHEWQQLQGEDAPQISVGSNSNAFKEAVKSVKTILRNLTDGEITISAYDTAWVALIDA
GDKTPAFPSAVKWIAENQLSDGSWGDAYLFSYHDRLINTLACVVALRSWNLFPHQCNKGITFFR
ENIGKLEDENDEHMPIGFEVAFPSLLEIARGINIDVPYDSPVLKDIYAKKELKLTRIPKEIMHK
IPTTLLHSLEGMRDLDWEKLLKLQSQDGSFLFSPSSTAFAFMQTRDSNCLEYLRNAVKRFNGGV
PNVFPVDLFEHIWIVDRLQRLGISRYFEEEIKECLDYVHRYWTDNGICWARCSHVQDIDDTAMA
FRLLRQHGYQVSADVFKNFEKEGEFFCFVGQSNQAVTGMFNLYRASQLAFPREEILKNAKEFSY
NYLLEKREREELIDKWIIMKDLPGEIGFALEIPWYASLPRVETRFYIDQYGGENDVWIGKTLYR
MPYVNNNGYLELAKQDYNNCQAQHQLEWDIFQKWYEENRLSEWGVRRSELLECYYLAAATIFES
ERSHERMVWAKSSVLVKAISSSFGESSDSRRSFSDQFHEYIANARRSDHHFNDRNMRLDRPGSV
QASRLAGVLIGTLNQMSFDLFMSHGRDVNNLLYLSWGDWMEKWKLYGDEGEGELMVKMIILMKN
NDLTNFFTHTHFVRLAEIINRICLPRQYLKARRNDEKEKTIKSMEKEMGKMVELALSESDTFRD
VSITFLDVAKAFYYFALCGDHLQTHISKVLFQKV

SEQ ID NO;19 – AtEKS

MSINLRSSGCSSPISATLERGLDSEVQTRANNVSFEQTKEKIRKMLEKVELSVSAYDTSWVAMV
PSPSSQNAPLFPQCVKWLLDNQHEDGSWGLDNHDHQSLKKDVLSSTLASILALKKWGIGERQIN
KGLQFIELNSALVTDETIQKPTGFDIIFPGMIKYARDLNLTIPLGSEVVDDMIRKRDLDLKCDS
EKFSKGREAYLAYVLEGTRNLKDWDLIVKYQRKNGSLFDSPATTAAAFTQFGNDGCLRYLCSLL
QKFEAAVPSVYPFDQYARLSIIVTLESLGIDRDFKTEIKSILDETYRYWLRGDEEICLDLATCA
LAFRLLLAHGYDVSYDPLKPFAEESGFSDTLEGYVKNTFSVLELFKAAQSYPHESALKKQCCWT
KQYLEMELSSWVKTSVRDKYLKKEVEDALAFPSYASLERSDHRRKILNGSAVENTRVTKTSYRL
HNICTSDILKLAVDDFNFCQSIHREEMERLDRWIVENRLQELKFARQKLAYCYFSGAATLFSPE
LSDARISWAKGGVLTTVVDDFFDVGGSKEELENLIHLVEKWDLNGVPEYSSEHVEIIFSVLRDT
ILETGDKAFTYQGRNVTHHIVKIWLDLLKSMLREAEWSSDKSTPSLEDYMENAYISFALGPIVL
PATYLIGPPLPEKTVDSHQYNQLYKLVSTMGRLLNDIQGFKRESAEGKLNAVSLHMKHERDNRS
KEVIIESMKGLAERKREELHKLVLEEKGSVVPRECKEAFLKMSKVLNLFYRKDDGFTSNDLMSL
VKSVIYEPVSLQKESLT

SEQ ID NO: 20 CfCYP76AH8
METITLLLALFFIALTYFISSRRRRNLPPGPFPLPIIGNMLQLGSKPHQSFAQLSKKYGPLMSI
HLGSLYTVIVSSPEMAKEILQKHGQVFSGRTIAQAVHACDHDKISMGFLPVANTWRDMRKICKE
QMFSHHSLEASEELRHQKLQQLLDYAQKCCEAGRAVDIREASFITTLNLMSATMFSTQATEFDS
EATKEFKEIIEGVATIVGVANFADYFPILKPFDLQGIKRRADGYFGRLLKLIEGYLNERLESRR
LNPDAPRKKDFLETLVDIIEANEYKLTTEHLTHLMLDLFVGGSETNTTSLEWIMSELVINPDKM
AKVKEELKSVVGDEKLVNESDMPRLPYLQAVIKEVLRIHPPGPLLLPRKAESDQVVNGYLIPKG
TQILFNAWAMGRDPTIWKDPESFEPERFLNQSIDFKGQDFELIPFGSGRRICPGMPLANRILHM
TTATLVHNFDWKLEEGTADADHKGELFGLAVRRATPLRIIPLKP

SEQ ID NO:21 CfCYP76AH11
MELVQVIAVVAVVVVLWSQLKRKGRKLPPGPSPLPIVGNIFQLSGKNINESFAKLSKIYGPVMS
LRLGSLLTVIISSPEMAKEVLTSKDFANRPLTEAAHAHGHSKFSVGFVPVSDPKWKQMRRVCQE
EMFASRILENSQQRRHQKLQELIDHVQESRDAGRAVTIRDPVFATTLNIMSLTLFSADATEFSS
SATAELRDIMAGVVSVLGAANLADFFPILKYFDPQGMRRKADLHYGRLIDHIKSRMDKRSELKK
ANPNHPKHDDFLEKIIDITIQRNYDLTINEITHLLVDLYLAGSESTVMTIEWTMAELMLRPESL
AKLKAELRSVMGERKMIQESDDISRLPYLNGAIKEALRLHPPGPLLFARKSEIDVELSGYFIPK
GTQILVNEWGMGRDPSVWPNPECFQPERFLDKNIDYKGQDPQLIPFGAGRRICPGIPIAHRVVH
SVVAALVHNFDWEFAPGGSQCNNEFFTGAALVREVPLKLIPLNPPSI

SEQ ID NO:22 CfCYP71D381
MEFDFPSALIFPAVSLLLLLWLTKTRKPKSDLDRIPGPRRLPLIGNLHHLISLTPPPRLFREMA
AKYGPLMRLQLGGVPFLIVSSVDVAKHVVKTNDVPFANRPPMHAARAITYNYTDIGFAPYGEYW
RNLRKICTLELLSARRVRSFRHIREEENAGVAKWIASKEGSPANLSERVYLSSFDITSRASIGK
ATEEKQTLTSSIKDAMKLGGFNVADLYPSSKLLLLITGLNFRIQRVFRKTDRILDDLLSQHRST
SATTERPEDLVDVLLKYQKEETEVHLNNDKIKAVIMDMFLAGGETSATAVDWAMAEMIRNPTTL
KKAQEEVRRVFDGKGYVDEEEFHELKYLKLVIKEMLRMHPPLPFLVPRMNSERCEINGYEIPAN
TRLLINAWAIGRPKYWNDAEKFIPERFENSSIDFKGNNLEYIPFGAGRRMCPGMTFGLASVEFT
LAMLLYHFDWKMPQGIKLDMTESFGASLKRKHDLLMIPTLKRPLRLAP

SEQ ID NO:23 CYP720B4
MAPMADQISLLLVVFTVAVALLHLIHRWWNIQRGPKMSNKEVHLPPGSTGWPLIGETFSYYRSM
TSNHPRKFIDDREKRYDSDIFISHLFGGRTVVSADPQFNKFVLQNEGRFFQAQYPKALKALIGN
YGLLSVHGDLQRKLHGIAVNLLRFERLKVDFMEEIQNLVHSTLDRWADMKEISLQNECHQMVLN

LMAKQLLDLSPSKETSDICELFVDYTNAVIAIPIKIPGSTYAKGLKARELLIKKISEMIKERRN
HPEVVHNDLLTKLVEEGLISDEIICDFILFLLFAGHETSSRAMTFAIKFLTYCPKALKQMKEEH
DAILKSKGGHKKLNWDDYKSMAFTQCVINETLRLGNFGPGVFREAKEDTKVKDCLIPKGWVVFA
FLTATHLHEKFHNEALTFNPWRWQLDKDVPDDSLFSPFGGGARLCPGSHLAKLELSLFLHIFIT
RFSWEARADDRTSYFPLPYLTKGFPISLHGRVENE

SEQ ID NO:24 CfCYP76AH9
MDFFTLLAALFLITLTFFLFFKSESKRRGGANLPPGPYPLPIVGNIFQLGKKPHQSLAQLAKIH
GPLMSLHFGSVYTVIVTSPEMAKEIFVKNDQAFLNRTVVEAVHAHDHDKISMAFMDVGTEWRTL
RRICKEQMFSTQSLETSQGLRQEKLQQLHDFVQRCCDSGRVVDIREASFVTTLNLMSATLFSIQ
ATEFDSNATEEFREIMEGVASIVGDPNFADYFPILKRFDPQGVKRKAELYFGKMLVLVEDLLQK
RQEERRRSPSYAKKDDLLERLVDVLNEKNEYKLTTKHITHLLLDLFVGGSETTTTSVEWIMSEL
LINPEKLAKLKEELKTVVGEKKQVQESDIPQLPYFEAVLKEVFRLHPPGPLLLPRKAECDVQVG
SYTIPKETQILVNAWAIGRDPAIWPNPEAFEPERFLSQKMDYKGQDFELIPFGSGRRICPGLSF
ANRMLPMTVATLIHNFDWKLEVEANAEDVHKGEMFGIAVRRAVPLRAYPIQP

SEQ ID NO:25 CfDXS2
ATGGCGTCTTGTGGAGCTATCGGGAGTAGTTTCTTGCCACTGCTCCATTCCGACGAGTCAAGCT
TGTTATCTCGGCCCACTGCTGCTCTTCACATCAAGAAGCAGAAGTTTTCTGTGGGAGCTGCTCT
GTACCAGGATAACACGAACGATGTCGTTCCGAGTGGAGAGGGTCTGACGAGGCAGAAACCAAGA
ACTCTGAGTTTCACGGGAGAGAAGCCTTCAACTCCAATTTTGGATACCATCAACTATCCAATCC
ACATGAAGAATCTGTCCGTGGAGGAACTGGAGAGATTGGCCGATGAACTGAGGGAGGAGATAGT
TTACACGGTGTCGAAAACGGGAGGGCATTTGAGCTCAAGCTTGGGTGTATCAGAGCTCACCGTT
GCACTGCATCATGTATTCAACACACCCGATGACAAAATCATCTGGGATGTTGGACATCAGGCGT
ATCCACACAAAATCTTGACAGGGAGGAGGTCCAGAATGCACACCATCCGAAAAACTTTCGGGCT
TGCAGGGTTCCCCAAGAGGGATGAGAGCCCGCACGACGCGTTCGGAGCTGGTCACAGCTCCACT
AGTATTTCAGCTGGTCTAGGGATGGCGGTGGGGAGGGACTTGCTACAGAAGAACAACCACGTGA
TCTCGGTGATCGGAGACGGAGCCATGACAGCGGGGCAGGCATACGAGGCCTTGAACAATGCAGG
ATTTCTTGATTCCAATCTGATCATCGTGTTGAACGACAACAAACAAGTGTCCCTGCCTACAGCC
ACCGTCGACGGCCCTGCTCCTCCCGTCGGAGCCTTGAGCAAAGCCCTCACCAAGCTGCAAGCAA
GCAGGAAGTTCCGGCAGCTACGAGAAGCAGCAAAAGGCATGACTAAGCAGATGGGAAACCAAGC
ACACGAAATTGCATCCAAGGTAGACACTTACGTTAAAGGAATGATGGGGAAACCAGGCGCCTCC
CTCTTCGAGGAGCTCGGGATTTATTACATCGGCCCTGTAGATGGACATAACATCGAAGATCTTG
TCTATATTTTCAAGAAAGTTAAGGAGATGCCTGCGCCCGGCCCTGTTCTTATTCACATCATCAC
CGAGAAGGGCAAAGGCTACCCTCCAGCTGAAGTTGCTGCTGACAAAATGCATGGTGTGGTGAAG
TTTGATCCAACAACGGGGAAACAGATGAAGGTGAAAACGAAGACTCAATCATACACCCAATACT
TCGCGGAGTCTCTGGTTGCAGAAGCAGAGCAGGACGAGAAAGTGGTGGCCATCCACGCGGCGAT
GGGAGGCGGAACGGGGCTGAACATCTTCCAGAAACGGTTTCCCGACCGATGTTTCGATGTCGGG
ATAGCCGAGCAGCATGCAGTCACCTTCGCCGCGGGTCTTGCAACGGAAGGCCTCAAGCCCTTCT
GCACAATCTACTCTTCCTTCCTGCAGCGAGGTTATGATCAGGTGGTGCACGATGTGGATCTTCA
GAAACTCCCGGTGAGATTCATGATGGACAGAGCTGGACTTGTGGGAGCTGACGGCCCAACCCAT
TGCGGCGCCTTCGACACCACCTACATGGCCTGCCTGCCCAACATGGTCGTCATGGCTCCCTCCG
ATGAGGCTGAGCTCATGCACATGGTCGCCACTGCCGCTGTCATTGATGATCGCCCTAGCTGCGT
TAGGTACCCTAGAGGAAACGGTATAGGGGTGCCCCTCCCTCCAAACAATAAAGGAATTCCATTA
GAGGTTGGGAAGGGAAGGATTTTGAAAGAGGGTAACCGAGTTGCCATTCTAGGCTTCGGAACTA
TCGTGCAAAACTGTCTAGCAGCAGCCCAACTTCTTCAAGAACACGGCATATCCGTGAGCGTAGC
CGAT
GCGAGATTCTGCAAGCCTCTGGATGGAGATCTGATCAAGAATCTTGTGAAGGAGCACGAAGTTC
TCATCACTGTGGAAGAGGGATCCATTGGAGGATTCAGTGCACATGTCTCTCATTTCTTGTCCCT
CAATGGACTCCTCGACGGCAATCTTAAGTGGAGGCCTATGGTGCTCCCAGATAGGTACATTGAT
CATGGAGCATACCCTGATCAGATTGAGGAAGCAGGGCTGAGCTCAAAGCATATTGCAGGAACTG
TTTTGTCACTTATTGGTGGAGGGAAAGACAGTCTTCATTTGATCAACATGTAA

SEQ ID NO:26 CfGGPPs1

CGCCCTCAACTGTTATGCAACAACATTTATATGGTAATGAGATGATAAATCCCAACTCCATAAC
CAAGTTTAAACAAAGTTTTAACCATCCAACATTAGTTTATTCAAGAATTAAGAATTCTTACAGT
TAAATATACAAATTCTACATCATAATGAAGAAAATGTTGCTTAGCAGAAGACTAAACACCAAGT
TTTAGATCATCATCATCATCATCACCAAAAAAAAGAATTGCAATGCTAATTTTGCAAGAATCTA
GTTTTGTATATATAACAGACACCAGACCTCATCAAGCAGAGCAAACTAGTTCTGCCTGTGAGCA
ATGTAATCGGCCAGCGCCACCAGCGGCGCCGCCTTCCGGGGGTCAAAATCCAGCAGCTGCTGCT
TGGCCTCCTCATTCAGCCTCTCAGCAAACTCCATAGCTTTCTCCAATCCCAGAAGCTTTGGGTA
GGTGGTCTTGTCGACGGCCAAGTCTTTGCCGGCCGTCTTCCCCAACTCCTCCGAGGATTTTGTG
ACATCTAAAATGTCATCCACCACTTGGAAGAGAAGCCCAATTTTTCTAGCAAAAGTTCTCAGTT
TCTCAACTTGATCACTGCTTCCTCCTCCCAAAATGGCCCCCAAAACTACAGAGGCCTCAAGCAA
TGCTGCAGTTTTGTGTATGTGTATGAATTCCAATGTGTCCAGTCCTACATTGGGATTGCCGGTG
CAATGCAAATCCACCACCTGCCCCGCCACCAGCCCCTCCGTCCCGATCGCCTTCGCCAACTCCG
CCACCGCCGCAAGAATCCTCTCAGGGGCCACCCCCGTGGTGGCAGTGGCGATGAATTCAAACGC
GAAGGCCAATAAAGCATCACCGGCGAGCACGGCGACGTTCTCGCCGAAGACTTTGTGATTGGTG
GGCTTGCCGCGGCGGAGGTCATCATTGTCCATACAGGGAAGATCATCGTGGATGAGAGACATGG
TGTGGATCATCTCCACCGCGCAGGCGGCGGGGATCGCCGCCGCTTGGGGGCCGCCCACCACCTC
GCAGGCGGCGATGCAGAGCATGGGGCGGACCCTCTTTCCGCCGGCGAGCAAGGAGTACCTCATG
GCCTCGTGGATCATCGGAGGGTTCTTCACCGCCACCGCGTCGTCGAGAGCCTTGTTCACGTGGG
TGGCTTTCTCGACGACGTAGGCGTTGAAATTGAAGGGCGCTTCATCTCCTCGGGTGAAGATTCT
TGCTTCCTCGCCGGTGAGGACGGCGGAGACGGCGAAGGAGGAAATGGGCCTCCGCGATTTGAGG
AAAGCGGGCTCGAATCTGGGATGGTGGATGAATTTGGAGGGGTGTGGTTGCTTGAAAATGGGGA
GGTTTTGAACCCAAGCATCGACCAGATTCATAGACCTCATTTTTTCTCTTAGGGGCTTTGTGTA
GTGGAAATGGTGAAGGCAAGAAACAGTTGTGGTTGTTTGCTGAGATC

SEQ ID NO:27 CfTPS1-wt

ATGTCATGGATGAACAACGGTAAAAACCTTAACTGCCAACTTACTCACAAGAAAATATCGAAAG
TAGCCGAGATTCGAGTTGCCACGGTGAACGCGCCGCCGGTGCACGATCAAGACGATTCCACAGA
AAATCAGTGCCATGACGCGGTGAATAATATTGAGGATCCGATCGAATACATAAGAACGCTGCTG
AGGACGACAGGGGACGGCCGAATAAGTGTGTCGCCGTATGACACTGCGTGGGTCGCTCTGATCA
AGGACTTGCAAGGACGCGATGCCCCCGAGTTTCCGTCGAGCCTGGAGTGGATCATACAGAATCA
GCTGGCCGATGGGTCGTGGGGCGATGCCAAGTTCTTCTGTGTGTATGATCGCCTCGTGAATACG
ATAGCATGCGTGGTGGCCTTGAGATCATGGGATGTTCATGCTGAAAAGGTGGAAAGAGGAGTGA
GATACATCAATGAAAATGTGGAAAAGCTTAGAGATGGAAATGAGGAACACATGACTTGTGGGTT
CGAAGTGGTGTTTCCTGCGCTTCTGCAGAGAGCTAAGAGCTTAGGGATCCAAGATCTTCCCTAT
GATGCTCCCGTCATTCAAGAGATATATCACTCCAGGGAACAAAAGTTGAAAAGGATTCCACTGG
AGATGATGCACAAAGTGCCAACTTCTTTATTATTTAGTCTGGAAGGGCTGGAGAATTTGGAGTG
GGATAAGCTTTTGAAACTGCAGTCAGCTGATGGCTCTTTCCTCACTTCTCCCTCCTCCACTGCC
TTCGCTTTTATGCAAACTCGTGATCCTAAATGCTACCAATTCATCAAAAACACTATTCAAACTT
TCAACGGAGGAGCACCACACACTTATCCTGTCGATGTTTTTGGAAGACTTTGGGCAATCGACAG
GCTGCAGCGCCTCGGGATTTCTCGCTTCTTTGAGTCCGAGATTGCTGATTGCATCGCCCACATC
CACAGGTTTTGGACAGAGAAGGGAGTTTTCAGTGGAAGAGAATCAGAGTTTTGCGACATTGATG
ATACATCCATGGGAGTCCGACTCATGAGAATGCATGGATACGATGTTGATCCAAATGTATTGAA
GAACTTCAAAAAGGATGACAAGTTTTCATGCTACGGTGGACAGATGATTGAGTCTCCGTCTCCC
ATTTACAATCTCTACAGGGCTTCCCAACTCCGCTTCCCCGGTGAGCAAATTCTCGAAGATGCCA
ACAAATTTGCCTACGATTTCTTACAAGAAAAGCTTGCCCACAACCAGATTCTTGATAAATGGGT
TATATCTAAGCACTTGCCTGATGAGATAAAACTGGGACTGGAGATGCCGTGGTACGCCACCCTA
CCCCGCGTGGAGGCAAGATACTACATACAGTACTATGCTGGTTCAGGCGATGTATGGATCGGAA
AGACTCTCTACAGGATGCCCGAGATCAGCAACGATACATATCATGAGCTTGCAAAAACAGACTT
CAAGAGATGCCAAGCTCAGCATCAGTTTGAGTGGATTTACATGCAAGAATGGTACGAGAGTTGC
AACATGGAAGAATTCGGGATAAGCAGAAAGGAGCTTCTGGTTGCTTACTTCTTGGCGACTGCAA

GCATATTCGAGCTGGAGAGGGCTAATGAGAGAATCGCCTGGGCCAAATCCCAAATCATTTCCAC
CATCATTGCATCTTTCTTCAATAACCAAAACACTTCACCGGAGGATAAACTTGCATTTTTAACA
GATTTCAAAAATGGCAACTCCACAAACATGGCTCTGGTGACCCTCACTCAATTCCTAGAGGGAT
TCGACAGATACACTAGCCATCAGTTGAAGAATGCCTGGAGCGTATGGCTGAGAAAGCTGCAGCA
AGGAGAAGGCAACGGCGGCGCAGACGCAGAGCTCCTAGTAAACACATTGAACATTTGTGCCGGC
CACATTGCCTTTAGGGAAGAAATACTCGCACACAACGACTACAAGACTCTCTCCAACCTGACTA
GCAAAATCTGTCGACAACTTTCTCAAATTCAAAATGAAAAGGAGTTGGAGACAGAGGGACAGAA
AACAAGCATAAAAAACAAGGAACTGGAAGAAGATATGCAAAGACTGGTGAAGTTGGTGTTGGAG
AAATCAAGGGTTGGAATCAACAGAGATATGAAGAAAACATTTCTTGCAGTGGTAAAAACTTATT
ACTACAAAGCATATCATTCTGCTCAGGCCATCGACAACCATATGTTCAAAGTACTTTTCGAACC
AGTCGCCctcgagTGCTG

SEQ ID NO:28 CfTPS1-codon optimised

AAAAATGGCTTGGATGAACAACGGTAAGAATTTGAATTGCCAATTGACCCACAAGAAGATCTCT
AAGGTTGCCGAAATTAGAGTTGCTACTGTTAATGCTCCACCAGTTCATGATCAAGATGACTCTA
CTGAAAATCAATGCCATGATGCCGTTAACAACATCGAAGATCCAATCGAATATATCAGAACCTT
GTTGAGAACTACCGGTGATGGTAGAATTTCTGTTTCTCCATATGATACTGCTTGGGTCGCTTTG
ATTAAGGACTTGCAAGGTAGAGATGCTCCAGAATTTCCATCTTCATTGGAATGGATCATCCAAA
ATCAATTGGCTGATGGTTCTTGGGGTGATGCTAAGTTTTTTTGCGTTTACGATAGATTGGTCAA
CACCATTGCTTGTGTTGTTGCTTTGAGATCTTGGGATGTTCATGCTGAAAAAGTTGAAAGAGGT
GTCAGATATATCAACGAAAACGTCGAAAAGTTGAGAGATGGTAACGAAGAACATATGACCTGTG
GTTTCGAAGTTGTTTTCCCAGCTTTGTTGCAAAGAGCTAAGTCTTTGGGTATTCAAGATTTGCC
ATATGATGCCCCAGTTATCCAAGAAATCTATCACTCTAGAGAACAAAAGTCCAAGAGAATCCCA
TTGGAAATGATGCATAAGGTCCCAACTAGTTTGTTGTTCTCTTTGGAAGGTTTGGAAAACTTGG
AATGGGACAAGTTGTTGAAGTTGCAATCAGCAGATGGTTCCTTTTTGACTTCTCCATCTTCTAC
TGCTTTCGCTTTCATGCAAACTAGAGATCCAAAGTGCTACCAATTCATCAAGAACACCATTCAA
ACTTTCAACGGTGGTGCTCCACATACTTATCCAGTTGATGTTTTTGGTAGATTGTGGGCCATTG
ACAGATTGCAAAGATTGGGTATTTCCAGATTCTTCGAATCCGAAATTGCTGACTGCATTGCCCA
TATTCATAGATTCTGGACTGAAAAGGGTGTTTTCTCTGGTAGAGAATCTGAATTCTGCGATATC
GATGATACCTCTATGGGTGTTAGATTGATGAGAATGCATGGTTACGATGTTGATCCAAACGTCT
TGAAGAATTTCAAGAAGGACGATAAGTTCTCTTGCTACGGTGGTCAAATGATTGAATCTCCATC
TCCAATCTACAACTTGTACAGAGCTTCCCAATTGAGATTTCCAGGTGAACAAATTTTGGAAGAT
GCCAACAAGTTCGCCTACGACTTTTTACAAGAAAAGTTGGCCCATAATCAAATCTTGGACAAGT
GGGTTATTTCCAAACATTTGCCAGACGAAATCAAGTTGGGTTTAGAAATGCCATGGTATGCTAC
TTTGCCAAGAGTTGAAGCCAGATATTACATCCAATATTACGCTGGTTCTGGTGATGTTTGGATT
GGTAAAACCTTGTATAGAATGCCAGAAATCTCCAACGATACCTATCATGAATTGGCTAAGACCG
ATTTCAAGAGATGTCAAGCTCAACATCAATTTGAATGGATCTACATGCAAGAATGGTACGAATC
TTGCAACATGGAAGAATTCGGTATCTCCAGAAAAGAATTATTGGTCGCTTACTTCTTGGCTACC
GCTTCTATTTTTGAATTGGAAAGAGCCAACGAAAGAATTGCTTGGGCTAAGTCTCAAATCATCT
CTACTATTATCGCCTCCTTCTTCAACAATCAAAACACCTCTCCAGAAGATAAGTTGGCTTTCTT
GACTGACTTTAAGAACGGTAACTCTACCAACATGGCTTTGGTTACTTTGACCCAATTCTTAGAA
GGTTTCGACAGATACACTTCCCACCAATTGAAAAATGCTTGGTCTGTTTGGTTGAGAAAGTTGC
AACAAGGTGAAGGTAATGGTGGTGCTGATGCTGAATTATTAGTTAACACCTTGAACATTTGCGC
CGGTCATATTGCTTTCAGAGAAGAAATTTTGGCTCACAACGATTACAAGACCTTGTCTAACTTG
ACCTCTAAGATCTGCAGACAATTGAGTCAAATCCAAAACGAAAAAGAATTGGAAACCGAAGGTC
AAAAGACCTCCATTAAGAACAAAGAATTAGAAGAAGATATGCAAAGATTAGTCAAGTTGGTCTT
GGAAAAGTCCAGAGTTGGTATCAACAGAGACATGAAGAAAACTTTCTTGGCCGTTGTTAAGACC
TACTACTACAAAGCTTATCATTCCGCTCAAGCCATCGATAACCATATGTTTAAGGTTTTGTTCG
AACCAGTCGCCTGA

SEQ ID NO:29 CfTPS3-wt

ATGATCACCTCTAAATCATCTGCAGCTGTTAAATGCAGCCTCACCACGCCAACAGATTTGATGG
GGAAAATAAAAGAGGTCTTCAACAGGGAAGTCGATACTTCTCCGGCAGCCATGACTACTCATTC
TACAGATATACCCTCTAATCTCTGCATAATCGACACCCTCCAGAGGCTGGGAATCGACCAATAC
TTCCAATCCGAAATCGACGCTGTTCTACATGATACATACAGGTTATGGCAACTGAAAAAGAAAG
ATATATTTTCGGATATTACTACTCATGCAATGGCGTTCAGACTTTTGCGAGTCAAAGGATATGA
AGTTGCATCAGACGAACTGGCTCCATACGCTGATCAAGAGCGCATTAACCTGCAAACCATTGAT
GTGCCGACAGTTGTTGAGCTATACAGAGCAGCACAGGAGAGATTAACTGAAGAAGATAGCACTC
TTGAGAAACTGTATGTTTGGACCAGCGCCTTTCTGAAGCAGCAGTTGCTCACTGATGCCATTCC
TGACAAGAAATTGCACAAACAAGTGGAATACTACTTGAAGAACTACCATGGCATATTAGATAGA
ATGGGAGTGAGACGAAACCTCGACCTATATGACATAAGCCATTATAAAAGTCTCAAAGCTGCTC
ACAGGTTCTATAATCTGAGTAATGAAGATATCCTAGCATTTGCGAGGCAAGATTTTAATATTAG
CCAAGCCCAACACCAGAAAGAACTTCAGCAGCTGCAAAGGTGGTATGCAGATTGTAGGTTGGAC
ACGTTGAAATTTGGAAGAGATGTAGTGCGTATAGGAAATTTTCTGACTTCAGCAATGATTGGTG
ATCCTGAATTGTCTGACCTCCGTCTAGCGTTTGCCAAACATATAGTGCTCGTAACACGTATTGA
TGATTTTTTCGATCACGGTGGGCCTAAAGAAGAATCATACGAGATCCTTGAATTAGTAAAAGAA
TGGAAAGAGAAGCCAGCAGGAGAATATGTTTCTGAAGAAGTTGAAATCCTATTTACAGCAGTAT
ACAATACAGTAAACGAGTTGGCAGAAATGGCTCATATCGAACAAGGACGAAGCGTTAAAGACCT
TCTAGTTAAACTGTGGGTTGAAATACTATCAGTTTTCAGAATAGAATTGGATACATGGACCAAC
GACACAGCACTTACCTTAGAAGAGTACTTGTCACAATCCTGGGTGTCCATTGGCTGCAGAATCT
GCATTCTCATATCAATGCAATTCCAAGGTGTAAAATTATCTGATGAAATGCTTCAGAGTGAAGA
ATGCACTGATTTGTGTCGGTATGTTTCAATGGTTGACCGGCTGCTCAACGATGTGCAAACTTTT
GAGAAGGAACGCAAGGAAAATACAGGAAATAGTGTGAGCCTTCTGCAAGCAGCTCACAAAGATG
AAAGAGTCATTAATGAAGAGGAAGCTTGTATAAAGGTAAAAGAATTGGCTGAATATAACAGGAG
AAAACTGATGCAGATTGTCTACAAAACAGGAACCATTTTCCCAAGAAAATGCAAAGATCTGTTT
TTGAAGGCATGCAGAATTGGTTGTTATTTGTACTCAAGTGGCGACGAATTTACTTCGCCTCAAC
AAATGATGGAAGATATGAAGTCACTGGTTTATGAACCCCTACCAATTTCTCCTCCTGAAGCTAA
TAATGCAAGTGGAGAAAAAATGAGTTGTGTCAGCAACTAG

SEQ ID NO:30 CfTPS3-codon optimized for expression in yeast
ATGATCACCTCCAAATCTTCCGCTGCTGTTAAGTGTTCTTTGACTACTCCAACTGATTTGATGG
GTAAGATCAAAGAAGTTTTCAACAGAGAAGTTGATACCTCTCCAGCTGCTATGACTACTCATTC
TACTGATATTCCATCCAACTTGTGCATCATCGATACCTTGCAAAGATTGGGTATCGACCAATAC
TTCCAATCCGAAATTGATGCTGTCTTGCATGATACTTACAGATTGTGGCAATTGAAGAAGAAGG
ACATCTTCTCTGATATTACCACTCATGCTATGGCCTTCAGATTATTGAGAGTTAAGGGTTACGA
AGTTGCCTCTGATGAATTGGCTCCATATGCTGATCAAGAAAGAATCAACTTGCAAACCATTGAT
GTTCCAACCGTCGTCGAATTATACAGAGCTGCACAAGAAAGATTGACCGAAGAAGATTCTACCT
TGGAAAAGTTGTACGTTTGGACTTCTGCTTTCTTGAAGCAACAATTATTGACCGATGCCATCCC
AGATAAGAAGTTGCATAAGCAAGTCGAATATTACTTGAAGAACTACCACGGTATCTTGGATAGA
ATGGGTGTTAGAAGAAACTTGGACTTGTACGATATCTCCCACTACAAATCTTTGAAGGCTGCTC
ATAGATTCTACAACTTGTCTAACGAAGATATTTTGGCCTTCGCCAGACAAGATTTCAACATTTC
TCAAGCCCAACACCAAAAAGAATTGCAACAATTGCAAAGATGGTACGCCGATTGCAGATTGGAT
ACTTTGAAATTCGGTAGAGATGTCGTCAGAATCGGTAACTTTTTAACCTCTGCTATGATCGGTG
ATCCAGAATTGTCTGATTTGAGATTGGCTTTTGCTAAGCACATCGTTTTGGTTACCAGAATCGA
TGATTTCTTCGATCATGGTGGTCCAAAAGAAGAATCCTACGAAATTTTGGAATTGGTCAAAGAA
TGGAAAGAAAAGCCAGCTGGTGAATACGTTTCTGAAGAAGTCGAAATCTTATTCACCGCTGTTT
ACAACACCGTTAACGAATTGGCTGAAATGGCCCATATTGAACAAGGTAGATCTGTTAAGGATTT
GTTGGTTAAGTTGTGGGTCGAAATATTGTCCGTTTTCAGAATCGAATTGGATACCTGGACTAAC
GATACTGCTTTGACTTTGGAAGAATACTTGTCCCAATCCTGGGTTTCTATTGGTTGCAGAATCT
GCATTTTGATCTCCATGCAATTCCAAGGTGTTAAGTTGAGTGACGAAATGTTGCAAAGTGAAGA
ATGTACCGATTTGTGCAGATACGTTTCCATGGTCGATAGATTATTGAACGATGTCCAAACCTTC
GAAAAAGAAAGAAAAGAAAACACCGGTAACTCCGTTTCTTTGTTGCAAGCTGCTCACAAAGACG
AAAGAGTTATCAACGAAGAAGAAGCCTGCATCAAGGTAAAAGAATTAGCCGAATACAATAGAAG
AAAGTTGATGCAAATCGTCTACAAGACCGGTACTATTTTCCCAAGAAAATGCAAGGACTTGTTC

TTGAAGGCTTGTAGAATTGGTTGCTACTTGTACTCTTCTGGTGATGAATTCACTTCCCCACAAC
AAATGATGGAAGATATGAAGTCCTTGGTCTATGAACCATTGCCAATTTCTCCACCTGAAGCTAA
CAATGCATCTGGTGAAAAAATGTCCTGCGTCAGTAACTGA

SEQ ID NO:31 OsCPSsyn
ATGCCGGCCTCGGAAGAAACACGCGAATGGGAAGCAGAAGGACAACATGAACACACGGATGAAC
TAAGAGAGACGACGACAACCATGATCGATGGCATCAGGACAGCACTGAGATCAATCGGAGAGGG
AGAGATTAGCATCTCAGCCTATGACACTTCGTTGGTTGCCCTTCTGAAGAGGCTAGATGGGGGT
GATGGTCCTCAGTTCCCATCAACCATCGACTGGATCGTTCAGAATCAGCTACCGGATGGTTCAT
GGGGTGATGCCTCCTTCTTCATGATGGGAGACCGGATCATGAGCACCCTCGCTTGTGTTGTAGC
GTTGAAGTCATGGAACATCCACACCGATAAATGCGAGAGAGGTTTGTTGTTTATCCAAGAAAAT
ATGTGGAGGTTGGCCCATGAGGAAGAAGACTGGATGCTAGTTGGATTTGAGATTGCCTTGCCCT
CGCTCCTAGACATGGCTAAGGACCTGGATCTTGACATCCCTTACGATGAGCCAGCCTTGAAAGC
AATATATGCCGAGAGAGAAAGGAAGCTTGCCAAGATTCCAAGAGACGTGCTACACGCTATGCCA
ACAACTTTACTTCATAGCCTTGAGGGAATGGTTGACTTGGACTGGGAAAAGCTCCTCAAGCTCC
GGTGTCTCGATGGGTCCTTCCATTGCTCCCCTGCTTCGACGGCTACTGCTTTCCAGCAAACAGG
AGACCAGAAATGCTTTGAATACCTCGATGGAATCGTCAAAAAGTTCAATGGAGGAGTTCCCTGT
ATCTACCCTTTGGATGTGTACGAACGCTTATGGGCCGTCGATAGGCTGACGAGGCTGGGCATAT
CAAGGCACTTCACAAGTGAAATTGAGGATTGCTTAGACTACATTTTCAGGAACTGGACTCCAGA
TGGATTAGCTCACACAAAGAACTGCCCGGTAAAAGATATCGATGACACGGCCATGGGTTTCCGT
CTCCTCCGACTTTACGGCTACCAAGTCGACCCATGCGTGTTGAAGAAGTTCGAAAAGGATGGCA
AGTTCTTCTGCTTGCACGGGGAGTCCAACCCATCCTCTGTCACCCCAATGTACAACACTTACCG
GGCCTCCCAGCTCAAATTTCCTGGCGATGACGGTGTCCTTGGGCGAGCTGAGGTGTTTTGCCGC
TCATTCCTCCAAGACAGGAGAGGCTCAAACAGAATGAAGGACAAGTGGGCCATCGCCAAGGATA
TCCCAGGCGAGGTTGAGTATGCTATGGACTACCCATGGAAAGCAAGTTTACCGCGTATTGAAAC
AAGGTTGTACTTGGATCAATATGGAGGTAGTGGCGATGTATGGATTGGGAAGGTCCTGCACAGG
ATGACTCTTTTCTGCAACGACCTGTACCTCAAGGCAGCTAAAGCTGACTTCAGTAATTTCCAGA
AAGAGTGCCGAGTTGAGTTGAATGGCCTTAGAAGGTGGTATTTGAGGAGTAATCTGGAGAGGTT
TGGAGGGACTGATCCACAGACTACACTGATGACATCCTACTTCTTAGCTTCAGCGAACATCTTC
GAGCCAAACCGAGCAGCGGAACGTCTTGGATGGGCTCGCGTAGCGTTGCTTGCCGATGCCGTCT
CCTCTCACTTCAGGAGAATCGGGGGACCAAAAAATTTGACCAGTAATCTTGAAGAGCTTATCAG
CCTTGTTCCATTTGACGACGCTTATTCTGGCAGTCTTCGTGAAGCTTGGAAGCAGTGGCTCATG
GCATGGACTGCAAAGGAGAGCAGCCAGGAGTCAATTGAAGGGGACACGGCAATATTGTTGGTTC
GTGCCATCGAGATTTTTGGAGGACGGCATGTTTTGACTGGGCAAAGACCGGACCTTTGGGAGTA
TTCCCAGCTCGAGCAGCTCACCTCCTCCATCTGCCGCAAACTGTACAGGAGGGTTCTTGCCCAG
GAGAATGGGAAAAGTACGGAGAAAGTTGAGGAGATAGACCAGCAATTGGATTTGGAGATGCAGG
AATTGACTCGGCGCGTTCTTCAGGGCTGCAGCGCTATTAACAGACTAACCCGGGAGACGTTTCT
CCATGTGGTGAAGAGCTTCTGCTATGTCGCCTACTGCTCACCTGAGACAATTGATAACCACATC
GACAAGGTCATATTCCAAGATGTGATTTAG

SEQ ID NO:32 SsSCS-wt
ATGGCGAAAATGAAAGAGAATTTCAAGAGGGAAGACGATAAATTTCCAACGACAACGACTCTTC
GATCCGAAGATATACCCTCTAATTTGTGTATAATCGACACCCTTCAAAGGTTGGGGGTCGATCA
ATTCTTCCAATATGAAATCAACACTATTCTAGATAACACATTCAGGTTGTGGCAAGAAAAACAC
AAAGTTATATATGGCAATGTTACTACTCATGCAATGGCATTTAGGCTTTTGCGAGTGAAAGGAT
ACGAAGTTTCATCAGAGGAGTTGGCTCCATATGGTAACCAAGAGGCTGTTAGCCAGCAAACAAA
TGACCTGCCGATGATTATTGAGCTTTATAGAGCAGCAAATGAGAGAATATATGAAGAAGAGAGG
AGTCTTGAAAAAATTCTTGCTTGGACTACCATCTTTCTCAATAAGCAAGTGCAAGATAACTCAA
TTCCCGACAAAAAACTGCACAAACTGGTGGAATTCTACTTGAGGAATTACAAAGGCATAACCAT
AAGATTGGGAGCTAGACGAAACCTCGAGCTATATGACATGACCTACTATCAAGCTCTGAAATCT
ACAAACAGGTTCTCTAATTTATGCAACGAAGATTTTCTAGTTTTCGCAAAGCAAGATTTCGATA

TACATGAAGCCCAGAACCAGAAAGGACTTCAACAACTGCAAAGGTGGTATGCAGATTGTAGGTT
GGACACCTTAAACTTTGGAAGAGATGTAGTTATTATTGCTAATTATTTGGCTTCATTAATTATT
GGTGATCATGCGTTTGACTATGTTCGTCTCGCATTTGCCAAAACATCTGTGCTTGTAACAATTA
TGGATGATTTTTTCGACTGTCATGGCTCTAGTCAAGAGTGTGACAAGATCATTGAATTAGTAAA
AGAATGGAAGGAGAATCCGGATGCAGAGTACGGATCTGAGGAGCTTGAGATCCTTTTTATGGCG
TTGTACAATACAGTAAATGAGTTGGCGGAGAGGGCTCGTGTTGAACAGGGGCGTAGTGTCAAAG
AGTTTCTAGTCAAACTGTGGGTTGAAATACTCTCAGCTTTCAAGATAGAATTAGATACATGGAG
CAATGGCACGCAGCAAAGCTTCGATGAATACATTTCTTCGTCGTGGTTGTCGAACGGTTCCCGG
CTGACAGGTCTCCTGACGATGCAATTCGTCGGAGTAAAATTGTCCGATGAAATGCTTATGAGTG
AAGAGTGCACTGATTTGGCTAGGCATGTCTGTATGGTCGGCCGGCTGCTCAACGACGTGTGCAG
TTCTGAGAGGGAGCGCGAGGAAAATATTGCAGGAAAAAGTTATAGCATTCTACTAGCAACTGAG
AAAGATGGAAGAAAAGTTAGTGAAGATGAAGCCATTGCAGAGATCAATGAAATGGTTGAATATC
ACTGGAGAAAAGTGTTGCAGATTGTGTATAAAAAAGAAAGCATTTTGCCAAGAAGATGCAAAGA
TGTATTTTTGGAGATGGCTAAGGGTACGTTTTATGCTTATGGGATCAACGATGAATTGACTTCT
CCTCAGCAATCCAAGGAAGATATGAAATCCTTTG

SEQ ID NO:33 SsSCS-codon optimized for expression in yeast
AAAAATGGCCAAGATGAAGGAAAACTTCAAGAGAGAAGATGACAAGTTCCCAACTACTACTACC
TTGAGATCTGAAGATATCCCATCCAACTTGTGCATTATCGATACCTTGCAAAGATTGGGTGTTG
ACCAATTCTTCCAATACGAAATCAACACCATCTTGGACAACACTTTCAGATTGTGGCAAGAAAA
GCACAAGGTTATCTACGGTAATGTTACTACACATGCTATGGCCTTCAGATTATTGAGAGTTAAG
GGTTACGAAGTTTCCTCCGAAGAATTAGCTCCATACGGTAATCAAGAAGCCGTTTCTCAACAAA
CTAACGACTTGCCAATGATCATCGAATTATACAGAGCTGCCAACGAAAGAATCTACGAAGAAGA
AAGATCCTTGGAAAAGATTTTGGCTTGGACCACCATTTTCTTGAACAAGCAAGTTCAAGACAAC
TCCATCCCAGATAAGAAGTTGCATAAGTTGGTCGAATTCTACTTGAGAAACTACAAGGGTATCA
CCATTAGATTAGGTGCCAGAAGAAACTTGGAATTATACGACATGACTTACTACCAAGCCTTGAA
GTCTACCAACAGATTCTCTAACTTGTGTAACGAAGATTTCTTGGTTTTCGCCAAGCAAGATTTC
GATATTCACGAAGCCCAAAATCAAAAGGGTTTACAACAATTACAAAGATGGTACGCCGATTGCA
GATTGGATACTTTGAATTTCGGTAGAGATGTCGTCATTATCGCTAACTATTTGGCCTCCTTGAT
TATTGGTGATCATGCCTTTGATTACGTCAGATTGGCTTTTGCTAAGACCTCTGTTTTGGTTACC
ATCATGGATGATTTCTTCGATTGCCATGGTTCTTCTCAAGAATGCGACAAGATAATCGAATTGG
TAAAAGAATGGAAAGAAAACCCAGATGCCGAATACGGTTCTGAAGAATTGGAAATTTTGTTCAT
GGCCTTGTACAACACCGTTAACGAATTGGCTGAAAGAGCTAGAGTTGAACAAGGTAGATCTGTC
AAAGAATTTTTGGTCAAGTTGTGGGTTGAAATCTTGTCCGCTTTCAAGATTGAATTGGATACCT
GGTCTAACGGTACTCAACAATCTTTCGACGAATATATCTCCTCCTCTTGGTTGTCTAATGGTTC
TAGATTGACTGGTTTGTTGACCATGCAATTTGTTGGTGTCAAATTGTCCGACGAAATGTTGATG
TCAGAAGAATGTACTGATTTGGCTAGACACGTATGTATGGTCGGTAGATTATTGAACGATGTCT
GCTCATCTGAAAGAGAAAGAGAAGAAAACATTGCCGGTAAGTCCTACTCTATTTTGTTGGCTAC
TGAAAAGGACGGTAGAAAGGTTTCTGAAGATGAAGCTATTGCTGAAATCAACGAAATGGTCGAA
TACCATTGGAGAAAGGTCTTGCAAATCGTCTACAAGAAAGAATCCATCTTGCCTAGAAGATGCA
AGGACGTTTTTTTGGAAATGGCTAAGGGTACTTTTTACGCCTACGGTATTAACGATGAATTGAC
CTCTCCACAACAATCCAAAGAAGATATGAAGTCCTTCGTTTTTTAA

SEQ ID NO:34 CfCYP71D381
ATGGAGTTCGATTTCCCGTCAGCACTCATCTTTCCCGCAGTTTCATTGCTTCTTCTTCTGTGGC
TAACAAAAACCAGGAAACCCAAGAGCGATCTGGATCGCATCCCAGGCCCGAGGAGGCTCCCGCT
GATCGGAAACCTGCACCATCTCATCAGCCTGACTCCGCCGCCGCGTCTCTTCCGAGAAATGGCC
GCGAAATACGGGCCCCTGATGCGGCTGCAGCTGGGCGGAGTCCCCTTCCTCATCGTCTCCTCTG
TCGACGTGGCGAAGCATGTGGTCAAAACCAACGACGTGCCCTTCGCCAATCGTCCCCCCATGCA
CGCGGCGCGCGCCATAACCTACAACTACACCGACATCGGCTTCGCCCCCTACGGCGAGTACTGG
CGGAACCTCAGAAAGATCTGCACGCTCGAGCTGCTCAGCGCGCGGCGCGTGAGGTCATTCCGCC
ACATAAGGGAGGAGGAGAACGCCGGCGTGGCGAAGTGGATCGCTTCCAAGGAAGGGTCGCCGGC
GAATCTGTCGGAGAGGGTTTATCTCTCTTCATTCGACATCACAAGCAGGGCGTCCATCGGGAAG

GCGACGGAGGAGAAGCAAACGCTAACTTCCTCCATCAAGGATGCTATGAAGCTGGGCGGATTCA
ACGTCGCCGATCTTTATCCCTCCAGCAAGCTGCTGCTTCTCATCACCGGGCTGAATTTCAGGAT
ACAGAGAGTTTTCCGGAAAACAGATCGGATTCTGGATGATTTACTCAGCCAGCACCGGTCCACC
AGCGCCACCACGGAAAGGCCCGAAGATCTCGTCGACGTCCTCCTCAAGTATCAAAAGGAAGAAA
CCGAAGTCCATCTCAATAACGATAAAATCAAAGCAGTGATAATGGATATGTTTCTGGCCGGCGG
CGAGACATCAGCAACAGCGGTGGACTGGGCCATGGCGGAGATGATAAGAAACCCCACCACACTC
AAAAAGGCGCAAGAAGAAGTGAGACGTGTCTTCGACGGCAAGGGATACGTAGACGAAGAAGAGT
TCCACGAGCTAAAATACCTGAAACTGGTGATCAAAGAGATGCTGAGGATGCACCCGCCGCTGCC
ATTTCTGGTCCCACGCATGAACAGCGAGAGGTGCGAAATCAATGGATACGAAATCCCAGCCAAC
ACGAGACTGCTGATCAACGCATGGGCCATCGGAAGAGACCCCAAGTACTGGAATGACGCAGAGA
AGTTCATACCAGAGAGATTCGAGAACAGCTCCATCGATTTCAAAGGGAACAATCTGGAATACAT
ACCGTTTGGCGCAGGAAGGAGAATGTGCCCCGGGATGACGTTCGGGCTGGCGAGCGTGGAGTTC
ACGCTGGCCATGCTTCTGTATCACTTCGACTGGAAGATGCCCCAAGGGATTAAGCTGGACATGA
CGGAGTCCTTCGGCGCATCTCTCAAGAGGAAGCATGATTTGCTCATGATTCCCACTCTCAAGAG
GCCTTTGCGACTTGCTCCATGA

SEQ ID NO:35 CfCYP76AH8 wt
ATGGAAACCATCACTCTTCTCCTTGCTCTTTTCTTCATCGCTCTCACATATTTCATCTCCTCGA
GGCGCCGGAGAAACCTTCCTCCGGGGCCTTTTCCTCTTCCGATCATCGGAAACATGCTGCAGCT
CGGCTCCAAACCCCACCAGTCATTCGCCCAACTCTCAAAGAAATATGGGCCTCTCATGTCCATC
CACCTCGGGAGTTTATACACCGTGATCGTCTCCTCCCCGGAGATGGCAAAAGAGATACTGCAAA
AACACGGCCAAGTATTTTCGGGGCGTACCATCGCACAGGCGGTGCACGCGTGCGACCACGACAA
GATCTCCATGGGATTTCTGCCGGTGGCAAACACATGGCGTGATATGCGCAAAATATGCAAAGAG
CAGATGTTCTCGCATCACAGTCTGGAAGCCAGCGAGGAGCTCCGGCACCAGAAGCTGCAGCAGC
TGCTGGACTACGCCCAGAAATGCTGCGAAGCCGGTCGTGCCGTCGATATTCGTGAGGCCTCCTT
CATTACAACGCTCAACCTCATGTCTGCCACCATGTTTTCGACTCAAGCTACCGAGTTCGACTCA
GAAGCTACCAAGGAGTTCAAGGAGATCATCGAGGGGGTGGCCACCATTGTGGGTGTGGCCAATT
TCGCTGACTACTTCCCAATCTTGAAGCCTTTCGATCTGCAGGGGATCAAGCGCAGGGCAGATGG
CTACTTTGGAAGATTGCTTAAATTGATAGAGGGATATCTGAATGAAAGGCTGGAATCGAGAAGG
TTGAACCCAGATGCCCCCAGGAAGAAGGACTTTTTGGAAACACTCGTGGACATCATCGAGGCTA
ATGAATACAAGTTGACGACCGAGCACCTCACGCACCTCATGCTTGACTTGTTTGTTGGAGGATC
GGAGACAAACACGACCTCACTGGAGTGGATAATGTCGGAATTAGTGATCAACCCGGACAAAATG
GCGAAGGTGAAAGAGGAGCTGAAGAGCGTGGTAGGAGACGAGAAACTGGTGAACGAGTCAGACA
TGCCGAGGCTGCCATACTTGCAAGCAGTGATCAAAGAAGTGCTGCGAATTCACCCTCCCGGCCC
TCTTCTGCTTCCTCGCAAGGCAGAGAGTGATCAAGTCGTGAATGGGTACTTGATCCCAAAGGGG
ACTCAAATACTCTTCAATGCATGGGCAATGGGCAGAGACCCCACTATCTGGAAGGACCCTGAAT
CTTTTGAGCCCGAACGCTTCCTCAATCAAAGCATAGACTTCAAAGGCCAAGATTTCGAGCTGAT
CCCATTCGGCTCGGGGAGAAGAATCTGCCCCGGGATGCCGTTGGCGAATCGGATTCTGCACATG
ACGACGGCCACTCTGGTTCACAACTTCGACTGGAAATTGGAAGAAGGAACAGCTGATGCGGATC
ACAAAGGAGAGCTCTTTGGGCTGGCCGTTCGCCGTGCAACTCCTCTCAGGATCATTCCACTTAA
GCCATGA

SEQ ID NO:36 CfCYP76AH8 codon optimized for expression in yeast
ATGGAAACCATCACCTTGTTGTTGGCCTTGTTTTTCATTGCTTTGACCTACTTCATCTCCTCCA
GAAGAAGAAGAAATTTGCCACCAGGTCCATTTCCATTGCCAATTATTGGTAACATGTTGCAATT
GGGTTCCAAGCCACATCAATCTTTTGCTCAATTGTCCAAAAAGTACGGTCCATTGATGTCCATT
CATTTGGGTTCCTTGTACACCGTTATAGTCTCTTCACCAGAAATGGCCAAAGAAATCTTGCAAA
AACACGGTCAAGTTTTCTCCGGTAGAACTATTGCTCAAGCTGTTCATGCTTGTGATCACGATAA
GATTTCTATGGGTTTTTTGCCAGTTGCCAACACTTGGAGAGATATGAGAAAGATCTGCAAAGAA
CAAATGTTCTCCCACCATTCTTTGGAAGCTAGTGAAGAATTGAGACACCAAAAGTTGCAACAAT
TATTAGACTACGCTCAAAAGTGTTGCGAAGCTGGTAGAGCTGTTGATATTAGAGAAGCCTCTTT
CATTACCACCTTGAACTTGATGTCTGCTACTATGTTTTCTACCCAAGCTACCGAATTTGATTCC

GAAGCTACAAAAGAATTCAAAGAAATTATCGAAGGTGTCGCCACTATAGTTGGTGTTGCTAATT
TTGCTGATTACTTCCCAATCTTGAAGCCATTTGACTTGCAAGGTATTAAGAGAAGAGCTGATGG
TTACTTCGGTAGATTATTGAAGTTGATCGAAGGTTACTTGAACGAAAGATTGGAATCTAGAAGA
TTGAACCCAGATGCTCCAAGAAAGAAGGATTTCTTGGAAACCTTGGTTGATATCATCGAAGCCA
ACGAATACAAGTTGACTACTGAACATTTGACCCACTTGATGTTGGATTTGTTTGTTGGTGGTTC
TGAAACTAACACCACATCCTTGGAATGGATCATGTCTGAATTGGTTATCAACCCAGATAAGATG
GCCAAGGTCAAAGAAGAATTGAAGTCTGTTGTTGGTGACGAAAAGTTGGTTAACGAATCTGATA
TGCCAAGATTGCCATACTTGCAAGCCGTTATCAAAGAAGTTTTGAGAATTCATCCACCTGGTCC
TTTGTTGTTGCCAAGAAAAGCTGAATCTGATCAAGTTGTTAACGGTTATTTGATCCCAAAGGGT
ACTCAAATTTTGTTCAATGCTTGGGCTATGGGTAGAGATCCAACTATTTGGAAAGATCCAGAAT
CCTTCGAACCAGAAAGATTCTTGAATCAATCCATCGACTTCAAGGGTCAAGACTTCGAATTGAT
TCCATTTGGTTCTGGTAGAAGAATCTGTCCAGGTATGCCATTGGCTAATAGAATCTTGCATATG
ACTACCGCCACTTTGGTTCATAATTTCGATTGGAAATTGGAAGAAGGTACTGCTGACGCTGATC
ATAAGGGTGAATTATTTGGTTTGGCTGTTAGAAGAGCTACCCCATTGAGAATCATTCCATTGAA
ACCATAA

SEQ ID NO:37 CfCYP76AH11 wt
ATGGAGTTGGTGCAAGTGATTGCGGTGGTGGCGGTGGTGGTTGTGTTGTGGTCTCAGTTGAAGC
GAAAGGGCAGAAAGCTTCCACCAGGGCCTTCCCCTCTGCCCATCGTGGGAAACATTTTCCAGCT
GTCGGGTAAAAACATTAATGAATCATTTGCGAAGCTCTCCAAGATATACGGCCCCGTCATGTCC
CTACGCCTCGGCAGCTTACTCACGGTCATCATTTCGTCGCCGGAAATGGCGAAAGAAGTACTGA
CAAGCAAAGATTTTGCAAACCGGCCCCTCACGGAGGCGGCCCATGCGCACGGCCACTCCAAGTT
CTCCGTCGGGTTTGTGCCCGTCAGCGACCCAAAATGGAAACAAATGCGCCGCGTTTGCCAGGAG
GAAATGTTTGCGAGCCGGATCTTGGAAAACAGCCAGCAACGCCGCCACCAGAAGCTGCAGGAGC
TGATCGACCACGTGCAGGAATCCCGCGACGCTGGGCGAGCAGTCACTATCCGCGACCCCGTGTT
CGCAACCACGCTCAACATTATGTCGCTCACGTTGTTTTCGGCTGATGCCACTGAGTTCAGTTCA
AGCGCTACCGCTGAGTTGAGGGATATCATGGCGGGCGTCGTTTCTGTTCTTGGTGCTGCAAATT
TGGCCGACTTCTTCCCCATCCTCAAATATTTCGATCCGCAAGGGATGAGGCGCAAAGCAGATCT
CCACTACGGGAGACTCATCGACCACATCAAGAGCCGGATGGACAAGCGATCTGAGTTGAAAAAG
GCAAACCCGAATCATCCCAAGCACGACGACTTCTTGGAGAAGATCATAGACATCACTATACAGA
GGAACTACGACTTGACCATCAATGAGATCACGCATTTACTGGTGGACTTGTATCTAGCTGGAAG
CGAATCAACGGTGATGACAATTGAATGGACAATGGCGGAACTGATGCTGCGCCCGGAGTCACTA
GCGAAGCTGAAAGCCGAGCTGAGAAGCGTGATGGGAGAGAGAAAGATGATCCAGGAATCAGATG
ACATATCAAGGCTTCCATATCTGAATGGGGCCATCAAAGAAGCCCTCCGCCTCCATCCGCCTGG
CCCCCTCCTTTTTGCCCGCAAAAGCGAAATCGATGTGGAACTCAGCGGCTATTTCATCCCAAAA
GGCACTCAGATACTTGTGAATGAATGGGGCATGGGCAGAGACCCCAGCGTTTGGCCTAATCCAG
AGTGTTTCCAGCCGGAACGCTTCTTGGACAAAAACATTGATTACAAGGGCCAAGATCCCCAACT
CATCCCATTCGGAGCCGGAAGAAGGATCTGCCCGGGGATACCCATCGCGCATCGGGTTGTGCAT
TCAGTAGTGGCCGCATTAGTCCACAATTTCGACTGGGAGTTTGCCCCAGGCGGCAGTCAATGTA
ACAACGAGTTCTTCACCGGGGCAGCACTCGTCAGGGAAGTTCCTCTCAAACTCATCCCACTCAA
CCCACCTTCTATTTGA

SEQ ID NO:38 CfCYP76AH11 codon optimized for expression in yeast
ATGGAATTGGTCCAAGTTATCGCTGTTGTTGCAGTTGTTGTTGTTTTGTGGTCCCAATTGAAAA
GAAAGGGTAGAAAATTGCCACCAGGTCCATCTCCATTGCCAATAGTTGGTAATATCTTCCAATT
GTCCGGTAAGAACATCAACGAATCTTTCGCTAAGTTGTCCAAAATCTACGGTCCAGTTATGTCT
TTGAGATTGGGTTCTTTGTTGACCGTCATTATCTCTTCACCAGAAATGGCCAAAGAAGTCTTGA
CTTCTAAGGATTTTGCTAACAGACCATTGACTGAAGCTGCTCATGCTCATGGTCATTCTAAATT
TTCTGTTGGTTTCGTTCCAGTCTCTGATCCAAAATGGAAACAAATGAGAAGAGTCTGCCAAGAA
GAAATGTTCGCCTCTAGAATTTTGGAAAACTCCCAACAAAGAAGACACCAAAAGTTGCAAGAAT
TGATCGACCACGTTCAAGAATCTAGAGATGCTGGTAGAGCTGTTACTATTAGAGATCCAGTTTT
CGCTACCACCTTGAACATTATGTCCTTGACTTTGTTTTCTGCCGATGCTACTGAATTCTCTTCT

TCTGCTACTGCTGAATTGAGAGATATTATGGCTGGTGTTGTTTCTGTTTTGGGTGCTGCTAATT
TGGCTGATTTCTTCCCAATCTTGAAATACTTCGATCCACAAGGTATGAGAAGAAAGGCTGACTT
GCATTACGGTAGATTGATTGACCATATCAAGTCCAGAATGGACAAGAGATCTGAATTGAAGAAG
GCTAATCCAAACCATCCAAAGCACGATGATTTCTTGGAAAAGATCATCGACATCACCATTCAAA
GAAACTACGACTTGACCATTAACGAAATCACCCATTTGTTGGTCGACTTGTATTTGGCTGGTTC
TGAATCTACTGTTATGACCATTGAATGGACCATGGCCGAATTGATGTTAAGACCAGAATCATTG
GCTAAATTGAAGGCAGAATTGAGATCCGTTATGGGTGAAAGAAAGATGATCCAAGAATCCGACG
ACATTTCTAGATTGCCATACTTAAACGGTGCTATCAAAGAAGCCTTAAGATTGCATCCACCTGG
TCCTTTGTTGTTTGCTAGAAAGTCTGAAATCGATGTTGAATTGTCTGGTTACTTCATCCCAAAG
GGTACTCAAATCTTGGTTAATGAATGGGGTATGGGTAGAGATCCTTCTGTTTGGCCTAATCCAG
AATGTTTTCAACCAGAAAGATTTTTGGATAAGAACATTGACTACAAGGGTCAAGACCCACAATT
GATTCCATTTGGTGCAGGTAGAAGAATTTGTCCAGGTATTCCAATTGCCCATAGAGTTGTTCAT
TCAGTTGTTGCTGCTTTGGTTCATAACTTCGATTGGGAATTTGCTCCTGGTGGTTCTCAATGTA
ACAACGAATTTTTCACTGGTGCTGCCTTGGTTAGAGAAGTTCCATTGAAGTTGATTCCTTTGAA
CCCACCATCCATCTGA

SEQ ID NO:39 PsCYP720B4
ATGGCGCCCATGGCAGACCAAATATCATTACTGTTGGTGGTGTTCACGGTAGCGGTGGCGCTCC
TCCACCTTATTCACAGGTGGTGGAATATCCAGAGAGGCCCAAAAATGAGTAATAAGGAGGTTCA
TCTGCCTCCTGGGTCGACTGGATGGCCGCTTATTGGCGAAACCTTCAGTTATTATCGCTCCATG
ACCAGCAATCATCCCAGGAAATTCATCGACGACAGAGAGAAAAGATATGATTCGGACATTTTCA
TATCTCATCTATTTGGAGGCCGGACGGTTGTATCAGCGGATCCCCAGTTCAACAAGTTTGTTCT
ACAAAACGAGGGGAGATTCTTTCAAGCCCAATACCCAAAGGCACTGAAGGCTTTGATAGGCAAC
TACGGGCTGCTCTCTGTGCATGGAGATCTCCAGAGAAAGCTCCACGGAATAGCTGTGAATTTGC
TGAGGTTTGAGAGACTGAAAGTCGATTTCATGGAGGAGATACAGAATCTCGTGCACTCCACGTT
GGATAGATGGGCAGATATGAAGGAAATTTCTCTGCAGAATGAATGTCACCAGATGGTTCTCAAC
TTGATGGCCAAACAACTGCTGGATTTATCTCCTTCCAAAGAGACGAGTGATATTTGCGAGCTAT
TCGTTGACTATACCAATGCAGTGATTGCCATTCCCATCAAAATCCCAGGTTCCACCTATGCAAA
GGGGCTTAAGGCAAGGGAGCTTCTCATAAAAAAGATTTCAGAAATGATAAAAGAGAGAAGGAAT
CATCCTGAAGTTGTTCATAATGATTTGTTAACTAAACTTGTGGAAGAGGGGCTCATTTCAGATG
AAATTATTTGTGATTTTATTTTATTTTTACTTTTTGCTGGACATGAGACTTCCTCTAGAGCCAT
GACATTTGCTATCAAGTTTCTTACCTATTGCCCCAAGGCATTGAAGCAAATGAAGGAAGAGCAT
GATGCTATATTAAAATCAAAGGGAGGTCATAAGAAACTTAATTGGGATGACTACAAATCAATGG
CATTCACTCAATGTGTTATAAATGAAACACTTCGATTAGGTAACTTTGGTCCAGGGGTGTTTAG
AGAAGCTAAAGAAGACACTAAAGTAAAAGATTGTCTCATTCCAAAAGGATGGGTGGTATTTGCT
TTTCTGACTGCAACACATCTACATGAAAAGTTTCATAATGAAGCTCTTACTTTTAACCCATGGC
GATGGCAATTGGATAAAGATGTACCAGATGATAGTTTGTTTTCACCTTTTGGAGGTGGAGCTAG
GCTTTGTCCAGGATCTCATCTAGCTAAACTTGAATTGTCACTTTTTCTTCACATATTTATCACA
AGATTCAGTTGGGAAGCGCGTGCAGATGATCGTACCTCATATTTTCCATTACCTTATTTAACTA
AAGGCTTTCCCATTAGCCTTCATGGTAGAGTAGAGAATGAATAA

SEQ ID NO:40 amino acid sequence of CfCYP76AH15:
METMTLLLPLFFIALTYFLSWRRRRNLPPGPFPLPIIGNLLQIGSKPHQSFAQLSKKYGPLMSV
QLGSVYTVIASSPEMAKEILQKHGQVFSGRTIAQAAQACGHDQISIGFLPVATTWRDMRKICKE
QMFSHHSLESSKELRHEKLQKLLDYAQKCCEAGRAVDIREAAFITTLNLMSATLFSTQATEFDS
EATKEFKEVIEGVAVIVGEPNFADYFPILKPFDLQGIKRRANSYFGRLLKLMERYLNERLESRR
LNPDAPKKNDFLETLVDIIQADEYKLTTDHVTHLMLDLFVGGSETSATSLEWIMSELVSNPSKL
AKVKAELKSVVGEKKVVSESEMARLPYLQAVIKEVLRLHPPGPLLLPRKAGSDQVVNGYLIPKG
TQLLFNVWAMGRDPSIWKNPESFEPERFLNQNIDYKGQDFELIPFGSGRRICPGMPLADRIMHM
TTATLVHNFDWKLEDGAGDADHKGDDPFGLAIRRATPLRIIPLKP

SEQ ID NO:41 CfCYP76AH17 (aa):
MESMNALVVGLLLIALTILFSLRRRRNLAPGPYPFPIIGNMLQLGTKPHQSFAQLSKKYGPLMS
IHLGSLYTVIVSSPEMAKEILQKHGQVFSGRTIAQAVHACDHDKISMGFLPVSNTWRDMRKICK
EQMFSHHSLEGSQGLRQQKLLQLLDYAQKCCETGRAVDIREASFITTLNLMSATMFSTQATEFE
SKSTQEFKEIIEGVATIVGVANFGDYFPILKPFDLQGIKRKADGYFGRLLKLIEGYLNERLESR
KSNPNAPRKNDFLETVVDILEANEYKLSVDHLTHLMLDLFVGGSETNTTSLEWTMSELVNNPDK
MAKLKQELKSVVGERKLVDESEMPRLPYLQAVIKESLRIHPPGPLLLPRKAETDQEVNGYLIPK
GTQILFNVWAMGRDPSIWKDPESFEPERFLNQNIDFKGQDFELIPFGSGRRICPGMPLANRILH
MATATMVHNFDWKLEQGTDEADAKGELFGLAVRRAVPLRIIPLQP

SEQ ID NO:42 cytochrome P450 CYP76AH1 from Salvia miltiorrhiza_
MDSSPLLAALFFIAATIIFLSFRRRRNLPPGPFPYPIVGNMLQLGANPHQVFAKLSKRYGPLMS
IHLGSLYTVIVSSPEMAKEILHRHGQVFSGRTIAQAVHACDHDKISMGFLPVASEWRDMRKICK
EQMFSNQSMEASQGLRRQKLQQLLDHVQKCSDSGRAVDIREAAFITTLNLMSATLFSSQATEFD
SKATMEFKEIIEGVATIVGVPNFADYFPILRPFDPQGVKRRADVFFGKLLAKIEGYLNERLESK
RANPNAPKKDDFLEIVVDIIQANEFKLKTHHFTHLMLDLFVGGSDTNTTSIEWAMSELVMNPDK
MARLKAELKSVAGDEKIVDESAMPKLPYLQAVIKEVMRIHPPGPLLLPRKAESDQEVNGYLIPK
GTQILINAYAIGRDPSIWTDPETFDPERFLDNKIDFKGQDYELLPFGSGRRVCPGMPLATRILH
MATATLVHNFDWKLEDDSTAAADHAGELFGVAVRRAVPLRIIPIVKS

SEQ ID NO:43 CYP76AH4 from Rosmarinus officinalis
MDSFPLLVALFFILAATWLISFRRRRNLPPGPFPLPIVGNMLQLGAKPHQTFAQLSKKYGPLMS
IHLGSLYTVIVSSPEMAKEILHKHGQVFSGRTIAQAVHACDHDNISMGFLPVASEWRDMRKICK
EQMFSHQSMEASQGLRKQKLQQLLEFAQKCSDSGRAVDIRDAAFITTLNLMSATLFSMQATEFD
SEVTMEFKEIIEGVATIVGVPNFADYFPILKPFDPQGVKRRADAYFGKLLAKIEGYFNERIASR
KANPNAPKKDDFLETLVDIIEANEYKLKTDHLTHLMLDLFVGGSETNTTSIEWTMSELVMNPEK
MAKLKAELKSVAGDEKIVDESEIAKLPYLQAVIKEVMRIHPPGPLLLPRKAESDQEVNGYLIPK
GTQVLINAWAIGRDPSVWKNPDSFEPERFLEQKIDFKGQDFELLPFGSGRRVCPGLPLASRILH
MTAATLVHNFDWKLEDEATAEADHAGELFGLAVRRAVPLRIIPIVKS

SEQ ID NO:44 CfCPR
MESTIEKLSPFDLMTAILKGVKLDNSNGSAGVEHPAVIAMLMENKDLVMMLTTSVAVLLGLAVY
LVWRRGAGSAKRVVEPPKLVIPKGPVDAEEEDDGKKKVTIFFGTQTGTAEGFAKALAEEAKARY
PLTNFKVVDLDDYAADDEEYEEKMKKETFAFFFLATYGDGEPTDNAARFYKWFSEGKERGEIFK
NLNYGVFGLGNRQYEHFNKIAIVVDDILLEQGGNRLVPVGLGDDDQCIEDDFSAWRDNVWPELD
KLLRDEDDATVATPYTAAVLEYRVVFHDQSDELHSENNLANGHANGNASYDAQHPCKVNVAVKR
ELHTPLSDRSCTHLEFDISGTGLEYETGDHVGVYCENLIETVEEAERLLGLSPQTFFSVHTDKA
DGTPLGGSALPPPFPPCTLRTALSRYADLLNAPKKSALTALAAYASDPSEADRLKHLASPDGKE
EYAQYVVSGQRSLLEVMADFPSAKPPLGVFFAAIAPRLQPRFYSISSSPKIAPSRIHVTCALVY
EKMPTGRIHKGVCSTWMKNAVPLEESPNCSSAPVFVRTSNFRLPADPKVPVIMIGPGTGLAPFR
GFLQERLALKESGAELGPAILFFGCRNSKMDFIYQDELDNFVKAGVVSELVLAFSREGPAKEYV
QHKMAQKASDVWNMISEGGYVYVCGDAKGMARDVHRTLHTIVQEQGSLDSSKTESFVKNLQMTG
RYLRDVW

SEQ ID NO:45 cDNA encoding CfCPR
ATGGAATCGACTATTGAGAAGCTTTCGCCCTTCGATTTGATGACTGCGATTCTCAAAGGAGTCA
AACTTGATAATTCGAACGGGTCTGCTGGGGTGGAGCATCCGGCTGTGATCGCGATGCTGATGGA
GAACAAGGATCTCGTGATGATGCTCACCACCTCCGTCGCGGTGCTTCTAGGACTTGCTGTGTAT
CTCGTGTGGCGGCGCGGAGCCGGATCGGCGAAGAGGGTGGTGGAGCCGCCGAAGCTGGTGATTC

CCAAGGGCCCGGTGGATGCGGAGGAAGAGGATGATGGGAAGAAGAAGGTTACCATCTTTTTTGG
GACGCAGACTGGAACTGCTGAAGGCTTTGCTAAGGCACTTGCCGAAGAAGCTAAAGCAAGATAT
CCGCTGACCAACTTTAAAGTAGTTGACTTGGATGATTATGCTGCCGATGATGAAGAGTATGAAG
AGAAGATGAAGAAGGAGACCTTTGCATTCTTCTTCTTGGCGACATATGGAGATGGTGAGCCTAC
CGACAATGCTGCGAGATTTTACAAGTGGTTTTCCGAGGGGAAAGAGAGAGGTGAGATATTCAAG
AATCTCAACTATGGTGTATTTGGTCTTGGAAACAGGCAGTATGAGCATTTCAACAAGATTGCTA
TAGTGGTGGATGACATTCTTCTTGAGCAAGGTGGAAATCGGCTTGTCCCTGTGGGTCTTGGAGA
TGACGATCAATGTATCGAAGATGATTTCTCAGCATGGCGTGATAATGTGTGGCCTGAGCTGGAT
AAGTTGCTCCGTGATGAGGATGATGCAACTGTTGCAACTCCATATACTGCAGCCGTTTTGGAGT
ATCGTGTTGTGTTCCATGACCAGTCAGATGAACTGCACTCGGAAAACAACTTAGCCAATGGTCA
TGCAAaTGGAAATGCTTCTTATGATGCTCAaCACCCCTGCAAAGTGAATGTTGCTGTAAAAaGG
GAGCTACATACTCCTCTATCCGATCGTTCTTGCACTCACTTGGAATTCGACATATCTGGCACTG
GATTAGAGTATGAAACAGGGGACCACGTTGGTGTTTACTGTGAGAACTTGATTGAAACTGTAGA
GGAAGCAGAAAGGCTTCTTGGTCTTTCTCCACAAACATTCTTTTCAGTTCACACTGATAAAGCG
GACGGCACACCACTTGGTGGAAGTGCCTTGCCTCCTCCCTTCCCGCCGTGCACTTTGAGGACAG
CGCTAAGTCGATATGCTGATCTTTTGAATGCTCCCAAAAAGTCTGCTTTGACTGCATTGGCTGC
TTATGCCTCTGACCCTAGTGAAGCTGATCGGCTCAAGCACCTTGCTTCCCCTGATGGAAAGGAG
GAATATGCTCAATATGTGGTTTCTGGTCAGAGAAGCCTACTTGAGGTGATGGCTGACTTCCCAT
CTGCCAAGCCTCCTCTTGGTGTTTTCTTTGCTGCAATTGCTCCTCGCTTGCAGCCTCGATTTTA
TTCAATCTCATCCTCACCAAAGATTGCACCTTCAAGAATTCACGTCACTTGTGCGTTGGTGTAT
GAGAAAATGCCCACTGGACGAATCCACAAGGGTGTCTGCTCAACATGGATGAAGAATGCTGTGC
CATTGGAGGAAAGCCCCAACTGCTCTTCAGCACCAGTTTTTGTACGGACCTCAAACTTCAGACT
CCCTGCTGATCCTAAAGTACCAGTTATAATGATTGGCCCTGGAACCGGTTTGGCTCCATTCAGG
GGTTTTCTTCAGGAAAGATTAGCCCTCAAGGAATCTGGAGCAGAACTTGGTCCTGCTATATTAT
TCTTCGGGTGCAGAAACAGTAAAATGGATTTCATTTACCAAGATGAACTGGATAACTTTGTTAA
AGCTGGAGTGGTTTCTGAGCTTGTCCTTGCGTTTTCACGCGAGGGTCCTGCTAAGGAATACGTG
CAGCATAAGATGGCACAGAAGGCCTCGGATGTGTGGAATATGATATCAGAAGGGGGCTACGTTT
ATGTATGTGGTGATGCTAAGGGCATGGCACGTGACGTTCACCGGACTCTTCACACCATTGTTCA
AGAACAGGGATCTCTGGACAGCTCGAAAACCGAGAGCTTCGTCAAGAATCTGCAGATGACCGGC
CGGTACCTGCGTGACGTGTGGTGA

>

SEQ ID NO:46 – MvTPS5

MSITFNLKIAPFSGPGIQRSKETFPATEIQITASTKSTMTTKCSFNASTDFMGKLREKVGGKAD
KPPVVIHPVDISSNLCMIDTLQSLGVDRYFQSEINTLLEHTYRLWKEKKKNIIFKDVSCCAIAF
RLLREKGYQVSSDKLAPFADYRIRDVATILELYRASQARLYEDEHTLEKLHDWSSNLLKQHLLN
GSIPDHKLHKQVEYFLKNYHGILDRVAVRRSLDLYNINHHHRIPDVADGFPKEDFLEYSMQDFN
ICQAQQQEELHQLQRWYADCRLDTLNYGRDVVRIANFLTSAIFGEPEFSDARLAFAKHIILVTR
IDDFFDHGGSREESYKILDLVQEWKEKPAEEYGSKEVEILFTAVYNTVNDLAEKAHIEQGRCVK
PLLIKLWVEILTSFKKELDSWTEETALTLDEYLSSSWVSIGCRICILNSLQYLGIKLSEEMLSS
QECTDLCRHVSSVDRLLNDVQTFKKERLENTINSVGLQLAAHKGERAMTEEDAMSKIKEMADYH
RRKLMQIVYKEGTVFPRECKDVFLRVCRIGYYLYSSGDEFTSPQQMKEDMKSLVYQPVKIHPLE
AINV

Examples

The invention is further illustrated by the following examples, which however, should not be construed as limiting for the invention.

Example 1

Various full length cDNAs were prepared and expressed in a *Nicotiana benthamiana*/Agrobacterium system as described below. cDNA encoding p19, Cfdxs cDNA of SEQ ID NO:25 and CfGGPPs cDNA of SEQ ID NO:26 were expressed together with the different combinations listed below. The following combinations of cDNAs were expressed:

Experiment 1

cDNA encoding the following enzymes

1) CfTPS1 of SEQ ID NO:5
2) CfTPS3 of SEQ ID NO:12
3) CfCYP76AH11 of SEQ ID:21

were expressed in a *Nicotiana benthamiana*/Agrobacterium system.

Experiment 1.1

cDNA encoding the following enzymes

1) TwTPS7 of SEQ ID NO:4
2) CfTPS3 of SEQ ID NO:12
3) CfCYP76AH11 of SEQ ID:21

are expressed in a *Nicotiana benthamiana*/Agrobacterium system.

Experiment 2

cDNA encoding the following enzymes

1) CfTPS1 of SEQ ID NO:5
2) CfTPS3 of SEQ ID NO:12
3) CfCYP71D381 of SEQ ID:22

were expressed in a *Nicotiana benthamiana*/Agrobacterium system.

Experiment 2.1

cDNA encoding the following enzymes

1) TwTPS7 of SEQ ID NO:4
2) CfTPS3 of SEQ ID NO:12
3) CfCYP71D381 of SEQ ID:22

are expressed in a *Nicotiana benthamiana*/Agrobacterium system.

Experiment 3

cDNA encoding the following enzymes

1) CfTPS1 of SEQ ID NO:5
2) SsSCS of SEQ ID NO:11
3) CfCYP76AH8 of SEQ ID:20

were expressed in a *Nicotiana benthamiana*/Agrobacterium system.

Experiment 4

cDNA encoding the following enzymes

1) CfTPS1 of SEQ ID NO:5
2) SsSCS of SEQ ID NO:11
3) CfCYP76AH11 of SEQ ID:21

were expressed in a *Nicotiana benthamiana*/Agrobacterium system.

Experiment 5

cDNA encoding the following enzymes

1) CfTPS1 of SEQ ID NO:5
2) SsSCS of SEQ ID NO:11
3) PsCYP720B4 of SEQ ID:23

were expressed in a *Nicotiana benthamiana*/Agrobacterium system.

Experiment 6

cDNA encoding the following enzymes

1) Ossyn-CPP of SEQ ID NO:1
2) SsSCS of SEQ ID NO:11
3) CfCYP71D381 of SEQ ID:22

were expressed in a *Nicotiana benthamiana*/Agrobacterium system.

Experiment 7

cDNA encoding the following enzymes

4) TwTPS7 of SEQ ID NO:4
5) CfTPS3 of SEQ ID NO:12
6) CYP76AH1 of SEQ ID:42

were expressed in a *Nicotiana benthamiana*/Agrobacterium system.

Experiment 8

cDNA encoding the following enzymes

1) TwTPS7 of SEQ ID NO:4
2) CfTPS3 of SEQ ID NO:12
3) CYP76AH4 of SEQ ID:43

were expressed in a *Nicotiana benthamiana*/Agrobacterium system.

The cDNAs described above were cloned into the pCAMBIA130035Su vector. pCAMBIA130035Su containing cDNA encoding the enzymes described above and T-DNA expression plasmid containing the anti-post transcriptional gene silencing protein p19 (35S:p19)(Voinnet, Rivas et al. 2003), were transformed into the AGL-1 - GV3850 Agrobacterium strain by electroporation using a 2mm electroporation cuvette in a Gene Pulser (Bio-Rad; Capacity 25 µF; 2.5 kV; 400 Ω). The transformed agrobacteria were subsequently transferred to 1mL YEP (yeast extract peptone) media and grown for 2-3 hours at 30°C in YEP media. 200µL were transferred to YEP-agar solid media containing 35µg/mL rifampicillin, 50µg/mL carbencillin and 50µg/mL kanamycin and grown for 2 days. Multiple colonies were transferred from the plate to 20mL YEP media in falcon tube containing 17.5 µg/mL rifampicillin, 25 µg/mL carbencillin and 25 µg/mL kanamycin and grown at 30°C over night (ON) at 225 rpm. Agrobacteria were spun down by centriguation at 3500xg for 10 min and resuspended in 5mL H_2O. OD_{600} was measured and H_2O was added to reach an OD_{600}=1. 3mL of agrobacteria culture containing the plasmids encoding the enzymes outlined above for each experiment and p19 gene respectively were mixed. Controls only containing cDNAs encoding CfCYP76AH8 of SEQ ID:20, CfCYP76AH11 of SEQ ID:21, CfCYP71D381 of SEQ ID:22, or PsCYP720B4 of SEQ ID:23 was mixed similarly. Each mix of agrobacteria cultures were infiltrated into independent 4-6 weeks old N. benthamiana plants. Plants were grown for 7 days in greenhouse before metabolite extraction.

Extraction and GC-MS analysis

2 or 3 infiltrated leafs from each N. benthamiana line were chosen and from each of these 1 or 2 leaf disc's (Ø = 3cm) were carved out and added to 1 mL n-hexane with 1ppm 1-eicosene as internal standard (IS). The 2 or 3 replicates served as experimental replicates. Extraction was done at RT for 1 hour in an orbital shaker set at 220 rpm. Plant material was spun down and extracts were transferred to new vials.

Extracts were analyzed on a Shimadzu GCMS-QP2010 Ultra using an Agilent HP-5MS column (30 m x 0.250 mm i.d., 0.25 µm film thickness). Injection volume and temperature was set at 1µL and 250°C. GC program: 50 °C for 2 min, ramp at rate 4 °C min-1 to 110 °C, ramp at rate 8 °C min-1 to 250 °C, ramp at rate 10 °C min-1 to 310 °C and hold for 5 min. Both He and H_2 were used as carrier gas and hence the retentions times were normalized with Kovat's retention index using 1ppm C_7 - C_{30} Saturated Alkanes as reference. Electron impact (Ei) was used as ionization method in the mass spectrometer (MS) with the ion source temperature set to 230°C and 70 eV. MS spectra's was recorded from 50m/z to 350m/z. Compound identification was done by comparison to authentic standards and comparison to reference spectra databases (Wiley Registry of Mass Spectral Data, 8th Edition, July 2006, John Wiley & Sons, ISBN: 978-0-470-04785-9).

Figure 7 summarises the outcome of experiments 1 and 2. Cytochrome P450 enzymes CfCYP76AH11 (top) and CfCYP71D381 (bottom) catalyzes oxidation of Dehydroabietadiene (1.1) and Miltiradiene (1.2) into novel compounds 1.1.1-1.1.4 and 1.2.1 corresponding to hydroxylated or ketonated derivatives of Dehydroabietadiene (1.1) and Miltiradiene (1.2). As shown in Danish patent application PA 2014 00056 then expression of CfTPS1 and CfTPS3 in the presence of GGPP leads to formation of Dehydroabietadiene and Miltiradiene in N. benthamiana. Arrows indicate enzymatic reactions. Compound class identifications given in parenthesis are only based on GC-MS mass spectral analysis and comparison with the Wiley compound library. Experiment 1.1 will give the same results as experiment 1, and experiment 2.1 will give the same results as experiment 2.

Figure 8A shows GC-MS chromatogram of *N. benthamiana* of experiment 1 as well as mass spectra of compounds 1.1.1 & 1.2.1 shown in figure 7.

Figure 8B shows GC-MS chromatogram of *N. benthamiana* of experiment 2 as well as mass spectra of compounds 1.1.2-1.1.4 shown in figure 7.

Figure 9 summarises the outcome of experiments 3, 4 and 5. Cytochrome P450 enzymes CfCYP76AH8 (top panel), CfCYPAH11 (middle panel) and PsCYP720B4 (lower panel) catalyzes oxidation of Manool (1.3) into novel compounds 1.3.1-1.3.10 corresponding to hydroxylated, dihydroxylated or ketonated derivatives of Manool (1.3). As shown in Danish patent application PA 2014 00056 then expression of CfTPS1 and

SsSCS in the presence of GGPP leads to formation of Manool in N. benthamiana. Arrows indicate enzymatic reactions. Compound class identifications given in parenthesis are only based on GC-MS mass spectral analysis and comparison with the Wiley compound library.

Figure 10A shows GC-MS chromatogram of *N. benthamiana* of experiment 3 as well as mass spectra of compounds 1.3.1 & 1.3.5 shown in figure 9.

Figure 10B shows GC-MS chromatogram of *N. benthamiana* of experiment 4 as well as mass spectra of compounds 1.3.6 & 1.3.9 shown in figure 9.

Figure 10C shows GC-MS chromatogram of *N. benthamiana* of experiment 5 as well as mass spectra of compound 1.3.10 shown in in figure 9.

Figure 11 summarises the outcome of experiment 6. Cytochrome P450 enzyme CfCYP71D381 catalyzes oxidation of syn-manool (1.4) into novel compounds 1.4.1-1.4.3 corresponding to a hydroxylated derivatives of syn-manool (1.4). As shown in Danish patent application PA 2014 00056 then expression of Ossyn-CPP and SsSCS in the presence of GGPP leads to formation of syn-Manool in N. benthamiana. Arrows indicate enzymatic reactions. Compound class identification given in parenthesis is a tentative identification based on GC-MS mass spectral analysis and comparison with the Wiley compound library.

Figure 12 shows GC-MS chromatogram of *N. benthamiana* of experiment 6 as well as mass spectra of compounds 1.4.1-1.4.3 shown in in figure 11.

Co-expression of TwTPS7, CfTPS3 and CYP76AH1 resulted in the formation of ferruginol (experiment 7)(see fig. 20).

Co-expression of TwTPS7, CfTPS3 and CYP76AH4 resulted in the formation of ferruginol (experiment 8)(see fig. 20).

Example 2

Heterologous nucleic acids encoding the proteins described in Table 2 were expressed in *S. cerevisiae*.

Table 2. CDSs used in this Example.

CDS	Full length Protein SEQ ID	CDS SEQ ID	Description
GGPPS7			*Synechococcus* geranylgeranyl diphosphate synthase
CfTPS1	SEQ ID NO:5	SEQ ID NO:28	*Coleus forskohlii* diterpene synthase 1, (+)-CPP type diTPS
CfTPS3	SEQ ID NO:12	SEQ ID NO:30	*Coleus forskohlii* diterpene synthase 3, diTPS of class I
OsCPSsyn	SEQ ID NO:1	SEQ ID NO:	*Oryza sativa* syn-CPP synthase, diTPS of class-II
SsSCS	SEQ ID NO:11	SEQ ID NO:33	*Salvia sclarea* diterpene synthase, diTPS of class-I
CfCYP76AH8	SEQ ID NO:20	SEQ ID NO:36	*Coleus forskohlii* cytochrome P450
CfCYP76AH11	SEQ ID NO:21	SEQ ID NO:38	*Coleus forskohlii* cytochrome P450
CfCYP71D381	SEQ ID NO:22	SEQ ID NO:34	*Coleus forskohlii* cytochrome P450
PsCYP720B4	SEQ ID NO:23	SEQ ID NO:39	*Picea sitchensis* cytochrome P450
CfCPR			*Coleus forskohlii* cytochrome P450 reductase

In this study GGPPS7 from *Synechococcus* was used. However, another GGPPS could also have been expressed in yeast, e.g. CfGGPPS of SEQ ID NO:26. The sequence of *GGPPS7* of *Synechococcus sp.* is available under Gene ID:86553638.

Table 3. List of plasmids used in this Example.

Plasmid name	Scaffold plasmid	CDS and promoter
pCYPCC-1	pROP196 XI-5 assembler 1	Rv #205 GGPPs7<-pTPI1 #219
pCYPCC-3	pROP196 XI-5 assembler 1	Rv #205 GGPPs7<-pPGK1 1c
pCYPCC-7	pROP197 XI-5 assembler 3	#-3 CfTPS3 <-#161pTDH3
pCYPCC-9	pVAN858 assembler 2	2c pTEF1->#-5 CfTPS1
pCYPCC-10	pVAN858 assembler 2	2c pTEF1->#-6 OsCPssyn
pCYPCC-18	pROP197 XI-5 assembler 3	#-8 SsSCS <-#161pTDH3
pCYPCC-21	pROP197 XI-5 assembler 3	Res# 236 CfTPS3 co<-#161pTDH3
pCYPCC-25	EPSB585	Rv #238 CfCYP76AH8 co <- 5p-pTDH3/pTEF2 ->Res# 233_Cf co

pCYPCC-26	EPSB585	Rv #237 CfCYP76AH11 co <- 5p-pTDH3/pTEF2 ->Res# 233_CfCPR_tefU co
pCYPCC-27	EPSB585	#10- CfCYP71D381 <- 5p-pTDH3/pTEF2 ->Res# 233_CfCPR_t
pCYPCC-28	EPSB585	#11- PsCYP720B4 <- 5p-pTDH3/pTEF2 ->Res# 233_CfCPR_te
pCYPCC-29	EPSB585	Res# 239 CfCYP76AH6 (16) co <- 5p-pTDH3/pTEF2 ->Res# 233_CfCPR_tefU co
pCYPCC-30	pROP196 XI-5 assembler 1	Rv #205 GGPPs7 <- 6p-pTPI1/pPDC1 ->Res# 233_CfCPR_tefU
pCYPCC-31	pVAN858 assembler 2	Rv #237 CfCYP76AH11 co <-3c pgk-pTEF-1 ->Rv #234 CfCYP7
pCYPCC-34	pVAN858 assembler 2	#10- CfCYP71D381 <-3c pgk-pTEF-1 ->Rv #234 CfCYP76AH8
pCYPCC-35	pVAN858 assembler 2	#11- PsCYP720B4 <-3c pgk-pTEF-1 ->Rv #234 CfCYP76AH8 c
pCYPCC-37	pVAN858 assembler 2	Rv #238 CfCYP76AH8 co<-pPGK1 1c
pCYPCC-38	pVAN858 assembler 2	Rv #237 CfCYP76AH11 co<-pPGK1 1c
pCYPCC-39	pVAN858 assembler 2	#10- CfCYP71D381<-pPGK1 1c
pCYPCC-40	pVAN858 assembler 2	#11- PsCYP720B4<-pPGK1 1c
pCYPCC-41	pVAN858 assembler 2	Res# 239 CfCYP76AH6 (16) co<-pPGK1 1c
pCYPCC-42	pVAN858 assembler 2	Res160 pTEF-2 ->CfTPS1, co
pCYPCC-44	pVAN858 assembler 2	Res160 pTEF-2 ->OsCPssyn
pCYPCC-45	pROP197 XI-5 assembler 3	Res# 236 CfTPS3 truncated co <- 5p-pTDH3/pTEF2 ->CfTPS1
pCYPCC-46	pROP197 XI-5 assembler 3	Res# 236 CfTPS3 truncated co <- 5p-pTDH3/pTEF2 ->Tw14,
pCYPCC-47	pROP197 XI-5 assembler 3	Res# 236 CfTPS3 truncated co <- 5p-pTDH3/pTEF2 ->OssynC
pCYPCC-48	pROP197 XI-5 assembler 3	SsSCS, trun, co <- 5p-pTDH3/pTEF2 ->CfTPS1, trun co
pCYPCC-50	pROP197 XI-5 assembler 3	SsSCS, trun, co <- 5p-pTDH3/pTEF2 ->OssynCPP
pCYPCC-51	pROP197 XI-5 assembler 3	SsSCS, co<-#161pTDH3

All diTPS enzymes were truncated to remove putative plastid targeting sequence (see CDS sequences referred to in Table 2)

DNA sequences encoding enzymes of interest were in general codon optimized for expression in *S. cerevisiae*. Codon optimization for expression in *Saccharomyzes cerevisae* was performed using the Geneart service from LifeTechnologies.
DNA fragments encoding the enzymes of interest were USER cloned into the pre-digested plasmid backbones. All plasmids constructed and used in this study are summarized in table 3. DNA fragments of interest were liberated from plasmids by NotI enzyme-digestion as linear DNA fragments suitable for yeast transformation. The plasmids are designed to accommodate integration of up to three NotI-digested fragments at the same site in the genome.

Table 4. Strains used and generated in this Example

Name	Genotype

EFSC4416	MATa HO::KO HIS3 LEU2 Δura3
IT-4, T-5	MATa HO::KO HIS3 LEU2 Chr XI-5::GGPPS7<-pTPI, pTEF1->CfTPS1 trun, SsSCS trun <- TDH3, DR-KlURA3-DR
IT-4, T-5, uracil loop out	MATa HO::KO HIS3 LEU2 Chr XI-5::GGPPS7<-pTPI, pTEF1->CfTPS1 trun, SsSCS trun <- TDH3,
IT-5, T-6	MATa HO::KO HIS3 LEU2 Chr XI-5::GGPPS7<-pPGK, pTEF1->OssynCPP trun, CfTPS3 trun Codon Optimized <- TDH3, DR-KlURA3-DR
IT-5, T-6, uracil loop out	MATa HO::KO HIS3 LEU2 Chr XI-5::GGPPS7<-pPGK, pTEF1->OssynCPP trun, CfTPS3 trun Codon Optimized <- TDH3,
IT-7, T-1	MATa HO::KO HIS3 LEU2 Chr XI-5::GGPPS7<-pTPI, pTEF1->CfTPS1 trun codon optimized, CfTPS3 trun codon optimized <- TDH3,DR-KlURA3-DR
IT-7, T-2	MATa HO::KO HIS3 LEU2 Chr XI-5::GGPPS7<-pTPI, pTEF1->OsCPSsyn, CfTPS3 trun codon optimized <- TDH3,DR-KlURA3-DR
IT-7, T-4	MATa HO::KO HIS3 LEU2 Chr XI-5::GGPPS7<-pTPI, pTEF1->CfTPS1 trun codon optimized, SsSCS trun codon optimized <- TDH3,DR-KlURA3-DR
IT-7, T-9	MATa HO::KO HIS3 LEU2 Chr XI-5::GGPPS7<-pTPI-pPDC -> CfCPR; CfTPS3 truncated co <- 5p-pTDH3/pTEF2 ->CfTPS1, trun co; PsCYP720B4 <-3c pgk-pTEF-1 ->Rv #234 CfCYP76AH8 co,DR-KlURA3-DR
IT-7, T-13	MATa HO::KO HIS3 LEU2 Chr XI-5::GGPPS7<-pTPI-pPDC -> CfCPR; CfTPS3 truncated co <- 5p-pTDH3/pTEF2 ->CfTPS1, trun co; PsCYP720B4 <- pPGK1,DR-KlURA3-DR
IT-7, T-31	MATa HO::KO HIS3 LEU2 Chr XI-5::GGPPS7<-pTPI-pPDC -> CfCPR; SsSCS truncated co <- 5p-pTDH3/pTEF2 ->CfTPS1, trun co; Rv #237 CfCYP76AH11 co <-3c pgk-pTEF-1 ->Rv #234 CfCYP76AH8 co,DR-KlURA3-DR
IT-7, T-32	MATa HO::KO HIS3 LEU2 Chr XI-5::GGPPS7<-pTPI-pPDC -> CfCPR; SsSCS truncated co <- 5p-pTDH3/pTEF2 ->CfTPS1, trun co; CfCYP71D381 <-3c pgk-pTEF-1 ->Rv #234 CfCYP76AH8 co,DR-KlURA3-DR
IT-7, T-33	MATa HO::KO HIS3 LEU2 Chr XI-5::GGPPS7<-pTPI-pPDC -> CfCPR; SsSCS truncated co <- 5p-pTDH3/pTEF2 ->CfTPS1, trun co; PsCYP720B4 <-3c pgk-pTEF-1 ->Rv #234 CfCYP76AH8 co,DR-KlURA3-DR
IT-7, T-34	MATa HO::KO HIS3 LEU2 Chr XI-5::GGPPS7<-pTPI-pPDC -> CfCPR; SsSCS truncated co <- 5p-pTDH3/pTEF2 ->CfTPS1, trun co; Rv #238 CfCYP76AH8 co<-pPGK1 1c ,DR-KlURA3-DR
IT-7, T-37	MATa HO::KO HIS3 LEU2 Chr XI-5::GGPPS7<-pTPI-pPDC -> CfCPR; SsSCS truncated co <- 5p-pTDH3/pTEF2 ->CfTPS1, trun co;PsCYP720B4 <- pPGK1,DR-KlURA3-DR
IT-7, T-50	MATa HO::KO HIS3 LEU2 Chr XI-5::GGPPS7<-pTPI, pTEF1->CfTPS1 trun,SsSCS trun <- TDH3, Chr X-3:: CfCYP76AH8 co <- 5p-pTDH3/pTEF2 -> CfCPR_tefU co,DR-KlURA3-DR

IT-7, T-58	MATa HO::KO HIS3 LEU2 Chr XI-5::GGPPS7<-pPGK, pTEF1->OssynCPP trun, CfTPS3 trun Codon Optimized <- TDH3. Chr X-3:: PsCYP720B4 <- 5p-pTDH3/pTEF2 ->Res# 233_CfCPR_tefU co; DR-KlURA3-DR

All strains were grown in 96 deep well plates as follows. Single colonies were inoculated in 500 µl selective (no Uracil) Yeast-Synthetic Complete Medium in 2.2 ml 96 deep well plates and grown o/n @ 30°C, 400 RPM. The following day 50 µl of the o/n culture was used as inoculum in 500 µl Yeast Synthetic Complete-Medium. These cultures were grown for 72 hours @ 30°C, 400 RPM.

Metabolite extraction GC-MS analysis

Metabolites were extracted from broth (500 ul) by addition of 500 µl 99 % Ethanol UV-grade followed by 20 minutes incubation @ 60 degrees. Samples were cleared by centrifugation @ 3.000 g, 10 min and 800 µl of the supernatant was transferred to clean GC/LC vial. For GC analysis, non-polar metabolites were extracted from the ethanol extract by partitioning with 500 ul hexane. After extraction, the solvent was transferred into new 1.5-mL glass vials and stored at –20°C until GC-MS analysis. For LC-analysis, the cleared ethanol extract were stored at -20°C until analysis and applied without further modification.

GC-MS analysis

One microliter of hexane extract was injected into a Shimadzu GC-MS-QP2010 Ultra. Separation was carried out using an Agilent HP-5MS column (20m 0.180mm i.d., 0.18µm film thickness) with purge flow of 4 mL min^{-1} for 1 min, using H_2 as carrier gas. The GC temperature program was 60°C for 1 min, ramp at rate 30°C min^{-1} to 180°C, ramp at rate 10°C min^{-1} to 250°C, ramp at rate 30°C min^{-1} to 320°C, and hold for 3 min. Injection temperature was set at 250°C in splitless mode. Column flow and pressure was set to 5. mL min^{-1} and 66.7 kPa yielding a linear velocity of 66.5 cm s^{-1}. Ion source and transfer line for mass spectrometer (MS) was set to 300°C and 280°C respectively. MS was set in scan mode from m/z 50 to m/z 350 with a scan width of 0.5s. Solvent cutoff was 4 min.

LC-MS analysis

Analytical LC-MS was carried out using an Agilent 1100 Series LC (Agilent Technologies, Germany) coupled to a Bruker HCT-Ultra ion trap mass spectrometer (Bruker Daltonics, Bremen, Germany). A Zorbax SB-C18 column (Agilent; 1.8 µm,

2.1 x 50 mm) maintained at 35°C was used for separation. The mobile phases were: A, water with 0.1% (v/v) HCOOH and 50mM NaCl; B, acetonitrile with 0.1% (v/v) HCOOH. The gradient program was: 0 to 1 min, isocratic 50% B; 1 to 10 min, linear gradient 50 to 95% B; 10 to 11.4 min, isocratic 98% B; 11.4 to 17 min, isocratic 50% B. The flow rate was 0.2 mL min^{-1}. The mass spectrometer was run in alternating positive/negative mode and the range m/z 100-800 was acquired

Results

The yeast strain IT-7 T13 expressing the diterpene synthases CfTPS1 and CfTPS3 as well as the cytochrome P450 PsCYP720B4 produces oxidized miltiradiene , which tentatively was identified as miltiradienic acid (see figure 13). LC-MS analysis of ethanol extracts generated from media as well as cell material of yeast strain IT-7 T13 shows that the extracted ion m/z matches the predicted mass of miltiradiene carrying a single carboxylic acid group (303 m/z), compound 1.2.1. Compound 1.2.1 was not produced by a control yeast strain expressing CfTPS1 and CfTPS, but not PsCYP720B4. Formation of a carboxylic acid group is consistent with previously reported activity of PsCYP720B4 (Hamberger et al, 2011). 1.2.1 was tentatively identified as miltiradienic acid based on the mass to charge ratio, fragmentation pattern and retention time. The 257 fragment is a signature ion of labdane-type diterpenes.

The yeast strain IT7-T50 expressing CfTPS1, SsSCS and CfCYP76AH8 produces various oxidised manool as shown in figure 14.

Extracted ion chromatograms recorded by LC-MS analysis of ethanol extracts generated from media as well as cell material of yeast strain IT7-T50, shows that the extracted ion trace combine the ions corresponding to the mass of manool (273 m/z), monohydroxy-manool (289 m/z, hydrogen adduct of water loss ion), dihydroxymanool (347 m/z) and keto-hydroxy-manool (345 m/z). Thus, IT7-T50 produces keto-hydroxy mannol (1.3.11), dihydroxy manool (1.3.12) and hydroxyl manool (1.3.13).

The yeast strain IT7-T37 expressing CfTPS1, SsSCS, PsCYP720B4 produces various oxidised manool as shown in figure 15.

Extracted ion chromatograms recorded by LC-MS analysis of ethanol extracts generated from media as well as cell material of yeast strain IT7-T37, shows that the extracted ion trace combine the ions corresponding to the mass of monohydroxy-manool (329 m/z, sodium adduct), dihydroxymanool (345 m/z, sodium adduct) and

keto-hydroxy-manool or a carboxylic acid of manool (343 m/z, sodium adduct). Thus, IT7-T37 produces the oxidized compounds 1.3.12-1.3.16.

The yeast strain IT7-T58 expressing OssynCPP, CfTPS3, and PsCYP720B4 produces various oxidised isopimaradienes as shown in figure 16.
Extracted ion chromatograms recorded by LC-MS analysis of ethanol extracts generated from media as well as cell material of IT7-T58, shows that the extracted ion trace combine the ions corresponding to the mass of syn-pimara-9,(11),15-diene or similar $C_{20}H_{32}$ diterpene oxidized to contain either one keto- and one hydroxyl group or a single carboxylic acid group (303 m/z). Formation of a carboxylic acid group would be consistent with previously reported activity of PsCYP720B4 (Hamberger et al, 2011). The 257 m/z fragment observed in all compounds 1.5.1-1.5.5 is a signature ion of diterpenes. This ion is observed as 253 m/z in 1.5.6 possibly indicative of introduction of two double bonds in the structure. (OssynCPP, CfTPS3, PsCYP720B4, strain name IT7-T58)

Cytochrome P450 enzyme pairs can catalyse production of higher oxidised diterpenes. CYP76AH8, CYP76AH11 and CYP76AH8, CYP720B4 catalyse oxidation of Manool into novel compounds 1.3.14-1.3.18 tentatively identified as hydroxylated, dihydroxylated or hydroxylated and ketonated derivatives of manool based on their parental masses.

The yeast strain IT7-T32 expressing CfTPS1, SsSCS, CfCYP76AH8, CfCYP71D381 produces various oxidised manool as shown in figure 17.
Extracted ion chromatograms recorded by LC-MS analysis of ethanol extracts generated from media as well as cell material of IT7-T32, shows that the extracted ion trace combine the ions corresponding to the mass of monohydroxy-manool (289 m/z = hydrogen adduct of water loss ion) and dihydroxymanool (345 m/z = sodium adduct, 305 = hydrogen adduct of water loss ion). Metabolites 1.3.14 and 1.3.15 were tentatively identified as dihydroxy manool derivatives arising from oxidation of monohydroxy manool by CfCYP71D381. 1.3.13 has a characteristic parental mass matching a monohydroxylated manool.

The yeast strain IT7-T31 expressing CfTPS1, SsSCS, CfCYP76AH8, CfCYP76AH11 produces oxidised manool as shown in figure 18. Extracted ion chromatograms

recorded by LC-MS analysis of ethanol extracts generated from media as well as cell material of IT7-T31, shows that the extracted ion trace combine the ions corresponding to the mass of monohydroxy-manool (289 m/z = hydrogen adduct of water loss ion) and dihydroxymanool (345 m/z = sodium adduct, 305 = hydrogen adduct of water loss ion). Metabolite 1.3.16 was tentavely identified as a unique dihydroxy manool variant. 1.3.17 and 1.3.11 correspond to monohydroxylated manool derivatives as evidenced by the identical mass spectra. However, 1.3.17 and 1.3.13 are indeed different metabolites as evidenced by the shift in retention time.

The yeast strain IT7-T9 expressing CfTPS1, SsSCS, CfCYP76AH8, PsCYP720B4 produces oxidized manool as shown in figure 19. Extracted ion chromatograms recorded by LC-MS analysis of ethanol extracts generated from media as well as cell material of IT7-T9 shows that the extracted ion traces combine the ions, 299 and 339 m/z, corresponding to an unknown derivative, observed only in the strain expressing CfCYP76AH8 and PsCYP720B4. Based on the fragmentation pattern, the metabolite 1.3.18 could potentially represent a dioxidized manool derivative with three additional degrees of desaturation, i.e. double bonds.

Claims

1. A method of producing a diterpenoid, said method comprising the steps of:

a) providing one or more host organisms, wherein
   I. at least one of the host organisms comprises a heterologous nucleic acid encoding a diTPS of class II,
   II. at least one of the host organisms comprises a heterologous nucleic acid encoding a diTPS of class I,
   III. at least one of the host organisms comprises a heterologous nucleic acid encoding a cytochrome P450 (CYP)

   with the proviso that said diTPS of class II, said diTPS of class I and said CYP are not all from the same species;

b) Incubating said one or more host organisms either simultaneously or sequentially under conditions allowing growth of said host organisms, wherein at least the host organism comprising a heterologous nucleic acid encoding a diTPS of class II is incubated in the presence of geranylgeranyl pyrophosphate (GGPP);

c) Optionally isolating diterpenoid.

2. The method according to claim 1, wherein step a) comprises providing one host organism comprising:
   I. a heterologous nucleic acid encoding a diTPS of class II,
   II. a heterologous nucleic acid encoding a diTPS of class I,
   III. a heterologous nucleic acid encoding a cytochrome P450 (CYP)

3. The method according to claim 1, wherein step a) comprises providing one host organism comprising:
   I. a heterologous nucleic acid encoding a diTPS of class II,
   II. a heterologous nucleic acid encoding a diTPS of class I,
   III. at least two heterologous nucleic acids each encoding a different cytochrome P450 (CYP)

4. The method according to claim 1, wherein step a) comprises providing two host organism, wherein the first host organism comprises:
   I. a heterologous nucleic acid encoding a diTPS of class II,
   II. a heterologous nucleic acid encoding a diTPS of class I,

   and the second host organism comprises:
   III. a heterologous nucleic acid encoding a cytochrome P450 (CYP)

5. The method according to claim 1, wherein step a) comprises providing two host organism, wherein the first host organism comprises:
   IV. a heterologous nucleic acid encoding a diTPS of class II,
   V. a heterologous nucleic acid encoding a diTPS of class I,

   and the second host organism comprises
   VI. at least two heterologous nucleic acids each encoding a different cytochrome P450 (CYP)

6. The method according to any one of claims 4 to 5, wherein step b) comprises incubating said two host organisms simultaneously in cultivation medium under conditions allowing growth of said host organisms in the presence of geranylgeranyl pyrophosphate.

7. The method according to any one of claims 4 to 5, wherein step b) comprises the steps of

   b1) incubating the first host organism under conditions allowing growth of said host organism in the presence of geranylgeranyl pyrophosphate, thereby producing a diterpene

   b2) incubating the second host organism under conditions allowing growth of said host organism in the presence of the diterpene produced in step b1)

8. The method according to claim 1, wherein step a) comprises providing three or more host organisms, wherein the first host organism comprises:
   I. a heterologous nucleic acid encoding a diTPS of class II,
   II. a heterologous nucleic acid encoding a diTPS of class I,

   and the second host organism comprises
   III. a heterologous nucleic acids encoding a cytochrome P450 (CYP)

   and the third and optionally further host organisms each comprises

IV. a heterologous nucleic acids encoding a cytochrome P450 (CYP) different to the CYP under III .

9. The method according to claim 8, wherein step b) comprises incubating said three or more host organisms simultaneously in cultivation medium under conditions allowing growth of said host organisms in the presence of geranylgeranyl pyrophosphate.

10. The method according to claim 8, wherein step b) comprises the steps of

b1) incubating the first host organism under conditions allowing growth of said host organism in the presence of geranylgeranyl pyrophosphate, thereby producing a diterpene

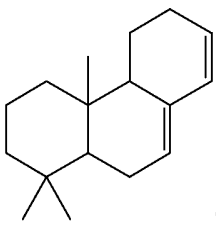
b2) incubating the second host organism under conditions allowing growth of said host organism in the presence of the diterpene produced in step b1), thereby producing an intermediate diterpenoid

b3) incubating the third host organism under conditions allowing growth of said host organism in the presence of intermediate diterpenoid produced in step b2), thereby producing a diterpenoid.

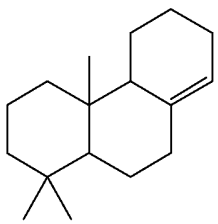
11. The method according to any one of the preceding claims, wherein the diterpenoid is a C_{20}-molecule substituted at one or more positions with substituent(s) selected from the group consisting of –OH, =O and –COOH.

12. The method according to any one of the preceding claims, wherein the diterpenoid is a C_{20}-molecule containing a decalin core and one or more substitutents selected from the group consisting of –OH, =O and/or –COOH.

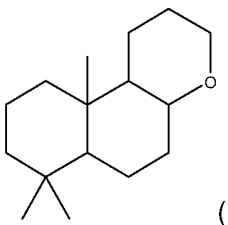
13. The method according to any one of the preceding claims, wherein the diterpenoid is a C_{20}-molecule containing a core structure of one of following formulas I, II, III, IV, V, VI, IX, X, XXII, XXIII or XXIV.



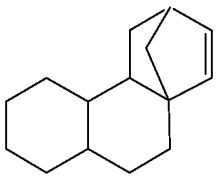
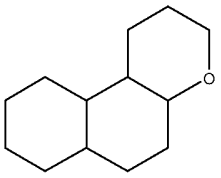
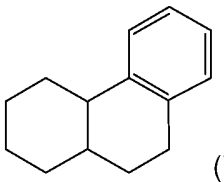
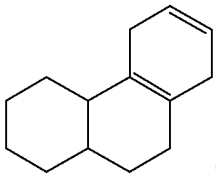
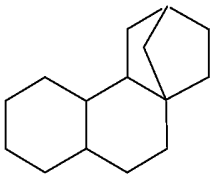
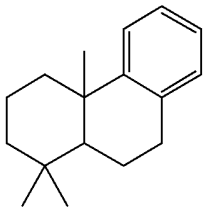
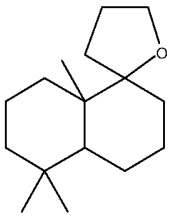
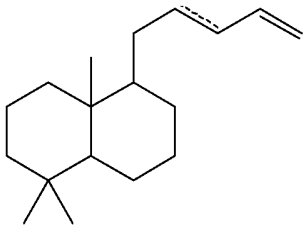
(I) or



(II) or



(III) or


(IV) or

(V) or

(VI) or

(IX) or

(X) or

(XXII) or

(XXIII) or

(XXIV).

14. The method according to claim 13, wherein diterpenoid is a C_{20}-molecule containing a cores structure of formulas I, II, III, IV, V, VI, IX, X, XXII, XXIII or XXIV substituted at one or more positions by one or more groups selected from the group consisting of:
    a) alkyl, such as C_{1-6}-alkyl, for example C_{1-3}, wherein said alkyl may be linear or branched, for example alkyl may be isopropyl or methyl, wherein said alkyl may be substituted with one or more groups selected from the group consisting of –OH, =O and –COOH;
    b) alkenyl, such as C_{1-6} alkenyl, such as C_{2-4}-alkenyl, such as C_{2-3}-alkenyl, wherein said alkyl may be substituted with one or more groups selected from the group consisting of –OH, =O and –COOH;
    c) –OH;
    d) =O; and
    e) –COOH.

15. The method according to any one of claims 1 to 10, wherein the diterpenoid is a diterpene, which has been substituted at one or more positions with substituent(s) selected from the group consisting of –OH, =O and –COOH, wherein said diterpene is the product of any of the reactions VII to XIX described herein above.
16. The method according to any one of claims 1 to 10, wherein the diterpenoid is a diterpene, which has been substituted at one or more positions with substituent(s)

selected from the group consisting of –OH, =O and –COOH, wherein said diterpene is the product of any of the reactions XX, XXII, XXIII, XXIV, XXV, XXVII, XXIX, LII, LIII, LIV, LV or LVI described herein above.

17. The method according to any one of claims 1 to 10, wherein the diterpenoid is a diterpene, which has been substituted at one or more positions with substituent(s) selected from the group consisting of –OH, =O and –COOH, wherein the diterpene is any of the compounds 1 to 47 of Table 1.

18. The method according to any one of claims 1 to 10, wherein the diterpenoid is a diterpene substituted at one or more positions with –OH, =O and/or –COOH, wherein the diterpene is a C_{20}-molecule consisting of 20 carbon atoms, up to three oxygen atoms and hydrogen atoms, and which contains a core structure of any of formulas I, II, III, IV, VI, IX, X, XII, XLII, XLIII, XLIV, XLV, XLVI, XXVII, XXVIII, XXIX, XXX, XXXI, XXXII, XXXIII, XXXIV, XXXV, XXXVI, XXXVII, XXXVIII, XXXIX, XL XLI and/or XLVI.

19. The method according to any one of claims 1 to 10, wherein the diterpenoid is dehydroabietadiene substituted at one or more positions with substituent(s) selected from the group consisting of –OH, =O and –COOH.

20. The method according to any one of claims 1 to 10, wherein the diterpenoid is miltiradiene substituted at one or more positions with substituent(s) selected from the group consisting of –OH, =O and –COOH.

21. The method according to any one of claims 1 to 10, wherein the diterpenoid is manool substituted at one or more positions with substituent(s) selected from the group consisting of –OH, =O and –COOH.

22. The method according to any one of claims 1 to 10, wherein the diterpenoid is syn-manool substituted at one or more positions with substituent(s) selected from the group consisting of –OH, =O and –COOH.

23. The method according to any one of claims 1 to 10, wherein the diterpenoid is syn-pimara-9,(11),15-diene substituted at one or more positions with substituent(s) selected from the group consisting of –OH, =O and –COOH.

24. The method according to any one of claims 1 to 10, wherein the diterpenoid is syn-isopimara-7,15-diene substituted at one or more positions with substituent(s) selected from the group consisting of –OH, =O and –COOH.

25. The method according to any one of claims 1 to 10, wherein the diterpenoid is ferruginol.

26. A host organism comprising

I. a heterologous nucleic acid encoding a diTPS of class II,

II. a heterologous nucleic acid encoding a diTPS of class I,

III. a heterologous nucleic acid encoding a cytochrome P450 (CYP)

with the proviso that said diTPS of class II, said diTPS of class I and said CYP are not all from the same species.

27. The host organism according to claim 26, wherein the host organism comprises at least the following heterologous nucleic acids:

a) a heterologous nucleic acid encoding CfTPS1 of SEQ ID NO:5 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith; and

b) a heterologous nucleic acid encoding CfTPS3 of SEQ ID NO:12 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith;

c) a heterologous nucleic acid encoding CYP76AH11 of SEQ ID NO:21 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith.

28. The host organism according to claim 26, wherein the host organism comprises at least the following heterologous nucleic acids:

a) a heterologous nucleic acid encoding TwTPS7 of SEQ ID NO:4 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith; and

b) a heterologous nucleic acid encoding CfTPS3 of SEQ ID NO:12 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith;

c) a heterologous nucleic acid encoding CYP76AH11 of SEQ ID NO:21 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith.

29. The host organism according to claim 26, wherein the host organism comprises at least the following heterologous nucleic acids:

a) a heterologous nucleic acid encoding CfTPS1 of SEQ ID NO:5 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith; and

b) a heterologous nucleic acid encoding CfTPS3 of SEQ ID NO:12 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith;

c) a heterologous nucleic acid encoding CYP71D381 of SEQ ID NO:22 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith.

30. The host organism according to claim 26, wherein the host organism comprises at least the following heterologous nucleic acids:

a) a heterologous nucleic acid encoding TwTPS7 of SEQ ID NO:4 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith; and

b) a heterologous nucleic acid encoding CfTPS3 of SEQ ID NO:12 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith;

c) a heterologous nucleic acid encoding CYP71D381 of SEQ ID NO:22 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith.

31. The host organism according to claim 26, wherein the host organism comprises at least the following heterologous nucleic acids:

a) a heterologous nucleic acid encoding CfTPS1 of SEQ ID NO:5 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith; and

b) a heterologous nucleic acid encoding CfTPS3 of SEQ ID NO:12 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith;

c) a heterologous nucleic acid encoding CYP720B4 of SEQ ID NO:23 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith.

32. The host organism according to claim 26, wherein the host organism comprises at least the following heterologous nucleic acids:

a) a heterologous nucleic acid encoding Ossyn-CPP of SEQ ID NO:1 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith; and

b) a heterologous nucleic acid encoding CfTPS3 of SEQ ID NO:12 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith;

c) a heterologous nucleic acid encoding PsCYP720B4 of SEQ ID NO:23 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith.

33. The host organism according to claim 26, wherein the host organism comprises at least the following heterologous nucleic acids:
a) a heterologous nucleic acid encoding CfTPS1 of SEQ ID NO:5 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith; and
b) a heterologous nucleic acid encoding SsSCS of SEQ ID NO:11 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith;
c) a heterologous nucleic acid encoding CYP76AH8 of SEQ ID NO:20 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith.

34. The host organism according to claim 26, wherein the host organism comprises at least the following heterologous nucleic acids:
a) a heterologous nucleic acid encoding CfTPS1 of SEQ ID NO:5 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith; and
b) a heterologous nucleic acid encoding SsSCS of SEQ ID NO:11 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith;
c) a heterologous nucleic acid encoding CYP76AH11 of SEQ ID NO:21 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith.

35. The host organism according to claim 26, wherein the host organism comprises at least the following heterologous nucleic acids:
    a) a heterologous nucleic acid encoding CfTPS1 of SEQ ID NO:5 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith; and
    b) a heterologous nucleic acid encoding SsSCS of SEQ ID NO:11 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith;
    c) a heterologous nucleic acid encoding CYP720B4 of SEQ ID NO:23 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith.

36. The host organism according to claim 26, wherein the host organism comprises at least the following heterologous nucleic acids:
    a) a heterologous nucleic acid encoding Ossyn-CPP of SEQ ID NO:1 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith; and
    b) a heterologous nucleic acid encoding SsSCS of SEQ ID NO:11 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith;
    c) a heterologous nucleic acid encoding CYP71D381 of SEQ ID NO:22 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith.

37. The host organism according to claim 26, wherein the host organism comprises at least the following heterologous nucleic acids:
    a) a heterologous nucleic acid encoding CfTPS1 of SEQ ID NO:5 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least

85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith; and

b) a heterologous nucleic acid encoding SsSCS of SEQ ID NO:11 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith;

c) a heterologous nucleic acid encoding CYP76AH8 of SEQ ID NO:20 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith; and

d) a heterologous nucleic acid encoding CYP71D381 of SEQ ID NO:22 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith.

38. The host organism according to claim 26, wherein the host organism comprises at least the following heterologous nucleic acids:

a) a heterologous nucleic acid encoding CfTPS1 of SEQ ID NO:5 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith; and

b) a heterologous nucleic acid encoding SsSCS of SEQ ID NO:11 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith;

c) a heterologous nucleic acid encoding CYP76AH8 of SEQ ID NO:20 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith; and

d) a heterologous nucleic acid encoding CYP76AH11 of SEQ ID NO:21 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith.

39. The host organism according to claim 26, wherein the host organism comprises at least the following heterologous nucleic acids:

a) a heterologous nucleic acid encoding CfTPS1 of SEQ ID NO:5 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith; and

b) a heterologous nucleic acid encoding SsSCS of SEQ ID NO:11 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith;

c) a heterologous nucleic acid encoding CYP76AH8 of SEQ ID NO:20 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith; and

d) a heterologous nucleic acid encoding CYP720B4 of SEQ ID NO:23 or a functional homologue thereof sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% sequence identity therewith.

40. The method according to any one of claims 1 to 25, wherein the host organism is a host organism according to any one of claims 26 to 39.

41. A collection of host organisms comprising at least two different host organisms, wherein

I. at least one of the host organisms comprises a heterologous nucleic acid encoding a diTPS of class II,

II. at least one of the host organisms comprises a heterologous nucleic acid encoding a diTPS of class I,

III. at least one of the host organisms comprises a heterologous nucleic acid encoding a cytochrome P450 (CYP)

with the proviso that said diTPS of class II, said diTPS of class I and said CYP are not all from the same species.

42. The collection of host organisms according to claim 41, said collection comprising two host organisms, wherein the first host organism comprises:

I. a heterologous nucleic acid encoding a diTPS of class II,

II. a heterologous nucleic acid encoding a diTPS of class I,

and the second host organism comprises:

III. a heterologous nucleic acid encoding a cytochrome P450 (CYP)

43. The collection of host organisms according to claim 41, said collection comprising two host organism, wherein the first host organism comprises:

I. a heterologous nucleic acid encoding a diTPS of class II,

II. a heterologous nucleic acid encoding a diTPS of class I,

and the second host organism comprises

III. at least two heterologous nucleic acids each encoding a different cytochrome P450 (CYP)

44. The collection of host organisms according to claim 41, said collection comprising three or more host organisms, wherein the first host organism comprises:

I. a heterologous nucleic acid encoding a diTPS of class II,

II. a heterologous nucleic acid encoding a diTPS of class I,

and the second host organism comprises

III. a heterologous nucleic acids encoding a cytochrome P450 (CYP)

and the third and optionally further host organisms each comprises

IV. a heterologous nucleic acids encoding a cytochrome P450 (CYP) different to the CYP under III.

45. The method according to any one of claims 1 to 25, the host organism according to claim 26 or the collection of host organism according to any one of claims 40 to 44, wherein the diTPS of class II is a polypeptide sharing at least 30%, such as at least 35% sequence identity to the sequence of SsLPPS (SEQ ID NO:6) or to the sequence of AtCPS (SEQ ID NO:18).

46. The method, the host organism or the collection of host organisms according to claim 45, wherein the said diTPS of class II contains the following motif of four amino acids:

D/E-X-D-D, wherein X may be any amino acid .

47. The method according to any one of claims 1 to 25, the host organism according to claim 26 or the collection of host organism according to any one of claims 40 to 44,

wherein the diTPS of class II is selected from the group consisting of syn-CPP type diTPS, ent-CPP type diTPS, (+)-CPP type diTPS, LPP type diTPS and LPP like type diTPS.

48. The method according to any one of claims 1 to 25, the host organism according to claim 26 or the collection of host organism according to any one of claims 40 to 44, wherein the diTPS of class I is a polypeptide sharing at least 30%, such as at least 35% sequence identity to the sequence of ScSCS (SEQ ID NO:11) or to the sequence of AtEKS (SEQ ID NO: 20).

49. The method, the host organism or the collection of host organisms according to claim 48, wherein the said diTPS of class I contains the following motif of five amino acids:

D-D-X-X-D/E, wherein X indicates any amino acids.

50. The method according to any one of claims 1 to 25, the host organism according to claim 26 or the collection of host organism according to any one of claims 40 to 44, wherein the diTPS of class I is selected from the group consisting of EpTPS8 like diTPS, EpTPS23 like diTPS, SsSCS like diTPS, CfTPS3 like diTPS, CfTPS4 like diTPS, TwTPS2 like diTPS, EpTPS1 like diTPS and CfTPS14 like diTPS.

51. The method according to any one of claims 1 to 25, or the host organism according to claim 26 or the collection of host organisms according to any one of claims 40 to 44, wherein the diTPS of class II is a polypeptide selected from the group consisting of:

a) EpTPS7 of SEQ ID NO:2
b) TwTPS7 of SEQ ID NO:4;
c) CfTPS1 of SEQ ID NO:5;
d) TwTPS21 of SEQ ID NO:7;
e) TwTPS14/28 of SEQ ID NO:8;
f) syn-CPP of SEQ ID NO:1;
g) ZmAN2 of SEQ ID NO:3;
h) SsLPPS of SEQ ID NO:6;
i) CfTPS2 of SEQ ID NO:17;and

j) functional homologues of the polypeptides of a) to i) sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95% sequence identity therewith.

52. The method according to any one of claims 1 to 25, the host organism according to claim 26 or the collection of host organism according to any one of claims 40 to 44, wherein the diTPS of class II is an enzyme capable of catalysing any of the reactions I to V.

53. The method according to any one of claims 1 to 25, the host organism according to claim 26 or the collection of host organism according to any one of claims 40 to 44, wherein the diTPS of class II is an enzyme capable of catalysing any of the reactions shown in figures 3A or 3B.

54. The method according to any one of claims 1 to 25, the host organism according to claim 26 or the collection of host organisms according to any one of claims 40 to 44, wherein the diTPS of class I is a polypeptide selected from the group consisting of:

a) EpTPS8 of SEQ ID NO:9;
b) EpTPS23 of SEQ ID NO:10;
c) TwTPS2 of SEQ ID NO:14;
d) EpTPS1 of SEQ ID NO:15;
e) CfTPS14 of SEQ ID NO:16;
f) SsSCS of SEQ ID NO:11;
g) CfTPS12 of SEQ ID NO:12;
h) CfTPS4 of SEQ ID NO:13;
i) MvTPS5 of SEQ ID NO:r46;
j) functional homologues of the polypeptides of a) to i) sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95% sequence identity therewith.

55. The method according to any one of claims 1 to 25, or the host organism according to claim 26 or the collection of host organisms according to any one of claims 40 to 44, wherein the diTPS of class I is an enzyme capable of catalysing any of the reactions VII to XIX.

56. The method according to any one of claims 1 to 25, or the host organism according to claim 26 or the collection of host organisms according to any one of claims 40 to 44, wherein the diTPS of class I is an enzyme capable of catalyzing at least one of the reactions LII, LIII, LIV, LV, LVI and X.

57. The method, the host organism or the collection of host organisms according to any one of the preceding claims, wherein the CYP is an enzyme capable of catalyzing one or more of the following reactions:
Hydroxylation
Oxidation leading to a formation of a carbonyl groups
Oxidation leading to formation of a carboxyl group.

58. The method, the host organism or the collection of host organisms according to any one of claims 1 to 57, wherein the CYP is a polypeptide comprising the following motif of 5 amino acids:

A/G-G-X-X-T/S, wherein X indicates any amino acids.

59. The method, the host organism or the collection of host organisms according to any one of claims 1 to 58, wherein the CYP is a polypeptide comprising the following motif of 4 amino acids:

E-X-X-R.

60. The method, the host organism or the collection of host organisms according to any one of claims 1 to 59, wherein the CYP is a polypeptide comprising the following motif of 10 amino acids:

F-X-X-G-X-X-X-C-X-G.

61. The method, the host organism or the collection of host organisms according to any one of claims 1 to 60, wherein the CYP is a polypeptide comprising the following motif of 3 amino acids:

P-F-G.

62. The method, the host organism or the collection of host organisms according to any one of the preceding claims, wherein the CYP is an enzyme capable of catalysing hydroxylation of a diterpene to form hydroxyl-diterpene.

63. The method, the host organism or the collection of host organisms according to any one of the preceding claims, wherein the CYP is an enzyme capable of catalysing hydroxylation of a hydroxyl-diterpene to form dihydroxy-diterpene and/or diterpene ketone and/or diterpene aldehyde.

64. The method, the host organism or the collection of host organisms according to any one of the preceding claims, wherein the CYP is an enzyme capable of catalysing hydroxylation of a dihydroxy-diterpene to form trihydroxy-diterpene and/or diterpene carboxylic acid.

65. The method, the host organism or the collection of host organisms according to any one of the preceding claims, wherein the CYP is an enzyme capable of catalysing hydroxylation of a diterpene ketone to form hydroxyl diterpene ketone or diterpene carboxylic acid.

66. The method, the host organism or the collection of host organisms according to any one of the preceding claims, wherein the CYP is an enzyme capable of catalysing hydroxylation of a diterpene carboxylic acid to form hydroxy-diterpene carboxylic acid.

67. The method, the host organism or the collection of host organisms according to any one of the claims 62 to 66, wherein the diterpene is a C_{20}-molecule consisting of 20 carbon atoms, up to three oxygen atoms and hydrogen atoms containing a core structure of one of following formulas I, II, III, IV, V, VI, IX, X, XXII, XXIII or XXIV.

68. The method, the host organism or the collection of host organisms according to any one of the claims 62 to 66, wherein the diterpene is a be a C_{20}-molecule consisting of 20 carbon atoms, up to three oxygen atoms and hydrogen atoms, and which contains a core structure of any of formulas I, II, III, IV, VI, IX, X, XII, XLII, XLIII,

XLIV, XLV, XLVI, XXVII, XXVIII, XXIX, XXX, XXXI, XXXII, XXXIII, XXXIV, XXXV, XXXVI, XXXVII, XXXVIII, XXXIX, XL, XLI and/or XLVI.

69. The method, the host organism or the collection of host organisms according to any one of the claims 62 to 66, wherein diterpene is selected from the group of diterpene listed in Table 1.

70. The method, the host organism or the collection of host organisms according to any one of claims 1 to 69, wherein the CYP is an enzyme capable of catalysing hydroxylation of dehydroabietadiene.

71. The method, the host organism or the collection of host organisms according to any one of claims 1 to 69, wherein the CYP is an enzyme capable of catalysing hydroxylation of miltiradiene.

72. The method, the host organism or the collection of host organisms according to any one of claims 1 to 69, wherein the CYP is an enzyme capable of catalysing hydroxylation of manool to form hydroxyl manool.

73. The method, the host organism or the collection of host organisms according to any one of claims 1 to 69, wherein the CYP is an enzyme capable of catalysing hydroxylation of hydroxyl manool, to form dihydoxy manool.

74. The method, the host organism or the collection of host organisms according to any one of claims 1 to 69, wherein the CYP is an enzyme capable of catalysing hydroxylation of 1,3 manool.

75. The method, the host organism or the collection of host organisms according to any one of claims 1 to 69, wherein the CYP is an enzyme capable of catalysing carboxylation of a diterpene.

76. The method, the host organism or the collection of host organisms according to any one of claims 1 to 75, wherein the CYP is a CYP of the family 76.

77. The method, the host organism or the collection of host organisms according to claim 76, wherein the CYP is a polypeptide sharing at least 40% sequence identity with the polypeptide of SEQ ID NO:20 and/or the polypeptide of SEQ ID NO:21.

78. The method, the host organism or the collection of host organisms according to claim 76, wherein the CYP is a polypeptide sharing at least 40% sequence identity with the polypeptide of SEQ ID NO:24.

79. The method, the host organism or the collection of host organisms according to any one of claims 76 to 78, wherein the CYP is an enzyme capable of catalysing at least one of reactions XXXIX, XL, XLI, XLV, XLVI, XLVII or LI.

80. The method, the host organism or the collection of host organisms according to any one of claims 1 to 75, wherein the CYP is a CYP of the family 71.

81. The method, the host organism or the collection of host organisms according to claim 80, wherein the CYP is a polypeptide sharing at least 40% sequence identity with the polypeptide of SEQ ID NO:22.

82. The method, the host organism or the collection of host organisms according to any one of claims 80 to 81, wherein the CYP is an enzyme capable of catalysing at least one of reactions XXXIX, XLII orLI.

83. The method, the host organism or the collection of host organisms according to any one of claims 1 to 75 wherein the CYP is a CYP of the family 720.

84. The method, the host organism or the collection of host organisms according to claim 83, wherein the CYP is a polypeptide sharing at least 40% sequence identity with the polypeptide of SEQ ID NO:23.

85. The method, the host organism or the collection of host organisms according to any one of claims 83 to 84, wherein the CYP is an enzyme capable of catalysing at least one of reactions XLI, XLIII, XLIV, XLVIII, XLIX, L, LI, or LII.

86. The method, the host organism or the collection of host organisms according to any one of claims 1 to 85, wherein the CYP is a polypeptide selected from the group consisting of:
    a) CYP76AH8 of SEQ ID NO:20;
    b) CYP76AH11 of SEQ ID NO:21;
    c) CYP71D381 of SEQ ID NO:22;
    d) CYP720B4 of SEQ ID NO:23;
    e) CYP76AH1 of SEQ ID NO:42;
    f) CYP76AH4 of SEQ ID NO:43;
    g) CYP76AH15 of SEQ ID NO:40;
    h) CYP76AH17 of SEQ ID NO:41;
    and
    i) functional homologues of the polypeptides of a) to h) sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95% sequence identity therewith.

87. The method according to any one of the preceding claims, wherein step b) involves incubating the host organism in a cultivation medium.

88. The method, the host organism or the collection of host organisms according to any one of the preceding claims, wherein the host organism further comprises one or more heterologous nucleic acids encoding enzymes involved in the biosynthesis of GGPP.

89. The method, the host organism or the collection of host organisms according to any one of the preceding claims, wherein said enzymes is selected from the group consisting of CfDXS of SEQ ID NO:25, CfGGPPS of SEQ ID NO:26 and functional homolgoues of any of the aforementioned sharing at least 70% sequence identity therewith.

90. The method, the host organism or the collection of host organisms according to any one of the preceding claims, wherein host organism further comprises a heterologous nucleic acid encoding a cytochrome P450 reductase (POR).

91. The method, the host organism or the collection of host organisms according to claim 90, wherein the POR is selected from the group consisting of:
    a. ATR1 from A thaliana encoded by Gene ID:3150037
    b. POR from *Stevia rebaudiana* encoded by Gene ID: 93211213
    c. CfCPR of SEQ ID NO:44
    d. functional homologues of the polypeptides of a) to c) sharing at least 70%, such as at least 80%, such as at least 85%, such as at least 90%, such as at least 95% sequence identity therewith.

92. The method, the host organism or the collection of host organisms according to any one of the preceding claims, wherein the host organism is a microorganism.

93. The method, the host organism or the collection of host organisms according to claim 92, wherein the microorganism is yeast.

94. The method according to any one of the preceding claims, wherein the host organism is a microorganism, and step b) involves incubating said microorganism in a cultivation medium.

95. The method according to claim 940, wherein the method comprises a step of isolating said diterpenoid from the microorganism and/or the cultivation medium.

96. The method, the host organism or the collection of host organisms according to any one of claims 1 to 91, wherein the host organisms is a plant.

97. The method according to claim 96, wherein the method comprises isolating the diterpenoid from said plant or parts thereof, for example from leaves, fruits, flowers, stems, seeds and/or roots of plants.

98. A method of producing a diterpenoid, said method comprising the steps of
    a) providing a host organism or a collection of host organism according to any one of claims 26 to 44;
    b) preparing an extract of said host organism(s) or preparing individual extracts or each host organism;
    c) providing GGPP

d) incubating one or more of said extracts with GGPP, wherein at least the extract of the host organism comprising a heterologous nucleic acid encoding a diTPS of class II is incubated in the presence of geranylgeranyl pyrophosphate (GGPP);

e) optionally, incubating the product of the incubation under d) with one or more of said extracts,

f) optionally, incubating the product of the incubation under e) with one or more of said extracts, wherein step f) may optionally be repeated;

thereby producing a diterpeneoid.

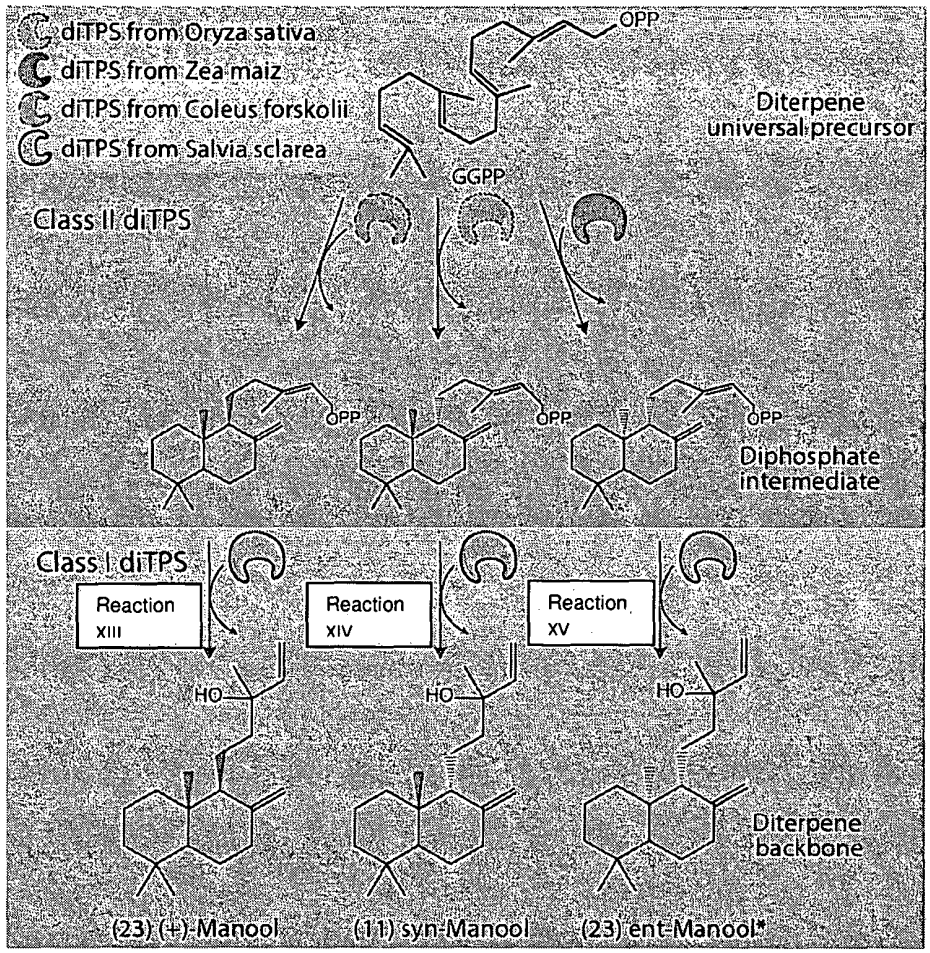

diTPS from Oryza sativa
diTPS from Zea maiz
diTPS from Coleus forskolii
diTPS from Salvia sclarea
Class II diTPS
OPP
GGPP
Diterpene universal precursor
OPP
OPP
OPP
Diphosphate intermediate
Class I diTPS
Reaction XIII
Reaction XIV
Reaction XV
HO
HO
HO
Diterpene backbone
(23) (+)-Manool
(11) syn-Manool
(23) ent-Manool*

Fig. 1

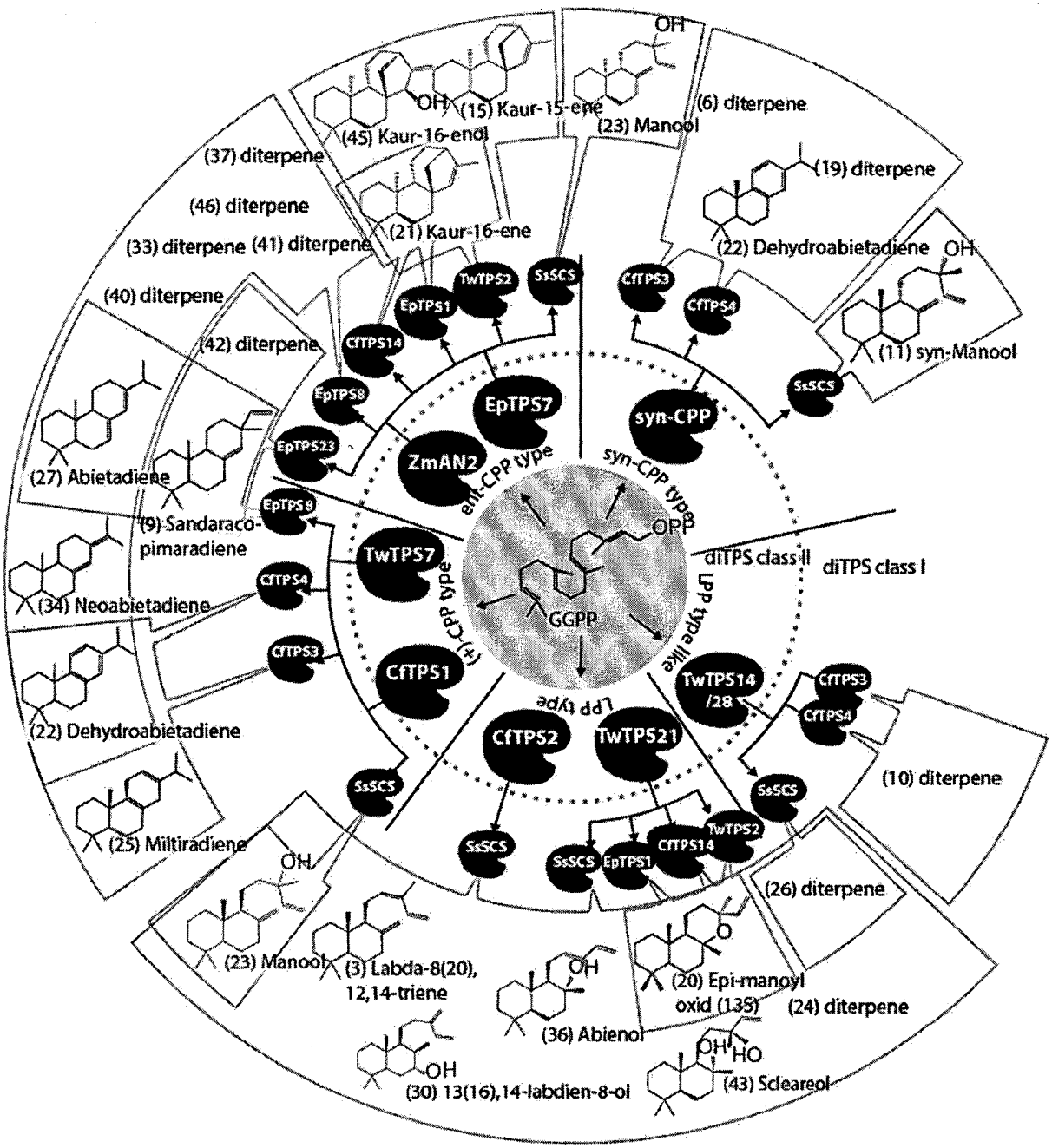

OH
(15) Kaur-15-ene
(45) Kaur-16-enol
(23) Manool
(6) diterpene
(37) diterpene
(46) diterpene
(33) diterpene
(41) diterpene
(21) Kaur-16-ene
(19) diterpene
(22) Dehydroabietadiene
OH
(40) diterpene
(42) diterpene
(11) syn-Manool
(27) Abietadiene
(9) Sandaraco-pimaradiene
(34) Neoabietadiene
(22) Dehydroabietadiene
(25) Miltiradiene
(10) diterpene
(26) diterpene
(23) Manool
(3) Labda-8(20), 12,14-triene
(36) Abienol
(20) Epi-manoyl oxid (135)
(24) diterpene
(30) 13(16),14-labdien-8-ol
(43) Sclareol
TwTPS2
SsSCS
EpTPS1
CfTPS14
EpTPS8
EpTPS7
EpTPS23
ZmAN2
ent-CPP type
syn-CPP type
syn-CPP
CfTPS3
CfTPS4
SsSCS
EpTPS8
TwTPS7
CfTPS4
CfTPS3
CfTPS1
(+)-CPP type
OPP
GGPP
diTPS class II
diTPS class I
LPP type like
LPP type
TwTPS14 /28
CfTPS3
CfTPS4
CfTPS2
TwTPS21
SsSCS
SsSCS
SsSCS
EpTPS1
CfTPS14
TwTPS2

Fig. 2A

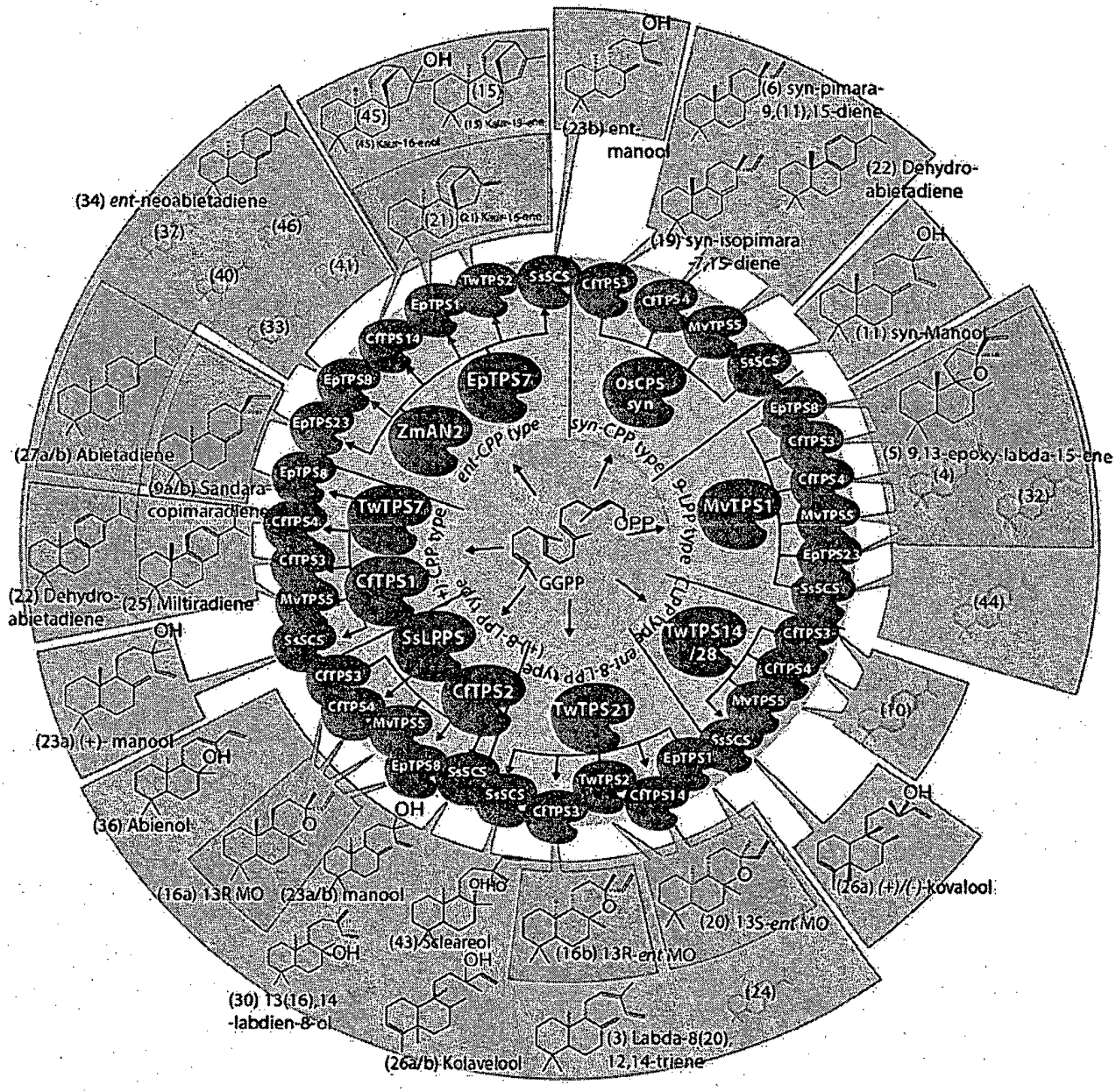

(6) syn-pimara-9,(11),15-diene
(23b) ent-manool
(22) Dehydro-abietadiene
(19) syn-isopimara-7,15-diene
(11) syn-Manool
(5) 9,13-epoxy-labda-15-ene
(4)
(32)
(44)
(10)
(26a) (+)/(-)-kovalool
(20) 13S-ent MO
(16b) 13R-ent MO
(24)
(3) Labda-8(20),12,14-triene
(26a/b) Kolavelool
(30) 13(16),14-labdien-8-ol
(43) Sclareol
(16a) 13R MO
(23a/b) manool
(36) Abienol
(23a) (+)- manool
(22) Dehydro-abietadiene
(25) Miltiradiene
(9a/b) Sandaracopimaradiene
(27a/b) Abietadiene
(34) ent-neoabietadiene
(37)
(46)
(40)
(41)
(33)
(45) Kaur-16-enol
(15) Kaur-15-ene
(21) Kaur-15-ene
OH
TwTPS2
SsSCS
CfTPS3
CfTPS4
MvTPS5
EpTPS1
CfTPS14
EpTPS8
EpTPS7
OsCPS4 syn
EpTPS23
ZmAN2
ent-CPP type
syn-CPP type
EpTPS8
TwTPS7
CfTPS4
CfTPS3
CfTPS1
MvTPS5
SsSCS
SsLPPS
CfTPS2
CfTPS3
CfTPS4
MvTPS5
EpTPS8
SsSCS
SsSCS
CfTPS3
OPP
GGPP
MvTPS1
CPP type
8-LPP type
ent-8-LPP type
(+)-8-LPP type
TwTPS14/28
TwTPS21
EpTPS8
CfTPS3
EpTPS23
SsSCS
CfTPS3
CfTPS4
MvTPS5
SsSCS
EpTPS1
TwTPS2
CfTPS14

Fig. 2B

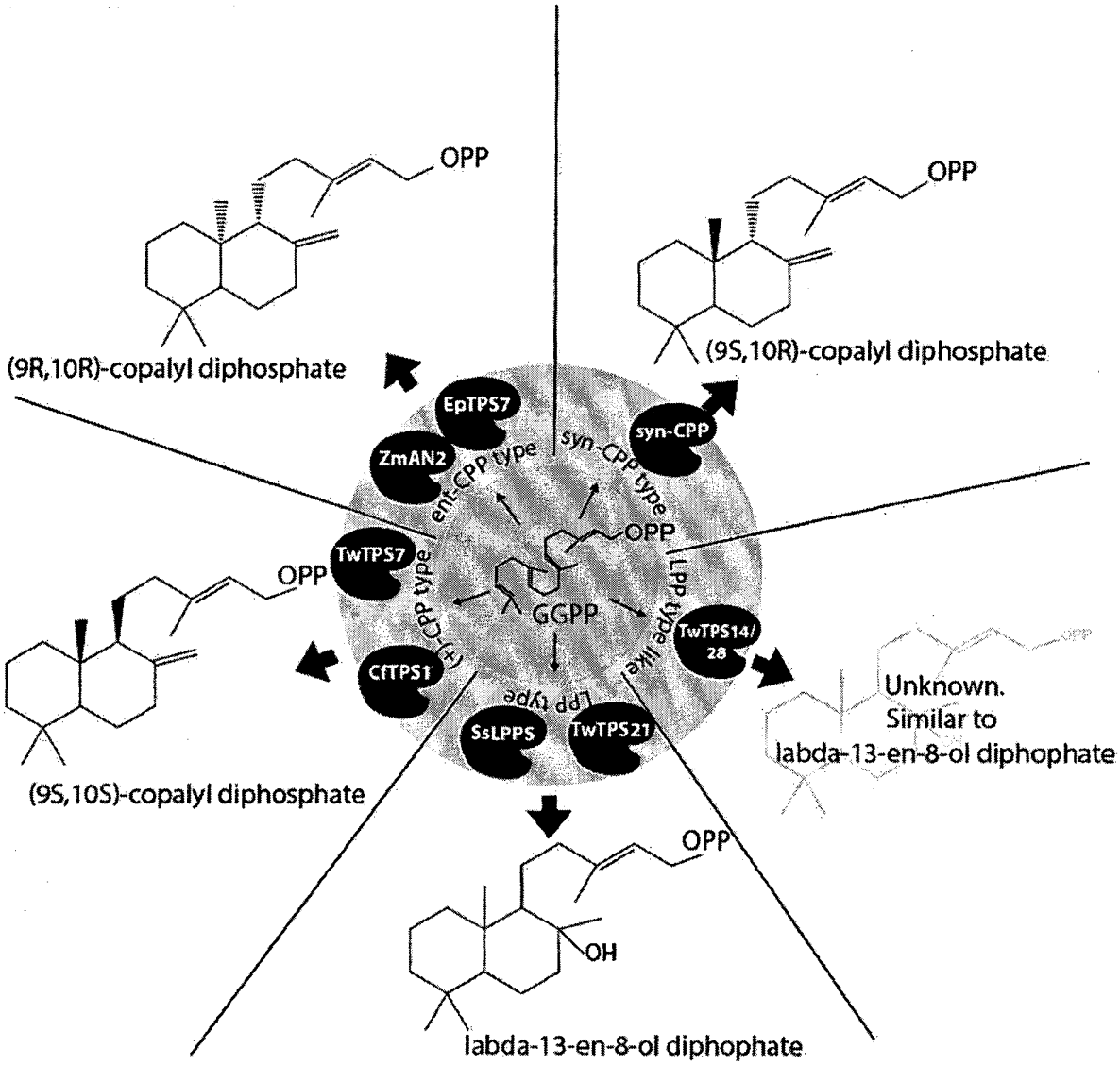

OPP
(9R,10R)-copalyl diphosphate
OPP
(9S,10R)-copalyl diphosphate
EpTPS7
ZmAN2
ent-CPP type
syn-CPP
syn-CPP type
OPP
TwTPS7
OPP
(-)-CPP type
GGPP
LPP type like
TwTPS14/28
CfTPS1
Unknown.
Similar to
labda-13-en-8-ol diphophate
LPP type
SsLPPS
TwTPS21
(9S,10S)-copalyl diphosphate
OPP
OH
labda-13-en-8-ol diphophate

Fig. 3A

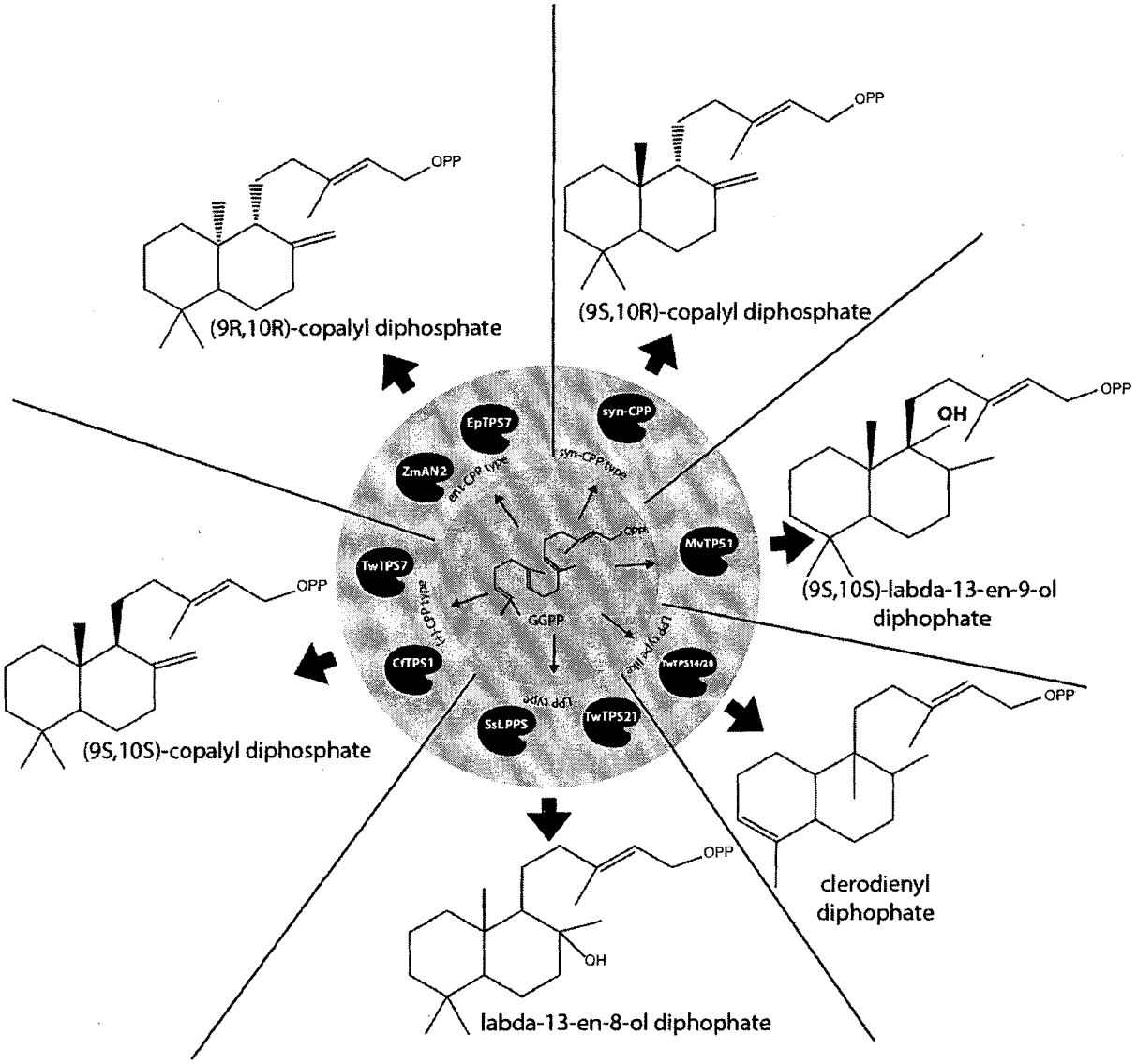

OPP
(9R,10R)-copalyl diphosphate
(9S,10R)-copalyl diphosphate
OH
(9S,10S)-labda-13-en-9-ol diphophate
(9S,10S)-copalyl diphosphate
clerodienyl diphophate
labda-13-en-8-ol diphophate
EpTPS7
syn-CPP
ZmAN2
syn-CPP type
ent-CPP type
MvTPS1
TwTPS7
(+)-CPP type
GGPP
CfTPS1
LPP type
TwTPS14/20
SsLPPS
TwTPS21

Fig. 3B

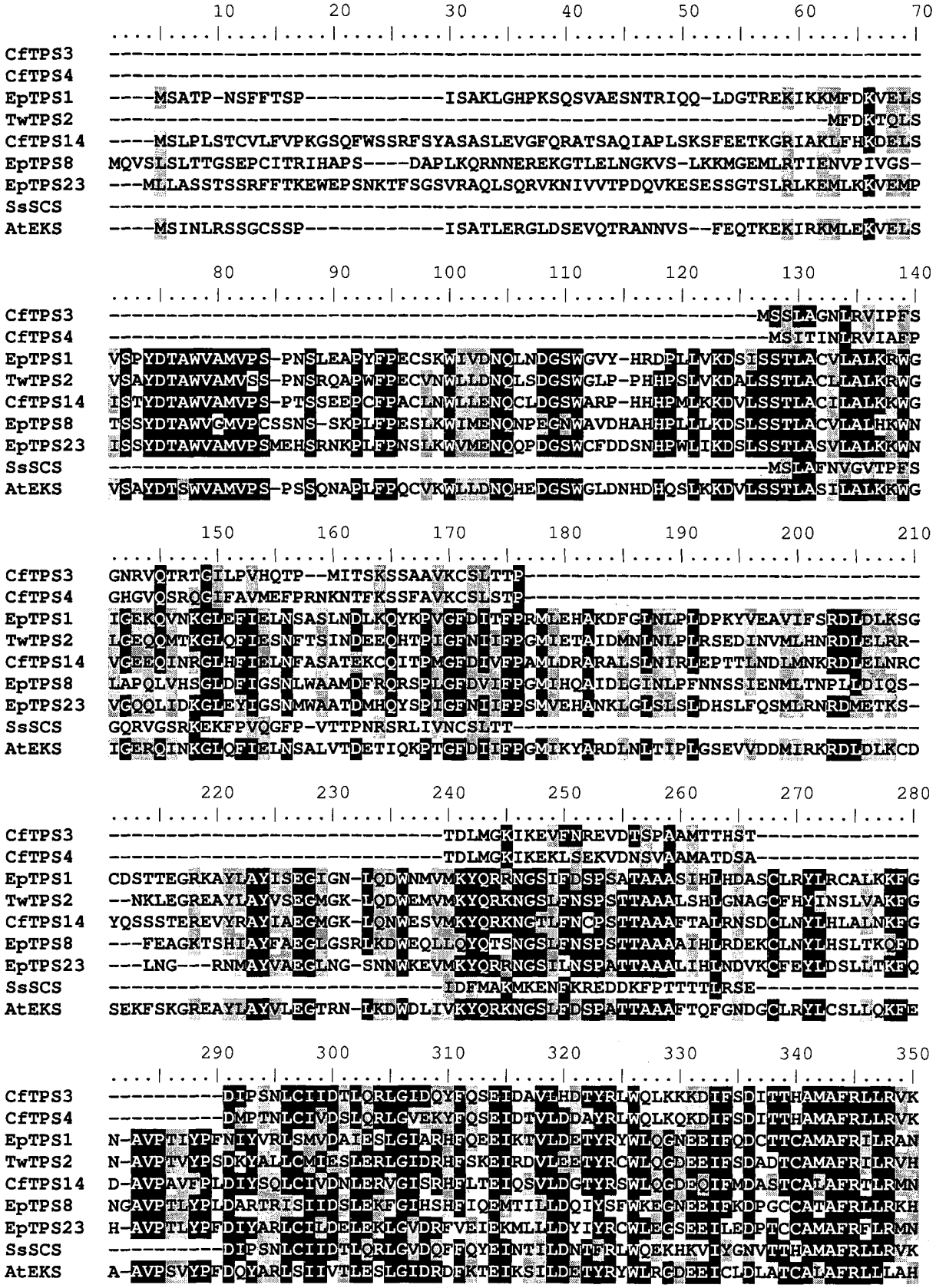

CfTPS3
CfTPS4
EpTPS1
TwTPS2
CfTPS14
EpTPS8
EpTPS23
SsSCS
AtEKS

Fig 4

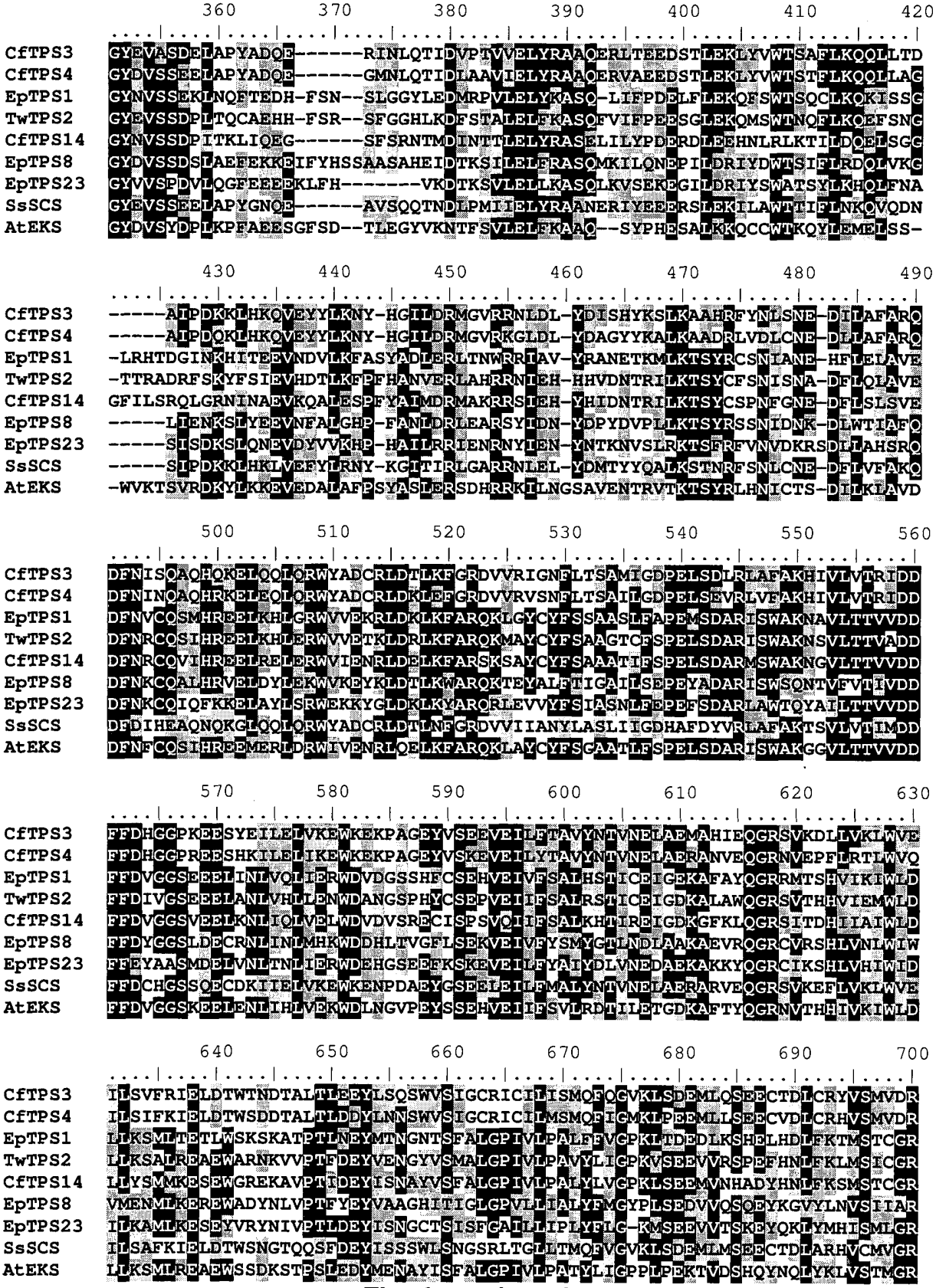


Fig. 4 continued

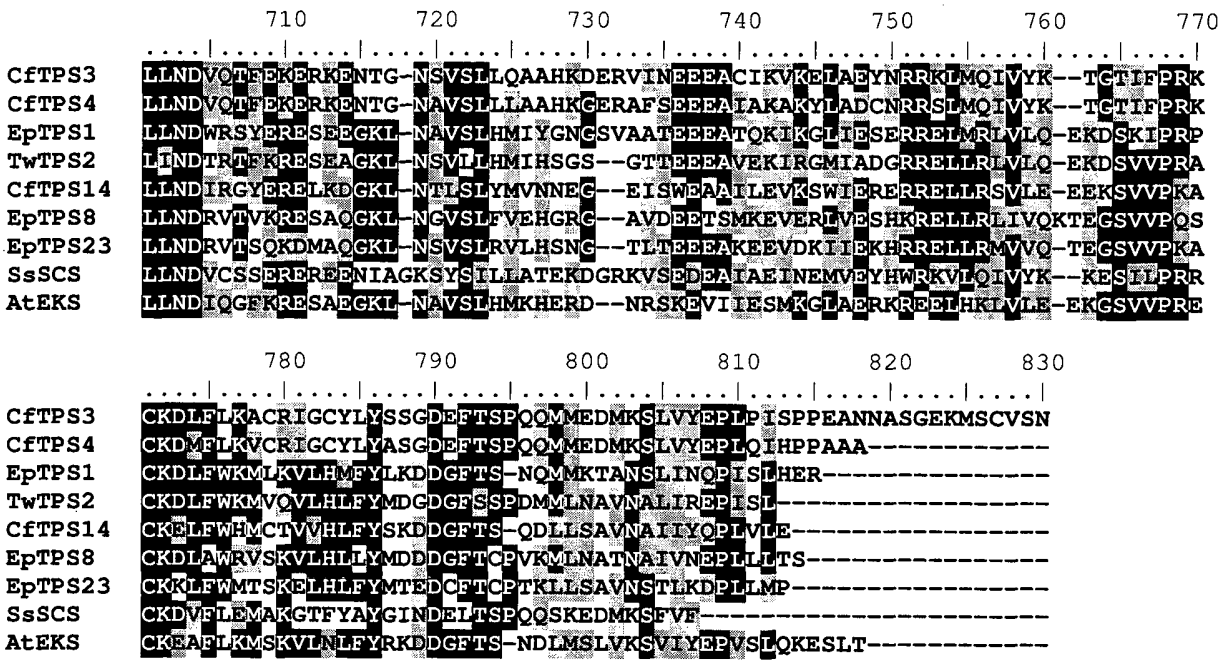


Fig. 4 continued

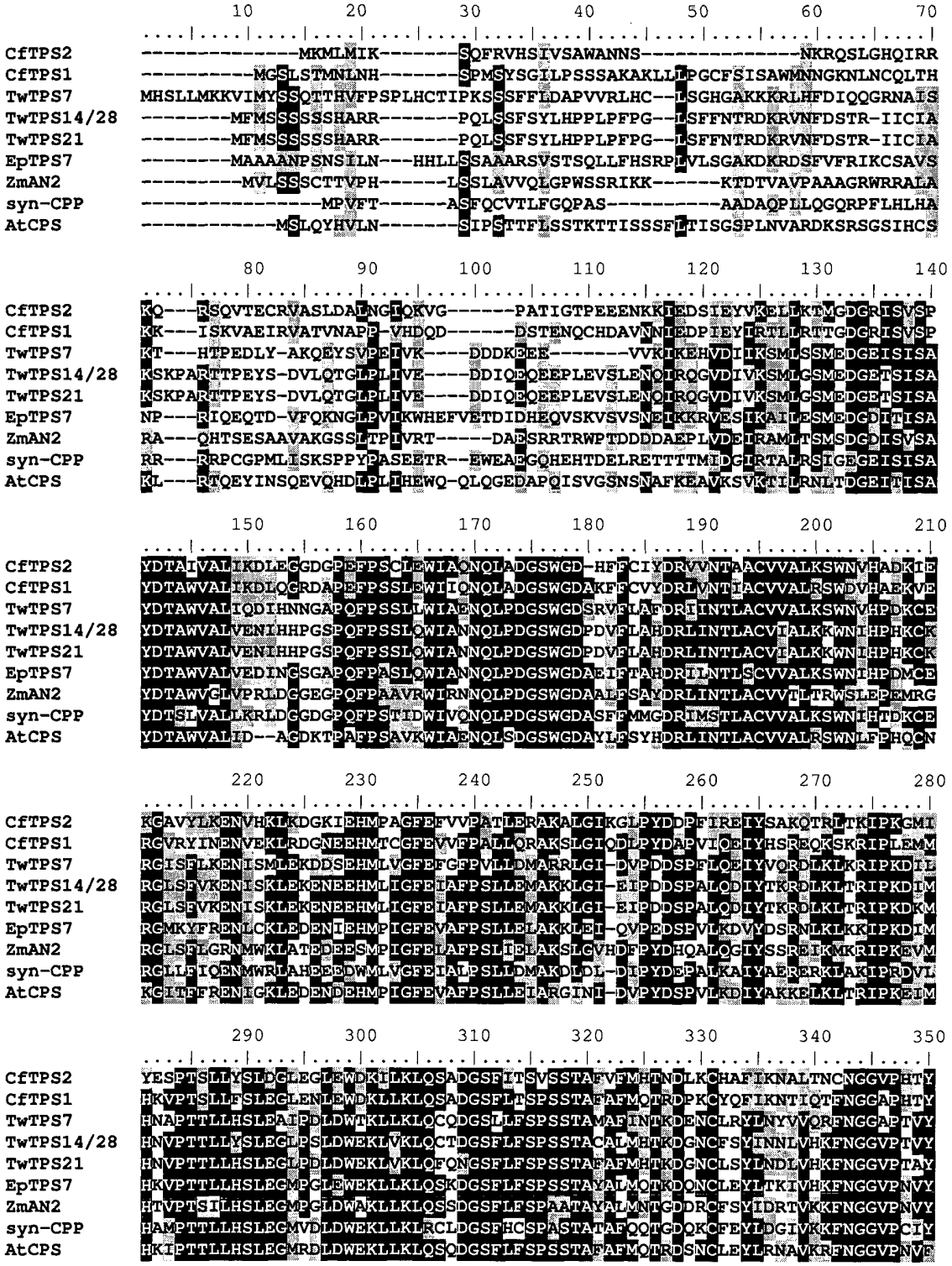

CfTPS2
CfTPS1
TwTPS7
TwTPS14/28
TwTPS21
EpTPS7
ZmAN2
syn-CPP
AtCPS

Fig. 5

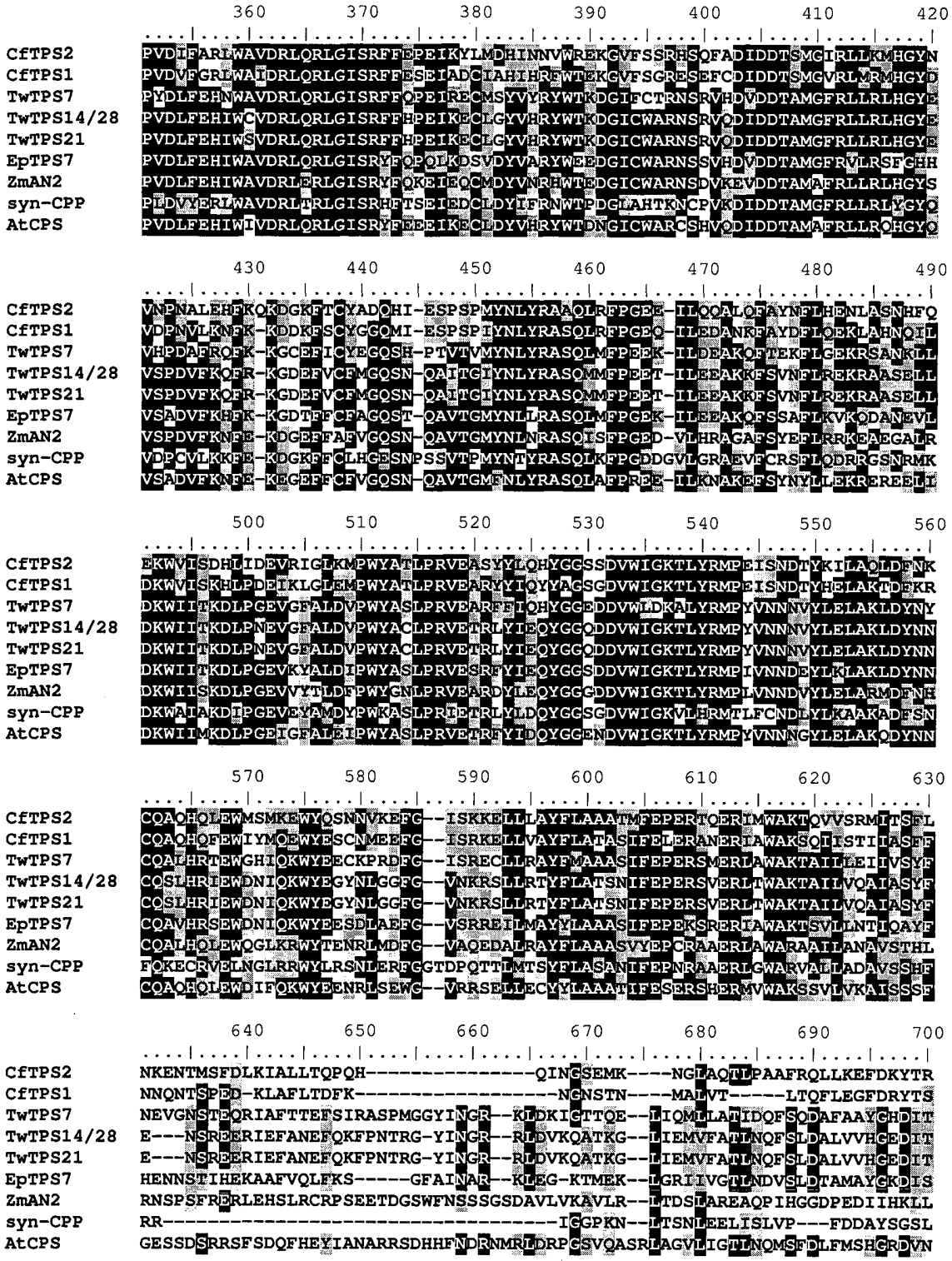


Fig. 5 continued

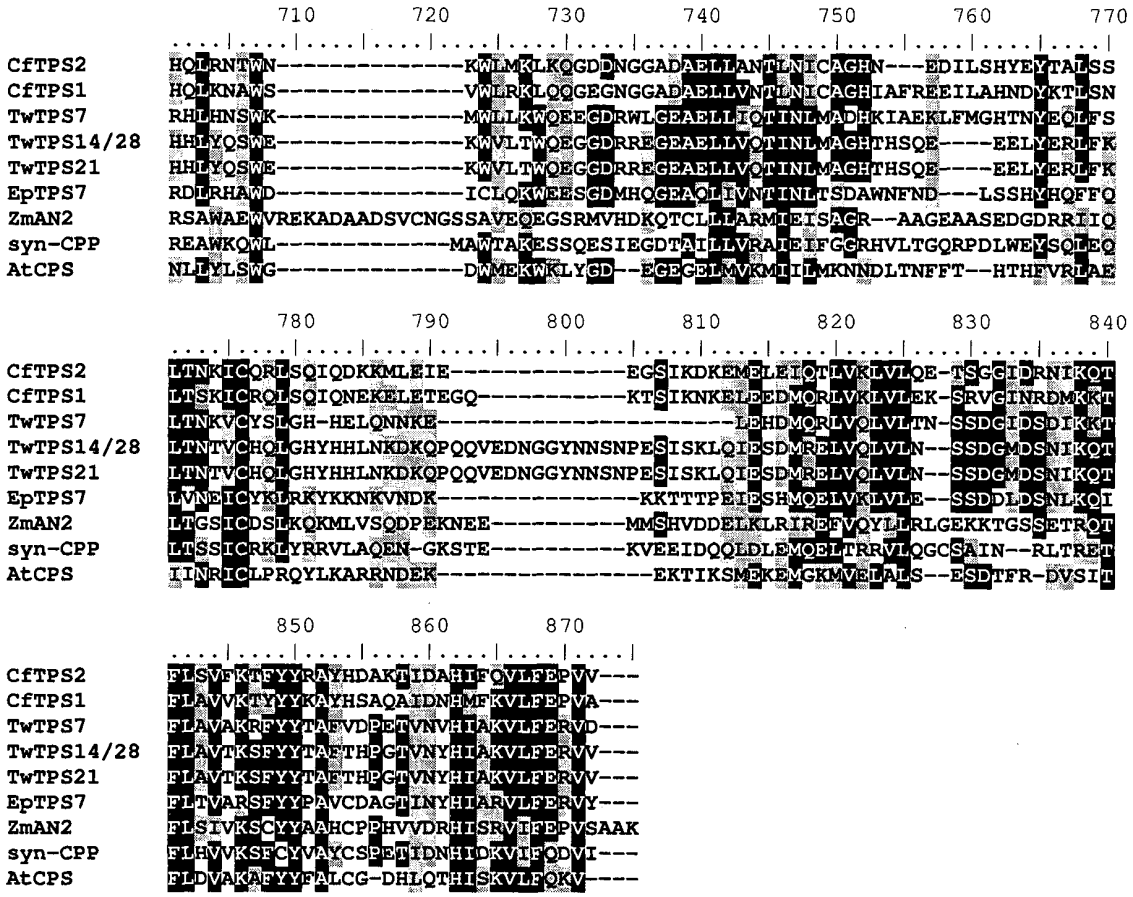


Fig. 5 continued

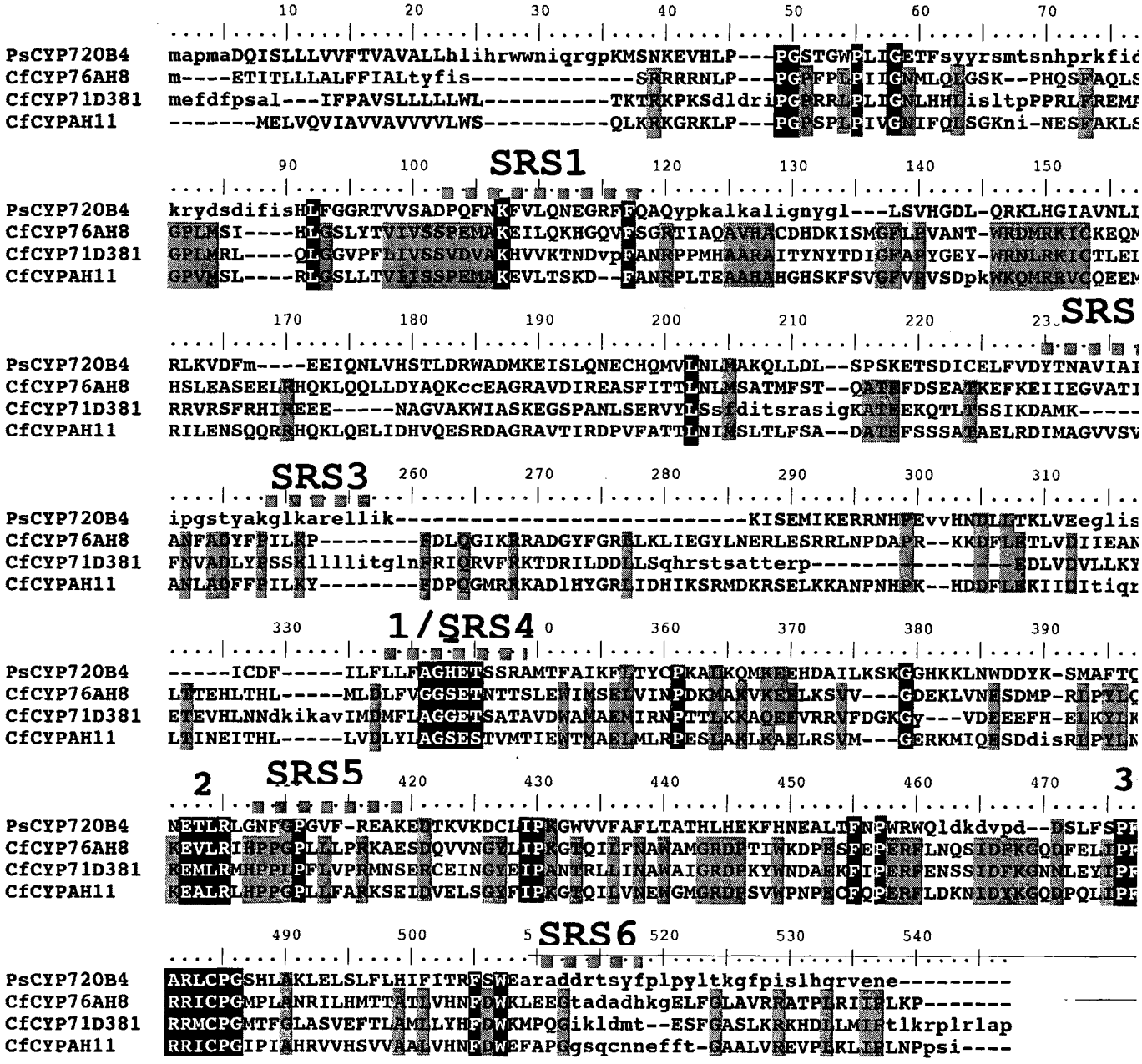

SRS1
SRS
SRS3
1/SRS4
2
SRS5
3
SRS6
PsCYP720B4
CfCYP76AH8
CfCYP71D381
CfCYPAH11

Fig. 6

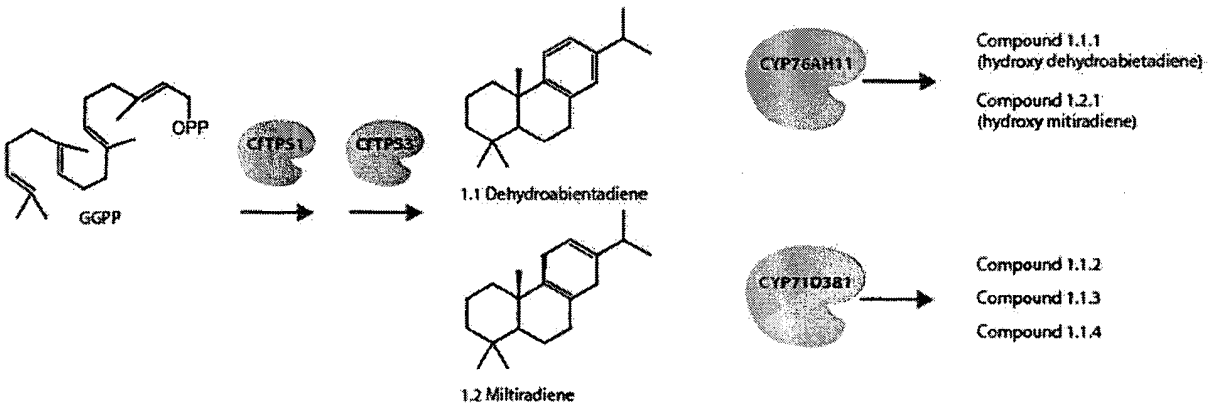

OPP
GGPP
CfTPS1
CfTPS3
1.1 Dehydroabientadiene
1.2 Miltiradiene
CYP76AH11
Compound 1.1.1
(hydroxy dehydroabietadiene)
Compound 1.2.1
(hydroxy mitiradiene)
CYP71D381
Compound 1.1.2
Compound 1.1.3
Compound 1.1.4

Fig. 7

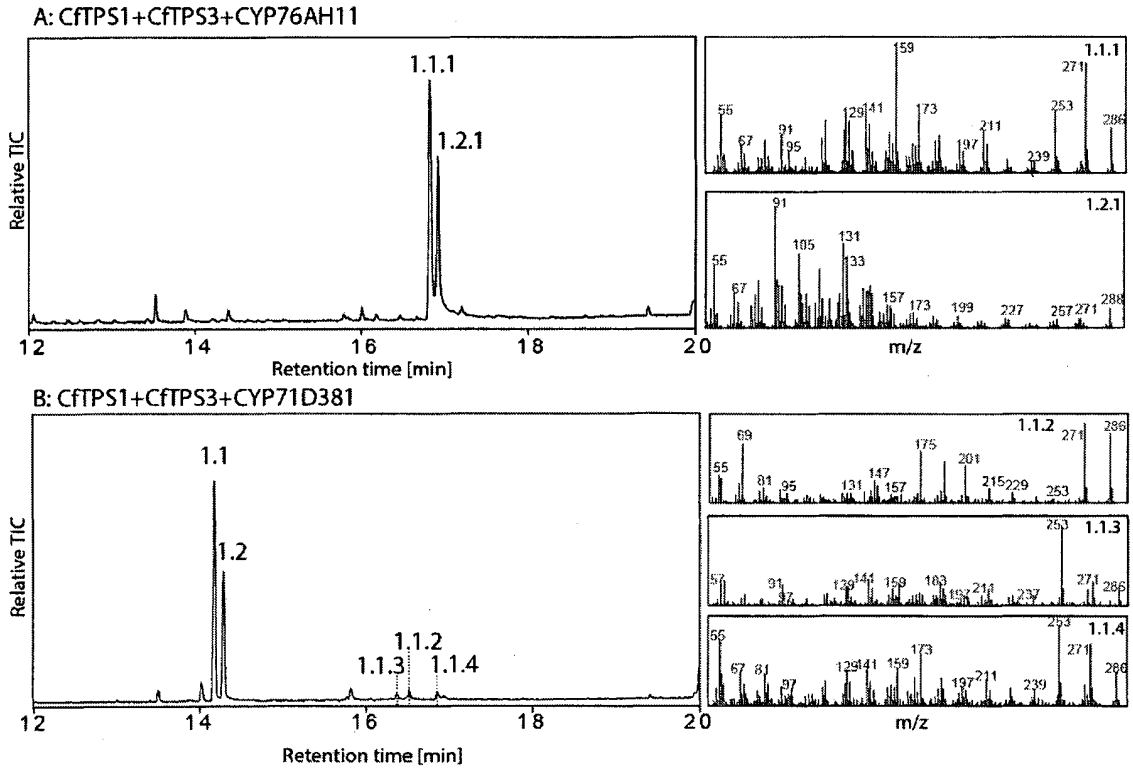

A: CfTPS1+CfTPS3+CYP76AH11
Relative TIC
1.1.1
1.2.1
Retention time [min]
m/z
B: CfTPS1+CfTPS3+CYP71D381
1.1
1.2
1.1.2
1.1.3
1.1.4

Fig. 8

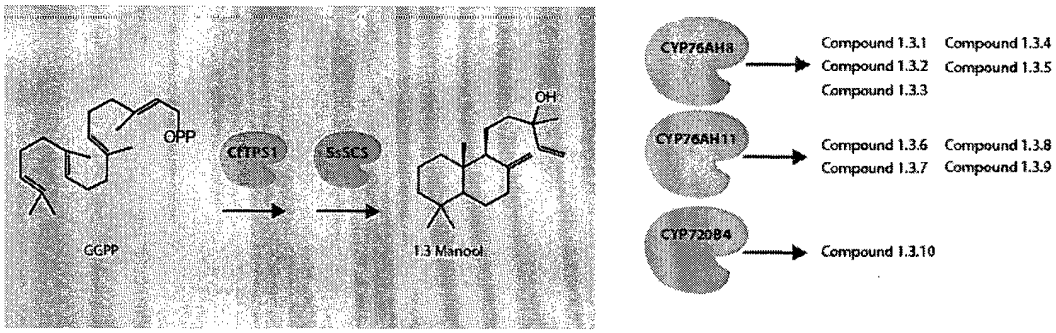

OPP
GGPP
CfTPS1
SsSCS
OH
1.3 Manool
CYP76AH8
Compound 1.3.1
Compound 1.3.2
Compound 1.3.3
Compound 1.3.4
Compound 1.3.5
CYP76AH11
Compound 1.3.6
Compound 1.3.7
Compound 1.3.8
Compound 1.3.9
CYP720B4
Compound 1.3.10

Fig. 9

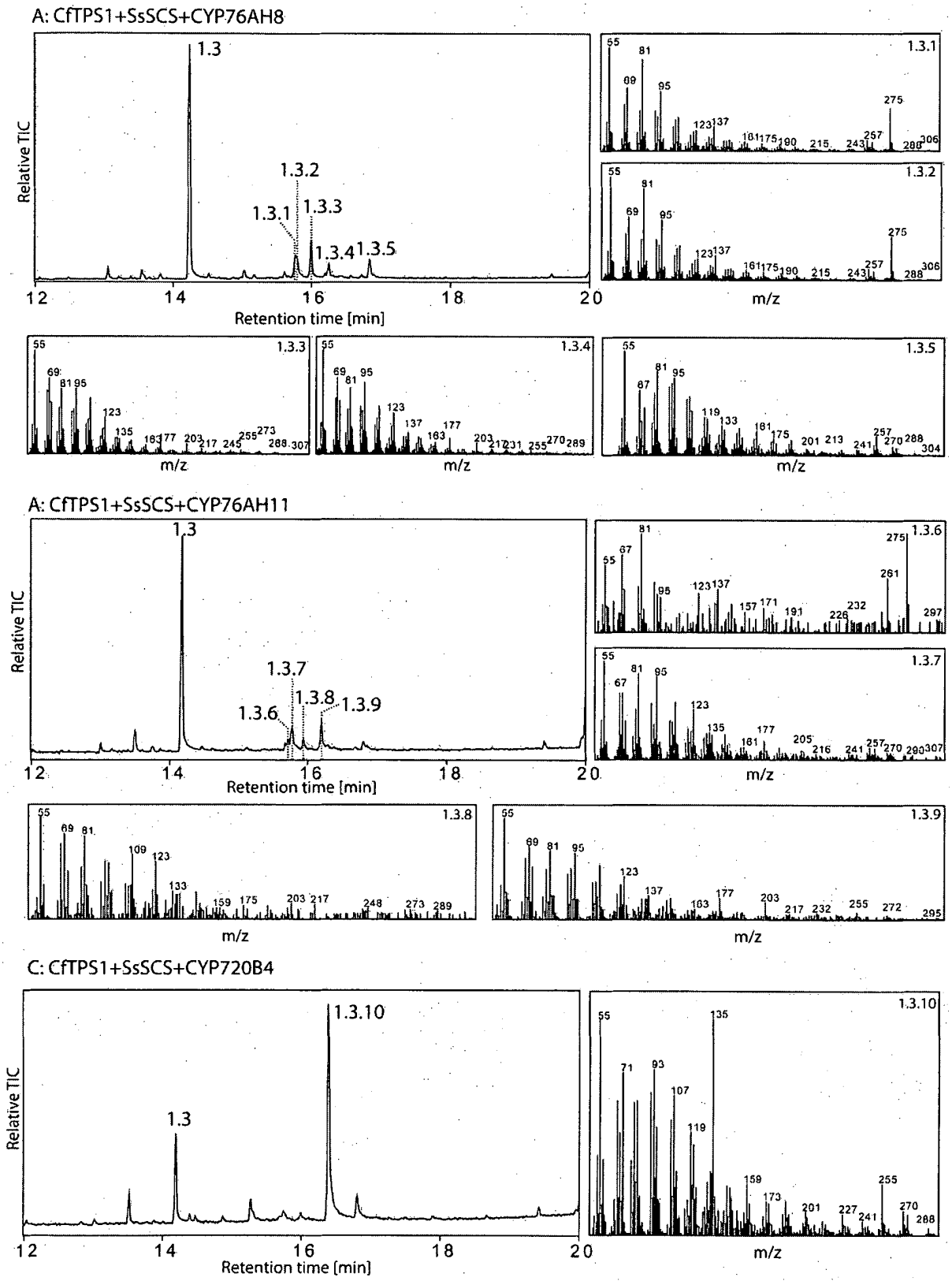

A: CfTPS1+SsSCS+CYP76AH8
Relative TIC
1.3
1.3.1
1.3.2
1.3.3
1.3.4
1.3.5
Retention time [min]
m/z
A: CfTPS1+SsSCS+CYP76AH11
1.3.6
1.3.7
1.3.8
1.3.9
C: CfTPS1+SsSCS+CYP720B4
1.3.10

Fig. 10

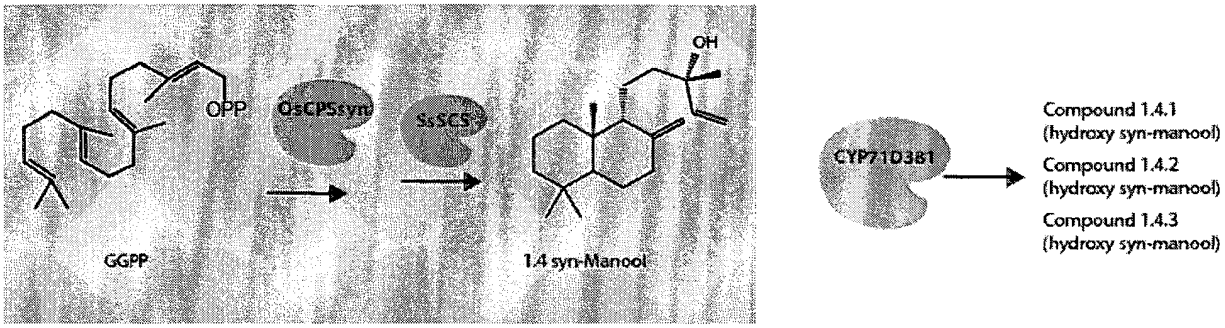

OPP
OsCPSsyn
SsSCS
OH
GGPP
1.4 syn-Manool
CYP71D381
Compound 1.4.1
(hydroxy syn-manool)
Compound 1.4.2
(hydroxy syn-manool)
Compound 1.4.3
(hydroxy syn-manool)

Fig. 11

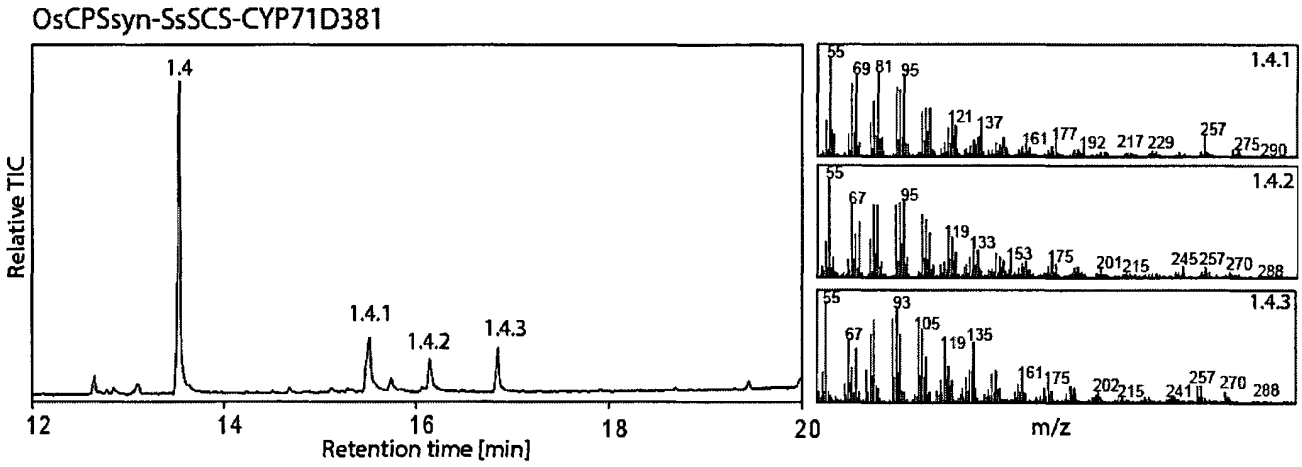

OsCPSsyn-SsSCS-CYP71D381
Relative TIC
1.4
1.4.1
1.4.2
1.4.3
12
14
16
18
20
Retention time [min]
1.4.1
1.4.2
1.4.3
m/z

Fig. 12

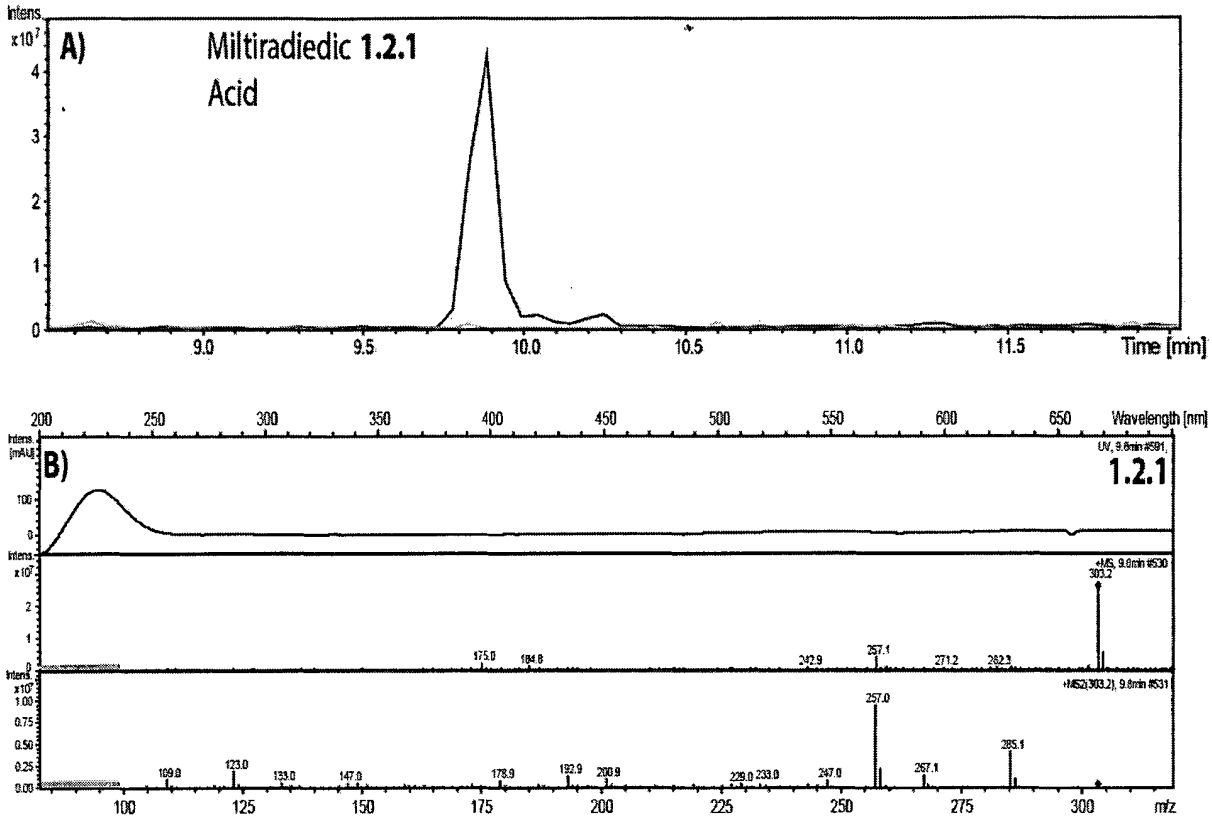
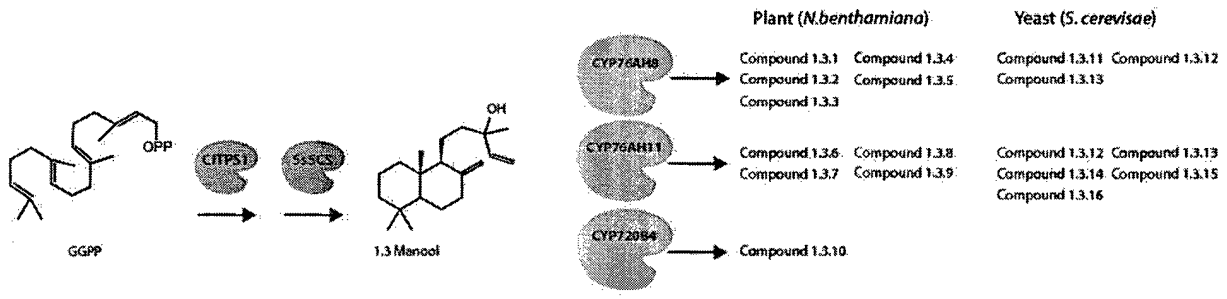
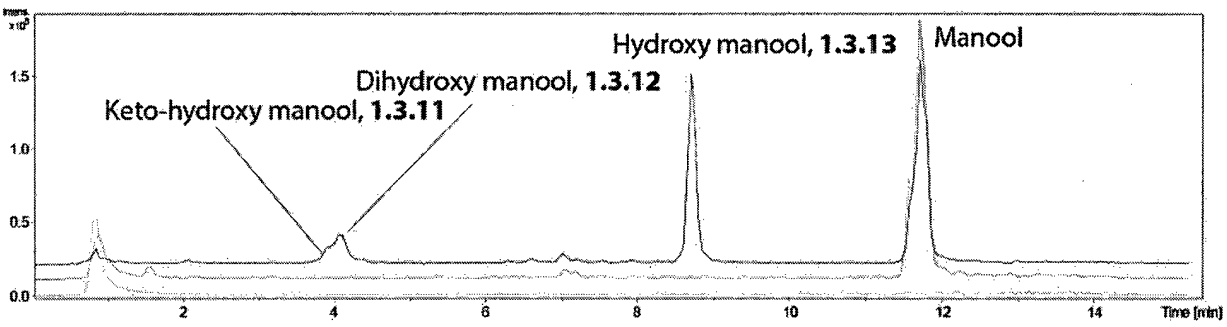

A)
Miltiradiedic 1.2.1
Acid
Time [min]
B)
1.2.1
Wavelength [nm]
m/z
C)
OPP
GGPP
CfTPS1
CfTPS3
1.1 Dehydroabientadiene
1.2 Miltiradiene
PsCYP72084
Yeast (S. cerevisae)
Compound 1.2.11
(miltiradiedic acid)

Fig. 13

A)


OPP
CiTPS1
SsSCS
OH
GGPP
1.3 Manool
Plant (N.benthamiana)
Yeast (S. cerevisae)
CYP76AH8
Compound 1.3.1 Compound 1.3.4
Compound 1.3.2 Compound 1.3.5
Compound 1.3.3
Compound 1.3.11 Compound 1.3.12
Compound 1.3.13
CYP76AH11
Compound 1.3.6 Compound 1.3.8
Compound 1.3.7 Compound 1.3.9
Compound 1.3.12 Compound 1.3.13
Compound 1.3.14 Compound 1.3.15
Compound 1.3.16
CYP72B64
Compound 1.3.10

B)


Keto-hydroxy manool, 1.3.11
Dihydroxy manool, 1.3.12
Hydroxy manool, 1.3.13
Manool
Time [min]

C)

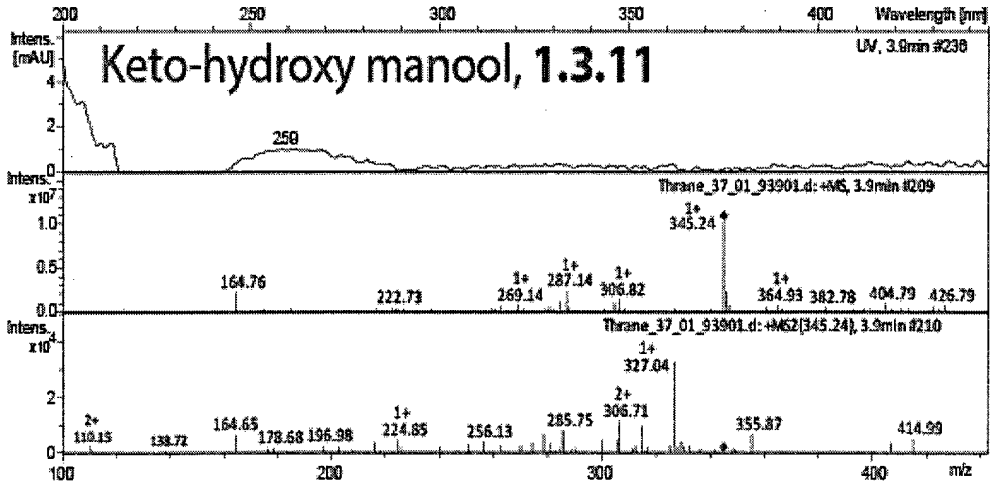
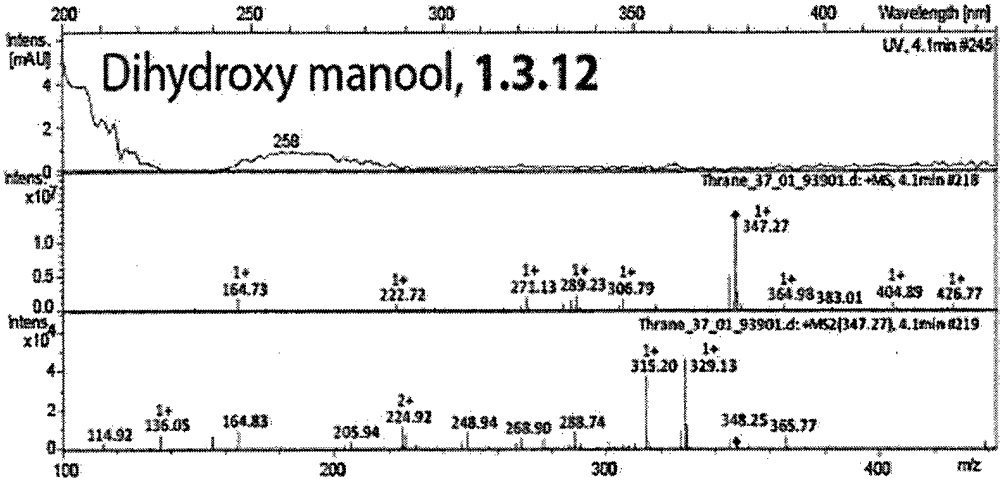

Keto-hydroxy manool, 1.3.11
Wavelength [nm]
UV, 3.9min #236
Intens. [mAU]
259
Thrane_37_01_93901.d: +MS, 3.9min #209
Thrane_37_01_93901.d: +MS2(345.24), 3.9min #210
m/z

Fig. 14

D)



E)

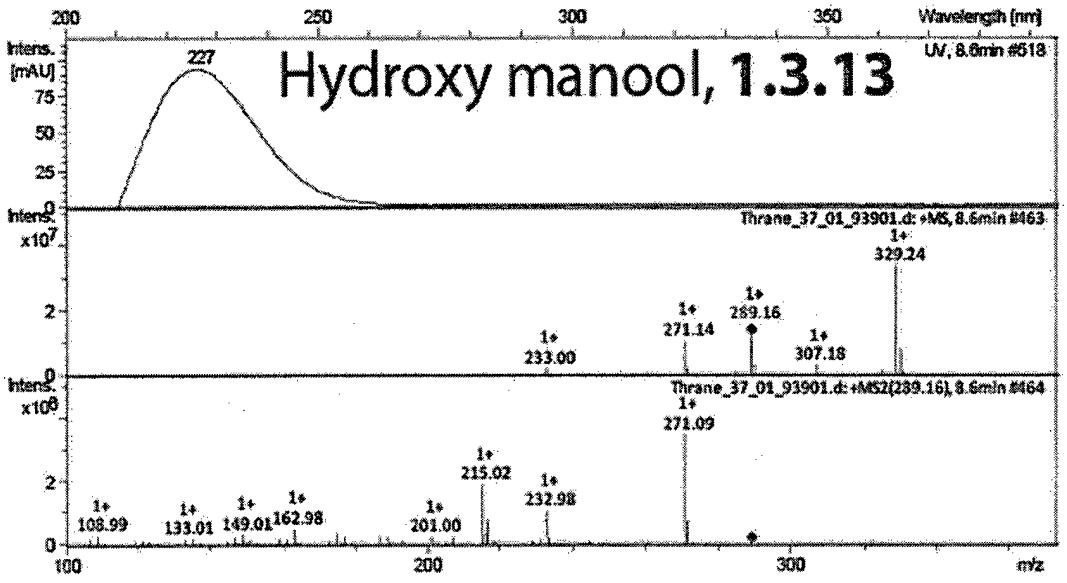
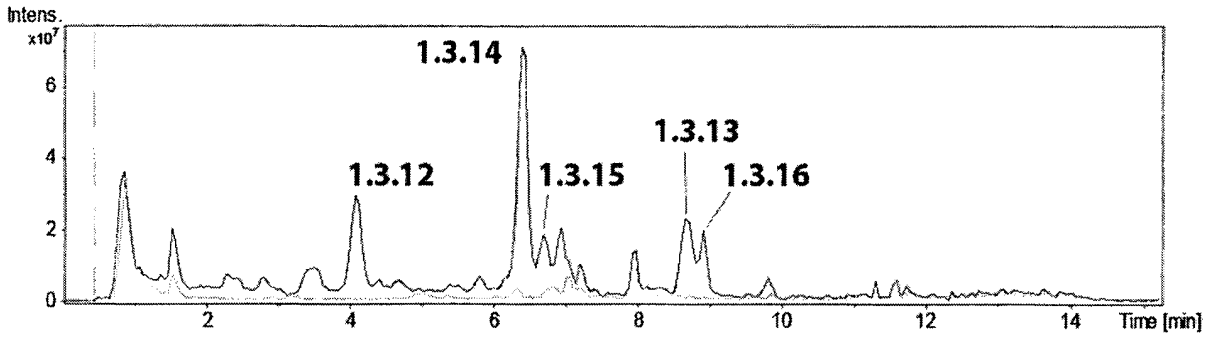


Fig. 14'cont

A)


Intens.
x10^7
1.3.14
1.3.12
1.3.15
1.3.13
1.3.16
Time [min]

B)

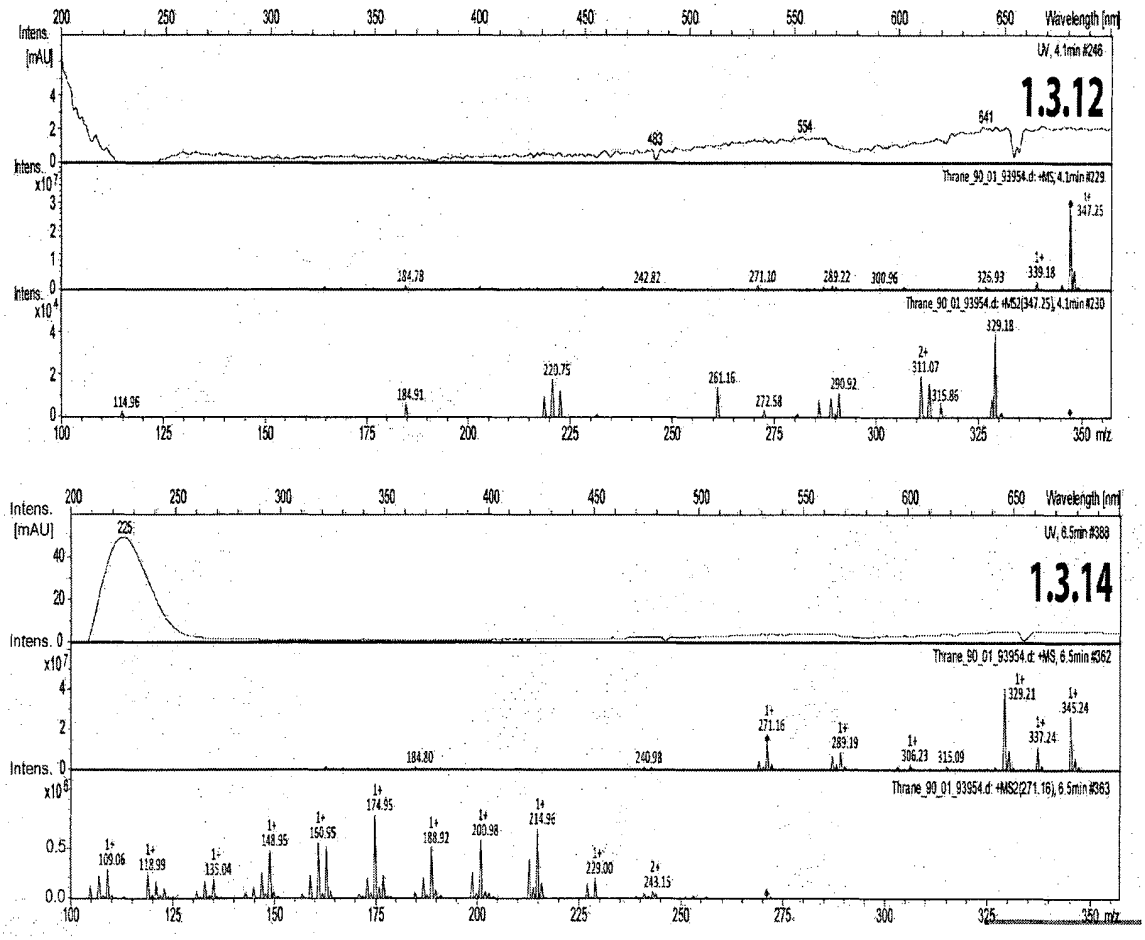

Wavelength [nm]
Intens.
[mAU]
UV, 4.1min #246
1.3.12
Thrane_90_01_93954.d: +MS, 4.1min #229
Thrane_90_01_93954.d: +MS2(347.25), 4.1min #230
m/z
UV, 6.5min #388
1.3.14
Thrane_90_01_93954.d: +MS, 6.5min #362
Thrane_90_01_93954.d: +MS2(271.16), 6.5min #363

Fig. 15

B)

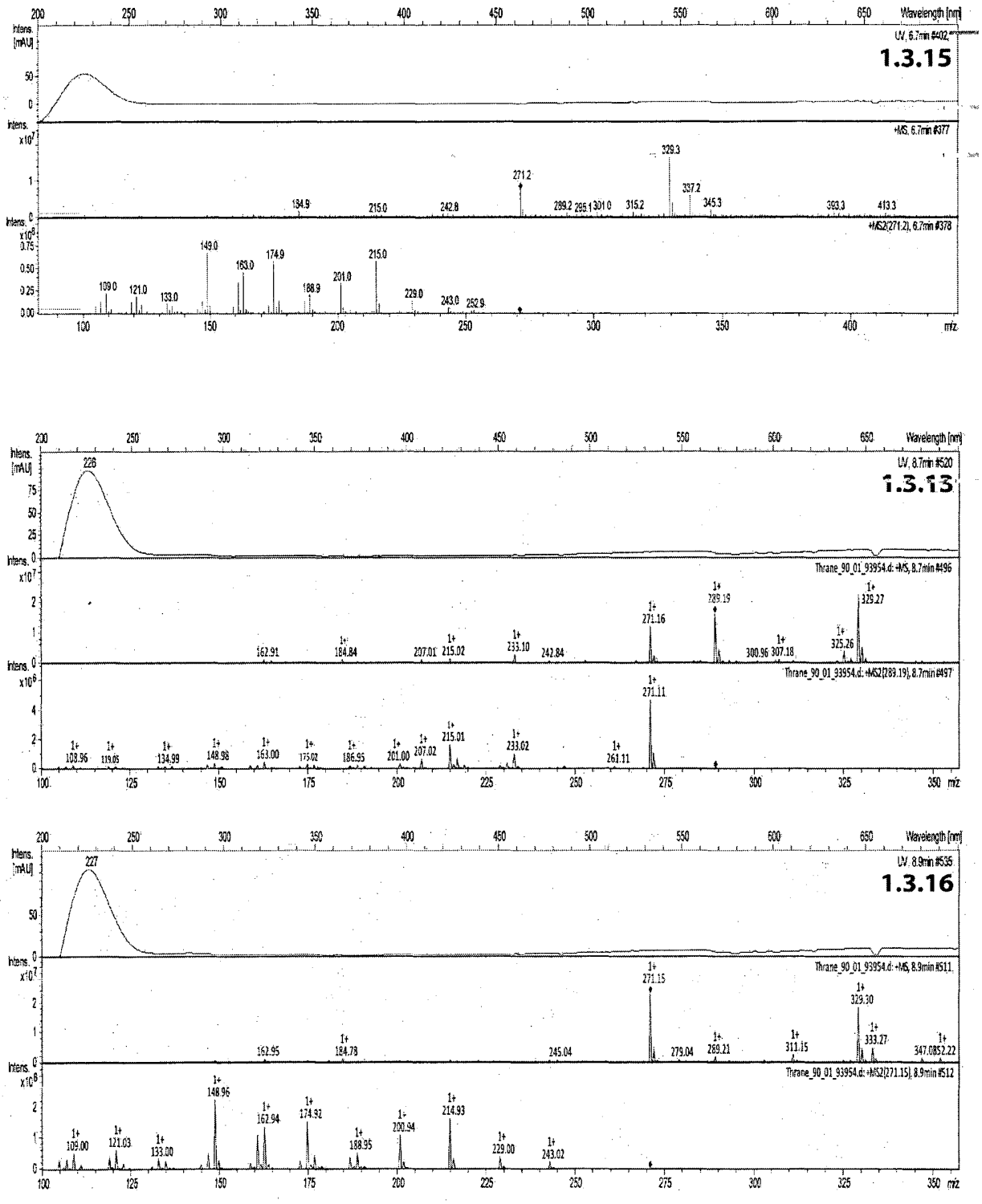
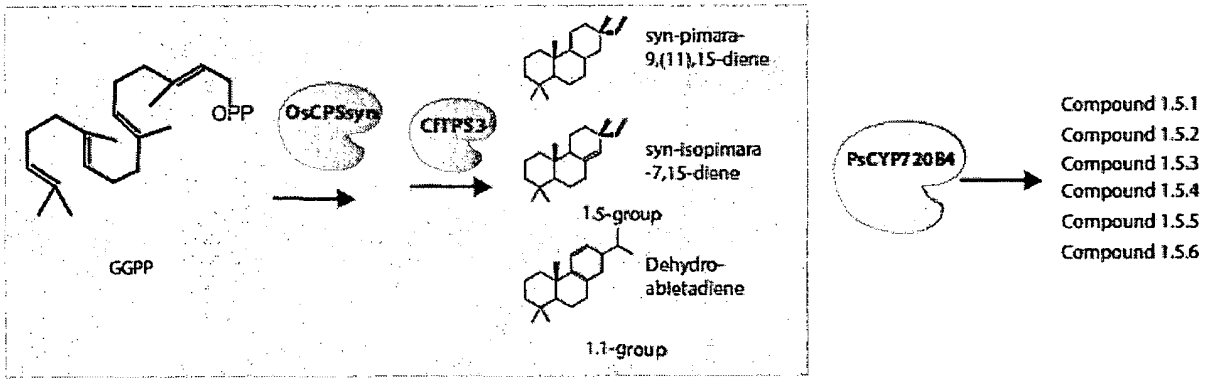
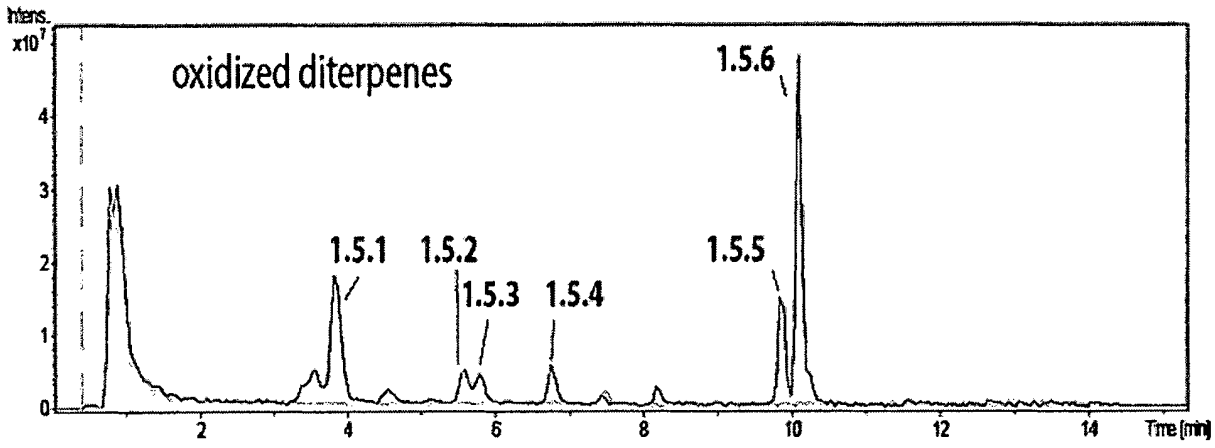

Wavelength [nm]
UV, 6.7min #402
1.3.15
+MS, 6.7min #377
+MS2(271.2), 6.7min #378
m/z
UV, 8.7min #520
1.3.13
Thrane_90_01_93954.d: +MS, 8.7min #496
Thrane_90_01_93954.d: +MS2(289.19), 8.7min #497
UV, 8.9min #535
1.3.16
Thrane_90_01_93954.d: +MS, 8.9min #511
Thrane_90_01_93954.d: +MS2(271.15), 8.9min #512

Fig. 15 – cont'

A)


OPP
GGPP
OsCPSsyn
CfTPS3
syn-pimara-9,(11),15-diene
syn-isopimara-7,15-diene
1.5-group
Dehydro-abietadiene
1.1-group
PsCYP720B4
Compound 1.5.1
Compound 1.5.2
Compound 1.5.3
Compound 1.5.4
Compound 1.5.5
Compound 1.5.6

B)


Intens.
x10^7
oxidized diterpenes
1.5.1
1.5.2
1.5.3
1.5.4
1.5.5
1.5.6
Time [min]

C)

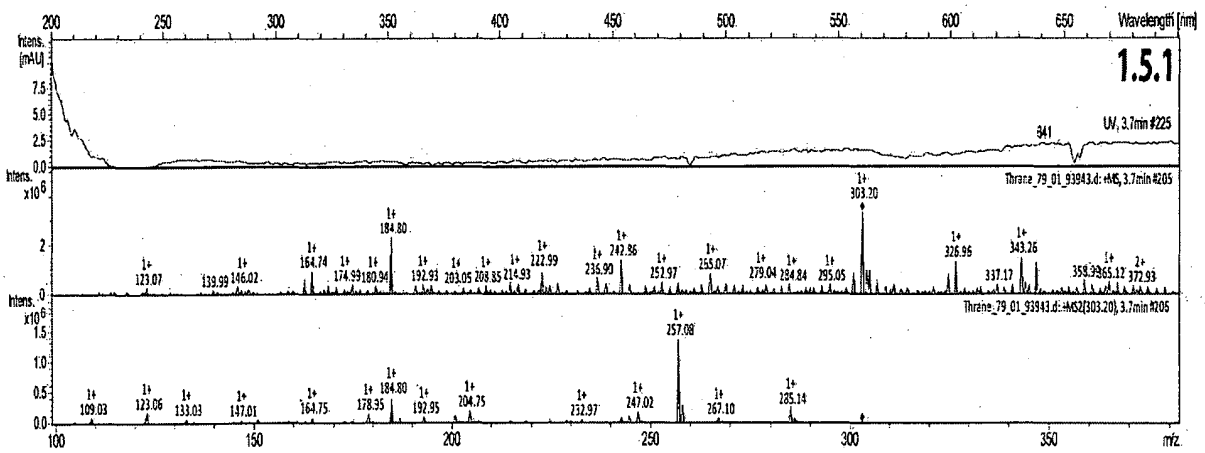

Wavelength [nm]
Intens. [mAU]
1.5.1
UV, 3.7min #225
Thrane_79_01_93943.d: +MS, 3.7min #205
Thrane_79_01_93943.d: +MS2(303.20), 3.7min #205
m/z

Fig. 16

C)

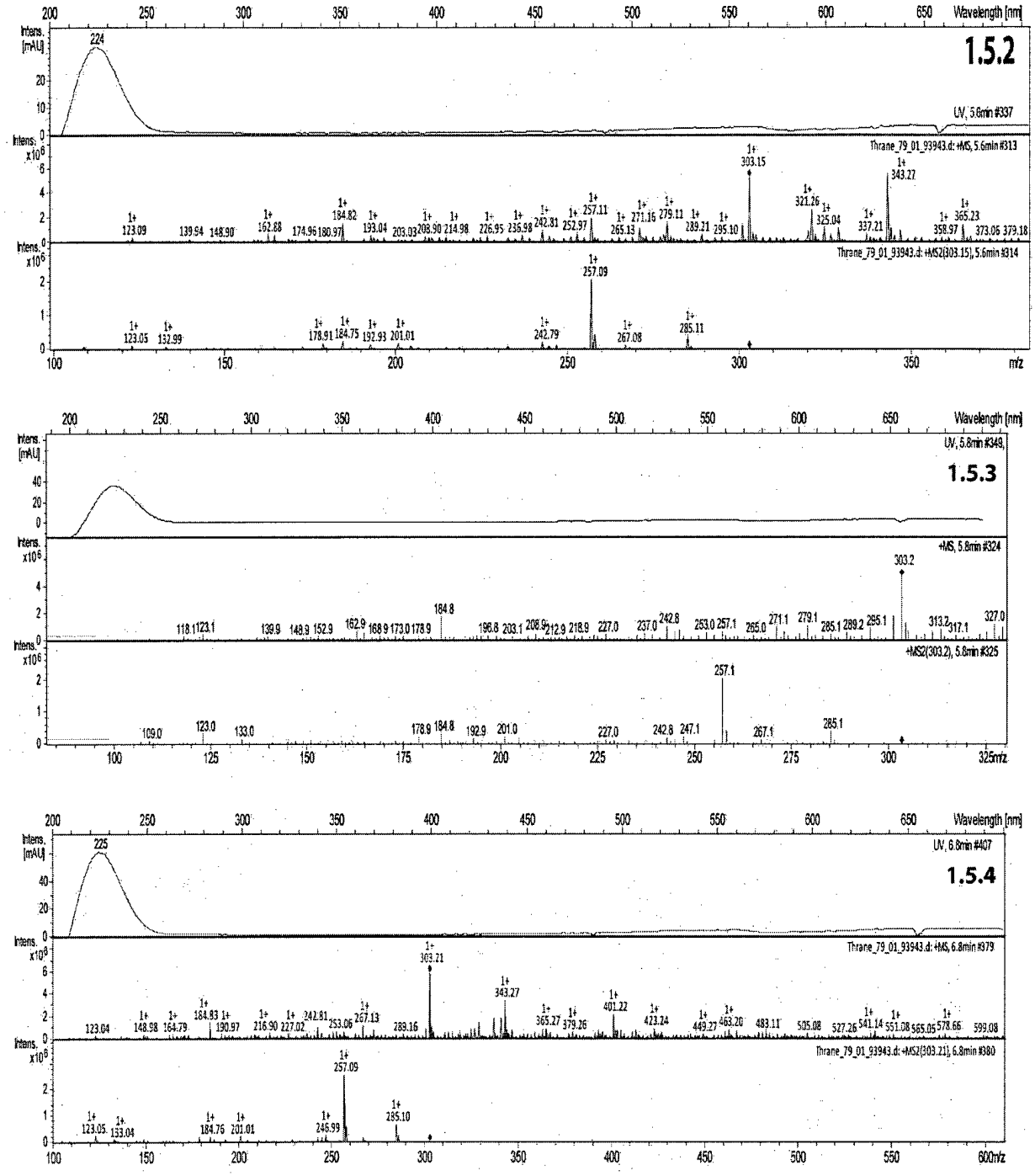

1.5.2
UV, 5.6min #337
Thrane_79_01_93943.d: +MS, 5.6min #313
Thrane_79_01_93943.d: +MS2(303.15), 5.6min #314
1.5.3
UV, 5.8min #349
+MS, 5.8min #324
+MS2(303.2), 5.8min #325
1.5.4
UV, 6.8min #407
Thrane_79_01_93943.d: +MS, 6.8min #379
Thrane_79_01_93943.d: +MS2(303.21), 6.8min #380
Wavelength [nm]
Intens. [mAU]
m/z

Fig. 16 – cont'

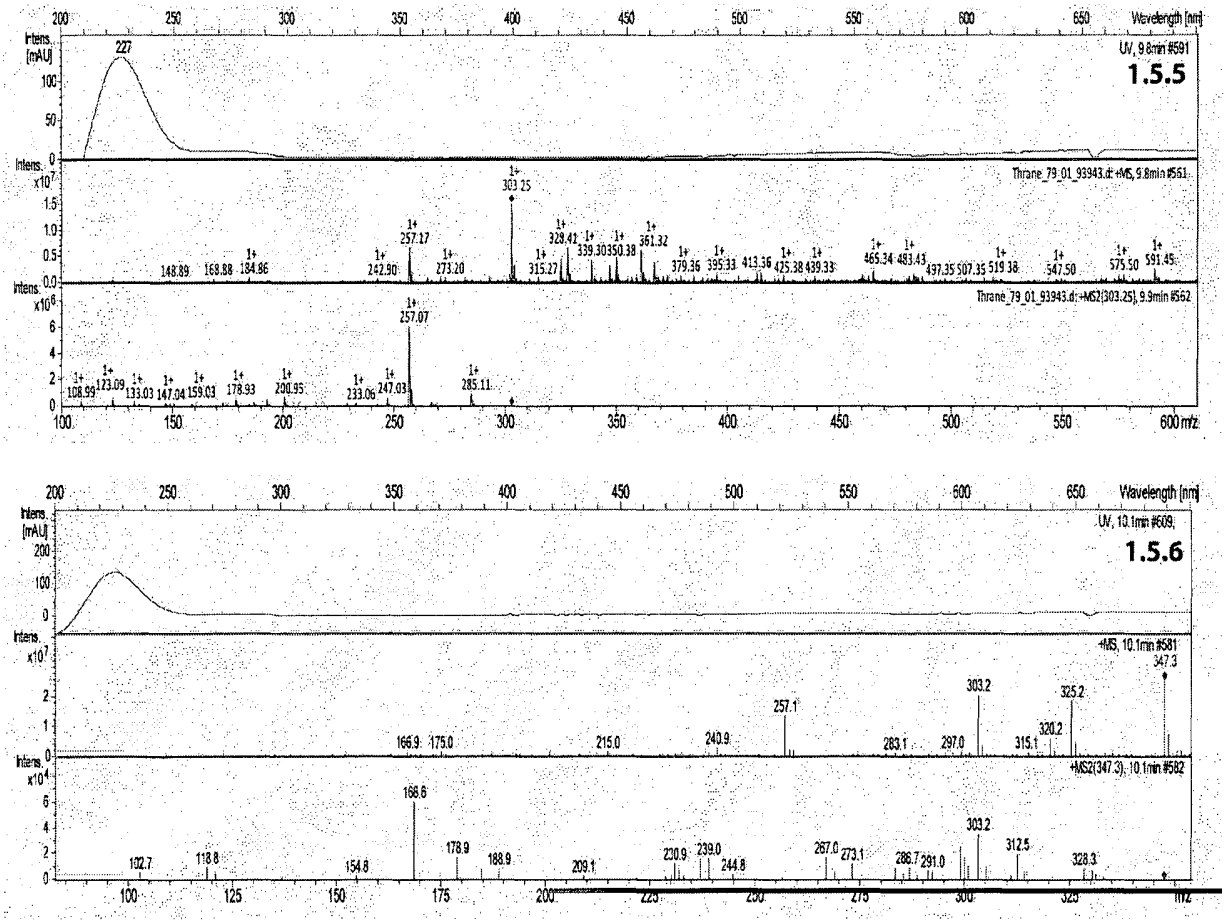
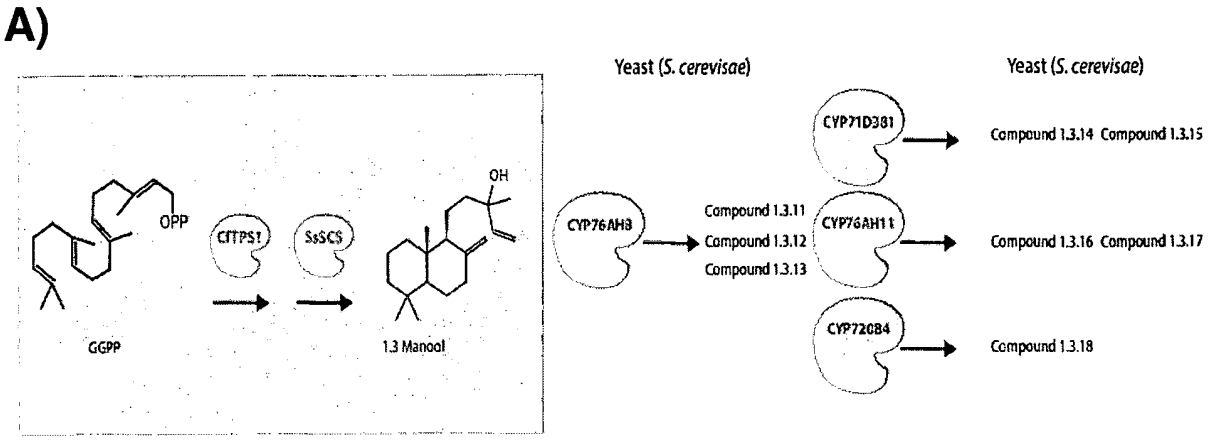
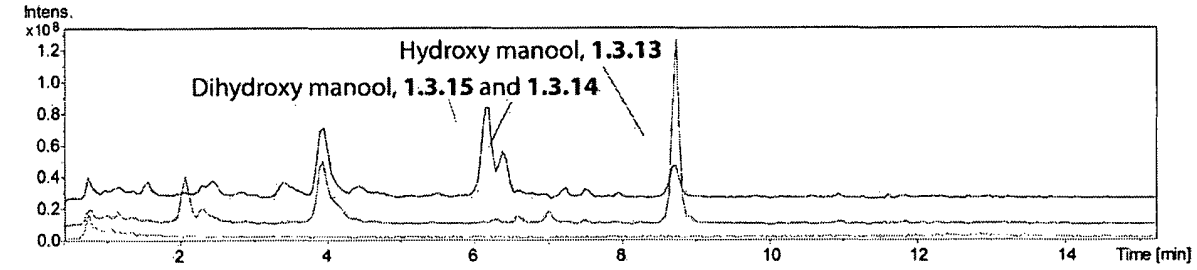

Wavelength [nm]
UV, 9.8min #591
1.5.5
Thrane_79_01_93943.d: +MS, 9.8min #551
Thrane_79_01_93943.d: +MS2(303.25), 9.9min #562
m/z
Wavelength [nm]
UV, 10.1min #609
1.5.6
+MS, 10.1min #581
+MS2(347.3), 10.1min #582
m/z

Fig. 16 – cont'


A)
Yeast (S. cerevisae)
Yeast (S. cerevisae)
OPP
CfTPS1
SsSCS
OH
GGPP
1.3 Manool
CYP76AH3
Compound 1.3.11
Compound 1.3.12
Compound 1.3.13
CYP71D381
CYP76AH11
CYP720B4
Compound 1.3.14 Compound 1.3.15
Compound 1.3.16 Compound 1.3.17
Compound 1.3.18


B)
Intens.
x10 8
Hydroxy manool, 1.3.13
Dihydroxy manool, 1.3.15 and 1.3.14
Time [min]

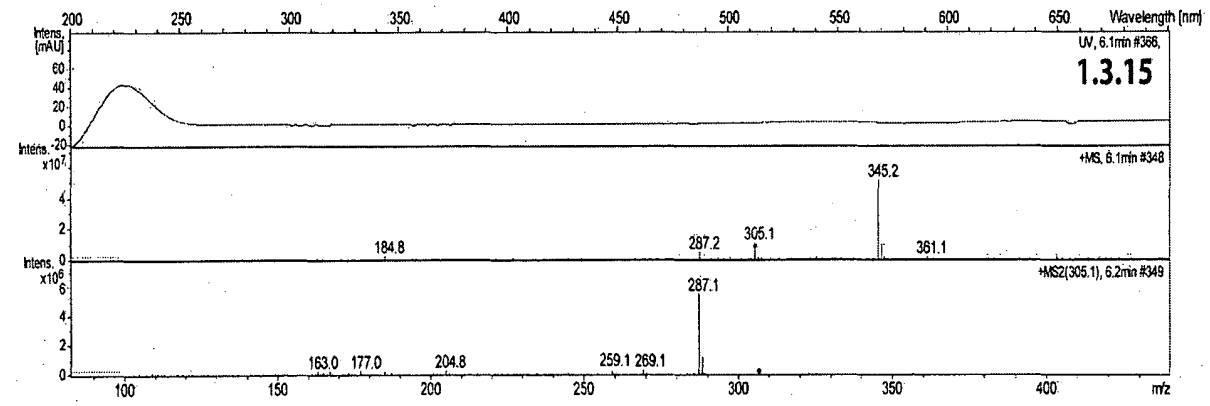

C)
Wavelength [nm]
UV, 6.1min #366,
1.3.15
+MS, 6.1min #348
345.2
184.8
287.2
305.1
361.1
+MS2(305.1), 6.2min #349
287.1
163.0
177.0
204.8
259.1
269.1
m/z

Fig. 17

C)

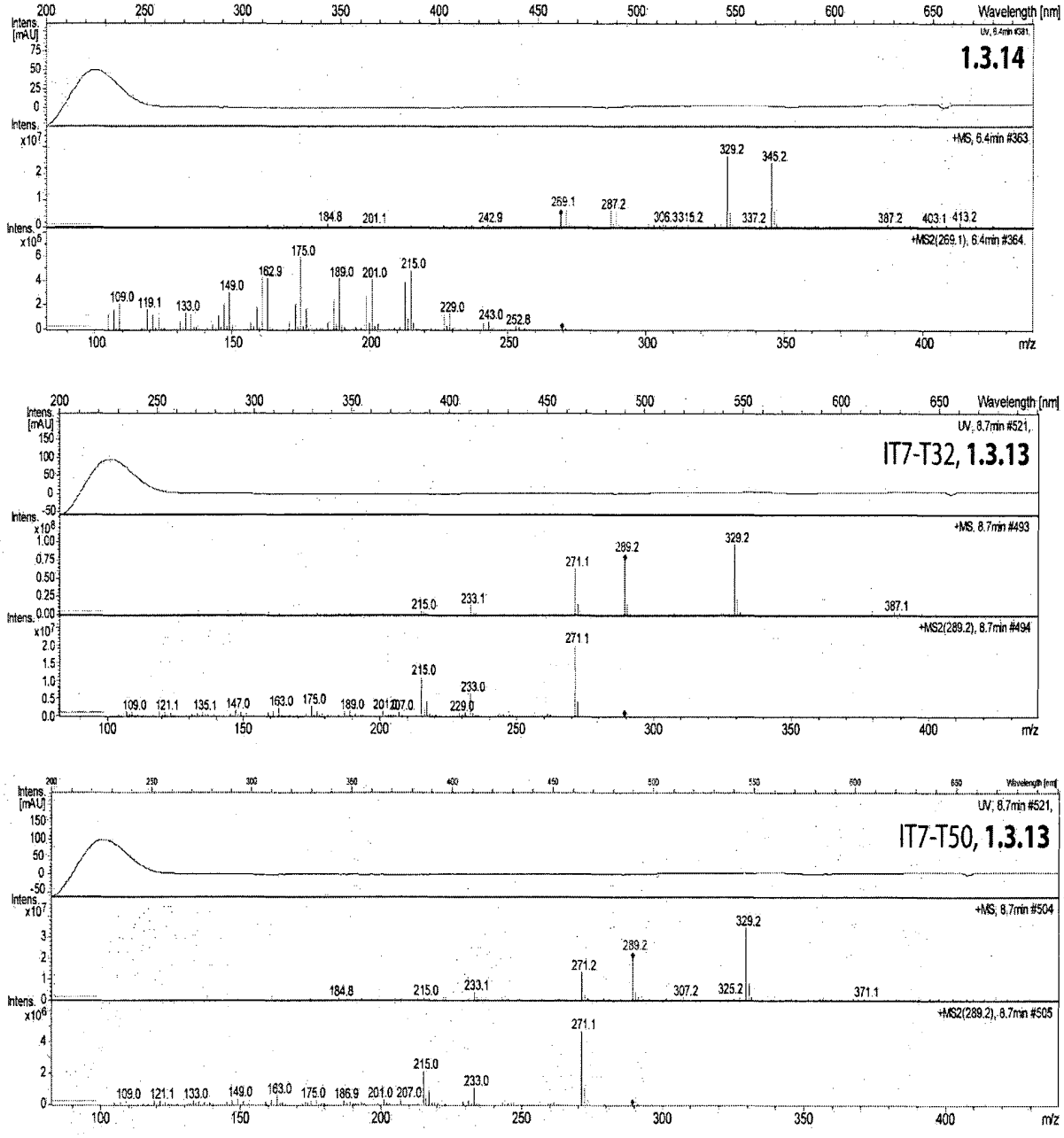
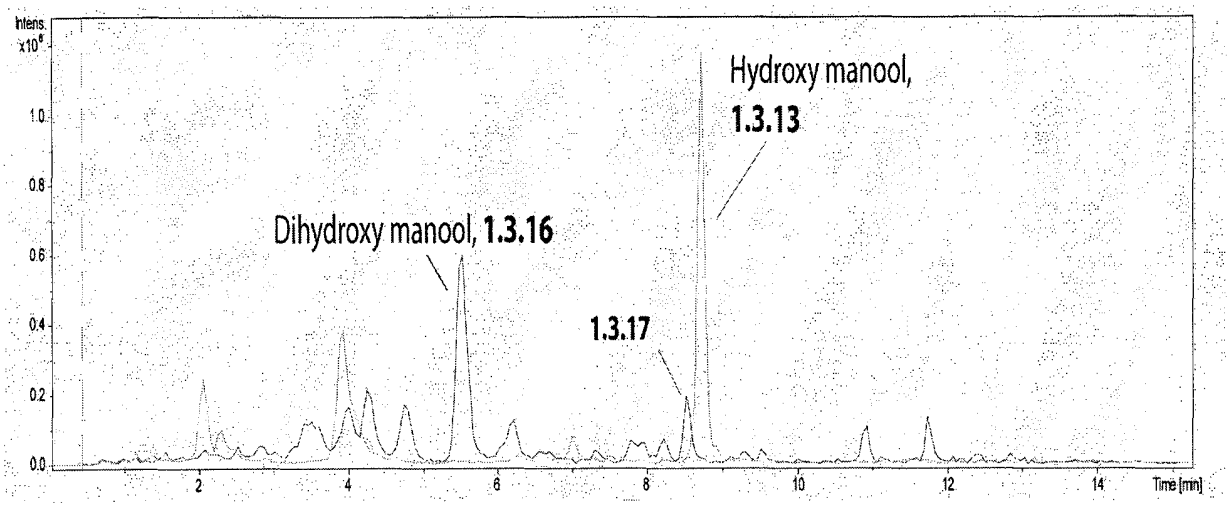

Wavelength [nm]
1.3.14
+MS, 6.4min #363
+MS2(269.1), 6.4min #364
m/z
IT7-T32, 1.3.13
UV, 8.7min #521,
+MS, 8.7min #493
+MS2(289.2), 8.7min #494
IT7-T50, 1.3.13
+MS, 8.7min #504
+MS2(289.2), 8.7min #505

Fig. 17 – cont'

A)


Intens.
x10^6
1.0
0.8
0.6
0.4
0.2
0.0
Hydroxy manool,
1.3.13
Dihydroxy manool, 1.3.16
1.3.17
2
4
6
8
10
12
14
Time [min]

B)

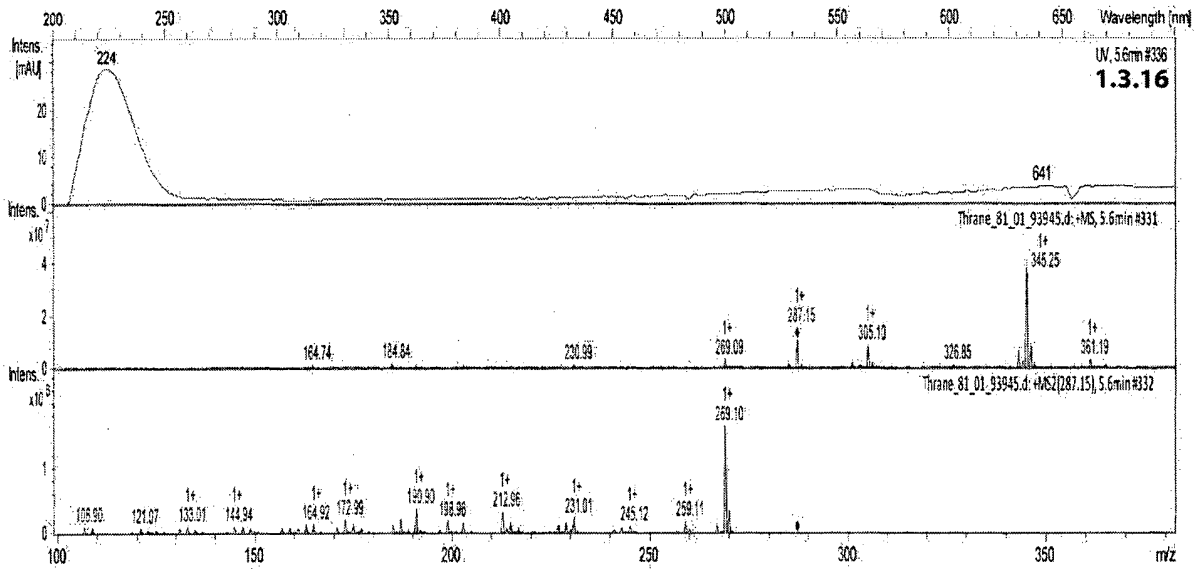

200
250
300
350
400
450
500
550
600
650
Wavelength [nm]
Intens.
[mAU]
224
UV, 5.6min #336
1.3.16
641
Intens.
Thirane_81_01_93945.d: +MS, 5.6min #331
345.25
287.15
305.10
269.09
361.19
326.85
164.74
184.84
230.99
Intens.
Thirane_81_01_93945.d: +MS2(287.15), 5.6min #332
269.10
105.90
121.07
133.01
144.94
164.92
172.99
189.90
198.96
212.96
231.01
245.12
259.11
100
150
200
250
300
350
m/z

Fig. 18

B)

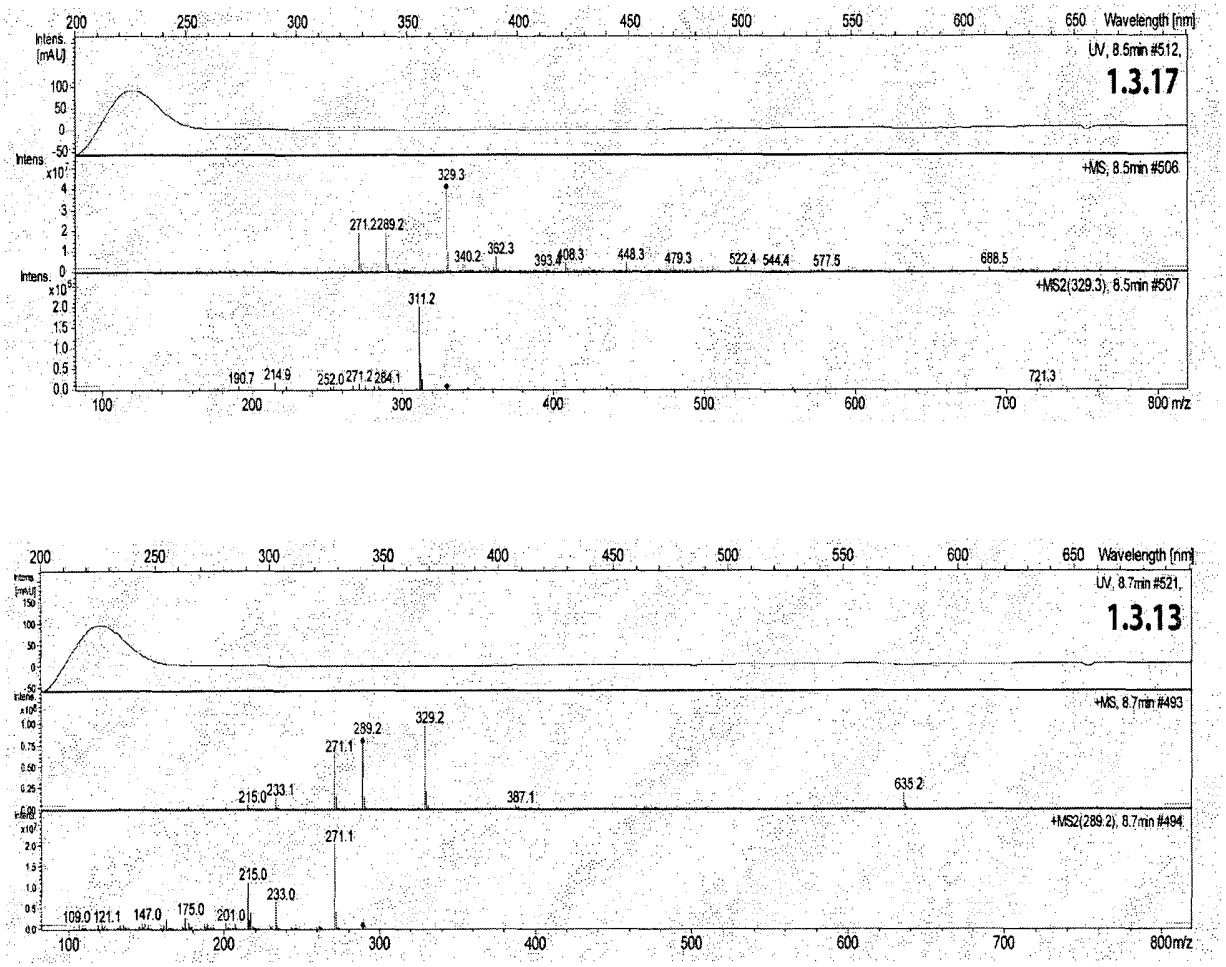
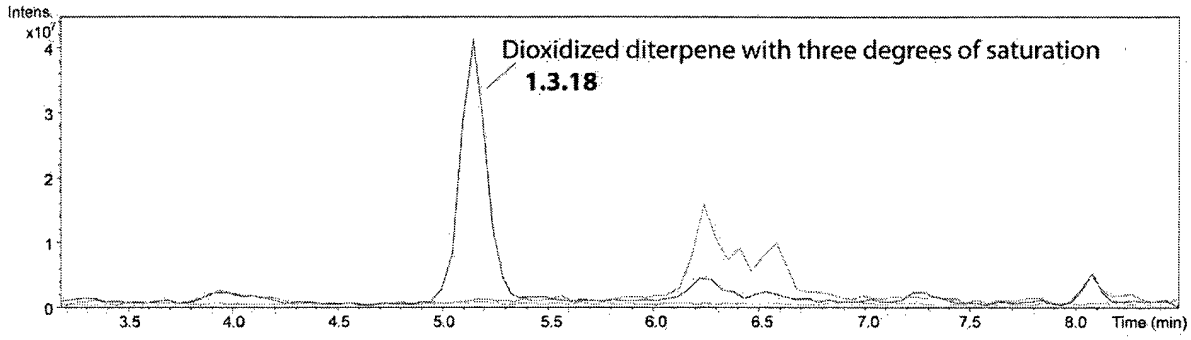

Wavelength [nm]
UV, 8.5min #512,
1.3.17
+MS, 8.5min #506
+MS2(329.3), 8.5min #507
m/z
UV, 8.7min #521,
1.3.13
+MS, 8.7min #493
+MS2(289.2), 8.7min #494

Fig. 18 - cont'

A)



B)

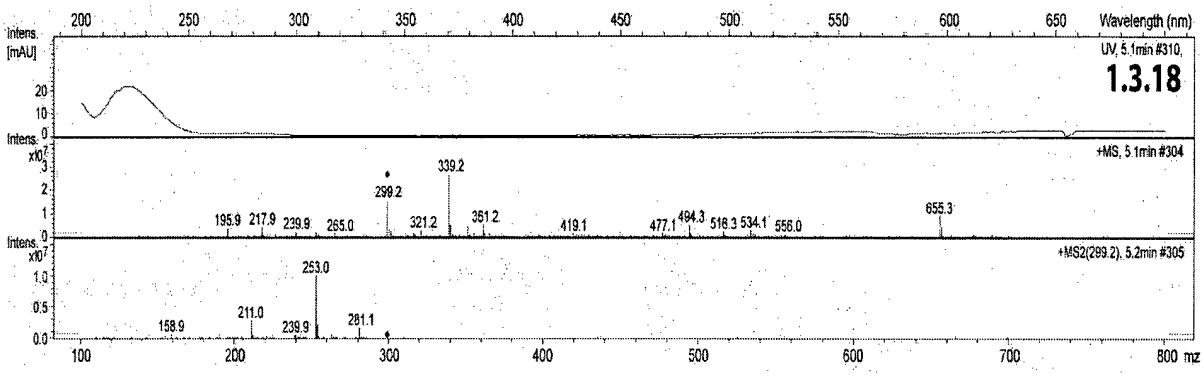


Fig. 19

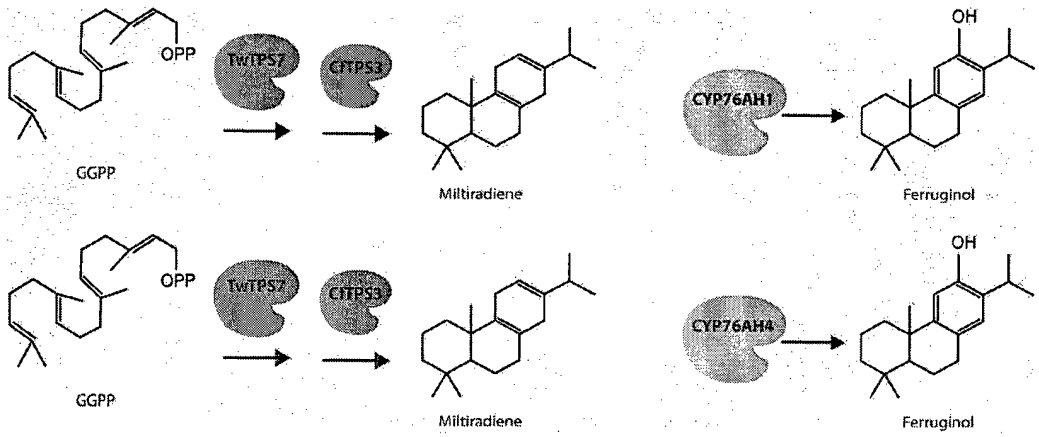


Fig. 20

INTERNATIONAL SEARCH REPORT

International application No
PCT/DK2015/050181

A. CLASSIFICATION OF SUBJECT MATTER
INV. C07K14/415 C12P5/00 C12P7/00
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C07K C12P

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPO-Internal, BIOSIS, EMBASE, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	PATERAKI IRINI ET AL: "Manoyl oxide (13R), the biosynthetic precursor of forskolin, is synthesized in specialized root cork cells in Coleus forskohlii", PLANT PHYSIOLOGY, AMERICAN SOCIETY OF PLANT PHYSIOLOGISTS, ROCKVILLE, MD, US, vol. 164, 1 March 2014 (2014-03-01), pages 1222-1236, XP002724033, ISSN: 0032-0889 abstract ----- -/--	1-98

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
28 September 2015	12/10/2015

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Griesinger, Irina

1

Form PCT/ISA/210 (second sheet) (April 2005)

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	MITCHELL ROD: "Identificati0n and characterization of diterpene synthases in the salvinorin A biosynthetic pathway", INTERNET CITATION, 1 January 2012 (2012-01-01), pages 1-2, XP002738984, Retrieved from the Internet: URL:http://theses.ucalgary.ca/bitstream/11023/171/5/ucalgary_2012_mitchell_rod.pdf [retrieved on 2015-04-27] figure 7	1-98
X	PHILIPP ZERBE ET AL: "Gene discovery of modular diterpene metabolism in nonmodel systems", PLANT PHYSIOLOGY, AMERICAN SOCIETY OF PLANT PHYSIOLOGISTS, ROCKVILLE, MD, US , vol. 162, no. 2 1 June 2013 (2013-06-01), pages 1073-1091, XP002724041, ISSN: 0032-0889, DOI: 10.1104/PP.113.218347 Retrieved from the Internet: URL:http://www.plantphysiol.org/content/162/2/1073 [retrieved on 2013-04-23] abstract	1-98
X	YONGJIN J. ZHOU ET AL: "Modular Pathway Engineering of Diterpenoid Synthases and the Mevalonic Acid Pathway for Miltiradiene Production", JOURNAL OF THE AMERICAN CHEMICAL SOCIETY, vol. 134, no. 6, 15 February 2012 (2012-02-15), pages 3234-3241, XP055058678, ISSN: 0002-7863, DOI: 10.1021/ja2114486 abstract	1-98
A	DE 10 2009 025996 A1 (UNIV DES SAARLANDES CAMPUS SAA [DE]) 23 December 2010 (2010-12-23) claim 1	1-98
A	LARISSA M. PODUST ET AL: "Diversity of P450 enzymes in the biosynthesis of natural products", NATURAL PRODUCT REPORTS, vol. 29, no. 10, 1 January 2012 (2012-01-01), page 1251, XP055215367, ISSN: 0265-0568, DOI: 10.1039/c2np20020a the whole document	1-98

-/--

1

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	GONG HAI-YAN ET AL: "Diterpene synthases and their responsible cyclic natural products", NATURAL PRODUCTS AND BIOPROSPECTING, SPRINGER OPEN, GERMANY , vol. 4, no. 2 1 April 2014 (2014-04-01), pages 59-72, XP002739034, ISSN: 2192-2209, DOI: 10.1007/S13659-014-0012-8 Retrieved from the Internet: URL:http://rd.springer.com/article/10.1007%2Fs13659-014-0012-8 [retrieved on 2014-04-18] the whole document	1-98

X,P	Johan Andersen-Ranberg: "Identification and Characterization of Biosynthetic Parts Involved in Plant Diterpenoid Biosyntheses - Employees", , 24 June 2014 (2014-06-24), XP055183341, University of Copenhagen Retrieved from the Internet: URL:http://plen.ku.dk/english/employees/?pure=en%2Fpublications%2Fidentification-and-characterization-of-biosynthetic-parts-involved-in-plant-diterpenoid-biosyntheses(9b672fe8-15a1-4685-8204-ce711d76b041).html [retrieved on 2015-04-15] abstract	1-98

X,P	DRAGANA BOZIC ET AL: "Towards Elucidating Carnosic Acid Biosynthesis in Lamiaceae: Functional Characterization of the Three First Steps of the Pathway in Salvia fruticosa and Rosmarinus officinalis", PLOS ONE, vol. 10, no. 5, 28 May 2015 (2015-05-28), page e0124106, XP055215828, DOI: 10.1371/journal.pone.0124106 abstract; figure 1	1-98

E	WO 2015/113570 A1 (UNIV COPENHAGEN [DK]; UNIV DANMARKS TEKNISKE [DK]) 6 August 2015 (2015-08-06) abstract	1-98

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/DK2015/050181

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
DE 102009025996 A1	23-12-2010	NONE	
WO 2015113570 A1	06-08-2015	NONE	