



(51) International Patent Classification:

C07K 14/415 (2006.01) C12N 15/79 (2006.01)
C12N 9/88 (2006.01) C12N 15/82 (2006.01)

(21) International Application Number:

PCT/EP2015/065218

(22) International Filing Date:

3 July 2015 (03.07.2015)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

14175719.5 4 July 2014 (04.07.2014) EP

(71) Applicant: EELCO WALLAART B.V [NL/NL]; Duurswoldlaan 22, 9727 DJ Groningen (NL).

(72) Inventor: WALLAART, Thorvald Eelco; Duurswoldlaan 22, 9727 DJ Groningen (NL).

(74) Agent: VAN KOOLIJ, Adriaan; Arnold + Siedsma, Bezuidenhoutseweg 57, 2594 AC Den Haag (NL).

(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:

- as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))
- of inventorship (Rule 4.17(iv))

Published:

- with international search report (Art. 21(3))

(54) Title: GENE AND POLYPEPTIDE INVOLVED IN VALENCENE SYNTHESIS AND USES THEREOF

(57) Abstract: The present invention relates to an isolated nucleic acid sequence with a nucleic acid sequence having 70%, preferably 80%, more preferably 85%, most preferably 90% nucleotide sequence identity with SEQ ID NO. 1. Specifically, the present invention relates to a gene, and the polypeptide encoded, involved in the valencene synthesis.



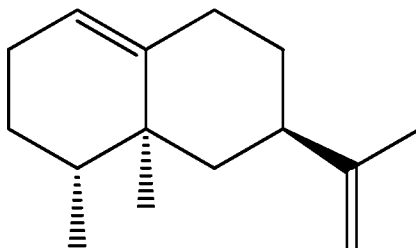
WO 2016/001412 A1

GENE AND POLYPEPTIDE INVOLVED IN VALENCENE SYNTHESIS AND USES THEREOF

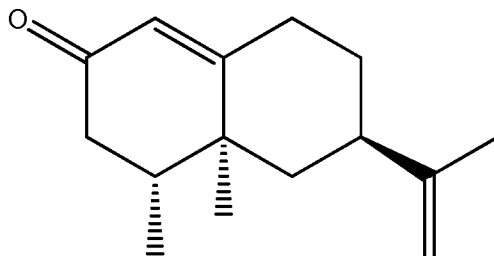
Description

The present invention relates to a nucleic acid sequence involved in the synthesis of valencene. Further, the present invention relates to a polypeptide, or enzyme, capable of converting farnesyl pyrophosphate (FPP) into valencene. Furthermore, the present invention relates to the use of the present nucleic acid sequence and polypeptide for the production of valencene, or its derivative nootkatone.

Valencene (1,2,3,5,6,7,8,8a-octahydro-7-isopropyl-1,8a-dimethyl-naphthalene) or represented by



and nootkatone (4,4a,5,6,7,8-hexahydro-6-isopropenyl-4,3a-dimethyl-2(3H)-naphthalenone) or represented by



are belonging to the group of sesquiterpenes and are natural constituents of citrus oils such as orange and grapefruit.

The flavour and aroma of citrus species is composed of complex combinations of mostly sugars, flavonoids, acids and volatile oils. Mono- and sesquiterpenes are the main components of this volatile, also called essential, oil. Although the monoterpene limonene normally accounts for over 90% of the content of essential oils obtained from citrus species,

several unique sesquiterpene compounds, which are present in very low quantities, have an intense effect on the flavor and aroma of the citrus species.

The sesquiterpenes valencene, nootkatone, alpha- and beta-sinensal, are examples of compounds present in minor quantities in oranges and having an important role in the overall flavor and aroma of citrus fruit. Valencene is obtained by the enzymatic activity of valencene synthase and is a major component of citrus essential oils. Nootkatone is a putative-derivative of valencene and can be produced by enzymatic or by synthetic conversion of valencene. Nootkatone, occupies a small part of the essential oil but has a dominant role in the flavor and aroma of grapefruit.

Because of their outstanding organoleptic characteristics and typical taste, valencene and nootkatone are extensively used as an aroma component of citrus fruit and citrus-derived odorants. Valencene is an important compound for the flavour and fragrance (F&F) industry. It is applied in beverage and food. It has also an important place in the cosmetic industry.

Valencene is present in a number of citrus plant species like grapefruit *Citrus paradisi* and orange *Citrus sinensis* but also in non citrus fruit plants like *Perilla frutescens* (beefsteak-mint) and *Vitis vinifera* L. (wine grape).

Valencene and nootkatone are produced in plants as an secondary metabolite and as most secondary metabolites they are produced only in relatively small quantities. For industrial applications natural valencene is isolated from Valencia oranges *Citrus sinensis* cv. Valencia. Due to small quantities present in these oranges natural valencene is quite expensive. Natural nootkatone, typically isolated from grapefruit, is even more expensive than valencene.

Biotechnological production of these compounds could be a rational approach to obtain natural valencene and nootkatone in big quantities and at lower costs. A valencene synthase with a high product specificity is a prerequisite for such an approach. However, not all valencene synthase enzymes have a very high product specificity, which is of paramount importance for biotechnological production of pure compounds.

Considering, amongst others, the economical importance of valencene and nootkatone, there is a need in the art to provide a novel gene and polypeptide involved in the valencene biosynthesis.

The above need of the art, amongst other needs, is met by the present invention by providing a novel gene involved in the valencene biosynthesis as outlined in the appended claim 1.

Specifically, the above need of the art, amongst other needs, is met by the present invention, according to a first aspect, by providing an isolated nucleic acid sequence with a nucleic acid sequence having 70%, preferably 80%, more preferably 85%, most preferably 90% nucleotide sequence identity with SEQ ID NO. 1. According to a further preferred embodiment, the present

invention relates to an isolated nucleic acid sequence with a nucleic acid sequence having 95% nucleotide sequence identity with SEQ ID NO.1.

Sequence identity, as used herein, is defined as the number of identical nucleic acids, over the full length of the present sequences, divided by the number of nucleic acids, of the full length and multiplied by 100%.

According to a further preferred embodiment of the present invention, the present isolated nucleic acid sequence has a nucleic acid sequence of SEQ ID NO. 1.

The present invention is based on the identification and isolation of a gene which is involved in the valencene synthesis. This gene is identified in the plant *Coleus forskohlii* and encodes valencene synthase. Surprisingly, the valencene synthase encoded by the nucleic acid sequence according to the present invention produces nearly exclusive valencene. Due to this very high product specificity, the present polypeptide, or enzyme, provides advantageous properties for biotechnological production of natural valencene and its putative derivative nootkatone.

Accordingly, according to a second aspect, the present invention relates to an isolated polypeptide, or enzyme, encoded by the isolated nucleic acid sequence of the present invention.

According to a preferred embodiment of the above second aspect, the present invention relates to isolated polypeptides, or enzymes, having 70%, preferably 80%, more preferably 85%, most preferably 90% or 95% amino acid sequence identity with SEQ ID NO. 2.

Sequence identity, as used herein, is defined as the number of identical amino acids, over the full length of the present sequence, divided by the number of amino acids, of the full length and multiplied by 100%.

According to an especially preferred embodiment of the above second aspect, the present invention relates to an isolated polypeptide, or enzyme, having an amino acid sequence of SEQ ID NO.2.

The present nucleic acids can be used to recombinantly express, or provide, the present polypeptide, or enzyme. Accordingly, according to a third aspect, the present invention relates to expression vectors comprising the present isolated nucleic acid sequence.

Generally, an expression vector according to the present invention will comprise, besides the present nucleic acid, transcription and translation regulating sequences suitable for expression of the present nucleic acids in a host organism desired.

Expression vectors generally also comprise selection and/or marker sequences. A variety of expression vectors for insect, mammal, plant, yeast, bacterium and/or fungus expression of the present nucleic acids are commercially available.

According to a preferred embodiment of the above third aspect, the present invention relates to an expression vector which further comprises a gene construct encoding for a

polypeptide selected from the group consisting of CoA carboxylase, hydroxyl-methyl-glutaryl CoA reductase, squalene synthase, acetoacetyl-CoA thiolase, 3-hydroxy-3-methyl-glutaryl-CoA synthase, 3-hydroxy-3-methyl-glutaryl-CoA reductase, mevalonate kinase, mevalonate phosphate kinase, mevalonate diphosphate decarboxylase, 1-deoxy-D-xylulose 5-phosphate synthase, 1-deoxy-D-xylulose 5-phosphate reductoisomerase, 4-diphosphocytidyl-2-C-methyl-D-erythritol synthase, 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase, 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase, (E)-4-hydroxy-3-methyl-but-2-enyl pyrophosphate synthase, (E)-4-hydroxy-3-methyl-but-2-enyl pyrophosphate reductase, isopentenyl pyrophosphate isomerase, geranyl pyrophosphate synthase and/or farnesyl pyrophosphatesynthase. This embodiment is particularly advantageous since the polypeptides, or enzymes, provide the production of a valencene precursor, preferably farnesyl pyrophosphate, in the host cell of interest. This valencene precursor can be converted into valence by the valencene synthase according to the present invention.

According to an especially preferred embodiment, the present expression vector further comprises a gene construct encoding for the cytochrome P450, or CYP. This embodiment is particularly advantageous if the production of nootkatone is desired since P450 provides the enzymatic conversion of the valence produced to nootkatone. Accordingly, one expression vector provides the genetic information needed for the production of nootkatone.

According to a fourth aspect, the present invention relates to transgenic organisms comprising the present expression vector or the present nucleic acid sequence. The organisms according to the present invention are preferably selected from the group consisting of plant, bacterium, fungus, such as *Mortierella isabellina* and yeast.

According to a preferred embodiment of the above fourth aspect, the present plant is selected from the group consisting of the genera Carum, Cichorium, Daucus, Juniperus, Chamomilla, Lactuca, Pogostemon, Vetiveria, Capsicum, Gossyium, Lycopersicon, Nicotiana, Phleum, Solanum and Ulmus or is selected from *Glycine max*, *Helianthus annuus* and *Brassica napus*. It is advantageous to use a plant already producing sesquiterpenes, as these plants already have the basic pathway and storage compartments available. For the production of nootkatone, a plant that already is producing nootkatone or is able to catalyse the conversion of valencene into nootkatone is advantageous because the enzyme(s) catalysing the conversion of valencene into nootkatone is already present.

According to yet another preferred embodiment of the above fourth aspect, the present yeast is selected from the group consisting of *Yarrowia lipolytica*, *Cryptococcus curvatus*, *Rhodospiridium toruloides*, *Lipomyces starkeyi*, *Rhodotorula glutinis*, *Geotrichum robustum* and *Saccharomyces cerevisiae* and *Pichia pastoris*. The advantage of using oleaginous yeast host cells

since they have the capacity to accumulate up to about 50% (dry weight) of storage carbohydrates in oil bodies, providing the production of large quantities of terpenes.

The present nucleic acid sequence and the present polypeptide, or enzyme, according to the invention, as outlined above, provide the production of valencene and/or
5 nootkatone.

Accordingly, according to a fifth aspect, the present invention relates to the use of the present nucleic acid, polypeptide, expression vectors or transgenic organism, as outlined above, for the production of valencene and or nootkatone.

Given that the advantageous gene according to the present invention is identified
10 in *Coleus forskohlii*, the present invention relates also to the use of the plant *Coleus forskohlii* for the production of valencene and or nootkatone.

The present invention will be further elucidated, and detailed, in the following non limiting examples. In the examples, reference is made to figures wherein:

15 **Figure 1:** shows valencene synthase encoding cDNA part between start and stop codon flanked by *NcoI* and *BamHI* sites, respectively, subcloned in the *NcoI/BamHI* site of the expression vector pET 11d.

Figure 2: SDS-PAGE gel: lanes 1 – 5 are all showing the pellets fractions: lane 1
20 and 2 valencene synthase-pET11d construct in BL21(DE3), lane 3 and 4 amorpha-4,11-diene synthase-pET11d construct in BL21(DE3) (positive control). Lane 5 pET11d without insert in BL21(DE3) (negative control). M: PageRuler™ 10-200 KDa ladder (Fermentas). Lanes 6 – 10 are all showing the supernatant fractions: lane 6 and 7 valencene synthase-
25 pET11d construct in BL21(DE3), lane 8 and 9 amorpha-4,11-diene synthase-pET11d construct in BL21(DE3) (positive control). Lane 10 pET11d without insert in BL21(DE3) (negative control). Amorpha-4,11-diene synthase-pET11d construct in BL21(DE3) (positive control) and valencene synthase-pET11d construct in BL21(DE3) showed a intens spot
30 which was not present in the negative control pET 11d.

Figure 3: Upper part: Flame Ionization Detector (FID) signals of the FPP assay with terpene synthases in expression vector pET11d in BL21(DE3).
lower part: Mass spectrum of the FPP assay with terpene synthases in
35 expression vector pET11d in BL21(DE3).

Figure 4: Mass spectrum of reference valencene compared with mass spectrum of the FPP assay with the terpene synthases of the invention (SEQ ID NO.2) in expression vector pET11d in BL21(DE3). This comparison yielded a quality score of 99%, corresponding with a maximum score of identicalness.

Figure 5: Schematic representation of pSYL092. Valencene synthase flanked by the TEF promoter and XPR2 terminator of *Y lipolytica* and with the URA3 auxotrophic selection marker.

Figure 6: Schematic representation of pSYL094. tHMG-CoA encoding gene flanked by the TEF promoter and XPR2 terminator of *Y lipolytica* and with the Leu2 auxotrophic selection marker.

EXAMPLES

Example 1

Isolation and Characterization of the Valencene Synthase Encoding Gene

A. Isolation of Total RNA

Total RNA was isolated from the roots of *Coleus forskohlii* by using the NucleoSpin® RNA Plant kit from Macherey & Nagel (Düren, Germany) according to the manufacturers instructions.

B. cDNA Synthesis

The reverse transcription reaction was carried out in a 20 µl reaction containing 2,5 µg total RNA, 0.2 µg Oligo dT primer (SEQ ID NO.10); 5'-AAT CGA TAG GCC GAG GCG GCC GAT CAG-(T)28-(A/C/G)-(A/T/C/G)-3', 0.5 mM each dATP, dTTP, dCTP and dGTP, 10 mM DTT, 2 U "RiboLock™" ribonuclease inhibitor (Fermentas), first strand synthesis buffer (Fermentas) and catalyzed with 200 U "RevertAid™" Moloney murine leukemia virus (M-MLV) reverse transcriptase RNase H minus (Fermentas). After 1 h incubation at 37°C. the reaction was stopped by placing the reaction mixture at -20°C.

C. PCR-based Sequence Generation

The terpene synthases of the biosynthetic pathways leading to production of mono-, sesqui-, and di-terpenes are ubiquitous to all plant species. Although sequence conservation is not very high among terpene synthases of different plant species, discrete conserved domains are present. These conserved domains have been the basis for isolation of a number of terpene synthases encoding genes from a variety of plant species using degenerate-primer based RT-PCR. Based on comparison of sequences of terpenoid synthases, two degenerated primers were designed for two conserved regions. The degenerated primers were designed in such a way that they have high homology with sesquiterpene synthase encoding genes. The sequence of the degenerated sense primer (primer A) was 5'-GA(C/T) GA(G/A) AA(C/T) GGI AA(G/A) TT(C/T) AA(G/A) GA-3' (SEQ ID NO.11) and the sequence of the degenerated anti sense primer (primer B) was 5'-CC (G/A)TA IGC (G/A)TC (G/A)AA IGT (G/A)TC (G/A)TC-3' (SEQ ID NO.12).

PCR was performed in a total volume of 100 µl containing 0.5 µM of each of these two primers, 0.2 mM each dNTP, 1.5 U Pfu DNA polymerase/1x Pfu PCR buffer (Fermentas) and 2 µl cDNA. The reaction was incubated in a thermocycler (Biometra T-gradient) with 45 seconds denaturation at 94°C, 45 seconds annealing at 40°C and 1 minute and 15 seconds elongation at 68°C during 40 cycles.

Agarose gel electrophoresis revealed a single specific PCR product of approximately 550 bp (530 bp). A 3' A overhang was made by using Taq DNA polymerase (Fermentas). The sticky DNA was "TA" subcloned in pGEM-Teasy (Promega) and E.coli DH5-alpha (Invitrogen) was transformed with this construct. The inserts of 5 individual transformants were sequenced and they all showed the same sequence, with exception for the primer parts, as shown in SEQ ID NO.3

D. RACE PCR and subcloning

3'RACE

The by PCR obtained sequence was used to design a specific sense (primer C) and anti sense (primer D) primer. The sequence of the sense primer C was 5'-CCA GGG GTT TGC TGA GTC TGT A-3' (SEQ ID NO.13) and the sequence of the anti sense primer D was 5'-GCG ATG GTC TTG GTG AGC ATG AT-3' (SEQ ID NO.14). The 3' part of the mature cDNA was amplified by PCR with the sense primer C and the oligo dT primer (5'-AAT CGA TAG GCC GAG GCG GCC GAT CAG-(T)28-(A/C/G)-(A/T/C/G)-3') as the anti sense primer.

PCR was performed in a total volume of 50 µl containing 0.5 µM of each of these two primers, 0.2 mM each dNTP, 1.5 U Pfu DNA polymerase/1x Pfu PCR buffer (Fermentas) and 2 µl cDNA. The reaction was incubated in a thermocycler (Biometra T-gradient) with 45 seconds

denaturation at 92°C, 45 seconds annealing at 58°C and 1 minute and 30 seconds elongation at 68°C during 40 cycles.

Agarose gel electrophoresis revealed a major specific PCR product of approximately 1200 bp (1359 bp). A 3'A overhang was introduced by using Taq DNA polymerase (Fermentas). The obtained sticky DNA was "TA" subcloned in pGEM-Teasy (Promega) and E.coli DH5-alpha (Invitrogen) was transformed with this construct. The identity of these clones was checked by colony PCR with the primers C and D yielding a PCR product of approximately 450bp. The insert was sequenced and revealed a sequence, as shown in SEQ ID NO.4.

10 5'RACE

To clone the encoding DNA sequence of the 5' part of the mRNA, new cDNA was synthesised with primer E (5'-GCCAACTCTTGTTCAAATTGAT-3') (SEQ ID NO.15) and under the conditions as described under B "cDNA Synthesis". The obtained cDNA was tailed by Terminal deoxynucleotidyl Transferase (TdT) yielding a poly A tail at the 3'part of the cDNA. The poly A tailing was performed in a total volume of 30 µl containing 20 µl of approximately 1 pmol cDNA, 6 µl 5x TdT buffer (Fermentas), and 3 µl 2mM dATP. After heat incubation for 3 minutes at 94°C followed by incubation at ice, 1 µl (20U) TdT (Fermentas) were added. The reaction mixture was incubated on ice for 60 minutes. Again 1 µl (20U) TdT was added and the reaction mixture was incubated at 37°C for 30 min. The reaction was stopped by incubating at 70°C for 10 minutes. The 3' part of the cDNA was amplified by nested PCR. The first PCR was performed with the primer combination; anti sense primer F (5'-CGGTGAAAACATCAAGTTCTTCAA-3') (SEQ ID NO.16) and the oligo dT primer, as sense primer.

The PCR was performed in a total volume of 50 µl containing 0.5 µM of each of these two primers, 0.2 mM each dNTP, 1.5 U Pfu DNA polymerase/1x Pfu PCR buffer (Fermentas) and 2 µl poly A tailed cDNA. The reaction was incubated in a thermocycler (Biometra T-gradient) with 45 seconds denaturation at 92°C, 45 seconds annealing at 58°C and 1 minute and 30 seconds elongation at 68°C during 40 cycles.

This yielded a product as SEQ ID NO.6 In the second PCR, the PCR-product of the first PCR SEQ ID NO.6 was used as template. As primers for this PCR primer D (upstream from primer F) and the oligo dT primer where used. This nested PCR was performed in a total volume of 50 µl containing 0.5 µM of each of these two primers, 0.2 mM each dNTP, 1.5 U Pfu DNA polymerase/1x Pfu PCR buffer (Fermentas) and 2 µl PCR DNA of the previous PCR as template. The reaction was incubated in a thermocycler (Biometra T-gradient) with 45 seconds denaturation at 92°C, 45 seconds annealing at 58°C and 1 minute and 30 seconds elongation at 68°C during 40 cycles. Agarose gel electrophoresis revealed a major PCR product of approximately 1000 bp (1044 bp). A 3'A overhang was introduced by using Taq DNA polymerase

(Fermentas). The obtained sticky DNA was “TA” subcloned in pGEM-Teasy (Promega) and E.coli DH5-alpha (Invitrogen) was transformed with this construct. The insert was sequenced and revealed a sequence, as shown in SEQ ID NO.8.

5 E. Amplification of full length cDNA clone

Based on the sequence data obtained from the 5' and 3' RACE PCR, primer G (5'-CCATGGCTCAAGTGCAATCGGAAATTG-3') (SEQ ID NO.17), with an NcoI site (underlined) introduced at the start codon ATG (Italic), and primer H (5'-

GGATCCTAAATTAGAATGGGGTCGACGAAC-3') (SEQ ID NO.18) with a BamHI site

10 (underlined) introduced at the stop codon TAG (Italic), were designed. With these primers a PCR was performed in a total volume of 50 µl containing 0.5 µM of each of these two primers, 0.2 mM each dNTP, 1.5 U Pfu DNA polymerase/1x Pfu PCR buffer (Fermentas) and 2 µl cDNA. The reaction was incubated in a thermocycler (Biometra T-gradient) with 45 seconds denaturation at 92°C, 45 seconds annealing at 58°C and 1 minute and 30 seconds elongation at 68°C during 40
15 cycles.

Agarose gel electrophoresis revealed a major specific PCR product of approximately 1200 bp (1359 bp). A 3'A overhang was introduced by using Taq DNA polymerase (Fermentas). The obtained sticky DNA was “TA” subcloned in pGEM-Teasy (Promega) and E.coli DH5-alpha (Invitrogen) was transformed with this construct. The insert was sequenced and
20 revealed a sequence, as shown in SEQ ID NO.1.

Example 2

Expression of the Valencene Synthase Encoding Gene in E.coli BL21 (DE3)

For functional expression the cDNA clone was subcloned, in frame, into the
25 expression vector pET 11d (Stratagene). After digestion with BamHI and NcoI of the PCR product of SEQ ID NO.1 and the expression vector pET 11d both were gel purified and ligated together to yield a construct as revealed in figure 1.

To obtain expression, this gene construct (figure 1), pET 11d without an insert as **negative control**, and pET 11d with the *Artemisia annua* L. amorpha-4,11-diene synthase (ADS) gene (Wallaart et al., Planta 212(3); 460-465 (2001)); as **positive control** were transformed to
30 *E.coli* BL21 (DE3) (Stratagene), and grown overnight on LB agar plates supplemented with ampicillin at 37°C. Cultures of 50 ml LB medium supplemented with ampicillin (100 [mu]g/ml) and 0.25 mM isopropyl-1-thio-[beta]-D-galactopyranoside (IPTG) were inoculated with these over night cultures to A600=0.5 and grown for 3 h at 27°C. The cells were harvested by centrifugation
35 during 8 minutes at 2000 g and resuspended in 2 ml assay buffer (10% glycerol, 10mM MgCl₂,

1mM Na ascorbate, 15mM MOPSO (pH 7,0), and 2mM DDT). An aliquot of 1 ml resuspended cells was sonicated on ice four times for 5 seconds with 30 second intervals at a duty output of 40%, and centrifuged for 5 minutes at 4°C in a microfuge at 13.000 rpm.

The obtained supernatant was used for terpene synthase enzyme activity determinations. Supernatant and pellet of the valencene synthase-pET 11d construct in E.coli BL21 (DE3) yielded on SDS-PAGE gel a protein band of approximately 50 to 60 kDa. This was about the same size as seen for, the positive control, amorpha-4,11-diene synthase. The pET 11d vector without an insert, the negative control, did not reveal such a protein band on the SDS-PAGE gel, as is shown in figure 2.

Example 3

Conversion of FPP into Valencene by Valencene Synthase Expressed in E.coli

Besides the supernatant of sonicated cells, also intact cells were used in the FPP assay. Farnesyl pyrophosphate (FPP) was purchased from (Sigma-Aldrich) as an ammonium salt solved in a methanol / water mixture (MeOH : H₂O (10mM NH₄OH) = 7 : 1). The methanol / water was removed by vacuum drying in a speed vac and the FPP salt was resolved in 200 µl of an ethanol / water mixture (EtOH : H₂O (200 mM NH₄HCO₃) = 1 : 1). To a total volume of 1 ml supernatant (EXAMPLE 2) 50 µM FPP was added (10.9 µl). After the addition of a 1mL redistilled pentane overlay to trap volatile products, the tubes were carefully mixed and incubated for 1 h at 30°C. Boiled samples were used as controls. Following the assay, the tubes were vigorously mixed. The organic layer was removed and passed over a short column of aluminum oxide (grade III) overlaid with anhydrous MgSO₄. The assay was extracted with another 1 mL of 20% (v/v) diethyl ether in pentane which was also passed over the aluminum oxide column, and the column washed with 1.5 mL of 20% (v/v) diethyl ether in pentane. For GC-MS analysis, the pooled pentane/diethyl-ether mixture was slowly concentrated under a stream of N₂ gas.

GC-MS analyses of the concentrated pentane/diethyl-ether mixtures were performed using a HP 5890 series II GC and HP 5972A Mass Selective Detector (Hewlett-Packard) equipped with an HP-5 MS or HP-Innowax column (both 30 m x 0.25 mm i.d., 0.25 µm df). The oven was programmed at an initial temperature of 70°C. for 1 min, with a ramp of 5°C. min⁻¹, to 210°C. and final time of 5 min. The injection port (splitless mode), interface and MS source temperatures were 175, 290 and 180°C., respectively, and the He inlet pressure was controlled by electronic pressure control to achieve a constant column flow of 1.0 ml min⁻¹. Ionization potential was set at 70 eV, and scanning was performed from 30-250 amu. GC-MS analysis on the two different GC-columns of FPP assay products by the sonicated preparation

showed that the main product had a mass spectrum and retention time equal to that of the reference valencene (see figure 3 and figure 4).

GC-MS analysis of the FPP assay with intact (not sonicated) transformed cells revealed the same GC-MS chromatogram as the supernatant of sonicated transformed cells did. In both assays valencene was produced. Identification of these assay products with the GC-MS gave a mass-spectrum identical to the mass-spectrum of the reference valencene (Fluka) with a quality score of 99% (maximum score), see figure 4. No valencene was found in assays performed with the positive and negative controls.

Example 4

Expression of valencene Synthase in Transgenic Tobacco

There are many ways to introduce DNA into a plant cell. Suitable methods by which DNA can be introduced into the plant cell include *Agrobacterium* infection or direct delivery of DNA such as, for example, by PEG-mediated transformation of protoplasts (Omirulleh et al., Plant Molecular Biology 21, 415-428 (1993)) or electroporation, by acceleration of DNA coated microprojectiles (for example, microprojectile bombardment) microinjection, etc.

Because *Agrobacterium tumefaciens*-mediated transformation of *Nicotiana tabacum* with a sesquiterpene cyclase gene is known in literature (Hohn and Ohlrogge, Plant Physiol. 97, 460-462 (1991)) delivery of expression units (cassettes), containing the valencene synthase encoding gene, mediated by *Agrobacterium* seemed to be a rational approach. In particular because it is not known in literature that "wild type" tobacco is producing valencene by its self.

There are several binary vector systems suitable to transfer the amorphadiene synthase encoding gene assembled in an expression cassette behind a suitable promoter (for example, the cauliflower mosaic virus (CaMV) 35S promoter) and upstream of a suitable terminator (for example, the nopaline synthase transcription terminator (nos-tail)) to tobacco.

Analogous to example 1, suitable restriction sites for subcloning were introduced by using PCR with a sense primer I 5'-GCGGATCCATGGCTCAAGTGCAATCGGAAATTG -3' (SEQ ID NO.19) introducing a *Bam*HI site (underlined) preceding the start codon (Italic) and an anti-sense primer J 5'-GCGGATCCTCTAAATTAGAATGGGGTCGACGAAC-3' (SEQ ID NO.20) (introducing a *Bam*HI site (underlined) directly behind the stop codon (Italic). After digestion with *Bam*HI the PCR product and the plant-expression cassette pLV399 were gel purified and ligated to provide the gene encoding valencene synthase with the cauliflower mosaic virus 35S promoter and a nopaline synthase transcription terminator. The plant-expression cassette pLV399 is a pUC 19 vector (Yanisch-Perron, C. et al., Gene 33, 103-119 (1985)) in which the multiple

cloning site (polylinker) is replaced by a CaMV 35 S promoter *Bam*HI fused to a nos-tail (terminator) flanked by the 'unique' sites; *Eco*RI, *Kpn*I, *Xho*I, and a *Hind*III site downstream from the promoter and *Eco*RI, *Xho*I, *Pst*I, *Sph*I, *Kpn*I, *Hind*III upstream from the terminator. The orientation of the valencene synthase encoding gene in pLV399 was checked by restriction analysis with *Sph*I. After digestion of this construct with *Kpn*I the valencene synthase encoding gene flanked by the 35S promoter and nos terminator was ligated into the *Kpn*I digested binary vector pCGN1548.

To mobilize the recombinant binary vector to *Agrobacterium tumefaciens* LBA4404 (Gibco BRL, Life Technologies), a triparental mating procedure was carried out by using *E.coli* (DH5[alpha]) carrying the recombinant binary vector and a helper *E.coli* carrying the plasmid pRK2013 to mobilize the recombinant binary vector to *A. tumefaciens* LBA4404.

This transformed *Agrobacterium* strain was used for transformation of explants from the target plant species. Only the transformed tissue carrying a resistance marker (kanamycin-resistance, present between the binary plasmid T-DNA borders) regenerated on a selectable (kanamycin containing) regeneration medium. (According to Rogers S G, Horsch R B, Fraley R T Methods Enzymol (1986)118: 627-640).

The plants regenerated out of the transformed tissue expressed the valencene synthase gene as followed from the presence therein of valencene as confirmed by GC-MS analyses.

Example 5

Expression of the Valencene Synthase Encoding cDNA clone in Yarrowia lipolytica

For expression of the valencene synthase encoding cDNA in *Yarrowia lipolytica* expression signals like a promoter and terminator sequence were cloned. The XPR2 terminator was amplified by PCR using primer K (5'-CACAAACTAGTTTGCCACCTACAAGCCAGAT-3') (SEQ ID NO.21) and primer L (5'-CACAAGCGGCCGCACGCGTGCAATTAACAGATAGTTTGCCGGTGATAATTC -3') (SEQ ID NO.22). Genomic DNA from wild type strain *Yarrowia lipolytica* SYL291 was isolated using the G-nome DNA kit (MP Biomedicals, Solon, USA) according to the manufactures protocol. The PCR was performed in a total volume of 50 µl containing 0.5 µM of each of these two primers, 0.2 mM of each dNTP, 1.5 U Pfu DNA polymerase/1x Pfu PCR buffer (Fermentas) and 2 µl genomic DNA of wild type strain *Yarrowia lipolytica* SYL291. The reaction was incubated in a thermocycler (Biometra T-gradient) with an initial denaturation step of 2 minutes at 92°C followed

by 40 cycles of 45 seconds denaturation at 92°C, 45 seconds annealing at 58°C and 1 minute and 30 seconds elongation at 68°C.

Agarose gel electrophoresis revealed a specific PCR product of approximately 200 bp. A 3'A overhang was introduced by using Taq DNA polymerase (Fermentas). The obtained sticky DNA was "TA" subcloned in pGEM-Teasy (Promega Corporation, Madison, USA) yielding pSYL076.

In analogy with the XPR2 terminator, the TEF promotor was amplified using the primers M (5'-CACAAGGTACCAGAGACCGGGTTGGCGGCGCATTTGTGTC-3') (SEQ ID NO.23) and N (5'-

CACAAGCGGCCGCGCTAGCGAATGATTCTTATACTCAGAAGGAAATG-3') (SEQ ID NO.24) yielding a product of approximately 400 bp. This product was also TA cloned in pGEM-Teasy (Promega Corporation, Madison, USA) yielding pSYL077.

The XPR2 terminator flanked by NotI and SpeI and subcloned into pGEM-Teasy pSYL076 was cut with NotI and SpeI yielding a 200 bp fragment containing the XPR2 terminator.

The TEF promoter flanked by and KpnI and NotI and subcloned into pGEM-Teasy pSYL077 was cut with KpnI and NotI yielding a 400 bp fragment containing the TEF promotor. Both fragments were ligated in a 3 points ligation in KpnI and SpeI cut pBluescriptII KS (-) yielding vector pSYL083.

The *Coleus forskohlii* valencene synthase encoding cDNA from example 1 (SEQ ID NO.1) was amplified by PCR as described above using the sequence of SEQ ID NO.1 as template with primers O (5'-CACACTCTAGACACAAAAATGGCTCAAGTGCAATCGGAAATTG-3') (SEQ ID NO.25) and P (5'-CACACACGCGTACACCTAAATTAGAATGGGGTCGACGAAC-3') (SEQ ID NO.26) yielding a product of approximately 1700 bp.

This product was also TA cloned in pGEM-Teasy (Promega Corporation, Madison, USA) yielding pSYL085.

This vector (pSYL085) was cut with XbaI and MluI yielding a fragment of approximately 1500 bp containing the valencene synthase encoding cDNA. The TEF promoter and XPR2 terminator containing vector pSYL083 was cut open with MluI and NheI. The valencene synthase encoding cDNA flanked by XbaI and MluI was ligated into MluI and NheI opened pSYL083. Yielding a valencene synthase encoding cDNA flanked by the TEF promoter and XPR2 terminator in pBluescriptII KS (-) pSYL086.

For selection purposes the gene of the auxotrophic selection marker "3-isopropylmalate dehydratase" (Leu2) was together with its own promoter and terminator amplified by PCR as described above with the primers Q (5'-CATCTCTAGACTACCCGTTGCTATCTCCACAC-3') (SEQ ID NO.27) and R (5'-

CACACTCTAGAATCGATCGTCTAACGGACTTGATATACAAC-3') (SEQ ID NO.28) and genomic DNA from wild type strain *Yarrowia lipolytica* SYL291 as template. The product was ligated in pGEM-T easy (Promega Corporation, Madison, USA) yielding pSYL090.

The Leu2 selection marker containing vector pSYL090 and the valencene synthase encoding cDNA flanked by the TEF promoter and XPR2 terminator pSYL086 were cut with XbaI. The 2.5 kb fragment from pSYL090 was ligated in pSYL086 yielding pSYL092 see figure 5.

A well known approach, in literature, to improve the flux of the mevalonate pathway is the introduction of truncated HMG CoA reductase (tHMG-CoA red, SEQ ID NO.11) in a terpenoids producing organism. For this purpose primers were designed to amplify a part of the HMG CoA reductase gene to obtain truncated HMG CoA reductase by PCR. The PCR was performed as described above with genomic DNA isolated from wild type strain *Yarrowia lipolytica* SYL291 as template. Primers used in this reaction were primer S (5'-CACACTCTAGACACAAAAATGACCCAGTCTGTGAAGGTGGTTGAGAAG-3') (SEQ ID NO.29) and primer T (5'-CACACACGCGTACACCTATGACCGTATGCAAAT-3') (SEQ ID NO.30) yielding a PCR product of about 1400 bp. The PCR product was ligated in pGem-T easy (Promega Corporation, Madison, USA) according to the manufactures protocol yielding pSYL075.

To flank the cloned tHMG-CoA red gene with a promoter and terminator the tHMG-CoA red gene was cut out pSYL075 with XbaI and MluI and ligated into MluI, NheI cut pSYL083 yielding pSYL087.

For selection purposes the gene of the auxotrophic selection marker "orotidine 5-phosphate decarboxylase" (URA3) was together with its own promoter and terminator amplified by PCR as described above with the primers U (5'-CATCTCTAGAGGTGTGTTCTGTGGAGCATTC-3') (SEQ ID NO.31) and V (5'-CACACTCTAGAATCGAGGTGTAGTGGTAGTGCAGTG-3') (SEQ ID NO.32) and genomic DNA from wild type strain *Yarrowia lipolytica* SYL291 as template yielding a PCR product of approximately 1900 bp. The URA3 fragment was ligated in pGEM-T easy (Promega Corporation, Madison, USA) yielding pSYL091.

The Ura3 selection marker containing vector pSYL091 and the tHMG-CoA encoding gene flanked by the TEF promoter and XPR2 terminator pSYL087 were cut with XbaI. The 1.9 kb fragment from pSYL091 was ligated in pSYL087 yielding pSYL094, see figure 6.

Yarrowia lipolytica SYL312 (Δ ura3, Δ leu2) was transformed with linearized pSYL092 according to Chen et al., 1997 yielding *Yarrowia lipolytica* SYL320 (valencene, Δ leu2). Possible transformants were checked by PCR with the primers O and P based on the valencene synthase encoding cDNA. Stability of the transformants was checked via culturing without selection pressure (YPD plates containing uracil and leucine) followed by culturing on YNB plates without uracil. *Yarrowia lipolytica* SYL320 (valencene, Δ leu2) was transformed with linearized

pSYL094 according to Chen et al., 1997 yielding *Yarrowia lipolytica* SYL321 (valencene, tHMG-CoA). These double transformants were checked by PCR with the primers T and S for the presence of the tHMG-CoA gene. Stability was checked as described above but with and without leucine.

Both *Yarrowia lipolytica* SYL320 as *Yarrowia lipolytica* SYL321 produced valencene while

5 *Yarrowia lipolytica* SYL312 did not.

CLAIMS

1. Isolated nucleic acid sequence with a nucleic acid sequence having 70%, preferably 80%, more preferably 85%, most preferably 90% nucleotide sequence identity with SEQ ID NO. 1.

2. Isolated nucleic acid sequence according to claim 1, having a nucleic acid sequence of SEQ ID NO. 1.

3. Isolated polypeptide encoded by the isolated nucleic acid sequence according to claim 1 or claim 2.

4. Isolated polypeptide according to claim 3, having 70%, preferably 80%, more preferably 85%, most preferably 90% amino acid sequence identity with SEQ ID NO. 2.

5. Expression vector comprising an isolated nucleic acid sequence according to claim 1 or claim 2.

6. Expression vector according to claim 5, further comprising a gene construct encoding for a polypeptide selected from the group consisting of CoA carboxylase, hydroxyl-methyl-glutaryl CoA reductase, squalene synthase, acetoacetyl-CoA thiolase, 3-hydroxy-3-methyl-glutaryl-CoA synthase, 3-hydroxy-3-methyl-glutaryl-CoA reductase, mevalonate kinase, mevalonate phosphate kinase, mevalonate diphosphate decarboxylase, 1-deoxy-D-xylulose 5-phosphate synthase, 1-deoxy-D-xylulose 5-phosphate reductoisomerase, 4-diphosphocytidyl-2-C-methyl-D-erythritol synthase, 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase, 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase, (E)-4-hydroxy-3-methyl-but-2-enyl pyrophosphate synthase, (E)-4-hydroxy-3-methyl-but-2-enyl pyrophosphate reductase, isopentenyl pyrophosphate isomerase, geranyl pyrophosphate synthase and farnesyl pyrophosphatesynthase.

7. Expression vector according to claim 5 or claim 6, further comprising a gene construct encoding for P450.

8. Transgenic organism comprising an expression vector according to any of the claims 5 to 7 or an isolated nucleic acid sequence according to claim 1 or claim 2.

9. Transgenic organism according to claim 8, wherein said organism is selected from the group consisting of plant, bacterium, fungus and yeast.

5 10. Transgenic organism according to claim 9, wherein the plant is selected from the group consisting of the genera *Carum*, *Cichorium*, *Daucus*, *Juniperus*, *Chamomilla*, *Lactuca*, *Pogostemon*, *Vetiveria*, *Capsicum*, *Gossyium*, *Lycopersicon*, *Nicotiana*, *Phleum*, *Solanum* and *Ulmus* or is selected from *Glycine max*, *Helianthus annuus* and *Brassica napus*.

10 11. Transgenic organism according to claim 9, wherein the yeast is selected from the group consisting of *Yarrowia lipolytica*, *Cryptococcus curvatus*, *Rhodospiridium toruloides*, *Lipomyces starkeyi*, *Rhodotorula glutinis*, *Geotrichum robustum* and *Saccharomyces cerevisiae* and *Pichia pastoris*.

15 12. Use of an isolated nucleic acid according to claim 1 or claim 2, an isolated polypeptide according to claim 3 or claim 4, an expression vector according to any of the claims 5 to 7 or a transgenic organism according to any of the claims 9 to 11 for the production of valencene and or nootkatone.

20 13. Use of the plant *Coleus forskohlii* for the production of valencene and or nootkatone.

FIGURE 1

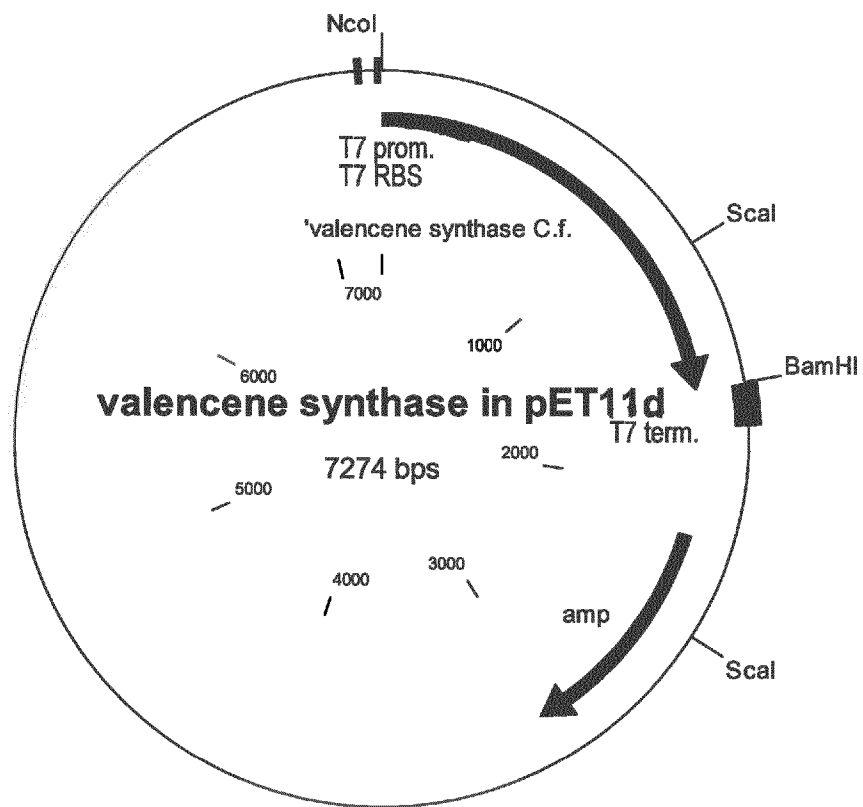


FIGURE 2

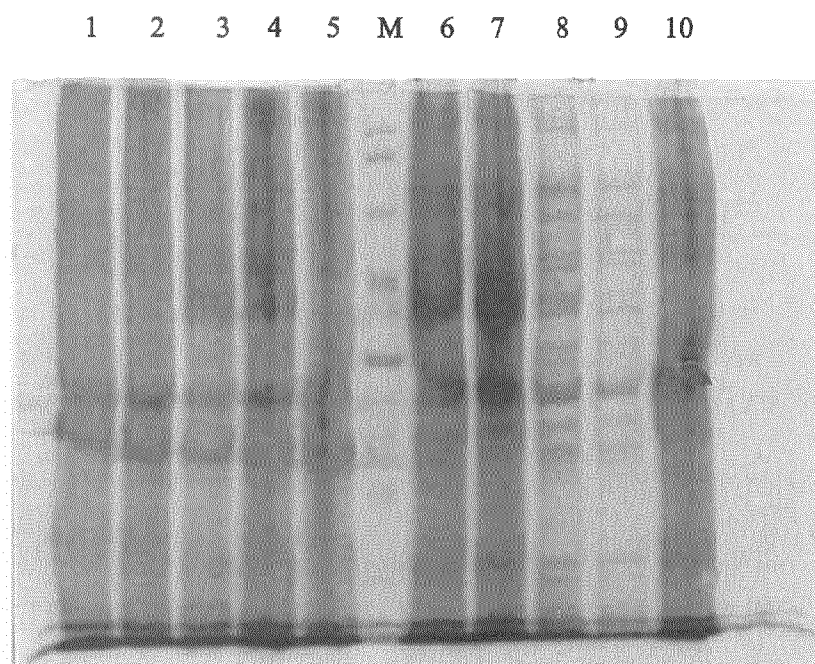


FIGURE 3

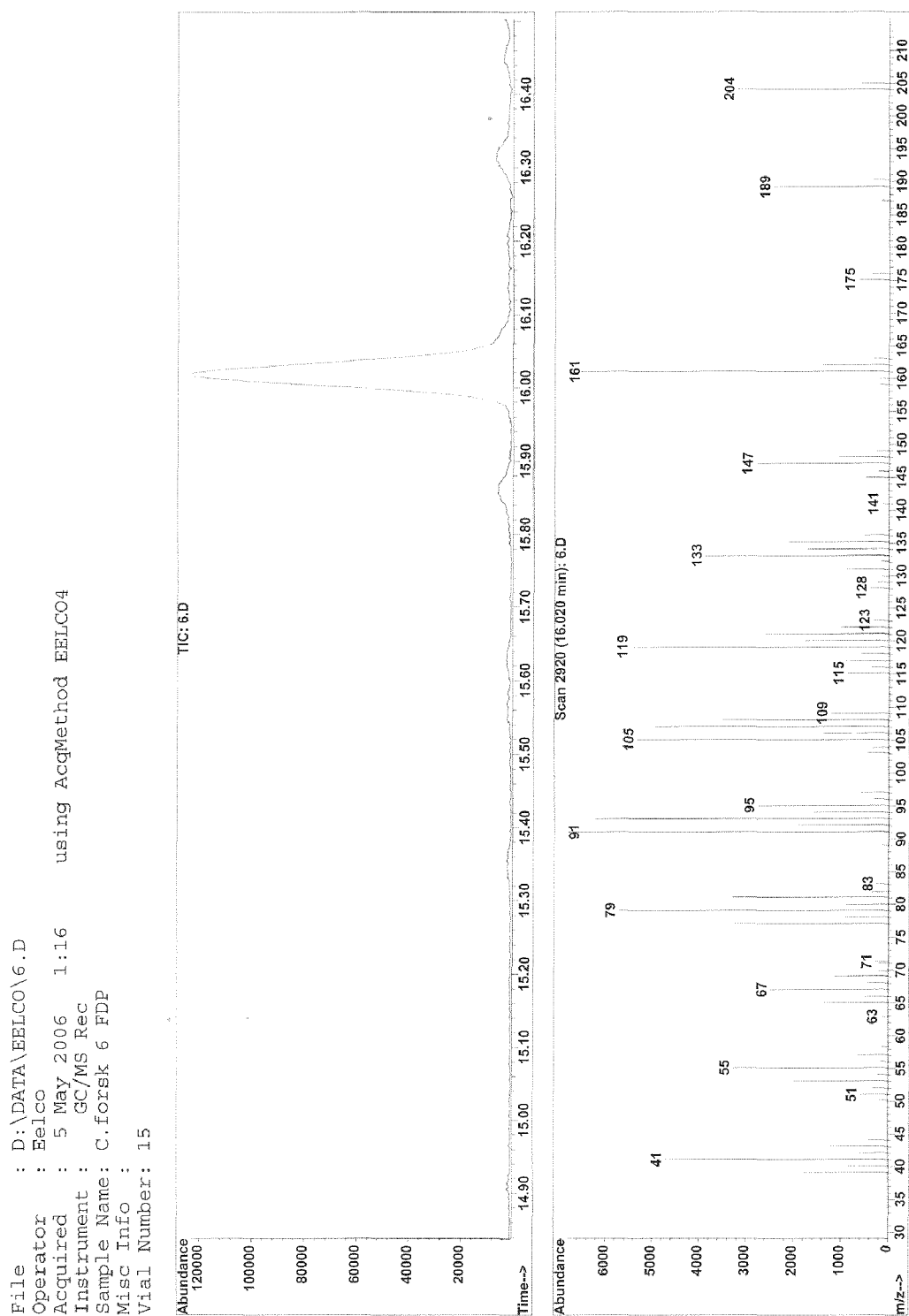


FIGURE 4

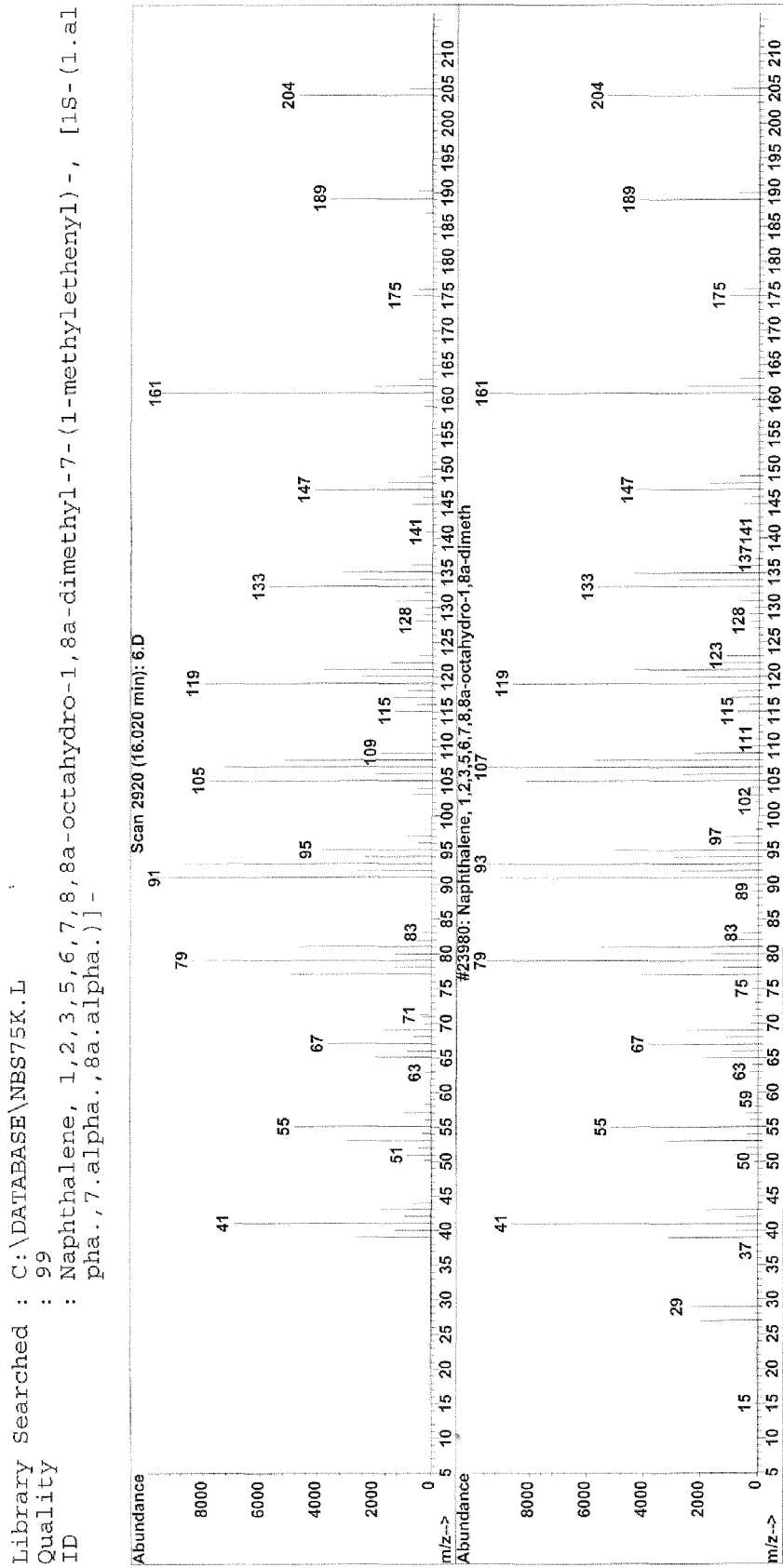


FIGURE 5

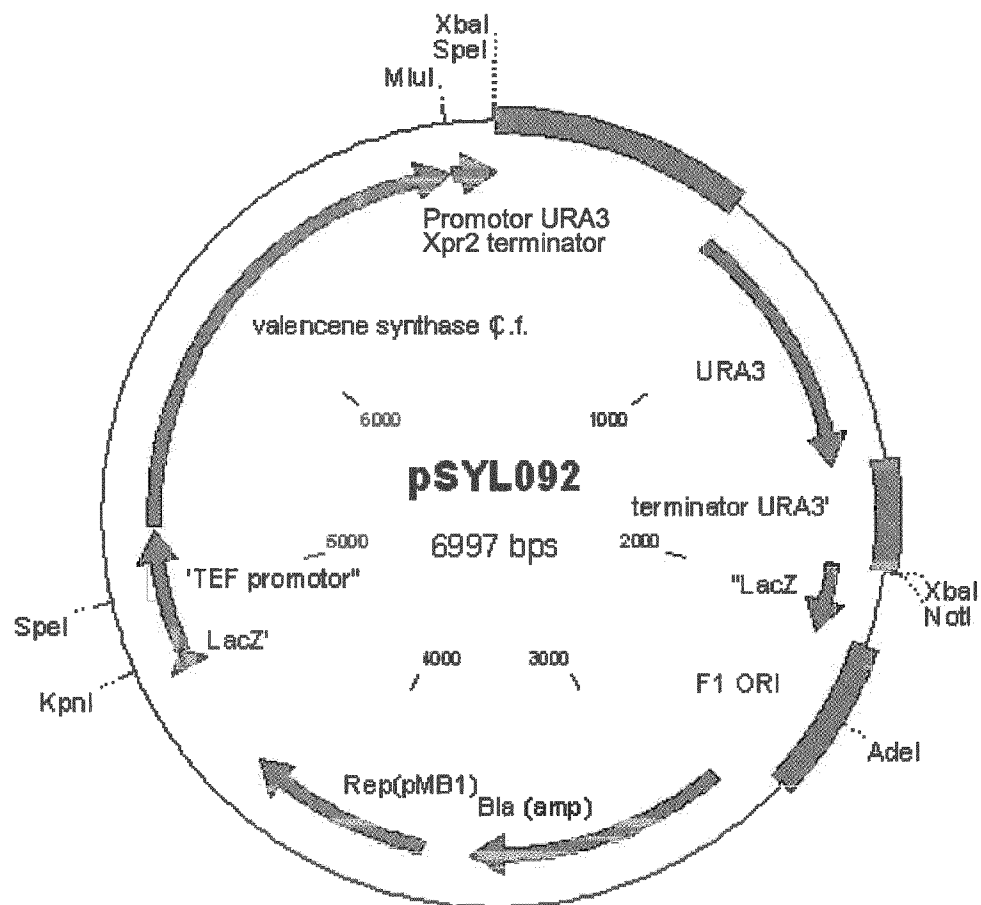
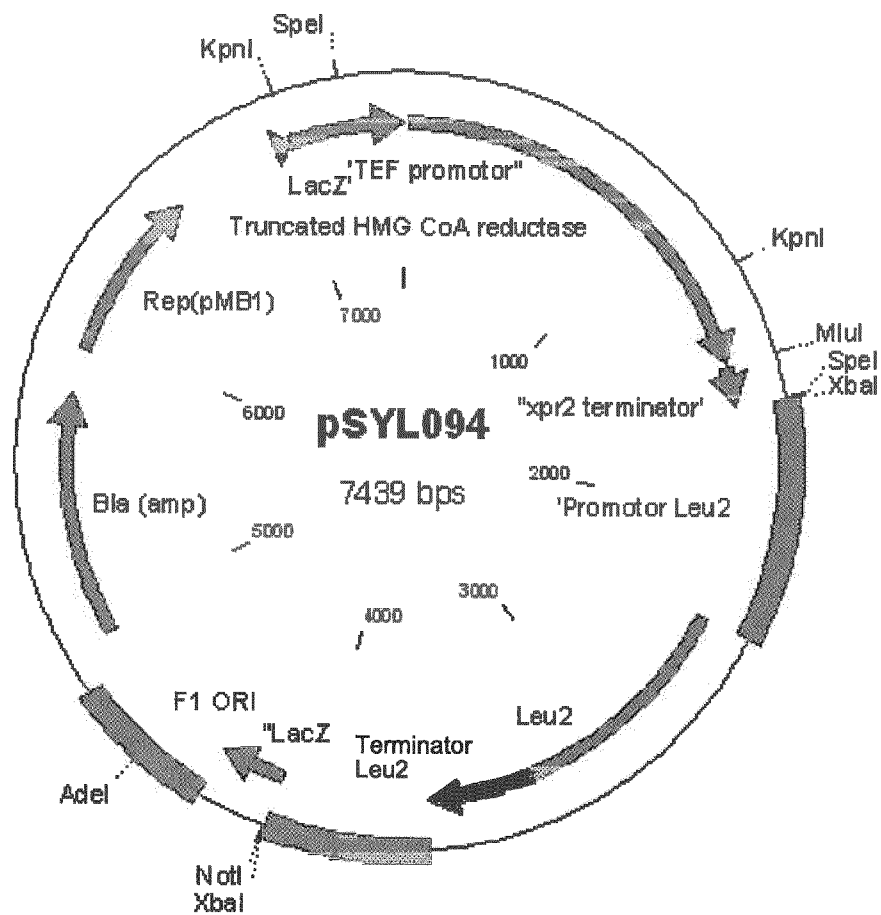


FIGURE 6



INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2015/065218

A. CLASSIFICATION OF SUBJECT MATTER
INV. C07K14/415 C12N9/88 C12N15/79 C12N15/82
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 C12N

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, Sequence Search, BIOSIS, EMBL, FSTA, 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	DATABASE UniProt [Online] 13 November 2013 (2013-11-13), "SubName: Full=Germacrene A {ECO:0000313 EMBL:AGN72800.1}";", XP002744142, retrieved from EBI accession no. UNIPROT:T1RRI5 Database accession no. T1RRI5 the whole document -----	1-4
X	WO 2012/058636 A1 (ALLYLIX INC [US]; AMICK JEAN DAVIN [US]; PARK EUNYOUNG [US]; JULIEN BR) 3 May 2012 (2012-05-03)	1-6,8-12
Y	claims 36,63,110; figure 5b; sequence 460 page 152, lines 1-21 page 149, lines 10-15 ----- -/-	7



Further documents are listed in the continuation of Box C.



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

2 October 2015

Date of mailing of the international search report

14/10/2015

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Obel, Nicolai

INTERNATIONAL SEARCH REPORT

International application No

PCT/EP2015/065218

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	J.M BANDEIRA ET AL: "Composição do óleo essencial de quatro espécies do gênero <i>Plectranthus</i> ", REVISTA BRASILEIRA DE PLANTAS MEDICINAIS, vol. 13, no. 2, 1 January 2011 (2011-01-01), pages 157-164, XP055211585, ISSN: 1516-0572, DOI: 10.1590/S1516-05722011000200006 table 1	13
Y	----- KATARINA CANKAR ET AL: "A chicory cytochrome P450 mono-oxygenase CYP71AV8 for the oxidation of (+)-valencene", FEBS LETTERS, vol. 585, no. 1, 1 January 2011 (2011-01-01), pages 178-182, XP055157852, ISSN: 0014-5793, DOI: 10.1016/j.febslet.2010.11.040 abstract	7
A	----- WO 2011/050432 A1 (FUNDO DE DEFESA DA CITRICULTURA FUNDECITRUS [BR]; IVIA [ES]; ALQUEZAR) 5 May 2011 (2011-05-05) claim 7	1-12
A	----- TAMARA WRIESSNEGGER ET AL: "Production of the sesquiterpenoid (+)-nootkatone by metabolic engineering of <i>Pichia pastoris</i> ", METABOLIC ENGINEERING, vol. 24, 1 July 2014 (2014-07-01), pages 18-29, XP055211611, ISSN: 1096-7176, DOI: 10.1016/j.ymben.2014.04.001 abstract	11

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EP2015/065218

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2012058636 A1	03-05-2012	AU 2011320127 A1	02-05-2013
		CA 2815829 A1	03-05-2012
		CN 103282492 A	04-09-2013
		EP 2633042 A1	04-09-2013
		JP 2013544084 A	12-12-2013
		US 2012246767 A1	27-09-2012
		WO 2012058636 A1	03-05-2012

WO 2011050432 A1	05-05-2011	CN 102933084 A	13-02-2013
		CR 20120281 A	13-11-2012
		EP 2493318 A1	05-09-2012
		MA 33648 B1	01-10-2012
		US 2012272405 A1	25-10-2012
		WO 2011050432 A1	05-05-2011
