



- (51) **International Patent Classification:**
C12N 9/88 (2006.01)
- (21) **International Application Number:**
PCT/US2015/063656
- (22) **International Filing Date:**
3 December 2015 (03.12.2015)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**
62/089,511 9 December 2014 (09.12.2014) US
- (71) **Applicant: DSM IP ASSETS B.V.** [NL/NL]; Het Overloon 1, 6411 Te Heerlen (NL).
- (72) **Inventor; and**
- (71) **Applicant (for US only): ROYER, John** [UZ/US]; 16 De-mar Road, Lexington, Massachusetts 02420 (US).
- (74) **Agent: CHEN, Xi;** 6480 Dobbin Road, Columbia, Maryland 21045 (US).
- (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).

Published:

- with international search report (Art. 21(3))
- with sequence listing part of description (Rule 5.2(a))

(54) **Title:** METHODS FOR PRODUCING ABIENOL

(57) **Abstract:** The present disclosure relates to a novel method, expression vectors, and host cells for producing abienol by converting geranylgeranyl diphosphate (GGPP) to abienol in the presence of a combination of a class II diterpene synthase and a bifunctional class I/II abienol synthase.



WO 2016/094178 A1

METHODS FOR PRODUCING ABIENOL

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims priority to and the benefit of the filing date of United States Provisional Patent Application No. 62/089,511 filed December 9, 2014, the disclosure of which is hereby incorporated herein by reference.

FIELD OF THE INVENTION

[0002] Embodiments of the present disclosure relate to novel methods for the production of abienol and expression vectors and host cells useful in such methods.

BACKGROUND OF THE INVENTION

[0003] With a growing world economy and request for higher living standard, the demand for fragrance is on the rise. However, supplies of animal or plant based fragrance are limited, due to restraints in natural resources and animal protection. Attempts have been made to produce fragrances or their precursors from renewable sources such as microorganisms.

[0004] Ambergris is a prized, traditional fragrance ingredient that is a byproduct of the whale intestine. Ambrox, a substitute for ambergris, is produced by a chemical conversion of the diterpene sclareol, which is currently obtained from clary sage. Ambrox can also be generated from the related diterpene abienol, which has been found in fir and tobacco (Barrero et al. 1993, Tetrahedron 49:10405-10412). Also see **Figure 1**.

[0005] The pathways to both abienol and sclareol in plants are proposed to involve two steps. The first step consists of the conversion of the isoprenoid pathway molecule geranylgeranyl diphosphate (GGPP) to a common intermediate named labda-13-en-8-ol diphosphate (LDPP) through the activity of a class II diterpene synthase (diTPS). The second step is catalyzed by a class I diTPS. There are several type of class I diTPS, each responsible for producing a specific end product. For example, abienol synthase (ABS) is for abienol production, and sclareol synthase (Scs) is responsible for sclareol production. See **Figure 2**.

[0006] The enzymes involved in the two-step conversion of GGPP to sclareol or GGPP to abienol are plant specific and can be in the form of two independent enzymes or a single enzyme with two active sites. For example, in abienol production by tobacco (Sallaud et al.

2012, *Plant J.*, 72(1):1-17), the class II diTPS of tobacco (referred to as NtCPS2 by Sallaud et al., and referred to as Nt-class II-diTPS by the present disclosure) and the class I diTPS synthase of tobacco (referred to as abienol synthase of tobacco or Nt-ABS by the present disclosure), are in the form of two different protein molecules. Similarly, in sclareol production by clary sage (Schalk et al. 2012, *Journal of Am. Chem. Soc.* 134:18900-18903);(Caniard et al. 2012, *BMC Plant Biology* 12:119), the class II diTPS of clary sage (referred to as Ss LPS by Schalk et al, and referred to as Ss-class II-diTPS by the present disclosure) and the class I diTPS synthase of clary sage (referred to as sclareol synthase of clary sage or Ss-Scs by the present disclosure), are also in the form of two independent protein molecules. In contrast, in the production of abienol by fir (Zerbe et al. 2012, *J. Biol. Chem.* 287:12121-12131), both class I and class II diTPS subunits reside on one bifunctional class I/II abienol synthase (referred to as AbCAS by Zerbe et al. and by the present disclosure). Therefore in fir, GGPP is converted to abienol in the presence of the single bifunctional class I/II abienol synthase. See **Figure 2**.

[0007] Plant sources for sclareol and abienol are considered to be unreliable; thus a process for microbial production of either product could have commercial value. The class II diTPS and sclareol synthase genes from clary sage have been isolated and simultaneously expressed in *E. coli*, resulting in titers of approximately 1.5 grams per liter of sclareol in lab scale fermenters. Measurable production in *E. coli* of abienol has been achieved by the simultaneous expression of the class II diTPS and abienol synthase genes of tobacco, or the expression of the individual class I/II abienol synthase of fir. However, production of abienol based on the existing methods is very low. It could therefore be desirable to produce abienol at a much higher titer.

SUMMARY OF THE INVENTION

[0008] We have now surprisingly found a novel method for significantly increasing the production rate of abienol from geranylgeranyl diphosphate (GGPP) in the presence of a combination of a class II diterpene synthase and a bifunctional class I/II abienol synthase. In one embodiment, the class II diTPS may be from tobacco or clary sage, and the bifunctional class I/II abienol synthase may be from fir. In another embodiment, the above class II diTPS may be a polypeptide comprising an amino acid sequence at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95% , at least 97%, or at least 99% identical to SEQ ID NO:1. In

another embodiment, the above class II diTPS may be a polypeptide comprising an amino acid sequence at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95% , at least 97%, or at least 99% identical to SEQ ID NO:2. In another embodiment, the above bifunctional class I/II abienol synthase may be a polypeptide comprising an amino acid sequence at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95% , at least 97%, or at least 99% identical to SEQ ID NO:3.

[0009] In another embodiment, the combination of the class II diTPS and the bifunctional class I/II abienol synthase is expressed by a recombinant host cell. In one embodiment, the recombinant cell is a genetically modified microorganism genetically modified to express at least one exogenous polypeptide selected from the group consisting of the above class II diTPS and the bifunctional class I/II abienol synthase. In one embodiment, the host cell comprises at least one nucleic acid encoding one or more of the amino acid sequences selected from the group consisting of said class II diTPS and said bifunctional class I/II abienol synthase.

[0010] In one embodiment, the recombinant host cell is a fungus. In another embodiment, the recombinant host cell is a *Yarrowia* fungus. In one specific embodiment, the recombinant host cell is *Yarrowia lipolytica*.

[0011] In another aspect of the disclosure, the present invention is directed to a recombinant host cell comprising a class II diterpene synthase and a bifunctional class I/II abienol synthase. In one embodiment, the class II diTPS may be from tobacco or clary sage, and the bifunctional class I/II abienol synthase may be from fir. In another embodiment, the above class II diTPS may be a polypeptide comprising an amino acid sequence at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95% , at least 97%, or at least 99% identical to SEQ ID NO:1. In another embodiment, the above class II diTPS may be a polypeptide comprising an amino acid sequence at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95% , at least 97%, or at least 99% identical to SEQ ID NO:2. In another embodiment, the above bifunctional class I/II abienol synthase may be a polypeptide comprising an amino acid sequence at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95% , at least 97%, or at least 99% identical to SEQ ID NO:3.

[0012] In one embodiment, the recombinant host cell is a fungus. In another embodiment, the recombinant host cell is a *Yarrowia* fungus. In one specific embodiment, the recombinant host cell is *Yarrowia lipolytica*.

[0013] In another aspect of the disclosure, the present invention is directed to an expression vector comprising a polynucleotide molecule encoding an amino acid sequence comprising SEQ ID NO:1 and a polynucleotide molecule encoding an amino acid sequence comprising SEQ ID NO:3.

[0014] In another aspect of the disclosure, the present invention is directed to an expression vector comprising a polynucleotide molecule encoding an amino acid sequence comprising SEQ ID NO:2 and a polynucleotide molecule encoding an amino acid sequence comprising SEQ ID NO:3.

[0015] In one embodiment, one or both of the above described polynucleotide molecules are operationally linked to a transcriptional control sequence.

OVERVIEW OF THE SEQUENCE LISTING

[0016] The nucleic acid sequences listed in the accompanying sequence listing are shown using standard letter abbreviation for nucleotide bases. Only one strand of each nucleic acid sequence is shown, but the complementary strand is understood to be included by any reference to the displayed strand. In the accompanying sequence listing:

[0017] SEQ ID NO:1 is the amino acid sequence encoding the class II diterpene synthase from *Nicotiana tabacum*

MQVIITSSHRFFCHHLHQLKSPTSLSAQKAEFKKHGPRNWL FQTEGSLLYKPVRLNCATS
 DASYLGNVNEYLESDHKNSEEKDIQVSRTIQMKGLTEEIKHMLNSMEDGRLNVLAYD
 TAWVSFIPNTTNGNDQRPMFPSC LQWIIDNQLSDGSWGEEIVFCIYDRLLNTLVCVIAL
 TLWNTCLHKRNKGVMFIKENLSKLETGEVENMTSGFELVFPTLLEKAQQLDIDIPYDAP
 VLKDIYARREVKLTRIPKDVIIHTIPTTVLFSLEGLRDDLDWQRLLKLQMPDGSFLISPAST
 AFAFMETNDEKCLAYLQNVVEKSNGGARQYPFDLVTRLWAIDRLQRLGISYYFAEEFK
 ELLNHVFRYWDEENGIFSGRNSNVSDVDDTCMAIRLLRLHGYDVSPDALNNFKDGDQF
 VCFRGEVDGSPTHMFNLYRCSQVLFPGEKILEEAKNFTYNFLQQCLANNRCLDKWVIA
 KDIPGEIWYALEFPWYASLPRVEARYYYIEQYGGADDIWIGKTLYRMPDVNNNVYLQAA
 KLDYNRCQSQHRFEWLIMQEWF EKCNFQQFGISKKYLLVSYFLAAASIFEVEKSRRERLA
 WAKSRIICKMITSYYNDEATTWTTNSLLMEFKVSHDPTRKNGNETKEILVLKNLRQFL
 RQLSEETFEDLGKDIHHQLQNAWETWLVFLREEKNACQEETELLVRTINLSGGYMT HDE

ILFDADYENLSNLTNKVCGKLNELQNDKVTGGSKNTNIELDMQALVKLVFGNTSSNINQ
DIKQTFFAVVKTFYYSAHVSEEIMNFHISKVLFQQV

[0018] SEQ ID NO:2 is the amino acid sequence encoding the class II diterpene synthase from *Salvia sclarea*

MTSVNLSRAPAAIIRRLQLQPEFHAECSWLKSSSKHAPFTLSCQIRPKQLSQIAELRVTS
LDASQASEKDISLVQTPHKVEVNEKIEESIEYVQNLLMTSGDGRISVSPYDTAVIALIKDL
KGRDAPQFPSCLEWIAHHQLADGSWGDEFFCIYDRILNTLACVVALKSWNLQSDIIEKG
VTYIKENVHKLKGANVEHRTAGFELVVPTFMQMATDLGIQGLPYDHPLIKEIADTKKQR
LKEIPKDLVYQMPTNLLYSLEGLGDLEWERLLKLQSGNGSFLTSPSSTAAVLMHTKDEK
CLKYIENALKNCDGGAPHTYPVDIFSRLWAIDRLQRLGISRFFQHEIKYFLDHIESVWEET
GVFSGRYTKFSDIDDTSMGVRLLKMHGYDVDPNVLKHFQKQDQKGFSCYIGQSVEASAP
MYNLYRAAQLRFPGEEVFEEATKFAFNFLQEMLVKDRLQERWVISDHLFDEIKLGLKM
PWYATLPRVEAAYYLDHYAGSGDVWIGKSFYRMPEISNDTYKELAILDFNRCQTQHQL
EWIHMQEWYDRCSLSEFGISKRELLRSYFLAAATIFEPERTQERLLLAKTRILSKMITSFV
NISGTTLSLDYNFNGLDEIISSANEDQGLAGTLLATFHQLLDGFDIYTLHQLKHVWSQWF
MKVQQGEGSGGEDAVLLANTLNICAGLNEDVLSNNEYTALSTLTNKICNRLAQIQDNKI
LQVVDGSIKDKELEQDMQALVKLVLQENGGAVDRNIRHTFLSVFKTFYYDAYHDDETT
DLHIFKVL

[0019] SEQ ID NO:3 is the amino acid sequence encoding the bifunctional class I/ II abienol synthase from *Abies balsamea*

MALPVYSLKSHIPITTASAKMNYTPNKGMITANGRSRRIRLSPNKIVACAGEADRTFPSQ
SLEKTALFPDQFSEKNGTPSNFTPPNREFPPSFWNNDIINSITASHKVQTGDRKRIQTLISEI
KNVFNSMGDGETSPSAYDTAWVARIPAVDGSEQPQFPQTLEWILQNQLKDGSGWGEFFY
FLAYDRLLATLACIITLTIWRTGNVQLHKGIEFFRKQVVRMDDEADNHRPSGFEIVFPAM
LNEAKSLGLDLPYELPFIEQMVKKREAKLKMITTNVLYTIQTTLYSLEGLHEIVDFDKII
KLQSKDGSFLGSPASTAAVFMQTGNTKCLEFLEFVLRKFRNHVPSDYPLDLFERLWVVD
TVERLGIDRHFKEIKDALDYVYSCWDERGIGWAKDSPIADIDDTAMGLRILRLHGYNV
SPDVLKTFKDENGEEFFCFMGQTQRGVTDMLNVYRCSQVAFPGETIMEEEAKLCTERYLR
NALENADAFDKWAIKKNIRGEVEYALKYPWHRSMRLEVRSYIGNYGPNDVWLGKSL
YMMPYISNEKYLELAKLDFNSVQSLHQEEIRELVRWCKSSGFTELKFTRDRVVETYFAV
ASSMFEPEFSTCRAVYTKISVLLVILDDLVDGYGSPDEIKLFSEAVKRWDLSLLEQMPDH

MKICFLGLYNTVNEVAEEGRKTQGHDLGYIRNLWEIQLAAFTREAEWSQGKYVPSFD
EYIENAQVSIGVATILLITILFTEEDDILSHIDYGSKFLRLASLTARLANDIKTYQEERAHGE
VVSIAIQCYMKDRPEITEEEALKYVYGRMVNDLAELNSEYLKSNEMPQNCKRLVFD TAR
VAQLFTMEGDGLTYSDTMEIKEHIKKCLFEPAT

[0020] SEQ ID NO:4 is the amino acid sequence encoding the abienol synthase gene from *Nicotiana tabacum*

MVLGLRSKIIPDPDHKLGNIKLGSVTNAICHRPCRVRCSHSTASSMEEAKERIRETFGKIE
LSPSSYDTAWVAMVPSRYSMNQPCFPQCLDWILENQREDGSWGLNPSHPLLVKDSLST
LASLLALRKWRIGDNQVQRGLGFIETHGWAVDNKDQISPLGFEIIFPCMINYAEKLNLDL
PLDPNLVNMMLCERELTIERALKNEFEGNMANVEYFAEGLGELCHWKEMMLRQRHNG
SLFDSPATTAALIIHQYDEKCFGYLNSILKLHDNWVPTICPTKIHSNLFVLDALQNLGV
DRYFKTEVKRVLDEIYRLWLEKNEEIFSDVAHCAMAFRLLRMNNYEVSSEEELEGFVDQE
HFFTTSSGKLMNHVAILELHRASQVAIHERKDHILDKISTWTRNFMEQKLLDKHIPDRSK
KEMEFAMRKFYGTDFRVETRRYIESYKMDSFKILKAAYRSSGINNIDLLKFSEHDFNLQ
TRHKEELQQMKRWFTDCKLEQVGLSQQYLYTSYFIIAAILFEPEYADARLAYAKYAIIT
AVDDFFDCFICKEELQNIIEVERWEGYSTVGFRSERVRIFFLALYKMVEEIAAKAETKQ
GRCVKDHLINLWIDMLKCMLVELDLWKIKSTPSIEEYLSVACVTIGVPCFVLTSLYLLG
PKLSKDVIESSEVSALCNCTAAVARLINDIHSYKREQAESSTNMVSILITQSQGTISEEEAI
RQIKEMMESKRRELLGMVLQNKESQLPQVCKDLFWTTINAAYSIIHTHGDGYRFPEEFK
NHINDVIYKPLNQYSP

[0021] SEQ ID NO:5 is the DNA sequence encoding the class II diterpene synthase from *Nicotiana tabacum*, as optimized for expression in *Yarrowia lipolytica*

atgcaggttattattacctcctcaccgattttctgccaccacctcaccagctcaagtcctcctccttctgctcagaaggctgagtta
agaagcacggcccccgaactggctttccagactgagggctctcctttacaagcctgtccgactcaactgtgctacttctgatgctttac
cttgtaacgtgaacgagtaccttgagtctgaccactctaagaactccgaggagaaggatattcaggtttcccgaactatccagatgaagggt
cttaccgaggagatcaagcacatgcttaactctatggaggacggacgacttaacgtcctcgctacgacactgctgggtttcctttattcctaa
cactaccaacaacggaaacgatcagcgacctatgtttccctctgtcttcagtggattattgacaaccagctttctgatgggtcttggggagagg
agattgtttctgcatttacgaccgactccttaacactctcgtttgtgtattgctctcactctctggaacactgccttcacaagcgaaacaagggt
gtgatgtttatcaaggagaacctttctaagctggagactggtgaggttgagaacatgacttctgggtttgagcttggtttccactctccttgaga
agggccagcagctcgatattgacattccctacgatgctctgtctgaaggatattacgctcgacgagaggttaagctcaccggaatcccta
aggacgttatccactattcccactaccgttctctttctcttgagggaactccgagatgacctcgactggcagcgactcctgaagctccagat

gcctgacgggtcttctgatttccctgctccactgcctttgcttcatggagactaacgatgagaagtgtcttgctaccttcagaacgttggt
gagaagtctaacggaggtgcccacagtaccccttcgaccttggtactcgacttgggccattgatcgactccagcgactcggaatctcttact
actttgccgaggaggtcaaggagcttctcaaccacgtgtccgatactgggacgaggagaacggaattttcttgacgaaactctaactgtt
ctgatgttgatgacacttgcattggctatccgacttctccgacttcacggttacgatgtttccctgacgcccttaacaactcaaggacggcgac
cagttcgtttgcttccgaggtgaggtggacggttctcctaccacatgtttaacctctaccgatgttccaggttctttcccgagagaagatt
cttgaggaggctaagaacttcacttacaacttcttcagcagtgcttgcataacaaccgatgcctcgacaagtgggtcattgctaaggacatc
cccgccgagatttggtacgtcttgagttccctggtagcctccctccccgagtgagggtcgatactacattgagcagtagcgccgagct
gacgatatgttgattggcaagactctctaccgaatcccgatgtcaacaacaacgtttaccttcaggctccaagctcgattacaaccgatgc
cagtcaccagcaccgatttgagtggtgattatgcaggagtggttgagaagtgaacttcagcagttcgggaattccaagaagtacctccttg
tttcttacttcttctgctgccgcttctatgttgaggtcgagaagtcccagagcgacttgcttgggctaagtctcgaattatctgtaagatgattact
tcttactacaacgatgaggccactacctggaccactcgaaactctctccttatggagtttaaggttctcacgacctacccgaaagaacggta
acgagactaaggagatccttggttctcaagaaccttcgacagttccttcgacagcttctgaggagactttcgaggaccttggaaggacatcc
accaccagcttcagaacgcttgggagacttggcttgtttccttcgagaggagaagaacgcttgcaggaggagactgagcttctcgtgcga
actattaacctctctggcggctacatgaccacgatgagattctttcgatgccgactacgagaacctgtccaaccttaccacaagggttggtg
gcaagctcaacgagcttcagaacgacaaggtcactggcggctctaagaacaccaacattgagcttgacatgcaggctctcgttaagctgggt
tttgtaaaccttctctaatcaaccaggacattaagcagactttcttctgtgtgtaagaccttctactacttgcccagtttctgaggaga
ttatgaactttcacattccaaggtgctctttcagcaggtctaa

[0022] SEQ ID NO:6 is the DNA sequence the class II diterpene synthase from *Salvia sclarea*, as optimized for expression in *Yarrowia lipolytica*

atgacttctgttaacctttcccgagcccccgctgccattatccgacgacgacttcagcttcagcctgagtttcacgccgagtggttcttggtgaa
gtcttctccaagcagcccccttcaccttctcctgccagattcgacctaaagcagctctccagattgccgagcttcgagttactccctcgatg
cttcccaggcttcgagaaggacatttccctgttcagactccccacaaggttgaggtaaacgagaagatcgaggagtctatcgagtacgtcc
agaaccttctcatgacttccggcgacggacgaatttctgtgtccccctacgacaccgctgtgatcgccctgattaaggacctcaagggtcga
gatgccccctcagtttccctctgttcttgagtggttggccaccaccagcttgcctgatggctcttggggcgacgagttcttctgtattacgaccga
atcctgaacactctcgcttgtgtcgtcgcctgaagtcttggaaccttcagctctgatattattgagaagggtgtgacctacatcaaggagaacgt
ccacaagctcaagggtgccaacgttgagcaccgaactgccgattcgagcttgtggttctacctttatgcagatggccactgaccttggcat
tcagggtcttccctacgatccccctcatcaaggagattgctgacactaagaagcagcgactcaaggagattccaaggatctcggttacc
agatgcctaccaaccttctctactcccttgagggactcggcgaccttgagtgggagcgactcctgaagctccagtcggcaacggctccttc
ctcacttccccctcctccaccgccgctccttatgcacaccaaggacgagaagtgtctgaagtacatcgagaacgccctcaagaactcgcg
acggaggtgctccccacacttacctgtcgatattttctctcgacttgggctattgatcgacttcagcgacttggatttctcgattcttcagca
cgagatcaagtacttctcgatcacatcgagtccgttgggaggagaccggagtttctctggacgatacactaagtttctgatattgatgaca

cctctatgggcttcgacttctcaagatgcacggatacgcgtcgatcccaacgttctcaagcacttcaagcagcaggatggttaagtttctg
ctacattggtcagctctgtcgagcttgcctctcctatgtacaacctttaccgagccgccagcttcgattccctggtgaggaggtttttaggagg
ccactaagtttgctttaacttccttcaggagatgcttgtcaaggatcgacttcaggagcgatgggtgatttccgaccaccttttcgatgagatta
agctgggctcaagatgccttggtacgccaccttccccgagtcgaggccgttactacctcgatcactacgctggttctggtgatgtttggat
tggcaagcttttctaccgaatgcctgagatttccaacgatacctacaaggagcttgccattctcgatttcaaccgatgccagactcagcaccag
cttgagtggattcacatgcaggagtgttacgaccgatgctctcttccgagttcgccatttccaagcgagagcttctccgatcttactttctgc
cgccgctaccattttcgagcctgagcgaactcaggagcgacttctccttgccaagacccgaatcctgtctaagatgattacttctttgtcaaca
tttccggtactaccctttctctcgactacaacttcaacggcctcgatgagattatctcctctgccaacgaggatcagggtctggctggtactctc
ctggctaccttccaccagcttctcgacggattcgatattacactctccaccagctcaagcacgtttggtccagtggttcatgaagggtgcagc
agggagaggggttctggcggcgaggacgccgtgctccttgccaacacctcaacatctgcgccggcctcaacgaggacgtgctctccaac
aacgagtacactgctctgtccacctcaccaacaagatttgaaccgactcgcccgatccaggacaacaagatttccaggttgtggacg
gtcccatcaaggataaggagcttgagcaggatattgcaggcccttgtaagctcgctccttcaggagaacggcggagccgttgaccgaaaca
tccgacacaccttcttccgttttcaagactttctactacgatgcctaccacgacgatgagactaccgaccttcacatcttcaaggttctcttgc
accgttgtgtaa

[0023] SEQ ID NO:7 is the DNA sequence encoding the bifunctional class I/ II abienol synthase from *Abies balsamea*, as optimized for expression in *Yarrowia lipolytica*

atggcctgcccgttactctctcaagtcccacattcccatcactaccatcgctctgctaagatgaactacaccccaacaagggtatgatta
ccgccaacggacgatctcgacgaatccgacttttcccaacaagatcgttgcttgtgctgggtgaggctgatcgaacttccccctctcagtcct
tgagaagaccgctctcttcccgatcagtttttgagaagaacgggtactccctctaacttcactcccccaaccgagagttccccctcttttg
gaacaacgatattatcaactctattactgcttctcacaaggttcagactggcgaccgaaagcgaatccagactctcatttctgagattaagaac
gtctttaactctatgggcgatggtgagacttctccctctgcttacgacaccgcttgggttgcgaatccccgccgttgatggctctgagcagc
cccagttccccagactcttgagtggattcttcagaaccagctcaaggacgggttcttggggtgaggagtctacttcttctgcttacgaccgactt
ctcgctacccttgctgcattattaccctcaccatttggcgaactggcaacgttcagcttcacaagggcattgagttctccgaaagcaggttgt
tcgaatggacgatgaggtgataaccaccgacctctggtttgagattgtcttcccgctatgcttaacgaggctaagtccttggcttgacc
tgccctacgagcttcccttcattgagcagatggtaagaagcgagaggccaagctcaagatgattaccaccaacgtcctgtacaccattcag
actaccctccttactctctggagggccttcacgagattgtgacttcgataagattatcaagctccagccaaggacgggttcttctcggtcc
cccgttctactgcccgtgtttcatgcagactggtaacactaagtgccttgagttccttgagttcgttctccgaaagtttcgaaccacgtgccc
tctgactacccccctgatctctttgagcgactttgggtcgttgacactgttgagcgacttggcattgatcgacacttcaagaaggagatcaagg
acgctcttgactacgtgtactcttgttgggacgagcgaggcattggctgggccaaggactctccatcgccgatattgatgacactgctatgg
gccttcgaatccttcgactgcacggatacaacgttccccgatgttctcaagactttcaaggacgagaacggagagttcttttctcatgggt
cagactcagcgaggagttaccgacatgcttaacgtttaccgatgttctcaggttgccttcccgagagactatcatggaggaggccaagctc

tgtactgagcgatacctgcgaaacgctctggagaacgccgacgctttgacaagtgggctattaagaagaacattcgaggcgagggtggag
 tacgctctcaagtacccttggcaccgatctatgccccgactggaggtgcgatcttacattggtaactacggccccaacgatgtctggcttgg
 aagtccctttacatgatgccctacatttctaacgagaagtaccttgagcttgccaagctggacttcaactctgtgcagtccttcaccaggagg
 agattcgagagcttgtccgatgggtgaagtctctggttactgagctgaagttcacccgagatcgagttgtgagacttacttcgctgttgcct
 cctccatgtttgagcccgagttctctacctgtcgagccgtttactaagatttccgttctcctcgtcattcttgacgaccttacgatggctacgg
 ttctcccgacgagatcaagctgttctccgaggtgtcaagcgatgggatctctcccttctgagcagatgcccgaccacatgaagatttgcttc
 ctgggtctttacaacactgttaacgaggttgctgaggaggacgaaagactcagggccacgatgttcttggtacattcgaaacctttggga
 gattcagctcgccgctttcacccgagaggtgagtggtcccagggcaagtacgtgccctcttcgatgagtacattgagaacgccaggttt
 ctattggagttgctactatcctccttattactattcttttactgaggaggacgatattctctccacattgattacgggtccaagttctccgactcg
 ctctcttaccgctcgacttgccaacgacatcaagacttaccaggaggagcgagcccacggcgaggtggtttccgctattcagtggttacatga
 aggaccgacccgagattactgaggaggaggctctcaagtacgtttacgggtgaatgggttaacgatctcgccgagcttaactctgagtacctc
 aagtctaacgagatgccccagaactgcaagcgactggttttgacactgcccagttgcccagcttttactatggaggggtgacggcctcac
 ctactctgacactatggagattaaggagcacatcaagaagtgcctctttgagcccgctacctaa

[0024] SEQ ID NO:8 is the DNA sequence encoding the abienol synthase gene from *Nicotiana tabacum*, as optimized for expression in *Yarrowia lipolytica*

atgggtcttggcctgcgatctaagatcattccccttcccgatcacaagctcggaacatcaagctcggttctgttaccaacgctatttgcaccg
 accctgtcgagtcgatgtctctactctactgttctctatggaggaggctaaggagcgaatccgagagactttcgaaagattgagctttct
 cctcctcttacgacactgcttgggttgcctatgggtcccctctcgatactctatgaaccagccctgttttccccagtgcccttgactggattcttgaga
 accagcgagaggacggttcttggggcctcaaccctctcacccttctcgtaaaggactccctttcttccactctcgcttctctcttgccttc
 gaaagtggcgaattggtgataaccaggtccagcgaggtcttggctttattgagactcacggttgggctgtcgataacaaggatcagattctc
 cccttggtttgagattatcttccctgcatgattaactacgctgagaagctcaaccttgacctcccccttgaccccaaccttgtaacatgatgct
 ctgcgagcgagagcttaccattgagcgagccctcaagaacgagtttgagggtaacatggctaaccgttgagtcttctgaggggactcgggt
 gagctttgtcactggaaggagatgatgcttcgacagcgacacaacggctctctcttctgactctcccgccaccactgccgctgcccttatttacc
 accagtacgatgagaagtgtttggctacctcaactctatctcaagctccacgataactgggttcccactatttggcccaactaagattcactct
 aacctttctcgttgatgcccttcagaacctcgaggtgaccgatactttaagactgagggttaagcgagttctcgatgagattaccgactttgg
 cttgagaagaacgaggagatttttctgacgttgctcactgtgctatggcttttcgactccttgaatgaacaactacgaggtttcctctgaggag
 cttgaggggtttgttgaccaggagcacttcttactacctcttggaagctcatgaaccacgttgctattctggagcttcaccgagcttctcag
 gtggctattcacgagcgaaaggaccacattcttgataagatttctacttggactcgaaactttatggagcagaagctccttgacaagcacattc
 ccgaccgatctaagaaggagatggagtttgctatgcgaaagttttacggcactttcgaccgagtgagactcgacgatacatcgagtcttac
 aagatggactcctttaagattctcaaggccgttaccgatctccggtattaacaacattgaccttctcaagttctctgagcacgatttcaacctct
 gccgagcccgacacaaggaggagcttcagcagatgaagcgatggttcaccgattgcaagctggagcaggtcggcttttctcagcagttacc

tctacacttcttacttcattatcgccgctatcctctttgagcccgagtacgctgatgctcgacttgcttacgctaagtacgccatcattatcactgc
 cgtggacgatttcttcgattgtttatttgaaggaggagcttcagaacatcatcgagcttgctgagcgatgggagggatactctaccgtcgga
 ttccgatctgagcgagttcgaatttctccttgccctttacaagatgggtgaggagattgctgccaaggccgagactaagcagggtcgatgtg
 tcaaggatcaccttattaacctttggattgatgctcaagtgtatgctgggtgagcttgacctttggaagattaagtcactacccccctatcga
 ggagtacctttctgttgctgtgttactattgggttcctgtttgttctcacttctctaccttctcggacccaagctgtccaaggacgtcattga
 gtctctgaggtttccgccccttgcaactgtactgccgctgtcggccgacttattaacgatattcactcctacaagcgagagcaggctgagtc
 tctactaacatggtttctatccttatcaccagtcaccagggtactatctctgaggaggaggctattcgacagattaaggagatgatggagttaa
 gcgacgagagcttctcggaatggttctccagaacaaggagtcaccagctccccaggtgtgcaaggacctcttttgactaccatcaacgcc
 gcttactctattcacactcacggcgatggataccgattccccgaggaggtcaagaaccacatcaacgatgttattacaagcccccaaccagt
 actccccctaa

BRIEF DESCRIPTION OF THE DRAWINGS

[0025] Embodiments of the invention will now be shown, by way of example only, with reference to Figures 1-3 in which:

[0026] **FIG. 1** shows that conversion of sclareol and abienol to Ambrox can be performed by a chemical process.

[0027] **FIG. 2** shows the biochemical pathways to convert geranylgeranyl diphosphate to abienol in tobacco and to sclareol in clary sage. A. pathways to abienol in tobacco and sclareol in clary sage involve the activity of two separate enzymes: a type II diTPS that converts geranylgeranyl diphosphate (GGPP) to labda-13-en-8-ol diphosphate (LDPP) and a type I diTPS that converts LDPP to the final product. In tobacco, an abienol synthase (Nt-ABS) which is a type I diTPS converts LDPP to abienol. In clary sage, a sclareol synthase which is another type I diTPS converts LDPP to sclareol. B. in fir, the 2-step conversion of GGPP to LDPP and LDPP to abienol is performed by a single, bifunctional class I/II *cis* abienol synthase (Ab-CAS) .

[0028] **FIG. 3** shows a HPLC chromatogram of dodecane overlay of an abienol producing strain. A C18 column was utilized. Mobile phase was methanol ethanol 4:1, 0.1% TFA with an addition of 11 % water during the first three minutes then straight methanol ethanol for 14 minutes. Samples were diluted 1:10 in mobile phase before injection. **FIG. 3a** shows the abienol peak of the commercial abienol reference sample. **FIG3b** shows the abienol peak of the abienol generated from the transformants.

DETAILED DESCRIPTION OF THE INVENTION

[0029] Unless otherwise defined herein, scientific and technical terms used herein will have the meanings that are commonly understood by those ordinary skilled in the art.

[0030] The term “class II diterpene synthase” or “class II diTPS” indicates an enzyme capable of catalyzing the conversion of geranylgeranyl diphosphate to labda-13-en-8-ol diphosphate. The class II diterpene synthase can be from various organisms, such as tobacco (*Nicotiana*) or clary sage (*Salvia*). Specific class II diterpene synthase utilized in the embodiments herein, derived from *Nicotiana tabacum* and *Salvia sclarea*, are referred to by an additional notation, e.g., “Nt-Class II-diTPS”, and “Ss-Class II-diTPS”, respectively. An example of a Nt-Class II-diTPS is the polypeptide having amino acid sequence SEQ ID NO:1. An example of a Ss-Class II-diTPS is the polypeptide having amino acid sequence SEQ ID NO:2

[0031] The term “class I diterpene synthase” or “class I diTPS” indicates an enzyme capable of catalyzing the conversion of labda-13-en-8-ol diphosphate to either abienol or sclareol. Specific class I diterpene synthase utilized in the embodiments herein, derived from *Nicotiana tabacum*, are referred to by an additional notation, e.g. “Nt-ABS”. An example of a Nt-ABS is the polypeptide having amino acid sequence SEQ ID NO:4.

[0032] The term “bifunctional class I/II abienol synthase” or “bifunctional class I/II CAS” indicates an enzyme capable of catalyzing the conversion of geranylgeranyl diphosphate to abienol. The bifunctional class I/II abienol synthase has two active sites, having the function of a class I diterpene synthase and a class II diterpene synthase, respectively. Specific bifunctional class I/II abienol synthase utilized in particular embodiments herein, derived from *Abies balsamea*, are referred to by an additional notation, e.g., “Ab-bifunctional class I/II CAS” or simple, Ab-CAS. An example of Ab-bifunctional class I/II abienol synthase is the polypeptide having amino acid sequence SEQ ID NO:3.

[0033] Sequence Identity: The relatedness between two amino acid sequences or between two nucleotide sequences is described by the parameter “sequence identity”.

[0034] For purposes of the present disclosure, the degree of sequence identity between two amino acid sequences is determined using the Needleman-Wunsch algorithm (Needleman and Wunsch, 1970, J. Mol. Biol. 48: 443-453) as implemented in the Needle program of the EMBOSS package (EMBOSS: The European Molecular Biology Open Software Suite, Rice et al., 2000, Trends Genet. 16: 276-277), preferably version 3.0.0 or later. The optional parameters

used are gap open penalty of 10, gap extension penalty of 0.5, and the EBLOSUM62 (EMBOSS version of BLOSUM62) substitution matrix. The output of Needle labeled “longest identity” (obtained using the –nobrief option) is used as the percent identity and is calculated as follows:
$$(\text{Identical Residues} \times 100) / (\text{Length of Alignment} - \text{Total Number of Gaps in Alignment})$$

[0035] For purposes of the present disclosure, the degree of sequence identity between two deoxyribonucleotide sequences is determined using the Needleman-Wunsch algorithm (Needleman and Wunsch, 1970, *supra*) as implemented in the Needle program of the EMBOSS package (EMBOSS: The European Molecular Biology Open Software Suite, Rice et al., 2000, *supra*), preferably version 3.0.0 or later. The optional parameters used are gap open penalty of 10, gap extension penalty of 0.5, and the EDNAFULL (EMBOSS version of NCBI NUC4.4) substitution matrix. The output of Needle labeled “longest identity” (obtained using the –nobrief option) is used as the percent identity and is calculated as follows:

$$(\text{Identical Deoxyribonucleotides} \times 100) / (\text{Length of Alignment} - \text{Total Number of Gaps in Alignment})$$

[0036] Nucleic acid construct: The term “nucleic acid construct” means a nucleic acid molecule, either single or double-stranded, which is isolated from a naturally occurring gene or is modified to contain segments of nucleic acids in a manner that would not otherwise exist in nature or which is synthetic. The term nucleic acid construct is synonymous with the term “expression cassette” when the nucleic acid construct contains the control sequences required for expression of a coding sequence of the present disclosure.

[0037] Control sequences: The term “control sequences” means all components necessary for the expression of a polynucleotide encoding a polypeptide of the present disclosure. Each control sequence may be native or foreign to the polynucleotide encoding the polypeptide or native or foreign to each other. Such control sequences include, but are not limited to, a leader, polyadenylation sequence, propeptide sequence, promoter, signal peptide sequence, and transcription terminator. At a minimum, the control sequences include a promoter, and transcriptional and translational stop signals. The control sequences may be provided with linkers for the purpose of introducing specific restriction sites facilitating ligation of the control sequences with the coding region of the polynucleotide encoding a polypeptide.

[0038] Operably linked: The term “operably linked” means a configuration in which a control sequence is placed at an appropriate position relative to the coding sequence of a polynucleotide such that the control sequence directs the expression of the coding sequence.

[0039] Expression: The term “expression” includes any step involved in the production of the polypeptide including, but not limited to, transcription, post-transcriptional modification, translation, post-translational modification, and secretion.

[0040] Expression vector: The term “expression vector” means a linear or circular DNA molecule that comprises a polynucleotide encoding a polypeptide and is operably linked to additional nucleotides that provide for its expression.

[0041] Host cell: The term “host cell” means any cell type that is susceptible to transformation, transfection, transduction, and the like with a nucleic acid construct or expression vector comprising a polynucleotide of the present disclosure. The term “host cell” encompasses any progeny of a parent cell that is not identical to the parent cell due to mutations that occur during replication.

[0042] Variant: The term “variant” means a polypeptide having enzyme activity comprising an alteration, i.e., a substitution, insertion, and/or deletion of one or more (several) amino acid residues at one or more (several) positions. A substitution means a replacement of an amino acid occupying a position with a different amino acid; a deletion means removal of an amino acid occupying a position; and an insertion means adding one or more amino acids adjacent to an amino acid occupying a position. In some embodiment, the above one or more amino acids are 1-3 amino acids.

[0043] In an embodiment of the method according to the first aspect of the invention, abienol is converted from geranylgeranyl diphosphate (GGPP) in the presence of a combination of class II diterpene synthase and bifunctional class I/II abienol synthase. The class II diterpene synthase according to embodiments herein may be from any organism that natively expresses an independent enzyme of class II diterpene synthase. In one embodiment, the class II diterpene synthase is from tobacco, or specifically, from *Nicotiana tabacum*. In another embodiment, the class II diterpene synthase is from clary sage, or specifically, from *Salvia sclarea*. The class II diterpene synthase according to embodiments herein may include, for example and without limitation, a polypeptide comprising an amino acid sequence having at least 50%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at

least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% identity to SEQ ID NO:1 or SEQ ID NO:2.

[0044] The bifunctional class I/II abienol synthase according to embodiments herein may be from any organism which natively expresses a single enzyme with two active sites, having class I diterpene synthase activity and class II diterpene synthase activities, respectively. In one embodiment, the bifunctional class I/II abienol synthase is from fir, or specifically, from *Abies balsamea*. The bifunctional class I/II abienol synthase according to embodiments herein may include, for example and without limitation, a polypeptide comprising an amino acid sequence having at least 50%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% identity to SEQ ID NO:3.

[0045] In an embodiment of the invention, the class II diTPS and the bifunctional class I/II abienol synthase are expressed by a recombinant host cell, such as a recombinant microorganism. Therefore, the steps of one aspect of the invention may take place within a host cell, *i.e.*, the method may be at least partially an *in vivo* method. The host cell may be recombinant and may, for example, be a genetically modified microorganism. Therefore, a microorganism may be genetically modified, *i.e.*, artificially altered from its natural state, to express both class II diTPS and bifunctional class I/II abienol synthase. In one embodiment, it expresses a combination of a Nt-Class II-diTPS and an Ab-bifunctional class I/II abienol synthase. In another embodiment, it expresses a combination of a Ss-Class II-diTPS and an Ab-bifunctional class I/II abienol synthase. Preferably, the enzymes are exogenous, *i.e.*, not present in the cell prior to modification, having been introduced using microbiological methods such as are described herein. Furthermore, in the method of the invention, the enzymes may each be expressed by a recombinant host cell, either within the same host cell or in separate host cells. The abienol may be secreted from the host cell in which it is formed.

[0046] The host cell may be genetically modified by any manner known to be suitable for this purpose by the person skilled in the art. This includes the introduction of the genes of interest, such as one or more genes encoding the bifunctional class I/II abienol synthase and class II diTPS enzymes, into a plasmid or cosmid or other expression vector which are capable of reproducing within the host cell. Alternatively, the plasmid or cosmid DNA or part of the plasmid or cosmid DNA or a linear DNA sequence may integrate into the host genome, for

example by homologous recombination or random integration. To carry out genetic modification, DNA can be introduced or transformed into cells by natural uptake or mediated by well-known processes such as electroporation. Genetic modification can involve expression of a gene under control of an introduced promoter. The introduced DNA may encode a protein which could act as an enzyme or could regulate the expression of further genes.

[0047] Such a host cell may comprise a nucleic acid sequence encoding a bifunctional class I/II abienol synthase and a class II diTPS. For example, the cell may comprise one nucleic acid sequence comprising SEQ ID NO. 3 and at least one nucleic acid sequence comprising SEQ ID NO:1 or SEQ ID NO:2, or a complement thereof, or a fragment of such a polynucleotide encoding a functional variant (which may be a fragment providing a functional variant) of any of the enzymes in bifunctional class I/II abienol synthase and class II diTPS, for example enzymes as described herein. The nucleic acid sequences encoding the enzymes may be exogenous, *i.e.*, not naturally occurring in the host cell.

[0048] Therefore, another aspect of the invention provides a recombinant host cell, such as a microorganism, comprising a first polypeptide which is a class II diTPS, for example, having an amino acid sequence at least 50%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% identical to SEQ ID NOs:1 or 2 and comprising a second polypeptide which is a bifunctional class I/II abienol synthase, for example, having an amino acid sequence at least 50%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% identical to SEQ ID NO:3. The cell may also comprise polypeptides which are functional variants or fragments of any of the above sequences.

[0049] A suitable polynucleotide may be introduced into the cell by random integration, homologous recombination and/ or may form part of an expression vector comprising a combination of polynucleotide sequences SEQ ID NOs:1 and 3, a combination of polynucleotide sequences SEQ ID NOs:1 and 2, or a complement thereof. Such an expression vector forms another aspect of the invention. Suitable vectors for construction of such an expression vector are well known in the art and may be arranged to comprise the polynucleotide operably linked to one or more expression control sequences, so as to be useful to express the required enzymes in a

host cell, for example a micro-organism as described above. For example, promoters including, but not limited to TEF1, HSP, and HYP promoters can be used in conjunction with endogenous genes and/or heterologous genes for modification of expression patterns of class II diTPS and a bifunctional class I/II abienol synthase. Similarly, exemplary terminator sequences include, but are not limited to, the use of *Yarrowia lipolytica* XPR2 terminator sequences.

[0050] In some embodiments, the recombinant or genetically modified host cell, as mentioned throughout this specification, may be any microorganism selected from the group consisting of yeast, fungi (such as members of the genus *Yarrowia*), protists, algae, bacteria, and archaea. The bacterium may comprise a gram-positive bacterium or a gram-negative bacterium including but not limited to the genera *Escherichia*, *Corynebacterium*, *Streptomyces*, *Bacillus*, *Pseudomonas*, *Paracoccus*, and *Rhodococcus*. In certain embodiments of the invention, yeast or fungi of genera including, but not limited to, *Aspergillus niger*, *Aspergillus terreus*, *Aspergillus nidulans*, *Aspergillus oryzae*, *Neurospora crassa*, *Blakeslea*, *Candida*, *Cryptococcus*, *Cunninghamella*, *Lipomyces*, *Mortierella*, *Mucor*, *Phycomyces*, *Pythium*, *Rhodosporidium*, *Rhodotorula*, *Trichosporon*, and *Yarrowia* are employed. In certain particular embodiments, organisms of species that include, but are not limited to, *Blakeslea trispora*, *Candida pulcherrima*, *C. revkaufi*, *C. tropicalis*, *Cryptococcus curvatus*, *Cunninghamella echinulata*, *C. elegans*, *C. japonica*, *Escherichia coli*, *Fusarium sporotrichioides*, *F.graminearum*, *Fusarium venenatum*, *Gibberrella zeae*, *G. fujikuroi*, *Lipomyces starkeyi*, *L. lipoferus*, *Mortierella alpina*, *M. isabellina*, *M. ramanniana*, *M. vinacea*, *Mucor circinelloides*, *Phycomyces blakesleanus*, *Pythium irregulare*, *Rhodosporidium toruloides*, *Rhodotorula glutinis*, *R. gracilis*, *R. graminis*, *R. mucilaginosa*, *R. pinicola*, *Saccharomyces cerevisiae*, *Trichosporon pullans*, *T. cutaneum*, and *Yarrowia lipolytica* are used.

[0051] Particularly suitable microorganisms, for example, include *Yarrowia lipolytica*, *Escherichia coli*, *Fusarium venenatum*, *Gibberrella fujikuroi* and *Saccharomyces cerevisiae*.

[0052] The *Yarrowia* platform has been optimized for flux through the isoprenoid pathway and has achieved production levels of greater than 10 grams per liter of total carotenoids. Since diterpenes are also derived from the isoprenoid pathway, *Yarrowia* is well suited for high level production of abienol.

[0053] The recombinant host cell or microorganism may be used to express the enzymes mentioned above and a cell free extract then obtained by standard methods, for use in the method according to the first aspect of the invention.

[0054] Embodiments of the present disclosure also encompass variants of the polypeptides as defined herein. As used herein, a “variant” means a polypeptide in which the amino acid sequence differs from the base sequence from which it is derived in that one or more amino acids within the sequence are substituted for other amino acids. For example, a variant of SEQ ID NO:1 may have an amino acid sequence at least about 50% identical to SEQ ID NO:1, for example, at least about 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical. The variants and/or fragments are functional variants/fragments in that the variant sequence has similar or identical functional enzyme activity characteristics to the enzyme having the non-variant amino acid sequence specified herein.

[0055] For example, a functional variant of SEQ ID NO:1 has similar or identical class II diTPS characteristics as SEQ ID NO:1. An example may be that the rate of conversion by a functional variant of SEQ ID NO: 1, of GGPP to LDPP, may be the same or similar, although said functional variant may also provide other benefits. For example, at least about 80%, 90%, 95%, 96%, 97%, 98%, 99% or at least about 100% the rate will be achieved when using the enzyme that is a functional variant of SEQ ID NO:1.

[0056] A functional variant or fragment of any of the above SEQ ID NO amino acid sequences, therefore, is any amino acid sequence which remains within the same enzyme category (*i.e.*, has the same EC number). Methods of determining whether an enzyme falls within a particular category are well known to the skilled person, who can determine the enzyme category without use of inventive skill. Suitable methods may, for example, be obtained from the International Union of Biochemistry and Molecular Biology.

[0057] Amino acid substitutions may be regarded as “conservative” where an amino acid is replaced with a different amino acid with broadly similar properties. Non-conservative substitutions are where amino acids are replaced with amino acids of a different type.

[0058] By “conservative substitution” is meant the substitution of an amino acid by another amino acid of the same class, in which the classes are defined as follows:

Class Amino Acid Examples:

Nonpolar: A, V, L, I, P, M, F, W

Uncharged polar: G, S, T, C, Y, N, Q

Acidic: D, E

Basic: K, R, H.

[0059] The invention extends to a novel combination of class II diTPS and a bifunctional class I/II abienol synthase for converting GGPP to abienol. Significant enhancement of conversion rate from GGPP to abienol is the result of such novel combination. In a particular embodiment as shown in Example 3, a 10-fold enhancement in abienol titer was observed when the class II diTPS from either *Nicotiana tabacum* or *Salvia sclarea* was used in combination with a bifunctional class I/II abienol synthase from *Abies balsamea*, when comparing with that a bifunctional class I/II abienol synthase being used alone.

[0060] The above-described result is unexpected. It is conventionally recognized in the art that a single bifunctional enzyme would be more efficient in catalyzing a substrate than two separate enzymes which have the same two functions, due to a scaffolding effect (Zerbe *et al.* 2012, J. Biol. Chem. 287:pp12121-12131). The reasoning is that expression of the two activities in the bifunctional enzyme is more balanced, and activities are maintained in close proximity to each other. This view is corroborated by the observation in this disclosure that a combination of class I and class II diTPS from *Nicotiana tabacum* produces little abienol compared to the class I/II bifunctional diTPS from *A. balsamea*. See Example 3.

[0061] Against the above teachings, the inventors of the present disclosure created transformants which contain a class II diTPS single function enzyme and a bifunctional class I/II abienol synthase. The results showed that the titer of abienol production is significantly greater than that produced by the two-enzyme system of class I and class II diTPS, or by a single bifunctional class I/II abienol synthase. Adding a class II diTPS enzyme in addition to a bifunctional class I/II abienol synthase does not produces merely an additive effect. Instead, a synergistic effect is achieved, which is unexpected.

[0062] The following examples are intended to illustrate the invention without limiting its scope in any way.

EXAMPLES

Example 1Construction of Vectors Containing the Class II diTPS gene, the Bifunctional Class I/II abienol synthase

[0063] Plasmids that were constructed and used in the present disclosure are shown in **Table 1**.

[0064] The bifunctional class I/II abienol synthase gene of *Abies balsamea* (Ab-CAS) was codon optimized according to *Yarrowia* codon bias, and synthesized *de novo*. During the *de novo* synthesis the sequence 5'-TGCTAGCCACAAAA, containing an *NheI* restriction site and a typical Kozak sequence for enabling efficient translation initiation, was added immediately upstream of the presumptive ATG start codon. The sequence ACGCGT-3', comprising an *MluI* restriction site, was added immediately downstream of the stop codon. This sequence was cleaved using *NheI* and *MluI* and ligated to pMB6655 cut with *NheI* and *MluI* to produce pMB6839. The resulting protein encoded by the Ab-CAS gene of pMB6839 is specified in SEQ ID No: 3. The class II diTPS gene of *Nicotinia tabacum* (Nt-Class II-diTPS) was cloned into expression vector pMB6674 as described above to create pMB6845. The protein sequence of Nt-Class II-diTPS gene encoded by pMB6845 is specified in SEQ ID No:2. The Ab-CAS gene and promoter and terminator in pMB6839 were excised by digestion with *PvuII* and *SspI* and cloned into the *EcoRV* site of pMB6845 to create pMB6847, which expresses both the Ab-CAS gene and the Nt-Class II-diTPS gene along with the hygromycin resistance gene *HPH*.

[0065] The abienol synthase gene of *Nicotinia tabacum* (Nt-ABS) was synthesized and cloned into pMB6655 as described for pMB6839, to create pMB6840. The protein sequence for the Nt-ABS gene of pMB6840 is specified in SEQ ID No:4. The Nt-ABS gene and promoter and terminator in pMB6840 was removed by digestion with *PvuII* and *SspI*, and cloned into the *EcoRV* site of pMB6845 to create pMB6849, which expresses both the Nt-ABS gene and the Nt-class II-diTPS gene along with the hygromycin resistance gene *HPH*.

[0066] The class II diTPS of *Salvia sclarea* (Ss-class II-diTPS) gene was synthesized and cloned into pMB6674 as described above to create pMB6874. The protein sequence of Ss-class II-diTPS is specified in SEQ ID No. 2. The Ab-CAS gene and expression signals of pMB6839 were excised by digestion with *PvuII* and *SspI* and cloned into the *EcoRV* site of

pMB6874 to create pMB6879, which expresses both the Ab-ABS gene and Ss-class II-diTPS gene with hygromycin selection.

[0067] The gene sequences encoding amino acid sequences SEQ ID NOs 1-4 were codon optimized according to *Yarrowia* codon bias and the resulting nucleic acid sequences are SEQ ID NOs 5-8, respectively.

Table 1. Plasmids

<i>Plasmid</i>	<i>Backbone</i>	<i>Insert (s)</i>	<i>Source</i>
pMB6839	pMB6655 (hygR)	Ab-CAS (class I/II)	Synthesized <i>NheI</i> – <i>MluI</i> fragment
pMB6845	pMB6674 (hygR)	Nt-Class II-diTPS	Synthesized <i>NheI</i> – <i>MluI</i> fragment
pMB6847	pMB6845 (hygR)	Ab-CAS (class I/II) + Nt-Class II-diTPS	Synthesized <i>NheI</i> – <i>MluI</i> fragment
pMB6840	pMB6655 (hygR)	Nt-ABS (class I)	Synthesized <i>NheI</i> – <i>MluI</i> fragment
pMB6849	pMB6845(hygR)	Nt-ABS(class I) + Nt-Class II-diTPS	Synthesized <i>NheI</i> – <i>MluI</i> fragment
pMB6874	pMB6674 (hygR)	Ss-Class II-diTPS	Synthesized <i>NheI</i> – <i>MluI</i> fragment
pMB6879	pMB6874 (hygR)	Ab-CAB(class I/II) + Ss-Class II-diTPS	Synthesized <i>NheI</i> – <i>MluI</i> fragment

[0068] All basic molecular biology and DNA manipulation procedures described herein are generally performed according to Sambrook *et al.* or Ausubel *et al.* (J. Sambrook, E.F. Fritsch, T. Maniatis (eds). 1989. *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press: New York; F.M. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, K. Struhl (eds.). 1998. *Current Protocols in Molecular Biology*. Wiley: New York).

Example 2

Construction of Recombinant Host Cells Containing Class II diTPS and Bifunctional Class I/II abienol synthase genes

[0069] In this example, the vectors containing class II diTPS and bifunctional class I/II abienol synthase genes were introduced into a host strain of *Y. lipolytica*.

[0070] The vectors described above were transformed into a *Y. lipolytica* strain (ML7206) which had previously been optimized for increased flux through the isoprenoid pathway. Host strain ML7206 is a prototrophic *Y. lipolytica* strain with the following genotype (*MATB erg9-4789::ura3 {HMG-tr GGS carB}*). Strain ML7206 was constructed by the introduction of heterologous genes under the control of constitutive promoters, coupled with several generations of crossbreeding, starting with MF350 and ATCC201249 as described in U.S. Patent No. 7,851,199. The *ERG9* gene, encoding squalene synthase, was replaced with a hypomorphic version harboring a point mutation (F317I) and an insertion of the *URA3* gene (subsequently inactivated by mutation) directly after the stop codon.

[0071] The *GGS* gene and the truncated *HMG* gene (“*HMG-tr*”) were derived from *Yarrowia* sequences corresponding to native geranylgeranyl pyrophosphate synthase and hydroxymethylglutaryl-CoA reductase genes, respectively. The *carB* gene was derived from *Mucor circinelloides*, and encodes a phytoene dehydrogenase activity, but has no bearing on the following example.

Example 3

Study of the Abienol Titer of the Recombinant Host Cells Containing a Class II diTPS gene and a Bifunctional Class I/II CAS gene

[0072] In this example, the productions of abienol in recombinant host cells described in Example 2 were examined.

[0073] The transformants described in Example 2 were grown in shake flasks on a rich medium (YPD) overlaid with a 10% volume of dodecane. Previous studies with isoprenoids have shown that the isoprenoid products are typically exported by microorganisms, and accumulate in a dodecane overlay. After growth at 28°C, 200 rpm, for 6 days, the dodecane fraction was removed from the shake flasks and analyzed by HPLC on a C18 column, with a photo-diode array detector. The HPLC set up consisted of a YMC PackPro C18 RS column [part

RS08503-1456WT 150x4.6mm S3 μ m] at a column temperature of 16°C, mobile phase consisted of a mixture of (400 mL Methanol, 100mL Ethanol, and 0.1% Trifluoro-Acetic Acid) using an isocratic flow rate of 1 mL/min. The transformants containing the abienol synthase gene generated a peak at 2.90 minutes with an optimum absorbance of 237 nm (**Figure 3b**). This peak was consistent with a reference abienol standard purchased from Toronto Research Chemicals Cat.# 107600 (**Figure 3a**).

Table 2. Abienol production in transformants of isoprenoid optimized strain MB7206.

Genes transformed into MB7206	Abienol absorbance (millions), YP+5% glucose	Abienol absorbance (millions), YP+10% glucose	Abienol absorbance (millions), YP+12.5% glucose
none	0	0	ND (Not done)
Tobacco Class I Nt ABS and Tobacco Class II Nt diTPS (plasmid pMB6849)	0.4	0.4	ND
Fir Bifunctional Class I/II Ab-CAS (plasmid pMB6839)	1.9	2.3	ND
Fir Bifunctional Class I/II Ab-CAS and tobacco Class II diTPS (plasmid pMB6847)	18.5	23.8	22.5
Fir Bifunctional Class I/II Ab-CAS and sage Class II diTPS (plasmid pMB6879) (4 clones tested)	ND	ND	17.5

[0074] As shown in **Table 2**, transformation containing both Class II diTPS gene and Class I diTPS abienol synthase (ABS) gene from tobacco into *Yarrowia* resulted in very low

production of abienol. Transformation of the bifunctional class I/II abienol synthase gene (Ab-CAS) of fir alone resulted in a slightly elevated production of abienol. Surprisingly, transformants which contain the fir bifunctional class I/II abienol synthase gene (Ab-CAS) together with either the tobacco class II di-TPS gene or the sage class II di-TPS gene produced significantly higher levels of abienol. This effect was repeatedly observed across three different medium compositions in the transformant containing the tobacco class II di-TPS gene and the Ab-CAS gene. It is unexpected to observe that the addition of the class II enzyme from either tobacco or sage could significantly surpass the activity of the bifunctional class I/II abienol synthase of fir alone, particularly given the fact that transformation of the Class II diTPS and Class I diTPS abienol synthase genes from tobacco resulted in very little product.

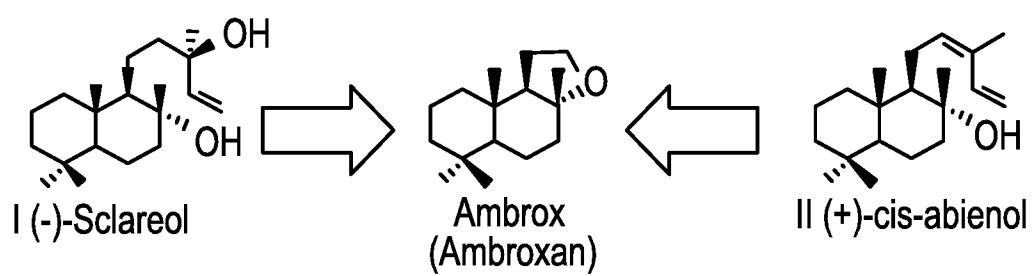
What is claimed is:

1. A method for producing abienol, which comprises converting geranylgeranyl diphosphate (GGPP) to abienol in the presence of a combination of class II diterpene synthase (diTPS) and bifunctional class I/II abienol synthase (CAS).
2. The method of claim 1, wherein the class II diTPS is a polypeptide comprising an amino acid sequence at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95% , at least 97%, or at least 99% identical to SEQ ID NO:1.
3. The method of claim 1, wherein the class II diTPS is a polypeptide comprising an amino acid sequence at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95% , at least 97%, or at least 99% identical to SEQ ID NO:2.
4. The method of claim 1, wherein the class II diTPS is from tobacco or clary sage.
5. The method of claim 1, wherein the bifunctional class I/II abienol synthase is a polypeptide comprising an amino acid sequence at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95% , at least 97%, or at least 99% identical to SEQ ID NO:3.
6. The method of claim 1, wherein the bifunctional class I/II abienol synthase is from fir.
7. The method of any of the proceeding claims, wherein at least one of said class II diTPS and said bifunctional class I/II abienol synthase is expressed by a recombinant host cell.
8. The method of claim 7, wherein the recombinant cell is a genetically modified microorganism genetically modified to express at least one exogenous polypeptide selected from the group consisting of said class II diTPS and said bifunctional class I/II abienol synthase.
9. The method of claim 8, wherein the host cell comprises at least one nucleic acid encoding one or more of the amino acid sequences selected from the group consisting of said class II diTPS and said bifunctional class I/II abienol synthase.
10. The method of claim 9, wherein the recombinant host cell is a fungus.
11. The method of claim 10, wherein the recombinant host cell is *Yarrowia lipolytica*.
12. A recombinant host cell comprising a class II diterpene synthase and bifunctional class I/II abienol synthase.
13. The recombinant host cell of claim 12, wherein the class II diTPS is a polypeptide comprising an amino acid sequence at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95% , at least 97%, or at least 99% identical to SEQ ID NO:1.

14. The recombinant host cell of claim 12, wherein the class II diTPS is a polypeptide comprising an amino acid sequence at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95% , at least 97%, or at least 99% identical to SEQ ID NO:2.
15. The recombinant host cell of claim 12, wherein the class II diTPS is from tobacco or clary sage.
16. The recombinant host cell of claim 12, wherein the bifunctional class I/II abienol synthase is a polypeptide comprising an amino acid sequence at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95% , at least 97%, or at least 99% identical to SEQ ID NO:3.
17. The recombinant host cell of claim 12, wherein bifunctional class I/II abienol synthase is from fir.
18. The recombinant host cell of claims 12-17, wherein the recombinant host cell is a fungus.
19. The recombinant host cell of claim 18, wherein the recombinant host cell is *Yarrowia lipolytica*.
20. An expression vector comprising a polynucleotide molecule encoding an amino acid sequence comprising SEQ ID NO:1 and a polynucleotide molecule encoding an amino acid sequence comprising SEQ ID NO:3.
21. An expression vector comprising a polynucleotide molecule encoding an amino acid sequence comprising SEQ ID NO:2 and a polynucleotide molecule encoding an amino acid sequence comprising SEQ ID NO:3.
22. The expression vector of claim 20 or claim 21, wherein one or both of the polynucleotide molecules are operationally linked to a transcription control sequence.

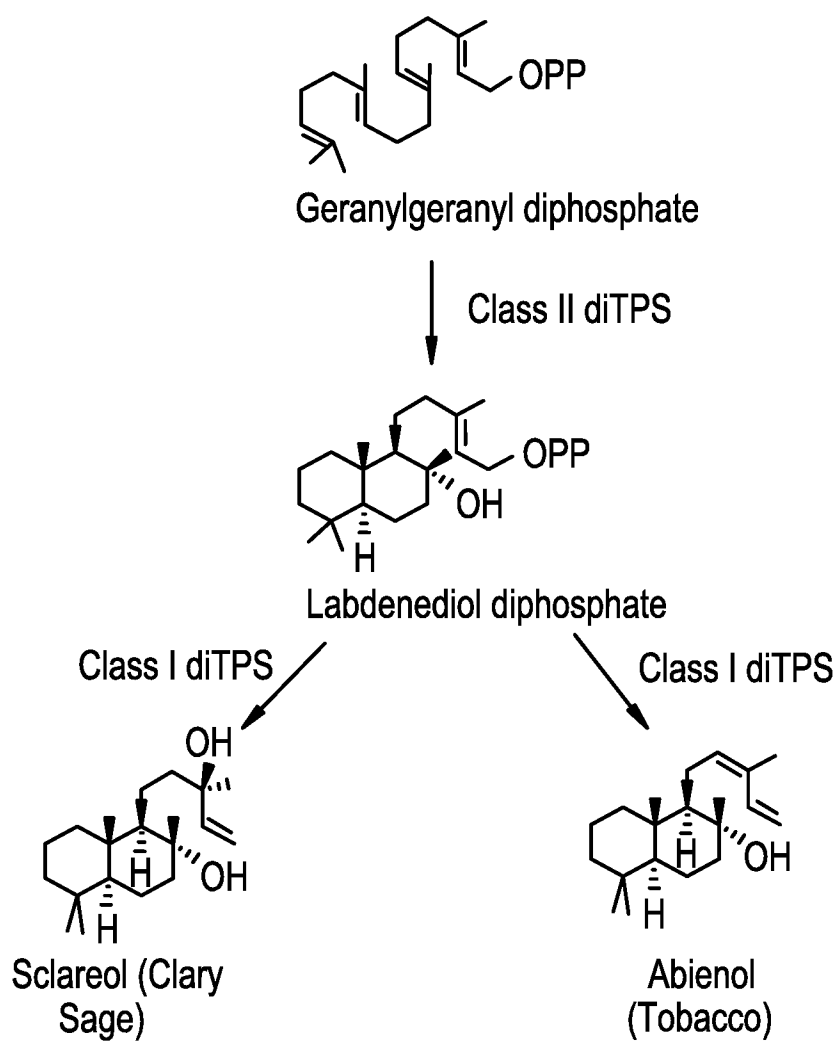
1/5

FIG. 1



2/5

FIG. 2A



3/5

FIG. 2B

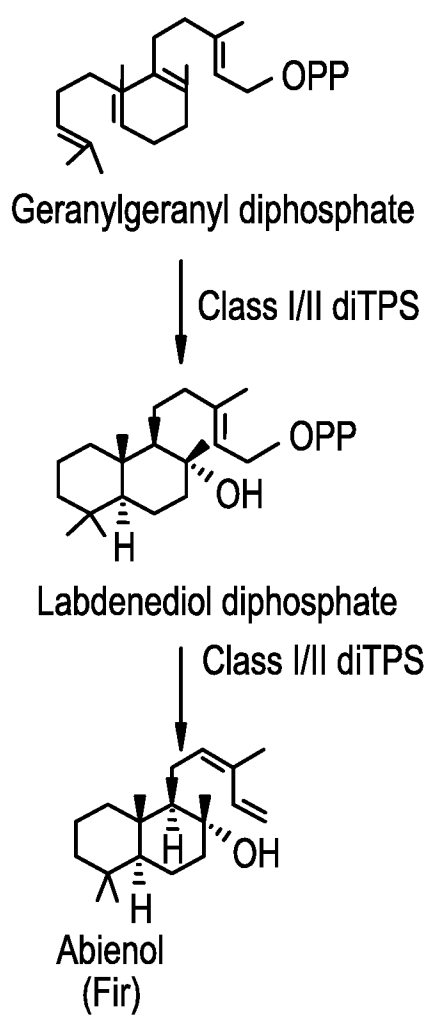


FIG. 3A

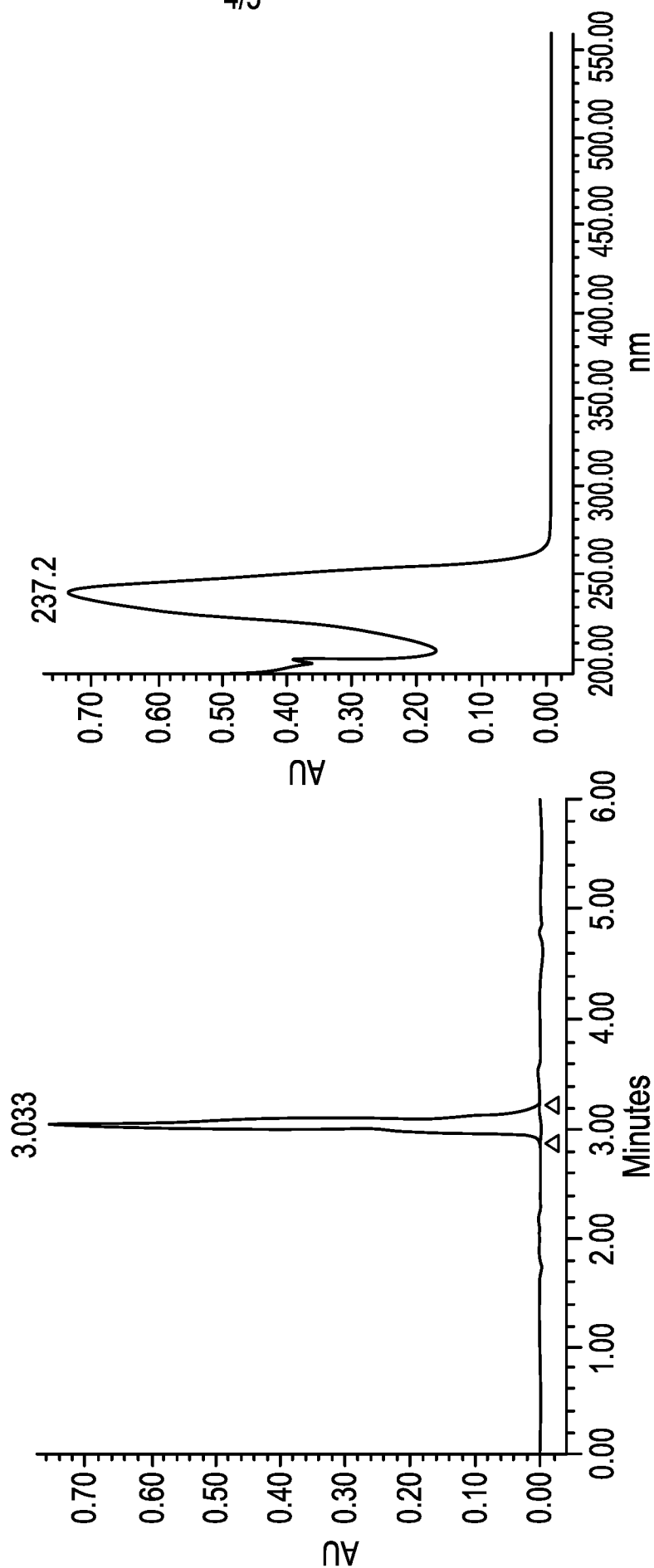
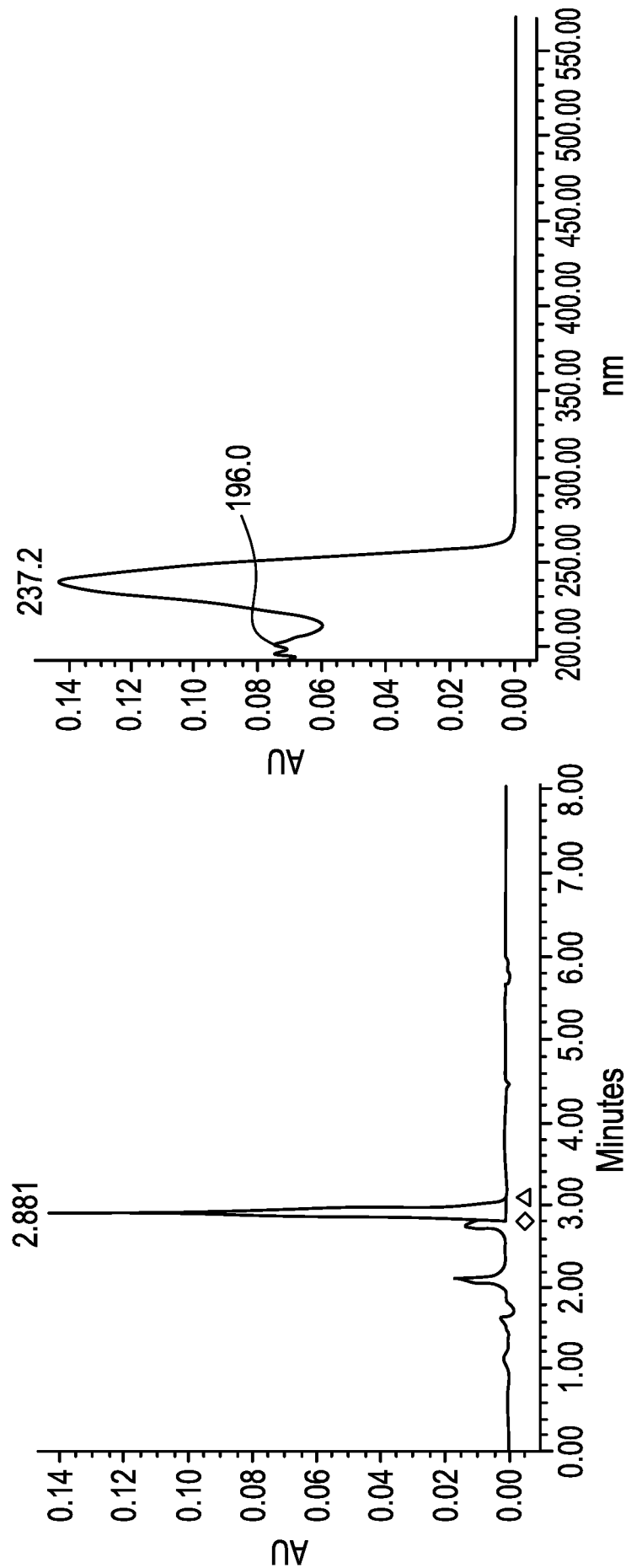


FIG. 3B



INTERNATIONAL SEARCH REPORT

International application No.

PCT/US15/63656

A. CLASSIFICATION OF SUBJECT MATTER

IPC(8) - C12N 9/88 (2016.01)

CPC - Y02P 20/52; C12N 9/88

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)

IPC(8): C12N 9/88 (2016.01)

CPC: Y02P 20/52; C12N 9/88; C12Y 402/03, 402/03032; C12P 7/02; C07D 307/92

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)

PatSeer (US, EP, WO, JP, DE, GB, CN, FR, KR, ES, AU, IN, CA, INPADOC Data); Google/Google Scholar; NCBI/BLAST/PubMed; Dialog ProQuest; abienol, 'GGPP,' 'diTPS,' 'CAS,' bifunctional, class, 'diterpene alcohol'

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X --- Y	US 2013/0224809 A1 (BOHLMANN, J et al.); August 29, 2013; paragraphs [0007], [0009], [0017], [0034], [0077], [0082], [0124]	1, 4, 6, 7/1, 7/4, 7/6, 8/7/1, 8/1/4, 8/7/6, 9/8/7/1, 9/8/7/4, 9/8/7/6, 10/9/8/7/1, 10/9/8/7/4, 10/9/8/7/6, 11/10/9/8/7/1, 11/10/9/8/7/4, 11/10/9/8/7/6, 12, 15, 17 ----- 2, 3, 5, 7/2, 7/3, 7/5, 8/7/2, 8/7/3, 8/7/5, 9/8/7/2, 9/8/7/3, 9/8/7/5, 10/9/8/7/2, 10/9/8/7/3, 10/9/8/7/5, 11/10/9/8/7/2, 11/10/9/8/7/3, 11/10/9/8/7/5, 13, 14, 16, 20-22
Y	SALLAUD, C et al. Characterization of Two Genes for the Biosynthesis of the Labdane Diterpene Z-abienol in Tobacco (Nicotiana tabacum) Glandular Trichomes: GenBank: CCD33018.1: 8-hydroxy-copalyl diphosphate synthase [Nicotiana tabacum]. 01 October 2012; GenBank: CCD33018.1.	2, 7/2, 8/7/2, 9/8/7/2, 10/9/8/7/2, 11/10/9/8/7/2, 13, 20, 22/20



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

10 February 2016 (10.02.2016)

Date of mailing of the international search report

19 FEB 2016

Name and mailing address of the ISA/

Mail Stop PCT, Attn: ISA/US, Commissioner for Patents
P.O. Box 1450, Alexandria, Virginia 22313-1450

Facsimile No. 571-273-8300

Authorized officer

Shane Thomas

PCT Helpdesk: 571-272-4300

PCT OSP: 571 272 7771

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US15/63656

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	✓ CANIARD, A et al. Discovery and Functional Characterization of Two Diterpene Synthases for Sclareol Biosynthesis in <i>Salvia sclarea</i> (L.) and Their Relevance for Perfume Manufacture: GenBank: AFU61897.1: labd-13-en-8-ol diphosphate synthase [<i>Salvia sclarea</i>]. 10 October 2012; GenBank: AFU61897.1.	3, 7/3, 8/7/3, 9/8/7/3, 10/9/8/7/3, 11/10/9/8/7/3, 14, 21, 22/21
Y	✓ ZERBE, P et al. Bifunctional cis-Abienol Synthase from <i>Abies balsamea</i> Discovered by Transcriptome Sequencing and Its Implications for Diterpenoid Fragrance Production: GenBank: AEL99953.1: diterpene synthase TPS4 [<i>Abies balsamea</i>]. 10 April 2012; GenBank: AEL99953.1.	5, 7/5, 8/7/5, 9/8/7/5, 10/9/8/7/5, 11/10/9/8/7/5, 16, 20-22

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US15/63656

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☒ Claims Nos.: 18, 19
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.