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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
12 December 2024 (12.12.2024)

WIPO | PCT

(10) International Publication Number
WO 2024/251968 A1

(51) International Patent Classification:
C12N 15/82 (2006.01) *C12N 15/70* (2006.01)

(21) International Application Number:
PCT/EP2024/065772

(22) International Filing Date:
07 June 2024 (07.06.2024)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
23386045.1 09 June 2023 (09.06.2023) EP

(71) Applicant: **UNIVERSITY OF COPENHAGEN** [DK/DK]; Nørregade 10, 1165 Copenhagen K (DK).

(72) Inventors: **KAMPRANIS, Sotirios**; C/O University of Copenhagen, Nørregade 10, 1165 Copenhagen K (DK). **ZHAO, Yong**; C/O University of Copenhagen, Nørregade 10, 1165 Copenhagen K (DK). **LIANG, Feiyan**; C/O University of Copenhagen, Nørregade 10, 1165 Copenhagen K (DK).

(74) Agent: **HØIBERG P/S**; Adelgade 12, 1304 Copenhagen K (DK).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, 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, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, 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, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:
— *of inventorship (Rule 4.17(iv))*

Published:
— *with international search report (Art. 21(3))*
— *with sequence listing part of description (Rule 5.2(a))*
— *in black and white; the international application as filed contained color or greyscale and is available for download from PATENTSCOPE*

(54) **Title:** PRODUCTION OF PLANT DITERPENOIDS

(57) **Abstract:** The present invention relates to the field of plant diterpenoid synthesis, and in particular methods and host cells therefor. More specifically, the invention relates to production of taxanes comprising an oxetane, taxanes comprising an oxetane and acetoxy group and taxanes hydroxylated at position 10, 11 and/or 13, as well as enzymes and host cells useful for such production.

WO 2024/251968 A1

Production of plant diterpenoids

Technical field

The present invention relates to the field of plant diterpenoid synthesis, and in particular methods and host cells therefore. More specifically, the invention relates to production of taxanes comprising an oxetane, taxanes comprising an oxetane and acetoxy group and taxanes hydroxylated at position 10, 11 and/or 13, as well as enzymes and host cells useful for such production.

Background

Paclitaxel, also known as taxol, is one of the most effective anticancer drugs ever developed. It is used for the treatment of melanoma, ovarian, breast, bladder, prostate and esophageal cancer (Cragg, G. M., 1998).

However, its current production is particularly costly, pressing for the development of sustainable and cost-effective production methods. Production of paclitaxel in microbial or plant cell factories will reduce its cost, increase its availability, and enable the efficient synthesis of derivatives with improved pharmacological properties. Nevertheless, progress in this direction has been hampered by incomplete knowledge of the biosynthetic pathway in the native host, the yew tree (*Taxus* spp.).

Despite extensive efforts in understanding paclitaxel biosynthesis (Guerra-Bubb et al., 2012), one key biosynthetic step, i.e. the formation of the characteristic oxetane, still remains unclear. The oxetane is one of four structural features regarded to be essential for biological activity of paclitaxel.

Summary

Hitherto paclitaxel has been semi-synthesised from more available precursors, such as 10-deacetylbaccatin III extracted from yew trees. The inventors of the present disclosure have realised production of the hitherto unrecognised paclitaxel pathway intermediates 4-hydroxy-5,20-epoxy-taxane, 4-acetoxy-5,20-epoxy-taxane, 10-hydroxy-taxadiene, 11-hydroxyl-4,12-taxadiene and 13-hydroxy-taxadiene. In particular, production of the novel intermediate 4-hydroxy-5,20-epoxy-taxane is of great importance, because it contains the characteristic oxetane of paclitaxel, as well as of the novel intermediate 4-acetoxy-5,20-epoxy-taxane, because it contains both the

characteristic oxetane ring as well as acetoxy group of paclitaxel. Thus, the invention allows for heterologous production of taxanes comprising oxetane in host cells, as well as 4-acetoxy-5,20-epoxy-taxane, 10-, 11-, and/or 13-hydroxy-taxane.

In particular, the invention provides novel methods, polypeptides and host cells for biosynthesis of the important and characteristic taxoid oxetane, taxanes comprising both the characteristic oxetane and acetoxy group, as well as taxanes hydroxylated at position 10, 11 and/or 13, are presented here.

It is a main aspect of the present disclosure to provide methods of producing a taxane comprising an oxetane, comprising the steps of:

i. providing a host cell comprising a heterologous nucleic acid encoding CYP1 of SEQ ID NO: 3 or functional homologues thereof having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto, and
ii. incubating said host cell in presence of taxadiene,

thereby producing the taxane comprising an oxetane.

A further main aspect is to provide methods of producing a 13-hydroxy-taxane, a 11-hydroxy-taxane, and/or a 10-hydroxy-taxane, comprising the steps of:

a. providing a host cell comprising a heterologous nucleic acid encoding CYP3 of SEQ ID NO: 7 or functional homologues thereof having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto, and
b. incubating said host cell in presence of taxadiene,

thereby producing the 13-hydroxy-taxane, the 11-hydroxy-taxane, and/or the 10-hydroxy-taxane.

It is also an aspect of the disclosure to provide host cells capable of producing taxane comprising an oxetane and optionally 10-hydroxy-taxane, 11-hydroxy-taxane and/or 13-hydroxy-taxane, in the presence of taxadiene, said host cell comprising a heterologous nucleic acid encoding CYP1 as set forth in SEQ ID NO: 3 and a heterologous nucleic acid encoding CYP3 as set forth in SEQ ID NO: 7, or functional homologues of any of the aforementioned having at least 70%, such as at least 80%,

for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto.

It is also an aspect of the present disclosure to host cells capable of producing 10-hydroxy-taxane, 11-hydroxy-taxane and/or 13-hydroxy-taxane, in the presence of taxadiene, said host cell comprising a heterologous nucleic acid encoding CYP3 as set forth in SEQ ID NO: 7 and a heterologous nucleic acid encoding a taxadiene synthase capable of catalysing production of taxadiene, such as TXS as set forth in SEQ ID NO: 1 and/or TXS _3xSG_MBP_SKL as set forth in SEQ ID NO: 2, or functional homologues of any of the aforementioned having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto.

Provided herein is also 4-hydroxy-5,20-epoxy-taxane, 10-hydroxy-taxadiene, 11-hydroxyl-4,12-taxadiene and/or 13-hydroxy-taxadiene, obtained by a method described herein.

Furthermore, provided is also vectors comprising one of the above nucleic acid constructs, as well as host cells comprising said nucleic acid constructs and/or vectors.

Also provided herein is a kit of parts comprising a host cell as described herein, and/or nucleic acid constructs as described herein, and/or a vector as described herein, and optionally instructions for use.

Also provided is the use of the nucleic acid constructs, vectors or host cells for production of a taxane comprising an oxetane, a 10-hydroxy-taxane, a 11-hydroxy-taxane and/or a 13-hydroxy-taxane.

Provided is also cell cultures obtained by the methods or comprising the host cells described herein.

Furthermore, also provided is a fermentation liquid comprising the taxane comprising an oxetane, 10-hydroxy-taxane, 11-hydroxy-taxane and/or 13-hydroxy-taxane, wherein said fermentation liquid is obtained by a method, comprised in a cell culture, and/or comprised within and/or secreted by a host cell described herein.

Provided is also compositions comprising a fermentation liquid described herein, and/or a taxane comprising an oxetane, a 10-hydroxy-taxane, a 11-hydroxy-taxane, and/or a 13-hydroxy-taxane obtained by a method described herein.

Further provided is methods for treating a disorder, such as cancer, comprising administration of a therapeutic sufficient amount of any one of the compounds obtained by the methods herein.

Description of Drawings

Figure 1. Production of the oxetane ring of taxol in tobacco by expressing CYP1 (SEQ ID NO: 3). UPLC-HRMS chromatograms (EIC, positive mode) of methanol extracts of tobacco (*Nicotiana benthamiana*) leaves transiently expressing taxadiene synthase (TXS, SEQ ID NO: 1), TcuCPR (SEQ ID NO: 4) and CYP1 reveal the production of 4-hydroxy-5,20-epoxy-taxane (compound **1**, taxologenic oxetane). The extract of tobacco leaves only expressing TXS was used as control and reveal no production of taxologenic oxetane (compound **1**). P19 (SEQ ID NO: 16) was co-expressed in all samples to help suppress gene silencing.

Figure 2. Production of the oxetane ring of taxol in the yeast *Saccharomyces cerevisiae* by expressing CYP1 (SEQ ID NO: 3). UPLC-HRMS chromatograms (EIC, positive mode) of ethyl acetate extracts of yeast cultures of *S. cerevisiae* TA004 expressing taxadiene synthase (TXS _3xSG_MBP_SKL, SEQ ID NO: 2), TcuCPR (SEQ ID NO: 4) and CYP1 reveal the production of taxologenic oxetane (compound **1**) in pH buffered/non-buffered media. Ethyl acetate extracts of yeast cultures of *S. cerevisiae* TA003 expressing only TXS _3xSG_MBP_SKL were used as controls and reveal no production of taxologenic oxetane (compound **1**).

Figure 3. CYP3 (SEQ ID NO: 7) oxidizes C-10, C-11 and C-13 position of taxadiene. UPLC-HRMS chromatograms (EIC, positive mode) of methanol extracts of tobacco leaves transiently expressing taxadiene synthase (TXS, SEQ ID NO: 1), TcuCPR (SEQ ID NO: 4) and CYP3 reveal the production of compounds **2**-**4**, harbouring oxidation at C-10, C-11 and C-13 position of taxadiene, respectively. The methanol extract of tobacco leaves expressing only TXS was used as control and reveal no production of

any of compounds **2-4**. P19 (SEQ ID NO: 16) was co-expressed in all samples to help suppress gene silencing.

Figure 4. (A) The key intermediate 4-hydroxy-5,20-epoxy-taxane (taxologenic oxetane, compound **1**) can be consumed by TAX19 (SEQ ID NO: 46 encoded by SEQ ID NO: 48) but not by TAT (SEQ ID NO: 45 encoded by SEQ ID NO: 47) when each of the genes encoding these was co-expressed with TXS and CYP1 in tobacco (*Nicotiana benthamiana*) leaves. **(B)** Three new peaks (**a-c**) were detected when TAX19 was co-expressed with TXS and CYP1 in tobacco (*N. benthamiana*) leaves. These peaks were not observed when TAT was co-expressed with TXS and CYP1 in tobacco (*N. benthamiana*) leaves. **(C)** Proposed biosynthetic step catalysed by TAX19 for taxol synthesis.

Figure 5. MS spectra of compounds (peaks) **a (A), b (B), and c (C)** acquired from UPLC-HRMS analysis. All of these three peaks show molecular ions corresponding to the acetylated product of the key intermediate taxologenic oxetane (compound **1**).

Detailed description

Definitions

The term “as set forth in” is herein used as equivalent to “as set out in”, “as described in”, “as depicted in”, “being the same as”, “of”, and/or “according to”. For example “[..] CYP1 as set forth in SEQ ID NO: 3” herein means, that CYP1 has the sequence of SEQ ID NO: 3. Another example is “[..] 4-hydroxy-5,20-epoxy-taxane as set forth in (IV)”, which herein implies that structure (IV) is the structure of the 4-hydroxy-5,20-epoxy-taxane. It follows that “as set forth in” denotes that something “is the same as” and is therefore to be construed as “being limited to”.

As used herein, the singular forms “a”, “an” and “the” include plural referents unless the context clearly states otherwise.

The terms “incubating” and “cultivating” are used interchangeably herein, and refers to maintaining host cells under culture conditions, which allow the cells to grow. Preferably, said culture conditions allow expression of the enzyme(s) encoded by the heterologous gene(s) contained in said host cells. Preferably, the host cells are incubated under culture conditions allowing said host cells to produce a taxane

comprising an oxetane, a 10-hydroxy-taxane, a 11-hydroxy-taxane and/or a 13-hydroxy-taxane. In embodiments where the host cell is contained within a multicellular organism (e.g. a plant), “cultivating” or “incubation” refers to maintaining said multicellular organisms under conditions allowing said multicellular organism to grow. In embodiments where the host cell is a unicellular organism, “cultivating” or “incubating” refers to maintaining said unicellular organism under conditions allowing said unicellular organism to grow and/or multiply.

The term "enzyme" as used herein refers to proteins or polypeptides, which are capable of catalysing biochemical reactions. Further, unless context dictates otherwise, as used herein "enzyme" includes protein fragments that retain the relevant catalytic activity, and may include artificial enzymes synthesized to retain the relevant catalytic activity.

The term "functional homologue" of an amino acid sequence, refers to a polypeptide comprising said amino acid sequence with the proviso that one or more amino acids are substituted, deleted, added, and/or inserted, and which polypeptide has (qualitatively) the same enzymatic functionality for substrate conversion. The term "functional homologue" of a nucleic acid encoding a polypeptide, refers to a nucleic acid comprising said nucleic acid sequence with the proviso that one or more nucleobases are substituted, deleted, added, and/or inserted, and which nucleic acid encodes a polypeptide, which polypeptide has (qualitatively) the same enzymatic functionality for substrate conversion as the polypeptide encoded by said nucleic acid. Nucleic acids or nucleic acid sequence may also be referred to as polynucleotides. Preferably, a functional homologue shares at least 70% sequence identity, preferably at least 80%, preferably at least 85% sequence identity, preferably at least 90% sequence identity, preferably at least 95% sequence identity, more preferred at least 98% sequence identity to said amino acid sequence.

The term “heterologous nucleic acid” refers to a nucleic acid, which has been inserted into a host cell or into a progenitor of the host cell, e.g. by recombinant or transgenic methods. The respective protein or RNA encoded by a heterologous nucleic acid is also referred to as "heterologous”. The heterologous nucleic acid may be part of a non-integrated nucleic acid, e.g. a vector, including but not limited to a plasmid. Preferably, the heterologous nucleic acid(s) are integrated into the host cell genome.

The term "host cell" refers to a cell, which comprises one or more heterologous nucleic acids.

The term "polypeptide" as used herein refers a sequential chain of amino acids linked together via peptide bonds. The term is used to refer to an amino acid chain of any length. As is known to those skilled in the art, polypeptides may be processed and/or modified, and the term polypeptide may refer to both unmodified or modified polypeptides.

The term “sequence identity” as used herein describes the relatedness between two amino acid sequences or between two nucleotide sequences, i.e. a candidate sequence (e.g. a mutant sequence) and a reference sequence (such as a wild type sequence) based on their pairwise alignment. For purposes of the present invention, the sequence identity between two amino acid sequences is determined using the Needleman-Wunsch algorithm (Needleman and Wunsch, 1970, J. Mo/. 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 5.0.0 or later (available at https://www.ebi.ac.uk/Tools/psa/emboss_needle/). The 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:

(Identical Residues x 100)/(Length of Alignment - Total Number of Gaps in Alignment)

The Needleman-Wunsch algorithm is also used to determine whether a given amino acid in a sequence other than the reference sequence corresponds to a given position in a reference sequence.

For purposes of the present invention, the sequence identity between two nucleotide 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, Trends Genet. 16: 276-277), preferably version 5.0.0 or later. The parameters used are gap open penalty of 10, gap extension penalty of 0.5, and the DNAFULL

(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:

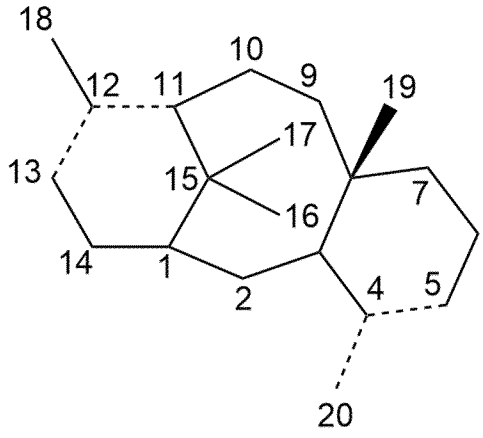
(Identical Deoxyribonucleotides x 100)/(Length of Alignment - Total Number of Gaps in Alignment). Sequence identity is calculated over the entire length of the reference sequence.

The term “oxetane” refers to the stable chemical motif or unit of the structure:

(I)

Oxetane is a cyclic ether, in other words a small, polar heterocycle. The terms “oxetane” and “oxetane ring” are used interchangeably throughout the description.

The term “taxane” herein comprises a class of diterpenoid compounds comprising a common core skeleton set forth in (II):



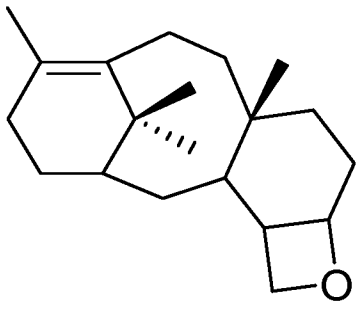
(II)

and wherein the core may further be substituted, preferably the core may be further substituted at positions 1, 2, 4, 7, 9, 10, 11, 13, and/or 20. For instance, position 1, 10, 11 and/or 13 may be substituted with a hydroxyl group. With regards to the structure set forth in (II), the dashed lines indicates either a single or a double bond, with the proviso that only one of the bonds between position 11 and 12 or position 12 and 13 may be a double bond simultaneously. The atom at position 20 may be carbon (C) or oxygen (O) and may further be substituted. Taxoids are taxadiene-derived diterpenoids. A preferred type of taxanes are taxoid compounds and/or taxoids. Taxanes have various structures and may be substituted with different groups, such as

different functional groups. Non-limiting examples of taxanes of structure I are paclitaxel, 10-deacetyltaxol, 10-deacetyl-baccatin III, and/or baccatin III, taxadiene, or 5,20-epoxy-taxane. Unless otherwise specified the numbering of atoms of taxanes used herein is as indicated in formula (II).

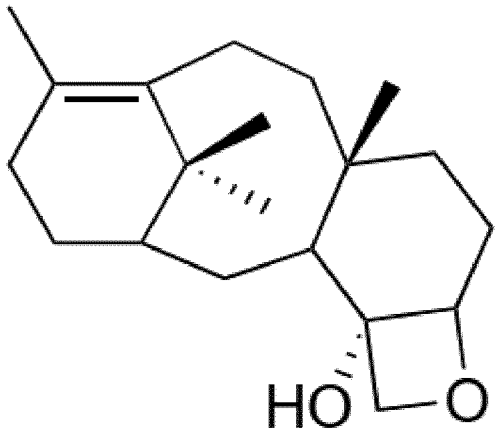
In agreement with the above definition, the term “a taxane comprising oxetane ” refers to a taxane of structure (II) comprising an oxetane moiety. Preferably, said oxetane is incorporating the carbons at position 4 to 5, more preferably the carbons at position 4, 5 and 20. In other words, the carbons at position 5 and 20 may also be linked to oxygen and form an oxetane moiety.

In agreement with the above definition, the term “5,20-epoxy-taxane” refers to a taxane with an oxetane incorporating carbons at position 4, 5 and 20 of structure (II), of the following structure:



(III)

5,20-epoxy-taxane may further be substituted, preferably the 5,20-epoxy-taxane may further be substituted at position 1, 2, 4, 7, 9, 10, 11, 13, and/or 20. In particular, a 5,20-epoxy-taxane may be 4-hydroxy-5,20-epoxy-taxane as set forth in (IV):

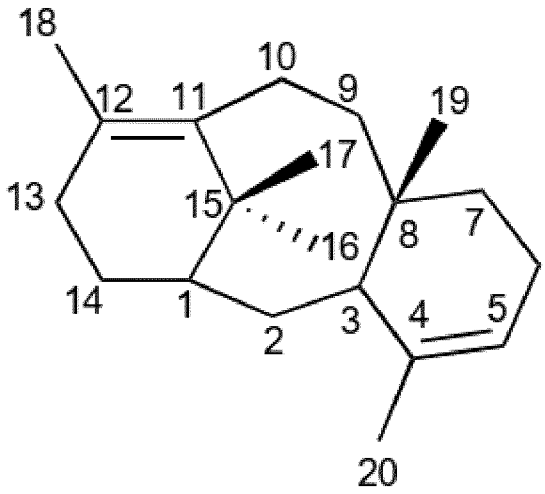


(IV)

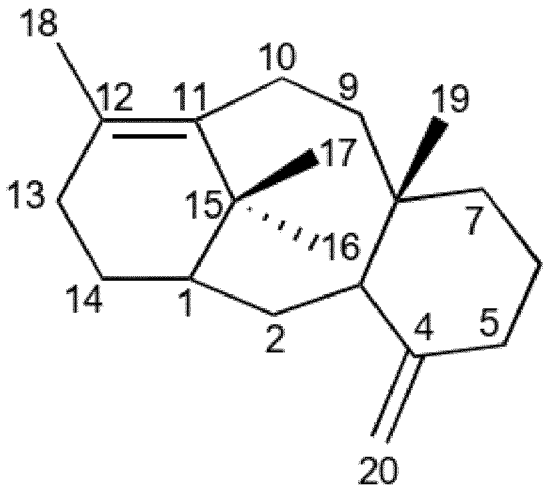
4-hydroxy-5,20-epoxy-taxane is herein also referred to as compound **1** and/or taxologenic oxetane herein, and the terms may be used interchangeably.

In agreement with the above definition, the terms “10-hydroxy-taxane”, “11-hydroxy-taxane” and “13-hydroxy-taxane” refers to taxanes with a hydroxyl group at position 10, 11 and 13 of the structure set forth in (II), respectively.10-hydroxy-taxane may further be substituted at position 1, 2, 4, 7, 9, 11, 13, and/or 20, 11-hydroxy-taxane may further be substituted at position 1, 2, 4, 7, 9, 10, 13, and/or 20, and 13-hydroxy-taxane may further be substituted at position 1, 2, 4, 7, 9, 10, 11, and/or 20.

The term “taxadiene” herein comprises the isomers taxa-4(5),11(12)-diene (endotaxadiene) and/or taxa-4(20),11(12)-diene (exotaxadiene) as set forth in (V) and (VI), respectively:

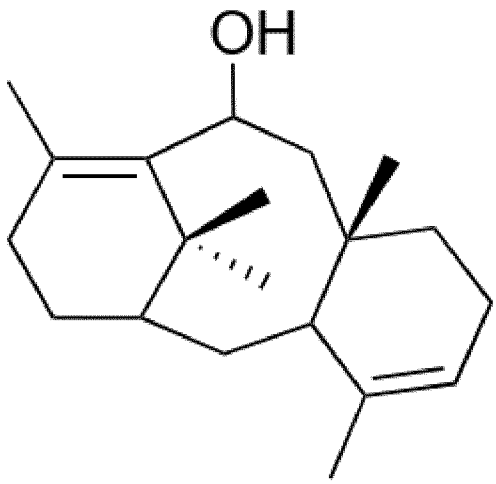


(V),



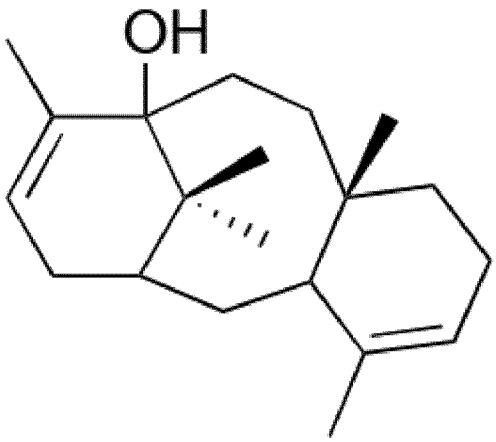
(VI)

The term “10-hydroxy-taxadiene” herein refers to the compound as set forth in (VII):



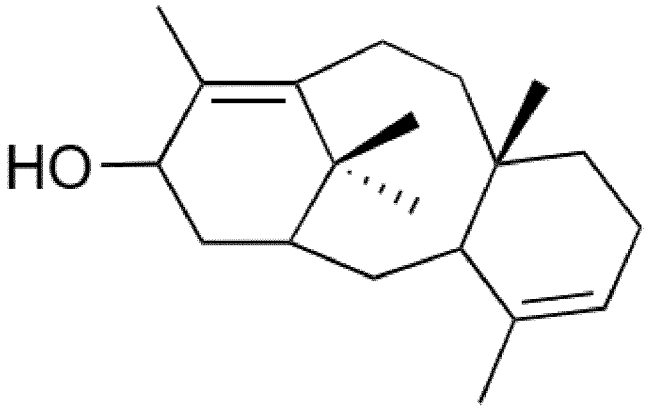
(VII)

The term “11-hydroxy-taxa-12(13)-diene” herein refers to the compound as set forth in (VIII):



(VIII)

The term “13-hydroxy-taxadiene” herein refers to the compound as set forth in (IX):



(IX)

The terms “cytochrome P450 reductase” and “CPR” are used interchangeably throughout the description, and comprises an enzyme capable of transferring electrons from NADPH to cytochrome P450 with the EC number 1.6.2.4. Alternative names are “P450 reductase”, “POR”, “CYPOR”, “NADPH:P450 oxidoreductase”, “NADPH:ferrihemoprotein oxidoreductase” and “NADPH:hemoprotein oxidoreductase”.

The terms “cytochrome P450”, “CYP”, “cytochrome P450 enzyme”, “CYP enzyme”, “P450” and “CYP450” are used interchangeably throughout the description.

Cytochrome P450 enzyme

The present disclosure provides methods, polypeptides and host cells for production of diterpenoids, notably taxanes comprising an oxetane and/or 10-hydroxy-taxanes, 11-hydroxy-taxanes and/or 13-hydroxy-taxanes. The host cell to be used with the present disclosure comprises a heterologous nucleic acid encoding an enzyme capable of

catalysing formation of a taxane comprising an oxetane and/or an enzyme capable of catalysing formation of a 10-hydroxy-taxane, a 11-hydroxy-taxane and/or a 13-hydroxy-taxane.

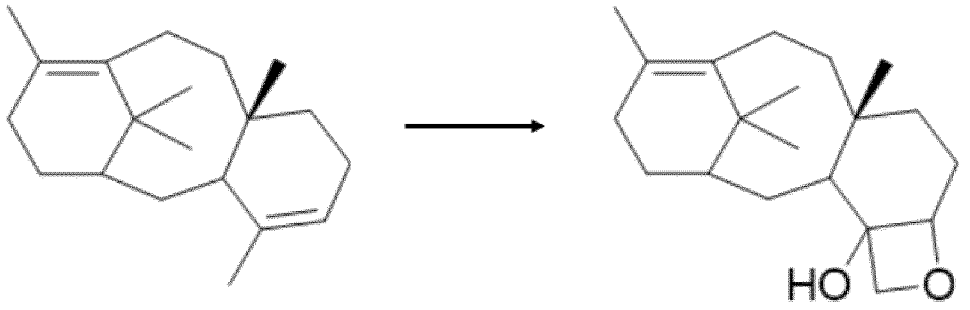
Production of said taxoid compounds may be obtained by expression of a cytochrome P450 enzyme, either alone or together with one or more polypeptides of different activity as disclosed herein below, for example in the section “Combination of polypeptides”, in a host cell of the present disclosure. The cytochrome P450 enzyme preferably have the enzyme activity described in this section.

The cytochrome P450 enzyme(s) described herein are capable of catalysing formation of a taxane comprising an oxetane and/or an enzyme capable of catalysing formation of a 10-hydroxy-taxane, a 11-hydroxy-taxane and/or a 13-hydroxy-taxane from a starting taxane. Said starting taxane may be taxadiene or a taxane as set forth in structure (II), which is substituted in one or more positions. Examples of starting taxanes are provided herein.

Enzymes catalysing formation of oxetane (CYP1)

The host cell to be used with the present disclosure may preferably comprise a heterologous nucleic acid encoding an enzyme capable of catalysing formation of a taxane comprising an oxetane. The invention also provides such enzymes as well as method using same.

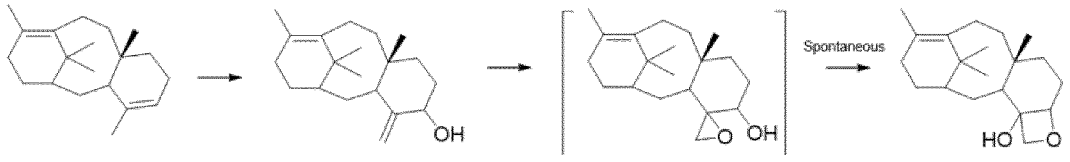
It is preferred that said enzyme is capable of catalysing formation of an oxetane incorporating carbons at position 4 to 5 of taxadiene. Even more preferably, said enzyme is capable of catalysing formation of an oxetane incorporating carbons at positions 4, 5 and 20 of taxadiene. In other words, it is preferred that said enzyme is capable of catalysing the following reaction (A):



(A)

Thus, the enzyme capable of catalysing formation of a taxane comprising an oxetane may be capable of catalysing reaction A outlined above.

It is comprised within the invention, that reaction A may be the sum of one or more reactions, possible also comprising one or more spontaneous reactions. Thus, the enzyme capable of catalysing formation of an oxetane incorporating carbons at position 4, 5 and 20 of taxadiene may be capable of catalysing the following reaction series (B):



(B)

Thus, the enzyme capable of catalysing formation of a taxane comprising an oxetane may be capable of catalysing reaction series B outlined above.

In some embodiments the enzyme is further capable of catalysing similar reactions to A or similar series to B, wherein the starting taxane is substituted at one or more positions compared to the structure above. For example, the starting taxane may be a taxane of structure (II) as set forth above, and further comprise a hydroxyl group at position 10, 11 and/or 13. In some embodiments the enzyme is further capable of catalysing a similar reaction, wherein the starting taxane is substituted at one or more positions compared to the structure above. For example, the starting taxane may be a taxane of structure (II), and further comprise a hydroxyl group at position 10, 11 and/or 13. In some embodiments, the starting taxane is a 10-hydroxy-taxane such as 10-hydroxy-taxadiene. In other embodiments the starting taxane is a 11-hydroxy-taxane

such as 11-hydroxyl-4,12-taxadiene. In further other embodiments, the starting taxane is a 13-hydroxy-taxane such as 13-hydroxy-taxadiene.

A taxane comprising an oxetane may be a 5,20-epoxy-taxane. In particular, a taxane comprising an oxetane is 4-hydroxy-5,20-epoxy-taxane as set forth in (IV) herein above.

The enzyme capable of catalysing formation of a taxane comprising an oxetane may be any useful enzyme with above-mentioned activities, in particular said enzyme may be a cytochrome P450 enzyme. The enzyme capable of catalysing formation of a taxane comprising an oxetane may be derived from any suitable source, but in a preferred embodiment, said enzyme is an enzyme from *Taxus cuspidata.* Thus, the enzyme capable of catalysing formation of a taxane comprising an oxetane may be a cytochrome P450 enzyme from *Taxus cuspidata.*

In some embodiments of the present disclosure, the host cell comprises a heterologous nucleic acid encoding CYP1. Said CYP1 is preferably CYP1 of SEQ ID NO: 3 or a functional homologue thereof. The person skilled in the art will appreciate that CYP1 or a functional homologue thereof preferably have the ability to convert taxadiene to 4-hydroxy-5,20-epoxy-taxane as outlined in reaction A and/or B depicted herein above.

A functional homologue of CYP1 of SEQ ID NO: 3 preferably has at least 70% sequence identity, preferably at least 80%, preferably at least 85% sequence identity, preferably at least 90% sequence identity, preferably at least 95% sequence identity, more preferred at least 98% sequence identity thereto.

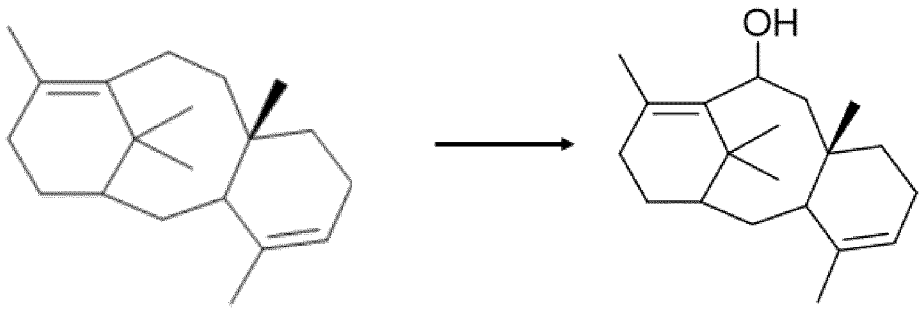
The heterologous nucleic acid encoding CYP1 of SEQ ID NO: 3 or a functional homologue thereof, may have any sequence encoding said CYP1. In some embodiments, CYP1 may be encoded by the nucleic acid with accession number AY289209.2 (NCBI GenBank, 4 June 2023) or a functional homologue thereof, for example an engineered or codon-optimised variant thereof. In some embodiments, the nucleic acid encoding CYP1 is a nucleic acid as set forth in SEQ ID NO: 10, or a functional homologue thereof encoding a functional homologue of CYP1 having at least 70% sequence identity to the CYP1 encoded by SEQ ID NO:10. In some embodiments, the nucleic acid encoding CYP1 or a functional homologue thereof

having at least 70% sequence identity thereto is a nucleic acid as set forth in SEQ ID NO: 11, or a functional homologue thereof encoding a functional homologue of CYP1 having at least 70% sequence identity to the CYP1 encoded by SEQ ID NO:11.Thus, CYP1 as set forth in SEQ ID NO: 3 and/or a functional homologue thereof having at least 70% sequence identity thereto may be encoded by the nucleic acid as set forth in SEQ ID NO: 10 and/or SEQ ID NO: 11, or a functional homologue of any of the aforementioned having at least 70% sequence identity thereto.

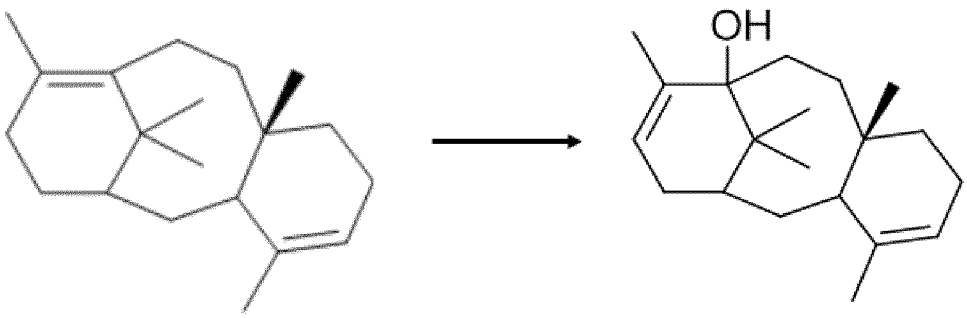
Enzymes catalysing hydroxylation at position 10, 11 and/or 13 (CYP3)

The host cell to be used with the present disclosure may preferably also comprise a heterologous nucleic acid encoding an enzyme capable of catalysing formation of a 10-hydroxy-taxane, a 11-hydroxy-taxane and/or a 13-hydroxy-taxane. The invention also provides such enzymes as well as methods using same.

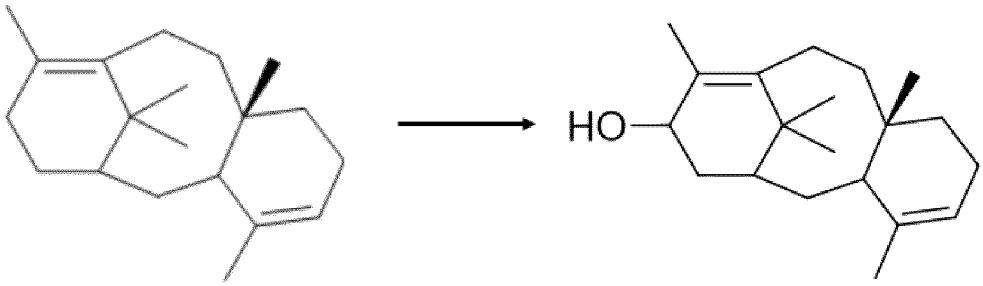
It is preferred that said enzyme is capable of catalysing hydroxylation of position 10, 11 and/or 13 of a starting taxane, such as taxadiene. Said hydroxylation may involve a rearrangement, e.g. a rearrangement of double bonds. In other words, it is preferred that said enzyme is capable of catalysing at least one of the following reactions (C, D and E):



(C)



(D)



(E)

In some embodiments the enzyme is further capable of catalysing similar reactions, wherein the starting taxane is substituted at one or more positions compared to the structure above. For example the starting taxane may comprise an oxetane, for example the starting taxane may be a 5,20-epoxy-taxane such as 4-hydroxy-5,20-epoxy-taxane.

Thus, the enzyme capable of catalysing formation of a 10-hydroxy-taxane, a 11-hydroxy-taxane and/or a 13-hydroxy-taxane may be capable of catalysing hydroxylation of taxadiene at position 10 to form 10-hydroxy-taxadiene as depicted in reaction C, hydroxylation of taxadiene at position 11 to form 11-hydroxyl-4,12-taxadiene as depicted in reaction D and/or hydroxylation of taxadiene at position 13 to form 13-hydroxy-taxadiene as depicted in reaction E.

The enzyme capable of catalysing formation of a 10-hydroxy-taxane, a 11-hydroxy-taxane and/or a 13-hydroxy-taxane may also be capable of catalysing hydroxylation of 4-hydroxy-5,20-epoxy-taxane at position 10 to form a 4-hydroxy-10-hydroxy-5,20-epoxy-taxane and/or a 10-hydroxy-5,20-epoxy-taxane, hydroxylation of 4-hydroxy-5,20-epoxy-taxane at position 11 to form a 4-hydroxy-11-hydroxy-5,20-epoxy-taxane, a 11-hydroxy-5,20-epoxy-taxane and/or 4-hydroxy-12-hydroxy-5,20-epoxy-12,13-taxane, and/or hydroxylation of 4-hydroxy-5,20-epoxy-taxane at position 13 to form a 4-hydroxy-13-hydroxy-5,20-epoxy-taxane and/or a 13-hydroxy-5,20-epoxy-taxane.

Examples of a 10-hydroxy-taxane, a 11-hydroxy-taxane and a 13-hydroxy-taxane are is 10-hydroxy-taxadiene as set forth in (VII), 11-hydroxyl-4,12-taxadiene as set forth in (VIII) and 13-hydroxy-taxadiene as set forth in (IX), respectively.

The enzyme capable of catalysing formation of a 10-hydroxy-taxane, a 11-hydroxy-taxane and/or a 13-hydroxy-taxane may be any useful enzyme with above-mentioned activities, in particular said enzyme may be a cytochrome P450 enzyme. The enzyme

capable of catalysing formation of a 10-hydroxy-taxane, a 11-hydroxy-taxane and/or a 13-hydroxy-taxane may be derived from any suitable source, but in a preferred embodiment, said enzyme is an enzyme from *Taxus cuspidata.* Thus, the enzyme capable of catalysing formation of a taxane comprising an oxetane may be a cytochrome P450 enzyme from *Taxus cuspidata.*

In some embodiments of the present disclosure, the host cell comprises a heterologous nucleic acid encoding CYP3. Said CYP3 is preferably CYP3 of SEQ ID NO: 7 or a functional homologue thereof. The person skilled in the art will appreciate that CYP3 or a functional homologue thereof preferably have the ability to convert taxadiene to 10-hydroxy-taxadiene, 11-hydroxyl-4,12-taxadiene and/or 13-hydroxy-taxadiene as outlined in reaction C, D and E, respectively, depicted herein above.

A functional homologue of CYP3 of SEQ ID NO: 7 preferably has at least 70% sequence identity, preferably at least 80%, preferably at least 85% sequence identity, preferably at least 90% sequence identity, preferably at least 95% sequence identity, more preferred at least 98% sequence identity thereto.

The heterologous nucleic acid encoding CYP3 of SEQ ID NO: 7 or a functional homologue thereof, may have any sequence encoding said CYP3. In some embodiments, the nucleic acid encoding CYP3 is a nucleic acid as set forth in SEQ ID NO: 15, or a functional homologue thereof encoding a functional homologue of CYP3 having at least 70% sequence identity to the CYP3 encoded by SEQ ID NO:15.

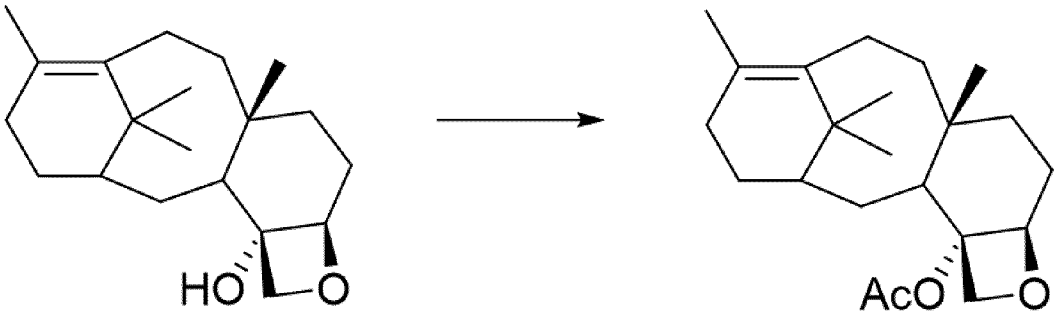
O-acetyltransferase

In another aspect, disclosed herein is a host cell comprising a nucleic acid, such as a heterologous nucleic acid, encoding an acetyltransferase capable of catalysing acetylation of an hydroxy-group of a taxane comprising an oxetane, i.e. capable of catalysing formation of an *O*-acetylated taxane comprising an oxetane from a taxane comprising an oxetane. The invention also provides such enzymes as well as method using same.

Said *O*-acetylated taxane comprising an oxetane may be a taxane comprising an oxetane according to structure (III) comprising an acetoxy group (AcO) at position 4. In

particular, said *O*-acetylated taxane comprising an oxetane may be 4-acetoxy-5,20-epoxy-taxane as set forth in structure (X). Structure (X) is shown in reaction (F).

It is preferred that said enzyme is capable of catalysing *O*-acetylation of a taxane comprising an oxetane, such as *O*-acetylation of the hydroxy group of 4-hydroxy-5,20-epoxy-taxane as set forth in (IV). It is preferred that said enzyme is capable of catalysing at least the following reaction (F):



(X) (F)

The enzyme capable of catalysing formation of an *O*-acetylated taxane comprising an oxetane may be any useful enzyme with above-mentioned activity, in particular said enzyme may be an acetyltransferase. The enzyme capable of catalysing formation of an *O*-acetylated taxane comprising an oxetane may be derived from any suitable source, but in a preferred embodiment, said enzyme is an enzyme from *Taxus chinensis.* Thus, the enzyme capable of catalysing formation of an *O*-acetylated taxane comprising an oxetane may be an acetyltransferase from *Taxus chinensis.*

In some embodiments of the present disclosure, the host cell comprises a heterologous nucleic acid encoding TAX19. Said TAX19 is preferably TAX19 of SEQ ID NO: 46 or a functional homologue thereof. According to the present disclosure, TAX19 or a functional homologue thereof is capable of acetylated a taxane comprising an oxetane, such as converting an hydroxy-group of a taxane comprising an oxetane into an acetoxy-group. In particular, TAX19 or a functional homologue thereof is capable of converting a 4-hydroxy-5,20-epoxy-taxane (structure IV) into a 4-acetoxy-5,20-epoxy-taxane (structure (X)).

Thus, host cells comprising a heterologous nucleic acid encoding TAX19 of SEQ ID NO: 46 or a functional homologue thereof are capable of producing an *O*-acetylated taxane comprising an oxetane, such as 4-acetoxy-5,20-epoxy-taxane, from a taxane comprising an oxetane.

A functional homologue of TAX19 of SEQ ID NO: 46 preferably has at least 70% sequence identity, preferably at least 80%, preferably at least 85% sequence identity, preferably at least 90% sequence identity, preferably at least 95% sequence identity, more preferred at least 98% sequence identity to SEQ ID NO: 46.

The heterologous nucleic acid encoding TAX19 of SEQ ID NO: 46 or a functional homologue thereof, may have any sequence encoding said TAX19. In some embodiments, the heterologous nucleic acid encoding TAX19 is a nucleic acid as set forth in SEQ ID NO: 48, or a functional homologue thereof encoding a functional homologue of TAX19 having at least 70% sequence identity to the TAX19 encoded by SEQ ID NO: 48.

Thus, in another aspect the present disclosure, relates to a host cell comprising a heterologous nucleic acid encoding TAX19 as set forth in SEQ ID NO: 46, or functional homologues thereof having at least 70% sequence identity, such as at least 80%, such as at least 90%, such as at least 95%, such as at least 99% sequence identity thereto.

For example, provided herein is a host cell capable of producing an *O*-acetylated taxane comprising an oxetane, wherein the host cell comprises a heterologous nucleic acid encoding an acetyltransferase, such as TAX19 as set forth in SEQ ID NO: 46, or functional homologues thereof having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto. Said *O*-acetylated taxane comprising an oxetane may be an *O*-acetylated 5,20-epoxy-taxane. In preferred embodiments, said *O*-acetylated taxane comprising an oxetane is 4-acetoxy-5,20-epoxy-taxane.

Method

In a main aspect, the present disclosure concerns methods for preparing a taxane comprising an oxetane, a 10-hydroxy-taxane, a 11-hydroxy-taxane and/or a 13-hydroxy-taxane. The methods of the invention generally comprise the steps of:

i. providing a host cell capable of producing a taxane comprising an oxetane, a 10-hydroxy-taxane, a 11-hydroxy-taxane and/or a 13-hydroxy-taxane, wherein said host cell comprises a heterologous nucleic acid encoding an enzyme capable of catalysing formation of a taxane comprising an oxetane and/or an enzyme capable of catalysing formation of a 10-hydroxy-taxane, a 11-hydroxy-taxane and/or a 13-hydroxy-taxane,

ii. incubating said host cell under conditions allowing said host cell to produce the taxane comprising an oxetane, the 10-hydroxy-taxane, the 11-hydroxy-taxane and/or the 13-hydroxy-taxane and/or growth of said host cell.

A taxane is defined herein above in the section “Definitions”.

Preferably said enzyme is a cytochrome P450 enzyme, most preferably any one of the enzymes described herein above in the section “Cytochrome P450 enzyme”.
Preferably, said enzyme capable of catalysing formation of a taxane comprising an oxetane is CYP 1 of SEQ ID NO: 3 or a functional homologue thereof having at least 70% sequence identity, preferably at least 80%, preferably at least 85% sequence identity, preferably at least 90% sequence identity, preferably at least 95% sequence identity, more preferred at least 98% sequence identity thereto. Preferably, said enzyme capable of catalysing formation of a 10-hydroxy-taxane, a 11-hydroxy-taxane and/or a 13-hydroxy-taxane is CYP3 of SEQ ID NO: 7 or a functional homologue thereof having at least 70% sequence identity, preferably at least 80%, preferably at least 85% sequence identity, preferably at least 90% sequence identity, preferably at least 95% sequence identity, more preferred at least 98% sequence identity thereto.

The host organism may be any of the host organisms described herein below in the section “Host cell”.

<u>Preferred embodiments</u>

In a preferred embodiment, the method is for producing a taxane comprising an oxetane, said method comprising the steps of;

i. providing a host cell comprising a heterologous nucleic acid encoding CYP1 of SEQ ID NO: 3 or functional homologues thereof having at least 70%,

such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto, and

ii. incubating said host cell in presence of taxadiene,

thereby producing the taxane comprising an oxetane.

In another preferred embodiment, the method is for producing a 5,20-epoxy-taxane, said method comprising the steps of:

i. providing a host cell comprising a heterologous nucleic acid encoding CYP1 of SEQ ID NO: 3 or functional homologues thereof having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto, and

ii. incubating said host cell in presence of taxadiene,

thereby producing the 5,20-epoxy-taxane.

In another preferred embodiment, the method is for producing 4-hydroxy-5,20-epoxy-taxane, said method comprising the steps of:

iii. providing a host cell comprising a heterologous nucleic acid encoding CYP1 of SEQ ID NO: 3 or functional homologues thereof having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto, and

iv. incubating said host cell in presence of taxadiene,

thereby producing the 4-hydroxy-5,20-epoxy-taxane.

In another preferred embodiment, the method is for producing 4-hydroxy-5,20-epoxy-taxane, said method comprising the steps of:

i. providing a host cell comprising a heterologous nucleic acid encoding CYP1 of SEQ ID NO: 3, and a heterologous nucleic acid encoding a taxadiene synthase (EC 4.2.3.17) capable of converting geranylgeranyl diphosphate (GGPP) into taxadiene, such as TXS of SEQ ID NO: 1 and/or TXS _3xSG_MBP-SKL of SEQ ID NO: 2,
or functional homologues of any of the aforementioned having at least 70% sequence identity, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto.

ii. incubating said host cell, preferably in the presence of geranylgeranyl diphosphate (GGPP),

thereby producing the 4-hydroxy-5,20-epoxy-taxane.

In another preferred embodiment, the method is for producing a 13-hydroxy-taxane, a 11-hydroxy-taxane, and/or a 10-hydroxy-taxane, said method comprising the steps of:

i. providing a host cell comprising a heterologous nucleic acid encoding CYP3 of SEQ ID NO: 7 or functional homologues thereof having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto, and

ii. incubating said host cell in presence of taxadiene,

thereby producing the 13-hydroxy-taxane, the 11-hydroxy-taxane, and/or the 10-hydroxy-taxane.

In another preferred embodiment, the method is for producing 10-hydroxy-taxadiene, 11-hydroxyl-4,12-taxadiene, and/or 13-hydroxy-taxadiene, said method comprising the steps of;

i. providing a host cell comprising a heterologous nucleic acid encoding CYP3 of SEQ ID NO: 7 or functional homologues thereof having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto, and

ii. incubating said host cell in presence of taxadiene,

thereby producing 10-hydroxy-taxadiene, 11-hydroxyl-4,12-taxadiene, and/or 13-hydroxy-taxadiene.

In a further preferred embodiment, the method is for producing a taxane comprising an oxetane, said method comprising the steps of;

i. providing a host cell comprising a heterologous nucleic acid encoding CYP1 of SEQ ID NO: 3 and CYP3 as set forth in SEQ ID NO: 7, or functional homologues of any of the aforementioned having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity to any of the aforementioned, and

ii. incubating said host cell in presence of taxadiene,

thereby producing the taxane comprising an oxetane.

Thus, in another preferred embodiment, the method is for producing 4-hydroxy-5,20-epoxy-taxane, 4-hydroxy-10-hydroxy-5,20-epoxy-taxane, 10-hydroxy-5,20-epoxy-

taxane, 11-hydroxy-5,20-epoxy-taxane, 4-hydroxy-11-hydroxy-5,20-epoxy-taxane, 4-hydroxy-12-hydroxy-5,20-epoxy-12,13-taxane, 13-hydroxy-5,20-epoxy-taxane, and/or 4-hydroxy-13-hydroxy-5,20-epoxy-taxane, said method comprising the steps of:

i. providing a host cell comprising a heterologous nucleic acid encoding CYP1 of SEQ ID NO: 3 and a heterologous nucleic acid encoding CYP3 of SEQ ID NO: 7, or functional homologues of any of the aforementioned having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto; and

ii. incubating said host cell in presence of taxadiene,

thereby producing said 4-hydroxy-10-hydroxy-5,20-epoxy-taxane, 10-hydroxy-5,20-epoxy-taxane, 11-hydroxy-5,20-epoxy-taxane, 4-hydroxy-11-hydroxy-5,20-epoxy-taxane, 4-hydroxy-12-hydroxy-5,20-epoxy-12,13-taxane, 13-hydroxy-5,20-epoxy-taxane, and/or 4-hydroxy-13-hydroxy-5,20-epoxy-taxane.

In another preferred embodiment, the method is for producing a taxane comprising an oxetane, said method comprising the steps of:

i. providing a host cell comprising a heterologous nucleic acid encoding CYP1 of SEQ ID NO: 3, and a heterologous nucleic acid encoding a taxadiene synthase (EC 4.2.3.17) capable of converting geranylgeranyl diphosphate (GGPP) into taxadiene, such as TXS of SEQ ID NO: 1 and/or TXS _3xSG_MBP-SKL of SEQ ID NO: 2,
or functional homologues of any of the aforementioned having at least 70% sequence identity, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto.

ii. incubating said host cell, preferably in the presence of geranylgeranyl diphosphate (GGPP),

thereby producing said taxane comprising an oxetane.

In another aspect, provided herein is a method of producing an *O*-acetylated taxane, such as 4-*O*-acetylated taxane, said method comprising the steps of:

i. providing a host cell comprising a heterologous nucleic acid encoding TAX19 of SEQ ID NO: 46, or functional homologues of any of the aforementioned having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity to SEQ ID NO: 46; and

ii. incubating said host cell in presence of taxadiene,

thereby producing the *O*-acetylated taxane, such as 4-*O*-acetylated taxane.

The taxadiene may be endotaxadiene. In other embodiments, the taxadiene is exotaxadiene. In preferred embodiments, the taxadiene is endotaxadiene.

Incubating

The step of incubating the host cell under conditions enabling growth may be performed by any method known to the skilled person. The term "growth" of a host cell or a multicellular organism comprising a host cell should be understood as a reference to proliferation, multiplication, differentiation and/or maintenance of viability of the subject host cell, or multicellular organism.

If the host cell is comprised in a multicellular organism, said conditions are usually conditions enabling maintenance of viability and/or growth of the multicellular organism. Thus, if the host cells are plant cells comprised in a plant, the cultivation conditions are conditions suitable for maintenance and/or growth of said plant. That could e.g. be sowing seeds or other regenerative parts of the said plant in a field or in a green house. Cultivation may further comprise watering and/or fertilising.

If the host cell is a unicellular organism, cultivation may be incubation in a medium comprising at least a carbon source and a nitrogen source at a temperature suitable for growth of said unicellular organism. The carbon source may e.g. be a carbohydrate, such as sugars or polysaccharides. The nitrogen source may for example be amino acids or polypeptides. The skilled person is well able of selecting a suitable cultivation medium based on the particular host cell.

Culture conditions

Incubating said host cell in the presence of taxadiene may be obtained in several manners. For example, taxadiene may be added to the host cell. If the host cell is a microorganism, then taxadiene may be added to the cultivation medium of said microorganism. If the host organism is a plant, then taxadiene may be added to the soil of the plant or it may be introduced into the plant by infiltration. Thus, if the heterologous nucleic acid(s) are introduced into the plant by infiltration, then taxadiene may be co-infiltrated together with the heterologous nucleic acid(s).

Thus, the method may further comprise a step of supplying said taxadiene to the host cell, such as incubating said host cell in a cultivation medium comprising taxadiene.

In some embodiments, the taxadiene is endotaxadiene. In other embodiments, the taxadiene is exotaxadiene. In preferred embodiments, the taxadiene is endotaxadiene.

It is also comprised within the disclosure, that the host cell is capable of producing taxadiene. Host cells capable of producing taxadiene is described further herein below, for example in section “Host cell”. In such embodiments incubating said host cell in the presence of taxadiene simply requires cultivating said host cell.

It is also comprised within the disclosure, that the host cell is capable of producing GGPP. Host cells capable of producing GGPP is described further herein. In such embodiments incubating said host cell in the presence of GGPP simply requires cultivating said host cell.

Taxane comprising an oxetane

As described herein the host cells of the invention are capable of producing a taxane comprising an oxetane. Said taxane comprising an oxetane may be a 5,20-epoxy-taxane, for example 4-hydroxy-5,20-epoxy-taxane or a derivative thereof.

Provided here is a method of producing 4-hydroxy-5,20-epoxy-taxane, said method comprising the steps of;

i. providing a host cell comprising a heterologous nucleic acid encoding CYP1 of SEQ ID NO: 3 or functional homologues thereof having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto, and
ii. incubating said host cell in presence of taxadiene,

thereby producing 4-hydroxy-5,20-epoxy-taxane.

In one embodiment of the invention the taxane comprising an oxetane is paclitaxel (taxol). It is comprised within the invention, that the host cells of the invention produces an intermediate compound, wherein the intermediate also is a taxane comprising an oxetane, and that paclitaxel (taxol) is produced by organic synthesis from said intermediate.

In one embodiment the taxane comprising an oxetane is a 10-hydroxy-5,20-epoxy-taxane.

In one embodiment the taxane comprising an oxetane is a 11-hydroxy-5,20-epoxy-taxane.

In one embodiment the taxane comprising an oxetane is a 13-hydroxy-5,20-epoxy-taxane.

In one embodiment the taxane comprising an oxetane is a 4-hydroxy-10-hydroxy-5,20-epoxy-taxane.

In one embodiment the taxane comprising an oxetane is a 4-hydroxy-11-hydroxy-5,20-epoxy-taxane.

In one embodiment the taxane comprising an oxetane is a 4-hydroxy-13-hydroxy-5,20-epoxy-taxane.

In one embodiment the taxane comprising an oxetane is a 4-hydroxy-12-hydroxy-5,20-epoxy-12,13-taxane.

In one embodiment the taxane comprising an oxetane is 10-deacetyl-baccatin III or baccatin III. Said 10-deacetyl-baccatin III or baccatin III may be produced directly by the host cell of the invention, or it may be produced by organic synthesis from an intermediate compound produced by the host cell.

In one embodiment the taxane comprising an oxetane comprises an acetoxy group, i.e. said taxane comprising an oxetane is *O*-acetylated. Thus, in one embodiment the taxane comprising an oxetane is 4-acetoxy-5,20-epoxy-taxane.

<u>*O*-acetylated taxane comprising an oxetane</u>

Provided herein is a method of producing an *O*-acetylated taxane comprising an oxetane, said method comprising the steps of:

i. providing a host cell comprising a heterologous nucleic acid encoding CYP1 of SEQ ID NO: 3 and a heterologous nucleic acid encoding TAX19 of SEQ ID NO: 46, or functional homologues of any of the aforementioned having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto; and
ii. incubating said host cell in presence of taxadiene;

thereby producing the *O*-acetylated taxane comprising an oxetane.

In some embodiments, the method of producing an *O*-acetylated taxane comprising an oxetane comprises the steps of:

i. providing a host cell comprising a heterologous nucleic acid encoding CYP1 of SEQ ID NO: 3, a heterologous nucleic acid encoding TAX19 of SEQ ID NO: 46, and a taxadiene synthase, such as TXS of SEQ ID NO: 1 and/or TXS _3xSG_MBP-SKL of SEQ ID NO: 2, or functional homologues of any of the aforementioned having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto; and
ii. incubating said host cell, preferably in the presence of geranylgeranyl diphosphate (GGPP);

thereby producing the *O*-acetylated taxane comprising an oxetane.

Said *O*-acetylated taxane comprising an oxetane may be an *O*-acetylated 5,20-epoxy-taxane. In preferred embodiments, said *O*-acetylated taxane comprising an oxetane is 4-acetoxy-5,20-epoxy-taxane.

In other embodiments, the method is for producing 4-acetoxy-5,20-epoxy-taxane, said method comprising the steps of:

i. providing a host cell comprising a heterologous nucleic acid encoding CYP1 of SEQ ID NO: 3, a heterologous nucleic acid encoding TAX19 of SEQ ID NO: 46, and a taxadiene synthase, such as TXS of SEQ ID NO: 1 and/or TXS _3xSG_MBP-SKL of SEQ ID NO: 2, or functional homologues of any of the aforementioned having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto; and

i. incubating said host cell, preferably in the presence of geranylgeranyl diphosphate (GGPP);

thereby producing the 4-acetoxy-5,20-epoxy-taxane.

Isolation

The method may further comprise a step of isolating the taxane comprising an oxetane, the 10-hydroxy-taxane, the 11-hydroxy-taxane and/or the 13-hydroxy-taxane.

The compounds may be isolated through any useful method known to the skilled person. For example, the taxane comprising an oxetane, 10-hydroxy-taxane, 11-hydroxy-taxane and/or 13-hydroxy-taxane may be isolated by a method comprising one or more of the following:

- extraction, e.g. solvent extraction,
- precipitation,
- chromatography, e.g. liquid chromatography (LC).

In some embodiments of the present disclosure, the step of isolating the taxane comprising an oxetane, 10-hydroxy-taxane, 11-hydroxy-taxane and/or 13-hydroxy-taxane comprises extraction, such as extraction with a solvent, for example extraction with methanol (MeOH).

In some other embodiments, the step of isolating the taxane comprising an oxetane, 10-hydroxy-taxane, 11-hydroxy-taxane and/or 13-hydroxy-taxane comprises chromatography, such as LC, for example column chromatography or preparative/semi-preparative high performance LC (HPLC).

Host cell

The present disclosure relates to host cells comprising one or more heterologous genes encoding enzymes of the biosynthetic pathway towards paclitaxel, i.e. taxoid pathway.

In one preferred embodiment of the present disclosure, the host cell is capable of producing taxane comprising an oxetane, in the presence of taxadiene, said host cell comprising a heterologous nucleic acid encoding CYP1 as set forth in SEQ ID NO: 3 or a functional homologue thereof having at least 70%, for example at least 75%, such as

at least 80%, for example at least 85%, such as at least 90%, such as at least 95%, for example at least 99% sequence identity thereto.

In another preferred embodiment, the host cell is capable of producing 10-hydroxy-taxane, 11-hydroxy-taxane and/or 13-hydroxy-taxane, in the presence of taxadiene, said host cell comprising a heterologous nucleic acid encoding CYP3 as set forth in SEQ ID NO: 7 or a functional homologue thereof having at least 70%, for example at least 75%, such as at least 80%, for example at least 85%, such as at least 90%, such as at least 95%, for example at least 99% sequence identity thereto.

In further another embodiment of the present disclosure, the host cell is capable of producing taxane comprising an oxetane and optionally 10-hydroxy-taxane, 11-hydroxy-taxane and/or 13-hydroxy-taxane, in the presence of taxadiene, said host cell comprising a heterologous nucleic acid encoding CYP1 as set forth in SEQ ID NO: 3 and a heterologous nucleic acid encoding CYP3 as set forth in SEQ ID NO: 7, or functional homologues of any of the aforementioned or having at least 70%, for example at least 75%, such as at least 80%, for example at least 85%, such as at least 90%, such as at least 95%, for example at least 99% sequence identity thereto.

In addition to the heterologous nucleic acid encoding a cytochrome P450 enzyme capable of catalysing formation of a taxane comprising oxetane and/or a 10-hydroxy-taxane, a 11-hydroxy-taxane and/or a 13-hydroxy-taxane, the host organism may also comprise one or more heterologous nucleic acids encoding one or more of the following:

i. taxadiene synthase (EC 4.2.3.17) capable of converting geranylgeranyl diphosphate (GGPP) into taxadiene, and/or
ii. cytochrome P450 reductase (CPR), such as an endogenous and/or a heterologous CPR.

It is preferred that the host cell is capable of producing taxadiene, which in general serves as the starting compound for production of taxane comprising oxetane, 10-hydroxy-taxane, 11-hydroxy-taxane and/or 13-hydroxy-taxane.

It is generally preferred that the host cell comprises a heterologous nucleic acid encoding a taxadiene synthase capable of catalysing synthesis of taxadiene. This is in particular the case, if the host cell does not produce taxadiene naturally. Thus, in one

preferred embodiment, the host cell further comprises a heterologous nucleic acid encoding taxadiene synthase as set forth in SEQ ID NO: 1 and/or SEQ ID NO: 2 or a functional homologue of any of the aforementioned having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto.

In some embodiments, the taxadiene synthase is TXS as set forth in SEQ ID NO: 1 and/or TXS _3xSG_MBP_SKL as set forth in SEQ ID NO: 2 or a functional homologue of any of the aforementioned having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity to SEQ ID NO: 1 or SEQ ID NO: 2, respectively. Said TXS and/or TXS _3xSG_MBP_SKL may be comprised within a larger polypeptide. Thus, in some embodiments the taxadiene synthase may comprise TXS (SEQ ID NO: 1) and/or TXS _3xSG_MBP_SKL (SEQ ID NO: 2), for example the taxadiene synthase is a taxadiene synthase native to *Taxus brevifolia* which comprises TXS (SEQ ID NO: 1), or a functional homologue thereof having at least 70% sequence identity thereto.

Thus, in some embodiments, the host cell is capable of producing taxane comprising an oxetane, wherein said host cell comprises a heterologous nucleic acid encoding CYP1 as set forth in SEQ ID NO: 3, and a heterologous nucleic acid encoding a taxadiene synthase capable of catalysing production of taxadiene, such as TXS as set forth in SEQ ID NO: 1 and/or TXS _3xSG_MBP-SKL as set forth in SEQ ID NO: 2, or functional homologues of any of the aforementioned having at least 70% sequence identity, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto.

In one embodiment, the CPR is a CPR from a plant, for example the CPR is a CPR from a *Taxus*, such as *Taxus cuspidata*, for example TcuCPR (NCBI GenBank accession number: AY571340.1 (4 June 2023) and/or SEQ ID NO: 4) or *Taxus baccata*,a CPR from *Nicotiana*, such as *Nicotiana benthamiana*, a CPR from *Arabidopsis*, such as *A. thaliana*, for example AtCPR (NCBI GenBank accession number: NP_194183, 4 June 2023), and/or a CPR from *Populus*, such as a hybrid poplar for example *Populus trichocarpa* × *Populus deltoids*, for example CPR2 (NCBI GenBank accession number: AAK15260.1 (4 June 2023) and/or SEQ ID NO: 6),

In another embodiment, the CPR is TcuCPR as set forth in SEQ ID NO: 4, or a functional homologue thereof having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto. In further another embodiment, the CPR is CPR2 as set forth in SEQ ID NO: 6, or a functional homologue thereof having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto. In another embodiment, the CPR is AtCPR (NCBI GenBank accession number: NP_194183, 4 June 2023), or a functional homologue thereof having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto.

In addition to the heterologous nucleic acid encoding a cytochrome P450 enzyme capable of catalysing formation of a taxane comprising oxetane, the host organism may also comprise one or more heterologous nucleic acids encoding one or more of the following:

i. taxadiene synthase (EC 4.2.3.17) capable of converting geranylgeranyl diphosphate (GGPP) into taxadiene;
ii. cytochrome P450 reductase (CPR), such as an endogenous and/or a heterologous CPR; and/or
iii. *O*-acetyltransferase capable of acetylating an hydroxy-group of a taxane comprising an oxetane.

Thus, in some embodiments, the host cell comprises:

i. a heterologous nucleic acid encoding CYP1 (SEQ ID NO: 3) or a functional homologue thereof having at least 70% sequence identity, such as at least 80%, such as at least 90%, such as at least 95%, such as at least 99% sequence identity to SEQ ID NO: 3; and
ii. a heterologous nucleic acid encoding TAX19 (SEQ ID NO: 46) or a functional homologue thereof having at least 70% sequence identity, such as at least 80%, such as at least 90%, such as at least 95%, such as at least 99% sequence identity to SEQ ID NO: 46.

In other embodiments, the host cell comprises:

i. a heterologous nucleic acid encoding a taxadiene synthase, preferably TXS (SEQ ID NO: 1) and/or TXS _3xSG_MBP-SKL (SEQ ID NO: 2), or functional

homologues of any of the TXS and/or TXS _3xSG_MBP-SKL having at least 70% sequence identity, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity SEQ ID NO: 1 and/or SEQ ID NO: 2);

ii. a heterologous nucleic acid encoding CYP1 (SEQ ID NO: 3) or a functional homologue thereof having at least 70% sequence identity, such as at least 80%, such as at least 90%, such as at least 95%, such as at least 99% sequence identity to SEQ ID NO: 3; and

iii. a heterologous nucleic acid encoding TAX19 (SEQ ID NO: 46) or a functional homologue thereof having at least 70% sequence identity, such as at least 80%, such as at least 90%, such as at least 95%, such as at least 99% sequence identity to SEQ ID NO: 46.

In some embodiments, the taxadiene is endotaxadiene. In other embodiments, the taxadiene is exotaxadiene. In preferred embodiments, the taxadiene is endotaxadiene.

Organisms

In some embodiments of the present disclosure, the host cell is selected from the group of plant cells, yeast cells, bacterial cells and fungal cells.

In some embodiments the host cell is comprised within a multicellular organism. In such embodiments, only some of the cells of said multicellular organism may comprise heterologous nucleic acid(s) and/or heterologous polypeptide(s). It is however preferred that all cells of said multicellular organism are host cells that comprise the same heterologous nucleic acids(s) and/or heterologous polypeptide(s).

In some embodiments of the present disclosure, the host cell is plant cells comprised within a plant, within a part of a plant or within the seeds of said plant. Preferably, all cells of said plant or part thereof are host cells comprising the same heterologous nucleic acid(s) and/or heterologous polypeptide(s).

In some embodiments of the present disclosure, the host cells are plant cells, e.g. plant cells from a species of *Nicotiana* such as *Nicotiana benthamiana* or *Nicotiana tabacum.* The person skilled in the art will appreciate that a "plant cell" as used within the present invention refers to a structural and physiological unit of a plant, e.g. a tobacco plant.

The plant cell may be in form of a protoplast without a cell wall, an isolated single cell or a cultured cell, or as a part of higher organized unit such as but not limited to, plant tissue, a plant organ, or a whole plant.

In some embodiments of the present disclosure, the host cell is a yeast cell, e.g. a yeast cell belonging to the genus of *Saccharomyces*, *Candida, Cryptococcus, Pichia (Komagataella), Lipomyces, Pseudozyma, Rhodosporidium, Rhodotorula, Trichosporon, Trigonopsis, Yarrowia* or *Saccharomycopsis*, such as a yeast cell of the species *Saccharomyces cerevisiae, Yarrowia lipolytica, Hansenula polymorpha (Ogataea polymorpha), Rhodotorula toruloides* or *Pichia pastoris* (*Komagataella phaffii*).

The person skilled in the art will appreciate that a fungi or fungal cell(s) as used herein refers to any cell present within or derived from an organism belonging to the Kingdom Fungi. The methods are applicable to all fungi and fungal cells that are susceptible of genetic modifications. Furthermore, the person skilled in the art will appreciate that a "yeast cell" is herein defined to include the group consisting of small, unicellular organisms capable of growth and reproduction through budding or direct division (fission), or by growth as simple irregular filaments (mycelium). The yeast cell may be transformed or transfected with a heterologous vector for expression of a nucleic acid sequence inserted into the heterologous vector. Examples of a yeast cell include, but are not limited to *Saccharomyces cerevisiae*, *Yarrowia lipolytica*, *Hansenula polymorpha (Ogataea polymorpha), Rhodotorula toruloides* and/or *Pichia pastoris* (*Komagataella phaffii*), commonly used for transfection and expression of heterologous proteins.

In some embodiments of the present disclosure, the host cell is a bacterial cell, e.g. a bacterial cell belonging to the genus of *Escherichia*, *Bacillus, Corynebacterium, Pseudomonas* or *Streptomyces*, such as a bacterial cell of the species *Escherichia coli, Bacillus subtilis, Corynebacterium glutamicum, Pseudomonas putida* or *Streptomyces sp*.

The person skilled in the art will appreciate that a bacterial cell includes prokaryotic cells that may be propagated in culture. The bacterial cell may act as a host cell for the recombinant expression of heterologous polypeptide(s). The bacterial cell may be

transformed, transfected or infected with a vector for expression of a nucleic acid sequence inserted into the vector. Examples of suitable bacterial cells include, but are not limited to *E. coli, Bacillus subtilis, Corynebacterium glutamicum, Pseudomonas putida* and/or *Streptomyces sp.*

The host cell or a progenitor thereof may be prepared by any useful method available to the skilled person. For example, the heterologous nucleic acid(s) may be inserted into a cell by direct uptake, transduction, f-mating, transfection, transformation, bacterial infiltration or any other methods known in the art useful for creating recombinant host cells.

Modifications

In addition to any one of the above heterologous polypeptides and/or nucleic acids, the host cell may further comprise additional modifications such as one or more mutations, for example mutations of the native genome of said host cell. Such modifications or mutations may be, but are not limited to, deletion, overexpression of endogenous or heterologous genes or point-mutations.

For example, the host cell may be capable of producing geranylgeranyl pyrophosphate (GGPP). Preferably, the host cell is capable of producing GGPP, either natively or due to expression of heterologous nucleic acids. It is well known in the art how to modify host cells to produce GGPP and/or to overproduce GGPP, see for example Ignea et al., 2015. In other words, a person skilled in the art is well capable of engineering a host cell to produce or overproduce GGPP. In some embodiments, the host cell is a yeast cell further comprising a heterologous nucleic acid encoding ERG20(F96C)-L as set forth in SEQ ID NO: 5, ERG20(Y95A) and/or ERG20(F96C), or a functional homologue of any of the aforementioned having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity to SEQ ID NO: 5 or ERG20(F96C), respectively. In some embodiments, the host cell is a yeast cell further comprising a nucleic acid encoding a geranylgeranyl diphosphate synthase (GGDPS), such as a GGDPS from *Cistus creticus* for example CcGGDPS1 as set forth in SEQ ID NO: 42 (NCBI GenBank accession number: AAM21638, 4 June 2023), or a functional homologue thereof having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto.

In some embodiments, the host cell is a yeast cell further comprising a nucleic acid encoding a farnesyl diphosphate synthase (FDPS), such as a FDPS from *Saccharomyces cerevisiae* for example ERG20 (NCBI GenBank accession number: NM_001181600, 3 June 2023), *Salvia fruticosa* for example SfFDPS1 as set forth in SEQ ID NO: 39, or a functional homologue thereof having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto.

In some embodiments, the host cell is a yeast cell further comprising a nucleic acid encoding a cytochrome B5 (CYB5), such as an CYB5 from *Salvia pomifera*, for example SpCytb5 as set forth in SEQ ID NO: 40, or a functional homologue thereof having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto.

In some embodiments, the host cell is a yeast cell further comprising a nucleic acid encoding a cytochrome B5 reductase (CBR), such as a CBR from *S. pomifera*, for example SpCBR as set forth in SEQ ID NO: 38, or a functional homologue thereof having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto.

In some embodiments, the host cell is a plant cell further comprising a nucleic acid encoding P19 as set forth in SEQ ID NO: 16, or functional homologue thereof having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99%sequence identity thereto. Such nucleic acid may for example be SEQ ID NO: 17, or functional homologue thereof having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99%sequence identity thereto.

In other embodiments, the host cell further comprises a nucleic acid encoding an acetyltransferase, such as an acetyltransferase from a *Taxus* cell, for example a *Taxus chinensis* cell. In preferred embodiments, said acetyltransferase is TAX19 as set forth in SEQ ID NO: 46, or a functional homologue thereof having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto. Host cells comprising a nucleic acid encoding TAX19 of

SEQ ID NO: 46 are further capable of producing an *O*-acetylated taxane comprising an oxetane, such as 4-acetoxy-5,20-epoxy-taxane.

Combination of polypeptides

In further another embodiment, the host cell capable of producing 10-hydroxy-taxane, 11-hydroxy-taxane and/or 13-hydroxy-taxane, in the presence of taxadiene, comprises one or more heterologous nucleic acids encoding CYP3 as set forth in SEQ ID NO: 7 and a heterologous nucleic acid encoding a taxadiene synthase capable of catalysing production of taxadiene, such as TXS as set forth in SEQ ID NO: 1 and/or TXS _3xSG_MBP_SKL as set forth in SEQ ID NO: 2, or functional homologues of any of the aforementioned having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto.

In yet another embodiment, the host cell comprises one or more heterologous nucleic acids encoding CYP3 as set forth in SEQ ID NO: 7 and TXS as set forth in SEQ ID NO: 1 and/or TXS _3xSG_MBP_SKL as set forth in SEQ ID NO: 2, or functional homologues of any of the aforementioned having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto.

In further another embodiment, the host cell comprises one or more heterologous nucleic acids encoding:

- CYP1 (SEQ ID NO: 3); TXS (SEQ ID NO: 1) and/or TXS _3xSG_MBP_SKL (SEQ ID NO: 2); or
- CYP1 (SEQ ID NO: 3); CYP3 (SEQ ID NO: 7); or
- CYP1 (SEQ ID NO: 3); CYP3 (SEQ ID NO: 7); TXS (SEQ ID NO: 1) and/or TXS _3xSG_MBP_SKL (SEQ ID NO: 2); or
- TcuCPR (SEQ ID NO: 4); CYP1 (SEQ ID NO: 3); TXS (SEQ ID NO: 1) and/or TXS _3xSG_MBP_SKL (SEQ ID NO: 2); or
- TcuCPR (SEQ ID NO: 4); CYP1 (SEQ ID NO: 3); CYP3 (SEQ ID NO: 7); or
- TcuCPR (SEQ ID NO: 4); CYP1 (SEQ ID NO: 3); CYP3 (SEQ ID NO: 7); TXS (SEQ ID NO: 1) and/or TXS _3xSG_MBP_SKL (SEQ ID NO: 2); or
- TcuCPR (SEQ ID NO: 4); CYP1 (SEQ ID NO: 3); or
- TcuCPR (SEQ ID NO: 4); CYP3 (SEQ ID NO: 7); or
- TcuCPR (SEQ ID NO: 4); CYP1 (SEQ ID NO: 3); CYP3 (SEQ ID NO: 7);

or functional homologues of any of the aforementioned having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto.

In some embodiments, the host cell comprises one or more nucleic acid constructs and/or vectors as described herein, for example in the section “Polypeptides and nucleic acids”.

Polypeptides and nucleic acids

In addition to the methods and host cells, the invention also provides enzymes useful in the production of a taxane comprising an oxetane, an *O*-acetylated taxane comprising an oxetane, a 10-hydroxy-taxane, a 11-hydroxy-taxane and/or a 13-hydroxy-taxane.

In some embodiments of the present disclosure, the invention provides a polypeptide, such as an isolated polypeptide selected from the group of polypeptides of SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 7, SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 16, SEQ ID NO: 20, SEQ ID NO: 6, SEQ ID NO: 42, SEQ ID NO: 39, SEQ ID NO: 40, SEQ ID NO: 38, and functional homologues of any of the aforementioned having at least 70% sequence identity, preferably at least 80%, preferably at least 85% sequence identity, preferably at least 90% sequence identity, preferably at least 95% sequence identity, more preferred at least 98% sequence identity thereto.

The invention further provides one or more nucleic acid constructs encoding one or more of any of the polypeptides of SEQ ID NO: 3, SEQ ID NO: 7, SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 16, SEQ ID NO: 4, SEQ ID NO: 6 SEQ ID NO: 20, SEQ ID NO: 42, SEQ ID NO: 39, SEQ ID NO: 40, SEQ ID NO: 38, and functional homologues of any of the aforementioned having at least 70% sequence identity, preferably at least 80%, preferably at least 85% sequence identity, preferably at least 90% sequence identity, preferably at least 95% sequence identity, more preferred at least 98% sequence identity thereto.

The skilled person will appreciate that many different nucleic acid sequences may encode the same polypeptide. In some embodiment, the nucleic acid construct may comprise or consist of a coding sequence (e.g. a cDNA sequence) from *Taxus*, e.g. a cDNA derived from a wild type *Taxus baccata*, *Taxus cuspidata, Taxus chinensis,*

and/or *Taxus brevifolia*. In other embodiments, the nucleic acid sequence may be codon optimised for improved expression in the host cell.

In one embodiment the nucleic acid construct comprises or consists of one or more of any one of SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 15, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 17, SEQ ID NO: 36, SEQ ID NO: 37, SEQ ID NO: 43, SEQ ID NO: 44, SEQ ID NO: 41, or functional homologues thereof encoding the same polypeptide or a polypeptide having at least 70% sequence identity, preferably at least 80%, preferably at least 85% sequence identity, preferably at least 90% sequence identity, preferably at least 95% sequence identity, more preferred at least 98% sequence identity thereto or codon optimised versions of any of the aforementioned.

Thus, provided herein is a nucleic acid construct for expression in a host cell, comprising:

i. a nucleic acid encoding CYP1 as set forth in SEQ ID NO: 3 and/or a functional variant thereof having at least 70% sequence identity thereto, such as SEQ ID NO: 10 and/or SEQ ID NO: 11; and/or
ii. a nucleic acid encoding CYP3 as set forth in SEQ ID NO: 7, such as SEQ ID NO: 15; and optionally
iii. a nucleic acid encoding TXS as set forth in SEQ ID NO: 1 and/or TXS _3xSG_MBP_SKL as set forth in SEQ ID NO: 2, such as SEQ ID NO: 8 and/or SEQ ID NO: 9, respectively;

and further optionally,

i. a nucleic acid encoding ERG20(F96C)-L as set forth in SEQ ID NO: 5, such as SEQ ID NO: 13, CcGGDPS1 as set forth in SEQ ID NO: 42, such as SEQ ID NO: 43, ERG20(F96C), and/or ERG20(Y95A),
ii. a nucleic acid encoding TcuCPR as set forth in SEQ ID NO: 4, such as SEQ ID NO: 12, CPR2 as set forth in SEQ ID NO: 6, such as SEQ ID NO: 44, and/or AtCPR (NCBI GenBank accession number: NP_194183, 4 June 2023),
iii. a nucleic acid encoding *S. cerevisiae* ERG20 (NCBI GenBank accession number: NM_001181600, 3 June 2023), and/or SfFDPS1 as set forth in SEQ ID NO: 39, such as SEQ ID NO: 14,
iv. a nucleic acid encoding SpCytb5 as set forth in SEQ ID NO: 40, such as SEQ ID NO: 36,

v. a nucleic acid encoding SpCBR as set forth in SEQ ID NO: 38, such as SEQ ID NO: 37,

iv. a nucleic acid encoding HMG2(K6R) as set forth in SEQ ID NO: 20, such as SEQ ID NO: 41,

or functional homologues of any of the aforementioned having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto.

Provided herein is a nucleic acid construct for expression in a host cell, comprising:

i. a nucleic acid encoding CYP1 as set forth in SEQ ID NO: 3 and/or a functional variant thereof having at least 70% sequence identity thereto, such as SEQ ID NO: 10 and/or SEQ ID NO: 11, and a nucleic acid encoding CYP3 as set forth in SEQ ID NO: 7, such as SEQ ID NO: 15; or

ii. a nucleic acid encoding CYP1 as set forth in SEQ ID NO: 3 and/or a functional variant thereof having at least 70% sequence identity thereto, such as SEQ ID NO: 10 and/or SEQ ID NO: 11, and/or
a nucleic acid encoding CYP3 as set forth in SEQ ID NO: 7, such as SEQ ID NO: 15; and one or more of the following:

a. a nucleic acid encoding TXS as set forth in SEQ ID NO: 1 and/or TXS _3xSG_MBP_SKL as set forth in SEQ ID NO: 2, such as SEQ ID NO: 8 and/or SEQ ID NO: 9, respectively,

b. a nucleic acid encoding ERG20(F96C)-L as set forth in SEQ ID NO: 5, such as SEQ ID NO: 13, CcGGDPS1 as set forth in SEQ ID NO: 42, such as SEQ ID NO: 43, ERG20(F96C), and/or ERG20(Y95A),

c. a nucleic acid encoding TcuCPR as set forth in SEQ ID NO: 4, such as SEQ ID NO: 12, CPR2 as set forth in SEQ ID NO: 6, such as SEQ ID NO: 44, and/or AtCPR (NCBI GenBank accession number: NP_194183, 4 June 2023),

d. a nucleic acid encoding *S. cerevisiae* ERG20 (NCBI GenBank accession number: NM_001181600, 3 June 2023), and/or SfFDPS1 as set forth in SEQ ID NO: 39, such as SEQ ID NO: 14,

e. a nucleic acid encoding SpCytb5 as set forth in SEQ ID NO: 40, such as SEQ ID NO: 36,

f. a nucleic acid encoding SpCBR as set forth in SEQ ID NO: 38, such as SEQ ID NO: 37,

g. a nucleic acid encoding HMG2(K6R) as set forth in SEQ ID NO: 20, such as SEQ ID NO: 41,

or functional homologues of any of the aforementioned having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto.

In some embodiments, the nucleic acid construct for expression in a host cell further comprises a heterologous nucleic acid encoding TAX19 as set forth in SEQ ID NO: 46 or a functional homologue thereof having at least 70% sequence identity thereto, such as SEQ ID NO: 48 or a functional having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto.

For example, in some embodiments, the nucleic acid construct comprises:

i. a nucleic acid encoding CYP1 as set forth in SEQ ID NO: 3, such as SEQ ID NO: 10 and/or SEQ ID NO: 11; and
ii. a nucleic acid encoding TXS as set forth in SEQ ID NO: 1 and/or TXS _3xSG_MBP_SKL as set forth in SEQ ID NO: 2, such as SEQ ID NO: 8 and/or SEQ ID NO: 9, respectively,

or functional homologues of any of the aforementioned having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto.

For example, in some embodiments, the nucleic acid construct comprises:

i. a nucleic acid encoding CYP1 as set forth in SEQ ID NO: 3, such as SEQ ID NO: 10 and/or SEQ ID NO: 11;
ii. a nucleic acid encoding TAX19 as set forth in SEQ ID NO: 46, such as SEQ ID NO: 48; and
iii. a nucleic acid encoding TXS as set forth in SEQ ID NO: 1 and/or TXS _3xSG_MBP_SKL as set forth in SEQ ID NO: 2, such as SEQ ID NO: 8 and/or SEQ ID NO: 9, respectively,

or functional homologues of any of the aforementioned having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto.

In other embodiments, the nucleic acid construct comprises:

i. a nucleic acid encoding CYP3 as set forth in SEQ ID NO: 7, such as SEQ ID NO: 15; and
ii. a nucleic acid encoding TXS as set forth in SEQ ID NO: 1 and/or TXS _3xSG_MBP_SKL as set forth in SEQ ID NO: 2, such as SEQ ID NO: 8 and/or SEQ ID NO: 9, respectively,

or functional homologues of any of the aforementioned having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto.

In further embodiments, the nucleic acid construct may further comprise a promotor capable of inducing expression of the heterologous nucleic acid in the host cell. The skilled person will be able to select suitable promoters for a given host cell. In some embodiment, the promoter is an inducible promoter, operably linked to any one or more of the nucleic acid sequences provided herein.

Provided herein is also a vector comprising at least one of the nucleic acid constructs provided herein.

Further provided herein is a kit of parts comprising:

i. the host cell as described herein, and optionally instructions for use, and/or
ii. the nucleic acid construct as described herein or a vector as described herein, and optionally instructions for use and/or a host cell to be modified,

preferably wherein the host cell is a microorganism, for example a yeast cell, or a plant cell.

Compounds, compositions and pharmaceutical use

The methods and host cells described herein can be used to produce different plant diterpenoid compounds at high titer. In particular, the present methods and host cell are useful for producing a taxane comprising an oxetane, an *O*-acetylated taxane comprising an oxetane, a 10-hydroxy-taxane, a 11-hydroxy-taxane and/or said a 13-hydroxy-taxane.

Compounds

Thus, provided herein is 4-hydroxy-5,20-epoxy-taxane obtained by a method described herein. Provided is also, 10-hydroxy-taxadiene obtained by a method described herein. Provided herein is also 11-hydroxyl-4,12-taxadiene obtained by a method described herein. Also provided herein is 13-hydroxy-taxadiene obtained by a method described herein. Provided is also 4-acetoxy-5,20-epoxy-taxane obtained by a method described herein.

Provided herein is a purified compound of the formula 4-acetoxy-5,20-epoxy-taxane.

Cell culture, fermentation liquid and titers

Provided herein is also a cell culture obtained by a method described herein. Provided is also a cell culture comprising a host cell described herein, and optionally a cultivation medium.

Further provided is a fermentation liquid comprising the taxane comprising an oxetane, the *O*-acetylated taxane comprising an oxetane, the 10-hydroxy-taxane, the 11-hydroxy-taxane and/or the 13-hydroxy-taxane.

In one embodiment, the fermentation liquid is obtained by a method described herein. In other embodiments, the method may further comprise a step of obtaining a fermentation liquid, wherein said fermentation liquid optionally comprise the host cell. In another embodiment, the fermentation liquid is comprised in a cell culture described herein. In other embodiments, the taxane comprising an oxetane, the *O*-acetylated taxane comprising an oxetane, the 10-hydroxy-taxane, the 11-hydroxy-taxane and/or the 13-hydroxy-taxane is comprised within a host cell described herein. In further embodiments, the taxane comprising an oxetane, the *O*-acetylated taxane comprising an oxetane, the 10-hydroxy-taxane, the 11-hydroxy-taxane and/or the 13-hydroxy-taxane is secreted by a host cell described herein, for example to the fermentation liquid, cultivation medium or broth. In some embodiments, at least 50% of the host cells are lysed, such as at least 75%, such as at least 80%, for example at least 90%, such as at least 95%, such as at least 99% of the host cells are lysed. In other embodiments, at least 50% of solid cellular material has been separated from the liquid, such as at least 75%, such as at least 80%, for example at least 90%, such as at

least 95%, such as at least 99% of solid cellular material has been separated from the liquid.

As described above, the produced taxane comprising an oxetane, *O*-acetylated taxane comprising an oxetane, 10-hydroxy-taxane, 11-hydroxy-taxane and/or 13-hydroxy-taxane may be secreted by the host cell, and thus be present in the extracellular fraction (supernatant), or it may be retained in the host cell, and thus be present in the intracellular fraction. The total titer of a compound is the sum of the intracellular titer and extracellular titer of the compound.

In some embodiments, the host cell may be capable of producing taxane comprising an oxetane, *O*-acetylated taxane comprising an oxetane, 10-hydroxy-taxane, 11-hydroxy-taxane and/or 13-hydroxy-taxane is produced with a titer of at least 50 µg/L, such as at least 0.75 µg/L, for example at least 100 µg/L, such as at least 250 µg/L, for example 500 µg/L, such as at least 750 µg/L, for example at least 900 µg/L, such as at least 1000 µg/L, for example at least 2.5 mg/L, such as at least 5 mg/L, for example at least 7.5 mg/L, for example at least 10 mg/L, or more.

Methods for determining titers of the diterpenoid compounds are known in the art. For example, the titers may determined by UPLC-HRMS, as in the Examples of the present disclosure.

Compositions

The plant diterpenoid compounds, and more particular the taxanes, obtainable by the present methods may be useful for obtaining compositions comprising any of the compounds produced by the host cell of the present disclosure.

Thus, provided herein is a composition comprising one or more of a taxane comprising an oxetane, an *O*-acetylated taxane comprising an oxetane, a 10-hydroxy-taxane, a 11-hydroxy-taxane, and a 13-hydroxy-taxane obtained by any method described herein, and optionally further comprising one or more agents, additives and/or excipients.

Provided is also a composition comprising the fermentation liquid described herein above.

In some embodiments, said composition, the composition, the fermentation liquid and/or taxane comprising an oxetane, *O*-acetylated taxane comprising an oxetane, 10-hydroxy-taxane, 11-hydroxy-taxane and/or 13-hydroxy-taxane may be processed into in a semi-dry or dry solid form, optionally in form of a powder, tablet, capsule, chewable, gel and/or gum. In further some embodiments, the composition, the fermentation liquid and/or taxane comprising an oxetane, *O*-acetylated taxane comprising an oxetane, 10-hydroxy-taxane, 11-hydroxy-taxane and/or 13-hydroxy-taxane may be in a liquid form, optionally in a stabilized liquid form.

Pharmaceutical use

The plant diterpenoid compounds, and more particular the taxanes, obtainable by the present methods may be useful for manufacturing pharmaceutical compounds, in particular taxoid compounds, such as paclitaxel, 10-deacetylbaccatin III and/or baccatin III. Thus, the methods may further comprise a step of producing a pharmaceutical compound and/or composition from any of the compounds produced by the host cell of the present disclosure.

Provided is also a method of treating a disorder such as cancer, comprising administration of a therapeutic sufficient amount of a taxane, in particular a taxane comprising an oxetane, *O*-acetylated taxane comprising an oxetane, a 10-hydroxy-taxane, a 11-hydroxy-taxane and/or a 13-hydroxy-taxane, a pharmaceutical compound and/or composition comprising said taxane and/or compound, obtained by a obtained by the methods described herein. Non-limited examples of such cancers are melanoma cancer, ovarian cancer, breast cancer, bladder cancer, prostate cancer and esophageal cancer.

Examples

Example 1: Materials and Methods for Examples 2, 3 and 4

This example comprises the materials and methods for examples 2-4 herein below.

RNA isolation, cDNA synthesis and gene cloning

Total RNA was extracted from -80 °C preserved *Taxus baccata* and *Taxus cuspidata* needles with an RNA extraction Kit (Spectrum Plant Total RNA Kit; Sigma Aldrich; product # STRN50). The RNA was reverse-transcribed into cDNA using SuperScript® III

First-Strand Synthesis System (ThermoFisher; Catalog No. 18080051). This cDNA was used as template to PCR amplify genes of interest wherein transcript-specific primers of candidate genes were used. The pJET1.2 cloning kit (Thermo Fisher Scientific Inc.; Cat# K1231) was used to clone the blunt-end PCR products. The resulting constructs were confirmed by sequencing and used for Uracil-Specific-Excision-Reaction (USER) cloning.

Constructs for both *N. benthamiana* (tobacco) transient expression and yeast expression were produced using USER cloning with specific primers (USER-GeneName-FP and USER-GeneName-RP, Table 1). All of the primers were ordered from TAG Copenhagen, Denmark. pLIFE33 vector (Forman et al., 2022) was used in tobacco transient expression.

Table 1. Primers used for construction of constructs for gene expression in tobacco and yeast.

Primer	**SEQ ID NO:**
US_TXS_F	18
US_TXS_R_3SG	19
US_MBPv_3SG_F	21
US_MBPv_R	22
US_MBPv_SKL_R	23
US_TcuCYP1_F	24
US_TcuCYP1_R	25
US_TbaCYP1_F	26
US_TbaCYP1_R	27
US_TcuCYP3_F	28
US_TcuCYP3_R	29
US_TcCPR_F	30
US_TcCPR_R	31
US_TcCPR_F10	32
US_TcCPR_RC	33
US_ScErg20_F	34
US_ScErg20_SKL_R	35

Transient co-expression of genes of interest in *N.benthamiana* leaves

Constructs for tobacco transient expression were electro-transformed into agrobacteria (*A. tumefaciens* strain AGL-1- GV3850). The overnight cultures of engineered agrobacteria strains were used for agro-infiltration. Briefly, OD_{600} of 1 culture mix was used with equal ratio to every agrobacteria strains. For infiltration, for to six weeks old tobacco leaves were used. After infiltration, the plants were kept in the greenhouse (16 h light at 20ºC, 8 h dark at 19 ºC) for 7 days before subjected to metabolites analysis.

Heterologous expression of candidate genes in yeast

The *S. cerevisiae* strain AM158 that has been engineered for boosting diterpenoid production was used as the parent strain for the gene expression (Table 2). Yeast transformation was carried out using a lithium-acetate protocol. For yeast chromosomal integration, cassettes containing genes of interest, with URA selective marker, were released from plasmids by NotI (New England Biolabs, USA) digestion before transformation. To prepare for detection of produced compounds, the yeast strains were cultivated overnight in selective glucose media at 30°C and 150 rpm and used as seed cultures. To induce the triterpenoid production, each seed culture was washed three times with sterilized MQ water and transferred into a 100 mL glass flask containing 10 mL galactose/raffinose media. The yeast cultures were incubated (20°C and 150 rpm) for 3 days before extraction.

Yeast media**:** Yeast glucose media: 2% (w/v) glucose, 0.13% (w/v) Yeast Synthetic Dropout Medium Supplements (all essential amino acids), 0.67% (w/v) Yeast Nitrogen Base w/o AA. Yeast galactose/raffinose medium: 2% (w/v) galactose, 1% (w/v) raffinose, 0.13% (w/v) Yeast Synthetic Dropout Medium Supplements (all essential amino acids), 0.67% (w/v) Yeast Nitrogen Base w/o AA (Y2025, US Biologicals). Yeast pH buffered media:yeast galactose/raffinose media containing 10% phosphate buffer solution (1M) and the pH adjusted to 7.

Table 2. Yeast strains. Abbreviations: MBP; maltose-binding protein, 3xSG: SGSGSG-linker. SEQ ID NOs of nucleic acids and/or polypeptides comprised in the yeasts are indicated as well.

S. cerevisiae	**Characteristics**	**SEQ ID NOs of heterologous polypeptides**	**SEQ ID NOs of**

			heterologous nucleic acids
AM158	Mat a/alpha, *HO1*::P*GAL1-HMG2(K6R)/HO1*::P*GAL1-HMG2(K6R)*, *ura3,his3, trp1*, *leu2*::P*TDH3-HMG2(K6R)/leu2*::P*TDH3-HMG2(K6R)*, *HO1*::P*TDH3-HMG2(K6R)*, *ERG9::A02B05*/erg9, *ubc7/UBC7*, *ssm4/SSM4*, *mct1/MCT1*, *whi2/WHI2*, *gdh1/GDH1*, *erg3/ERG3*, *rox1/rox1*, *FLO8*::P*TDH3-CcGGDPS1*, *FLO1*::P*TDH3-SfFDPS1*, *FLO5*::P*TDH3-HEM3*, *YAL065*::P*TDH3-SpCytb5*, *PDC1*::P*TDH3-ERG20(F96C)*, *dos2*::P*TDH3-CPR2/dos2*, *YOR292c*::P*GAL1-SpCBR*	SEQ ID NO: 20, SEQ ID NO: 6, SEQ ID NO: 42, SEQ ID NO: 39, SEQ ID NO: 40, SEQ ID NO: 38	SEQ ID NO: 36, SEQ ID NO: 37, SEQ ID NO: 14, SEQ ID NO: 43, SEQ ID NO: 44, SEQ ID NO: 41
TA003	AM158, XI-2::[P*TEF1_ERG20(F96C)-L*, P*CCW12_TXS*_3xSG_*MBP-SKL,URA3*]	SEQ ID NO: 5, SEQ ID NO: 2	SEQ ID NO: 9, SEQ ID NO: 13
TA004	AM158, XI-2::[P*TEF1_ERG20(F96C)-L*, P*CCW12_TXS*_3xSG_*MBP-SKL, URA3*] , pESC-H_P*GAL1-10*_TcuCYP1_*Tcu*CPR	SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5	SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13

Sample preparation and UPLC-HRMS analysis

For tobacco metabolites extraction, two leaf discs (Ø3 cm, 1 disc per leaf) were used for each sample and ground in liquid nitrogen before extracting with 1 mL methanol (MeOH). For yeast samples, 5 mL ethyl acetate (EtOAc) was added to each induced yeast cultures (10 mL cultures in 100 mL flasks). The mixtures were centrifuged at 4000 RPM

for 10 min. Afterwards, 1 mL EtOAc supernatant was dried by blowing nitrogen and the dried metabolites were re-dissolved in 100 μL MeOH. All the samples were filtered with 0.22 μM filter before Ultra-Performance-Liquid-Chromatography-High-Resolution-Mass-Spectrometry (UPLC-HRMS) analysis.

UPLC-HRMS analysis was performed on the 731 Dionex UltiMate® 3000 Quaternary Rapid Separation UHPLC focused system (Thermo Fisher Scientific, 732 Germering, Germany) equipped with a Phenomenex Kinetex XB-C18 column (100 mm × 2.1 mm i.d., 1.7 μm 733 particle size, 100 Å pore size) (Phenomenex, Inc., Torrance, CA, USA). The column was operated at 40 °C, and the flow rate was maintained at 0.3 mL min^{-1}. The mobile phases were water (A) and 100% acetonitrile (B), both acidified with 0.05% formic acid. Separations were performed using the following gradient profile: 0 min, 20% B; 11 min, 80% B; 21 min, 90% B; 22 min, 100% B; 27min, 100% B; 28 min, 20% B. The column outlet was connected to a Bruker Daltonics Compact QqTOF mass spectrometer equipped with an electrospray ionization (ESI) interface (Bruker Daltonics, Bremen, Germany). Mass spectra were acquired in positive ion mode, using a capillary voltage of 4000V, an end plate offset of –500V, a drying temperature of 220 °C, a nebulizer pressure of 2.0 bar, and a drying gas flow of 8 L min^{-1}. Sodium formate solution (internal standard) was injected at the beginning of each chromatographic run, and the UPLC–HRMS raw data was calibrated against these sodium clusters using the Data Analysis 4.3 (Bruker Daltonics) software program.

Extraction, purification and identification of compounds 1-4

Compound **1** was purified from tobacco leaves infiltrated with agrobacterium carrying nucleic acids encoding P19 (SEQ ID NO: 16), TXS (SEQ ID NO: 1), CYP1 (SEQ ID NO: 3) and *Tcu*CPR (SEQ ID NO: 4). Compounds 2-4 were purified from tobacco leaves infiltrated with agrobacterium carrying P19 (SEQ ID NO: 16), TXS (SEQ ID NO: 1), CYP3 (SEQ ID NO: 7) and *Tcu*CPR (SEQ ID NO: 4). Tobacco leaves were extracted with MeOH. For purifying the individual compounds, tobacco extracts were pre-purified by silica gel column and eluted with hexane-EtOAc (from 100:1 to 50:50, v/v). Each resulting fraction were analyzed by UPLC-HRMS, and the fractions containing the target compounds were combined and further purified by repeated Semi-preparative HPLC separations.

The isolation of individual compounds were performed with a Shimadzu Prominence LC-20A system, consisting of a SIL-10AP autosampler, a LC-20AT quaternary pump, a CTO-10ASvp thermostatted column compartment, a SPD-M20A diode array detector detector, and a FRC-10A fraction collector. Consecutive injections of crude extract (0.2 mL per injection, 50 mg/mL in MeOH) were separated at a flow rate of 2 mL/min using the above-mentioned solvents with a Phenomenex Luna C18 (2) column (250 × 21.2 mm, 5 µm, 100 Å; Phenomenex, Torrance, CA, USA).

NMR analysis

NMR experiments were performed on a 600 MHz Bruker Avance III instrument (operating frequency of 600.13 MHz) equipped with a cryogenically cooled 1.7-mm TCI probe head and a Bruker SampleJet sample changer (Bruker Biospin, Karlsruhe, Germany). All experiments were acquired in automation (temperature equilibration to 300 K, optimization of lock parameters, gradient shimming, and setting of receiver gain). ^{1}H-NMR spectra were acquired with 30°-pulses and 64k data points. 2D homo- and heteronuclear experiments were acquired with 2048 data points in the direct dimension and 128 (HMBC) or 256 (multiplicity edited HSQC and NOESY) data points in the indirect dimension. IconNMR ver. 4.2 (Bruker Biospin, Karlsruhe, Germany) was used for controlling automated sample change and acquisition of NMR data, whereas Topspin ver. 4.0 (Bruker Biospin, Karlsruhe, Germany) was used for acquisition and processing of NMR data.

Example 2: Production of the oxetane ring of taxol in tobacco using CYP1

The present example demonstrates that the oxetane ring of taxol can be synthesized in tobacco by using CYP1.

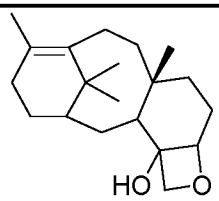
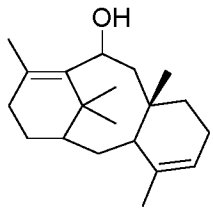
See Example 1 herein above for Material and Methods.

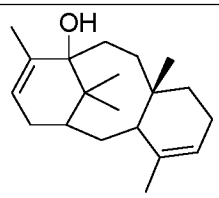
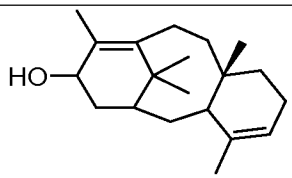
Agrobacterium-mediated transient gene expression in tobacco (*Nicotiana benthamiana*) is an efficient and reliable method for producing terpenoid compounds. To produce taxadiene derived compounds that contain the oxetane ring of taxol, we infiltrated tobacco leaves with agrobacteria carrying nucleic acids encoding CYP1 (SEQ ID NO: 3), together with the previously identified enzymes, taxadiene synthase (TXS, SEQ ID NO: 1), *Tcu*CPR (SEQ ID NO: 4), and P19 (SEQ ID NO: 16 encoded by SEQ ID NO: 17), was used to suppress gene silencing). Prior to the co-expression, the nucleic acid

sequence encoding CYP1 as set forth in SEQ ID NO: 3 from *Taxus cuspidata* was codon optimized to *S. cerevisiae* codon usage (TcuCYP1, SEQ ID NO: 10) and synthesized from Thermo Fisher. Meanwhile, cDNA of *Taxus baccata* CYP1 (TbaCYP1, SEQ ID NO: 11) was synthesised from RNA extracts of *Taxus baccata* needles followed by cloning as described in Example 1. The product profile obtained using the two nucleic acid sequences encoding CYP1 (SEQ ID NO: 3) and/or a functional variant or functional homologue thereof were compared in tobacco and was found to be similar.

UPLC-HRMS analysis of the MeOH extract of the resulted tobacco leaves reveals the production of a new compound **1** (**Figure 1**), exhibiting a parental $[M+H]^+$ ions of *m/z* 305.2463 (*calcd.* For 305.2475, $C_{20}H_{33}O_2^+$, ΔM 3.9 ppm). To obtain sufficient amount of compound **1** for structural elucidation, we extracted this compound from 120 infiltrated tobacco plants and purified it by silica gel column chromatography and HPLC separations. By comprehensive NMR analysis, compound **1** was identified as 4-hydroxy-5,20-epoxy-taxane (and named here as taxologenic oxetane), containing the characteristic oxetane ring of taxol (Table 3).

Table 3: Overview of produced diterpenoid compounds, compound **1-4**, in Examples 2-4

Peak No./ Compound	**HRMS** (*m/z*, calcd., ΔM)	**Molecular formula**	**Structure**
1	305.2463 $[M+H]^+$ (*calcd.* for 305.2475, $C_{20}H_{33}O_2^+$, ΔM 3.9 ppm)	$C_{20}H_{32}O_2$	 taxologenic oxetane
2	271.2424$[M-H_2O+H]^+$ (*calcd.* for 271.2420, $C_{20}H_{31}^+$, ΔM -1.2 ppm)	$C_{20}H_{32}O$	 10-hydroxy-taxadiene

3	271.2421 $[M\text{-}H_2O+H]^+$ (*calcd.* for 271.2420, $C_{20}H_{31}^+$, ΔM -0.2 ppm)	$C_{20}H_{32}O$	 11-hydroxyl-4,12-taxadiene
4	271.2421 $[M\text{-}H_2O+H]^+$ (*calcd.* for 271.2420, $C_{20}H_{31}^+$, ΔM -0.2 ppm)	$C_{20}H_{32}O$	 13-hydroxy-taxadiene

These results demonstrates that CYP1 (SEQ ID NO: 3) expression together with expression of TXS (SEQ ID NO: 1) and a CPR, here *Tcu*CPR (SEQ ID NO: 4), is sufficient to synthesize the oxetane ring of taxol in *N. benthamiana* and thereby produce taxologenic oxetane.

Example 3: CYP1 produces the oxetane ring of taxol in yeast

The present example demonstrates that the oxetane ring of taxol can be synthesized in yeast (*Saccharomyces cerevisiae*).

See Example 1 herein above for Material and Methods.

To investigate whether CYP1 (SEQ ID NO: 3) can catalyze the synthesis of the oxetane ring on the taxadiene scaffold in another heterologous bio-production system, we introduced CYP1 in the *S. cerevisiae* strain TA003 that is engineered for producing taxadiene to obtain strain TA004 (Table 2). Prior to the co-expression, CYP1 from *Taxus cuspidata* (SEQ ID NO: 10) was codon-optimized to *S. cerevisiae* codon usage and synthesized from Thermo Fisher.

UPLC-HRMS analysis showed that *S. cerevisiae* TA004 cells produce compound **1** in different yeast media but TA003 cells, which lack CYP1 (SEQ ID NO: 3), do not produce compound 1 (**Figure 2**).

These results demonstrates that CYP1 (SEQ ID NO: 3) expression together with expression of TXS_3xSG_MBP-SKL (SEQ ID NO: 2) and a CPR, here *Tcu*CPR (SEQ ID NO: 4), is sufficient to synthesize the oxetane ring of taxol in yeast, here *S. cerevisiae*, and thereby produce taxologenic oxetane.

Example 4: CYP3 oxidizes C-10, C-11 and C-13 positions of taxadiene

The present example demonstrates that the oxidation of C-10, C-11 and C-13 position of taxadiene can be catalyzed by CYP3 (SEQ ID NO: 7).

See Example 1 herein above for Material and Methods.

To synthesize C-10, C-11 and C-13 oxidized taxadiene, we infiltrated tobacco leaves with agrobacteria carrying nucleic acids encoding CYP3 (SEQ ID NO: 7), together with TXS (SEQ ID NO: 1), *Tcu*CPR (SEQ ID NO: 4) and P19 (SEQ ID NO: 16 encoded by SEQ ID NO: 17). Prior to the co-expression, cDNA of CYP3 was synthesised from RNA extracts of *Taxus cuspidata* needles (TcuCYP3, SEQ ID NO: 15) followed by cloning as described in Example 1.

UPLC-HRMS analysis reveals the production of three new compounds **2-4** in the tobacco leaves (**Figure 3**). Compounds **2-4** were extracted and isolated from 120 infiltrated tobacco plants, and subsequently identified as 10-hydroxy-taxadiene (**2**), 11-hydroxyl-4,12-taxadiene (**3**), 13-hydroxy-taxadiene (**4**) by NMR (Table 3), harbouring oxidation at C-10, C-11 and C-13 position of taxadiene, respectively.

Conclusion: The invention shows that CYP3 (SEQ ID NO: 7) expression together with expression of TXS (SEQ ID NO: 1), a CPR, here *Tcu*CPR (SEQ ID NO: 4), and P19 (SEQ ID NO: 16) is sufficient to synthesize the hydroxylated taxadiene derivatives 10-hydroxy-taxadiene (**2**), 11-hydroxyl-4,12-taxadiene (**3**), 13-hydroxy-taxadiene (**4**).

Example 5: The CYP1-produced taxologenic oxetane is further acetylated by TAX19

The present example demonstrates that the key intermediate 4-hydroxy-5,20-epoxy-taxane (taxologenic oxetane) can be further acetylated by TAX19 (SEQ ID NO: 46) in tobacco (*N. benthamiana*), in what appears to be the next step in taxol biosynthesis.

See Example 1 herein above for Material and Methods.

In the structure of taxol, the C4 hydroxyl group of 4-hydroxy-5,20-epoxy-taxane (taxologenic oxetane) is acetylated. To investigate if taxologenic oxetane can be acetylated by downstream enzymes en route to taxol, we studied the ability of two acetyltransferases that have been reported to acetylate taxadiene-5a-ol, TAT (SEQ ID NO: 45) and TAX19 (SEQ ID NO: 46).

We infiltrated tobacco leaves with agrobacteria carrying nucleic acids encoding CYP1 (SEQ ID NO: 3), taxadiene synthase (TXS, SEQ ID NO: 1), *Tcu*CPR (SEQ ID NO: 4), P19 (SEQ ID NO: 16 encoded by SEQ ID NO: 17), and TAT (SEQ ID NO: 45 encoded by SEQ ID NO: 47) or TAX19 (SEQ ID NO: 46 encoded by SEQ ID NO: 48). Prior to the co-expression, the nucleic acid sequence encoding TAT that was codon optimized to *S. cerevisiae* codon usage and TAX19 were synthesized from Twist Biosciences Inc.

UPLC-HRMS analysis of the MeOH extract of the resulted tobacco leaves revealed that when TAX19 was present, 4-hydroxy-5,20-epoxy-taxane (taxologenic oxetane, compound **1**) was entirely consumed (**Figure 4a**), suggesting that it is a substrate for TAX19 which catalyzes acetylation at its C4 hydroxyl group. On the contrary, when TAT was present, taxologenic oxetane was not consumed, suggesting that it is not a favourable substrate for TAT.

We observed three new peaks (a, b, and c) when TAX19 was present (**Figure 4b**), exhibiting similar parental $[M+H]^+$ ions of *m/z* 347.2580, 347.2571 and 347.2580 respectively (all calculated for 347.2580, $C_{22}H_{35}O_3^+$, which is the expected ion for 4-acetoxy-5,20-epoxy-taxane (4-acetyl-taxologenic oxetane) (**Figure 4c**), ΔM 0.2 ppm (for peak **a**); 2.7 ppm (for peak **b**); 0.2 ppm (for peak **c**)). It is unclear at this point which of the three peaks corresponds to 4-acetoxy-5,20-epoxy-taxane. The predicted and calculated *m/z* for each of the peaks is shown in Table 4, while the HRMS data of the corresponding compounds are shown in **Figure 5**.

Table 4. The predicted and calculated *m/z* for peaks **a**-**c** corresponding to 4-acetoxy-5,20-epoxy-taxane.

Peak No./Compound	HRMS (*m/z*, calcd., ΔM)	Molecular formula

a	347.2580 $[M+H]^+$ (*calcd.* for 347.2580, $C_{20}H_{35}O_3^+$, ΔM 0.2 ppm)	$C_{22}H_{34}O_3$
b	347.2571 $[M+H]^+$ (*calcd.* for 347.2580, $C_{20}H_{35}O_3^+$, ΔM 2.7 ppm)	$C_{22}H_{34}O_3$
c	347.2580 $[M+H]^+$ (*calcd.* for 347.2580, $C_{20}H_{35}O_3^+$, ΔM 0.2 ppm)	$C_{22}H_{34}O_3$

Sequence overview

SEQ ID NO:	Characteristics	Sequence
1	TXS, synthetic construct, polypeptide	MSSSTGTSKVVSETSSTIVDDIPRLSANYHGDLWHHNVIQTLET PFRESSTYQERADELVVKIKDMFNALGDGDISPSAYDTAWVARL ATISSDGSEKPRFPQALNWVFNNQLQDGSWGIESHFSLCDRLL NTTNSVIALSVWKTGHSQVQQGAEFIAENLRLLNEEDELSPDFQ IIFPALLQKAKALGINLPYDLPFIKYLSTTREARLTDVSAAADNIPA NMLNALEGLEEVIDWNKIMRFQSKDGSFLSSPASTACVLMNTG DEKCFTFLNNLLDKFGGCVPCMYSIDLLERLSLVDNIEHLGIGRH FKQEIKGALDYVYRHWSERGIGWGRDSLVPDLNTTALGLRTLR MHGYNVSSDVLNNFKDENGRFFSSAGQTHVELRSVVNLFRAS DLAFPDERAMDDARKFAEPYLREALATKISTNTKLFKEIEYVVEY PWHMSIPRLEARSYIDSYDDNYVWQRKTLYRMPSLSNSKCLEL AKLDFNIVQSLHQEELKLLTRWWKESGMADINFTRHRVAEVYFS SATFEPEYSATRIAFTKIGCLQVLFDDMADIFATLDELKSFTEGV KRWDTSLLHEIPECMQTCFKVWFKLMEEVNNDVVKVQGRDML AHIRKPWELYFNCYVQEREWLEAGYIPTFEEYLKTYAISVGLGP CTLQPILLMGELVKDDVVEKVHYPSNMFELVSLSWRLTNDTKTY QAEKARGQQASGIACYMKDNPGATEEDAIKHICRVVDRALKEA SFEYFKPSNDIPMGCKSFIFNLRLCVQIFYKFIDGYGIANEEIKDYI RKVYIDPIQV
2	TXS _3xSG_MBP-SKL, synthetic	MSSSTGTSKVVSETSSTIVDDIPRLSANYHGDLWHHNVIQTLET PFRESSTYQERADELVVKIKDMFNALGDGDISPSAYDTAWVARL ATISSDGSEKPRFPQALNWVFNNQLQDGSWGIESHFSLCDRLL

	construct, polypeptide	NTTNSVIALSVWKTGHSQVQQGAEFIAENLRLLNEEDELSPDFQ IIFPALLQKAKALGINLPYDLPFIKYLSTTREARLTDVSAAADNIPA NMLNALEGLEEVIDWNKIMRFQSKDGSFLSSPASTACVLMNTG DEKCFTFLNNLLDKFGGCVPCMYSIDLLERLSLVDNIEHLGIGRH FKQEIKGALDYVYRHWSERGIGWGRDSLVPDLNTTALGLRTLR MHGYNVSSDVLNNFKDENGRFFSSAGQTHVELRSVVNLFRAS DLAFPDERAMDDARKFAEPYLREALATKISTNTKLFKEIEYVVEY PWHMSIPRLEARSYIDSYDDNYVWQRKTLYRMPSLSNSKCLEL AKLDFNIVQSLHQEELKLLTRWWKESGMADINFTRHRVAEVYFS SATFEPEYSATRIAFTKIGCLQVLFDDMADIFATLDELKSFTEGV KRWDTSLLHEIPECMQTCFKVWFKLMEEVNNDVVKVQGRDML AHIRKPWELYFNCYVQEREWLEAGYIPTFEEYLKTYAISVGLGP CTLQPILLMGELVKDDVVEKVHYPSNMFELVSLSWRLTNDTKTY QAEKARGQQASGIACYMKDNPGATEEDAIKHICRVVDRALKEA SFEYFKPSNDIPMGCKSFIFNLRLCVQIFYKFIDGYGIANEEIKDYI RKVYIDPIQVSGSGSGSKIEEGKLVIWINGDKGYNGLAEVGKKF EKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIFWAHDRFGGYA QSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLIY NKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYFTWPLI AADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVDLIKNKHM NADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLP TFKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEA VNKDKPLGAVALKSYEEELAKDPRIAATMENAQKGEIMPNIPQM SAFWYAVRTAVINAASGRQTVDEALKDAQTRITKSKL
3	CYP1, *Taxus cuspidata*, polypeptide, GenBank accession number: AAQ56240.2	MDALYKSTVAKFNEVTQLDCSTESFSIALSAIAGILLLLLLFRSKR HSSLKLPPGKLGIPFIGESFIFLRALRSNSLEQFFDERVKKFGLVF KTSLIGHPTVVLCGPAGNRLILSNEEKLVQMSWPAQFMKLMGE NSVATRRGEDHIVMRSALAGFFGPGALQSYIGKMNTEIQSHINE KWKGKDEVNVLPLVRELVFNISAILFFNIYDKQEQDRLHKLLETIL VGSFALPIDLPGFGFHRALQGRAKLNKIMLSLIKKRKEDLQSGSA TATQDLLSVLLTFRDDKGTPLTNDEILDNFSSLLHASYDTTTSPM ALIFKLLSSNPECYQKVVQEQLEILSNKEEGEEITWKDLKAMKYT WQVAQETLRMFPPVFGTFRKAITDIQYDGYTIPKGWKLLWTTYS THPKDLYFNEPEKFMPSRFDQEGKHVAPYTFLPFGGGQRSCV

		GWEFSKMEILLFVHHFVKTFSSYTPVDPDEKISGDPLPPLPSKG FSIKLFPRP
4	TcuCPR, *Taxus cuspidata*, polypeptide, GenBank accession number: AAT76449.1	MQANSNTVEGASQGKSLLDISRLDHIFALLLNGKGGDLGAMTG SALILTENSQNLMILTTALAVLVACVFFFVWRRGGSDTQKPAVR PTPLVKEEDEEEEDDSAKKKVTIFFGTQTGTAEGFAKALAEEAK ARYEKAVFKVVDLDNYAADDEQYEEKLKKEKLAFFMLATYGDG EPTDNAARFYKWFLEGKEREPWLSDLTYGVFGLGNRQYEHFN KVAKAVDEVLIEQGAKRLVPVGLGDDDQCIEDDFTAWREQVWP ELDQLLRDEDDEPTSATPYTAAIPEYRVEIYDSVVSVYEETHALK QNGQAVYDIHHPCRSNVAVRRELHTPLSDRSCIHLEFDISDTGLI YETGDHVGVHTENSIETVEEAAKLLGYQLDTIFSVHGDKEDGTP LGGSSLPPPFPGPCTLRTALARYADLLNPPRKAAFLALAAHASD PAEAERLKFLSSPAGKDEYSQWVTASQRSLLEIMAEFPSAKPPL GVFFAAIAPRLQPRYYSISSSPRFAPSRIHVTCALVYGPSPTGRI HKGVCSNWMKNSLPSEETHDCSWAPVFVRQSNFKLPADSTTPI VMVGPGTGFAPFRGFLQERAKLQEAGEKLGPAVLFFGCRNRQ MDYIYEDELKGYVEKGILTNLIVAFSREGATKEYVQHKMLEKAS DTWSLIAQGGYLYVCGDAKGMARDVHRTLHTIVQEQESVDSSK AEFLVKKLQMDGRYLRDIW
5	ERG20(F96C)-L, synthetic construct, polypeptide	MASEKEIRRERFLNVFPKLVEELNASLLAYGMPKEACDWYAHS LNYNTPGGKLNRGLSVVDTYAILSNKTVEQLGQEEYEKVAILGW CIELLQAYCLVADDMMDKSITRRGQPCWYKVPEVGEIAINDAFM LEAAIYKLLKSHFRNEKYYIDITELFHEVTFQTELGQLMDLITAPE DKVDLSKFSLKKHSFIVTFKTAYYSFYLPVALAMYVAGITDEKDL KQARDVLIPLGEYFQIQDDYLDCFGTPEQIGKIGTDIQDNKCSW VINKALELASAEQRKTLDENYGKKDSVAEAKCKKIFNDLKIEQLY HEYEESIAKDLKAKISQVDESRGFKADVLTAFLNKVYKRSKL
6	CPR2, polypeptide, *Populus trichocarpa* × *Populus deltoids*, GenBank accession	MQSSSSSMKVSPLELMQAIIKGKVDPTNVSSESGGSAAEMATLI RENREFVIILTTSIAVLIGYVVVLIWRRSSGYQKPKVPVPPKPLIVK DLEPEVDDGKKKVTIFFGTQTGTAEGFAKALAEEAKARYEKAIF KTVDLDDYAEDDDEYEEKLKKESLAIFFLATYGDGEPTDNAARF YKWFTDGNERGEWLKELPYAVFGLGNRQYEHFNKIAIVVDKILG NQGGKQLVPVGLGDDDQCMEDDFAAWRELLWPELDQLLLDG DDPTGVSTPYTAAVAEYRVVLHDPEDAPLEDDNWSNANGHAIY DAQHPCRANVTVRRELHTPASDRSCTHLEFDISGTGLVYGTGD

	number: AAK15260.1	HVGVYCENLSEIVEEALQLLGLSPDIYFTIHTDNEDGTPLSGSAL PPPFPSSTLRTALTRYADLLSSPKKSALMALAAHATNPTEADRL RHLASPAGKDEYAQWIVANHRSLLEVMAEFPSAKPPLGVFFAS VAPRLLPRYYSISSSPSMAPSRIHVTCALVLEKTPAGRIHKGVCS TWMKNAVPLEKSHDCSWAPIFVRQSNFKLPADTKVPIIMIGPGT GLAPFRGFLQERLAQKEAGAELGSSVLFFGCRNRQMDFIYEDE LNNFVESGALSELSVAFSREGPTKEYVQHKMMQKASDIWNMIS QGGYLYVCGDAKGMAKDVHRTLHTIVQEQGSLDNSKTESFVK GLQMNGRYLRDVW
7	CYP3, *Taxus cuspidata*, polypeptide	MDAFNVLMGPLAKFDNFMQLGPYSENLSVTVTVTAIAVITLLLVL IRSKPQSCVNLPPGKLGYPFIGETLQLLQAFRSNRPQQFFDERQ KKFGSVFKTSLIGDRTVVLCGPSGNRLLLSNENKLVEASWPSSS IKLIGEDSIAGKNGEKHRILRAAVNRYLGPGALQNYMAKMRSEIE HHMNEKWKGKEQVKVLPLVKENVFSIATSLFFGVNDDGERERL HDLLETALAGVFSIPLDFPGTNYRKALEARLKLDKVLSSLIERRR SDLRSGVASGNEDLLSVLLTFKDEGGNPLTDKEILDNFSTLLHA SYDTTTSALTLTLKLMSSSTECYHKVVQEQLRIVSNKKEGEEISL KDLKDMKYTWQVVQETLRMFPPLFGSFRKAITDIHYDGYTIPKG WKVLWTTYSTHGREEYFNEPEKFMPSRFEEEGRHVAPYTFLPF GAGVRTCPGWEFSKTQILLFLHYFVKTFSGYIPLDPDEKVLGNP VPPLPANGFAIKLFPRPSFDQGSPME
8	TXS, synthetic construct, nucleic acid	ATGTCTAGCTCTACGGGTACGTCTAAAGTCGTGAGTGAAACC TCATCGACGATCGTGGACGATATTCCACGCTTGTCGGCGAA CTATCATGGAGATCTGTGGCATCATAACGTCATTCAGACATT GGAAACCCCGTTTCGCGAAAGTAGCACCTACCAGGAACGGG CAGATGAATTAGTCGTGAAAATCAAAGATATGTTTAATGCATT AGGAGATGGAGACATCTCGCCCAGCGCATATGATACGGCGT GGGTGGCTCGGTTGGCCACGATTAGCTCCGATGGCAGTGAA AAGCCGCGTTTCCCGCAGGCGCTGAACTGGGTGTTTAATAAT CAATTGCAGGATGGCAGCTGGGGCATTGAATCTCACTTTAGC CTCTGTGACCGGTTACTCAACACGACAAACTCCGTAATTGCG TTGTCAGTTTGGAAAACGGGCCATAGCCAGGTTCAACAGGG CGCGGAATTTATCGCTGAAAATCTGCGCCTGCTGAACGAGG AGGACGAACTGTCACCCGATTTTCAGATTATTTTTCCGGCTTT ACTCCAGAAAGCCAAAGCCTTAGGCATCAACCTGCCATATGA

		TCTGCCGTTCATCAAGTATCTGTCTACTACCCGCGAAGCCCG TCTCACTGACGTCTCTGCGGCGGCGGACAATATTCCAGCGA ACATGCTGAACGCACTGGAAGGGCTGGAAGAGGTTATCGAC TGGAATAAAATCATGCGCTTCCAAAGCAAGGACGGTAGCTTC TTAAGCAGCCCAGCATCTACTGCTTGTGTTCTGATGAATACC GGAGACGAAAAGTGCTTTACGTTTCTGAACAATCTGCTGGAC AAATTTGGGGGTTGTGTTCCTTGTATGTATTCCATTGATCTGT TGGAACGTCTGTCGCTGGTCGATAACATTGAACACTTAGGTA TCGGCCGCCACTTCAAACAAGAAATCAAGGGGGCGTTGGAT TATGTATACCGTCATTGGAGCGAGCGTGGTATTGGTTGGGG GCGCGATAGCTTGGTACCTGATCTGAACACCACTGCTTTGG GACTGCGCACTCTTCGTATGCACGGATACAACGTTAGTTCCG ATGTCCTCAATAATTTCAAGGACGAGAACGGCCGTTTTTTCA GCTCGGCCGGTCAGACGCATGTTGAACTGCGGTCCGTAGTC AATCTCTTTCGCGCTAGTGATCTGGCCTTCCCCGACGAGCG CGCTATGGACGATGCACGGAAGTTTGCCGAGCCGTATCTCC GCGAAGCCCTGGCCACCAAAATTTCAACCAACACCAAGCTTT TCAAAGAAATTGAGTATGTAGTAGAGTATCCGTGGCATATGT CTATTCCGCGCCTGGAAGCCCGCTCGTATATCGATTCTTACG ATGACAATTATGTGTGGCAACGCAAAACACTGTACCGTATGC CCAGCCTGTCAAATAGTAAGTGTCTGGAGCTGGCGAAACTG GATTTCAACATTGTGCAATCCCTGCACCAAGAAGAGCTGAAA TTACTGACTCGCTGGTGGAAGGAATCCGGCATGGCAGACAT CAATTTTACGCGTCACCGTGTTGCAGAGGTGTACTTCTCCTC GGCGACCTTTGAGCCGGAGTATTCGGCCACACGTATTGCAT TTACCAAGATTGGCTGCCTTCAGGTGCTTTTTGACGATATGG CGGATATTTTTGCGACACTTGATGAGCTTAAATCATTTACCGA AGGCGTGAAGCGTTGGGATACCTCTCTGTTGCATGAAATCCC CGAATGTATGCAGACCTGCTTCAAAGTTTGGTTCAAACTGAT GGAAGAAGTGAACAACGACGTCGTGAAAGTTCAGGGTCGTG ATATGTTAGCACACATCCGCAAGCCGTGGGAACTCTATTTCA ATTGCTATGTGCAGGAGCGTGAATGGTTAGAAGCGGGCTAC ATTCCTACCTTCGAAGAGTACTTAAAAACCTATGCCATTTCCG TCGGTTTAGGCCCGTGCACTCTGCAGCCTATCTTGCTGATGG GTGAGCTGGTAAAGGATGATGTGGTGGAAAAAGTTCACTACC

		CGTCGAATATGTTTGAACTGGTAAGTCTGAGTTGGCGTCTGA CAAACGACACCAAAACGTACCAGGCAGAAAAGGCACGTGGG CAACAGGCAAGCGGTATCGCGTGTTATATGAAGGATAATCCG GGCGCTACTGAGGAAGATGCCATTAAGCATATCTGCCGTGTT GTGGATCGCGCTCTTAAAGAAGCGTCATTCGAATATTTTAAA CCTAGTAATGATATTCCGATGGGTTGTAAGTCATTCATTTTCA ATCTTCGCCTGTGCGTGCAAATTTTTTACAAATTTATTGACGG CTACGGAATCGCCAACGAAGAAATCAAAGACTATATTCGTAA AGTTTACATCGATCCAATCCAGGTCTAA
9	TXS_3xSG_MBP-SKL, synthetic construct, nucleic acid	ATGTCTAGCTCTACGGGTACGTCTAAAGTCGTGAGTGAAACC TCATCGACGATCGTGGACGATATTCCACGCTTGTCGGCGAA CTATCATGGAGATCTGTGGCATCATAACGTCATTCAGACATT GGAAACCCCGTTTCGCGAAAGTAGCACCTACCAGGAACGGG CAGATGAATTAGTCGTGAAAATCAAAGATATGTTTAATGCATT AGGAGATGGAGACATCTCGCCCAGCGCATATGATACGGCGT GGGTGGCTCGGTTGGCCACGATTAGCTCCGATGGCAGTGAA AAGCCGCGTTTCCCGCAGGCGCTGAACTGGGTGTTTAATAAT CAATTGCAGGATGGCAGCTGGGGCATTGAATCTCACTTTAGC CTCTGTGACCGGTTACTCAACACGACAAACTCCGTAATTGCG TTGTCAGTTTGGAAAACGGGCCATAGCCAGGTTCAACAGGG CGCGGAATTTATCGCTGAAAATCTGCGCCTGCTGAACGAGG AGGACGAACTGTCACCCGATTTTCAGATTATTTTTCCGGCTTT ACTCCAGAAAGCCAAAGCCTTAGGCATCAACCTGCCATATGA TCTGCCGTTCATCAAGTATCTGTCTACTACCCGCGAAGCCCG TCTCACTGACGTCTCTGCGGCGGCGGACAATATTCCAGCGA ACATGCTGAACGCACTGGAAGGGCTGGAAGAGGTTATCGAC TGGAATAAAATCATGCGCTTCCAAAGCAAGGACGGTAGCTTC TTAAGCAGCCCAGCATCTACTGCTTGTGTTCTGATGAATACC GGAGACGAAAAGTGCTTTACGTTTCTGAACAATCTGCTGGAC AAATTTGGGGGTTGTGTTCCTTGTATGTATTCCATTGATCTGT TGGAACGTCTGTCGCTGGTCGATAACATTGAACACTTAGGTA TCGGCCGCCACTTCAAACAAGAAATCAAGGGGGCGTTGGAT TATGTATACCGTCATTGGAGCGAGCGTGGTATTGGTTGGGG GCGCGATAGCTTGGTACCTGATCTGAACACCACTGCTTTGG GACTGCGCACTCTTCGTATGCACGGATACAACGTTAGTTCCG

		ATGTCCTCAATAATTTCAAGGACGAGAACGGCCGTTTTTTCA GCTCGGCCGGTCAGACGCATGTTGAACTGCGGTCCGTAGTC AATCTCTTTCGCGCTAGTGATCTGGCCTTCCCCGACGAGCG CGCTATGGACGATGCACGGAAGTTTGCCGAGCCGTATCTCC GCGAAGCCCTGGCCACCAAAATTTCAACCAACACCAAGCTTT TCAAAGAAATTGAGTATGTAGTAGAGTATCCGTGGCATATGT CTATTCCGCGCCTGGAAGCCCGCTCGTATATCGATTCTTACG ATGACAATTATGTGTGGCAACGCAAAACACTGTACCGTATGC CCAGCCTGTCAAATAGTAAGTGTCTGGAGCTGGCGAAACTG GATTTCAACATTGTGCAATCCCTGCACCAAGAAGAGCTGAAA TTACTGACTCGCTGGTGGAAGGAATCCGGCATGGCAGACAT CAATTTTACGCGTCACCGTGTTGCAGAGGTGTACTTCTCCTC GGCGACCTTTGAGCCGGAGTATTCGGCCACACGTATTGCAT TTACCAAGATTGGCTGCCTTCAGGTGCTTTTTGACGATATGG CGGATATTTTTGCGACACTTGATGAGCTTAAATCATTTACCGA AGGCGTGAAGCGTTGGGATACCTCTCTGTTGCATGAAATCCC CGAATGTATGCAGACCTGCTTCAAAGTTTGGTTCAAACTGAT GGAAGAAGTGAACAACGACGTCGTGAAAGTTCAGGGTCGTG ATATGTTAGCACACATCCGCAAGCCGTGGGAACTCTATTTCA ATTGCTATGTGCAGGAGCGTGAATGGTTAGAAGCGGGCTAC ATTCCTACCTTCGAAGAGTACTTAAAAACCTATGCCATTTCCG TCGGTTTAGGCCCGTGCACTCTGCAGCCTATCTTGCTGATGG GTGAGCTGGTAAAGGATGATGTGGTGGAAAAAGTTCACTACC CGTCGAATATGTTTGAACTGGTAAGTCTGAGTTGGCGTCTGA CAAACGACACCAAAACGTACCAGGCAGAAAAGGCACGTGGG CAACAGGCAAGCGGTATCGCGTGTTATATGAAGGATAATCCG GGCGCTACTGAGGAAGATGCCATTAAGCATATCTGCCGTGTT GTGGATCGCGCTCTTAAAGAAGCGTCATTCGAATATTTTAAA CCTAGTAATGATATTCCGATGGGTTGTAAGTCATTCATTTTCA ATCTTCGCCTGTGCGTGCAAATTTTTTACAAATTTATTGACGG CTACGGAATCGCCAACGAAGAAATCAAAGACTATATTCGTAA AGTTTACATCGATCCAATCCAGGTCAGTGGCAGTGGTTCCGG TATGTCTAAAATTGAAGAAGGTAAATTGGTCATCTGGATCAAC GGTGATAAGGGTTATAATGGTTTGGCTGAAGTCGGTAAGAAG TTCGAAAAAGATACTGGTATCAAGGTTACCGTAGAACACCCA

		GATAAGTTGGAAGAAAAATTCCCACAAGTTGCTGCTACTGGT GATGGTCCAGATATTATCTTTTGGGCTCATGATAGATTTGGT GGTTACGCTCAATCTGGTTTGTTGGCTGAAATTACTCCAGAT AAGGCTTTCCAGGATAAGTTGTATCCTTTTACTTGGGATGCC GTTAGGTACAACGGTAAATTGATTGCTTATCCAATTGCCGTC GAAGCCTTGAGTTTGATCTATAACAAAGACCTGTTGCCAAAT CCACCTAAGACTTGGGAAGAAATTCCAGCTTTGGACAAAGAG CTAAAGGCTAAGGGTAAATCTGCTCTGATGTTTAACTTGCAA GAACCTTACTTTACCTGGCCATTGATTGCAGCTGATGGTGGT TATGCTTTTAAGTACGAGAATGGCAAGTACGATATCAAGGAT GTTGGTGTTGATAATGCTGGTGCTAAAGCTGGTTTGACCTTT TTGGTTGATCTGATCAAAAACAAGCACATGAACGCCGATACC GATTACTCTATTGCTGAAGCTGCTTTCAACAAAGGTGAAACT GCTATGACTATTAACGGTCCATGGGCTTGGTCTAACATTGAT ACTTCTAAGGTTAACTACGGTGTCACTGTTTTGCCAACTTTCA AAGGTCAACCATCTAAGCCATTCGTTGGTGTTTTATCTGCTG GTATTAACGCTGCTTCTCCAAACAAAGAATTGGCTAAAGAGT TCTTGGAGAACTACTTGTTGACCGACGAAGGTTTGGAAGCAG TTAACAAAGATAAGCCATTGGGTGCTGTTGCCTTGAAATCTTA CGAAGAAGAATTAGCCAAGGATCCAAGAATTGCTGCCACTAT GGAAAATGCACAAAAGGGCGAAATTATGCCAAACATCCCACA AATGTCTGCTTTTTGGTATGCAGTTAGAACTGCTGTTATCAAT GCTGCATCAGGTAGACAAACAGTTGATGAAGCTTTGAAGGAT GCTCAAACTAGAATCACCAAGAGCAAACTGTAA
10	TcuCYP1, synthetic construct, nucleic acid	ATGGACGCCCTGTACAAGAGCACCGTGGCCAAGTTCAACGA GGTGACCCAGCTGGACTGCAGCACCGAGAGCTTCTCTATCG CCCTGTCTGCTATCGCCGGCATCCTTTTGTTGCTGCTGCTGT TCAGGTCCAAGAGGCACAGCAGCCTGAAGCTGCCACCAGGC AAGCTTGGCATCCCATTCATCGGCGAGTCCTTCATCTTCCTG AGGGCCCTGAGGTCCAACAGCCTCGAGCAGTTCTTCGACGA GAGGGTGAAGAAGTTCGGCCTGGTTTTCAAGACCAGCCTGA TCGGCCATCCGACCGTGGTTCTTTGCGGCCCAGCTGGCAAC AGGCTGATCCTGAGCAACGAAGAGAAGCTGGTGCAGATGAG CTGGCCGGCTCAGTTCATGAAGCTCATGGGCGAGAACTCCG TGGCTACTAGAAGAGGCGAGGACCACATCGTGATGAGGTCT

		GCTCTGGCTGGCTTCTTCGGTCCAGGCGCTCTGCAGTCTTA CATCGGCAAGATGAACACCGAGATCCAGAGCCACATCAACG AGAAGTGGAAGGGCAAAGACGAGGTGAACGTGCTGCCACTG GTGCGCGAGCTGGTGTTCAACATCAGCGCCATCCTGTTCTTC AACATCTACGACAAGCAAGAGCAGGACAGGCTGCACAAGCT GCTCGAGACTATCCTCGTGGGCAGCTTCGCTCTGCCAATCG ATCTTCCAGGCTTCGGCTTCCACAGGGCTCTGCAAGGTAGG GCCAAGCTGAACAAGATCATGCTGAGCCTCATCAAGAAGAG GAAAGAGGACCTGCAGTCCGGCAGCGCTACCGCTACTCAGG ATCTGCTGTCTGTGCTGCTGACCTTCAGGGACGATAAGGGC ACCCCACTGACCAACGACGAGATCCTGGACAACTTCAGCAG CCTCCTGCACGCCTCTTACGACACCACCACCTCTCCGATGG CTCTGATCTTCAAGCTGCTGAGCAGCAACCCAGAGTGCTACC AGAAGGTCGTCCAAGAGCAGCTCGAGATCCTCTCCAACAAA GAGGAAGGCGAGGAAATCACCTGGAAGGACCTGAAGGCCAT GAAGTACACCTGGCAGGTCGCCCAAGAAACCCTGAGGATGT TCCCACCAGTGTTCGGCACCTTCCGCAAGGCCATCACCGAC ATCCAGTACGACGGCTACACGATCCCGAAAGGCTGGAAGCT GCTGTGGACCACCTACAGCACCCATCCGAAGGACCTCTACT TCAACGAGCCAGAGAAGTTCATGCCGAGCAGGTTCGACCAA GAGGGCAAACATGTGGCCCCGTACACCTTCCTTCCATTCGG TGGCGGTCAGAGGTCTTGCGTTGGCTGGGAGTTCAGCAAGA TGGAAATCCTGCTGTTCGTGCACCACTTCGTCAAGACCTTCT CCAGCTACACCCCGGTGGACCCGGATGAGAAGATTTCTGGC GATCCACTGCCGCCGCTGCCAAGCAAGGGCTTCTCAATCAA GCTGTTCCCGAGGCCGTGA
11	TbaCYP1, synthetic construct, nucleic acid	ATGGACGCCCTGTATAAGAGCACAGTTGCAAAATTTAATGAG GTCACACAGCTGGACTGTTCCACTGAATCTTTTTCCATTGCC CTCTCAGCTATTGCTGGCATTCTTCTGCTTCTCCTGCTCTTCC GTTCTAAACGCCACTCCTCCCTTAAACTTCCTCCTGGGAAAT TAGGCATCCCTTTCATTGGCGAGTCGTTTATCTTCCTGAGGG CTCTTCGATCGAACTCGCTGGAGCAATTTTTTGACGAGAGAG TGAAAAAATTCGGTCTCGTGTTCAAAACCTCCTTAATTGGGC ATCCCACAGTAGTACTCTGCGGCCCTGCGGGAAACCGGCTT ATTCTGTCCAACGAGGAGAAGCTGGTGCAAATGTCGTGGCC

		CGCTCAATTTATGAAGCTCATGGGAGAGAATTCCGTTGCCAC CAGGAGGGGTGAAGACCATATAGTTATGCGCTCTGCTCTTG CAGGTTTTTTCGGCCCTGGTGCGCTGCAGAGTTACATTGGTA AAATGAATACAGAGATCCAGAGTCATATCAACGAAAAATGGA AGGGAAAAGATGAGGTGAATGTACTTCCTTTGGTAAGAGAGC TCGTCTTCAACATTTCGGCCATCTTGTTTTTCAACATATATGA TAAGCAGGAACAGGATCGTCTGCATAAGCTTTTGGAAACTAT TCTGGTCGGAAGTTTTGCTCTTCCGATTGACTTGCCCGGATT TGGTTTCCATAGAGCACTCCAGGGACGGGCCAAGCTCAACA AAATTATGCTGTCTTTAATTAAAAAGAGAAAAGAAGATCTGCA GTCTGGATCGGCAACAGCCACTCAGGATCTGCTCTCTGTTTT GCTCACTTTCAGAGATGACAAAGGGACTCCACTCACCAATGA CGAGATACTCGACAACTTTTCTTCTCTGCTCCATGCCTCCTAT GACACCACCACTTCGCCAATGGCTTTGATTTTCAAGCTCTTG TCTTCCAATCCAGAATGCTATCAAAAAGTAGTTCAAGAGCAAT TGGAGATACTTTCCAACAAAGAGGAGGGCGAAGAGATCACA TGGAAGGATCTGAAAGCCATGAAATACACATGGCAAGTAGCT CAGGAAACGCTGCGGATGTTTCCTCCAGTTTTCGGAACATTT CGCAAGGCCATCACTGACATTCAGTATGATGGTTACACAATT CCAAAAGGGTGGAAGCTGTTGTGGACAACTTACAGTACACAT CCCAAGGACTTGTATTTCAATGAACCAGAGAAATTCATGCCT TCAAGATTCGATCAGGAAGGAAAGCATGTAGCTCCTTACACA TTTTTACCCTTCGGTGGAGGGCAGCGGTCATGTGTGGGATG GGAATTTTCAAAGATGGAGATTTTACTGTTCGTTCATCATTTT GTCAAAACTTTTAGCAGCTACACCCCTGTTGATCCCGACGAA AAAATATCAGGGGATCCACTCCCTCCTCTTCCTTCCAAGGGA TTTTCCATTAAACTGTTTCCGAGACCATAG
12	*Tcu*CPR, synthetic construct, nucleic acid, GenBank accession number: AY571340.1	ATGCAGGCCAACAGCAACACTGTTGAGGGCGCTAGCCAGGG CAAGAGCCTGCTGGATATCTCTAGGCTGGACCACATCTTCGC CCTGCTGCTGAATGGCAAAGGCGGCGATCTTGGCGCTATGA CCGGCTCTGCTCTGATCCTGACCGAGAACAGCCAGAACCTG ATGATCCTCACCACCGCTCTGGCTGTGCTGGTGGCTTGCGT GTTCTTCTTTGTTTGGCGTCGCGGCGGTAGCGACACCCAGA AGCCAGCTGTTAGGCCAACGCCGCTCGTGAAAGAAGAGGAC GAAGAGGAAGAGGACGACAGCGCGAAGAAGAAGGTCACCAT

		TTTCTTCGGCACCCAGACCGGCACTGCTGAGGGTTTCGCTA AGGCTCTCGCCGAGGAAGCCAAGGCCAGATACGAGAAGGC CGTGTTCAAGGTGGTGGACCTGGACAACTACGCCGCTGACG ACGAGCAGTACGAGGAAAAGCTCAAGAAAGAGAAGCTGGCC TTCTTCATGCTGGCCACCTACGGTGATGGCGAGCCAACCGA TAACGCCGCCAGGTTCTACAAGTGGTTCCTCGAGGGCAAAG AGCGCGAGCCTTGGCTGTCTGATCTGACCTACGGCGTTTTC GGCCTCGGCAACAGGCAATACGAGCACTTCAACAAGGTGGC CAAGGCCGTGGACGAGGTGCTGATTGAGCAAGGCGCTAAGA GGCTGGTGCCAGTTGGCTTGGGCGACGACGATCAGTGCATC GAGGACGACTTTACCGCCTGGCGCGAGCAAGTTTGGCCAGA GCTGGATCAACTGCTGCGCGACGAGGATGATGAGCCAACCT CTGCTACCCCGTACACCGCTGCTATTCCAGAGTACCGCGTG GAAATCTACGACAGCGTGGTGAGCGTGTACGAAGAGACTCA CGCCCTGAAGCAGAACGGCCAGGCTGTGTACGACATCCACC ATCCGTGCAGGTCCAATGTGGCTGTGAGGCGCGAGCTGCAT ACCCCACTGTCTGACAGGTCTTGCATCCACCTCGAGTTCGAC ATCAGCGACACCGGCCTGATCTACGAGACTGGCGATCATGT GGGCGTGCACACCGAGAACTCCATCGAGACTGTGGAAGAGG CCGCTAAGCTGCTGGGCTACCAGCTGGACACCATCTTCTCT GTGCACGGCGACAAAGAGGACGGCACTCCACTTGGCGGTTC TAGCCTGCCACCACCATTTCCAGGACCGTGCACTCTTAGGAC CGCTCTCGCTAGGTACGCCGACCTTCTTAACCCTCCTAGGAA GGCCGCTTTCCTGGCTCTCGCTGCTCACGCTTCAGATCCAG CTGAGGCTGAGAGGCTGAAGTTCCTGTCAAGCCCAGCTGGC AAGGACGAGTACTCTCAGTGGGTGACAGCTAGCCAGCGCAG CCTGCTTGAGATCATGGCTGAGTTCCCAAGCGCCAAGCCTC CACTGGGCGTTTTCTTCGCTGCTATCGCTCCAAGGCTCCAGC CGAGGTACTACAGCATCAGCAGCTCTCCAAGGTTCGCCCCA TCTCGCATCCATGTGACTTGCGCTCTGGTGTACGGCCCATCT CCAACCGGCAGGATTCATAAGGGCGTGTGCAGCAACTGGAT GAAGAACAGCCTGCCGAGCGAGGAAACCCACGACTGTTCTT GGGCTCCAGTGTTCGTGAGGCAGAGCAACTTCAAGCTGCCG GCCGACAGCACCACTCCGATCGTTATGGTTGGTCCAGGCAC CGGCTTCGCTCCGTTCAGAGGTTTCCTTCAAGAGAGGGCCA

		AGCTGCAAGAGGCCGGCGAAAAACTTGGCCCAGCCGTTCTG TTCTTCGGCTGCAGAAACAGGCAGATGGACTACATCTACGAG GACGAGCTGAAGGGCTACGTCGAGAAGGGCATCCTGACCAA CCTGATCGTGGCTTTCTCACGCGAGGGCGCCACCAAAGAGT ACGTCCAGCATAAGATGCTCGAGAAGGCGAGCGACACTTGG AGCCTGATTGCTCAAGGCGGCTACCTGTACGTTTGCGGCGA CGCTAAAGGCATGGCTAGGGATGTGCACAGGACCCTGCACA CCATCGTGCAAGAGCAAGAGAGCGTGGACAGCAGCAAGGC CGAGTTCCTGGTGAAGAAACTCCAGATGGACGGCCGCTACC TGAGGGACATTTGGTGA
13	ERG20(F96C)-L, synthetic construct, nucleic acid	ATGGCTTCAGAAAAAGAAATTAGGAGAGAGAGATTCTTGAAC GTTTTCCCTAAATTAGTAGAGGAATTGAACGCATCGCTTTTG GCTTACGGTATGCCTAAGGAAGCATGTGACTGGTATGCCCA CTCATTGAACTACAACACTCCAGGCGGTAAGCTAAATAGAGG TTTGTCCGTTGTGGACACGTATGCTATTCTCTCCAACAAGAC CGTTGAACAATTGGGGCAAGAAGAATACGAAAAGGTTGCCAT TCTAGGTTGGTGCATTGAGTTGTTGCAGGCTTACTGTTTGGT CGCCGATGATATGATGGACAAGTCCATTACCAGAAGAGGCC AACCATGTTGGTACAAGGTTCCTGAAGTTGGGGAAATTGCCA TCAATGACGCATTCATGTTAGAGGCTGCTATCTACAAGCTTTT GAAATCTCACTTCAGAAACGAAAAATACTACATAGATATCACC GAATTGTTCCATGAGGTCACCTTCCAAACCGAATTGGGCCAA TTGATGGACTTAATCACTGCACCTGAAGACAAAGTCGACTTG AGTAAGTTCTCCCTAAAGAAGCACTCCTTCATAGTTACTTTCA AGACTGCTTACTATTCTTTCTACTTGCCTGTCGCATTGGCCAT GTACGTTGCCGGTATCACGGATGAAAAGGATTTGAAACAAGC CAGAGATGTCTTGATTCCATTGGGTGAATACTTCCAAATTCAA GATGACTACTTAGACTGCTTCGGTACCCCAGAACAGATCGGT AAGATCGGTACAGATATCCAAGATAACAAATGTTCTTGGGTA ATCAACAAGGCATTAGAACTTGCTTCCGCAGAACAAAGAAAG ACTTTAGACGAAAATTACGGTAAGAAGGACTCAGTCGCAGAA GCCAAATGCAAAAAGATTTTCAATGACTTGAAAATTGAACAGC TATACCACGAATATGAAGAGTCTATTGCCAAGGATTTGAAGG CTAAAATTTCTCAGGTCGATGAGTCTCGTGGCTTCAAAGCTG ATGTCTTAACTGCGTTTTTGAACAAAGTTTACAAGAGATAA

14	SfFDPS1, synthetic construct, nucleic acid	atggcgaatctgaacggagccgcgtcggatctgagggcgacgtttctgggggtttattcga cgcttaaatccgagctcttgagcgaccctgctttcgagtggactgatggctctcgtcaatgg gtcgagcgtatgttggactacaatgtacctggagggaagttaaaccgaggcctgtcagtta ttgatagctacaagttactaaaagaaggaaaagatctgactgatgatgaagtgttcctagc tagtgctctcggctggtgtattgaatggctccaggcatactttctcgtccttgatgatattatgga taattctcacacacgacgtggtcagccatgctggtttaaagtcccaaaggttggtatgattgc cataaatgatggaatcattctccggaaccatatccccagaattcttaagaagcacttcaga acaaagccttactatgtcgatctgcttgatttgttcaatgaggtggaatttcaaactgcttctgg acagatgatagatttaattaccactattgaaggagaaaaagatctgtcaaaatactcattgc ctctccatcgccgcattgtacagtacaagactgcctactactcattttacctcccagttgcttgt gcgttgctcatggcgggtgaggacctggagaaacatccaacagtaaaagatgtgcttatt aatatgggaatatactttcaagtacaggatgactatttagattgctttggtgagcctgaaaag attggaaagattggaacagatattgaagatttcaaatgttcgtggctggttgtaaaggccct ggagctttgtaatgaagaacaaaagaaaaatctttttgagcactatggaaaggaagatcc aactgctgttgcaaagatcaaagtcctctacaatgagcttaatctccaaggtgcgtttgctga gttcgaaagccagagctacgagaaactaaatagctcgattgaagctcatcccagcaaat ctgtgcaagcagtgctcaagtctttcttgggcaagatatacaagaggcagaaaTAA
15	*Tcu*CYP3, synthetic construct, nucleic acid	ATGGACGCTTTTAATGTTTTAATGGGCCCTCTAGCAAAATTTG ATAATTTCATGCAGCTCGGCCCTTACTCTGAAAATCTTTCCGT TACAGTTACCGTCACAGCTATTGCCGTCATTACTCTTCTCCTG GTGTTGATCCGTTCCAAACCCCAATCTTGTGTAAACCTTCCT CCGGGAAAGCTTGGCTACCCTTTCATCGGCGAAACATTACAA TTGTTGCAGGCATTTCGATCGAACAGGCCGCAACAGTTCTTT GATGAGAGGCAGAAGAAATTTGGGTCTGTTTTCAAGACTTCA CTAATTGGGGACCGCACAGTGGTGCTGTGCGGTCCCTCAGG AAACCGTTTGCTGCTCTCCAACGAAAACAAGCTGGTGGAGG CATCCTGGCCGAGTTCTTCCATTAAATTGATCGGAGAGGATT CCATTGCTGGGAAAAACGGAGAGAAGCATCGGATCTTACGC GCCGCGGTAAACCGTTACCTGGGACCCGGAGCATTACAGAA TTATATGGCGAAGATGAGGTCAGAAATCGAACATCATATGAA TGAGAAATGGAAGGGGAAAGAGCAAGTGAAGGTGCTTCCTT TGGTAAAAGAGAATGTCTTCTCCATCGCAACCAGCTTGTTTTT CGGTGTCAATGATGACGGAGAACGGGAACGGCTTCATGACC TTTTGGAAACCGCACTTGCGGGTGTTTTTTCTATTCCACTGG ATTTTCCAGGAACAAATTATCGGAAAGCCCTTGAAGCGCGGT

		TAAAACTGGATAAAGTCCTTTCTTCTCTGATAGAAAGGAGAA GAAGCGATCTGCGATCAGGCGTGGCATCTGGTAATGAGGAT CTGCTCTCTGTGTTGCTCACTTTCAAAGACGAAGGAGGGAAT CCTCTGACAGACAAGGAGATCCTCGACAACTTCTCCACCCTG CTTCATGCATCATATGACACCACAACCTCAGCACTCACCTTG ACATTAAAGCTCATGTCCTCCTCTACTGAATGCTATCACAAAG TAGTTCAAGAGCAACTGAGAATAGTTTCCAACAAAAAGGAGG GAGAAGAAATCAGCTTGAAAGATCTGAAAGACATGAAATATA CATGGCAAGTTGTGCAGGAAACTCTGAGGATGTTCCCTCCG CTTTTTGGATCATTTCGTAAGGCCATCACTGACATTCATTATG ATGGTTATACAATCCCAAAAGGATGGAAAGTTTTATGGACAA CTTATAGTACACATGGGAGAGAAGAGTATTTCAATGAACCAG AGAAATTCATGCCTTCAAGATTCGAAGAGGAAGGAAGGCATG TTGCTCCTTACACATTTTTACCCTTCGGAGCAGGCGTGCGCA CCTGCCCAGGATGGGAATTTTCAAAAACCCAGATATTACTGT TCTTACATTATTTTGTTAAAACTTTCAGTGGCTACATCCCACT CGACCCTGACGAAAAAGTGTTAGGGAATCCAGTCCCTCCTCT CCCTGCCAATGGATTTGCTATAAAACTTTTCCCCAGGCCCTC ATTCGATCAAGGATCACCCATGGAATAA
16	P19, Tobacco virus 1 isolate AnHui, polypeptide, GenBank accession number: AKT94764.1	MLSDYEVGSNNSSHYLAIAVVCSENSDELLMFESEDSNENAIFF EDSCVISKSPSGGALNCAENYYLTSNAGLLDFLKSTNFKSPKPL PVANRCSNGPNHVDFHMRPLHLGLIFRVNLSLKSLDDLDTFQD SLVKIKGFWNNGAQHQSMFHYLSNFINYEKYSIVSFEYFKL
17	P19, Tobacco virus 1 isolate AnHui, synthetic construct, nucleic acid, GenBank	ATGTTAAGTGATTACGAAGTGGGATCCAATAATTCATCACACT ACTTGGCTATTGCCGTAGTGTGTTCCGAAAATTCCGATGAAT TACTAATGTTCGAATCGGAAGACAGCAACGAAAACGCAATTT TCTTTGAAGACAGTTGCGTTATTTCTAAGTCACCGTCCGGCG GTGCTTTAAATTGCGCTGAAAATTATTACTTGACTAGTAACGC AGGTTTGTTAGATTTTTTAAAATCTACTAATTTTAAATCACCCA AGCCTTTACCCGTGGCTAACAGATGCAGTAACGGACCTAACC ACGTTGACTTTCATATGAGACCTCTTCATTTAGGGCTTATATT

	accession number: KT203917.1	CCGGGTGAACCTATCTTTGAAGTCTCTCGACGATCTTGACAC TTTTCAGGATTCTTTAGTTAAGATAAAAGGTTTTTGGAATAAC GGTGCGCAGCACCAGTCTATGTTTCACTATCTATCTAACTTC ATAAACTACGAAAAGTATTCTATAGTGTCCTTTGAGTATTTTA AGTTATGA
18	US_TXS_F, synthetic construct, nucleic acid	ATCAACGGGUAAAATGTCTAGCTCTACGGGTAC
19	US_TXS_R_3SG, synthetic construct, nucleic acid	ACTGCCACUGACCTGGATTGGATCGATGTAA
20	HMG2(K6R), synthetic construct, polypeptide	MSLPLRTIVHLVKPFACTARFSARYPIHVIVVAVLLSAAAYLSVTQ SYLNEWKLDSNQYSTYLSIKPDELFEKCTHYYRSPVSDTWKLLS SKEAADIYTPFHYYLSTISFQSKDNSTTLPSLDDVIYSVDHTRYLL SEEPKIPTELVSENGTKWRLRNNSNFILDLHNIYRNMVKQFSNK TSEFDQFDLFIILAAYLTLFYTLCCLFNDMRKIGSKFWLSFSALSN SACALYLSLYTTHSLLKKPASLLSLVIGLPFIVVIIGFKHKVRLAAF SLQKFHRISIDKKITVSNIIYEAMFQEGAYLIRDYLFYISSFIGCAIY ARHLPGLVNFCILSTFMLVFDLLLSATFYSAILSMKLEINIIHRSTVI RQTLEEDGVVPTTADIIYKDETASEPHFLRSNVAIILGKASVIGLLL LINLYVFTDKLNATILNTVYFDSTIYSLPNFINYKDIGNLSNQVIISV LPKQYYTPLKKYHQIEDSVLLIIDSVSNAIRDQFISKLLFFAFAVSI SINVYLLNAAKIHTGYMNFQPQSNKIDDLVVQQKSATIEFSETRS MPASSGLETPVTAKDIIISEEIQNNECVYALSSQDEPIRPLSNLVE LMEKEQLKNMNNTEVSNLVVNGKLPLYSLEKKLEDTTRAVLVR RKALSTLAESPILVSEKLPFRNYDYDRVFGACCENVIGYMPIPVG VIGPLIIDGTSYHIPMATTEGCLVASAMRGCKAINAGGGATTVLT KDGMTRGPVVRFPTLIRSGACKIWLDSEEGQNSIKKAFNSTSRF ARLQHIQTCLAGDLLFMRFRTTTGDAMGMNMISKGVEYSLKQM VEEYGWEDMEVVSVSGNYCTDKKPAAINWIEGRGKSVVAEATI PGDVVKSVLKSDVSALVELNISKNLVGSAMAGSVGGFNAHAAN LVTALFLALGQDPAQNVESSNCITLMKEVDGDLRISVSMPSIEVG TIGGGTVLEPQGAMLDLLGVRGPHPTEPGANARQLARIIACAVL

		AGELSLCSALAAGHLVQSHMTHNRKTNKANELPQPSNKGPPCK TSALL
21	US_MBPv_3S G_F, synthetic construct, nucleic acid	AGTGGCAGUGGTTCCGGTATGTCTAAAATTGAAGAAGGTAAA TTGGTC
22	US_MBPv_R, nucleic acid	CGTGCGAUTTACTTGGTGATTCTAGTTTGAGC
23	US_MBPv_SK L_R, synthetic construct, nucleic acid	CGTGCGAUTTACAGCTTGGACTTGGTGATTCTAGTTTGAGC
24	US_TcuCYP1 _F, synthetic construct, nucleic acid	ATCAACGGGUATGGACGCCCTGTACAAGAGC
25	US_TcuCYP1 _R, synthetic construct, nucleic acid	CGTGCGAUTCACGGCCTCGGGAACAGC
26	US_TbaCYP1 _F, synthetic construct, nucleic acid	ATCAACGGGUAAAATGGACGCCCTGTATAAGAGC
27	US_TbaCYP1 _R, synthetic construct, nucleic acid	CGTGCGAUCTATGGTCTCGGAAACAGTTTAATG
28	US_TcuCYP3 _F, synthetic construct, nucleic acid	ATCAACGGGUAAAATGGACGCTTTTAATGTTTTAATGGG
29	US_TcuCYP3 _R, synthetic	CGTGCGAUTTATTCCATGGGTGATCCTTGATC

	construct, nucleic acid	
30	US_TcCPR_F, synthetic construct, nucleic acid	ATCAACGGGUAAAATGCAGGCCAACAGCAACAC
31	US_TcCPR_R, synthetic construct, nucleic acid	CGTGCGAUTCACCAAATGTCCCTCAGGTA
32	US_TcCPR_F 10, synthetic construct, nucleic acid	AGCGATACGUAAAATGCAGGCCAACAGCAACAC
33	US_TcCPR_R C, synthetic construct, nucleic acid	CACGCGAUTCACCAAATGTCCCTCAGGTA
34	US_ScErg20_ F, synthetic construct, nucleic acid	ATCAACGGGUAAAATGGCTTCAGAAAAAGAAATTAGGAG
35	US_ScErg20_ SKL_R, synthetic construct, nucleic acid	CGTGCGAUTTACAGTTTGCTTTTGCTTCTCTTGTAAACTTTG
36	SpCytb5, synthetic construct, nucleic acid	atggcgaaatctcacacatttgaggaggtcgcgaagcacaacaagactaaggattgctg gctcattatcagcggaaaggtttatgatgtgaccccattcatggaggatcatcctggaggtg atgaagttttgctgtccgcgaccggaaaagacgctaccaatgattttgaagatgttgggca cagcgattctgctcgggagatgatggacaaatacttcatcggtgagatagacatggcaac ggttcccctaaaacgaagctacatagctccacaacaaccatcatacaatccagacaaga ccccagagtttgtgatcaagatcttgcagttcctcgtgcctctcctgatcttgggattggcattt gctgtccggctctacaccaaagagaaaTGA

37	SpCBR, synthetic construct, nucleic acid	atggatttcttggcaacatctgaaggccaattggtagttggagttgctgttggtctcgtagctat tggtgctgcgtacttcttgttttcctccaagaagcccaaagtgtgcttagatgccgagaacttc aaggaatttaagcttgtgaagaaaacacaactaagtcataatgtggccaagttccggtttg ctctccctacgcctacttctgtattaggccttcctattggacagcacataagctgcaggggta aggacagtcaaggtgaagaggttattaaaccatacactccaactactttggattctgatgtt ggatgctttgaattagttataaagatgtatccacaagggagaatgtctcatcattttagagaa atgaaggaaggtgattatatggctgttaaaggaccaaagggccgttttaggtatcaacca ggccaagtgagagcatttggtatgcttgctggaggctctgggattacaccaatgttccaggt tgccagagcaatcctggagaatcccagtgacaaaacaaaggtgcatctgatttatgctaat gttacggctgatgacattcttctaaaggatgaaatcgatggtcttgccaaaaattatccgga ccagttccaagtttactacgtgctcaatcagcctccggaagtatggaacggtggagttggttt tgtgtcaaaggaaatgattcaggcgcactgccctcgtccagcctctgacattaagattctga gatgcggtccgcccccgatgaacaaggccatggctgctcatcttgaggcactcgaatact cagctgacatgcagttccagttcTAA
38	SpCBR, *Salvia pomifera*, polypeptide	MDFLATSEGQLVVGVAVGLVAIGAAYFLFSSKKPKVCLDAENFK EFKLVKKTQLSHNVAKFRFALPTPTSVLGLPIGQHISCRGKDSQ GEEVIKPYTPTTLDSDVGCFELVIKMYPQGRMSHHFREMKEGD YMAVKGPKGRFRYQPGQVRAFGMLAGGSGITPMFQVARAILE NPSDKTKVHLIYANVTADDILLKDEIDGLAKNYPDQFQVYYVLNQ PPEVWNGGVGFVSKEMIQAHCPRPASDIKILRCGPPPMNKAMA AHLEALEYSADMQFQF
39	SfFDPS1, *Salvia fruticosa*, polypeptide	MANLNGAASDLRATFLGVYSTLKSELLSDPAFEWTDGSRQWVE RMLDYNVPGGKLNRGLSVIDSYKLLKEGKDLTDDEVFLASALG WCIEWLQAYFLVLDDIMDNSHTRRGQPCWFKVPKVGMIAINDGI ILRNHIPRILKKHFRTKPYYVDLLDLFNEVEFQTASGQMIDLITTIE GEKDLSKYSLPLHRRIVQYKTAYYSFYLPVACALLMAGEDLEKH PTVKDVLINMGIYFQVQDDYLDCFGEPEKIGKIGTDIEDFKCSWL VVKALELCNEEQKKNLFEHYGKEDPTAVAKIKVLYNELNLQGAF AEFESQSYEKLNSSIEAHPSKSVQAVLKSFLGKIYKRQK
40	SpCytb5, *Salvia pomifera*, polypeptide	MAKSHTFEEVAKHNKTKDCWLIISGKVYDVTPFMEDHPGGDEV LLSATGKDATNDFEDVGHSDSAREMMDKYFIGEIDMATVPLKR SYIAPQQPSYNPDKTPEFVIKILQFLVPLLILGLAFAVRLYTKEK
41	HMG2(K6R), synthetic	ATGTCACTTCCCTTAAGAACGATAGTACATTTGGTAAAGCCCT TTGCTTGCACTGCTAGGTTTAGTGCGAGATACCCAATCCACG

	construct, nucleic acid	TCATTGTTGTTGCTGTTTTATTGAGTGCCGCTGCTTATCTATC CGTGACACAATCTTACCTTAACGAATGGAAGCTGGACTCTAA TCAGTATTCTACATACTTAAGCATAAAGCCGGATGAGTTGTTT GAAAAATGCACACACTACTATAGGTCTCCTGTGTCTGATACA TGGAAGTTACTCAGCTCTAAAGAAGCCGCCGATATTTATACC CCTTTTCATTATTATTTGTCTACCATAAGTTTTCAAAGTAAGGA CAATTCAACGACTTTGCCTTCCCTTGATGACGTTATTTACAGT GTTGACCATACCAGGTACTTATTAAGTGAAGAGCCAAAGATA CCAACTGAACTAGTGTCTGAAAACGGAACGAAATGGAGATTG AGAAACAACAGCAATTTTATTTTGGACCTGCATAATATTTACC GAAATATGGTGAAGCAATTTTCTAACAAAACGAGCGAATTTG ATCAGTTCGATTTGTTTATCATCCTAGCTGCTTACCTTACTCT TTTTTATACTCTCTGTTGCCTGTTTAATGACATGAGGAAAATC GGATCAAAGTTTTGGTTAAGCTTTTCTGCTCTTTCAAACTCTG CATGCGCATTATATTTATCGCTGTACACAACTCACAGTTTATT GAAGAAACCGGCTTCCTTATTAAGTTTGGTCATTGGACTACC ATTTATCGTAGTAATTATTGGCTTTAAGCATAAAGTTCGACTT GCGGCATTCTCGCTACAAAAATTCCACAGAATTAGTATTGAC AAGAAAATAACGGTAAGCAACATTATTTATGAGGCTATGTTTC AAGAAGGTGCCTACTTAATCCGCGACTACTTATTTTATATTAG CTCCTTCATTGGATGTGCTATTTATGCTAGACATCTTCCCGGA TTGGTCAATTTCTGTATTTTGTCTACATTTATGCTAGTTTTCGA CTTGCTTTTGTCTGCTACTTTTTATTCTGCCATTTTATCAATGA AGCTGGAAATTAACATCATTCACAGATCAACCGTCATCAGAC AGACTTTGGAAGAGGACGGAGTTGTCCCAACTACAGCAGAT ATTATATATAAGGATGAAACTGCCTCAGAACCACATTTTTTGA GATCTAACGTGGCTATCATTCTGGGAAAAGCATCAGTTATTG GTCTTTTGCTTCTGATCAACCTTTATGTTTTCACAGATAAGTT AAATGCTACAATACTAAACACGGTATATTTTGACTCTACAATT TACTCGTTACCAAATTTTATCAATTATAAAGATATTGGCAATCT CAGCAATCAAGTGATCATTTCCGTGTTGCCAAAGCAATATTAT ACTCCGCTGAAAAAATACCATCAGATCGAAGATTCTGTTCTA CTTATCATTGATTCCGTTAGCAATGCTATTCGGGACCAATTTA TCAGCAAGTTACTTTTTTTTGCATTTGCAGTTAGTATTTCCATC AATGTCTACTTACTGAATGCTGCAAAAATTCACACAGGATACA

		TGAACTTCCAACCACAATCAAATAAGATCGATGATCTTGTTGT TCAGCAAAAATCGGCAACGATTGAGTTTTCAGAAACTCGAAG TATGCCTGCTTCTTCTGGCCTAGAAACTCCAGTGACCGCGAA AGATATAATTATCTCTGAAGAAATCCAGAATAACGAATGCGTC TATGCTTTGAGTTCCCAGGACGAGCCTATCCGTCCTTTATCG AATTTAGTGGAACTTATGGAGAAAGAACAATTAAAGAACATGA ATAATACTGAGGTTTCGAATCTTGTCGTCAACGGTAAACTGC CATTATATTCCTTAGAGAAAAAATTAGAGGACACAACTCGTGC GGTTTTAGTTAGGAGAAAGGCACTTTCAACTTTGGCTGAATC GCCAATTTTAGTTTCCGAAAAATTGCCCTTCAGAAATTATGAT TATGATCGCGTTTTTGGAGCTTGCTGTGAAAATGTCATCGGC TATATGCCAATACCAGTTGGTGTAATTGGTCCATTAATTATTG ATGGAACATCTTATCACATACCAATGGCAACCACGGAAGGTT GTTTAGTGGCTTCAGCTATGCGTGGTTGCAAAGCCATCAATG CTGGTGGTGGTGCAACAACTGTTTTAACCAAAGATGGTATGA CTAGAGGCCCAGTCGTTCGTTTCCCTACTTTAATAAGATCTG GTGCCTGCAAGATATGGTTAGACTCGGAAGAGGGACAAAATT CAATTAAAAAAGCTTTTAATTCTACATCAAGGTTTGCACGTTT GCAACATATTCAAACCTGTCTAGCAGGCGATTTGCTTTTTATG AGATTTCGGACAACTACCGGTGACGCAATGGGTATGAACATG ATATCGAAAGGTGTCGAATACTCTTTGAAACAAATGGTAGAA GAATATGGTTGGGAAGATATGGAAGTTGTCTCCGTATCTGGT AACTATTGTACTGATAAGAAACCTGCCGCAATCAATTGGATT GAAGGTCGTGGTAAAAGTGTCGTAGCTGAAGCTACTATTCCT GGTGATGTCGTAAAAAGTGTTTTAAAGAGCGATGTTTCCGCT TTAGTTGAATTAAATATATCCAAGAACTTGGTTGGATCCGCAA TGGCTGGATCTGTTGGTGGTTTCAACGCGCACGCAGCTAATT TGGTCACTGCACTTTTCTTGGCATTAGGCCAAGATCCTGCGC AGAACGTCGAAAGTTCCAACTGTATAACTTTGATGAAGGAAG TTGATGGTGATTTAAGGATCTCTGTTTCCATGCCATCTATTGA AGTTGGTACGATTGGCGGGGGTACTGTTCTGGAGCCTCAGG GCGCCATGCTTGATCTTCTCGGCGTTCGTGGTCCTCACCCC ACTGAACCTGGAGCAAATGCTAGGCAATTAGCTAGAATAATC GCGTGTGCTGTCTTGGCTGGTGAACTGTCTCTGTGCTCCGC ACTTGCTGCCGGTCACCTGGTACAAAGCCATATGACTCACAA

		CCGTAAAACAAACAAAGCCAATGAACTGCCACAACCAAGTAA CAAAGGGCCCCCCTGTAAAACCTCAGCATTATTATAA
42	CcGGDPS1, *Cistus creticus*, polypeptide	Msymiqkansvnqaletavslreplkiheamrysvlaggkrvrpvlciaacelvggeesm ampaacavemihtmslihddppcmdnddlrrgkptnhkvfgediavlagdallafsfeh mtgstvgvsparivravgelaksigseglvagqvvdicsegltdvglehlefihvhktaalle gsvvlgailgggsdeeveklrkfarciglllfqvvddildvtksstelgktagkdlvadkvtypkl mgieksrefaeelkndaveqlrgfdqekaaplialanyiayrqn
43	CcGGDPS1, synthetic construct, nucleic acid	atgtcctacatgattcagaaggcgaattccgtgaatcaagccctagaaaccgccgtttcgc tccgggagccgctgaagatccacgaggcgatgaggtactcggttttggccggcgggaag cgagttcggccggtgctgtgcatcgccgcatgcgagctcgtcggcggtgaggaatccatg gcgatgcccgcggcttgcgccgtggaaatgattcacaccatgtcgctcatccacgacgac cctccctgtatggataacgacgatctccgccgcggaaaaccaacaaaccacaaggtatt cggcgaggatatcgcggtgttagccggggacgcgcttctcgctttctcattcgagcacatg acggggtccacagttggcgtatcccccgcccgaattgtgcgggcggttggggaattagcg aaatcgattggttcggaaggattggtagctggccaagttgtggatatctgcagtgagggcct tacagatgtaggcttggaacatttagagttcatccatgtgcataaaacggcggccttactcg aaggctcggtggttctgggagcaattcttgggggcggatcggatgaggaagtggagaaa ttgaggaaatttgcgaggtgcattggattattgtttcaggtagtggatgatattctggatgtgac gaaatcttcaacagaattggggaaaacggctgggaaagatctggtggcggataaggtg acatatccgaagctaatggggatcgagaaatcgagggagtttgcggaggagctgaaga atgatgcggtggagcagttgagagggtttgatcaggagaaggcggcgcctttgattgcttt ggctaattacattgcttataggcagaat
44	CPR2, synthetic construct, nucleic acid	atgcaatcatcaagcagctcgatgaaagtgtcaccacttgaacttatgcaagccataatca aaggcaaagtggacccaacaaatgtttcatcggaatccggtggttctgctgctgagatggc aactttgatccgcgagaatcgtgagtttgttattatcttaactacttccatagcggttttgatcgg ctacgttgtcgttttaatttggagaagatcatccggctatcagaaacctaaagtccctgtccct cctaagccgttgattgttaaagacctcgaacctgaagttgatgatggcaagaaaaaggtc accatctttttcggcacccaaactggtactgctgaaggatttgctaaggctctagctgagga ggcaaaagctcggtatgagaaggctatatttaaaactgttgatttggatgattatgcggagg atgacgatgaatacgaagagaaattgaagaaagagtctctggccattttcttcttggccac atatggagatggtgagcctacagataacgccgcgaggttttataaatggtttacagatggc aatgagaggggggaatggcttaaggaacttccatatgctgtttttggtcttggcaacaggca atacgagcattttaataagattgccatagtggtggataaaatccttggcaaccagggtggg aagcagcttgttccagtgggtcttggtgatgatgatcaatgcatggaagatgactttgccgc atggcgagaattgttgtggcctgagttggaccagttgcttcttgatggggatgatccaactgg

		tgtttctacccttatactgctgccgtggcagaatatcgggttgtattgcatgaccctgaagat gcaccattagaggatgataactggagtaatgcgaatggtcatgctatttatgatgctcagca tccatgcagggctaatgttactgtgaggagggagcttatacccctgcatctgatcgttcatg tacccatttggagttcgacatatctggcactggacttgtatatggaactggtgatcatgttggt gtgtactgtgaaaatctaagtgaaattgttgaggaagcactgcagttgttgggtttatcgcca gatatttacttcactatccatactgataatgaggatggcacaccacttagtggaagtgccttg ccacctccattcccatcgtccaccttaagaacagctctaactcgatatgctgatcttttgagtt cacccaaaaagtctgctttaatggctttagcagctcatgctactaatccaaccgaagctgat cggctaagacatcttgcatcacctgctggaaaggatgaatatgcacaatggatagttgca aatcatagaagcctcctggaagtcatggctgaatttccatcagccaaaccccccacttggag tcttctttgcttcagttgccccgcgattgctgccaagatactattctatttcatcatctccaagcat ggcaccttcaaggattcatgttacatgtgcactggttcttgagaaaacaccagcaggtcga attcacaaaggagtgtgctcaacttggatgaagaatgctgtgcctttagagaaaagccatg attgcagctgggcacctatttttgttagacaatcaaacttcaaacttccagcagatactaaag ttcccatcattatgattggccctggaactggtttagctcctttcaggggtttccttcaggaaaga ttagcccagaaagaagcaggagcagaactgggatcctctgtattattctttggttgcagga accgtcaaatggattttatctatgaagatgagctcaacaatttcgttgaaagtggtgcactttc tgaactatctgtagccttctcacgtgagggacctaccaaggaatatgtgcagcataagatg atgcagaaggcttctgatatctggaacatgatttctcaaggaggatatttatatgtttgtggag atgccaagggcatggctaaagatgtccacagaactctccacactatcgtgcaagagcag ggatctcttgacaactccaagacagagagctttgtgaagggtctgcaaatgaatggcagg tatctgcgtgatgtatggtaa
45	TAT, *Taxus cuspidata*, Accession: Q9M6F0.1, polypeptide	MEKTDLHVNLIEKVMVGPSPPLPKTTLQLSSIDNLPGVRGSIFNA LLIYNASPSPTMISADPAKPIREALAKILVYYPPFAGRLRETENGD LEVECTGEGAMFLEAMADNELSVLGDFDDSNPSFQQLLFSLPL DTNFKDLSLLVVQVTRFTCGGFVVGVSFHHGVCDGRGAAQFLK GLAEMARGEVKLSLEPIWNRELVKLDDPKYLQFFHFEFLRAPSI VEKIVQTYFIIDFETINYIKQSVMEECKEFCSSFEVASAMTWIART RAFQIPESEYVKILFGMDMRNSFNPPLPSGYYGNSIGTACAVDN VQDLLSGSLLRAIMIIKKSKVSLNDNFKSRAVVKPSELDVNMNHE NVVAFADWSRLGFDEVDFGWGNAVSVSPVQQQSALAMQNYF LFLKPSKNKPDGIKILMFLPLSKMKSFKIEMEAMMKKYVAKV*
46	TAX19, *Taxus chinensis*,	MENSVWKDLNVKSFDPVMVKPSIPLPKTVLHLSTVDNLPVLRG NLFNSLIVYKASDKISADPVKVIREALSKVLVYYFPFAGRLRYKE NGDLEVDCNGEGAAFVEAMVDCNLSVLGDLDDLNPSYEDLLFA

	Accession: KAH9291366.1, polypeptide	LPQNTDIVDLHLLVVQVTCFACGGFVVGVSFHHSICDGRGAGQ FLQSLAEIARGEDKLSCEPIWNRELLKPEDPIHLQFYHLYSLRPS GPTFEEWVHASLVINPATIKHMKQSIMEECNKVCSSFEIVAALG WRARTKALQIPQTQNVKLLFAVDMRKSFNPPFPKGYYGNAIGF ACAMDNAHDLINGSLLHAVNIIRKAKAYLFEQCSKSSVAVNPSAL DANTGQESVVALTDWRRLGFNEVNFGWGDAVNVCPVQRMTN GLVMPNYFLFLRPSEDMPDGIKILMCMASSMVKSFKFEVEDMIN KYVTAV*
47	TAT, synthetic construct, nucleic acid	ATGGAAAAGACCGACCTGCACGTGAACCTGATCGAGAAGGT GATGGTGGGCCCATCTCCGCCGCTGCCAAAGACTACTCTGC AGCTGAGCAGCATCGACAACCTGCCAGGCGTTAGGGGCAGC ATCTTCAACGCCCTGCTGATCTACAACGCCTCTCCATCTCCG ACCATGATCAGCGCCGATCCGGCTAAGCCAATCAGAGAGGC CCTGGCCAAGATCCTGGTGTACTACCCACCATTCGCTGGCA GGCTGCGCGAGACTGAGAATGGCGATCTTGAGGTTGAGTGC ACCGGCGAGGGCGCTATGTTCCTTGAGGCTATGGCTGACAA CGAGCTGAGCGTGCTGGGCGACTTCGACGACAGCAACCCAT CTTTCCAGCAGCTCCTCTTCAGCCTGCCGCTGGACACCAACT TCAAGGACCTGAGCCTGCTGGTGGTGCAGGTCACCAGATTC ACTTGCGGCGGCTTCGTGGTGGGCGTGTCATTCCATCATGG CGTTTGCGACGGTAGGGGCGCTGCCCAATTCCTTAAGGGCC TTGCTGAAATGGCTCGCGGCGAGGTGAAGCTGTCTCTCGAG CCAATTTGGAACCGCGAGCTGGTGAAGCTCGACGACCCAAA GTACCTGCAGTTCTTCCACTTCGAGTTCCTCAGGGCCCCGA GCATCGTCGAGAAGATCGTGCAGACCTACTTCATCATCGACT TCGAGACTATCAACTACATCAAGCAGAGCGTGATGGAAGAGT GCAAAGAGTTCTGCAGCAGCTTCGAGGTGGCCAGCGCTATG ACTTGGATCGCTAGGACCAGGGCCTTTCAGATCCCAGAGTC CGAGTACGTGAAGATCCTGTTCGGCATGGACATGAGGAACA GCTTCAACCCGCCACTGCCGTCTGGCTACTACGGCAACTCT ATCGGCACTGCCTGCGCCGTGGATAACGTGCAGGATCTGCT TAGCGGCAGCCTGCTGAGGGCCATCATGATCATCAAGAAGT CCAAGGTCAGCCTGAACGATAACTTCAAGAGCAGGGCCGTC GTGAAGCCGTCTGAGCTGGACGTGAACATGAACCACGAGAA CGTGGTGGCCTTCGCCGACTGGTCTAGGCTTGGCTTCGATG

		AGGTTGACTTCGGCTGGGGCAACGCTGTGTCTGTGTCTCCA GTTCAGCAGCAGAGCGCCCTGGCTATGCAGAATTACTTCCT GTTCCTGAAGCCGAGCAAGAACAAGCCGGACGGGATCAAGA TCCTCATGTTCCTGCCACTGAGCAAGATGAAGTCCTTCAAGA TCGAGATGGAAGCCATGATGAAGAAATACGTGGCCAAGGTG TGA
48	TAX19, synthetic construct, nucleic acid, Accession: JAHRHJ020003660.1	ATGGAGAATTCAGTGTGGAAAGACTTAAATGTAAAGAGCTTT GATCCAGTCATGGTGAAGCCCTCCATTCCTCTGCCCAAAACT GTGCTCCATCTCTCCACTGTGGACAACCTTCCAGTGTTAAGG GGAAATCTTTTTAACTCCTTAATTGTCTACAAAGCCTCTGACA AAATTTCTGCAGATCCTGTGAAAGTAATTCGGGAGGCTCTTT CCAAGGTGTTGGTGTATTATTTCCCATTTGCTGGACGGCTCA GATACAAAGAAAATGGGGACCTTGAAGTGGACTGCAATGGG GAGGGTGCTGCCTTTGTGGAAGCCATGGTGGACTGCAACCT TTCTGTGTTGGGAGATTTGGATGACCTCAATCCATCATATGA AGACTTGCTTTTTGCTCTTCCTCAGAATACAGACATTGTGGAC CTTCATCTTCTGGTTGTTCAGGTAACATGTTTTGCATGTGGG GGTTTTGTTGTGGGGGTGAGTTTCCACCATAGTATATGCGAT GGACGAGGAGCTGGTCAATTTCTGCAAAGCCTTGCAGAGAT AGCGAGAGGAGAAGATAAGCTATCATGTGAACCAATATGGAA CAGAGAACTGCTGAAGCCTGAAGATCCTATACACCTCCAATT TTATCACTTGTATTCGCTACGCCCTTCTGGCCCTACATTTGAG GAATGGGTCCATGCCTCTCTTGTTATAAACCCTGCGACAATA AAACACATGAAACAGTCTATTATGGAAGAATGTAATAAAGTTT GCTCTTCATTCGAAATTGTGGCAGCATTAGGTTGGCGAGCGA GGACAAAAGCTCTTCAAATCCCACAAACTCAGAATGTGAAGC TTCTGTTTGCGGTGGACATGAGGAAATCATTTAATCCCCCGT TTCCAAAAGGATACTATGGTAATGCCATTGGTTTTGCATGTG CAATGGATAATGCACATGATCTCATAAATGGGTCTCTTTTGCA TGCTGTAAATATTATAAGGAAAGCAAAGGCTTATTTATTCGAA CAGTGTTCAAAGTCAAGCGTCGCGGTTAACCCATCTGCATTA GATGCGAACACAGGACAAGAAAGTGTAGTTGCATTGACTGAT TGGAGGCGACTGGGATTTAATGAAGTGAACTTTGGGTGGGG AGATGCAGTGAATGTATGTCCTGTGCAACGGATGACAAATGG ACTAGTTATGCCAAACTATTTTCTATTCCTCCGACCCTCCGAG

		GACATGCCTGATGGAATCAAGATACTAATGTGCATGGCCTCA TCAATGGTGAAATCATTCAAATTTGAAGTGGAAGACATGATAA ACAAATATGTGACTGCAGTATGA

References

Cragg, G. M. Paclitaxel (Taxol): a success story with valuable lessons for natural product drug discovery and development. *Medicinal research reviews* 18, 315-331 (1998).

Guerra-Bubb, J., et al. The early stages of taxol biosynthesis: an interim report on the synthesis and identification of early pathway metabolites. *Natural product reports* 29, 683-696, doi:10.1039/c2np20021j (2012)

Forman, V., et al. A gene cluster in Ginkgo biloba encodes unique multifunctional cytochrome P450s that initiate ginkgolide biosynthesis. *Nat. Commun.* **13**, 5143 (2022).

Ignea et al. Efficient diterpene production in yeast by engineering Erg20p into a geranylgeranyl diphosphate synthase. *Metabolic Engineering* Jan;27:65-75. doi: 10.1016/j.ymben.2014.10.008 (2015)

Items

1. A method of producing a taxane comprising an oxetane, said method comprising the steps of;
   i. providing a host cell comprising a heterologous nucleic acid encoding CYP1 of SEQ ID NO: 3 or functional homologues thereof having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto, and
   ii. incubating said host cell in presence of taxadiene,
   thereby producing the taxane comprising an oxetane, or
   a method of producing a 13-hydroxy-taxane, a 11-hydroxy-taxane, and/or a 10-hydroxy-taxane, said method comprising the steps of:
   a. providing a host cell comprising a heterologous nucleic acid encoding CYP3 of SEQ ID NO: 7 or functional homologues thereof having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto, and
   b. incubating said host cell in presence of taxadiene,
   thereby producing the 13-hydroxy-taxane, the 11-hydroxy-taxane, and/or the 10-hydroxy-taxane.

2. A method of producing 4-hydroxy-5,20-epoxy-taxane, said method comprising the steps of;
   i. providing a host cell comprising a heterologous nucleic acid encoding CYP1 of SEQ ID NO: 3 or functional homologues thereof having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto, and
   ii. incubating said host cell in presence of taxadiene,
   thereby producing 4-hydroxy-5,20-epoxy-taxane.

3. A method of producing 10-hydroxy-taxadiene, 11-hydroxyl-4,12-taxadiene, and/or 13-hydroxy-taxadiene, said method comprising the steps of;
   i. providing a host cell comprising a heterologous nucleic acid encoding CYP3 of SEQ ID NO: 7 or a functional homologue thereof having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto, and
   ii. incubating said host cell in presence of taxadiene,

thereby producing 10-hydroxy-taxadiene, 11-hydroxyl-4,12-taxadiene, and/or 13-hydroxy-taxadiene.

4. A method of producing 4-hydroxy-10-hydroxy-5,20-epoxy-taxane, 10-hydroxy-5,20-epoxy-taxane, 11-hydroxy-5,20-epoxy-taxane, 4-hydroxy-11-hydroxy-5,20-epoxy-taxane, 4-hydroxy-12-hydroxy-5,20-epoxy-12,13-taxane, 13-hydroxy-5,20-epoxy-taxane, and/or 4-hydroxy-13-hydroxy-5,20-epoxy-taxane, said method comprising the steps of:
    i. providing a host cell comprising a heterologous nucleic acid encoding CYP1 of SEQ ID NO: 3 and a heterologous nucleic acid encoding CYP3 of SEQ ID NO: 7, or functional homologues of any of the aforementioned having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto; and
    ii. incubating said host cell in presence of taxadiene,

    thereby producing said 4-hydroxy-10-hydroxy-5,20-epoxy-taxane, 10-hydroxy-5,20-epoxy-taxane, 11-hydroxy-5,20-epoxy-taxane, 4-hydroxy-11-hydroxy-5,20-epoxy-taxane, 4-hydroxy-12-hydroxy-5,20-epoxy-12,13-taxane, 13-hydroxy-5,20-epoxy-taxane, and/or 4-hydroxy-13-hydroxy-5,20-epoxy-taxane.

5. A method of producing an *O*-acetylated taxane comprising an oxetane, said method comprising the steps of:
    i. providing a host cell comprising a heterologous nucleic acid encoding CYP1 of SEQ ID NO: 3 and a heterologous nucleic acid encoding TAX19 of SEQ ID NO: 46, or functional homologues of any of the aforementioned having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto; and
    ii. incubating said host cell in presence of taxadiene,

    thereby producing the *O*-acetylated taxane comprising an oxetane.

6. A method of producing 4-acetoxy-5,20-epoxy-taxane, said method comprising the steps of:

i. providing a host cell comprising a heterologous nucleic acid encoding CYP1 of SEQ ID NO: 3 and a heterologous nucleic acid encoding TAX19 of SEQ ID NO: 46, or functional homologues of any of the aforementioned having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto; and

ii. incubating said host cell in presence of taxadiene,

thereby producing 4-acetoxy-5,20-epoxy-taxane.

7. The method according to item 1 for production of the taxane comprising an oxetane, wherein the host cell further comprises CYP3 as set forth in SEQ ID NO: 7, or functional homologues thereof having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto.

8. The method according to any one of the preceding items, further comprising a step of supplying said taxadiene to the host cell, such as incubating said host cell in a cultivation medium comprising taxadiene.

9. The method according to any one of the preceding items, further comprising a step of isolating the taxane comprising an oxetane, *O*-acetylated taxane comprising an oxetane, 10-hydroxy-taxane, 11-hydroxy-taxane and/or said 13-hydroxy-taxane.

10. The method according to item 9, wherein the method comprises the step of isolating the taxane comprising an oxetane, *O*-acetylated taxane comprising an oxetane, 10-hydroxy-taxane, 11-hydroxy-taxane and/or said 13-hydroxy-taxane, wherein said step comprises extraction, precipitation and/or chromatography.

11. The method according to any one of items 9 to 10, wherein the method of isolating the taxane comprising an oxetane, *O*-acetylated taxane comprising an oxetane, 10-hydroxy-taxane, 11-hydroxy-taxane and/or said 13-hydroxy-taxane comprises a step of extraction, such as extraction with a solvent, for example methanol (MeOH).

12. The method according to any one of items 9 to 11, wherein the method of isolating the taxane comprising an oxetane, *O*-acetylated taxane comprising an oxetane, 10-hydroxy-taxane, 11-hydroxy-taxane and/or said 13-hydroxy-taxane comprises a step of chromatography, such as liquid chromatography (LC), for example column chromatography or preparative/semi-preparative high performance LC (HPLC).

13. The method according to any one of the preceding items, wherein the host cell is a host cell according to any one of items 14 to 32, 40 or 56 to 70.

14. A host cell capable of producing taxane comprising an oxetane, in the presence of taxadiene, said host cell comprising a heterologous nucleic acid encoding CYP1 as set forth in SEQ ID NO: 3, or
a host cell capable of producing 10-hydroxy-taxane, 11-hydroxy-taxane and/or said 13-hydroxy-taxane, in the presence of taxadiene, said host cell comprising a heterologous nucleic acid encoding CYP3 as set forth in SEQ ID NO: 7.

15. A host cell capable of producing taxane comprising an oxetane and optionally 10-hydroxy-taxane, 11-hydroxy-taxane, 13-hydroxy-taxane and/or said 13-hydroxy-taxane, in the presence of taxadiene, said host cell comprising a heterologous nucleic acid encoding CYP1 as set forth in SEQ ID NO: 3 and a heterologous nucleic acid encoding CYP3 as set forth in SEQ ID NO: 7, or functional homologues of any of the aforementioned having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto, or
a host cell capable of producing 10-hydroxy-taxane, 11-hydroxy-taxane, 13-hydroxy-taxane and/or said 13-hydroxy-taxane, in the presence of taxadiene, said host cell comprising a heterologous nucleic acid encoding CYP3 as set forth in SEQ ID NO: 7 and a heterologous nucleic acid encoding a taxadiene synthase capable of catalysing production of taxadiene, such as TXS as set forth in SEQ ID NO: 1 and/or TXS _3xSG_MBP-SKL as set forth in SEQ ID NO: 2, or functional homologues of any of the aforementioned having at least

70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto.

16. A host cell comprising a heterologous nucleic acid encoding TAX19 as set forth in SEQ ID NO: 46, or functional homologues thereof having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto.

17. A host cell capable of producing an *O*-acetylated taxane comprising an oxetane, wherein the host cell comprises a heterologous nucleic acid encoding an acetyltransferase, such as TAX19 as set forth in SEQ ID NO: 46, or functional homologues thereof having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto.

18. The method and/or host cell according to any one of the preceding items, wherein said host cell further comprises a heterologous nucleic acid encoding taxadiene synthase (EC 4.2.3.17) capable of converting geranylgeranyl diphosphate (GGPP) into taxadiene, preferably said taxadiene synthase is TXS as set forth in SEQ ID NO: 1 and/or TXS _3xSG_MBP-SKL as set forth in SEQ ID NO: 2 or a functional homologue of any of the aforementioned having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto.

19. The method and/or host cell according to any one of the preceding items, wherein said host cell further comprises a nucleic acid encoding cytochrome P450 reductase (CPR), such as an endogenous and/or a heterologous CPR, for example a CPR from a *Taxus*, such as *Taxus cuspidata*, a CPR from *Nicotiana*, such as *Nicotiana benthamiana*, a CPR from *Arabidopsis*, such as *A. thaliana*, and/or a CPR from *Populus*, such as *Populus trichocarpa* × *Populus deltoides*.

20. The method and/or host cell according to item 19, wherein the CPR is TcuCPR as set forth in SEQ ID NO: 4 and/or CPR2 as set forth in SEQ ID NO: 6, or functional homologues of any of the aforementioned having at least

70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto.

21. The method and/or host cell according to any one of the preceding items, wherein the host cell further comprises a nucleic acid encoding an acetyltransferase, such as an acetyltransferase from a *Taxus* cell, for example a *Taxus chinensis* cell.

22. The method and/or host cell according to item 21, wherein the acetyltransferase is TAX19 as set forth in SEQ ID NO: 46, or functional homologues thereof having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto.

23. The method and/or host cell according to any one of the preceding items, wherein the host cell is further capable of producing an *O*-acetylated taxane comprising an oxetane.

24. The method and/or host cell according to any one of the preceding items, wherein the host cell is selected from the group of plant cells, yeast cells, bacterial cells and fungal cells.

25. The method and/or host cell according to any one of the preceding items, wherein the host cell is plant cells, such as plant cells comprised within a plant or within a part of a plant.

26. The method and/or host cell according to item 25, wherein the plant cells are from a species of *Nicotiana*, such as *Nicotiana benthamiana* or *Nicotiana tabacum*.

27. The method and/or host cell according to any one of the preceding items, wherein the host cell is a yeast cell belonging to the genus of *Saccharomyces*, *Pichia*, *Candida, Cryptococcus, Pichia (Komagataella), Lipomyces, Pseudozyma, Rhodosporidium, Rhodotorula, Trichosporon, Trigonopsis, Yarrowia* or *Saccharomycopsis,* such as a yeast cell of the

species *Saccharomyces cerevisiae Yarrowia lipolytica, Hansenula polymorpha (Ogataea polymorpha), Rhodotorula toruloides* or *Pichia pastoris* (*Komagataella phaffii*).

28. The method and/or host cell according to any one of the preceding items, wherein the host cell is a bacterial cell belonging to the genus of *Escherichia*, *Bacillus, Corynebacterium, Pseudomonas* or *Streptomyces*, such as a bacterial cell of the species *Escherichia coli, Bacillus subtilis, Corynebacterium glutamicum, Pseudomonas putida* or *Streptomyces sp*.

29. The method and/or host cell according to any one of the preceding items, wherein the host cell is capable of production of GGPP.

30. The method and/or host cell according to any one of items 27 to 29, wherein the host cell further comprises a heterologous nucleic acid encoding ERG20(F96C)-L as set forth in SEQ ID NO: 5 and/or ERG20(F96C), or a functional homologue of any of the aforementioned having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto.

31. The method and/or host cell according to any one of items 27 to 30, wherein the host cell further comprises a heterologous nucleic acid encoding HMG reductase, such as an endogenous and/or a heterologous HMG reductase, for example a HMG reductase from *Saccharomyces cerevisiae*, for example HMG2(K6R) as set forth in SEQ ID NO: 20, or a functional homologue thereto having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto.

32. The method and/or host cell according to any one of the preceding items, wherein the taxane comprising an oxetane, *O*-acetylated taxane comprising an oxetane, 10-hydroxy-taxane, 11-hydroxy-taxane and/or 13-hydroxy-taxane is produced with a titer of at least 50 µg/L, such as at least 0.75 µg/L, for example at least 100 µg/L, such as at least 250 µg/L, for example 500 µg/L, such as at least 750 µg/L, for example at least 900 µg/L, such as at least

1000 µg/L, for example at least 2.5 mg/L, such as at least 5 mg/L, for example at least 7.5 mg/L, for example at least 10 mg/L, or more.

33. A nucleic acid construct for expression in a host cell, comprising:

i. a nucleic acid encoding CYP1 as set forth in SEQ ID NO: 3 and/or a functional variant thereof having at least 70% sequence identity thereto, such as SEQ ID NO: 10 and/or SEQ ID NO: 11; and/or

ii. a nucleic acid encoding CYP3 as set forth in SEQ ID NO: 7, such as SEQ ID NO: 15; and optionally

iii. a nucleic acid encoding TXS as set forth in SEQ ID NO: 1 and/or TXS _3xSG_MBP-SKL as set forth in SEQ ID NO: 2, such as SEQ ID NO: 8 and/or SEQ ID NO: 9, respectively;

and further optionally,

iv. a nucleic acid encoding ERG20(F96C)-L as set forth in SEQ ID NO: 5, such as SEQ ID NO: 13, CcGGDPS1 as set forth in SEQ ID NO: 42, such as SEQ ID NO: 43, ERG20(F96C), and/or ERG20(Y95A),

v. a nucleic acid encoding TcuCPR as set forth in SEQ ID NO: 4, such as SEQ ID NO: 12, CPR2 as set forth in SEQ ID NO: 6, such as SEQ ID NO: 44, and/or AtCPR (NCBI GenBank accession number: NP_194183, 4 June 2023),

vi. a nucleic acid encoding *S. cerevisiae* ERG20 (NCBI GenBank accession number: NM_001181600, 3 June 2023), and/or SfFDPS1 as set forth in SEQ ID NO: 39, such as SEQ ID NO: 14,

vii. a nucleic acid encoding SpCytb5 as set forth in SEQ ID NO: 40, such as SEQ ID NO: 36,

viii. a nucleic acid encoding SpCBR as set forth in SEQ ID NO: 38, such as SEQ ID NO: 37,

ix. a nucleic acid encoding HMG2(K6R) as set forth in SEQ ID NO: 20, such as SEQ ID NO: 41,

or functional homologues of any of the aforementioned having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto.

34. A nucleic acid construct for expression in a host cell, comprising:

i. a nucleic acid encoding CYP1 as set forth in SEQ ID NO: 3 and/or a functional variant thereof having at least 70% sequence identity thereto, such as SEQ ID NO: 10 and/or SEQ ID NO: 11, and a nucleic acid encoding CYP3 as set forth in SEQ ID NO: 7, such as SEQ ID NO: 15; or

ii. a nucleic acid encoding CYP1 as set forth in SEQ ID NO: 3 and/or a functional variant thereof having at least 70% sequence identity thereto, such as SEQ ID NO: 10 and/or SEQ ID NO: 11, and/or a nucleic acid encoding CYP3 as set forth in SEQ ID NO: 7, such as SEQ ID NO: 15; and one or more of the following:

a. a nucleic acid encoding TXS as set forth in SEQ ID NO: 1 and/or TXS _3xSG_MBP_SKL as set forth in SEQ ID NO: 2, such as SEQ ID NO: 8 and/or SEQ ID NO: 9, respectively,

b. a nucleic acid encoding ERG20(F96C)-L as set forth in SEQ ID NO: 5, such as SEQ ID NO: 13, CcGGDPS1 as set forth in SEQ ID NO: 42, such as SEQ ID NO: 43, ERG20(F96C), and/or ERG20(Y95A),

c. a nucleic acid encoding TcuCPR as set forth in SEQ ID NO: 4, such as SEQ ID NO: 12, CPR2 as set forth in SEQ ID NO: 6, such as SEQ ID NO: 44, and/or AtCPR (NCBI GenBank accession number: NP_194183, 4 June 2023),

d. a nucleic acid encoding *S. cerevisiae* ERG20 (NCBI GenBank accession number: NM_001181600, 3 June 2023), and/or SfFDPS1 as set forth in SEQ ID NO: 39, such as SEQ ID NO: 14,

e. a nucleic acid encoding SpCytb5 as set forth in SEQ ID NO: 40, such as SEQ ID NO: 36,

f. a nucleic acid encoding SpCBR as set forth in SEQ ID NO: 38, such as SEQ ID NO: 37,

g. a nucleic acid encoding HMG2(K6R) as set forth in SEQ ID NO: 20, such as SEQ ID NO: 41,

or functional homologues of any of the aforementioned having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto.

35. The nucleic acid construct according to any one of items 33 to 34, further comprising a promotor, such as an inducible promoter, operably linked to any one or more of the nucleic acid sequences.

36. The nucleic acid construct according to any one of items 33 to 35, further comprising a nucleic acid encoding TAX19 as set forth in SEQ ID NO: 46 or a functional homologue thereof having at least 70% sequence identity thereto, such as SEQ ID NO: 48 or a functional having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto.

37. A polypeptide selected from SEQ ID NO: 3 and SEQ ID NO: 7.

38. A nucleic acid encoding any one of the polypeptides according to item 33, item 34 or item 37.

39. A vector comprising at least one of the nucleic acid constructs according to any one of items 33, 34, 35, 36 or 38.

40. A host cell according to any one of items 14 to 32 or 56 to 70, comprising the nucleic acid construct according to any one of items 33 to 36, a nucleic acid according to item 38 or a vector according to item 39.

41. A kit of parts comprising:

i. the host cell according to item 40, and optionally instructions for use, and/or
ii. the nucleic acid construct according to any one of items 33 to 36 and instructions for use, and optionally a host cell to be modified,

preferably wherein the host cell is a microorganism or a plant cell.

42. 4-hydroxy-5,20-epoxy-taxane obtained by a method according to any one of items 1 to 13, 18 to 32, or 56 to 70.

43. 10-hydroxy-taxadiene obtained by a method according to any one of items 1 to 13, 18 to 32, or 56 to 70.

44. 11-hydroxy-taxane is 11-hydroxyl-4,12-taxadiene obtained by a method according to any one of items 1 to 13, 18 to 32, or 56 to 70.

45. 13-hydroxy-taxadiene obtained by a method according to any one of items 1 to 13, 18 to 32, or 56 to 70.

46. A purified compound of the formula 4-acetoxy-5,20-epoxy-taxane.

47. A cell culture obtained by a method according to any one of items 1 to 13, 18 to 32, or 56 to 70.

48. A cell culture, comprising a host cell according to any one of items 14 to 32, 40, or 56 to 70, and optionally a cultivation medium.

49. A fermentation liquid comprising the taxane comprising an oxetane, *O*-acetylated taxane comprising an oxetane, 10-hydroxy-taxane, 11-hydroxy-taxane, 13-hydroxy-taxane and/or said 13-hydroxy-taxane, wherein said fermentation liquid is:

i. obtained by the method of any one of items 1 to 13, 18 to 32, or 56 to 70;
ii. comprised in the cell culture according to any one of items 47 to 48, and/or
iii. comprised within and/or secreted by the host cell according to any one of items 14 to 32, 40, or 56 to 70.

50. The fermentation liquid of item 49, wherein at least 50%, such as at least 75%, such as at least 95%, such as at least 99% of the host cells are lysed.

51. The fermentation liquid according to any one of items 49 to 50, wherein at least 50%, such as at least 75%, such as at least 95%, such as at least 99% of solid cellular material has been separated from the liquid.

52. A composition comprising:

i. the fermentation liquid according to any one of items 49 to 51;

ii. a taxane comprising an oxetane obtained by the method of any one of items 1 to 13, 18 to 32, or 56 to 70;
iii. an *O*-acetylated taxane comprising an oxetane obtained by the method of any one of items 1 to 13, 18 to 32, or 56 to 70;
iv. a 10-hydroxy-taxane obtained by the method of any one of items 1 to 13, 18 to 32, or 56 to 70;
v. a 11-hydroxy-taxane obtained by the method of any one of items 1 to 13, 18 to 32, or 56 to 70; and/or
vi. a 13-hydroxy-taxane obtained by the method of any one of items 1 13, 18 to 32, or 56 to 70,

and optionally one or more agents, additives and/or excipients.

53. The composition of item 52, wherein the composition, the fermentation liquid and/or taxane comprising an oxetane, an *O*-acetylated taxane comprising an oxetane, 10-hydroxy-taxane, 11-hydroxy-taxane and/or said 13-hydroxy-taxane have been processed into in a semi-dry or dry solid form, optionally in form of a powder, tablet, capsule, chewable, gel and/or gum.

54. The composition of item 52 or 53, wherein the composition, the fermentation liquid and/or taxane comprising an oxetane, an *O*-acetylated taxane comprising an oxetane, 10-hydroxy-taxane, 11-hydroxy-taxane and/or said 13-hydroxy-taxane is in a liquid form, optionally in a stabilized liquid form.

55. A method for treating a disorder such as cancer, comprising administration of a therapeutic sufficient amount of a taxane comprising oxetane, an *O*-acetylated taxane comprising an oxetane, a 10-hydroxy-taxane, a 11-hydroxy-taxane and/or a 13-hydroxy-taxane obtained by the method of any one of items 1 to 13, 18 to 32, or 56 to 70.

56. The composition, the fermentation liquid, the method for treating a disorder, the method and/or host cell according to any one of the preceding items, wherein the taxane comprising an oxetane is a 5,20-epoxy-taxane.

57. The composition, the fermentation liquid, the method for treating a disorder, the method and/or host cell according to any one of the preceding items,

wherein the taxane comprising an oxetane is 4-hydroxy-5,20-epoxy-taxane or a derivative thereof.

58. The composition, the fermentation liquid, the method for treating a disorder, the method and/or host cell according to any one of the preceding items, wherein the taxane comprising an oxetane is paclitaxel (taxol).

59. The composition, the fermentation liquid, the method for treating a disorder, the method and/or host cell according to any one of the preceding items, wherein the 10-hydroxy-taxane is 10-hydroxy-taxadiene or a derivative thereof.

60. The composition, the fermentation liquid, the method for treating a disorder, the method and/or host cell according to any one of the preceding items, wherein the 11-hydroxy-taxane is 11-hydroxyl-4,12-taxadiene or a derivative thereof.

61. The composition, the fermentation liquid, the method for treating a disorder, the method and/or host cell according to any one of the preceding items, wherein the 13-hydroxy-taxane is 13-hydroxy-taxadiene or a derivative thereof.

62. The composition, the fermentation liquid, the method for treating a disorder, the method and/or host cell according to any one of the preceding items, wherein the taxane comprising an oxetane is a 10-hydroxy-5,20-epoxy-taxane.

63. The composition, the fermentation liquid, the method for treating a disorder, the method and/or host cell according to any one of the preceding items, wherein the taxane comprising an oxetane is a 11-hydroxy-5,20-epoxy-taxane.

64. The composition, the fermentation liquid, the method for treating a disorder, the method and/or host cell according to any one of the preceding items,

wherein the taxane comprising an oxetane is a 13-hydroxy-5,20-epoxy-taxane.

65. The composition, the fermentation liquid, the method for treating a disorder, the method and/or host cell according to any one of the preceding items, wherein the taxane comprising an oxetane is 10-deacetyl-baccatin III.

66. The composition, the fermentation liquid, the method for treating a disorder, the method and/or host cell according to any one of the preceding items, wherein the taxane comprising an oxetane is baccatin III.

67. The composition, the fermentation liquid, the method for treating a disorder, the method and/or host cell according to any one of the preceding items, wherein the taxadiene is endotaxadiene.

68. The composition, the fermentation liquid, the method for treating a disorder, the method and/or host cell according to any one of the preceding items, wherein the taxadiene is exotaxadiene.

69. The composition, the fermentation liquid, the method for treating a disorder, the method and/or the host cell according to any one of the preceding items, wherein said *O*-acetylated taxane comprising an oxetane is an *O*-acetylated 5,20-epoxy-taxane.

70. The composition, the fermentation liquid, the method for treating a disorder, the method and/or the host cell according to any one of the preceding items, wherein said *O*-acetylated taxane comprising an oxetane is 4-acetoxy-5,20-epoxy-taxane.

Claims

1. A method of producing a taxane comprising an oxetane, said method comprising the steps of;
   i. providing a host cell comprising a heterologous nucleic acid encoding CYP1 of SEQ ID NO: 3 or functional homologues thereof having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto, and
   ii. incubating said host cell in presence of taxadiene,
   thereby producing the taxane comprising an oxetane, or
   a method of producing a 13-hydroxy-taxane, a 11-hydroxy-taxane, and/or a 10-hydroxy-taxane, said method comprising the steps of:
   a. providing a host cell comprising a heterologous nucleic acid encoding CYP3 of SEQ ID NO: 7 or functional homologues thereof having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto, and
   b. incubating said host cell in presence of taxadiene,
   thereby producing the 13-hydroxy-taxane, the 11-hydroxy-taxane, and/or the 10-hydroxy-taxane.

2. The method according to claim 1, wherein the host cell comprises both a heterologous nucleic acid encoding CYP1 of SEQ ID NO: 3 and a heterologous nucleic acid encoding CYP3 as set forth in SEQ ID NO: 7, or functional homologues of the aforementioned having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto.

3. A host cell capable of producing taxane comprising an oxetane and optionally 10-hydroxy-taxane, 11-hydroxy-taxane and/or 13-hydroxy-taxane, in the presence of taxadiene, said host cell comprising a heterologous nucleic acid encoding CYP1 as set forth in SEQ ID NO: 3 and a heterologous nucleic acid encoding CYP3 as set forth in SEQ ID NO: 7, or functional homologues of any of the aforementioned having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto, or

a host cell capable of producing 10-hydroxy-taxane, 11-hydroxy-taxane and/or said 13-hydroxy-taxane, in the presence of taxadiene, said host cell comprising a heterologous nucleic acid encoding CYP3 as set forth in SEQ ID NO: 7 and a heterologous nucleic acid encoding a taxadiene synthase capable of catalysing production of taxadiene, such as TXS as set forth in SEQ ID NO: 1 and/or TXS _3xSG_MBP-SKL as set forth in SEQ ID NO: 2, or functional homologues of any of the aforementioned having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto.

4. The method and/or host cell according to any one of the preceding claims, wherein the host cell is a yeast cell belonging to the genus of *Saccharomyces*, *Pichia*, *Candida, Cryptococcus, Pichia (Komagataella), Lipomyces, Pseudozyma, Rhodosporidium, Rhodotorula, Trichosporon, Trigonopsis, Yarrowia* or *Saccharomycopsis,* such as a yeast cell of the species *Saccharomyces cerevisiae Yarrowia lipolytica, Hansenula polymorpha (Ogataea polymorpha), Rhodotorula toruloides* or *Pichia pastoris* (*Komagataella phaffii*).

5. The method and/or host cell according to any one of the preceding claims, wherein said host cell further comprises a heterologous nucleic acid encoding taxadiene synthase (EC 4.2.3.17) capable of converting geranylgeranyl diphosphate (GGPP) into taxadiene, preferably said taxadiene synthase is TXS as set forth in SEQ ID NO: 1 and/or TXS _3xSG_MBP-SKL as set forth in SEQ ID NO: 2 or a functional homologue of any of the aforementioned having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto.

6. The method and/or host cell according to any one of the preceding claims, wherein said host cell further comprises a nucleic acid encoding cytochrome P450 reductase (CPR), such as an endogenous and/or a heterologous CPR, for example a CPR from a *Taxus*, such as *Taxus cuspidata*, a CPR from *Nicotiana*, such as *Nicotiana benthamiana*, a CPR from *Arabidopsis*, such as *A. thaliana*, and/or a CPR from *Populus*, such as *Populus trichocarpa* × *Populus deltoides*.

7. The method and/or host cell according to claim 6, wherein the CPR is TcuCPR as set forth in SEQ ID NO: 4 and/or CPR2 as set forth in SEQ ID NO: 6, or functional homologues of any of the aforementioned having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto.

8. The method and/or host cell according to any one of the preceding items, wherein the host cell further comprises a nucleic acid encoding an acetyltransferase, such as an acetyltransferase from a *Taxus* cell, for example a *Taxus chinensis* cell.

9. The method and/or host cell according to claim 8, wherein the acetyltransferase is TAX19 as set forth in SEQ ID NO: 46, or functional homologues thereof having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto.

10. The method and/or host cell according to any one of the preceding claims, wherein the host cell is further capable of producing an *O*-acetylated taxane comprising an oxetane.

11. A nucleic acid construct for expression in a host cell, comprising:
    i. a nucleic acid encoding CYP1 as set forth in SEQ ID NO: 3 and/or a functional variant thereof having at least 70% sequence identity thereto, such as SEQ ID NO: 10 and/or SEQ ID NO: 11, and a nucleic acid encoding CYP3 as set forth in SEQ ID NO: 7, such as SEQ ID NO: 15; or
    ii. a nucleic acid encoding CYP1 as set forth in SEQ ID NO: 3 and/or a functional variant thereof having at least 70% sequence identity thereto, such as SEQ ID NO: 10 and/or SEQ ID NO: 11, and/or
    a nucleic acid encoding CYP3 as set forth in SEQ ID NO: 7, such as SEQ ID NO: 15; and one or more of the following:

a. a nucleic acid encoding TXS as set forth in SEQ ID NO: 1 and/or TXS _3xSG_MBP_SKL as set forth in SEQ ID NO: 2, such as SEQ ID NO: 8 and/or SEQ ID NO: 9, respectively,
b. a nucleic acid encoding ERG20(F96C)-L as set forth in SEQ ID NO: 5, such as SEQ ID NO: 13, CcGGDPS1 as set forth in SEQ ID NO: 42, such as SEQ ID NO: 43, ERG20(F96C), and/or ERG20(Y95A),
c. a nucleic acid encoding TcuCPR as set forth in SEQ ID NO: 4, such as SEQ ID NO: 12, CPR2 as set forth in SEQ ID NO: 6, such as SEQ ID NO: 44, and/or AtCPR (NCBI GenBank accession number: NP_194183, 4 June 2023),
d. a nucleic acid encoding *S. cerevisiae* ERG20 (NCBI GenBank accession number: NM_001181600, 3 June 2023), and/or SfFDPS1 as set forth in SEQ ID NO: 39, such as SEQ ID NO: 14,
e. a nucleic acid encoding SpCytb5 as set forth in SEQ ID NO: 40, such as SEQ ID NO: 36,
f. a nucleic acid encoding SpCBR as set forth in SEQ ID NO: 38, such as SEQ ID NO: 37,
g. a nucleic acid encoding HMG2(K6R) as set forth in SEQ ID NO: 20, such as SEQ ID NO: 41,

or functional homologues of any of the aforementioned having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto.

12. The nucleic acid construct according to claim 11, further comprising a nucleic acid encoding TAX19 as set forth in SEQ ID NO: 46 or a functional homologue thereof having at least 70% sequence identity thereto, such as SEQ ID NO: 48 or a functional having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto.

13. A fermentation liquid comprising the taxane comprising an oxetane, *O*-acetylated taxane comprising an oxetane, 10-hydroxy-taxane, 11-hydroxy-taxane and/or said 13-hydroxy-taxane, wherein said fermentation liquid is:

i. obtained by the method of any one of claims 1, 0, 4 to 10, or 14 to 24;
ii. comprised within and/or secreted by the host cell according to any one of claims 3 to 10, or 14 to 24.

14. The method, host cell and/or fermentation liquid according to any one of the preceding claims, wherein the taxane comprising an oxetane is 4-hydroxy-5,20-epoxy-taxane or a derivative thereof.

15. The method, host cell and/or fermentation liquid according to any one of the preceding claims, wherein the 10-hydroxy-taxane is 10-hydroxy-taxadiene or a derivative thereof.

16. The method, host cell and/or fermentation liquid according to any one of the preceding claims, wherein the 11-hydroxy-taxane is 11-hydroxyl-4,12-taxadiene or a derivative thereof.

17. The method,host cell and/or fermentation liquid according to any one of the preceding claims, wherein the 13-hydroxy-taxane is 13-hydroxy-taxadiene or a derivative thereof.

18. The method, host cell and/or fermentation liquid according to any one of the preceding claims, wherein the taxane comprising an oxetane is a 10-hydroxy-5,20-epoxy-taxane,11-hydroxy-5,20-epoxy-taxane or 13-hydroxy-5,20-epoxy-taxane.

19. The method, host cell according to any one of the preceding claims, wherein the taxane comprising an oxetane is baccatin III.

20. The method, the host cell and/or fermentation liquid according to any one of the preceding claims, wherein said *O*-acetylated taxane comprising an oxetane is an *O*-acetylated 5,20-epoxy-taxane.

21. The method, the host cell and/or fermentation liquid according to any one of the preceding claims, wherein said *O*-acetylated taxane comprising an oxetane is 4-acetoxy-5,20-epoxy-taxane.

22. The method, the host cell and/or fermentation liquid according to any one of the preceding claims, wherein the host cell comprises both a heterologous nucleic acid encoding CYP1 of SEQ ID NO: 3 and a heterologous nucleic acid encoding CYP3 as set forth in SEQ ID NO: 7, or functional homologues of the aforementioned having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto, and wherein said taxane comprising an oxetane is one or more of 4-hydroxy-10-hydroxy-5,20-epoxy-taxane, 10-hydroxy-5,20-epoxy-taxane, 11-hydroxy-5,20-epoxy-taxane, 4-hydroxy-11-hydroxy-5,20-epoxy-taxane, 4-hydroxy-12-hydroxy-5,20-epoxy-12,13-taxane, 13-hydroxy-5,20-epoxy-taxane, and/or 4-hydroxy-13-hydroxy-5,20-epoxy-taxane.

23. The method, the host cell and/or fermentation liquid according to any one of the preceding claims, wherein the host cell comprises a heterologous nucleic acid encoding CYP1 of SEQ ID NO: 3 or a functional homologue thereof having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto, and wherein said taxane comprising an oxetane is 4-hydroxy-5,20-epoxy-taxane.

24. The method, the host cell and/or fermentation liquid according to any one of the preceding claims, wherein the host cell comprises a heterologous nucleic acid encoding CYP3 as set forth in SEQ ID NO: 7 or a functional homologue thereof having at least 70%, such as at least 80%, for example at least 90%, such as at least 95%, for example at least 99% sequence identity thereto, and wherein said taxane comprising an oxetane one or more of 10-hydroxy-taxadiene, 11-hydroxyl-4,12-taxadiene, and/or 13-hydroxy-taxadiene.

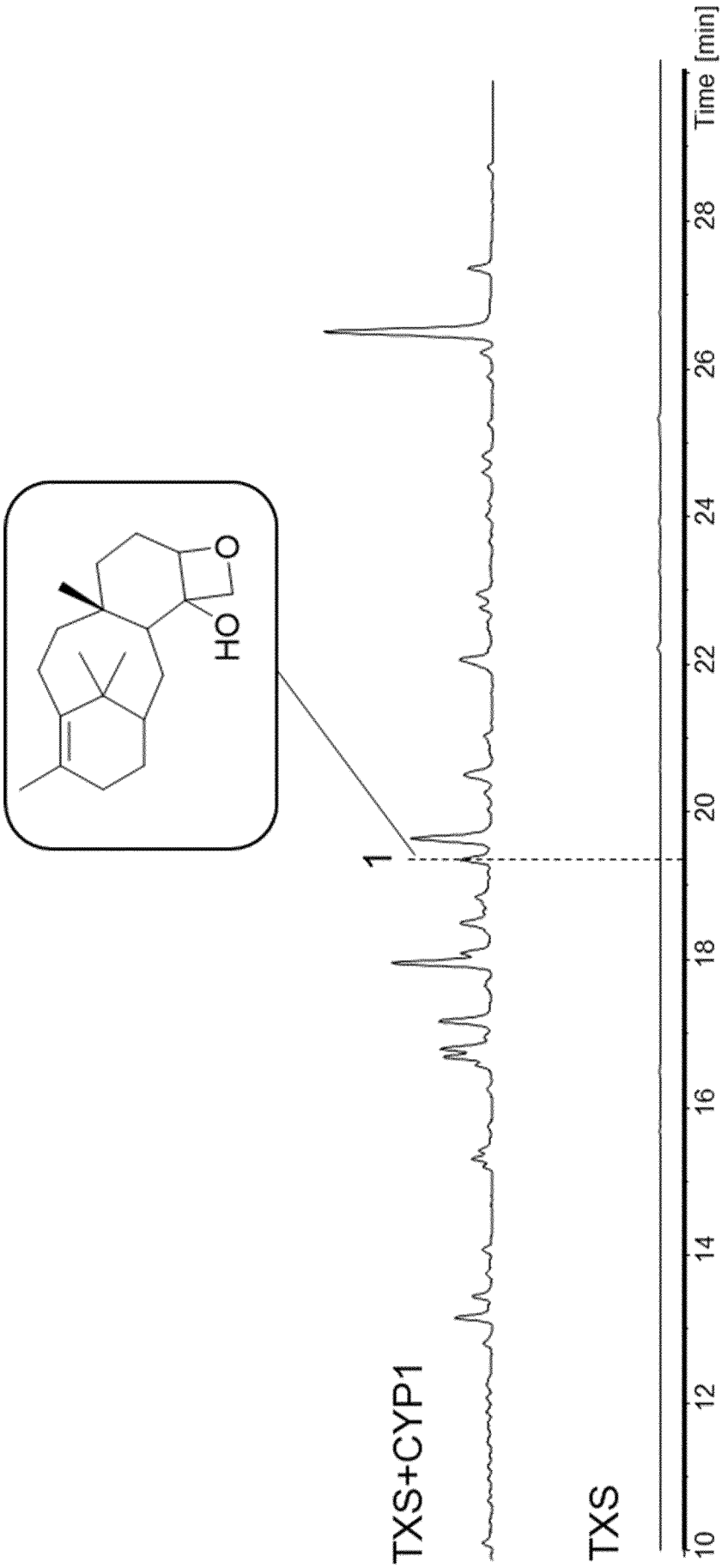


Figure 1

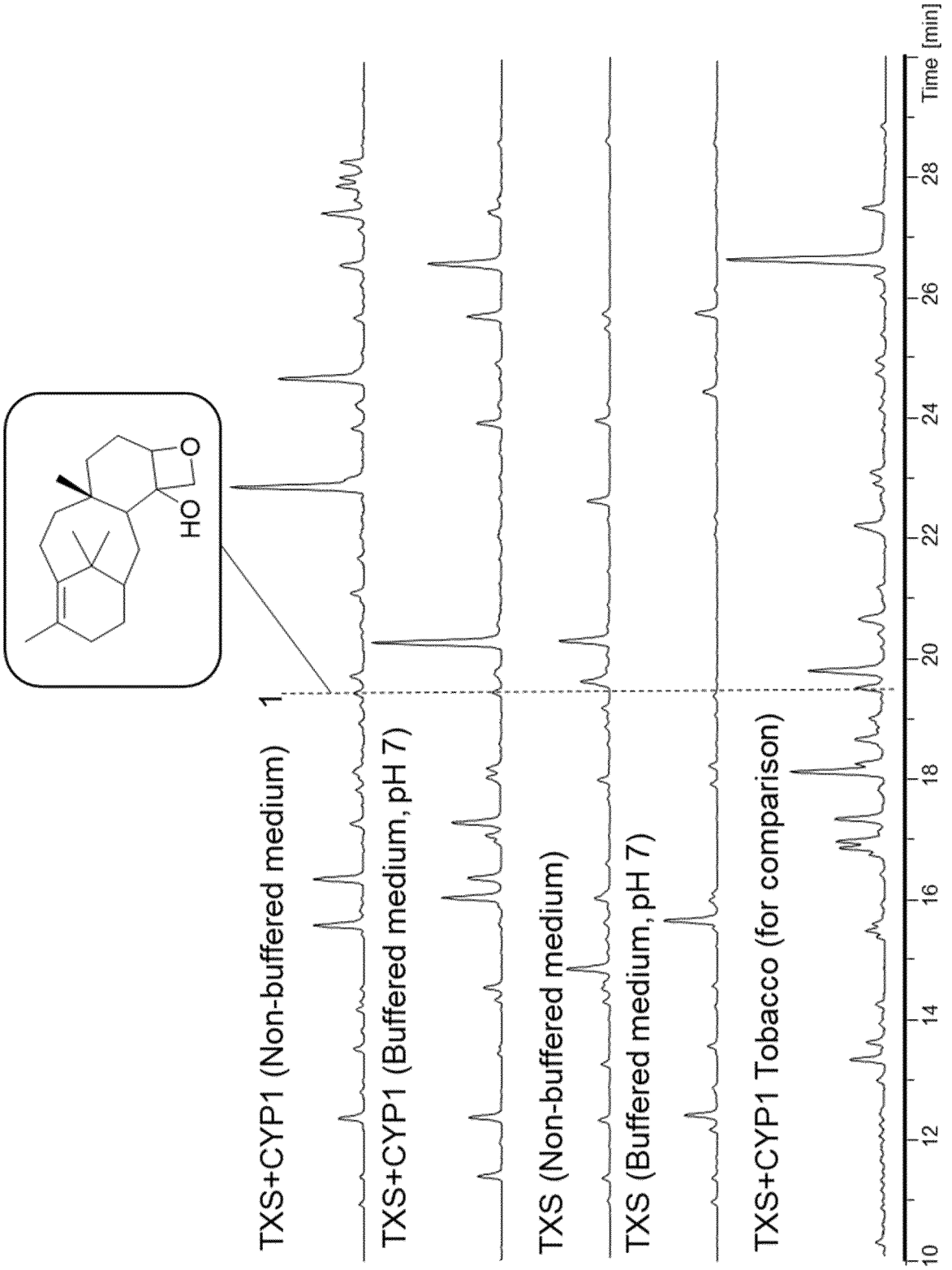

HO
O
1
TXS+CYP1 (Non-buffered medium)
TXS+CYP1 (Buffered medium, pH 7)
TXS (Non-buffered medium)
TXS (Buffered medium, pH 7)
TXS+CYP1 Tobacco (for comparison)
10
12
14
16
18
20
22
24
26
28
Time [min]

Figure 2

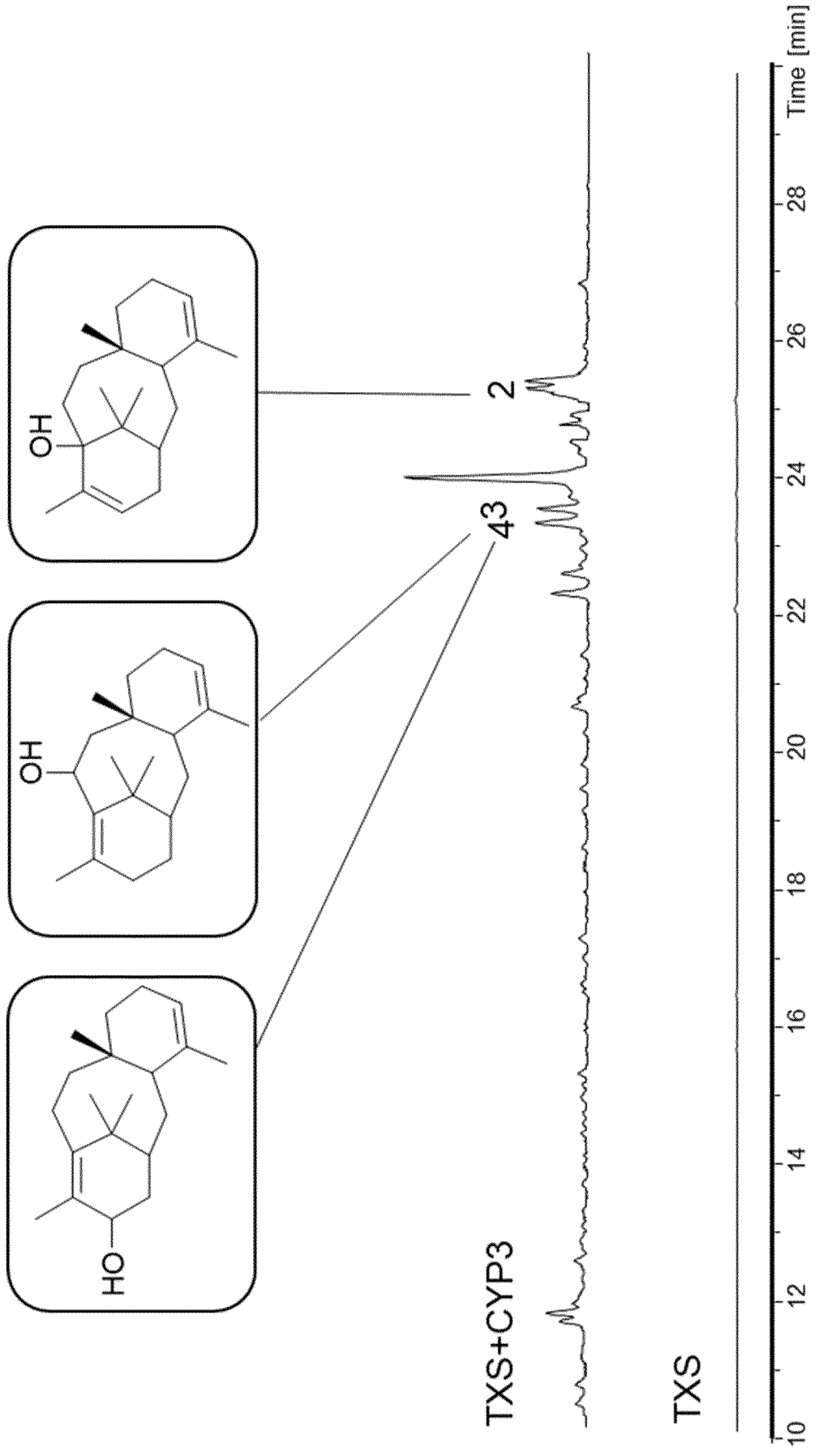


Figure 3

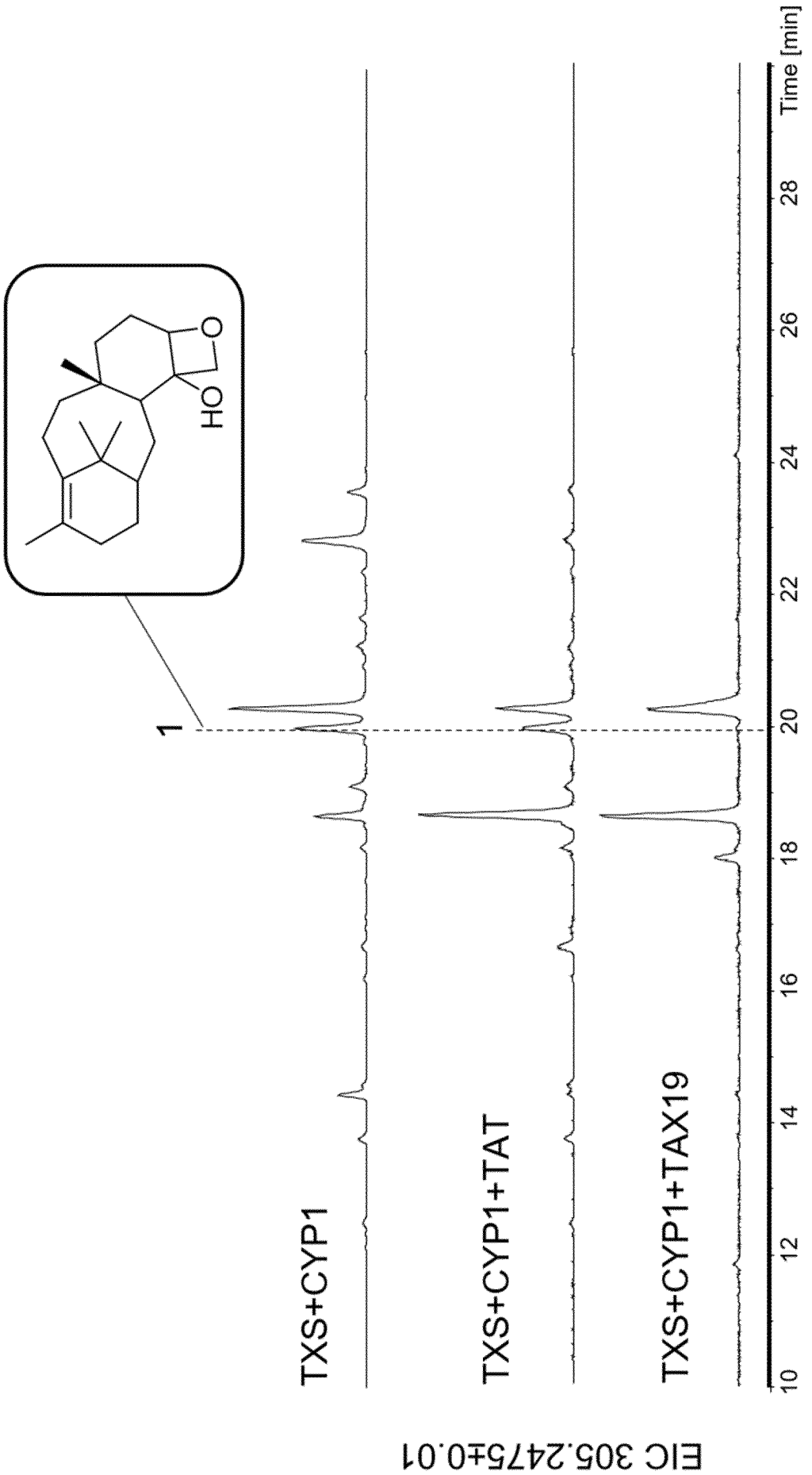

EIC 305.24754±0.01
TXS+CYP1
TXS+CYP1+TAT
TXS+CYP1+TAX19
1
HO
O
10
12
14
16
18
20
22
24
26
28
Time [min]

Figure 4A

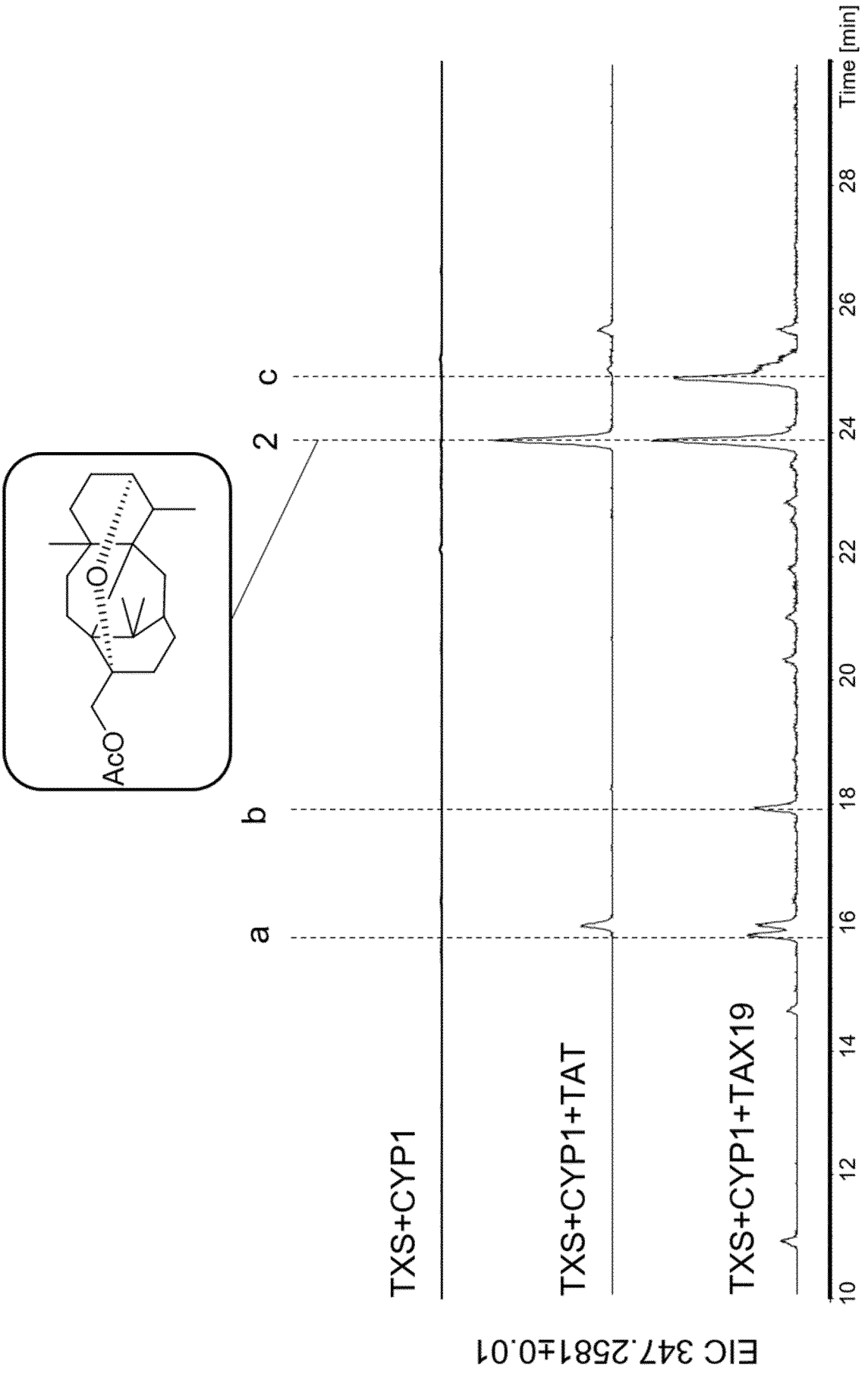

EIC 347.2581±0.01
TXS+CYP1
TXS+CYP1+TAT
TXS+CYP1+TAX19
a
b
c
2
AcO
O
10
12
14
16
18
20
22
24
26
28
Time [min]

Figure 4B

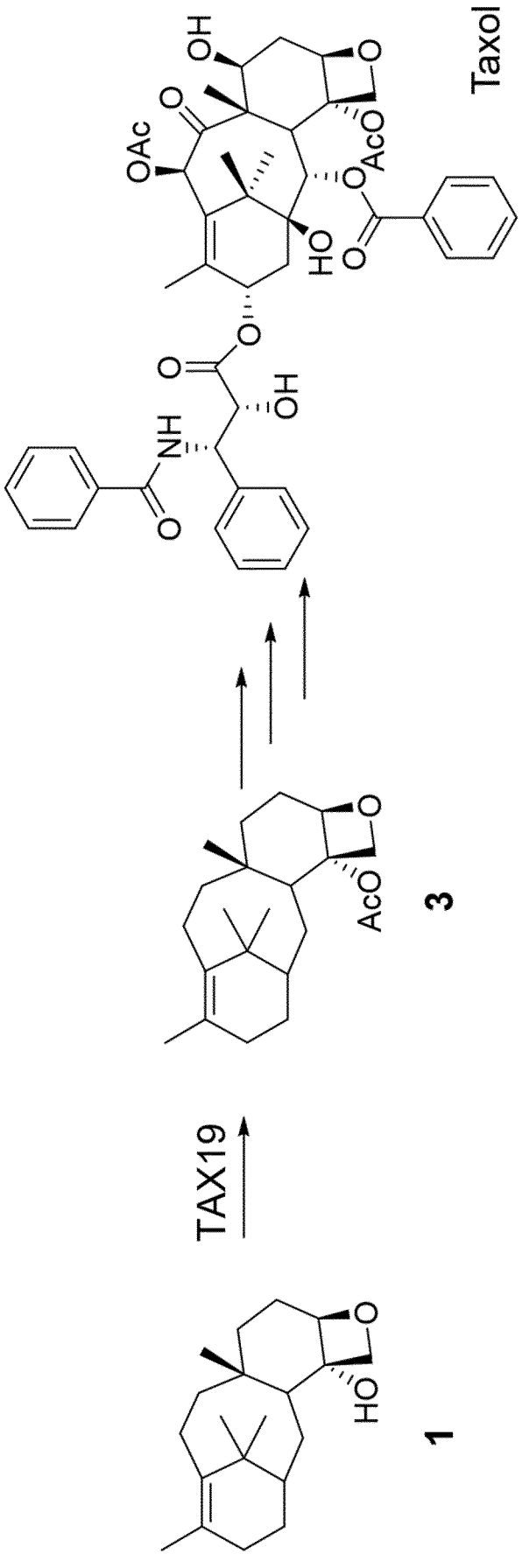


Figure 4C

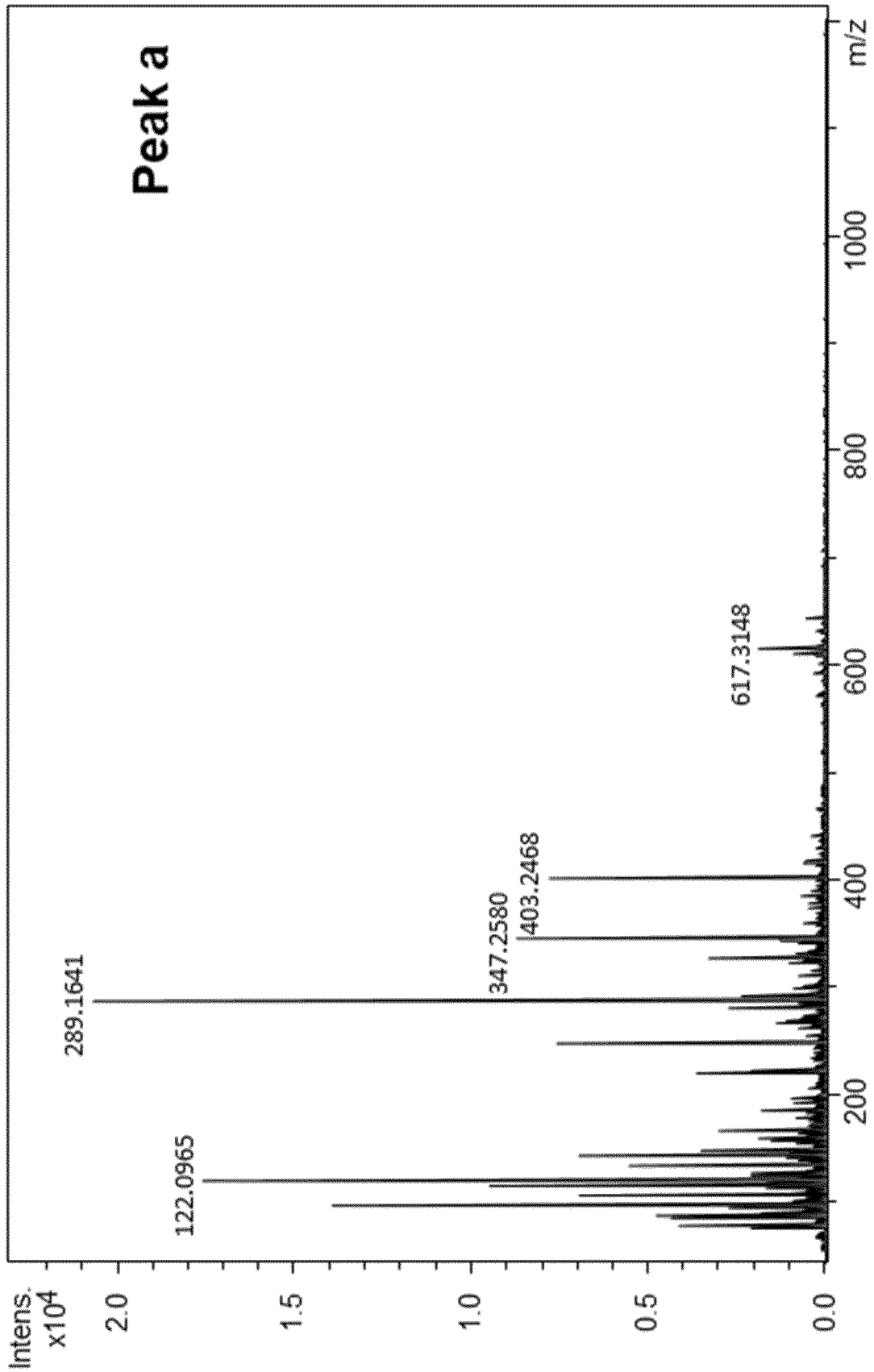

Peak a
Intens.
x10^4
2.0
1.5
1.0
0.5
0.0
122.0965
289.1641
347.2580
403.2468
617.3148
200
400
600
800
1000
m/z

Figure 5A

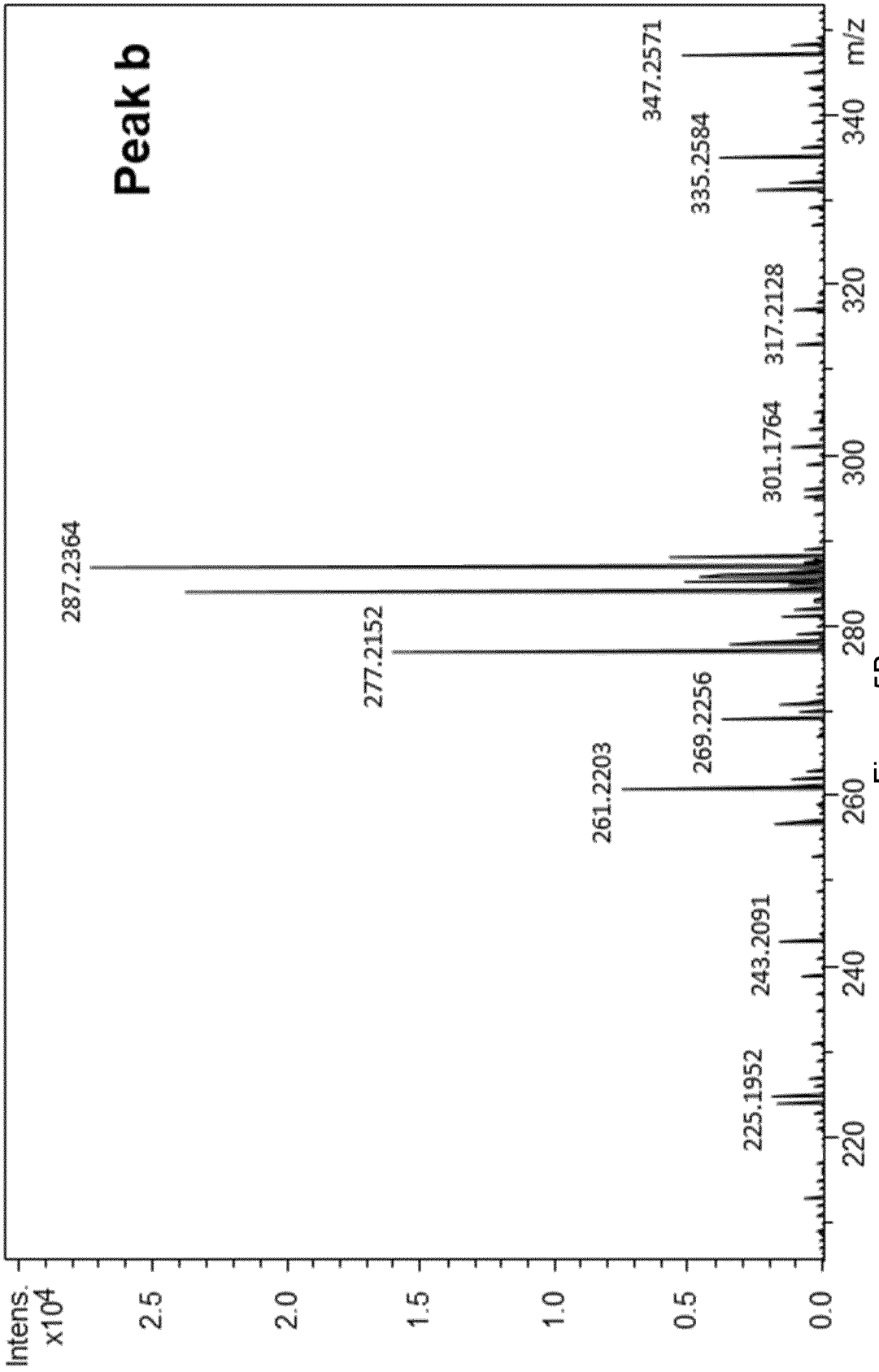

Peak b
Intens.
x10^4
2.5
2.0
1.5
1.0
0.5
0.0
287.2364
277.2152
269.2256
261.2203
243.2091
225.1952
301.1764
317.2128
335.2584
347.2571
220
240
260
280
300
320
340
m/z

Figure 5B

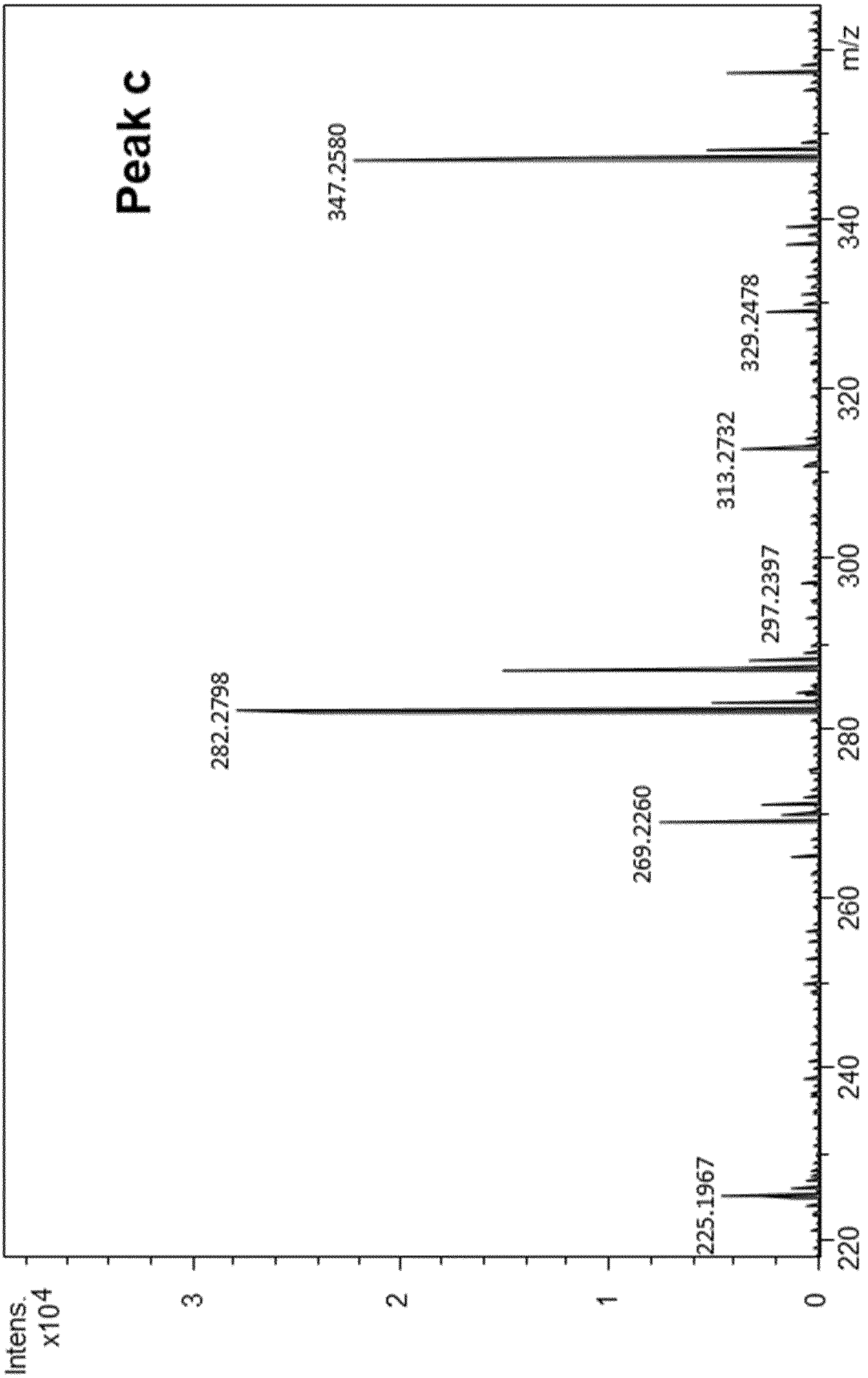

Peak c
Intens.
x10^4
3
2
1
0
220
240
260
280
300
320
340
m/z
225.1967
269.2260
282.2798
297.2397
313.2732
329.2478
347.2580

Figure 5C

INTERNATIONAL SEARCH REPORT

International application No.

PCT/EP2024/065772

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:

   a. [X] forming part of the international application as filed.

   b. [ ] furnished subsequent to the international filing date for the purposes of international search (Rule 13*ter*.1(a)).

      [ ] accompanied by a statement to the effect that the sequence listing does not go beyond the disclosure in the international application as filed.

2. [ ] With regard to any nucleotide and/or amino acid sequence disclosed in the international application, this report has been established to the extent that a meaningful search could be carried out without a WIPO Standard ST.26 compliant sequence listing.

3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2024/065772

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12N15/82 C12N15/70
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)
C12N C40B

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPO-Internal, BIOSIS, Sequence Search, EMBASE, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2011/060057 A1 (MASSACHUSETTS INST TECHNOLOGY [US]; AJIKUMAR PARAYIL K [US] ET AL.) 19 May 2011 (2011-05-19) the whole document	1-24
X	US 2003/077796 A1 (CROTEAU RODNEY B [US] ET AL) 24 April 2003 (2003-04-24) the whole document sequences 3, 7	1-24

-/--

[x] Further documents are listed in the continuation of Box C.

[x] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance;; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance;; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
16 July 2024	31/07/2024

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Keller, Yves

1

Form PCT/ISA/210 (second sheet) (April 2005)

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
T	RODNEY CROTEAU ET AL: "Taxol Biosynthesis and Molecular Genetics", PHYTOCHEMISTRY REVIEWS, KLUWER ACADEMIC PUBLISHERS, DO, vol. 5, no. 1, 1 June 2006 (2006-06-01), pages 75-97, XP019405391, ISSN: 1572-980X, DOI: 10.1007/S11101-005-3748-2 the whole document - - - - -	
T	WALKER KEVIN ET AL: "Partial purification and characterization of acetyl coenzyme A: Taxa-4(20),11(12)-dien-5alpha-ol O-acetyl transferase that catalyzes the first acylation step of Taxol biosynthesis", ARCHIVES OF BIOCHEMISTRY AND BIOPHYSICS, ACADEMIC PRESS, US, vol. 364, no. 2, 15 April 1999 (1999-04-15), pages 273-279, XP002175244, ISSN: 0003-9861, DOI: 10.1006/ABBI.1999.1125 the whole document - - - - -	
T	WALKER KEVIN ET AL: "Taxol biosynthesis: Molecular cloning of a benzoyl-CoA:taxane 2alpha-O-benzoyltransferase cDNA from Taxus and functional expression in Escherichia coli", PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES, NATIONAL ACADEMY OF SCIENCES, vol. 97, no. 25, 5 December 2000 (2000-12-05), pages 13591-13596, XP002175248, ISSN: 0027-8424, DOI: 10.1073/PNAS.250491997 the whole document - - - - -	
T	WALKER KEVIN ET AL: "Molecular cloning of a 10-deacetylbaccatin III-10-O-acetyl transferase cDNA from Taxus and functional expression in Escherichia coli", PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES, NATIONAL ACADEMY OF SCIENCES, vol. 97, no. 2, 18 January 2000 (2000-01-18), pages 583-587, XP002175247, ISSN: 0027-8424, DOI: 10.1073/PNAS.97.2.583 the whole document - - - - -	

-/--

1

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
T	WILDUNG M R ET AL: "CDNA CLONE FOR TAXADIENE SYNTHASE, THE DITERPENE CYCLASE THAT CATALYZES THE COMMITTED STEP OF TAXOL BIOSYNTHESIS", JOURNAL OF BIOLOGICAL CHEMISTRY, AMERICAN SOCIETY FOR BIOCHEMISTRY AND MOLECULAR BIOLOGY, US, vol. 271, no. 16, 19 April 1996 (1996-04-19), pages 9201-9204, XP000938117, ISSN: 0021-9258, DOI: 10.1074/JBC.271.38.23262 the whole document -----	
T	RÜDIGER KASPERA ET AL: "Cytochrome P450 oxygenases of Taxol biosynthesis", PHYTOCHEMISTRY REVIEWS, KLUWER ACADEMIC PUBLISHERS, DO, vol. 5, no. 2-3, 28 October 2006 (2006-10-28), pages 433-444, XP019453404, ISSN: 1572-980X, DOI: 10.1007/S11101-006-9006-4 the whole document -----	
T	BRADLEY WALTERS BIGGS ET AL: "Overcoming heterologous protein interdependency to optimize P450-mediated Taxol precursor synthesis in Escherichia coli", PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES, vol. 113, no. 12, 7 March 2016 (2016-03-07), pages 3209-3214, XP055525453, ISSN: 0027-8424, DOI: 10.1073/pnas.1515826113 -----	
T	Vikramaditya Ganapati Yadav: "Biosynthetic Engineering for the Assembly of Better Drugs", , 1 January 2013 (2013-01-01), XP093186110, Retrieved from the Internet: URL:https://core.ac.uk/download/pdf/183209 01.pdf the whole document -----	

-/--

1

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
T	AJIKUMAR PARAYIL KUMARAN ET AL: "Isoprenoid Pathway Optimization for Taxol Precursor Overproduction in Escherichia coli", SCIENCE, AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE, US, vol. 330, no. 6000, 1 October 2010 (2010-10-01), pages 70-74, XP002618146, ISSN: 0036-8075, DOI: 10.1126/SCIENCE.1191652 the whole document -----	
T	YECHUN WANG ET AL: "Transformation of taxol-producing endophytic fungi by restriction enzyme-mediated integration (REMI)", FEMS MICROBIOLOGY LETTERS, vol. 273, no. 2, 1 August 2007 (2007-08-01), pages 253-259, XP055538501, ISSN: 0378-1097, DOI: 10.1111/j.1574-6968.2007.00801.x the whole document -----	
T	WALLS LAURA ELLEN ET AL: "Enhancing Saccharomyces cerevisiae Taxane Biosynthesis and Overcoming Nutritional Stress-Induced Pseudohyphal Growth", MICROORGANISMS, vol. 10, no. 1, 13 January 2022 (2022-01-13), page 163, XP093186150, ISSN: 2076-2607, DOI: 10.3390/microorganisms10010163 the whole document -----	
T	WANG TAO ET AL: "Recent Research Progress in Taxol Biosynthetic Pathway and Acylation Reactions Mediated by Taxus Acyltransferases", MOLECULES, vol. 26, no. 10, 12 May 2021 (2021-05-12), page 2855, XP093080583, DOI: 10.3390/molecules26102855 Retrieved from the Internet: URL:https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8151764/pdf/molecules-26-02855.pdf> the whole document -----	

-/--

1

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2024/065772

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
T	Anonymous: "The Oxford Dictionary and thesaurus", The Oxford Dictionary and Thesaurus, 1 January 1997 (1997-01-01), pages 1414-1414, XP093186395, ISBN: 978-0-19-860171-5 page 1414, column 1, line 21 - line 26	
T	DE LA PEÑA RICARDO ET AL: "Rerouting plant terpene biosynthesis enables momilactone pathway elucidation", NATURE CHEMICAL BIOLOGY, vol. 17, no. 2, 1997, pages 205-212, XP037344760, ISSN: 1552-4450, DOI: 10.1038/S41589-020-00669-3 the whole document	

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/EP2024/065772

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2011060057 A1	19-05-2011	BR 112012011073 A2	06-10-2020
		CA 2780415 A1	19-05-2011
		CN 102812129 A	05-12-2012
		EP 2499257 A1	19-09-2012
		EP 2594648 A1	22-05-2013
		ES 2575790 T3	01-07-2016
		JP 5844738 B2	20-01-2016
		JP 2013510573 A	28-03-2013
		KR 20120101445 A	13-09-2012
		MY 162232 A	31-05-2017
		SG 181422 A1	30-07-2012
		US 2011189717 A1	04-08-2011
		US 2014024089 A1	23-01-2014
		US 2017002366 A1	05-01-2017
		US 2018080035 A1	22-03-2018
		WO 2011060057 A1	19-05-2011
US 2003077796 A1	24-04-2003	US 2003077796 A1	24-04-2003
		US 2004236089 A1	25-11-2004
		US 2007067864 A1	22-03-2007