Glycosyltransferase glycosylating flavokermesic acid and/or kermesic acid

Kannangara, Rubini Maya; Bennedsen, Mads; Madsen, Bjørn; Binderup, Kim; Thrane, Ulf; Frandsen, Rasmus John Normand ; Mortensen, Uffe Hasbro; Møller, Birger Lindberg; Okkels, Finn Thyge

Publication date:
2016

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

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 3 083 950 B1

(12) EUROPEAN PATENT SPECIFICATION

(45) Date of publication and mention of the grant of the patent:
29.11.2017 Bulletin 2017/48

(21) Application number: **14828027.4**

(22) Date of filing: **18.12.2014**

(51) Int Cl.:
C12N 9/10 *(2006.01)* ***C12P 7/42*** *(2006.01)*
C12P 7/66 *(2006.01)*

(86) International application number:
PCT/EP2014/078540

(87) International publication number:
WO 2015/091843 (25.06.2015 Gazette 2015/25)

(54) **GLYCOSYLTRANSFERASE GLYCOSYLATING FLAVOKERMESIC ACID AND/OR KERMESIC ACID**

FLAVOKERMESIC SÄURE UND / ODER KERMESSÄURE GLYKOSYLIERENDEN GLYCOSYLTRANSFERASE

GLYCOSYLTRANSFÉRASE GLYCOSYLER D'ACIDE FLAVOKERMESIC ET / OU D'ACIDE KERMÉSIQUE

(84) Designated Contracting States:
AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR

(30) Priority: **18.12.2013 EP 13198110**

(43) Date of publication of application:
26.10.2016 Bulletin 2016/43

(73) Proprietors:
- **Københavns Universitet**
  1165 Copenhagen (DK)
- **Danmarks Tekniske Universitet**
  2800 Lyngby (DK)

(72) Inventors:
- **KANNANGARA, Rubini Maya**
  DK-2000 Frederiksberg (DK)
- **BENNEDSEN, Mads**
  DK-3230 Græsted (DK)
- **MADSEN, Bjørn**
  DK-3000 Helsingør (DK)
- **BINDERUP, Kim**
  DK-2920 Charlottelund (DK)
- **THRANE, Ulf**
  DK-3200 Helsinge (DK)
- **FRANDSEN, Rasmus John Normand**
  DK-3450 Allerød (DK)
- **MORTENSEN, Uffe Hasbro**
  DK-2200 Copenhagen N (DK)
- **MØLLER, Birger Lindberg**
  DK-2700 Brønshøj (DK)
- **OKKELS, Finn Thyge**
  DK-4000 Roskilde (DK)

(74) Representative: **Plougmann Vingtoft a/s**
Rued Langgaards Vej 8
2300 Copenhagen S (DK)

(56) References cited:
WO-A1-92/01664

- **CHAVEZ-MORENO C K ET AL: "Metabolic profiling of Dactylopius (Hemiptera: Dactylopiidae) species pigments by geographical origin and hosts using multivariate data analysis", BIOCHEMICAL SYSTEMATICS AND ECOLOGY, PERGAMON PRESS, GB, vol. 38, no. 4, 1 August 2010 (2010-08-01) , pages 671-679, XP027437752, ISSN: 0305-1978, DOI: 10.1016/J.BSE.2010.04.003 [retrieved on 2010-06-01]**
- **ROBYN MEECH ET AL: "The glycosidation of xenobiotics and endogenous compounds: Versatility and redundancy in the UDP glycosyltransferase superfamily", PHARMACOLOGY AND THERAPEUTICS, vol. 134, no. 2, February 2012 (2012-02), pages 200-218, XP028477559, ISSN: 0163-7258, DOI: 10.1016/J.PHARMTHERA.2012.01.009 [retrieved on 2012-02-01]**

Printed by Jouve, 75001 PARIS (FR)

(Cont. next page)

- IRFAN BAIG ET AL: "On the Acceptor Substrate of C-Glycosyltransferase UrdGT2: Three Prejadomycin C-Glycosides from an Engineered Mutant of Streptomyces globisporus 1912 DELTA-IndE(urdGT2)", ANGEWANDTE CHEMIE INTERNATIONAL EDITION, WILEY - V C H VERLAG GMBH & CO. KGAA, DE, vol. 45, no. 46, 27 November 2006 (2006-11-27), pages 7842-7846, XP008133847, ISSN: 1433-7851, DOI: 10.1002/ANIE.200603176 [retrieved on 2006-10-24]

Description

FIELD OF THE INVENTION

[0001] The present invention relates to an isolated glycosyltransferase (GT) polypeptide capable of: (I): conjugating glucose to flavokermesic acid (FK); and/or (II): conjugating glucose to kermesic acid (KA) and the use of this GT to make e.g. Carminic acid.

BACKGROUND OF THE INVENTION

[0002] The natural pigment carminic acid is one of the most frequently used colorants of food, medicine, cosmetics and textiles.

[0003] Carminic acid is a colorant, which can be extracted from the female insect bodies of *Dactylopius coccus costa* (alternative name *Coccus cacti* L.). The insects live on *Nopalea coccinellifera, Opuntia fidus indica* and other plants of the family Cactaceae cultivated for instance in the desert areas of Mexico, Central and South America and Canary Islands. Depending on the pH the colorant may be a color in a spectrum from orange over red to purple and is generally known as cochineal or cochineal color. Carmine colorant is widely used in foods and beverages.

[0004] As known in the art *Porphyrophora polonica* is also producing carminic acid and was cultured for production of carminic acid in e.g. Poland.

[0005] In relation to current industrial relevant production, carminic acid is harvested by extraction from the insect's dried bodies with water or alcohol.

[0006] The insects (*Dactylopius coccus*) are cultured on cacti and the supply may therefore be relatively expensive and subject to undesirable variations and price fluctuations.

[0007] In order to try to resolve the problem of undesirable variations and price fluctuations - US5424421 (European Colour, published 1995) describes chemical synthesis of carminic acid by a route of synthesis involving different intermediates.

[0008] As discussed in e.g. WO2006/056585A1 (Chr. Hansen A/S) - during the aqueous based extraction of carminic acid from the insect, an amount of insect protein is also released from the insect and will be contained in the color extract and it has been reported that the cochineal insect proteins could create some allergy related problems. In WO2006/056585A1 a special process to reduce the amount of insect protein from the insect extract solution is described - however, the final produced color composition/product of WO2006/056585A1 will still comprise some amounts *Dactylopius coccus* costa insect proteins.

[0009] The structure of carminic acid is shown in Figure 1 - as can be seen it is a so-called C-glucoside (i.e. wherein the glucose is joined/conjugated to the aglucon by a carbon-carbon linkage).

[0010] According to the art - the term "aglycon" denotes the non-carbohydrate part of the corresponding glycosylated form of the aglycon. When the sugar is glucose the aglycon may be termed aglucon.

[0011] According to the art - the term "glycoside" denotes a compound which by hydrolysis results in a sugar and a non-sugar (aglycon) residue, e.g. glucosides can give glucose, galactosides can give galactose. As shown in Figure 1 - hydrolysis of the C-glucoside carminic acid results in glucose and the aglucon kermesic acid (KA).

[0012] The *in vivo* insect (*Dactylopius coccus*) biosynthetic pathway involved in carmine production is currently not described in details - accordingly, based on the prior art the skilled person does not know which compound is the aglucon during the *in vivo Dactylopius coccus* biosynthetic production of carminic acid.

[0013] Analysis of *D. coccus* has shown that a broad range of compounds related to carminic acid are present in extract from *D. coccus* and numerous of these compounds could in principle be glucosylated during the *in vivo Dactylopius coccus* biosynthetic production of carminic acid.

For instance, the article of Stathopoulou et al (Analytica Chimica Acta 804 (2013) 264-272) describes six new anthraquinones that are present in extract from *D. coccus* and any of these six new anthraquinones (see e.g. Figure 1 of the article) could in principle be the molecule which is glucosylated during the *in vivo Dactylopius coccus* biosynthetic production of carminic acid.

Furthermore, and as known in the art, the primary glucosylated compound formed during the *in vivo* biosynthetic production of the glucoside end product may be an unstable intermediate compound that will not be identified in an isolated extract from *D. coccus* as e.g. analyzed in the above discussed article of Stathopoulou et al.

Based on the prior art, it could be speculated that a relevant primary glucosylated compound during the *in vivo Dactylopius coccus* biosynthetic production of carminic acid could e.g. be an unstable intermediate polyketide compound with around the same number of carbon atoms as e.g. flavokermesic acid.

[0014] According to the art - the term "glycosyltransferase" (GT) denotes a glycosyltransferase enzyme capable of transferring a sugar from an activated nucleotide sugar to an aglycon to form a glycoside.

[0015] A herein relevant DNA or amino acid sequence of a glycosyltransferase involved in the *in vivo* insect (*Dactylopius*

coccus) biosynthetic pathway of carminic acid is not explicitly described in the prior art.

[0016] As known in the art, for insects that accumulate low molecular weight chemicals the relevant biosynthetic pathway genes are sometimes not present in the insect genome. For instance, some insects take up glycosides from the plants they feed on - see e.g. the article of Zagrobelny et al (Cyanogenic glucosides and plant-insect interactions; Phytochemistry. 2004 Feb;65(3):293-306) or the article of Geuder et al (Journal of Chemical Ecology, Vol. 23, No. 5, 1997). Also, the relevant biosynthetic pathway genes are sometimes found in the microorganisms living in the insects, see e.g. the article of Genta et al, (Potential role for gut microbiota in cell wall digestion and glucoside detoxification in Tenebrio molitor larvae), Journal of Insect Physiology 52 (2006) 593-601.

Dactylopius coccus insects feed on cactus plants and it could be that *D. coccus* insects (like other insects) take up relevant glycosides from the cactus they feed on

[0017] Accordingly, based on the prior art the skilled person could not know if the genome of *Dactylopius coccus* actually would comprise a gene encoding a glycosyltransferase involved in the *in vivo* biosynthetic pathway leading to carminic acid.

[0018] WO2004/111254A1 (Poalis A/S) describes *in vivo* production of a glucosylated form of vanillin in e.g. eukaryotic cell yeast cells and/or prokaryotic *E. coli* cells by using a glucosyltransferase for conjugating glucose to the vanillin aglucon *in vivo* within a microorganism cell. Natural vanillin is obtained from the plant vanilla bean. Accordingly, in the prior art successful *in vivo* production has been described in microorganism cells of plant glycoside compounds (such as e.g. vanillin glucoside).

SUMMARY OF THE INVENTION

[0019] The problem to be solved by the present invention relates to the provision of a glycosyltransferase (GT) involved in a biosynthetic pathway that may lead to carminic acid and the use of this glycosyltransferase to make e.g. carminic acid.

[0020] As discussed in working examples herein - the present inventors sequenced the entire genome and transcriptome (i.e. set of RNA molecules including mRNA) of *Dactylopius coccus* and microbial symbionts.

The identified oligonucleotide sequences obtained from the genome and transcriptome were analyzed for similarity to public known C-glycosyltransferase sequences and the result was negative. None of the identified gene sequences of the genome/transcriptome showed significant similarity to publicly known C-glycosyltransferase sequences.

[0021] As discussed above - based on the prior art the skilled person could not know if the genome of *Dactylopius coccus* actually would comprise a gene encoding a glycosyltransferase involved in the *in vivo* biosynthetic pathway leading to carminic acid. However, the present inventors continued to investigate the matter.

[0022] As discussed in working examples herein - the present inventors identified a *Dactylopius coccus* extract (including extracts of the endosymbionts present in *D. coccus*) with relevant GT activity and by a combination of relevant purification and testing steps the inventors were finally able to obtain a relatively pure fraction/composition wherefrom it was possible to obtain several partial amino acid sequences of putative GT enzyme candidates.

[0023] The partial amino acid sequences of these enzyme candidates were compared to the identified gene sequences of the transcriptome and after further detailed work, a sequence encoding a glycosyltransferase enzyme sequence was identified - the polynucleotide sequence encoding this isolated/cloned novel glycosyltransferase is shown in SEQ ID NO: 1 and the polypeptide amino acid sequence is shown in SEQ ID NO: 2.

The glycosyltransferase enzyme of SEQ ID NO: 2 may be termed "DcUGT2".

It is believed that the described isolated/cloned glycosyltransferase is the first described insect derived glycosyltransferase. As described, the identified gene sequences of the *Dactylopius coccus* transcriptome were analyzed for similarity to relevant public known glycosyltransferase sequences and the result was negative.

The present inventors found, that the publicly known prior art glycosyltransferase sequences have less than 45% identity to the novel glycosyltransferase polypeptide sequence shown as SEQ ID NO: 2.

[0024] As discussed in working examples herein - the present inventors tested the activity of the isolated/cloned novel glycosyltransferase and found that it was able to conjugate glucose to the aglycons flavokermesic acid (FK) and kermesic acid (KA) - see Figure 1.

Analysis of the glucosylated product showed that among the many potential O- and C-glucoside products potentially formed only a single glucoside product was formed with each of the two substrates. The analysis showed that each of the products were C-glucosylated at position 7 of the anthraquinone structure, more precisely 7-α-D-Glucopyranosyl-9,10-dihydro-3,6,8-trihydroxy-1-methyl-9,10-dioxoanthracenecarboxylic acid (DcII - see Figure 1) and 7-α-D-Glucopyranosyl-9,10-dihydro-3,5,6,8-tetrahydroxy-1-methyl-9,10-dioxoanthracenecarboxylic acid (carminic acid - see Figure 1) by the GT when UDP-glucose was used as the sugar donor substrate.

[0025] The article of Gutmann et al (Pure Appl. Chem, 2013-07-09) describes that even though a number of C-glycosides have been isolated from natural sources, the enzymes responsible for their biosynthesis are only known in

very few cases, and the biocatalytic approaches for C-glycoside production has yet to be established

[0026] The article of Baig et al (Angew Chem Int Ed Engl. 2006 Nov 27;45(46):7842-6) describes the glycosyltransferase (GT) termed UrdGT2 and explains that it is able to conjugate a sugar to a number of aglycons that may herein be considered relatively similar to flavokermesic acid (FK) and kermesic acid (KA).

Accordingly, it may be said that UrdGT2 prima facie would be a qualified guess for a GT that could be capable of conjugating sugar to flavokermesic acid (FK) and/or kermesic acid (KA). As discussed in working Example herein - the present inventors cloned the UrdGT2 and tested it for flavokermesic acid (FK) and/or kermesic acid (KA) GT activity and it was found that the UrdGT2 was able to use UDP-glucose as a sugar donor, but UrdGT2 did not glucosylate any of the tested putative aglycons - i.e. no GT activity was identified in relation to these aglycons.

The UrdGT2 has around 15-20% amino acid identity with SEQ ID NO:2 disclosed herein.

[0027] Based on both publicly known GT sequences and not publicly known GT sequences - the present inventors made different detailed sequence alignment investigations.

Based on these sequence alignment investigations - it is believed that amino acids 20 to around 468 of SEQ ID NO:2 comprise the essential parts of the catalytic domain.

Based on these sequence alignment investigations - it is believed that amino acids from around 291 to around 383 of SEQ ID NO:2 comprise the so-called activated nucleotide sugar binding site.

Based on these sequence alignment investigations - it is believed that amino acids from around 1 to around 20 of SEQ ID NO:2 comprise the so-called signal peptide and it is believed that this signal peptide may be removed without significantly affecting the herein relevant GT activity of the enzyme.

Furthermore, it is believed that the activated nucleotide sugar binding site may be substituted by similar (e.g. prior art known) GT activated nucleotide sugar binding site sequences - such as e.g. the activated nucleotide sugar binding site as described in Radominska-Pandya A, Bratton SM, Redinbo MR, Miley MJ. Drug Metab Rev. 2010 Feb;42(1): 133-44) and Plant Physiology, November 2008, Vol. 148, pp. 1295-1308.

[0028] Accordingly, a first aspect of the present invention relates to an isolated glycosyltransferase polypeptide capable of:

(I): conjugating glucose to flavokermesic acid (FK); and/or
(II): conjugating glucose to kermesic acid (KA);

and wherein the glycosyltransferase polypeptide is at least one polypeptide selected from the group consisting of:

(a) a polypeptide comprising an amino acid sequence which has at least 70% identity with amino acids 1 to 515 of SEQ ID NO:2;
(b) a polypeptide comprising an amino acid sequence which has at least 70% identity with amino acids 20 to 468 of SEQ ID NO:2;
(c) a polypeptide which is encoded by a polynucleotide that hybridizes under at least medium stringency conditions with (i) nucleotides 1 to 1548 of SEQ ID NO:1 or (ii) a complementary strand of (i); and
(d) a fragment of amino acids 1 to 515 of SEQ ID NO:2, which has the glycosyltransferase activity as specified in (I) and/or (II).

[0029] As understood by the skilled person in the present context - the term "a glycosyltransferase polypeptide capable of" of the first aspect relates to that the glycosyltransferase shall be capable of performing the glycosyltransferase (I) and/or (II) activity - but it may or may not also be capable of performing other glycosyltransferase activities.

[0030] As understood by the skilled person in the present context - the disclosure of the herein described novel glycosyltransferase sequence is an important tool to identify similar glycosyltransferases in e.g. other insects than *Dactylopius coccus* and without being limited to theory - it is believed that a sequence with at least 70% identity with SEQ ID NO:2 would be a plausible good candidate for a another herein relevant glycosyltransferase.

[0031] A second aspect of the present invention relates to an isolated polynucleotide comprising a nucleotide sequence which encodes the polypeptide of the first aspect and/or herein relevant embodiments thereof.

[0032] A third aspect of the present invention relates to a nucleic acid construct comprising the isolated polynucleotide of the second aspect and/or herein relevant embodiments thereof operably linked to one or more control sequences that direct the production of the polypeptide in an expression host.

[0033] A fourth aspect of the present invention relates to a recombinant expression vector comprising the nucleic acid construct of the third aspect and/or herein relevant embodiments thereof.

[0034] A fifth aspect of the present invention relates to a recombinant host cell comprising the nucleic acid construct of the third aspect and/or herein relevant embodiments thereof.

[0035] As discussed above - based on the prior art the skilled person does not know which compound is the primary glycosylated compound during the biosynthetic production of carminic acid *in vivo* in *Dactylopius coccus*.

[0036] It has been shown that *D. coccus* contains a GT able to C-glycosylate flavokermesic acid (FK) and/or kermesic acid (KA).
It is evident that this important knowledge is sufficient in order to e.g. produce carminic acid without the need to make an extraction from insects and thereby be able to make a carminic acid color composition/product essentially free of e.g. unwanted *Dactylopius coccus* costa insect proteins.
[0037] Since the skilled person did not know which compound is glycosylated during the *in vivo Dactylopius coccus* biosynthetic production of carminic acid it was actually unknown to the skilled person if there in nature actually existed a glycosyltransferase capable of C-glycosylating flavokermesic acid aglycon and/or the kermesic acid aglycon.
It is believed that the herein disclosed novel glycosyltransferase represents the first isolated glycosyltransferase capable of glycosylating flavokermesic acid aglycon and/or kermesic acid.
Accordingly, based on the technical disclosure herein - it is believed that the skilled person would be able to identify other suitable glycosyltransferases capable of glycosylating flavokermesic acid (FK) and/or kermesic acid (KA).
[0038] The skilled person would appreciate that one way to try to identify if an organism/plant would comprise a relevant glycosyltransferase would be to contact relevant aglycons (i.e. FK and/or KA) to the organism/plant (*in vivo* and/or *in vitro*) and then measure if the organism/plant produces relevant FK and/or KA glycosides.
[0039] As understood herein, if the organism/plant produces relevant FK and/or KA glycosides then the organism/plant will comprise a relevant glycosyltransferase - i.e. a glycosyltransferase which is capable of glycosylating flavokermesic acid in order to produce flavokermesic acid glycoside; and/or capable of glycosylating kermesic acid in order to produce kermesic acid glycoside.
[0040] As discussed below - based on the strategy above, the present inventors found that relevant glycosyltransferases may be identified in *Aloe* plants, *Haworthia* plants, Sorghum or rice plants.
[0041] Accordingly, a sixth aspect of the present invention relates to a method for producing flavokermesic acid (FK) glycoside and/or kermesic acid (KA) glycoside, wherein the method comprises following steps:

(A): contacting *in vitro* or *in vivo* in a recombinant host cell comprising a glycosyltransferase gene encoding a glycosyltransferase:

(a1): flavokermesic acid (FK) with a glycosyltransferase capable of glycosylating the flavokermesic acid under suitable conditions wherein there is produced the flavokermesic acid glycoside; and/or
(a2): kermesic acid (KA) with a glycosyltransferase capable of glycosylating the kermesic acid under suitable conditions wherein there is produced the kermesic acid glycoside.

[0042] The term "recombinant host cell" should herein be understood according to the art. As known in the art, recombinant polynucleotide (e.g. DNA) molecules are polynucleotide (e.g. DNA) molecules formed by laboratory methods of genetic recombination (such as molecular cloning) to bring together genetic material from multiple sources, creating sequences that would not otherwise be found in biological organisms. As understood by the skilled person - a recombinant host cell comprises recombinant polynucleotide (e.g. DNA) molecules and a recombinant host cell will therefore not be understood as covering a natural wildtype cell as such - such as e.g. a natural wildtype *Dactylopius coccus* cell.
Said in other words and as understood by the skilled person - for instance a natural wildtype *Dactylopius coccus* cell as such does not contain a recombinant glycosyltransferase gene encoding a glycosyltransferase.
[0043] It may be preferred that the recombinant host cell in step (A) is a recombinant host cell comprising a recombinant glycosyltransferase gene encoding a glycosyltransferase
[0044] As discussed herein - in working Examples there was made a contacting *in vitro* of flavokermesic acid (FK) and/or kermesic acid (KA) with the glycosyltransferase of SEQ ID NO:2. It may be seen as routine work for the skilled person to perform such an *in vitro* contacting step.
[0045] The glycosyltransferase of SEQ ID NO:2 was recombinantly expressed in a yeast cell (see working Example) - accordingly, a recombinant yeast host cell comprising a recombinant glycosyltransferase gene encoding a glycosyltransferase of SEQ ID NO:2 was made.
It is believed that if flavokermesic acid (FK) and/or kermesic acid (KA) would be added under suitable conditions to a fermentation medium the FK and/or KA compound(s) would enter into e.g. yeast cells fermented in the medium - accordingly, if e.g. the yeast cells are recombinant yeast host cells comprising a recombinant glycosyltransferase gene encoding a glycosyltransferase then there would be made a contacting *in vivo* in a recombinant host cell of FK and/or KA with a glycosyltransferase.
[0046] In e.g. above discussed WO2004/111254A1 (Poalis A/S) such *in vivo* contacting of different aglycon compounds in different recombinant host cells were made and the skilled person would know how to perform such contacting *in vivo* in a recombinant host cell of a relevant aglycon (here flavokermesic acid (FK) and/or kermesic acid (KA)) and a recombinantly expressed glycosyltransferase.
[0047] As discussed above - it is believed that the herein disclosed novel glycosyltransferase represents the first time

that there has been disclosed an isolated glycosyltransferase capable of glycosylating flavokermesic acid aglycon and/or kermesic acid aglycon.
It is believed that relevant partial sequences of herein disclosed novel glycosyltransferase of SEQ ID NO:2 may be recombinantly introduced into another glycosyltransferase sequence in order to construct a new hybrid glycosyltransferase sequence capable of glucosylating flavokermesic acid and/or kermesic acid. Such GTs with reduced k_M or increased V_{max} may prove important in securing rapid glucosylation of the substrates that may show toxic effects inhibiting yeast growth if accumulating at high levels (Esben Halkjaer Hansen et al. Phytochemistry 70(4): 473-482). Likewise, if so desired it is envisioned possible to modify the substrate specificity towards glucosylation of earlier pathway intermediates.

[0048] Accordingly, a further aspect of the present invention relates to a method for constructing a novel isolated hybrid glycosyltransferase polypeptide capable of:

(I): conjugating glucose to flavokermesic acid (FK); and/or
(II): conjugating glucose to kermesic acid (KA),

wherein the method comprises following steps:

(i): inserting a polynucleotide sequence encoding a fragment of an amino acid sequence which has at least 70% identity with amino acids 1 to 515 of SEQ ID NO:2 (preferably a fragment of an amino acid sequence which has at least 90% identity with amino acids 1 to 515 of SEQ ID NO:2, more preferably a fragment of an amino acid sequence which has at least 99% identity with amino acids 1 to 515 of SEQ ID NO:2) wherein the fragment comprises at least 75 amino acids (preferably at least 100 amino acids, more preferably at least 150 amino acids and even more preferably at least 468 amino acids), into another polynucleotide sequence derived from a glycosyltransferase in order to thereby construct a novel recombinant hybrid polynucleotide sequence;
(ii): expressing the novel hybrid polypeptide which is encoded by the novel recombinant hybrid polynucleotide sequence of step (i);
(iii): isolating the expressed novel hybrid polypeptide of step (ii);
(iv): testing if the isolated novel hybrid polypeptide of step (iii) is capable of:

(I): conjugating glucose to flavokermesic acid (FK); and/or
(II): conjugating glucose to kermesic acid (KA); and

(v) if positive in test of step (iv) then has been constructed the novel isolated hybrid glycosyltransferase polypeptide capable of:

(I): conjugating glucose to flavokermesic acid (FK); and/or
(II): conjugating glucose to kermesic acid (KA).

DEFINITIONS

[0049] All definitions of relevant terms are in accordance with what would be understood by the skilled person in relation to the relevant technical context.

[0050] The term "aglycon" denotes non-carbohydrate part of the corresponding glycosylated form of the aglycon. When the sugar is glucose the aglycon may be termed aglucon. Further, when the sugar is glucose the term glucosylated may be used instead of glycosylated. When the aglycon is glycosylated at a hydroxy group there is generally created a so-called O-glycosidic bond - i.e. a so-called O-Glycoside (or O-Glucoside if the sugar is glucose). When the aglycon is glycosylated by a carbon-carbon linkage it may be termed a C-glycosidic bond - i.e. a so-called C-Glycoside (or C-Glucoside if the sugar is glucose).

[0051] The term "glycoside" denotes a compound which on hydrolysis can give a sugar and a non-sugar (aglycon) residue, e. g. glucosides can give glucose and galactosides can give galactose.

[0052] The term "glycosyltransferase" denotes an enzyme capable of conjugating a nucleotide activated sugar to a compound (e.g. an aglycon compound). The sugar may e.g. be D and L isomers of galactose, glucosamine, N-acetylglusamine, xylose, glucuronic acid, rhamnose, arabinose, mannose or glucose. Alternatively the sugar may be a carbohydrate derivative such as e. g. inositol, olivose, rhodinose and etc available as nucleotide diphosphates. Further the sugar may e.g. be a monosaccharide, a disaccharide or a trisaccharide. In the case of oligo-and polysaccharides the sugars are linked one by one, by e. g. involving use of one or several glycosyltransferases. If the sugar is glucose the glycosyltransferase may be termed a glucosyltransferase.
When the glycosyltransferase conjugates a nucleotide activated sugar to a compound via a C-glycosidic bond it may

be termed a C-glycosyltransferase.

When the glycosyltransferase conjugates a sugar to an aglycon via an O-glycosidic bond it may be termed an O-glycosyltransferase.

[0053] The term "hybridizes" in relation to a polynucleotide which hybridizes under at least medium stringency conditions with (i) nucleotides 1 to 1548 of SEQ ID NO:1 or (ii) a complementary strand of (i) relates to the nucleotide sequence which hybridizes to a labeled nucleic acid probe corresponding to the nucleotide sequence shown in SEQ ID NO:1 or its complementary strand under medium to very high stringency conditions. Molecules to which the nucleic acid probe hybridizes under these conditions can be detected using e.g. X-ray film.

Herein relevant hybridization stringency conditions are stringency conditions that the skilled person normally would understand are relevant - i.e. for medium stringency conditions the conditions that skilled person would understand are medium stringency conditions. The skilled person knows relevant hybridization stringency conditions - see e.g. (J. Sambrook, E.F. Fritsch, and T. Maniatus, 1989, Molecular Cloning, A Laboratory Manual, 2d edition, Cold Spring Harbor, New York).

According to the art - for long probes of at least 100 nucleotides in length, very low to very high stringency conditions are defined as prehybridization and hybridization at 42°C in 5X SSPE, 0.3% SDS, 200 μg/ml sheared and denatured salmon sperm DNA, and either 25% formamide for very low and low stringencies, 35% formamide for medium and medium-high stringencies, or 50% formamide for high and very high stringencies, following standard Southern blotting procedures for 12 to 24 hours optimally.

For long probes of at least 100 nucleotides in length, the carrier material is finally washed three times each for 15 minutes using 2X SSC, 0.2% SDS preferably at least at 45°C (very low stringency), more preferably at least at 50°C (low stringency), more preferably at least at 55°C (medium stringency), more preferably at least at 60°C (medium-high stringency), even more preferably at least at 65°C (high stringency), and most preferably at least at 70°C (very high stringency).

[0054] The term "*in vitro*" (Latin: in glass) relates to studies that are conducted using components of an organism that have been isolated from their usual biological surroundings in order to permit a more detailed or more convenient analysis than can be done with whole organisms.These experiments are commonly called "test tube experiments". In contrast, *in vivo* studies are those that are conducted with living organisms in their normal intact state.

[0055] The term "*in vivo*" (Latin for "within the living") relates to experimentation using a whole, living organism as opposed to a partial or dead organism, or an *in vitro* ("within the glass", e.g., in a test tube or petri dish) controlled environment.

[0056] The term "isolated polynucleotide" essentially relates herein to that the polynucleotide is isolated from its natural environment - said in other words that the polynucleotide preparation is essentially free of other polynucleotide material with which it is natively associated. The polynucleotide sequence encoding the herein described isolated/cloned novel glycosyltransferase is shown in SEQ ID NO: 1 and it was isolated from the insect *Dactylopius coccus.* Accordingly, as understood by the skilled person- the term isolated polynucleotide does not cover the polynucleotide of SEQ ID NO: 1 when it is naturally present in the genome of *Dactylopius coccus.* The term "isolated polynucleotide" essentially relates to that the isolated polynucleotide is in a form suitable for use within genetically engineered protein production systems. Thus, an isolated polynucleotide contains at most 10%, preferably at most 8%, more preferably at most 6%, more preferably at most 5%, more preferably at most 4%, more preferably at most 3%, even more preferably at most 2%, most preferably at most 1%, and even most preferably at most 0.5% by weight of other polynucleotide material with which it is natively associated. The term "isolated polynucleotide" may herein alternatively be termed "cloned polynucleotide".

[0057] The term "isolated polypeptide" essentially relates herein to that the polypeptide is isolated from its natural environment. The novel glycosyltransferase polypeptide as shown in SEQ ID NO: 2 was isolated from the insect *Dactylopius coccus.* Accordingly, as understood by the skilled person - the term "isolated polypeptide" does not cover the glycosyltransferase polypeptide of SEQ ID NO: 2 when it is naturally present in the genome of *Dactylopius coccus.* The term "isolated polypeptide" denotes a polypeptide preparation which contains at most 10%, preferably at most 8%, more preferably at most 6%, more preferably at most 5%, more preferably at most 4%, at most 3%, even more preferably at most 2%, most preferably at most 1%, and even most preferably at most 0.5% by weight of other polypeptide material with which it is natively associated. The term "other polypeptide material with which it is natively associated" may in relation to the novel glycosyltransferase polypeptide as shown in SEQ ID NO: 2 be seen in relation to other polypeptide material with which it is natively associated in *Dactylopius coccus.* In some case - it may be preferred that the "isolated polypeptide" refers to a polypeptide which is at least 20% pure, preferably at least 40% pure, more preferably at least 60% pure, even more preferably at least 80% pure, most preferably at least 90% pure, and even most preferably at least 95% pure, as determined by SDS-PAGE.

[0058] The term "nucleic acid construct" as used herein refers to a nucleic acid molecule, either single- or double-stranded, which is isolated from a naturally occurring gene or which is modified to contain segments of nucleic acids in a manner that would not otherwise exist in nature. The term nucleic acid construct is synonymous with the term "expression

cassette" when the nucleic acid construct contains the control sequences required for expression of a coding sequence of the present invention. As known in the art control, sequences include all components, which are necessary or advantageous for the expression of a polynucleotide encoding a polypeptide of the present invention. Each control sequence may be native or foreign to the nucleotide sequence encoding the polypeptide. Such control sequences include, but are not limited to, a leader, polyadenylation sequence, propeptide sequence, promoter, signal peptide sequence, and transcription terminator. At a minimum, the control sequences include a promoter, and transcriptional and translational stop signals. The control sequences may be provided with linkers for the purpose of introducing specific restriction sites facilitating ligation of the control sequences with the coding region of the nucleotide sequence encoding a polypeptide.

[0059] The term "recombinant expression vector" relates to recombinant expression vectors comprising a polynucleotide of the present invention, a promoter, and transcriptional and translational stop signals. The various nucleic acids and control sequences described above may be joined together to produce a recombinant expression vector which may include one or more convenient restriction sites to allow for insertion or substitution of the nucleotide sequence encoding the polypeptide at such sites.

[0060] The term "recombinant host cell" should herein be understood according to the art. As known in the art, recombinant polynucleotide (e.g. DNA) molecules are polynucleotide (e.g. DNA) molecules formed by laboratory methods of genetic recombination (such as molecular cloning) to bring together genetic material from multiple sources, creating sequences that would not otherwise be found in biological organisms. As understood by the skilled person - a recombinant host cell comprises recombinant polynucleotide (e.g. DNA) molecules and a recombinant host cell will therefore not be understood as covering a natural wildtype cell, such as e.g. a natural wildtype *Dactylopius coccus* cell.

[0061] The term "Sequence Identity" relates to the relatedness between two amino acid sequences or between two nucleotide sequences.

For purposes of the present invention, the degree of sequence identity between two amino acid sequences is determined using the Needleman-Wunsch algorithm (Needleman and Wunsch, 1970, J. Mol. Biol. 48: 443-453) as implemented in the Needle program of the EMBOSS package (EMBOSS: The European Molecular Biology Open Software Suite, Rice et al., 2000, Trends Genet. 16: 276-277), preferably version 3.0.0 or later. The optional parameters used are gap open penalty of 10, gap extension penalty of 0.5, and the EBLOSUM62 (EMBOSS version of BLOSUM62) substitution matrix. The output of Needle labeled "longest identity" (obtained using the -nobrief option) is used as the percent identity and is calculated as follows:

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

[0062] For purposes of the present invention, the degree of 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, *supra*), preferably version 3.0.0 or later. The optional parameters used are gap open penalty of 10, gap extension penalty of 0.5, and the EDNAFULL (EMBOSS version of NCBI NUC4.4) substitution matrix. The output of Needle labeled "longest identity" (obtained using the - nobrief option) is used as the percent identity and is calculated as follows:

(Identical Deoxyribonucleotides x 100)/(Length of Alignment – Total Number of Gaps in Alignment).

[0063] Embodiments of the present invention are described below, by way of examples only.

DRAWINGS

[0064]

Figure 1: Schematic presentation of the relevant glycosyltransferase activity of the herein described isolated/cloned novel glycosyltransferase of SEQ ID NO:2 - as illustrated in the figure it was found to be able to conjugate glucose to the aglycons flavokermesic acid (FK) and kermesic acid (KA).

Figure 2: Production of glucosides of flavokermesic acid and kermesic acid using OsCGT and SbUGT85B1. LC-MS analyses of glucosylated products formed in assays containing UDP-glucose and flavokermesicc acid (FK) or

kermesic acid (KA). Crude lysate from the *E.coli* strain Xjb (negative control) incubated with FK (**A**) or KA (**C**). Crude lysate from Xjb cells expressing OsCGT incubated with FK (**B**) or KA (**D**). Crude lysate from Xjb cells expressing SbUGT85B1 incubated with FK (**E**) or KA (**F**). FK (**G**) and KA (**H**) substrates alone. The total ion chromatograms (TIC) and extracted ion chromatograms for *m/z* 313[M-H]$^-$, *m/z* 329[M-H]$^-$, *m/z* 475 [M-H]$^-$, *m/z* 491 [M-H]$^-$, corresponding to FK, KA, FK-monoglucoside, and KA-monoglucoside are indicated. Peak retention times are indicated in minutes.

DETAILED DESCRIPTION OF THE INVENTION

A novel isolated glycosyltransferase polypeptide as described herein

[0065] When there herein is referred to an isolated glycosyltransferase polypeptide as described herein there is referred to an isolated glycosyltransferase polypeptide of the first aspect of the invention and/or herein relevant embodiments thereof.

[0066] As discussed above - the term "isolated polypeptide" essentially relates to that the polypeptide is isolated from its natural environment. The herein described novel glycosyltransferase polypeptide as shown in SEQ ID NO: 2 was isolated from the insect *Dactylopius coccus.* Accordingly, as understood by the skilled person in the present context - the term "isolated polypeptide" does not cover the glycosyltransferase polypeptide of SEQ ID NO: 2 when it is naturally present in the genome of *Dactylopius coccus.*

[0067] Preferably, the isolated glycosyltransferase polypeptide as described herein denotes a polypeptide preparation which contains at most 10%, preferably at most 8%, more preferably at most 6%, more preferably at most 5%, more preferably at most 4%, at most 3%, even more preferably at most 2%, most preferably at most 1%, and even most preferably at most 0.5% by weight of other polypeptide material with which it is natively associated.

[0068] As understood by the skilled person, the term "other polypeptide material with which it is natively associated" may in relation to the novel glycosyltransferase polypeptide as shown in SEQ ID NO: 2 be seen as relation to other polypeptide material with which it is natively associated in *Dactylopius coccus.*

[0069] In some case - it may be preferred that the isolated glycosyltransferase polypeptide as described herein refers to a polypeptide which is at least 20% pure, preferably at least 40% pure, more preferably at least 60% pure, even more preferably at least 80% pure, most preferably at least 90% pure, and even most preferably at least 95% pure, as determined by SDS-PAGE.

[0070] Based on e.g. the sequence information disclosed herein - it is routine work for the skilled person to obtain an isolated glycosyltransferase polypeptide as described herein.
This may e.g. be done by recombinant expression in a suitable recombinant host cell according to procedures known in the art.
Accordingly, it is not believed necessary to describe such standard known recombinant expression procedures in many details herein.

[0071] Preferably, the isolated glycosyltransferase polypeptide as described herein is capable of:

(I): conjugating glucose to flavokermesic acid (FK); and
(II): conjugating glucose to kermesic acid (KA).

[0072] A preferred embodiment relates to wherein the glycosyltransferase polypeptide of the first aspect is:

(a) a polypeptide comprising an amino acid sequence which has at least 80% identity with amino acids 1 to 515 of SEQ ID NO:2; more preferably
(a) a polypeptide comprising an amino acid sequence which has at least 90% identity with amino acids 1 to 515 of SEQ ID NO:2; even more preferably (a) a polypeptide comprising an amino acid sequence which has at least 95% identity with amino acids 1 to 515 of SEQ ID NO:2; and most preferably
(a) a polypeptide comprising an amino acid sequence which has at least 98% identity with amino acids 1 to 515 of SEQ ID NO:2.

[0073] It may be preferred that the glycosyltransferase polypeptide of the first aspect is a polypeptide comprising an amino acid sequence with amino acids 1 to 515 of SEQ ID NO:2.

[0074] A preferred embodiment relates to wherein the glycosyltransferase polypeptide of the first aspect is:

(b) a polypeptide comprising an amino acid sequence which has at least 80% identity with amino acids 20 to 468 of SEQ ID NO:2; more preferably
(b) a polypeptide comprising an amino acid sequence which has at least 90% identity with amino acids 20 to 468

of SEQ ID NO:2; even more preferably
(b) a polypeptide comprising an amino acid sequence which has at least 95% identity with amino acids 20 to 468 of SEQ ID NO:2; and most preferably
(b) a polypeptide comprising an amino acid sequence which has at least 98% identity with amino acids 20 to 468 of SEQ ID NO:2.

[0075] It may be preferred that the glycosyltransferase polypeptide of the first aspect is a polypeptide comprising an amino acid sequence with amino acids 20 to 468 of SEQ ID NO:2.
[0076] A preferred embodiment relates to wherein the glycosyltransferase polypeptide of the first aspect is:

(c) a polypeptide which is encoded by a polynucleotide which hybridizes under at least medium-high stringency conditions with (i) nucleotides 1 to 1548 of SEQ ID NO:1 or (ii) a complementary strand of (i); more preferably
(c) a polypeptide which is encoded by a polynucleotide which hybridizes under at least high stringency conditions with (i) nucleotides 1 to 1548 of SEQ ID NO:1 or (ii) a complementary strand of (i); and even more preferably
(c) a polypeptide which is encoded by a polynucleotide which hybridizes under at least very stringency conditions with (i) nucleotides 1 to 1548 of SEQ ID NO:1 or (ii) a complementary strand of (i).

[0077] It is routine work for the skilled person to make a variant of an isolated glycosyltransferase polypeptide as described herein - i.e. a variant, wherein e.g. one or more amino acids of e.g. SEQ ID NO:2 have been modified/altered. Further - as known to the skilled person if such variant changes are not too drastic it will be plausible that the enzyme would maintain its relevant GT activity.
[0078] A preferred embodiment relates to wherein the glycosyltransferase polypeptide of the first aspect is:

(a) a polypeptide comprising an amino acid sequence with amino acids 1 to 515 of SEQ ID NO:2 or a variant thereof, wherein the variant comprises an alteration at one or more (several) positions of SEQ ID NO:2 and wherein the variant comprises less than 50 alterations, more preferably less than 40 alterations, even more preferably less than 20 alterations and most preferably less than 10 alterations.

[0079] In a preferred embodiment the term "an alteration at one or more (several) positions of SEQ ID NO:2" refers to 1 to 10 alterations in SEQ ID NO:2.
[0080] According to the art - the term "variant" means herein a peptide having the relevant GT activity comprising an alteration, i.e., a substitution, insertion, and/or deletion, at one or more (several) positions. A substitution means a replacement of an amino acid occupying a position with a different amino acid; a deletion means removal of an amino acid occupying a position; and an insertion means adding 1-3 amino acids adjacent to an amino acid occupying a position. The amino acid may be natural or unnatural amino acids - for instance, substitution with e.g. a particularly D-isomers (or D-forms) of e.g. D-alanine could theoretically be possible.
[0081] In a preferred embodiment the glycosyltransferase polypeptide of the first aspect is a GT which is membrane bound or insoluble in water.

Isolated polynucleotide comprising a nucleotide sequence which encodes the glycosyltransferase polypeptide as described herein

[0082] As discussed above - a second aspect of the present invention relates to an isolated polynucleotide comprising a nucleotide sequence which encodes the polypeptide of the first aspect and/or herein relevant embodiments thereof.
[0083] The term "isolated polynucleotide" may herein alternatively be termed "cloned polynucleotide".
[0084] As discussed above - the term "isolated polynucleotide" essentially relates to that the polynucleotide is isolated from its natural environment - said in other words that the polynucleotide preparation is essentially free of other polynucleotide material with which it is natively associated. The polynucleotide sequence encoding the described isolated/cloned novel glycosyltransferase is shown in SEQ ID NO: 1 and it was isolated from the insect *Dactylopius coccus.* Accordingly, as understood by the skilled person - the term isolated polynucleotide does not cover the polynucleotide of SEQ ID NO: 1 when it is naturally present in the genome of *Dactylopius coccus.*
[0085] The term "isolated polynucleotide" essentially relates to that the isolated polynucleotide is in a form suitable for use within genetically engineered protein production systems.
[0086] Thus, an isolated polynucleotide contains at most 10%, preferably at most 8%, more preferably at most 6%, more preferably at most 5%, more preferably at most 4%, more preferably at most 3%, even more preferably at most 2%, most preferably at most 1%, and even most preferably at most 0.5% by weight of other polynucleotide material with which it is natively associated.
[0087] Based on e.g. the sequence information disclosed herein - it is routine work for the skilled person to obtain an

isolated polynucleotide as described herein.
This may e.g. be done by recombinant expression in a suitable recombinant host cell according to procedures known in the art.
Accordingly, it is not believed necessary to describe such standard known recombinant expression procedures in many details herein.

A nucleic acid construct comprising the isolated polynucleotide as described herein

[0088] As discussed above - a third aspect of the present invention relates to a nucleic acid construct comprising the isolated polynucleotide of the second aspect and/or herein relevant embodiments thereof operably linked to one or more control sequences that direct the production of the polypeptide in an expression host.
[0089] According to the art - the term "nucleic acid construct" as used herein refers to a nucleic acid molecule, either single- or double-stranded, which is isolated from a naturally occurring gene or which is modified to contain segments of nucleic acids in a manner that would not otherwise exist in nature.
[0090] The term nucleic acid construct is synonymous with the term "expression cassette" when the nucleic acid construct contains the control sequences required for expression of a coding sequence of the present invention. As known in the art control sequences include all components, which are necessary or advantageous for the expression of a polynucleotide encoding a polypeptide of the present invention.
[0091] Based on e.g. the sequence information disclosed herein - it is routine work for the skilled person to make a relevant nucleic acid construct - for instance, based on the prior art the skilled person knows numerous different suitable control sequences for the expression of a polynucleotide encoding a polypeptide of the present invention.
Accordingly, it is not believed necessary to describe such standard known technical elements in many details herein.

A recombinant expression vector comprising the nucleic acid construct as described herein

[0092] As discussed above - a fourth aspect of the present invention relates to a recombinant expression vector comprising the nucleic acid construct of the third aspect and/or herein relevant embodiments thereof.
[0093] According to the art - the term "recombinant expression vector" relates to recombinant expression vectors comprising a polynucleotide of the present invention, a promoter, and transcriptional and translational stop signals. The various nucleic acids and control sequences described above may be joined together to produce a recombinant expression vector which may include one or more convenient restriction sites to allow for insertion or substitution of the nucleotide sequence encoding the polypeptide at such sites.
[0094] Based on e.g. the sequence information disclosed herein - it is routine work for the skilled person to make a relevant recombinant expression vector - for instance, based on the prior art the skilled person knows numerous different suitable promoter, and transcriptional and translational stop signals.
Accordingly, it is not believed necessary to describe such standard known technical elements in many details herein.

A recombinant host cell comprising the nucleic acid construct as described herein

[0095] As discussed above - a fifth aspect of the present invention relates to a recombinant host cell comprising the nucleic acid construct of the third aspect and/or herein relevant embodiments thereof.
[0096] The term "recombinant host cell" should herein be understood according to the art. As known in the art, recombinant polynucleotide (e.g. DNA) molecules are polynucleotide (e.g. DNA) molecules formed by laboratory methods of genetic recombination (such as molecular cloning) to bring together genetic material from multiple sources, creating sequences that would not otherwise be found in biological organisms. As understood by the skilled person - a recombinant host cell comprises recombinant polynucleotide (e.g. DNA) molecules and a recombinant host cell will therefore not be understood as covering a natural wildtype cell, such as e.g. a natural wildtype *Dactylopius coccus* cell.
[0097] Based on e.g. the sequence information disclosed herein - it is routine work for the skilled person to make a relevant recombinant host cell - for instance, based on the prior art the skilled person knows numerous different suitable recombinant host cells that for years have been used as recombinant host cells for e.g. expression of different polypeptides of interest.
[0098] The recombinant host cell may be any suitable cell such as any eukaryotic cell [e.g. mammalian cells (such as e.g. Chinese hamster ovary (CHO) cells) or a plant cell] or any prokaryotic cell.
[0099] Particularly preferred is wherein the recombinant host cell is a plant cell producing flavokermesic acid/kermesic acid or other related compound such as e.g. rhubarb plant cell.
[0100] Preferably the recombinant host cell is a cell selected from the group consisting of a filamentous fungal cell and a microorganism cell.
[0101] Filamentous fungi include all filamentous forms of the subdivision Eumycota and Oomycota (as defined by

Hawksworth et al. , 1995, supra). The filamentous fungi are characterized by a vegetative mycelium composed of chitin, cellulose, glucan, chitosan, mannan, and other complex polysaccharides. Vegetative growth is by hyphal elongation and carbon catabolism is obligately aerobic. In contrast, vegetative growth by yeasts such as *Saccharomyces cerevisiae* is by budding of a unicellular thallus and carbon catabolism may be fermentative.

[0102] It may be preferred that the filamentous fungal cell is a cell of a species of, but not limited to, *Acremonium, Aspergillus, Fusarium, Humicola, Mucor, Myceliophthora, Neurospora, Penicillium, Thielavia, Tolypocladium,* and *Trichoderma* or a teleomorph or synonym thereof.

[0103] A preferred *Aspergillus* cell is *Aspergillus niger* or *Aspergillus oryzae.*

[0104] A preferred microorganism cell herein is a microorganism cell selected from the group consisting of a yeast cell and prokaryotic cell.

[0105] A preferred yeast cell is a yeast cell selected from the group consisting of Ascomycetes, Basidiomycetes and fungi imperfecti. Preferably a yeast cell selected from the group consisting of Ascomycetes.

[0106] Preferred Ascomycetes yeast cell selected from the group consisting of *Ashbya, Botryoascus, Debaryomyces, Hansenula, Kluveromyces, Lipomyces, Saccharomyces spp* e. g. *Saccharomyces cerevisiae, Pichia spp., Schizosaccharomyces* spp.

[0107] A preferred yeast cell is a yeast cell selected from the group consisting of *Saccharomyces spp,* e. g. *Saccharomyces cerevisiae,* and *Pichia spp.*

[0108] A preferred prokaryotic cell is selected from the group consisting of *Bacillus, Streptomyces, Corynebacterium, Pseudomonas,* lactic acid bacteria and an *E. coli* cell.

[0109] A preferred *Bacillus* cell is *B. subtilis, B. amyloliquefaciens* or *B. licheniformis.*
A preferred *Streptomyces* cell is *S. setonii* or *S. coelicolor.*
A preferred *Corynebacterium* cell is *C. glutamicum.*
A preferred *Pseudomonas* cell is *P. putida* or *P. fluorescens.*

A method for producing flavokermesic acid (FK) glycoside and/or kermesic acid (KA) glycoside

[0110] As discussed above - a sixth aspect of the present invention relates to a method for producing flavokermesic acid (FK) glycoside and/or kermesic acid (KA) glycoside, wherein the method comprises following steps:

(A): contacting *in vitro* or *in vivo* in a recombinant host cell comprising a glycosyltransferase gene encoding a glycosyltransferase:

(a1): flavokermesic acid (FK) with a glycosyltransferase capable of glycosylating the flavokermesic acid under suitable conditions wherein there is produced the flavokermesic acid glycoside; and/or
(a2): kermesic acid (KA) with a glycosyltransferase capable of glycosylating the kermesic acid under suitable conditions wherein there is produced the kermesic acid glycoside.

[0111] It may be preferred that the recombinant host cell in step (A) is a recombinant host cell comprising a recombinant glycosyltransferase gene encoding a glycosyltransferase.

[0112] Preferably, the glycosyltransferase in step (a2) is a glucosyltransferase and there thereby in step (a2) is produced kermesic acid glucoside, preferably wherein the produced kermesic acid glucoside is Carminic acid (Figure 1 herein shows the structure of Carminic acid).

[0113] It may be preferred that the glycosyltransferase in step (a1) is a glucosyltransferase and there thereby in step (a1) is produced flavokermesic acid glucoside, preferably wherein the produced flavokermesic acid glucoside is the compound DcII (Figure 1 herein shows the structure of the compound DcII).

[0114] When the produced compound in step (a1) is DcII it may be preferred to use this DcII as an intermediate to make Carminic acid.
This may be done by chemical synthesis and the skilled person knows based on his common general knowledge how to do this.
Alternatively, it may be done enzymatically by e.g. using a suitable oxygenase. An example of a suitable oxygenase is cytochrome P450 superfamily of monooxygenases (officially abbreviated as CYP) enzyme. Other examples are flavine monooxygenases or different types of dioxygenases, this list not to be considered excluding the involvement of other classes of enzymes
As known in the art - the most common reaction catalyzed by cytochromes P450 is a monooxygenase reaction, e.g., insertion of one atom of oxygen into a substrate.

[0115] As understood by the skilled person in the present context - the terms flavokermesic acid (FK) and/or kermesic acid (KA) aglycons of step (a) of the method of the sixth aspect as discussed herein should be understood as the FK and/or KA specific compounds shown in Figure 1 and equivalent analogs of these specific compounds with minor

substituents (e.g. a FK methyl ester).
As understood in by the skilled person- if FK methyl ester is used as aglycon in step (a) of the method of the sixth aspect then there will via the glycosylation step be generated a FK methyl ester glycoside, which by routine removal of the methyl group will generate DcII - accordingly FK methyl ester aglycon may be seen as equivalent to FK aglycon in relation to the method of the sixth aspect as discussed herein.
In step (a) of the method of the sixth aspect is specified that there is used a glycosyltransferase capable of glycosylating FK and/or KA - accordingly it is understood that the GT must be capable of doing this.
[0116] It may be preferred to purify the glycoside produced in step (A) - i.e. in step (a1) and/or in step (a2).
Accordingly it may be preferred that the method of the sixth aspect comprises a further step (B) with following steps:

(B): purifying the produced glycoside in step (a1) and/or in step (a2) whereby one gets a composition, wherein at least 5% w/w (preferably at least 10% w/w, more preferably at least 50% w/w and most preferably at least 80% w/w) of the compounds in the composition is the produced flavokermesic acid glycoside and/or kermesic acid glycoside.

[0117] The skilled person knows how to purify such glycoside compounds and it may be done according to the art.
[0118] The purifying step (B) may be particularly preferred when:

- the produced glycoside in step (a2) is Carminic acid;
- the produced glycoside in step (a1) is compound DcII; and/or
- the produced glycoside in step (a1) is compound DcII and it is used as an intermediate to make Carminic acid.

[0119] As discussed herein - in working Examples there was made a contacting *in vitro* of flavokermesic acid (FK) and/or kermesic acid (KA) with the glycosyltransferase of SEQ ID NO:2. It may be seen as routine work for the skilled person to perform such an *in vitro* contacting step.
[0120] The glycosyltransferase of SEQ ID NO:2 was recombinantly expressed in a yeast cell (see working Example herein) - accordingly, in a working Example herein there was made a recombinant yeast host cell comprising a recombinant glycosyltransferase gene encoding a glycosyltransferase of SEQ ID NO:2.
It is believed that if flavokermesic acid (FK) and/or kermesic acid (KA) would be added under suitable condition to a fermentation medium the FK and/or KA compound(s) would enter into e.g. yeast cells fermented in the medium - accordingly, if e.g. the yeast cells are recombinant yeast host cells comprising a recombinant glycosyltransferase gene encoding a glycosyltransferase then there would be made a contacting *in vivo* in a recombinant host cell of FK and/or KA with a glycosyltransferase.
[0121] In a preferred embodiment the contacting in step (A) is *in vivo* and the recombinant host cell is a yeast cell, preferably wherein the yeast cell is selected from the group consisting of *Saccharomyces spp* (e.g. *Saccharomyces cerevisiae)* and *Pichia spp.*
[0122] Above is described preferred recombinant host cells - these preferred recombinant host cells may also be preferred recombinant host cells in relation to the method of the sixth aspect of the present invention.
[0123] In the present context - it may be said that it is within the skilled person's common knowledge to identify a suitable recombinant host cell to perform the *in vivo* contacting step (A) of the method of the sixth aspect and it is not believed that it is necessary to describe this in many details herein.
[0124] Above is discussed that preferred recombinant host cells may e.g. be a microorganism cell or a filamentous fungal cell - these cells may be preferred recombinant host cells in relation to the method of the sixth aspect.
[0125] It may be possible to make a recombinant host cell (e.g. a recombinant host microorganism cell) which comprises a gene encoding a product involved in the biosynthesis pathway leading to flavokermesic acid (FK) and/or kermesic acid (KA) and such a recombinant host cell could be preferred herein.
[0126] Accordingly, it may be preferred that the contacting in step (A) is contacting *in vivo* in a recombinant host cell comprising a recombinant glycosyltransferase gene encoding a glycosyltransferase and a gene encoding a product involved in the biosynthesis pathway leading to flavokermesic acid (FK) and/or kermesic acid (KA).
[0127] As discussed in working Example herein - the GT of SEQ ID NO:2 is membrane bound or hydrophobic/insoluble *in vivo* and in water. When production cells or fractions of cells containing the membrane bound GT are separated from the product (e.g. carminic acid), the GT can essentially not be present in the fraction where the more soluble product/hydrophilic product is present. This is an advantage for obtaining a final product (e.g. carminic acid product/composition) which is essentially totally free of the recombinant GT.
Because the substrates glycosylated by the GT may be hydrophobic aglycons, the aglycons would be expected to partly accumulate in membranes and other hydrophobic parts of the production cells. By the use of a membrane bound GT a more efficient glycosylation of hydrophobic compounds present in e.g. membranes is obtained
[0128] Accordingly, in a preferred embodiment the glycosyltransferase used in the method of the sixth aspect is a GT which is membrane bound or insoluble in water.

[0129] In a preferred embodiment - the glycosyltransferase in step (A) of the method of the sixth aspect is a glycosyltransferase of the first aspect and/or herein relevant embodiments thereof.

[0130] As discussed herein - the identified data/results of working Examples 4 show that herein relevant GT enzymes can be identified in e.g. Sorghum and rice plants.

The Sorghum polypeptide sequence (Genbank ID number: AAF17077.1) is shown as SEQ ID NO: 4 herein.

The rice polypeptide sequence (Genbank ID number: CAQ77160.1) is shown as SEQ ID NO: 5 herein.

[0131] It may be relevant that the glycosyltransferase in step (A) of the method of the sixth aspect is a glycosyltransferase comprising an amino acid sequence which has at least 70% (preferably at least 80%, more preferably at least 90% and even more preferably at least 98%) identity with amino acids 1 to 492 of SEQ ID NO:4.

[0132] It may be relevant that the glycosyltransferase in step (A) of the method of the sixth aspect is a glycosyltransferase comprising an amino acid sequence which has at least 70% (preferably at least 80%, more preferably at least 90% and even more preferably at least 98%) identity with amino acids 1 to 471 of SEQ ID NO:5.

EXAMPLES

Example 1 - *Cloning of D. coccus GT and test of its FK and KA activity*

Materials and methods

Purification of DNA and mRNA

[0133] Fresh frozen *Dactylopius coccus* (were obtained from Lanzarote). Fresh frozen *Porphyrophora polonica* were obtained from Poland. The frozen insects were ground into powder under liquid nitrogen and DNA/RNA was purified: DNA was purified using DNeasy Blood & Tissue kit (Qiagen), according to the protocol of the supplier. RNA was purified using RNeasy mini kit (Qiagen) according to the protocol of the supplier.

[0134] Eucaryote mRNA was made into cDNA using RT^2 Easy First Strand Kit (Qiagen) according to the protocol of the supplier using poly-dT priming of the revers transcriptase reaction.

Sequencing of DNA and RNA:

[0135] DNA and cDNA were sent for sequencing at BGI (Shenzen, China) for sequencing using 100 bp paired-end Illumina technology according to the protocol of Illumina at a coverage of approximately 60-100X and the output in fastq data format.

Analysis of DNA and RNA/cDNA sequences:

[0136] The obtained fastq-sequences of DNA and RNA/cDNA were imported into Genomic Workbench version 5.4 (CLC-bio, Århus, Denmark) and assembled using the *de novo* assembling algorithm into contigs. Output files were exported as FASTA format.

[0137] RNA/cDNA FASTA files were then imported into IOGMA v. 10 (Genostar, Grenoble, France) and putitative genes were identified using the "hidden Markov-Matrix-based prokaryote gene-finder.

[0138] Putative genes were annotated using BLAST (basic local alignment sequence tool) against genbank (NCBI) using as well the nucleotide sequence as the translated protein sequence. The putative genes were also annotated by similarity comparison to PFAM databases of protein families.

Preparation of protein fractions from D. coccus

[0139] Three grams of fresh *D. coccus insects* were homogenized in 120 mL of isolation buffer [350 mM sucrose, 20 mM Tricine (pH 7.9), 10 mM NaCl, 5 mM DTT, 1 mM PMSF) containing 0.3 g polyvinylpolypyrrolidone. The homogenate was filtered through a nylon cloth (22 μm mesh) and centrifuged for (10 min, 10,000 x g at 4 °C). The supernatant was centrifuged (1h, 105,000xg, at 4 °C), yielding a soluble and a membrane bound protein fraction. The soluble protein fraction was concentrated to 1 mL and buffer-exchanged with 20 mM Tricine (pH 7.9), 5 mM DTT by using Amicon ultrafugation-3K devices (Millipore). The membrane bound protein pellet was washed 3 times by resuspending the pellet in 60 mL of 20 mM Tricine (pH 7.9), 5 mM DTT using a marten paintbrush followed by re-centrifugation. The membrane bound protein pellet was finally resuspended in 1 mL 20 mM Tricine (pH 7.9), 5 mM DTT. The soluble protein fraction and the membrane bound protein fraction were analyzed for glycosylation activity.

Purification of a flavokermesic acid/kermesic acid-specific GT activity from *D. coccus* membrane proteins

[0140] A membrane bound protein fraction isolated from 3 g fresh *D. coccus insects* was solubilized by adding 1 % (v/v) Triton x-100 (reduced form) and gently stirring for 1.5 h in the cold. The Triton x-100 treated solution was centrifuged (1h, 105,000xg, at 4 °C) and the supernatant was isolated and applied to a column packed with 2 mL Q-sepharose Fast flow (Pharmacia). The column was washed in 4 mL of buffer A [20 mM Tricine (pH 7.9), 0.1 % (v/v) Triton x-100 (reduced form), 50 mM NaCl] and proteins were eluted with 20 mM Tricine (pH 7.9), 0.1 % (v/v) Triton x-100 (reduced form)] using a discontinuous NaCl gradient from 100 mM-500 mM,(with 50 mM increments). 0.5-ml-fractions were collected, desalted, analyzed by SDS-PAGE and monitored for glucosylation activity using the described radiolabeled glucosylation enzyme assay. A fraction containing enriched flavokermesic acid/kermesic acid-specific GT activity was subjected to peptide mass fingerprinting analysis.

Enzyme assays and glucoside product detection

[0141] Glucosylation of flavokermesic acid and kermesic acid was performed in assay mixtures of 60 μL containing 20 mM Tricine(pH 7.9), 3.3 μm UDP[14C]glucose and 20 uL protein extract (membrane bound or soluble protein). Reactions were incubated for 0.5 h at 30 °C and terminated by adding 180 μL of methanol. Samples were centrifuged at 16,000xg for 5 min at 4 °C and supernatant was spotted on TLC plates (silica gel 60 F254 plates; Merck). Assay products were resolved in dichloromethane:methanol:formic acid (7:2:2, by volume). Radiolabeled products were visualized using a STORM 840 PhosphorImager (Molecular Dynamics, http://www.moleculardynamics.com).

Expression of codon optimized DcUGT2, DcUGT4 and DcUGT5 in S. cerevisiae

[0142] A synthetic codon optimized version of DcUGT2 and two other putative GT sequences from the D. coccus transcriptome termed DcUGT4 and DcUGT5 for yeast expression was purchased from GenScript with flanking gateway recombination attL sites. The synthetic fragments were used as PCR templates with specific primers to generate the corresponding C-terminal StrepII-tagged versions. The six gene constructs (tagged and non-tagged fragments) were cloned into the gateway destination plasmid pYES-DEST52 (Invitrogen) using LR clonaseII enzyme mix. The six pYES-DEST52 plasmid constructs were transformed separately into the Invsc1 yeast strain (Invitrogen) and positive transformants were verified by PCR. Heterologous protein production was performed according to the instructions of the pYES-DEST52 gateway vector (Invitrogen). Production of heterologous StrepII-tagged protein was verified by western blotting using anti-Strep antibody. A membrane bound protein fraction was prepared from verified yeast transformants as described in (D. Pompon, B. Louerat, A. Bronine, P. Urban, Yeast expression of animal and plant P450s in optimized redox environments, Methods Enzymol. 272 (1996) 51-64) and screened for glucosylation activity towards flavokermesic acid/kermesic acid. The yeast optimized sequence is shown in SEQ ID NO: 3 herein.

LC-MS-MS

[0143] LC/MS was performed on an Agilent Q-TOF with the following HPLC system:

Column Kinetix 2.6μ XB-C18 100A (100 x 4.60 mm, Phenomenex). Solvent A (900 ml deionized water, 100 ml methanol and 50 ml formic acid). Solvent B (700 ml methanol, 300 ml deionized water and 50 ml formic acid). Flow 0.8 ml/min. 35°C.

[0144] Gradient elution. 0-1 min 100 % A. Linear gradient to 83 % A 3 min. linear gradient to 63 % A 6 min, linear gradient to 45 % A 9 min, linear gradient to 27 % A 12 min, linear gradient to 10 % A 15 min, linear gradient to 3 % A 17 min, linear gradient to 2 % A 19 min, linear gradient to 0 % A 20 min, 0 %A 22 min, linear gradient to 100 % A 25 min. Retention times were 7.6 min for carminic acid, 7.8 min for DC II, 13.7 min for flavokermesic acid and 13.9 min for kermesic acid.

Results:

[0145] The ability to glycosylate flavokermesic acid/kermesic acid using C14-UDP-glucose as a substrate was detected in homogenized *D. coccus* insects. The activity was shown to be membrane bound and the activity was purified and the purified proteins were submitted to proteomics analysis. It was shown that the enzymatic activity was to come from a polypeptide with a sequence corresponding to our candidate gene DcUGT2.

[0146] As discussed above - the herein relevant glycosyltransferase enzyme of SEQ ID NO: 2 may herein be termed "DcUGT2".

[0147] The amino acid sequence of DcUGT2 shows less than 45% homology to any known glycosyl transferase.

[0148] Knowing that cloning the wildtype sequence into yeast had given no relevant enzyme activity, we redesigned the nucleotide sequence of DcUGT2 to a sequence coding for the same polypeptide but using nucleotide codons optimized for *S. cerevisiae,* a process called codon optimization (the *S. cerevisiae* optimized sequence is shown as SEQ ID No. 3 herein). Subsequently the codon optimized sequence of DcUGT2 was cloned and expressed in yeast. The heterologous yeast strain contains a membrane bound enzyme activity capable of glucosylating kermesic acid and flavokermesic acid.

[0149] After obtaining peptide mass fingerprinting data from a D. coccus protein fraction enriched with GT activity towards flavokermesic acid/kermesic acid, we matched the peptide masses to the transcriptomic dataset and identified three putative UGTs (DcUGT2, DcUGT4 and DcUGT5) .

[0150] Heterologous expression of the three candidates in yeast revealed that only one of these UGTs, namely DcUGT2 was responsible for the observed glucosylation activity towards flavokermesic acid/kermesic acid in the *D. coccus* protein fraction.

[0151] A viscozyme treatment of the generated C-14 radiolabelled glucoside, showed that it was resistant towards hydrolysis, further suggesting that the DcUGT2 is a C-GT, responsible for producing DCII and carminic acid.

[0152] A LC-MS-MS showed formation of products with the same retention time, spectrum, molecular mass and molecular degradation pattern as DcII and carminic acid respectively.

Conclusion

[0153] The result of this example 1 demonstrated that it was not an easy task to isolate/clone the herein relevant glycosyltransferase enzyme of SEQ ID NO: 2, which may herein be termed "DcUGT2".

[0154] For instance, the identified gene sequences of the genome and transcriptome of *D. coccus* insects were analyzed for similarity to herein relevant public known C-glycosyltransferase sequences and the result was negative in the sense that none of the identified gene sequences of the genome/transcriptome showed herein significant similarity to publicly known herein relevant C-glycosyltransferase sequences.

[0155] However, even though the bioinformatic sequence similarity analysis could be said to indicate that the genome of *Dactylopius coccus* would not comprise a gene encoding a herein relevant glycosyltransferase - the present inventors continued to investigate the matter and the present inventors identified a *Dactylopius coccus* extract (including extracts of the endosymbionts present in *D. coccus*) with herein relevant GT activity and by a combination of herein relevant purification and testing steps the inventors were finally able to get a relatively pure fraction/composition wherefrom it was possible to obtain several partial amino acid sequences of possible GT enzyme candidates.

[0156] The present inventors tested the activity of the herein described isolated/cloned novel glycosyltransferase of SEQ ID NO: 2 (DcUGT2) and found that it was able to conjugate glucose to the aglycons flavokermesic acid (FK) and kermesic acid (KA) - see Figure 1 herein.

Example 2 *Testing KA GT activity* of prior art known UrdGT2

[0157] As discussed above - the UrdGT2 is described in the article Baig et al (Angew Chem Int Ed Engl. 2006 Nov 27;45(46):7842-6).

[0158] As discussed above - this article describes that UrdGT2 is capable of glycosylating different aglycon molecules that may be considered structurally similar to the herein relevant Kermesic acid (KA) and Flavokermesic acid (FK) aglycons.

[0159] A codon optimized synthetic version of UrdGT2 for *E. coli* expression was cloned and recombinantly expressed in *E. coli.* A crude soluble protein extract containing the the recombinant UrdGT2 was obtained - i.e. an extract comprising the UrdGT2

[0160] The UrdGT2 GT activity was analyzed *in vitro* using either UDP-glucose or TDP-glucose as a sugar donor and FA/KA as aglycone substrates. No activity was detected towards these aglycons - i.e. no herein relevant GT activity was identified in relation to these aglycons.

[0161] However, it was confirmed that the recombinant UrdGT2 was active, as demonstrated by the *in vitro* formation of a C14-radiolabelled glucoside derived from the glucosylation of an unidentified compound in the crude *E. coli* extract.

Example 3 *GT activity in Aloe plant and Haworthia* plant

Isolation and test of GT activity from *Aloe*

[0162]

1) The plant was washed from soil particles and separated into : A) Root, B) Green leaf tissue and C) the gel material

from the leaf
2) 5 g of tissue was frozen immediately in liquid nitrogen and ground in a cold mortar with a pestle to a fine powder.
3) 20 mL of cold extraction buffer [20 mM Tricine-HCl, 10 mM NaCl, 5 mM DTT, 1 mM PMSF, pH 7.9] containing a Complete protease inhibitor without EDTA (Roche), 0.1 % (w/v) proteamine sulfate and 0.5 g of PVPP were added to the powder and vortexed.
4) The homogenate was gently stirred at 4 °C for 10 min and then centrifuged at 12,000 xg at 4 °C for 5 min.
5) Supernatant was isolated and 1 mL of 2% (w/v) proteamine sulfate in 20 mM Tricine-HCl, pH 7.9 was added dropwise over 2 min at 4 °C under constant stirring.
6) The supernatant was filtered through 2 pieces of nylon mesh. The filtered supernatant was then centrifuged at at 12,000 xg at 4 °C for 5 min.
7) The supernatant was isolated and ultracentrifuged at 110,000 xg at 4 °C for 1 h.
8) The soluble protein fraction (supernatant) was isolated and buffer-exchanged 5 times with 20 mM Tricine-HCl, pH 7.9 containing 5 mM DTT using a Amicon Ultra centrifugal filter device-3K (Millipore)
9) 20 μL soluble protein extract was incubated in a total reaction volume of 60 μL containing UDP-glucose (1.25 mM final conc.) and either FK (50 μM final conc.), KA (50 μM final conc) or MeO-FK/EtO-FK (50 μM/50 μM final conc) for 2 h at 30 °C, shaking at 650 rpm.
10) Enzyme reactions were terminated with 180 μL cold methanol and filtered through a 0.45 micron filter and subjected to HPLC-MS analysis.

Table 1. Glucosides formed in *in vitro* glucosylation assays using enzyme extracts from *Aloe.*

m/z $[M-H]^-$ values / Aloe Soluble protein	475 m/z $[M-H]^-$FK-monoglc	491 m/z $[M-H]^-$KA-monoglc	489 m/z MeOFK-m $[M-H]^-$onoglc	503 m/z $[M-H]^-$EtOFK-monoglc
Leaf	3.73	3.71	5.81	6.63
Gel				
Root		3.71		

[0163] Crude soluble enzyme extracts of three *Aloe* tissues, green leaf material (Leaf), gel material from the leaf (Gel) and Root were tested for glucosylation activity towards flavokermesic acid (FK), kermesic acid (KA), methyl ester of flavokermesic acid (MeOFK) and ethyl ester of flavokermesic acid (EtOFK). Numbers correspond to retention times (min) after HPLC-MS separation of the novel glucosides formed *in vitro* (Table 1).
[0164] The m/z values 475 and 491 are the same m/z values as are obtained for DcII and CA, respectively, solubilized in similar solutions. Both m/z values are 162 (m/z value of glucose in a glucoside) higher than the m/z values of the FK and KA indicating that the glucose moiety from UDP-glucose in the reaction buffer has been transferred to the aglycone by a GT in the extract. The m/z [M-H] values 489 and 503 are also 162 higher than the m/z values obtained with MeOFK and EtOFK, respectively, indicating that a glucose unit has been added to both MeOFK and EtOFK by a GT present in the extract.

Isolation and test of GT activity from *Haworthia limifolia*

[0165] The procedure was as described for *Aloe* but plant tissue analyzed were following: A) Green leaf tissue, B) Gel material from the leaf, C) Base tissue (pink part between root and stem) and D) Root tissue.
[0166] Crude soluble enzyme extracts of four *Haworthia limifolia* tissues, green leaf material (Leaf), gel material from the leaf (Gel), pink tissue between root and stem (Base) and Root were tested for glucosylation activity towards flavokermesic acid (FK), kermesic acid (KA), methyl ester of flavokermesic acid (MeOFK) and ethyl ester of flavokermesic acid (EtOFK). Numbers correspond to retention times (min) after HPLC-MS separation of the novel glucosides formed *in vitro* (Table 2).

m/z [M-H]⁻ values ---------- Haworthia Soluble protein	475 m/z [M-H]⁻ FK-monoglc	491 m/z [M-H]⁻ KA-monoglc	489 m/z [M-H]⁻ MeOFK-monoglc	503 m/z [M-H]⁻ EtOFK-monoglc
Leaf	3.73	3.71	5.81	6.63
Gel				
Base	3.73	3.71	5.81	6.63
Root	3.73	3.71	5.81	6.63

Table 2. Glucosides formed in *in vitro* glucosylation assays using enzyme extracts from *Haworthia limifolia.*

[0167] The m/z values 475 and 491 are the same m/z values as are obtained for DcII and CA, respectively, solubilized in similar solutions. Both m/z values are 162 (m/z value of glucose in a glucoside) higher than the m/z values of the FK and KA indicating that the glucose moety from UDP-glucose in the reaction buffer has been transferred to the aglycone by a GT in the extract. The m/z [M-H] values 489 and 503 are also 162 higher than the m/z values obtained with MeOFK and EtOFK, respectively, indicating that a glucose unit has been added to both MeOFK and EtOFK by a GT present in the extract.

Conclusion

[0168] The results of this example demonstrate that herein relevant glycosyltransferase (GT) enzymes can be identified in *Aloe* plants and *Haworthia* plants.

[0169] Said in other words, *Aloe* plants and *Haworthia* plants comprise a glycosyltransferase which is capable of glycosylating flavokermesic acid in order to produce flavokermesic acid glycoside; and/or capable of glycosylating kermesic acid in order to produce kermesic acid glycoside.

Example 4 GT activity in Sorghum and rice plant

[0170] As known the art - Sorghum and rice plants comprise glycosyltransferases.

[0171] As known in the art - some of the Sorghum and rice glycosyltransferases may glycosylate low molecular weight aglycone compounds.

[0172] The in the art described glycosyltransferases from Sorghum and rice plants have significant less than 70% identity with amino acids 1 to 515 of SEQ ID NO:2 as disclosed herein.

[0173] It is not known in the art if glycosyltransferases of Sorghum and/or rice plants would be a herein relevant glycosyltransferase - i.e. a glycosyltransferase which is capable of glycosylating flavokermesic acid in order to produce flavokermesic acid glycosides; and/or capable of glycosylating kermesic acid in order to produce kermesic acid glycosides.

[0174] The known glycosyltransferases from Sorghum (*Sorghum bicolor*), SbUGT85B1, with Genbank ID number AF199453.1 (nucleotide seq.) / AAF17077.1 (polypeptide seq) and rice (*Oryza sativa*), OsCGT, with Genbank ID number FM179712.1 (nucleotide seq.) / CAQ77160.1 (polypeptide seq) were expressed in *E.coli* strain Xjb and crude *E.coli* proteins extracts were prepared and tested for glucosylation activity on the substrates kermesic acid and flavokermisic acid as described by Kannangara *et al.* (2011) and Augustin *et al.* (2012).

[0175] Figure 2 shows LC-MS analyses of glucosylated products formed in assays containing crude lysate of *E.coli* strain Xjb expressing either SbUGT85B1 or OsCGT, UDP-glucose and flavokermesicc acid (FK) or kermesic acid (KA). As a negative control crude extract from the *E.coli* strain Xjb was used in the assays.

[0176] There were identified KA glycosides (491 m/z [M-H] - the m/z[M-H] value of CA) for both glycosyltransferases and FK glycosides (475 m/z [M-H] the m/z[M-H] value of DcII) for OsCGT.

Conclusion

[0177] The result of this example demonstrated that herein relevant glycosyltransferase (GT) enzymes can be identified in Sorghum and/or rice plants.

[0178] Said in other words, Sorghum and/or rice plants comprise a glycosyltransferase which is capable of glycosylating flavokermesic acid in order to produce flavokermesic acid glycoside; and/or capable of glycosylating kermesic acid in order to produce kermesic acid glycoside.

Example 5 Use of endogenous GT gene or GT activity

[0179] As known in the art glycosyltransferases able to glycosylate low molecular weight are present in a lot of different organisms. A method to contact the glycosyltransferase of the cells of an organism with a low molecular weight compound is to introduce one or more genes directing the biosynthesis of the low molecular weight compound and thus enabling the cells to glycosylate the low molecular weight compound. The low molecular weight compound may be e.g. flavokermesic acid or kermersic acid or decorated versions of these molecules.

[0180] One or more genes directing the biosynthesis of flavokermesic acid or kermesic acid or decorated version of these molecules are introduced into a glycosyltransferase containing organism, e.g. the tobacco plant, *Nicotiana benthamiana*.

When the gene/genes is/are transiently expressed according to the methods described in D'Aoust *et al.* (2008) in e.g. plant tissue the low molecular weight compound or compounds is/are produced. Cells stably expressing the gene/genes are produced and selected according to the methods described in Gelvin (2003).

In cells containing either stably expressed and/or transiently expressed gene/genes the low molecular weight compounds come into contact with the endogenous glycosyltransferases, resulting in the formation of one or more glycosides of flavokermesic acid, kermesic acid or decorated versions of these molecules.

The presence of the glycosides is demonstrated by the extraction and the analytical methods described in example 3. Samples are prepared for LC/MS by the method for extraction described by Rauwald and Sigler (1994).

Conclusion

[0181] The results of this example demonstrate that endogenous glycosyltransferases present in the cells of a recombinant organism can be used to convert flavokermesic acid, kermesic acid or decorated versions of these molecules into glycosides when a gene/genes directing the biosynthesis of the aglycons are introduced into the organism.

[0182] Said in other words introduction of a gene or genes directing the biosynthesis of flavokermesic acid, kermesic acid, decorated versions of these molecules, or related low molecular weight compounds is a method to bring the low molecular weight compound in contact with glycosyltransferases and thus a method to produced glycosides of flavokermesic acid, kermesic acid or decorated version of these compounds.

REFERENCES

[0183]

1: US5424421 (European Colour, published 1995)

2: WO2006/056585A1 (Chr. Hansen A/S)

3: Stathopoulou et al (Analytica Chimica Acta 804 (2013) 264- 272)

4: Zagrobelny et al (Cyanogenic glucosides and plant-insect interactions; Phytochemistry. 2004 Feb;65(3):293-306)

5: Geuder et al (Journal of Chemical Ecology, Vol. 23, No. 5, 1997)

6: Genta et al, (Potential role for gut microbiota in cell wall digestion and glucoside detoxification in Tenebrio molitor larvae), Journal of Insect Physiology 52 (2006) 593-601

7: WO2004/111254A1 (Poalis A/S)

8: Gutmann et al (Pure Appl. Chem, 2013-07-09)

9: Pompon et al (Methods Enzymol. 272 (1996):51-64)

10: Baig et al (Angew Chem Int Ed Engl. 2006 Nov 27;45(46):7842-6)

11: Radominska-Pandya A, Bratton SM, Redinbo MR, Miley MJ. Drug Metab Rev. 2010 Feb;42(1):133-44)

12: Plant Physiology, November 2008, Vol. 148, pp. 1295-1308

13: Esben Halkjaer Hansen et al. Phytochemistry 70(4): 473-482

14: Kannangara et al. (Plant Journal. 68 (2011): 287-301)

15: Augustin et al. (Plant Physiology. 160 (2012): 1881-1895)

16: D'Aoust et al. (Methods Mol Biol 483 (2009): 41-50)

17: Gelvin (Microbiol Mol Biol Rev 67(1) (2003): 16-37)

18: Rauwald and Sigler (Phytochemical Analysis 5 (1994):266-270)

SEQUENCE LISTING

[0184]

<110> Chr. Hansen A/S

<120> Glycosyltransferase - glycosylating flavokermesic acid and/or kermesic acid

<130> P4547PC00

<150> EP13198110.2
<151> 2013-12-18

<160> 5

<170> PatentIn version 3.5

<210> 1
<211> 1548
<212> DNA
<213> Dactylopius coccus costa

<220>
<221> CDS
<222> (1)..(1548)

<400> 1

```
atg gaa ttt cgt tta cta atc ctg gct ctt ttt tct gta ctt atg agt      48
Met Glu Phe Arg Leu Leu Ile Leu Ala Leu Phe Ser Val Leu Met Ser
1               5                   10                  15

act tca aac gga gca gaa att tta gct ctt ttc cct att cac ggt atc      96
Thr Ser Asn Gly Ala Glu Ile Leu Ala Leu Phe Pro Ile His Gly Ile
            20                  25                  30

agt aat tat aat gtt gct gaa gca ctg ctg aag acc tta gct aac cgg     144
Ser Asn Tyr Asn Val Ala Glu Ala Leu Leu Lys Thr Leu Ala Asn Arg
        35                  40                  45

ggt cat aat gtt aca gtt gtc aca tct ttt cct caa aaa aaa cct gta     192
Gly His Asn Val Thr Val Val Thr Ser Phe Pro Gln Lys Lys Pro Val
    50                  55                  60

cct aat ttg tac gaa att gac gta tct gga gct aaa ggc ttg gct act     240
Pro Asn Leu Tyr Glu Ile Asp Val Ser Gly Ala Lys Gly Leu Ala Thr
65                  70                  75                  80

aat tca ata cat ttt gaa aga tta caa acg att att caa gat gta aaa     288
Asn Ser Ile His Phe Glu Arg Leu Gln Thr Ile Ile Gln Asp Val Lys
                85                  90                  95

tcg aac ttt aag aac atg gta cga ctt agc aga aca tac tgt gag att     336
Ser Asn Phe Lys Asn Met Val Arg Leu Ser Arg Thr Tyr Cys Glu Ile
            100                 105                 110

atg ttt tct gat ccg agg gtt ttg aac att cga gac aag aaa ttc gat     384
Met Phe Ser Asp Pro Arg Val Leu Asn Ile Arg Asp Lys Lys Phe Asp
        115                 120                 125

ctc gta ata aac gcc gta ttt ggc agt gac tgc gat gcc gga ttc gca     432
Leu Val Ile Asn Ala Val Phe Gly Ser Asp Cys Asp Ala Gly Phe Ala
    130                 135                 140
```

```
tgg aaa agt caa gct cca ttg att tca att ctc aat gct aga cat act      480
Trp Lys Ser Gln Ala Pro Leu Ile Ser Ile Leu Asn Ala Arg His Thr
145                 150                 155                 160

cct tgg gcc cta cac aga atg gga aat cca tca aat cca gcg tat atg      528
Pro Trp Ala Leu His Arg Met Gly Asn Pro Ser Asn Pro Ala Tyr Met
                165                 170                 175

cct gtc att cat tct aga ttt cct gta aaa atg aat ttc ttc caa aga      576
Pro Val Ile His Ser Arg Phe Pro Val Lys Met Asn Phe Phe Gln Arg
            180                 185                 190

atg ata aat acg ggt tgg cat ttg tat ttt ctg tac atg tac ttt tat      624
Met Ile Asn Thr Gly Trp His Leu Tyr Phe Leu Tyr Met Tyr Phe Tyr
        195                 200                 205

tat ggt aat gga gaa gat gcc aac aaa atg gcg aga aaa ttt ttt ggc      672
Tyr Gly Asn Gly Glu Asp Ala Asn Lys Met Ala Arg Lys Phe Phe Gly
    210                 215                 220

aac gac atg ccc gac ata aat gaa atg gtt ttt aat aca tct tta tta      720
Asn Asp Met Pro Asp Ile Asn Glu Met Val Phe Asn Thr Ser Leu Leu
225                 230                 235                 240

ttc gta aat act cac ttt tcg gtt gat atg cca tat cct ttg gtt cca      768
Phe Val Asn Thr His Phe Ser Val Asp Met Pro Tyr Pro Leu Val Pro
                245                 250                 255

aac tgc att gaa ata gga gga ata cat gta aaa gag cca caa cca ctg      816
Asn Cys Ile Glu Ile Gly Gly Ile His Val Lys Glu Pro Gln Pro Leu
            260                 265                 270

cct ttg gaa ata caa aaa ttc atg gac gaa gca gaa cat ggg gtc att      864
Pro Leu Glu Ile Gln Lys Phe Met Asp Glu Ala Glu His Gly Val Ile
        275                 280                 285

ttc ttc acg cta gga tca atg gtg cgt act tcc acg ttt cca aat caa      912
Phe Phe Thr Leu Gly Ser Met Val Arg Thr Ser Thr Phe Pro Asn Gln
    290                 295                 300

act att caa gca ttt aag gaa gct ttt gcc gaa tta cct caa aga gtc      960
Thr Ile Gln Ala Phe Lys Glu Ala Phe Ala Glu Leu Pro Gln Arg Val
305                 310                 315                 320

tta tgg aag ttt gag aat gaa aat gag gat atg cca tca aat gta ctc     1008
Leu Trp Lys Phe Glu Asn Glu Asn Glu Asp Met Pro Ser Asn Val Leu
                325                 330                 335

ata agg aaa tgg ttt cca caa aat gat ata ttc ggt cat aag aat atc     1056
Ile Arg Lys Trp Phe Pro Gln Asn Asp Ile Phe Gly His Lys Asn Ile
            340                 345                 350

aaa gca ttc att agt cac ggt gga aat tct gga gct ctg gag gct gtt     1104
Lys Ala Phe Ile Ser His Gly Gly Asn Ser Gly Ala Leu Glu Ala Val
        355                 360                 365

cat ttc gga gta ccg ata att gga att cct tta ttc tac gat cag tac     1152
His Phe Gly Val Pro Ile Ile Gly Ile Pro Leu Phe Tyr Asp Gln Tyr
    370                 375                 380

agg aat att ttg agt ttc gtt aaa gaa ggt gtt gcc gtt ctt ttg gat     1200
Arg Asn Ile Leu Ser Phe Val Lys Glu Gly Val Ala Val Leu Leu Asp
385                 390                 395                 400
```

```
gtg aat gat ctg acg aaa gat aat att tta tct tct gtc agg act gtt      1248
Val Asn Asp Leu Thr Lys Asp Asn Ile Leu Ser Ser Val Arg Thr Val
                405                 410                 415

gtt aat gat aag agt tac tca gaa cgt atg aaa gca ttg tca caa cta      1296
Val Asn Asp Lys Ser Tyr Ser Glu Arg Met Lys Ala Leu Ser Gln Leu
            420                 425                 430

ttc cga gat cga cca atg agt cct ctt gac aca gct gtt tac tgg aca      1344
Phe Arg Asp Arg Pro Met Ser Pro Leu Asp Thr Ala Val Tyr Trp Thr
        435                 440                 445

gaa tat gtc atc cgc cat aga gga gcc cat cac ctc aag acc gct ggc      1392
Glu Tyr Val Ile Arg His Arg Gly Ala His His Leu Lys Thr Ala Gly
    450                 455                 460

gca ttt ttg cat tgg tat cag tat tta ctt ttg gac gtt att acc ttc      1440
Ala Phe Leu His Trp Tyr Gln Tyr Leu Leu Leu Asp Val Ile Thr Phe
465                 470                 475                 480

tta tta gtc aca ttc tgc gct ttt tgt ttt att gtg aaa tat ata tgt      1488
Leu Leu Val Thr Phe Cys Ala Phe Cys Phe Ile Val Lys Tyr Ile Cys
                485                 490                 495

aaa gct ctc att cat cat tat tgg agc agt tcg aaa tct gaa aag ttg      1536
Lys Ala Leu Ile His His Tyr Trp Ser Ser Ser Lys Ser Glu Lys Leu
            500                 505                 510

aaa aaa aat taa                                                      1548
Lys Lys Asn
        515
```

<210> 2
<211> 515
<212> PRT
<213> Dactylopius coccus costa

<400> 2

```
Met Glu Phe Arg Leu Leu Ile Leu Ala Leu Phe Ser Val Leu Met Ser
1               5                   10                  15

Thr Ser Asn Gly Ala Glu Ile Leu Ala Leu Phe Pro Ile His Gly Ile
            20                  25                  30

Ser Asn Tyr Asn Val Ala Glu Ala Leu Leu Lys Thr Leu Ala Asn Arg
        35                  40                  45

Gly His Asn Val Thr Val Val Thr Ser Phe Pro Gln Lys Lys Pro Val
    50                  55                  60

Pro Asn Leu Tyr Glu Ile Asp Val Ser Gly Ala Lys Gly Leu Ala Thr
65                  70                  75                  80

Asn Ser Ile His Phe Glu Arg Leu Gln Thr Ile Ile Gln Asp Val Lys
                85                  90                  95
```

Ser Asn Phe Lys Asn Met Val Arg Leu Ser Arg Thr Tyr Cys Glu Ile
100 105 110

Met Phe Ser Asp Pro Arg Val Leu Asn Ile Arg Asp Lys Lys Phe Asp
115 120 125

Leu Val Ile Asn Ala Val Phe Gly Ser Asp Cys Asp Ala Gly Phe Ala
130 135 140

Trp Lys Ser Gln Ala Pro Leu Ile Ser Ile Leu Asn Ala Arg His Thr
145 150 155 160

Pro Trp Ala Leu His Arg Met Gly Asn Pro Ser Asn Pro Ala Tyr Met
165 170 175

Pro Val Ile His Ser Arg Phe Pro Val Lys Met Asn Phe Phe Gln Arg
180 185 190

Met Ile Asn Thr Gly Trp His Leu Tyr Phe Leu Tyr Met Tyr Phe Tyr
195 200 205

Tyr Gly Asn Gly Glu Asp Ala Asn Lys Met Ala Arg Lys Phe Phe Gly
210 215 220

Asn Asp Met Pro Asp Ile Asn Glu Met Val Phe Asn Thr Ser Leu Leu
225 230 235 240

Phe Val Asn Thr His Phe Ser Val Asp Met Pro Tyr Pro Leu Val Pro
245 250 255

Asn Cys Ile Glu Ile Gly Gly Ile His Val Lys Glu Pro Gln Pro Leu
260 265 270

Pro Leu Glu Ile Gln Lys Phe Met Asp Glu Ala Glu His Gly Val Ile
275 280 285

Phe Phe Thr Leu Gly Ser Met Val Arg Thr Ser Thr Phe Pro Asn Gln
290 295 300

Thr Ile Gln Ala Phe Lys Glu Ala Phe Ala Glu Leu Pro Gln Arg Val
305 310 315 320

Leu Trp Lys Phe Glu Asn Glu Asn Glu Asp Met Pro Ser Asn Val Leu
325 330 335

Ile Arg Lys Trp Phe Pro Gln Asn Asp Ile Phe Gly His Lys Asn Ile

340 345 350

Lys Ala Phe Ile Ser His Gly Gly Asn Ser Gly Ala Leu Glu Ala Val
355 360 365

His Phe Gly Val Pro Ile Ile Gly Ile Pro Leu Phe Tyr Asp Gln Tyr
370 375 380

Arg Asn Ile Leu Ser Phe Val Lys Glu Gly Val Ala Val Leu Leu Asp
385 390 395 400

Val Asn Asp Leu Thr Lys Asp Asn Ile Leu Ser Ser Val Arg Thr Val
405 410 415

Val Asn Asp Lys Ser Tyr Ser Glu Arg Met Lys Ala Leu Ser Gln Leu
420 425 430

Phe Arg Asp Arg Pro Met Ser Pro Leu Asp Thr Ala Val Tyr Trp Thr
435 440 445

Glu Tyr Val Ile Arg His Arg Gly Ala His His Leu Lys Thr Ala Gly
450 455 460

Ala Phe Leu His Trp Tyr Gln Tyr Leu Leu Leu Asp Val Ile Thr Phe
465 470 475 480

Leu Leu Val Thr Phe Cys Ala Phe Cys Phe Ile Val Lys Tyr Ile Cys
485 490 495

Lys Ala Leu Ile His His Tyr Trp Ser Ser Ser Lys Ser Glu Lys Leu
500 505 510

Lys Lys Asn
515

<210> 3
<211> 1548
<212> DNA
<213> Artificial Sequence

<220>
<223> Yeast Optimized Sequence

<400> 3

atggaattca gattgttgat attggccttg ttctccgtat tgatgtctac ctctaatggt 60
gccgaaatct tggctttatt ccctattcat ggtatatcta actacaacgt agctgaagca 120
ttgttgaaga ctttggctaa cagaggtcac aacgttaccg ttgtaacttc atttccacaa 180
aagaaaccag ttcctaattt gtacgaaatt gatgtatcag gtgcaaaggg tttagccaca 240

aactccatcc atttcgaaag attgcaaacc atcatccaag atgtcaagag taacttcaag 300
aacatggtta gattgtctag aacatactgt gaaatcatgt tctcagaccc aagagttttg 360
aacatcagag ataaaaagtt tgacttggtt ataaacgccg tattcggttc agattgcgac 420
gctggttttg catggaaaag tcaagctcct ttaatatcta tcttgaatgc cagacataca 480
ccatgggctt tgcacagaat gggtaatcct tccaacccag catatatgcc tgtaatccat 540
agtagattcc cagtcaagat gaatttcttt caaagaatga taaacaccgg ttggcactta 600
tactttttgt acatgtactt ctactacggt aatggtgaag atgctaacaa aatggcaaga 660
aagtttttcg gtaatgatat gcctgacata aacgaaatgg tttttaacac ctccttgttg 720
ttcgtaaaca ctcatttcag tgtcgatatg ccataccctt tagtcccaaa ctgtatcgaa 780
atcggtggta tccatgttaa ggaaccacaa cctttgccat tggaaatcca aaagtttatg 840
gatgaagcag aacatggtgt aatctttttc accttgggta gtatggtcag aacttctaca 900
ttccctaatc aaactattca agcctttaaa gaagccttcg ctgaattacc acaaagagtt 960
ttgtggaagt tcgaaaacga aaacgaagat atgccttcca acgttttgat cagaaagtgg 1020
ttcccacaaa acgacatctt cggtcataag aacatcaagg ctttcatttc acacggtggt 1080
aattccggtg ccttggaagc tgtccatttc ggtgttccta tcataggtat cccattgttt 1140
tatgatcaat acagaaacat cttgtctttc gttaaagaag gtgtagctgt cttgttggat 1200
gtaaacgact taactaagga taacatcttg tcttcagtta gaacagtcgt taacgacaag 1260
tcatactccg aaagaatgaa ggcattgtct caattgttta gagatagacc tatgtcacca 1320
ttagacacag ctgtttattg gaccgaatac gtaattagac atagaggtgc acatcactta 1380
aaaactgcag gtgccttttt gcactggtat caatacttgt tgttggatgt catcacattt 1440
ttgttggtta cattctgtgc attctgcttc atcgttaagt acatctgcaa ggccttaatc 1500
catcactact ggtccagttc taaatctgaa aagttgaaaa agaattaa 1548

<210> 4
<211> 492
<212> PRT
<213> Sorghum bicolor

<400> 4

Met Gly Ser Asn Ala Pro Pro Pro Pro Thr Pro His Val Val Leu Val
1 5 10 15

Pro Phe Pro Gly Gln Gly His Val Ala Pro Leu Met Gln Leu Ala Arg
20 25 30

Leu Leu His Ala Arg Gly Ala Arg Val Thr Phe Val Tyr Thr Gln Tyr
35 40 45

Asn Tyr Arg Arg Leu Leu Arg Ala Lys Gly Glu Ala Ala Val Arg Pro
50 55 60

Pro Ala Thr Ser Ser Ala Arg Phe Arg Ile Glu Val Ile Asp Asp Gly
65 70 75 80

Leu Ser Leu Ser Val Pro Gln Asn Asp Val Gly Gly Leu Val Asp Ser
85 90 95

Leu Arg Lys Asn Cys Leu His Pro Phe Arg Ala Leu Leu Arg Arg Leu
100 105 110

Gly Gln Glu Val Glu Gly Gln Asp Ala Pro Pro Val Thr Cys Val Val
115 120 125

Gly Asp Val Val Met Thr Phe Ala Ala Ala Ala Ala Arg Glu Ala Gly
130 135 140

Ile Pro Glu Val Gln Phe Phe Thr Ala Ser Ala Cys Gly Leu Leu Gly
145 150 155 160

Tyr Leu His Tyr Gly Glu Leu Val Glu Arg Gly Leu Val Pro Phe Arg
165 170 175

Asp Ala Ser Leu Leu Ala Asp Asp Asp Tyr Leu Asp Thr Pro Leu Glu
180 185 190

Trp Val Pro Gly Met Ser His Met Arg Leu Arg Asp Met Pro Thr Phe
195 200 205

Cys Arg Thr Thr Asp Pro Asp Asp Val Met Val Ser Ala Thr Leu Gln
210 215 220

Gln Met Glu Ser Ala Ala Gly Ser Lys Ala Leu Ile Leu Asn Thr Leu
225 230 235 240

Tyr Glu Leu Glu Lys Asp Val Val Asp Ala Leu Ala Ala Phe Phe Pro
245 250 255

Pro Ile Tyr Thr Val Gly Pro Leu Ala Glu Val Ile Ala Ser Ser Asp
260 265 270

Ser Ala Ser Ala Gly Leu Ala Ala Met Asp Ile Ser Ile Trp Gln Glu
275 280 285

Asp Thr Arg Cys Leu Ser Trp Leu Asp Gly Lys Pro Ala Gly Ser Val

290 295 300

Val Tyr Val Asn Phe Gly Ser Met Ala Val Met Thr Ala Ala Gln Ala
305 310 315 320

Arg Glu Phe Ala Leu Gly Leu Ala Ser Cys Gly Ser Pro Phe Leu Trp
325 330 335

Val Lys Arg Pro Asp Val Val Glu Gly Glu Glu Val Leu Leu Pro Glu
340 345 350

Ala Leu Leu Asp Glu Val Ala Arg Gly Arg Gly Leu Val Val Pro Trp
355 360 365

Cys Pro Gln Ala Ala Val Leu Lys His Ala Ala Val Gly Leu Phe Val
370 375 380

Ser His Cys Gly Trp Asn Ser Leu Leu Glu Ala Thr Ala Ala Gly Gln
385 390 395 400

Pro Val Leu Ala Trp Pro Cys His Gly Glu Gln Thr Thr Asn Cys Arg
405 410 415

Gln Leu Cys Glu Val Trp Gly Asn Gly Ala Gln Leu Pro Arg Glu Val
420 425 430

Glu Ser Gly Ala Val Ala Arg Leu Val Arg Glu Met Met Val Gly Asp
435 440 445

Leu Gly Lys Glu Lys Arg Ala Lys Ala Ala Glu Trp Lys Ala Ala Ala
450 455 460

Glu Ala Ala Ala Arg Lys Gly Gly Ala Ser Trp Arg Asn Val Glu Arg
465 470 475 480

Val Val Asn Asp Leu Leu Leu Val Gly Gly Lys Gln
485 490

<210> 5
<211> 471
<212> PRT
<213> Oryza sativa

<400> 5

Met Pro Ser Ser Gly Asp Ala Ala Gly Arg Arg Pro His Val Val Leu
1 5 10 15

Ile Pro Ser Ala Gly Met Gly His Leu Val Pro Phe Gly Arg Leu Ala

20 25 30

Val Ala Leu Ser Ser Gly His Gly Cys Asp Val Ser Leu Val Thr Val
35 40 45

Leu Pro Thr Val Ser Thr Ala Glu Ser Lys His Leu Asp Ala Leu Phe
50 55 60

Asp Ala Phe Pro Ala Val Arg Arg Leu Asp Phe Glu Leu Ala Pro Phe
65 70 75 80

Asp Ala Ser Glu Phe Pro Gly Ala Asp Pro Phe Phe Leu Arg Phe Glu
85 90 95

Ala Met Arg Arg Ser Ala Pro Leu Leu Gly Pro Leu Leu Thr Gly Ala
100 105 110

Gly Ala Ser Ala Leu Ala Thr Asp Ile Ala Leu Thr Ser Val Val Ile
115 120 125

Pro Val Ala Lys Glu Gln Gly Leu Pro Cys His Ile Leu Phe Thr Ala
130 135 140

Ser Ala Ala Met Leu Ser Leu Cys Ala Tyr Phe Pro Thr Tyr Leu Asp
145 150 155 160

Ala Asn Ala Gly Gly Gly Gly Gly Val Gly Asp Val Asp Ile Pro Gly
165 170 175

Val Tyr Arg Ile Pro Lys Ala Ser Ile Pro Gln Ala Leu His Asp Pro
180 185 190

Asn His Leu Phe Thr Arg Gln Phe Val Ala Asn Gly Arg Ser Leu Thr
195 200 205

Ser Ala Ala Gly Ile Leu Val Asn Thr Phe Asp Ala Leu Glu Pro Glu
210 215 220

Ala Val Ala Ala Leu Gln Gln Gly Lys Val Ala Ser Gly Phe Pro Pro
225 230 235 240

Val Phe Ala Val Gly Pro Leu Leu Pro Ala Ser Asn Gln Ala Lys Asp
245 250 255

Pro Gln Ala Asn Tyr Met Glu Trp Leu Asp Ala Gln Pro Ala Arg Ser
260 265 270

Val Val Tyr Val Ser Phe Gly Ser Arg Lys Ala Ile Ser Arg Glu Gln
275 280 285

Leu Arg Glu Leu Ala Ala Gly Leu Glu Gly Ser Gly His Arg Phe Leu
290 295 300

Trp Val Val Lys Ser Thr Val Val Asp Arg Asp Asp Ala Ala Glu Leu
305 310 315 320

Gly Glu Leu Leu Asp Glu Gly Phe Leu Glu Arg Val Glu Lys Arg Gly
325 330 335

Leu Val Thr Lys Ala Trp Val Asp Gln Glu Glu Val Leu Lys His Glu
340 345 350

Ser Val Ala Leu Phe Val Ser His Cys Gly Trp Asn Ser Val Thr Glu
355 360 365

Ala Ala Ala Ser Gly Val Pro Val Leu Ala Leu Pro Arg Phe Gly Asp
370 375 380

Gln Arg Val Asn Ser Gly Val Val Ala Arg Ala Gly Leu Gly Val Trp
385 390 395 400

Ala Asp Thr Trp Ser Trp Glu Gly Glu Ala Gly Val Ile Gly Ala Glu
405 410 415

Glu Ile Ser Glu Lys Val Lys Ala Ala Met Ala Asp Glu Ala Leu Arg
420 425 430

Met Lys Ala Ala Ser Leu Ala Glu Ala Ala Ala Lys Ala Val Ala Gly
435 440 445

Gly Gly Ser Ser His Arg Cys Leu Ala Glu Phe Ala Arg Leu Cys Gln
450 455 460

Gly Gly Thr Cys Arg Thr Asn
465 470

Claims

1. An isolated glycosyltransferase polypeptide capable of:

(I): conjugating nucleotide activated glucose to flavokermesic acid (FK); and/or
(II): conjugating nucleotide activated glucose to kermesic acid (KA);

and wherein the glycosyltransferase polypeptide is at least one polypeptide selected from the group consisting of:

(a) a polypeptide comprising an amino acid sequence which has at least 70% identity with amino acids 1 to 515 of SEQ ID NO:2;
(b) a polypeptide comprising an amino acid sequence which has at least 70% identity with amino acids 20 to 468 of SEQ ID NO: 2;
(c) a polypeptide which is encoded by a polynucleotide that hybridizes under at least medium stringency conditions with (i) nucleotides 1 to 1548 of SEQ ID NO:1 or (ii) a complementary strand of (i); and
(d) a fragment of amino acids 1 to 515 of SEQ ID NO:2, which has the glycosyltransferase activity as specified in (I) and/or (II).

2. The isolated glycosyltransferase polypeptide of claim 1, wherein the isolated glycosyltransferase polypeptide of claim 1 is capable of:

(I): conjugating glucose to flavokermesic acid (FK); and
(II): conjugating glucose to kermesic acid (KA).

3. The isolated glycosyltransferase polypeptide of any of the preceding claims, wherein the glycosyltransferase polypeptide of claim 1 is:

(a) a polypeptide comprising an amino acid sequence which has at least 80% identity with amino acids 1 to 515 of SEQ ID NO:2.

4. The isolated glycosyltransferase polypeptide of any of the preceding claims, wherein the glycosyltransferase polypeptide of claim 1 is:

(a) a polypeptide comprising an amino acid sequence which has at least 95% identity with amino acids 1 to 515 of SEQ ID NO:2.

5. An isolated polynucleotide comprising a nucleotide sequence which encodes the polypeptide of any of the preceding claims.

6. A nucleic acid construct comprising the isolated polynucleotide of claim 5 operably linked to one or more control sequences that direct the production of the polypeptide in an expression host.

7. A recombinant expression vector comprising the nucleic acid construct of claim 6.

8. A recombinant host cell comprising the nucleic acid construct of claim 7.

9. The recombinant host cell of claim 8, wherein the recombinant host cell is a cell selected from the group consisting of a filamentous fungal cell and a microorganism cell, preferably wherein the microorganism cell is a yeast cell, preferably wherein the yeast cell is selected from the group consisting of *Saccharomyces spp* (e.g. *Saccharomyces cerevisiae),* and *Pichia spp.*

10. A method for producing flavokermesic acid (FK) glycoside and/or kermesic acid (KA) glycoside, wherein the method comprises following steps:

(A): contacting *in vitro* or *in vivo* in a recombinant host cell comprising a glycosyltransferase gene encoding a glycosyltransferase:

(a1): flavokermesic acid (FK) with a glycosyltransferase capable of glycosylating the flavokermesic acid under suitable conditions wherein there is produced the flavokermesic acid glycoside; and/or
(a2): kermesic acid (KA) with a glycosyltransferase capable of glycosylating the kermesic acid under suitable conditions wherein there is produced the kermesic acid glycoside, and

wherein the glycosyltransferase in step (A) is a glycosyltransferase of any of claims 1 to 4.

11. The method of claim 10, wherein

- the glycosyltransferase in step (a1) is a glucosyltransferase and there thereby in step (a1) is produced fla-

vokermesic acid glucoside and wherein the produced flavokermesic acid glucoside is the compound DcII; and/or - the glycosyltransferase in step (a2) is a glucosyltransferase and there thereby in step (a2) is produced kermesic acid glucoside and the produced kermesic acid glucoside is carminic acid.

12. The method of any of claims 10 to 11, wherein the recombinant host cell in step (A) is a recombinant host cell comprising a recombinant glycosyltransferase gene encoding a glycosyltransferase and wherein the method comprises a further step (B) with following steps:

(B): purifying the produced glycoside in step (a1) and/or in step (a2) whereby one gets a composition, wherein at least 50% w/w of the compounds in the composition is the produced flavokermesic acid glycoside and/or kermesic acid glycoside.

13. The method of any of claims 10 to 12, wherein the contacting in step (A) of claim 1 is *in vivo* and the recombinant host cell is a yeast cell, preferably wherein the yeast cell is selected from the group consisting of *Saccharomyces spp* (e.g. *Saccharomyces cerevisiae),* and *Pichia spp.*

14. The method of any of claims 10 to 13, wherein the glycosyltransferase is membrane bound.

Patentansprüche

1. Isoliertes Glycosyltransferasepolypeptid, das zu Folgendem fähig ist:

(I): Konjugieren von nukleotidaktivierter Glucose zu Flavokermessäure (FK); und/oder
(II): Konjugieren von nukleotidaktivierter Glucose zu Kermessäure (KS);

und wobei das Glycosyltransferasepolypeptid zumindest ein Polypeptid ist, ausgewählt aus der Gruppe bestehend aus:

(a) einem Polypeptid, umfassend eine Aminosäuresequenz, die zumindest 70 % Übereinstimmung mit Aminosäure 1 bis 515 von SEQ ID NO:2 hat;
(b) einem Polypeptid, umfassend eine Aminosäuresequenz, die zumindest 70 % Übereinstimmung mit Aminosäure 20 bis 468 von SEQ ID NO:2 hat;
(c) einem Polypeptid, das durch ein Polynukleotid kodiert ist, das unter zumindest mittleren Stringenzbedingungen mit (i) Nukleotid 1 bis 1548 von SEQ ID NO:1 oder (ii) einem komplementären Strang von (i) hybridisiert; und
(d) einem Fragment von Aminosäure 1 bis 515 von SEQ ID NO:2, das die Glycosyltransferaseaktivität hat, wie in (I) und/oder (II) spezifiziert.

2. Isoliertes Glycosyltransferasepolypeptid nach Anspruch 1, wobei das isolierte Glycosyltransferasepolypeptid nach Anspruch 1 zu Folgendem fähig ist:

(I): Konjugieren von Glucose zu Flavokermessäure (FK); und
(II): Konjugieren von Glucose zu Kermessäure (KS).

3. Isoliertes Glycosyltransferasepolypeptid nach einem der vorstehenden Ansprüche, wobei das Glycosyltransferasepolypeptid nach Anspruch 1 Folgendes ist:

(a) ein Polypeptid, umfassend eine Aminosäuresequenz, die zumindest 80% Übereinstimmung mit Aminosäure 1 bis 515 von SEQ ID NO:2 hat.

4. Isoliertes Glycosyltransferasepolypeptid nach einem der vorstehenden Ansprüche, wobei das Glycosyltransferasepolypeptid nach Anspruch 1 Folgendes ist:

(a) ein Polypeptid, umfassend eine Aminosäuresequenz, die zumindest 95% Übereinstimmung mit Aminosäure 1 bis 515 von SEQ ID NO:2 hat.

5. Isoliertes Polynukleotid, umfassend eine Nukleotidsequenz, die das Polypeptid nach einem der vorstehenden Ansprüche kodiert.

6. Nukleinsäurekonstrukt, umfassend das isolierte Polynukleotid nach Anspruch 5, wirkverknüpft mit einer oder mehreren Steuersequenzen, die die Produktion des Polypeptids in einem Expressionswirt leiten.

7. Rekombinanter Expressionsvektor, umfassend das Nukleinsäurekonstrukt nach Anspruch 6.

8. Rekombinante Wirtszelle, umfassend das Nukleinsäurekonstrukt nach Anspruch 7.

9. Rekombinante Wirtszelle nach Anspruch 8, wobei die rekombinante Wirtszelle eine Zelle ist, ausgewählt aus der Gruppe bestehend aus einer filamentösen Pilzzelle und einer Mikroorganismuszelle, wobei bevorzugt die Mikroorganismuszelle eine Hefezelle ist, wobei bevorzugt die Hefezelle ausgewählt ist aus der Gruppe bestehend aus *Saccharomyces spp.* (z. B. *Saccharomyces cerevisiae)* und *Pichia spp.*

10. Verfahren zum Produzieren von Flavokermessäure- (FK-)Glykosid und/oder Kermessäure-(KS-)Glykosid, wobei das Verfahren die folgenden Schritte umfasst:

(A): Inkontaktbringen, *in vitro* oder *in vivo,* in einer rekombinanten Wirtszelle, umfassend ein Glycosyltransferasegen, das eine Glycosyltransferase kodiert, von:

(a1): Flavokermessäure (FK) mit einer Glycosyltransferase, die fähig ist, die Flavokermessäure unter geeigneten Bedingungen zu glycosylieren, wobei dabei das Flavokermessäureglycosid produziert wird; und/oder
(a2): Kermessäure (KS) mit einer Glycosyltransferase, die fähig ist, die Kermessäure unter geeigneten Bedingungen zu glycosylieren, wobei dabei das Kermessäureglycosid produziert wird, und

wobei die Glycosyltransferase in Schritt (A) eine Glycosyltransferase nach einem der Ansprüche 1 bis 4 ist.

11. Verfahren nach Anspruch 10, wobei

- die Glycosyltransferase in Schritt (a1) eine Glucosyltransferase ist und dadurch in Schritt (a1) Flavokermessäureglucosid produziert wird, und wobei das produzierte Flavokermessäureglucosid die Verbindung DcII ist; und/oder
- die Glycosyltransferase in Schritt (a2) eine Glucosyltransferase ist und dadurch in Schritt (a2) Kermessäureglucosid produziert wird und das produzierte Kermessäureglucosid Karminsäure ist.

12. Verfahren nach einem der Ansprüche 10 bis 11, wobei die rekombinante Wirtszelle in Schritt (A) eine rekombinante Wirtszelle ist, die ein rekombinantes Glycosyltransferasegen umfasst, das eine Glycosyltransferase kodiert, und wobei das Verfahren einen weiteren Schritt (B) mit den folgenden Schritten umfasst:

(B): Reinigen des in Schritt (a1) und/oder in Schritt (a2) produzierten Glycosids, wobei zumindest 50 % (w/w) der Verbindungen in der Zusammensetzung das produzierte Flavokermessäureglycosid und/oder Kermessäureglycosid sind.

13. Verfahren nach einem der Ansprüche 10 bis 12, wobei das Inkontaktbringen in Schritt (A) nach Anspruch 1 *in vivo* erfolgt und die rekombinante Wirtszelle eine Hefezelle ist, wobei bevorzugt die Hefezelle ausgewählt ist aus der Gruppe bestehend aus *Saccharomyces spp.* (z. B. *Saccharomyces cerevisiae)* und *Pichia spp.*

14. Verfahren nach einem der Ansprüche 10 bis 13, wobei die Glycosyltransferase membrangebunden ist.

Revendications

1. Polypeptide de glycosyltransférase isolé pouvant :

(I) : conjuguer un glucose à activation nucléotidique à de l'acide flavokermésique (FK) ; et/ou
(II) : conjuguer un glucose à activation nucléotidique à de l'acide kermésique (KA) ;

et ledit polypeptide de glycosyltransférase étant au moins un polypeptide choisi dans le groupe constitué par :

(a) un polypeptide comprenant une séquence d'acides aminés qui présente une identité d'au moins 70 % avec les acides aminés 1 à 515 de SEQ ID n° : 2 ;
(b) un polypeptide comprenant une séquence d'acides aminés qui présente une identité d'au moins 70 % avec les acides aminés 20 à 468 de SEQ ID n° : 2 ;
(c) un polypeptide qui est codé par un polynucléotide qui s'hybride dans des conditions de stringence au moins moyenne (i) aux nucléotides 1 à 1548 de SEQ ID n° : 1 ou (ii) à un brin complémentaire de (i) ; et
(d) un fragment des acides aminés 1 à 515 de SEQ ID n° : 2, qui présente l'activité glycosyltransférase telle qu'indiquée en (I) et/ou (II).

2. Polypeptide de glycosyltransférase isolé selon la revendication 1, ledit polypeptide de glycosyltransférase isolé selon la revendication 1 pouvant :

(I) : conjuguer un glucose à de l'acide flavokermésique (FK) ; et
(II) : conjuguer un glucose à de l'acide kermésique (KA).

3. Polypeptide de glycosyltransférase isolé selon l'une quelconque des revendications précédentes, ledit polypeptide de glycosyltransférase selon la revendication 1 étant :

(a) un polypeptide comprenant une séquence d'acides aminés qui présente une identité d'au moins 80 % avec les acides aminés 1 à 515 de SEQ ID n° : 2.

4. Polypeptide de glycosyltransférase isolé selon l'une quelconque des revendications précédentes, ledit polypeptide de glycosyltransférase selon la revendication 1 étant :

(a) un polypeptide comprenant une séquence d'acides aminés qui présente une identité d'au moins 95 % avec les acides aminés 1 à 515 de SEQ ID n° : 2.

5. Polynucléotide isolé comprenant une séquence nucléotidique qui code le polypeptide selon l'une quelconque des revendications précédentes.

6. Construction d'acide nucléique comprenant le polynucléotide isolé selon la revendication 5 fonctionnellement lié à une ou plusieurs séquences de contrôle qui dirige la production du polypeptide dans un hôte d'expression.

7. Vecteur d'expression recombinant comprenant la construction d'acide nucléique selon la revendication 6.

8. Cellule hôte recombinante comprenant la construction d'acide nucléique selon la revendication 7.

9. Cellule hôte recombinante selon la revendication 8, la cellule hôte recombinante étant une cellule choisie dans le groupe constitué par une cellule fongique filamenteuse et une cellule de microorganisme, de préférence ladite cellule de microorganisme étant une cellule de levure, de préférence ladite cellule de levure étant choisie dans le groupe constitué par *Saccharomyces spp* (par ex. *Saccharomyces cerevisiae)* et *Pichia spp.*

10. Procédé de production d'un glycoside d'acide flavokermésique (FK) et/ou d'un glycoside d'acide kermésique (KA), ledit procédé comprenant les étapes suivantes :

(A) : la mise en contact *in vitro* ou *in vivo* dans une cellule hôte recombinante comprenant un gène de glycosyltransférase codant une glycosyltransférase :

(a1) : de l'acide flavokermésique (FK) avec une glycosyltransférase pouvant glycosyler l'acide flavokermésique dans des conditions appropriées dans lesquelles est produit le glycoside d'acide flavokermésique ; et/ou
(a2) : de l'acide kermésique (KA) avec une glycosyltransférase pouvant glycosyler l'acide kermésique dans des conditions appropriées dans lesquelles est produit le glycoside d'acide kermésique, et

ladite glycosyltransférase dans l'étape (A) étant une glycosyltransférase selon l'une quelconque des revendications 1 à 4.

11. Procédé selon la revendication 10,

- ladite glycosyltransférase dans l'étape (a1) étant une glycosyltransférase et ainsi dans l'étape (a1) étant produit un glucoside d'acide flavokermésique et ledit glucoside d'acide flavokermésique produit étant le composé DcII ; et/ou
- ladite glycosyltransférase dans l'étape (a2) étant une glycosyltransférase et ainsi dans l'étape (a2) étant produit un glucoside d'acide kermésique et ledit glucoside d'acide kermésique produit étant l'acide carminique.

12. Procédé selon l'une quelconque des revendications 10 à 11, ladite cellule hôte recombinante dans l'étape (A) étant une cellule hôte recombinante comprenant un gène de glycosyltransférase recombinant codant une glycosyltransférase et ledit procédé comprenant une étape (B) supplémentaire avec les étapes suivantes :

(B) : la purification du glycoside produit dans l'étape (a1) et/ou dans l'étape (a2) grâce à laquelle on obtient une composition, dans laquelle au moins 50 % en p/p des composés présents dans la composition représentent le glycoside d'acide flavokermésique et/ou le glycoside d'acide kermésique produits.

13. Procédé selon l'une quelconque des revendications 10 à 12, ladite mise en contact dans l'étape (A) de la revendication 1 se faisant *in vivo* et ladite cellule hôte recombinante étant une cellule de levure, de préférence ladite cellule de levure étant choisie dans le groupe constitué par *Saccharomyces spp* (par ex. *Saccharomyces cerevisiae)* et *Pichia spp.*

14. Procédé selon l'une quelconque des revendications 10 à 13, ladite glycosyltransférase étant liée à une membrane.

Figure 1

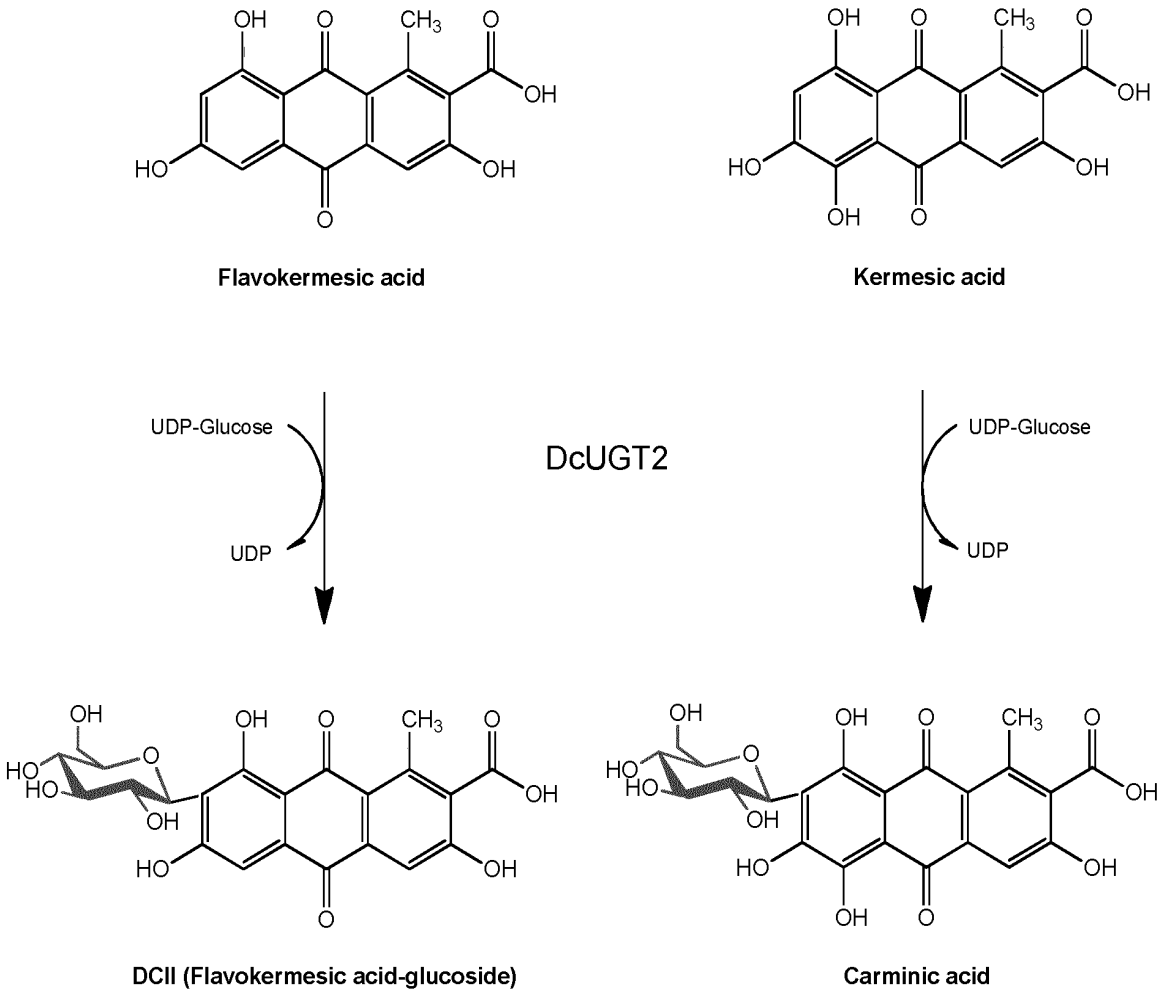

Flavokermesic acid
Kermesic acid
UDP-Glucose
UDP
DcUGT2
UDP-Glucose
UDP
DCII (Flavokermesic acid-glucoside)
Carminic acid

Figure 2 (A)

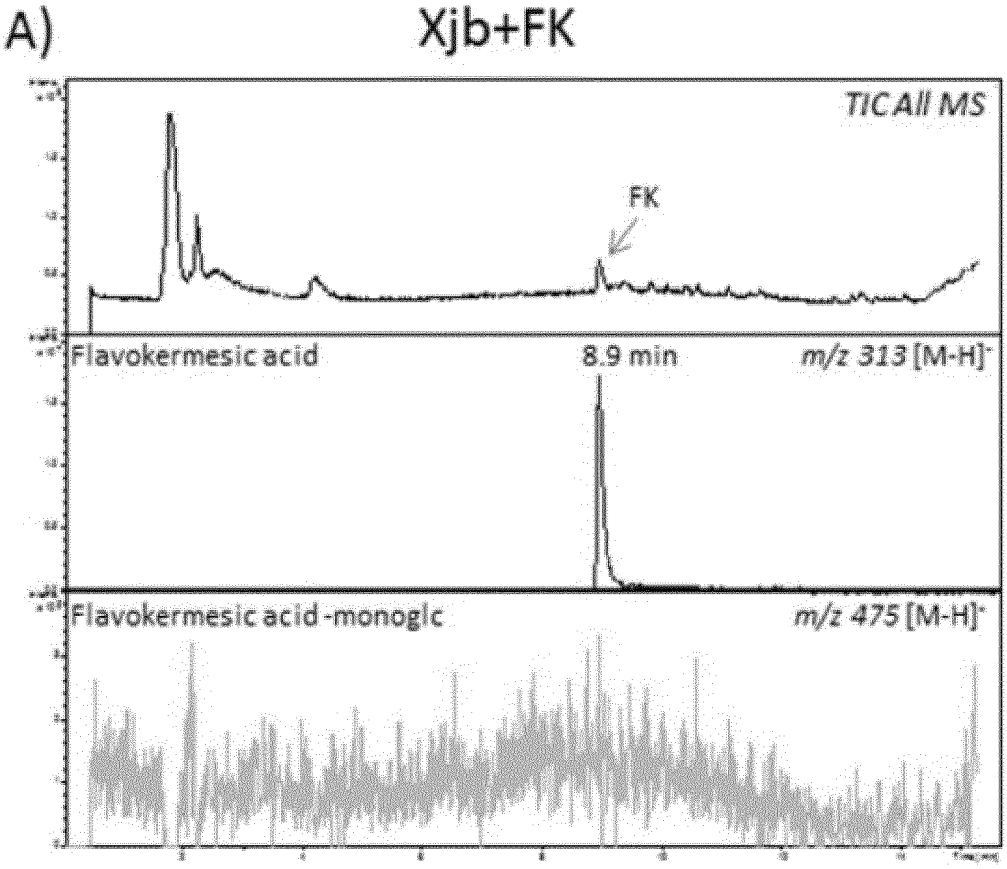

A)
Xjb+FK
TIC All MS
FK
Flavokermesic acid
8.9 min
m/z 313 [M-H]-
Flavokermesic acid -monoglc
m/z 475 [M-H]-

Figure 2 (B)

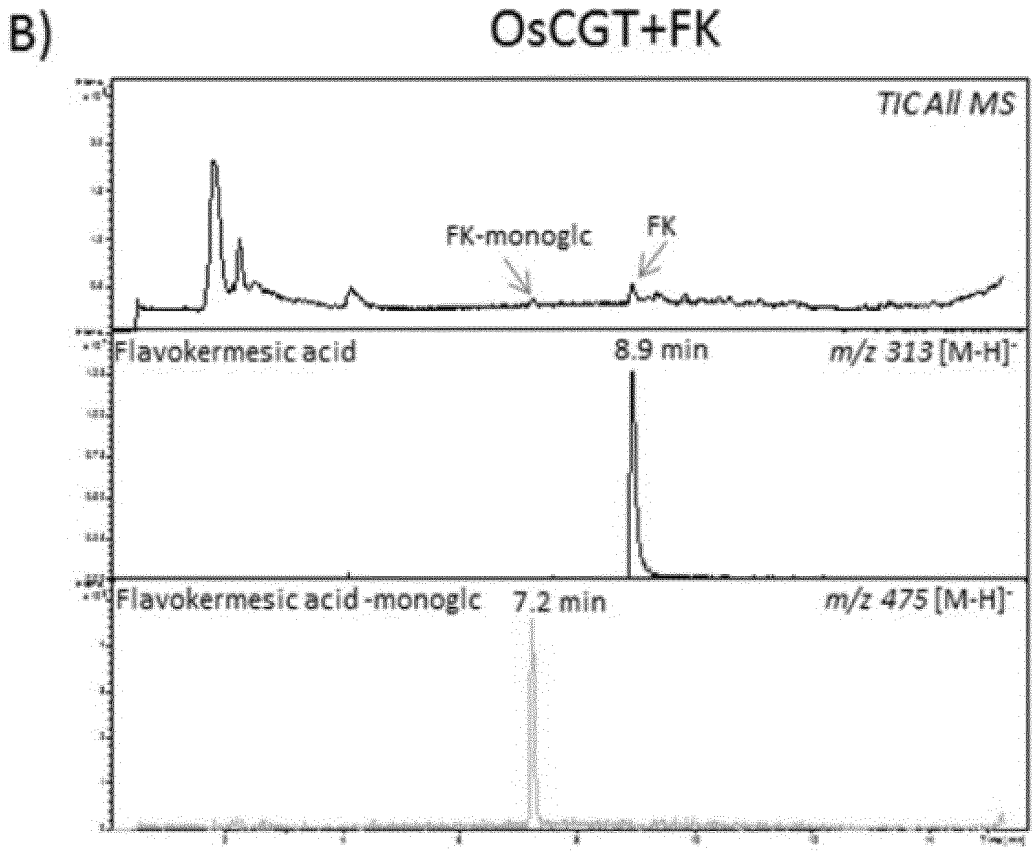

B)
OsCGT+FK
TIC All MS
FK-monoglc
FK
Flavokermesic acid
8.9 min
m/z 313 [M-H]-
Flavokermesic acid -monoglc
7.2 min
m/z 475 [M-H]-

Figure 2 (C)

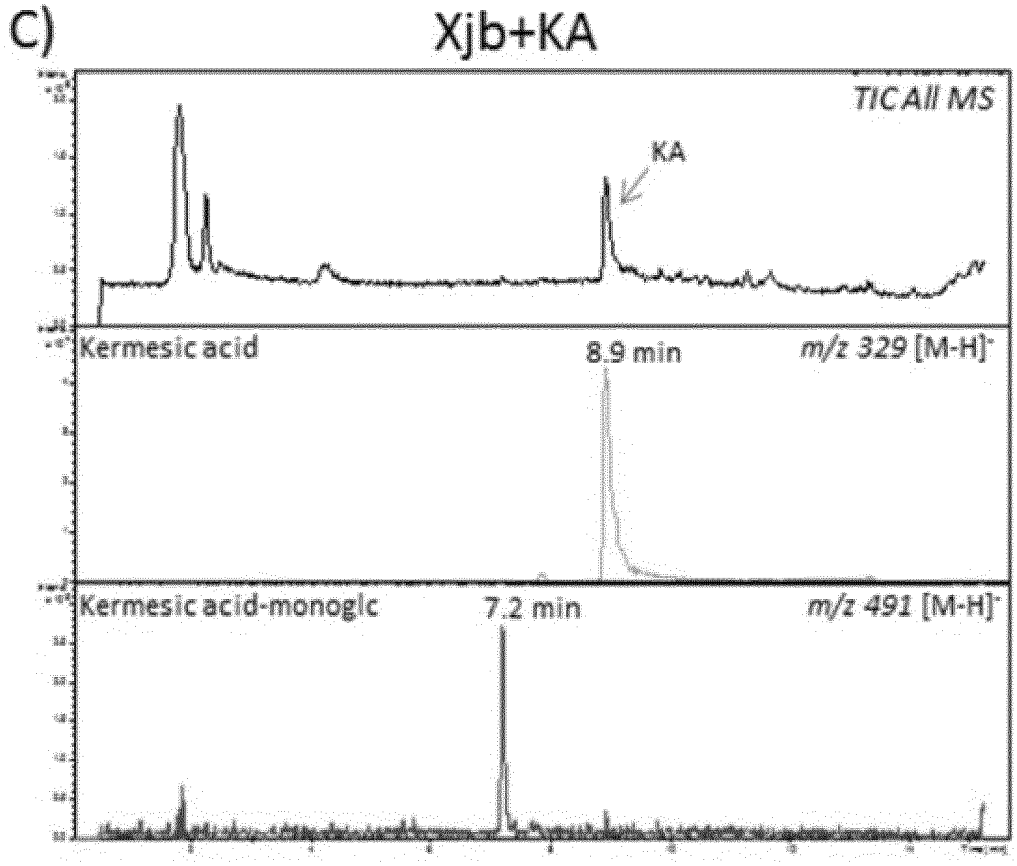

C)
Xjb+KA
TIC All MS
KA
Kermesic acid
8.9 min
m/z 329 [M-H]-
Kermesic acid-monoglc
7.2 min
m/z 491 [M-H]-

Figure 2 (D)

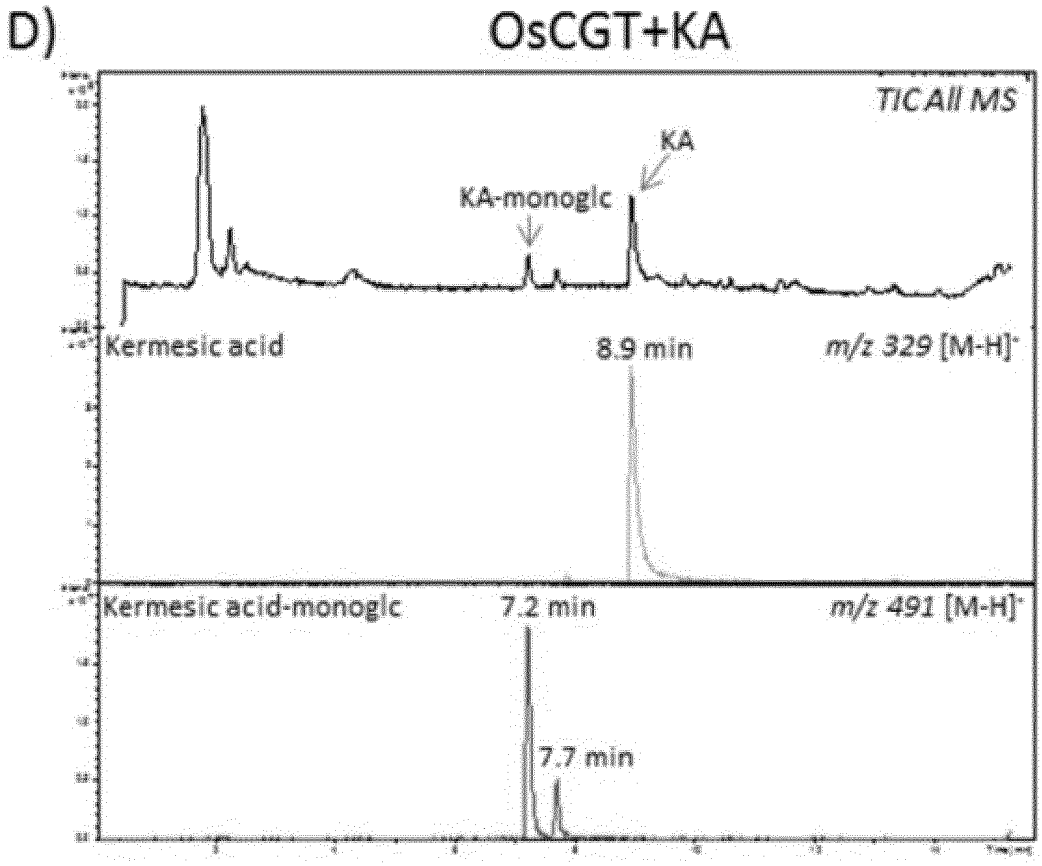

D)
OsCGT+KA
TIC All MS
KA
KA-monoglc
Kermesic acid
8.9 min
m/z 329 [M-H]-
Kermesic acid-monoglc
7.2 min
m/z 491 [M-H]-
7.7 min

Figure 2 (E)

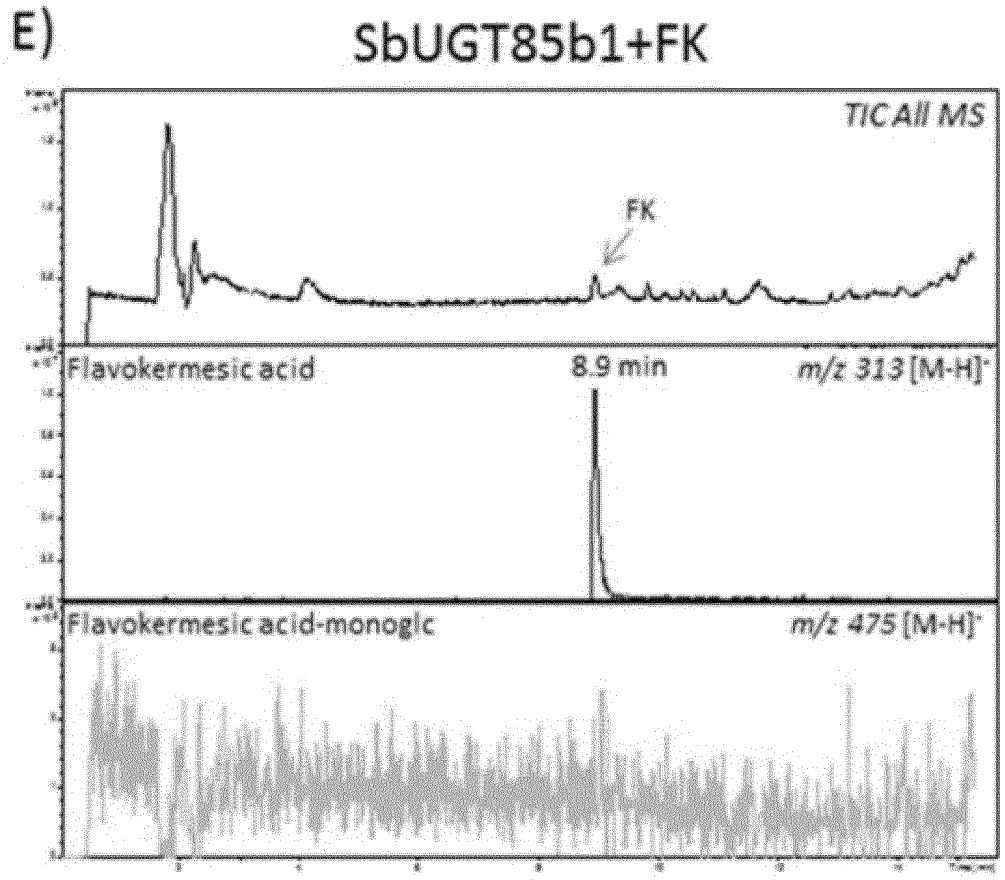

E)
SbUGT85b1+FK
TIC All MS
FK
Flavokermesic acid
8.9 min
m/z 313 [M-H]⁻
Flavokermesic acid-monoglc
m/z 475 [M-H]⁻

Figure 2 (F)

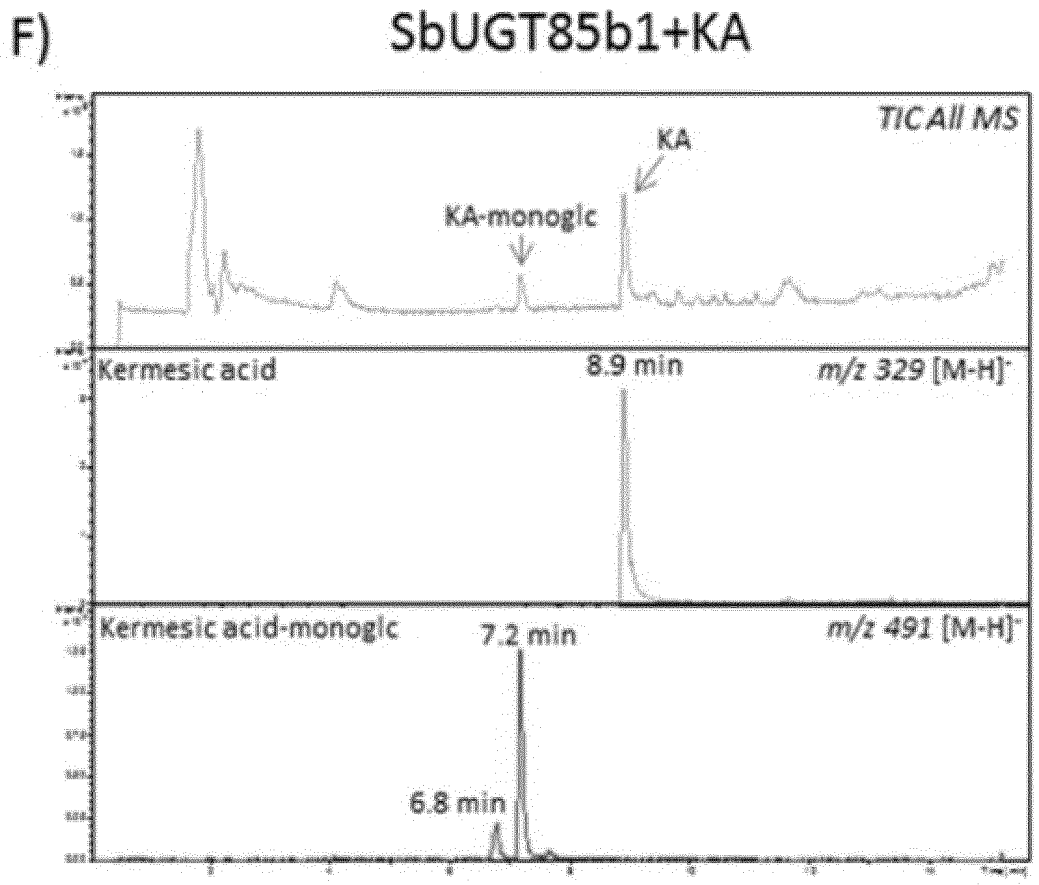

F)
SbUGT85b1+KA
TIC All MS
KA
KA-monoglc
Kermesic acid
8.9 min
m/z 329 [M-H]⁻
Kermesic acid-monoglc
7.2 min
m/z 491 [M-H]⁻
6.8 min

Figure 2 (G)

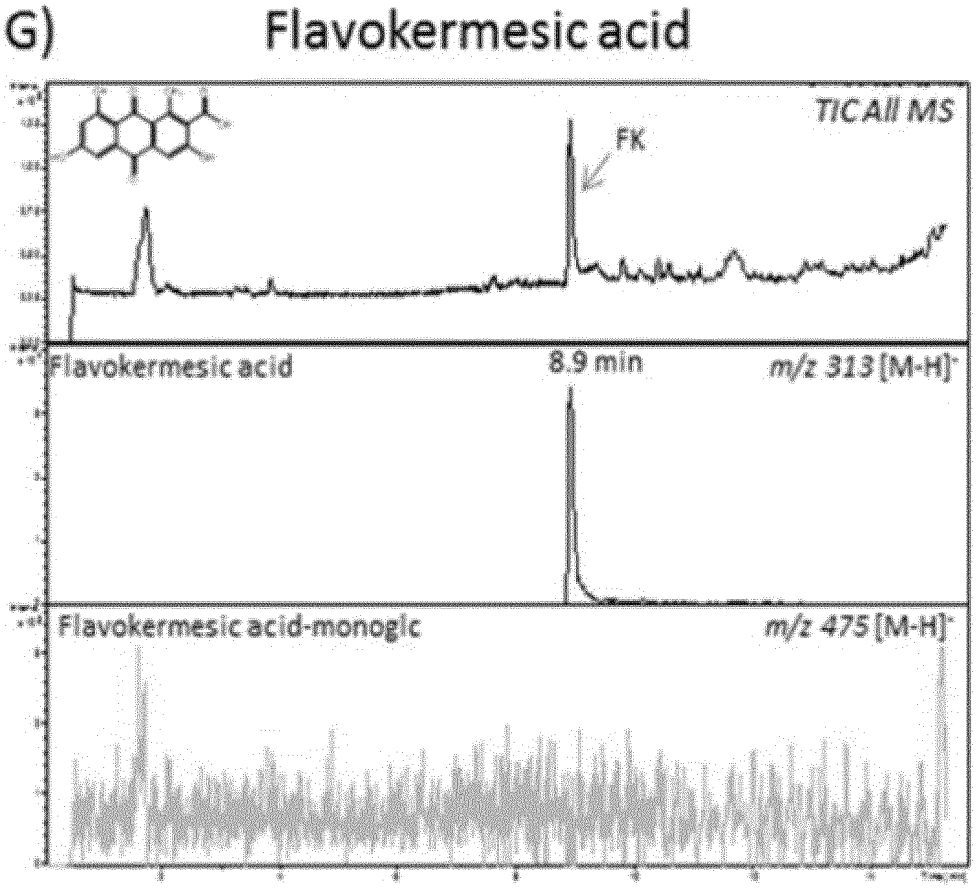
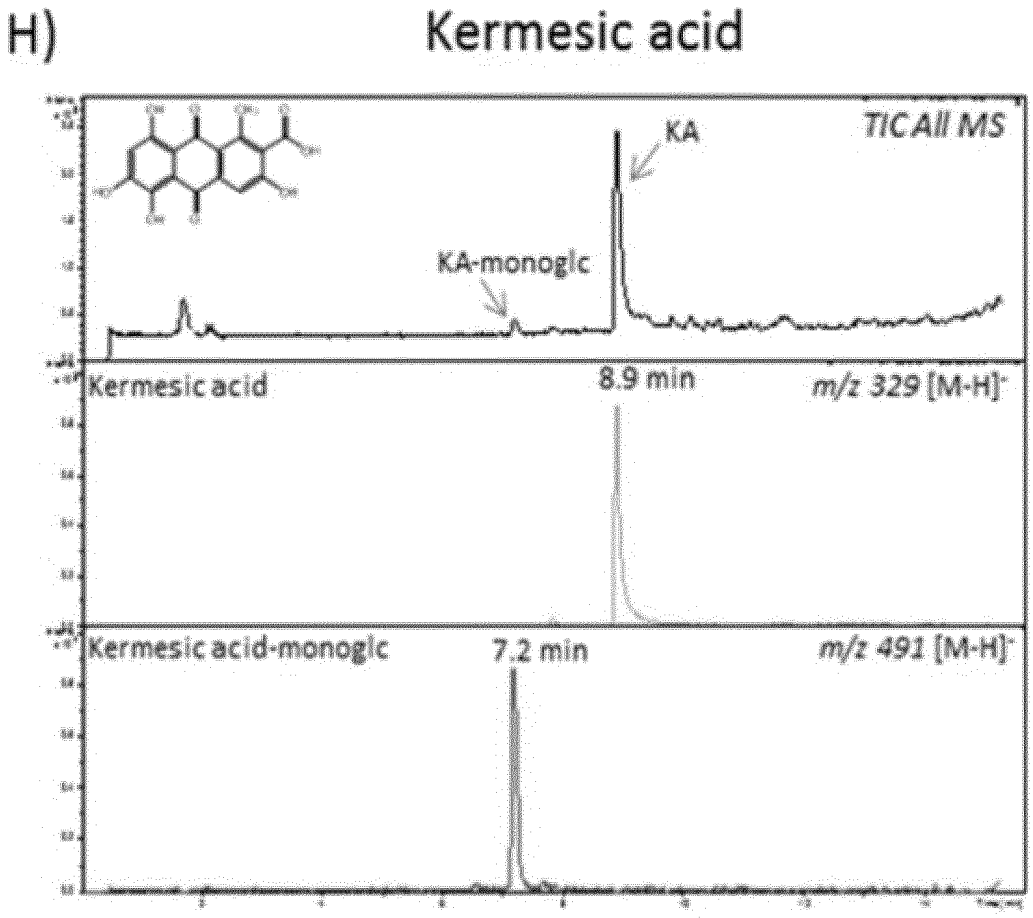

G)
Flavokermesic acid
TIC All MS
FK
Flavokermesic acid
8.9 min
m/z 313 [M-H]-
Flavokermesic acid-monoglc
m/z 475 [M-H]-

Figure 2 (H)


H)
Kermesic acid
KA
TIC All MS
KA-monoglc
Kermesic acid
8.9 min
m/z 329 [M-H]-
Kermesic acid-monoglc
7.2 min
m/z 491 [M-H]-

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- US 5424421 A **[0007] [0183]**
- WO 2006056585 A1 **[0008] [0183]**
- WO 2004111254 A1 **[0018] [0046] [0183]**
- EP 13198110 A **[0184]**

Non-patent literature cited in the description

- **STATHOPOULOU et al.** *Analytica Chimica Acta,* 2013, vol. 804, 264-272 **[0013]**
- **ZAGROBELNY et al.** Cyanogenic glucosides and plant-insect interactions. *Phytochemistry,* February 2004, vol. 65 (3), 293-306 **[0016] [0183]**
- **GEUDER et al.** *Journal of Chemical Ecology,* 1997, vol. 23 (5 **[0016] [0183]**
- *Journal of Insect Physiology,* 2006, vol. 52, 593-601 **[0016]**
- **GUTMANN et al.** *Pure Appl. Chem,* 09 July 2013 **[0025] [0183]**
- **BAIG et al.** *Angew Chem Int Ed Engl.,* 27 November 2006, vol. 45 (46), 7842-6 **[0026] [0157] [0183]**
- **RADOMINSKA-PANDYA A ; BRATTON SM ; REDINBO MR ; MILEY MJ.** *Drug Metab Rev.,* February 2010, vol. 42 (1), 133-44 **[0027] [0183]**
- *Plant Physiology,* November 2008, vol. 148, 1295-1308 **[0027] [0183]**
- **ESBEN HALKJAER HANSEN et al.** *Phytochemistry,* vol. 70 (4), 473-482 **[0047] [0183]**
- **J. SAMBROOK ; E.F. FRITSCH ; T. MANIATUS.** Molecular Cloning, A Laboratory Manual. Cold Spring Harbor, 1989 **[0053]**
- **NEEDLEMAN ; WUNSCH.** *J. Mol. Biol.,* 1970, vol. 48, 443-453 **[0061]**
- **RICE et al.** *Trends Genet.,* 2000, vol. 16, 276-277 **[0061]**
- **D. POMPON ; B. LOUERAT ; A. BRONINE ; P. URBAN.** Yeast expression of animal and plant P450s in optimized redox environments. *Methods Enzymol.,* 1996, vol. 272, 51-64 **[0142]**
- **STATHOPOULOU et al.** *Analytica Chimica Acta,* 2013, vol. 804, 264-272 **[0183]**
- **GENTA et al.** Potential role for gut microbiota in cell wall digestion and glucoside detoxification in Tenebrio molitor larvae. *Journal of Insect Physiology,* 2006, vol. 52, 593-601 **[0183]**
- **POMPON et al.** *Methods Enzymol.,* 1996, vol. 272, 51-64 **[0183]**
- **KANNANGARA et al.** *Plant Journal,* 2011, vol. 68, 287-301 **[0183]**
- **AUGUSTIN et al.** *Plant Physiology,* 2012, vol. 160, 1881-1895 **[0183]**
- **D'AOUST et al.** *Methods Mol Biol,* 2009, vol. 483, 41-50 **[0183]**
- **GELVIN.** *Microbiol Mol Biol Rev,* 2003, vol. 67 (1), 16-37 **[0183]**
- **RAUWALD ; SIGLER.** *Phytochemical Analysis,* 1994, vol. 5, 266-270 **[0183]**