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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
19 December 2024 (19.12.2024)

WIPO | PCT

(10) International Publication Number
WO 2024/259172 A1

(51) International Patent Classification:
C12N 9/10 (2006.01) *A01H 6/28* (2018.01)
C12N 5/04 (2006.01) *C12N 15/52* (2006.01)

(21) International Application Number:
PCT/US2024/033912

(22) International Filing Date:
13 June 2024 (13.06.2024)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
63/472,961 14 June 2023 (14.06.2023) US

(71) Applicant: **TRAIT BIOSCIENCES, INC.** [US/US]; 100 Entrada Dr., Los Alamos, NM 87544 (US).

(72) Inventor: **HUSSAIN, Syed, Ammar**; 120 Shadow Mountain LN, Morrisville, NC 27560 (US).

(74) Agent: **KERR, David, S.**; Berg Hill Greenleaf Ruscitti LLP, 1712 Pearl Street, Boulder, CO 80302 (US).

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

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

Published:
— *with international search report (Art. 21(3))*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*
— *with sequence listing part of description (Rule 5.2(a))*

(54) Title: IMPROVED SYNTHESIS OF CANNABINOID GLYCOSIDE COMPOUNDS

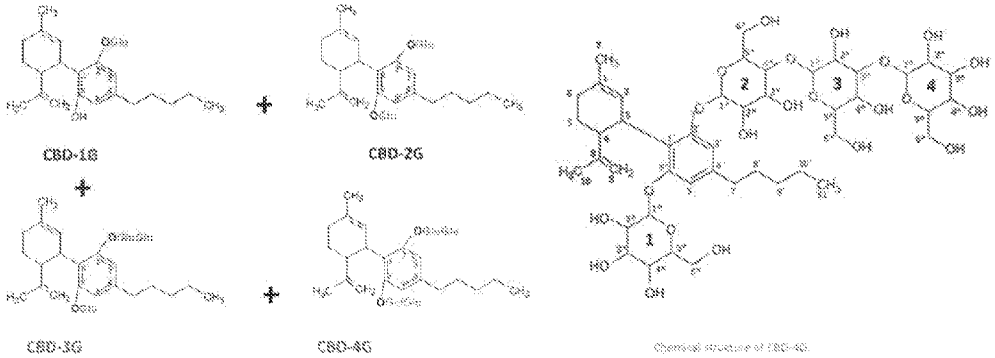


FIG. 1

(57) Abstract: The present invention relates generally to the identification novel UDP -glucosyltransferases enzymes having specific activity towards cannabinoid compounds. The present invention further relates generally to the use of novel UGT enzymes having specific activity towards cannabinoid compounds to generate water-soluble cannabinoid glycoside compounds.

WO 2024/259172 A1

IMPROVED SYNTHESIS OF CANNABINOID GLYCOSIDE COMPOUNDS

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims the benefit of and priority to U.S. Provisional Application No. 63/472,961, filed June 14, 2023. The entire specification and figures of the above-referenced application are hereby incorporated, in their entirety by reference.

SEQUENCE LISTING

The instant application contains contents of the electronic sequence listing (90425.00401-Sequence-Listing.xml; Size: 20,155 bytes; and Date of Creation: June 13, 2024) is herein incorporated by reference in its entirety.

TECHNICAL FIELD

The present invention relates generally to the identification novel UDP-glucosyltransferase enzymes having specific activity towards cannabinoid compounds. The present invention further relates generally to the use of novel UGT enzymes having specific activity towards cannabinoid compounds to generate water-soluble cannabinoid glycoside compounds.

BACKGROUND

Cannabinoids are a class of specialized compounds synthesized by *Cannabis*. They are formed by condensation of terpene and phenol precursors. They include these more abundant forms: Δ^9-tetrahydrocannabinol (THC), cannabidiol (CBD), cannabichromene (CBC), and cannabigerol (CBG). Another cannabinoid, cannabinol (CBN), is formed from THC as a degradation product and can be detected in some plant strains. Typically, THC, CBD, CBC, and CBG occur together in different ratios in the various plant strains. These cannabinoids are generally lipophilic, nitrogen-free, mostly phenolic compounds and are derived biogenetically from a monoterpene and phenol, the acid cannabinoids from a monoterpene and phenol carboxylic acid and have a C21 base. Cannabinoids also find their corresponding carboxylic acids in plant products. In general, the carboxylic acids have the function of a biosynthetic precursor. For example, the tetrahydrocannabinols Δ^9 – and Δ^8 -THC arise *in vivo* from the THC carboxylic acids by decarboxylation and likewise, CBD from the associated cannabidiolic acid.

Importantly, cannabinoids are hydrophobic small molecules and, as a result, are highly insoluble. Due to this insolubility, cannabinoids such as THC and CBD may need to be efficiently solubilized to facilitate transport, storage, and adsorption through certain tissues and organs. For

example, the metabolism of cannabinoids in the human body goes through the classic two-phases detoxification process of oxidation followed by glucuronidation – which is a form of glycosylation involving the addition of a sugar from UDP-Glucuronic Acid to a cannabinoid. As shown below, the chemical structures of UDP-glucuronic acid and UDP-glucose are similar. As described in, US8410064 by Pandya *et al.*, cannabinoids may be subject to cytochrome P450 oxidation and subsequent UDP-glucuronosyltransferase dependent glucuronidation in the body after consumption. (see figure 12) The resulting glucuronide of the oxidized cannabinoids is the main metabolite found in urine, and thus, this solubilization process plays a critical role in the metabolic clearance of cannabinoids. In another embodiment outlined in PCT/US18/24409 and PCT/US18/41710 (both of which are incorporated herein in their entirety by reference, including examples 1-19, and all specific materials and method), by Sayre *et al.*, cannabinoids may be glycosylated to form water-soluble glycoside compounds. In preferred embodiment, such water-soluble cannabinoid glycoside may include one or more sugar moieties, and preferably 1-3 sugar moieties, also referred to a glycosylation sites.

One area where water-soluble cannabinoids has seen renewed interest is in the fields of cannabinoid-infused consumer products. However, the ability to effectively solubilize cannabinoids has limited their applicability. To overcome these limitations, many manufacturers of cannabinoid-infused products have adopted the use of traditional pharmaceutical delivery methods of using nanoemulsions of cannabinoids. This nanoemulsion process essentially coats the cannabinoid in a hydrophilic compound, such as oil or other similar compositions. However, the use of nanoemulsions is limited both technically, and from a safety perspective:

First, a large number of surfactants and cosurfactants are required for nanoemulsion stabilization. Moreover, the stability of nanoemulsions is inherently unstable, and may be disturbed by slight fluctuations in temperature and pH and is further subject to the "oswald ripening effect" or ORE. ORE describes the process whereby molecules on the surface of particles are more energetically unstable than those within. Therefore, the unstable surface molecules often go into solution shrinking the particle over time and increasing the number of free molecules in solution. When the solution is supersaturated with the molecules of the shrinking particles, those free molecules will redeposit on the larger particles. Thus, small particles decrease in size until they disappear, and large particles grow even larger. This shrinking and growing of particles will result

in a larger mean diameter of a particle size distribution (PSD). Over time, this causes emulsion instability and eventually phase separation.

Second, nanoemulsions may not be safe for human consumption. For example, nanoemulsions were first developed as a method to deliver small quantities of pharmaceutical compounds having poor solubility. However, the ability to "hide" a compound, such as a cannabinoid, in a nanoemulsion may allow the cannabinoid to be delivered to parts of the body where it was previously prevented from entering, as well as accumulating in tissues and organs where cannabinoids and nanoparticles would not typically be found. Additionally, such nanoemulsions, as well as other water-compatible strategies, do not address one of the major-shortcomings of cannabinoid-infused commercial consumables, namely the strong unpleasant smell and taste. Moreover, such water-compatible strategies deliver inconsistent and delayed cannabinoid uptake in the body which may result in consumers ingesting a higher dose of cannabinoid-infused product than is recommended, as well as delayed, inconsistent, and unpredictable medical and/or psychotropic experiences. As will be discussed in more detail below, the current inventive technology overcomes the limitations of traditional cannabinoid emulsion systems while meeting the objectives of a truly effective and scalable cannabinoid production, solubilization, and isolation system.

SUMMARY OF THE INVENTION

One aspect of the present invention relates generally to the identification novel UDP-glucosyltransferase (UDP-UGTs or UGTs) enzymes having glycosylation activity towards one or more cannabinoid compounds. In one preferred aspect, the present invention includes the identification of novel UGTs according to the nucleotide sequences identified as SEQ ID NOs. 1, and UGTs having at least 90% sequence identity with SEQ ID NOs. 1, that have glycosylation activity towards one or more cannabinoid compounds, and preferably CBD. In this preferred embodiment, the UGT of the invention cany generate a cannabinoid, and preferably CBD having between 1-4 UDP sugar moieties, referred to generally as CBD-1G, CBD-2G, CBD-3G, and CBD-4G.

In another preferred aspect, the present invention includes the identification of novel UGTs according to the amino acid sequences identified as SEQ ID NOs. 2, and UGTs having at least 90% sequence identity with SEQ ID NOs. 2, that have glycosylation activity towards one or more cannabinoid compounds, and preferably CBD. In this preferred embodiment, the UGT of the

invention cany generate a cannabinoid, and preferably CBD having between 1-4 UDP sugar moieties, referred to generally as CBD-1G, CBD-2G, CBD-3G, and CBD-4G.

In another preferred aspect, the present invention includes the identification of novel UGTs according to the amino acid sequences identified as SEQ ID NOs. 4, 6, 8 and 10, and UGTs having at least 90% sequence identity with SEQ ID NOs. 4, 6, 8 and 10, that have glycosylation activity towards one or more cannabinoid compounds, and preferably CBN. In this preferred embodiment, the UGT of the invention cany generate a cannabinoid, and preferably CBD having at least one UDP sugar moiety, referred to generally as CBN-1G.

In another preferred aspect, the present invention includes the identification of novel UGTs according to the nucleotide sequences identified as SEQ ID NOs. 3, 5, 7, and 9, and UGTs having at least 90% sequence identity with SEQ ID NOs. 3, 5, 7, and 9, that have glycosylation activity towards one or more cannabinoid compounds, and preferably CBN. In this preferred embodiment, the UGT of the invention cany generate a cannabinoid, and preferably CBD having at least one UDP sugar moieties, referred to generally as CBN-1G.

One aspect of the present invention further relates generally to the use of novel UGT enzymes having specific activity towards one or more cannabinoid compounds to generate water-soluble cannabinoid glycoside compounds in *in vitro*, *ex vivo*, and *in vivo* systems. In one preferred aspect, the present invention use of novel UGT enzymes according to the amino acid sequences identified as SEQ ID NO. 1, and UGTs having at least 90% sequence identity with SEQ ID NO. 1, that have glycosylation activity towards one or more cannabinoid compounds, and preferably CBD, in *in vitro*, *ex vivo*, and *in vivo* systems. In one preferred aspect of the invention, an *in vivo* system may include a whole organism system, such as a plant, or cell culture, such as a plant cell culture, an algal cell culture, a fungi cell culture, or a microorganism cell culture, such as a bacterial or yeast cell culture.

Another aspect of the current inventive technology includes the isolated amino acid sequences encoding one or more UGTs that have glycosylation activity towards one or more cannabinoid compounds according to SEQ ID NOs. 1, and amino acid sequence having at least 90% sequence identity with SEQ ID NO. 1.

Another aspect of the current inventive technology includes a nucleotide sequence encoding one or more UGTs that have glycosylation activity towards one or more cannabinoid compounds according to SEQ ID NOs. 1, and a nucleotide sequence encoding an amino acid

sequence having at least 90% sequence identity with SEQ ID NOs. 1. Another aspect of the current inventive technology includes an expression vector having a nucleotide sequence encoding one or more UGTs that have glycosylation activity towards one or more cannabinoid compounds according to SEQ ID NOs. 1, and a nucleotide sequence encoding an amino acid sequence having at least 90% sequence identity with SEQ ID NOs. 2, operably linked to a promoter.

Another aspect of the current inventive technology includes one or more organisms, such as a plant, plant cell, bacteria, algae, fungi, or yeast cell, transformed by an expression vector having a nucleotide sequence encoding one or more UGTs that have glycosylation activity towards one or more cannabinoid compounds according to SEQ ID NOs. 1, and a nucleotide sequence encoding an amino acid sequence having at least 90% sequence identity with SEQ ID NOs. 2, operably linked to a promoter. One aspect of the current inventive technology includes improved systems and methods for the bioconversion of cannabinoid compounds into water-soluble cannabinoid glycosides, or water-soluble cannabinoid glycosides in a bacterial, yeast, or plant cell culture system. In another preferred aspect, a preferred plant cell culture system may include a *Cannabis* suspension cell culture, or a tobacco plant cell culture.

Additional aspects of the invention may become evident based on the specification and figures presented below.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1: shows exemplary chemical structures for cannabidiol (CBD) glycosides having between 1-4 UDP-glucose moieties, specifically CBD-1G, CBD-2G, CBD-3G, and CBD-4G generated by GT3 (SEQ ID NO. 1, 2)

Figure 2: shows HPLC analysis of CBD glycosides having between 1-4 UDP-glucose moieties, specifically CBD-1G, CBD-2G, CBD-3G, and CBD-4G generated by GT3 (SEQ ID NO. 1, 2) from a CBD substrate.

Figure 3: shows bioconversion of CBD into a CBD glycoside having a single UDP sugar moiety (CBD-1G) by UGTs 1-12.

Figure 4: shows bioconversion of CBD into a CBD glycoside having two UDP sugar moiety (CBD-2G) by UGTs 1-12, where GT3 demonstrated highest production of CBD-2G.

Figure 5: shows bioconversion of CBD into a CBD glycoside having three UDP sugar moiety (CBD-3G) by UGTs 1-12, where GT3 demonstrated highest production of CBD-3G.

Figure 6: shows bioconversion of CBD into a CBD glycoside having four UDP sugar moiety (CBD-4G) by UGTs 1-12, where GT3 demonstrated highest production of CBD-4G.

Figure 7A-E: (A) shows HPLC analysis of CBN glycosides generated having 1 UDP-glucose moiety (CBN1) from a CBN substrate and a UDP-glucose catalyzed by GT4 (SEQ ID NO. 3, 4); (B) shows HPLC analysis of CBN glycosides generated having 1 UDP-glucose moiety (CBN1) from a CBN substrate and a UDP-glucose catalyzed by GT5 (SEQ ID NO. 5, 6); (C) shows HPLC analysis of CBN glycosides generated having 1 UDP-glucose moiety (CBN1) from a CBN substrate and a UDP-glucose catalyzed by GT6 (SEQ ID NO. 7, 8); (D) shows HPLC analysis of CBN glycosides generated having 1 UDP-glucose moiety (CBN1) from a CBN substrate and a UDP-glucose catalyzed by GT8 (SEQ ID NO. 9, 10); (E) shows bioconversion of CBN into a CBN glycoside having a single UDP sugar moiety (CBD-1G) by UGTs 1-12.

Figure 8: show phylogenetic relationship between UGTs 1-12 (designated GT1-GT12).

DETAILED DESCRIPTION OF THE INVENTION

One embodiment of the present invention relates generally to the identification novel UDP-glucosyltransferases (UDP-UGTs or UGTs) enzymes having glycosylation activity towards one or more cannabinoid compounds. In one preferred embodiment, the present invention includes the identification of novel UGTs identified as GT3, that catalyzes the bioconversion of a CBD substrate and UDP-glucose to a CBD glycoside having between 1-4 UDP sugar moieties, referred to generally as CBD-1G, CBD-2G, CBD-3G, and CBD-4G.

Another embodiment of the present invention relates generally to the identification novel UDP-glucosyltransferases (UDP-UGTs or UGTs) enzymes having glycosylation activity towards one or more cannabinoid compounds. In one preferred embodiment, the present invention includes the identification of novel UGTs according to SEQ ID NOs. 1, and UGTs having at least 90% sequence identity with SEQ ID NOs. 1, that catalyzes the bioconversion of a CBD substrate and UDP-glucose to a CBD glycoside having between 1-4 UDP sugar moieties, referred to generally as CBD-1G, CBD-2G, CBD-3G, and CBD-4G.

Another embodiment of the present invention relates generally to the identification novel UDP-glucosyltransferases (UDP-UGTs or UGTs) enzymes having glycosylation activity towards one or more cannabinoid compounds. In one preferred embodiment, the present invention includes the identification of novel UGTs according to SEQ ID NOs. 2, or a fragment thereof, and UGTs having at least 90% sequence identity with SEQ ID NOs. 2, that catalyzes the bioconversion of a

CBD substrate and UDP-glucose to a CBD glycoside having between 1-4 UDP sugar moieties, referred to generally as CBD-1G, CBD-2G, CBD-3G, and CBD-4G.

In another embodiment, the present invention includes the identification of novel UGTs identified as GT4, GT5, GT6, and GT8, that catalyzes the bioconversion of a CBN substrate and UDP-glucose to a CBD glycoside having at least one UDP sugar moiety, referred to generally as CBN-1G.

Another embodiment of the present invention relates generally to the identification novel UDP-glucosyltransferases (UDP-UGTs or UGTs) enzymes having glycosylation activity towards one or more cannabinoid compounds. In one preferred embodiment, the present invention includes the identification of novel UGTs according to SEQ ID NOs. 4, 6, 8, and 10, and UGTs having at least 90% sequence identity with SEQ ID NOs. 1, that catalyzes the bioconversion of a CBN substrate and UDP-glucose to a CBD glycoside having at least one UDP sugar moiety, referred to generally as CBN-1G.

Another embodiment of the present invention relates generally to the identification novel UDP-glucosyltransferases (UDP-UGTs or UGTs) enzymes having glycosylation activity towards one or more cannabinoid compounds. In one preferred embodiment, the present invention includes the identification of novel UGTs according to SEQ ID NOs. 4, 6, 8 and 10, or a fragment thereof, and UGTs having at least 90% sequence identity with SEQ ID NOs. 2, that catalyzes the bioconversion of a CBN substrate and UDP-glucose to a CBD glycoside having at least one UDP sugar moiety, referred to generally as CBN-1G.

Another embodiment of the present invention relates generally to the use of novel UGT enzymes having specific activity towards one or more cannabinoid compounds to generate water-soluble cannabinoid glycoside compounds in *in vitro*, *ex vivo*, and *in vivo* systems. In one preferred embodiment, the present invention includes a novel UGT enzyme according to the amino acid sequences identified as SEQ ID NOs. 2, or a fragment thereof, or UGTs having at least 90% sequence identity with SEQ ID NOs. 2, that have glycosylation activity towards one or more cannabinoid compounds, and preferably CBD, in *in vitro*, *ex vivo*, and *in vivo* systems. In one preferred embodiment of the invention, an *in vivo* system may include a whole organism system, such as a plant, or cell culture, such as a plant cell culture, an algal cell culture, a fungi cell culture, or a microorganism cell culture, such as a bacterial or yeast cell culture.

One embodiment of the present invention further relates generally to novel methods of generating water-soluble cannabinoid glycoside compounds. In one preferred embodiment, the invention novel methods of generating water-soluble CBD-glycoside comprising the step of introducing a CBD substrate to a UGT enzyme according to the amino acid sequences identified as SEQ ID NOs. 2, or a fragment thereof, or UGTs having at least 90% sequence identity with SEQ ID NOs. 2, in *in vitro*, *ex vivo*, and *in vivo* systems. In another preferred embodiment, the invention novel methods of generating water-soluble CBN-glycoside comprising the step of introducing a CBN substrate to a UGT enzyme according to the amino acid sequences identified as SEQ ID NOs. 4, 6, 8 and 10, or a fragment thereof, or UGTs having at least 90% sequence identity with SEQ ID NOs. 4, 6, 8 and 10, in *in vitro*, *ex vivo*, and *in vivo* systems.

In one preferred embodiment of the invention, an *in vivo* system may include a whole organism system, such as a plant, or cell culture, such as a plant cell culture, an algal cell culture, a fungi cell culture, or a microorganism cell culture, such as a bacterial or yeast cell culture. In other embodiments, an *ex vivo* system may include a bioreactor system. In other embodiments, an *in vitro* system may include chemical conversion of cannabinoids into water-soluble cannabinoid glycoside compounds.

Yet, another embodiment of the current inventive technology may include the generation of genetically modified organisms configured to produce water-soluble cannabinoid glycoside compounds. In one preferred embodiment, a plant, a plant cell, an algal cell, a fungi, a bacteria, or a yeast cell, may be genetically modified to express a nucleotide sequence encoding one or more UGTs that have glycosylation activity towards one or more cannabinoid compounds, and preferably a UGT selected from the group of nucleotide sequences selected from the amino acid sequence according to SEQ ID NOs. 2, 4, 6, 8 and 10, or a fragment thereof, and/or an amino acid sequence having at least 90% sequence identity with SEQ ID NOs. 2, 4, 6, 8 and 10.

Yet, another embodiment of the current inventive technology may include the generation of genetically modified organisms configured to produce water-soluble cannabinoid glycoside compounds. In one preferred embodiment, a plant, a plant cell, an algal cell, a fungi, a bacteria, or a yeast cell, may be genetically modified to express a nucleotide sequence encoding one or more UGTs that have glycosylation activity towards one or more cannabinoid compounds, and preferably a UGT selected from the group of nucleotide sequences selected from SEQ ID NOs. 1,

3, 5, 7, and 9, and/or a nucleotide sequence having at least 90% sequence identity with SEQ ID NOs. 1, 3, 5, 7, and 9.

Another embodiment of the current inventive technology may include the generation of genetically modified organisms configured to produce water-soluble CBD glycoside compounds having between 1-4 UDP-sugar moieties. In one preferred embodiment, a plant, a plant cell, an algal cell, a fungi, a bacteria, or a yeast cell, may be genetically modified to express a nucleotide sequence according to SEQ ID NO. 1, and/or a nucleotide sequence having at least 90% sequence identity with SEQ ID NOs. 1.

Another embodiment of the current inventive technology may include the generation of genetically modified organisms configured to produce water-soluble CBD glycoside compounds having between 1-4 UDP-sugar moieties. In one preferred embodiment, a plant, a plant cell, an algal cell, a fungi, a bacteria, or a yeast cell, may be genetically modified to express a nucleotide sequence encoding a peptide having an amino acid sequence according to SEQ ID NO. 2, and/or amino acid sequence having at least 90% sequence identity with SEQ ID NO. 2.

Another embodiment of the current inventive technology may include the generation of genetically modified organisms configured to produce water-soluble CBN glycoside compounds having at least one UDP-sugar moiety. In one preferred embodiment, a plant, a plant cell, an algal cell, a fungi, a bacteria, or a yeast cell, may be genetically modified to express a nucleotide sequence selected from SEQ ID NO. 3, 5, 7 and 9, and/or a nucleotide sequence having at least 90% sequence identity with SEQ ID NOs. 3, 5, 7 and 9.

Another embodiment of the current inventive technology may include the generation of genetically modified organisms configured to produce water-soluble CBD glycoside compounds having least one UDP-sugar moiety. In one preferred embodiment, a plant, a plant cell, an algal cell, a fungi, a bacteria, or a yeast cell, may be genetically modified to express a nucleotide sequence encoding a peptide having an amino acid sequence according to SEQ ID NO. 4, 6, 8 and 10, and/or amino acid sequence having at least 90% sequence identity with SEQ ID NOs. 4, 6, 8 and 10.

Another embodiment of the current inventive technology includes the isolated peptide having an amino acid sequence encoding a UGT having glycosylation activity towards one or more cannabinoid compounds, and preferably CBD according to SEQ ID NO. 2, or a fragment thereof, and/or an amino acid sequence having at least 90% sequence identity with SEQ ID NO. 1.

Another embodiment of the current inventive technology includes the isolated nucleotide sequence encoding a UGT having glycosylation activity towards one or more cannabinoid compounds, and preferably CBD according to SEQ ID NO. 1, and/or a nucleotide sequence having at least 90% sequence identity with SEQ ID NO. 1.

Another embodiment of the current inventive technology includes an expression vector having a nucleotide sequence, operably linked to a promoter, encoding a UGT having glycosylation activity towards one or more cannabinoid compounds, and preferably CBD according to SEQ ID NO. 2, or a fragment thereof, and/or an amino acid sequence having at least 90% sequence identity with SEQ ID NO. 2.

Another embodiment of the current inventive technology includes an expression vector having a nucleotide sequence, operably linked to a promoter, encoding a UGT having glycosylation activity towards one or more cannabinoid compounds, and preferably CBD according to SEQ ID NO. 1, and/or a nucleotide sequence having at least 90% sequence identity with SEQ ID NO. 1.

Another embodiment of the current inventive technology includes one or more organisms, such as a plant, plant cell, bacteria, algae, fungi, or yeast cell, transformed by an expression vector having a nucleotide sequence encoding a UGT having glycosylation activity towards one or more cannabinoid compounds, and preferably CBD according to SEQ ID NO. 2, or a fragment thereof, and/or an amino acid sequence having at least 90% sequence identity with SEQ ID NO. 2.

Another embodiment of the current inventive technology includes one or more organisms, such as a plant, plant cell, bacteria, algae, fungi, or yeast cell, transformed by an expression vector having a nucleotide sequence encoding a UGT having glycosylation activity towards one or more cannabinoid compounds, and preferably CBD according to SEQ ID NO. 1, and/or a nucleotide sequence having at least 90% sequence identity with SEQ ID NO. 1.

Another embodiment of the current inventive technology includes the isolated peptide having an amino acid sequence encoding a UGT having glycosylation activity towards one or more cannabinoid compounds, and preferably CBN according to SEQ ID NO. 4, 6, 8, and 10, or a fragment thereof, and/or an amino acid sequence having at least 90% sequence identity with SEQ ID NO. 4, 6, 8, and 10.

Another embodiment of the current inventive technology includes the isolated nucleotide sequence encoding a UGT having glycosylation activity towards one or more cannabinoid

compounds, and preferably CBN according to SEQ ID NO. 3, 5, 7, and 9, and/or a nucleotide sequence having at least 90% sequence identity with SEQ ID NO. 3, 5, 7, and 9.

Another embodiment of the current inventive technology includes an expression vector having a nucleotide sequence, operably linked to a promoter, encoding a UGT having glycosylation activity towards one or more cannabinoid compounds, and preferably CBN according to SEQ ID NO. 4, 6, 8, and 10, or a fragment thereof, and/or an amino acid sequence having at least 90% sequence identity with SEQ ID NO. 4, 6, 8, and 10.

Another embodiment of the current inventive technology includes an expression vector having a nucleotide sequence, operably linked to a promoter, encoding a UGT having glycosylation activity towards one or more cannabinoid compounds, and preferably CBN according to SEQ ID NO. 3, 5, 7, and 9, and/or a nucleotide sequence having at least 90% sequence identity with SEQ ID NO. 3, 5, 7, and 9.

Another embodiment of the current inventive technology includes one or more organisms, such as a plant, plant cell, bacteria, algae, fungi, or yeast cell, transformed by an expression vector having a nucleotide sequence encoding a UGT having glycosylation activity towards one or more cannabinoid compounds, and preferably CBN according to SEQ ID NO. 4, 6, 8, and 10, or a fragment thereof, and/or an amino acid sequence having at least 90% sequence identity with SEQ ID NO. 4, 6, 8, and 10.

Another embodiment of the current inventive technology includes one or more organisms, such as a plant, plant cell, bacteria, algae, fungi, or yeast cell, transformed by an expression vector having a nucleotide sequence encoding a UGT having glycosylation activity towards one or more cannabinoid compounds, and preferably CBN according to SEQ ID NO. 3, 5, 7, and 9, and/or a nucleotide sequence having at least 90% sequence identity with SEQ ID NO. 3, 5, 7, and 9.

In another embodiment, the invention novel systems, methods and compositions for the production of water-soluble cannabinoid glycosides in yeast or bacteria cells. In one preferred embodiment, a plant, such as yeast or bacterial cell, may be genetically modified to express one or more heterologous UGTs, preferably a UGT selected from the group consisting of SEQ ID NO. 2, 4, 6, 8 and 10, or a fragment thereof. In one preferred embodiment, a culture of yeast cells, such as *Saccharomyces cerevisiae*, *Kluyveromyces marxianus,* or *Pichia pastoris* or other suitable yeast species, may be established in a fermenter or other similar apparatus. It should be noted that the use of the above identified example in this embodiment is exemplary only, as various yeast strains,

mixes of strains, hybrids of different strains or clones may be used to generate a suspension culture. In certain cases, such fermenters may include large industrial-scale fermenters allowing for a large quantity of yeast and bacterial cells to be grown. In this embodiment, it may be possible to culture a large quantity of cells from a single-strain of, for example, *S. cerevisiae*, *P. pastoris*, or *K. marxianus*, or a bacterium such as *E. Coli*, or *Bacillus subtilis*, which may establish a cell culture having a consistent rate of cannabinoid modification. Such cultured growth may be continuously sustained with the continual addition of nutrient and other growth factors being added to the culture. Such features may be automated or accomplished manually.

The water-soluble cannabinoids, such as CBD-1G, CBD-2G, CBD-3G, and CBD-4G, may be extracted from the cell culture's supernatant/media, or from the cells. In this preferred embodiment, a transformed yeast or bacteria cells may be lysed such that accumulated cannabinoid glycosides are released to the surrounding lysate. Additional steps may include treating this lysate. Examples of such treatment may include filtering, centrifugation or screening to remove extraneous cellular material as well as chemical treatments to improve later cannabinoid glycoside yields.

The cannabinoid glycosides, such as CBD-1G, CBD-2G, CBD-3G, and CBD-4G may be further isolated and purified. In one preferred embodiment, the culture's supernatant/media, or the cell's lysate may be processed utilizing affinity chromatography or other purification methods. In this preferred embodiment, an affinity column having a ligand configured to bind with one or more of the cannabinoid glycosides such that they may be immobilized or coupled to a solid support. The material may then be passed over the column such that the cannabinoid glycosides, having specific binding affinity to the ligand become bound and immobilized. In some embodiments, non-binding and non-specific binding proteins that may have been present in the lysate may be removed. Finally, the cannabinoid glycosides may be eluted or displaced from the affinity column by, for example, a corresponding sugar or other compound that may displace or disrupt the cannabinoid-ligand bond. The eluted cannabinoid glycosides may be collected and further purified or processed.

Another embodiment of the current invention may include systems, methods and compositions for the generation of water soluble cannabinoid glycoside compounds in whole plants and plant cell cultures. In this preferred embodiment, this N-terminal trichome targeting sequence or domain may generally include the first 28 amino acid residues of a generalized

synthase and may be coupled with a UGT, and preferably a UGT selected from the group consisting of SEQ ID NO. 2, 4, 6, 8 and 10, or a fragment thereof. Exemplary N-terminal trichome targeting sequence for THCA synthase and CBDA synthase are identified by Sayre et al., PCT/US18/41710, such sequences being specifically incorporated here by reference. This extracellular targeting sequence may be recognized by the plant cell and cause the transport of the UGT from the cytoplasm to the plant's trichrome, and in particular the storage compartment of the plant trichrome where extracellular cannabinoid glycosylation may occur. More specifically, in this preferred embodiment, one or more UGT, and preferably a UGT selected from the group consisting of SEQ ID NO. 2, 4, 6, 8 and 10, or a fragment thereof., may either be engineered to express all or part of the N-terminal extracellular targeting sequence as present in an exemplary synthase enzyme.

Generally, a trichome structure, such as in *Cannabis*, will have limited substrate for a UGT to use to effectuate glycosylation. To resolve this problem, in one embodiment, the invention may include systems, methods and compositions to increase substrates for UGTs in a plant trichome structure. In this preferred embodiment, an exogenous or endogenous UDP-glucose/UDP-galactose transporter may be expressed in a trichome producing plant, such as *Cannabis* plant. exemplary sequences being identified by Sayre et al., PCT/US18/41710, such sequences being specifically incorporated here by reference. In this embodiment, the UDP-glucose/UDP-galactose transporter may be modified to include a plasma-membrane targeting sequence and/or domain, exemplary sequences being identified by Sayre et al., PCT/US18/41710, such sequences being specifically incorporated here by reference. With this targeting domain, the UDP-glucose/UDP-galactose transporter may allow the artificial fusion protein to be anchored to the plasma membrane. In this configuration, sugar substrates from the cytosol may pass through the plasma membrane bound UDP-glucose/UDP-galactose transporter into the trichome. In this embodiment, substrates for UGTs may be localized to the trichome and allowed to accumulate further allowing enhanced glycosylation of cannabinoids in the trichome.

In additional embodiments, such plants or cell cultures may be genetically modified to direct cannabinoid synthesis to the cytosol, as opposed to a trichome structure. In one preferred embodiment, cannabinoid biosynthesis may be redirected from the plant's trichome to be localized in the plant cell's cytosol. In certain embodiments, a cytosolic cannabinoid production system may be established as described in PCT/US18/24409 and PCT/US18/41710, both by Sayre et al. (these applications are both incorporated by reference with respect to their disclosure related to cytosolic

cannabinoid production and/or modification in whole, and plant cell systems). In one embodiment, a cytosolic cannabinoid production system may include the *in vivo* creation of one or more recombinant proteins that may allow cannabinoid biosynthesis to be localized to the cytosol where one or more heterologous UGT proteins may also be expressed and present in the cytosol. This inventive feature allows not only higher levels of cannabinoid production and accumulation, but efficient production of cannabinoids in suspension cell cultures. Even more importantly, this inventive feature allows cannabinoid glycoside production and accumulation without a trichome structure in whole plants, allowing cells that would not traditionally produce cannabinoids, such as cells in *Cannabis* leaves and stalks, to become cannabinoid-producing cells

More specifically, in this preferred embodiment, one or more cannabinoid synthases may be modified to remove all or part of an N-terminal extracellular trichome targeting. Exemplary N-terminal trichome targeting sequence for THCA synthase and CBDA synthase are identified by Sayre et al., PCT/US18/41710. Co-expression with this cytosolic-targeted synthase with a heterologous UGT may allow the localization of cannabinoid synthesis to the cytosol. As noted below, in certain embodiments cannabinoid biosynthesis may be coupled with cannabinoid glycosylation in a cell cytosol. For example, in one preferred embodiment a UGT (for example SEQ ID NOs. 2, 4, 6, 8 and 10, or a fragment thereof) may be expressed in a cell, preferably a cannabinoid producing cell, and even more preferably a *Cannabis* cell and be further engineered to be directed to the cell's cytosol. Such cytosolic targeted UGT enzymes may be co-expressed with heterologous catalase and cannabinoid transporters or other genes that may reduce cannabinoid biosynthesis toxicity and/or facilitate transport through or out of the cell. In certain embodiments, a catalase enzyme may be co-expressed with a UGT of the invention. In one embodiment a heterologous catalase is selected from the group of catalase sequences identified in PCT/US18/24409 and PCT/US18/41710, both by Sayre et al., such catalase sequences being incorporated herein by reference. Such cytosolic targeted enzymes may also be co-expressed with one or more myb transcriptions factors that may enhance metabolite flux through the cannabinoid biosynthetic pathway which may increase cannabinoid production. In one embodiment a myb transcription factor may be endogenous to *Cannabis*, or an ortholog thereof. Examples of endogenous or endogenous like, myb transcription factor may include those identified by Sayre et al., PCT/US18/41710, such specific sequences being incorporated herein by reference.

Notably, in a preferred embodiment, one or more endogenous cannabinoid synthase genes may be disrupted and/or knocked out and replaced with cytosolic-targeted cannabinoid synthase proteins as described herein. The disrupted endogenous cannabinoid synthase gene(s) may be the same or different than the expressed cytosolic-targeted cannabinoid synthase protein. Methods of disrupting or knocking-out a gene are known in the art and could be accomplished by one of ordinary skill without undue experimentation, for example through CRISPR, Talen, and zinc-finger exonuclease systems, as well as heterologous recombination techniques.

In another embodiment, one or more endogenous cannabinoid synthase genes may be disrupted and/or knocked out in a *Cannabis* plant or suspension cell culture wherein one or more cannabinoid synthase genes has been disrupted and/or knocked out is selected from the group consisting of: a CBG synthase gene; a THCA synthase, a CBDA synthase, and a CBCA synthase. In this embodiment, the *Cannabis* plant or suspension cell culture may express a polynucleotide encoding one or more cannabinoid synthases having its trichome targeting sequence disrupted and/or removed which may be selected from the group consisting of: a CBG synthase gene having its trichome targeting sequence disrupted and/or removed; a THCA synthase having its trichome targeting sequence disrupted and/or removed; a CBDA synthase having its trichome targeting sequence disrupted and/or removed; and a CBCA synthase having its trichome targeting sequence disrupted and/or removed.

As used herein, a "cannabinoid" is a chemical compound (such as cannabinol, THC or cannabidiol) that is found in the plant species *Cannabis* among others like: *Echinacea*; *Acmella Oleracea*; *Helichrysum Umbraculigerum*; *Radula Marginata* (Liverwort) and *Theobroma Cacao*, and metabolites and synthetic analogues thereof that may or may not have psychoactive properties. Cannabinoids therefore include (without limitation) compounds (such as THC) that have high affinity for the cannabinoid receptor (for example Ki<250 nM), and compounds that do not have significant affinity for the cannabinoid receptor (such as cannabidiol, CBD). Cannabinoids also include compounds that have a characteristic dibenzopyran ring structure (of the type seen in THC) and cannabinoids which do not possess a pyran ring (such as cannabidiol). Hence a partial list of cannabinoids includes THC, CBD, dimethyl heptylpentyl cannabidiol (DMHP-CBD), 6,12-dihydro-6-hydroxy-cannabidiol (described in U.S. Pat. No. 5,227,537, incorporated by reference); (3S,4R)-7-hydroxy-Δ6-tetrahydrocannabinol homologs and derivatives described in U.S. Pat. No. 4,876,276, incorporated by reference; (+)-4-[4-DMH-2,6-diacetoxy-phenyl]-2-carboxy-6,6-

dimethylbicyclo[3.1.1]hept-2-en, and other 4-phenylpinene derivatives disclosed in U.S. Pat. No. 5,434,295, which is incorporated by reference; and cannabidiol (−)(CBD) analogs such as (−)CBD-monomethylether, (−)CBD dimethyl ether; (−)CBD diacetate; (−)3′-acetyl-CBD monoacetate; and ±AF11, all of which are disclosed in Consroe et al., J. Clin. Phannacol. 21:428S-436S, 1981, which is also incorporated by reference. Many other cannabinoids are similarly disclosed in Agurell et al., Pharmacol. Rev. 38:31-43, 1986, which is also incorporated by reference.

Examples of cannabinoids are tetrahydrocannabinol, cannabidiol, cannabigerol, cannabichromene, cannabicyclol, cannabivarin, cannabielsoin, cannabicitran, cannabigerolic acid, cannabigerolic acid monomethylether, cannabigerol monomethylether, cannabigerovarinic acid, cannabigerovarin, cannabichromenic acid, cannabichromevarinic acid, cannabichromevarin, cannabidolic acid, cannabidiol monomethylether, cannabidiol-C4, cannabidivarinic acid, cannabidiorcol, delta-9-tetrahydrocannabinolic acid A, delta-9- tetrahydrocannabinolic acid B, delta-9-tetrahydrocannabinolic acid-C4, delta-9- tetrahydrocannabivarinic acid,delta-9-tetrahydrocannabivarin, delta-9- tetrahydrocannabiorcolic acid, delta-9-tetrahydrocannabiorcol,delta-7-cis-iso- tetrahydrocannabivarin, delta-8-tetrahydrocannabiniolic acid, delta-8- tetrahydrocannabinol, cannabicyclolic acid, cannabicylovarin, cannabielsoic acid A, cannabielsoic acid B, cannabinolic acid, cannabinol methylether, cannabinol-C4, cannabinol-C2, cannabiorcol, 10-ethoxy-9-hydroxy-delta-6a-tetrahydrocannabinol, 8,9-dihydroxy-delta-6a-tetrahydrocannabinol, cannabitriolvarin, ethoxy- cannabitriolvarin, dehydrocannabifuran, cannabifuran, cannabichromanon, cannabicitran, 10-oxo-delta-6a-tetrahydrocannabinol, delta-9-cis- tetrahydrocannabinol, 3, 4, 5, 6-tetrahydro-7-hydroxy-alpha-alpha-2-trimethyl-9-n- propyl-2, 6-methano-2H-1-benzoxocin-5-methanol-cannabiripsol,trihydroxy-delta-9-tetrahydrocannabinol, and cannabinol. Examples of cannabinoids within the context of this disclosure include tetrahydrocannabinol and cannabidiol. The term “cannabinoid” may also include different modified forms of a cannabinoid such as a hydroxylated cannabinoid or cannabinoid carboxylic acid. For example, if a UGT were to be capable of glycosylating a cannabinoid, it would include the term cannabinoid as defined elsewhere, as well as the aforementioned modified forms. It may further include multiple glycosylation moieties.

A protein has “homology” or “homology” to a second protein if the amino acid sequence encoded by a gene has a similar amino acid sequence to that of the second gene. Alternatively, a

protein has homology to a second protein if the two proteins have "similar" amino acid sequences. (Thus, the term "homologous proteins" is defined to mean that the two proteins have similar amino acid sequences). More specifically, in certain embodiments, the term "homologous" with regard to a contiguous nucleic acid sequence, refers to contiguous nucleotide sequences that hybridize under appropriate conditions to the reference nucleic acid sequence. For example, homologous sequences may have from about 75%-100, or more generally 80% to 100% sequence identity, such as about 81%; about 82%; about 83%; about 84%; about 85%; about 86%; about 87%; about 88%; about 89%; about 90%; about 91%; about 92%; about 93%; about 94% about 95%; about 96%; about 97%; about 98%; about 98.5%; about 99%; about 99.5%; and about 100%. The property of substantial homology is closely related to specific hybridization. For example, a nucleic acid molecule is specifically hybridizable when there is a sufficient degree of complementarity to avoid non-specific binding of the nucleic acid to non-target sequences under conditions where specific binding is desired, for example, under stringent hybridization conditions, and would fall within the range of a homolog. In another embodiment, expression optimization, for example for a mammalian lipocalin or odorant binding protein, to be expressed in yeast may be considered homologous and having a variable sequence identity due to the variable codon positions. Additional embodiments may also include homology to include redundant nucleotide codons.

The term "operably linked," when used in reference to a regulatory sequence and a coding sequence, means that the regulatory sequence affects the expression of the linked coding sequence. "Regulatory sequences," or "control elements," refer to nucleotide sequences that influence the timing and level/amount of transcription, RNA processing or stability, or translation of the associated coding sequence. Regulatory sequences may include promoters; translation leader sequences; introns; enhancers; stem-loop structures; repressor binding sequences; termination sequences; polyadenylation recognition sequences; etc. Particular regulatory sequences may be located upstream and/or downstream of a coding sequence operably linked thereto. Also, particular regulatory sequences operably linked to a coding sequence may be located on the associated complementary strand of a double-stranded nucleic acid molecule.

As used herein, the term "promoter" refers to a region of DNA that may be upstream from the start of transcription, and that may be involved in recognition and binding of RNA polymerase and other proteins to initiate transcription. A promoter may be operably linked to a coding sequence for expression in a cell, or a promoter may be operably linked to a nucleotide sequence

encoding a signal sequence which may be operably linked to a coding sequence for expression in a cell. An "inducible" promoter may be a promoter which may be under environmental control. Tissue-specific, tissue-preferred, cell type specific, and inducible promoters constitute the class of "non-constitutive" promoters. A "constitutive" promoter is a promoter which may be active under most environmental conditions or in most cell or tissue types.

As used herein, the term "transformation" or "genetically modified" refers to the transfer of one or more nucleic acid molecule(s) into a cell. A plant is "transformed" or "genetically modified" by a nucleic acid molecule transduced into the plant when the nucleic acid molecule becomes stably replicated by the plant. As used herein, the term "transformation" or "genetically modified" encompasses all techniques by which a nucleic acid molecule can be introduced into, such as a plant.

The term "vector" refers to some means by which DNA, RNA, a protein, or polypeptide can be introduced into a host. The polynucleotides, protein, and polypeptide which are to be introduced into a host can be therapeutic or prophylactic in nature; can encode or be an antigen; or can be regulatory in nature, etc. There are various types of vectors including virus, plasmid, bacteriophages, cosmids, and bacteria. An "expression vector" is nucleic acid capable of replicating in a selected host cell or organism. An expression vector can replicate as an autonomous structure, or alternatively can integrate, in whole or in part, into the host cell chromosomes or the nucleic acids of an organelle, or it is used as a shuttle for delivering foreign DNA to cells, and thus replicate along with the host cell genome. Thus, an expression vector are polynucleotides capable of replicating in a selected host cell, organelle, or organism, e.g., a plasmid, virus, artificial chromosome, nucleic acid fragment, and for which certain genes on the expression vector (including genes of interest) are transcribed and translated into a polypeptide or protein within the cell, organelle or organism; or any suitable construct known in the art, which comprises an "expression cassette." In contrast, as described in the examples herein, a "cassette" is a polynucleotide containing a section of an expression vector of this invention. The use of a cassette assists in the assembly of the expression vectors. An expression vector is a replicon, such as plasmid, phage, virus, chimeric virus, or cosmid, and which contains the desired polynucleotide sequence operably linked to the expression control sequence(s).

As is known in the art, different organisms preferentially utilize different codons for generating polypeptides. Such "codon usage" preferences may be used in the design of nucleic

acid molecules encoding the proteins and chimeras of the invention in order to optimize expression in a particular host cell system. For example, all nucleotides of the present invention may be optimized for expression in a select organisms, such as a *Cannabis* plant, yeast, algae, fungi, and bacteria.

A polynucleotide sequence is operably linked to an expression control sequence(s) (e.g., a promoter and, optionally, an enhancer) when the expression control sequence controls and regulates the transcription and/or translation of that polynucleotide sequence.

Unless otherwise indicated, a particular nucleic acid sequence also implicitly encompasses conservatively modified variants thereof (e.g., degenerate codon substitutions), the complementary (or complement) sequence, and the reverse complement sequence, as well as the sequence explicitly indicated. Specifically, degenerate codon substitutions may be achieved by generating sequences in which the third position of one or more selected (or all) codons is substituted with mixed-base and/or deoxyinosine residues (see e.g., Batzer et al., Nucleic Acid Res. 19:5081 (1991); Ohtsuka et al., J. Biol. Chem. 260:2605-2608 (1985); and Rossolini et al., Mol. Cell. Probes 8:91-98 (1994)). Because of the degeneracy of nucleic acid codons, one can use various different polynucleotides to encode identical polypeptides. The Table below, contains information about which nucleic acid codons encode which amino acids.

Amino acid Nucleic acid codons

Amino Acid	Nucleic Acid Codons
Ala/A	GCT, GCC, GCA, GCG
Arg/R	CGT, CGC, CGA, CGG, AGA, AGG
Asn/N	AAT, AAC
Asp/D	GAT, GAC
Cys/C	TGT, TGC
Gln/Q	CAA, CAG
Glu/E	GAA, GAG
Gly/G	GGT, GGC, GGA, GGG
His/H	CAT, CAC
Ile/I	ATT, ATC, ATA
Leu/L	TTA, TTG, CTT, CTC, CTA, CTG

Lys/K	AAA, AAG
Met/M	ATG
Phe/F	TTT, TTC
Pro/P	CCT, CCC, CCA, CCG
Ser/S	TCT, TCC, TCA, TCG, AGT, AGC
Thr/T	ACT, ACC, ACA, ACG
Trp/W	TGG
Tyr/Y	TAT, TAC
Val/V	GTT, GTC, GTA, GTG

Moreover, because the proteins are described herein, one can chemically synthesize a polynucleotide which encodes these polypeptides/chimeric proteins. Oligonucleotides and polynucleotides that are not commercially available can be chemically synthesized e.g., according to the solid phase phosphoramidite triester method first described by Beaucage and Caruthers, Tetrahedron Letts. 22:1859-1862 (1981), or using an automated synthesizer, as described in Van Devanter et al., Nucleic Acids Res. 12:6159- 6168 (1984). Other methods for synthesizing oligonucleotides and polynucleotides are known in the art. Purification of oligonucleotides is by either native acrylamide gel electrophoresis or by anion-exchange HPLC as described in Pearson & Reanier, J. Chrom. 255:137-149 (1983).

The term "plant" or "plant system" includes whole plants, plant organs, progeny of whole plants or plant organs, embryos, somatic embryos, embryo-like structures, protocorms, protocorm-like bodies (PLBs), and culture and/or suspensions of plant cells. Plant organs comprise, e.g., shoot vegetative organs/structures (e.g., leaves, stems and tubers), roots, flowers and floral organs/structures (e.g., bracts, sepals, petals, stamens, carpels, anthers and ovules), seed (including embryo, endosperm, and seed coat) and fruit (the mature ovary), plant tissue (e.g., vascular tissue, ground tissue, and the like) and cells (e.g., guard cells, egg cells, trichomes and the like). The invention may also include *Cannabaceae* and other *Cannabis* strains, such as *C. sativa* generally.

The term "expression," as used herein, or "expression of a coding sequence" (for example, a gene or a transgene) refer to the process by which the coded information of a nucleic acid transcriptional unit (including, e.g., genomic DNA or cDNA) is converted into an operational, non-operational, or structural part of a cell, often including the synthesis of a protein. Gene expression

can be influenced by external signals; for example, exposure of a cell, tissue, or organism to an agent that increases or decreases gene expression. Expression of a gene can also be regulated anywhere in the pathway from DNA to RNA to protein. Regulation of gene expression occurs, for example, through controls acting on transcription, translation, RNA transport and processing, degradation of intermediary molecules such as mRNA, or through activation, inactivation, compartmentalization, or degradation of specific protein molecules after they have been made, or by combinations thereof. Gene expression can be measured at the RNA level or the protein level by any method known in the art, including, without limitation, Northern blot, RT-PCR, Western blot, or *in vitro*, *in situ*, or *in vivo* protein activity assay(s).

The term "nucleic acid" or "nucleic acid molecules" include single- and double-stranded forms of DNA; single-stranded forms of RNA; and double-stranded forms of RNA (dsRNA). The term "nucleotide sequence" or "nucleic acid sequence" refers to both the sense and antisense strands of a nucleic acid as either individual single strands or in the duplex. The term "ribonucleic acid" (RNA) is inclusive of iRNA (inhibitory RNA), dsRNA (double stranded RNA), siRNA (small interfering RNA), mRNA (messenger RNA), miRNA (micro-RNA), hpRNA (hairpin RNA), tRNA (transfer RNA), whether charged or discharged with a corresponding acetylated amino acid), and cRNA (complementary RNA). The term "deoxyribonucleic acid" (DNA) is inclusive of cDNA, genomic DNA, and DNA-RNA hybrids. The terms "nucleic acid segment" and "nucleotide sequence segment," or more generally "segment," will be understood by those in the art as a functional term that includes both genomic sequences, ribosomal RNA sequences, transfer RNA sequences, messenger RNA sequences, operon sequences, and smaller engineered nucleotide sequences that encoded or may be adapted to encode, peptides, polypeptides, or proteins.

The term "gene" or "sequence" refers to a coding region operably joined to appropriate regulatory sequences capable of regulating the expression of the gene product (e.g., a polypeptide or a functional RNA) in some manner. A gene includes untranslated regulatory regions of DNA (e.g., promoters, enhancers, repressors, etc.) preceding (up-stream) and following (down-stream) the coding region (open reading frame, ORF) as well as, where applicable, intervening sequences (i.e., introns) between individual coding regions (i.e., exons). The term "structural gene" as used herein is intended to mean a DNA sequence that is transcribed into mRNA which is then translated into a sequence of amino acids characteristic of a specific polypeptide.

As used herein, "fragment" refers to a portion of a peptide or nucleotide sequence that still retains the activity of the whole.

Unless other stated, disclosure of a DNA sequence also include the corresponding RNA and amino acid sequence including all redundant codons and conservative amino acid substitutions, disclosure of a RNA sequence also include the corresponding DNA and amino acid sequence including all redundant codons and conservative amino acid substitutions, and finally disclosure of amino acid sequence also include the corresponding RNA and DNA sequence including all redundant codons and conservative amino acid substitutions and vice versa.

A nucleic acid molecule may include either or both naturally occurring and modified nucleotides linked together by naturally occurring and/or non-naturally occurring nucleotide linkages. Nucleic acid molecules may be modified chemically or biochemically, or may contain non-natural or derivatized nucleotide bases, as will be readily appreciated by those of skill in the art. Such modifications include, for example, labels, methylation, substitution of one or more of the naturally occurring nucleotides with an analog, internucleotide modifications (e.g., uncharged linkages: for example, methyl phosphonates, phosphotriesters, phosphoramidates, carbamates, etc.; charged linkages: for example, phosphorothioates, phosphorodithioates, etc.; pendent moieties: for example, peptides; intercalators: for example, acridine, psoralen, etc.; chelators; alkylators; and modified linkages: for example, alpha anomeric nucleic acids, etc.). The term "nucleic acid molecule" also includes any topological conformation, including single-stranded, double-stranded, partially duplexed, triplexed, hair-pinned, circular, and padlocked conformations.

The term "sequence identity" or "identity," as used herein in the context of two nucleic acid or polypeptide sequences, refers to the residues in the two sequences that are the same when aligned for maximum correspondence over a specified comparison window.

As used herein, "heterologous" or "exogenous" in reference to a nucleic acid is a nucleic acid that originates from a foreign species, or is synthetically designed, or, if from the same species, is substantially modified from its native form in composition and/or genomic locus by deliberate human intervention. A heterologous protein may originate from a foreign species or, if from the same species, is substantially modified from its original form by deliberate human intervention.

Unless otherwise defined, all terms of art, notations and other scientific terminology used herein are intended to have the meanings commonly understood by those of skill in the art to which this invention pertains. In some cases, terms with commonly understood meanings are defined herein

for clarity and/or for ready reference, and the inclusion of such definitions herein should not necessarily be construed to represent a substantial difference over what is generally understood in the art. The techniques and procedures described or referenced herein are generally well understood and commonly employed using conventional methodology by those skilled in the art, such as, for example, the widely utilized molecular cloning methodologies described in Sambrook et al., Molecular Cloning: A Laboratory Manual 3rd. edition (2001) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. and Current Protocols in Molecular Biology (Ausbel et al., eds., John Wiley & Sons, Inc. 2001. As appropriate, procedures involving the use of commercially available kits and reagents are generally carried out in accordance with manufacturer defined protocols and/or parameters unless otherwise noted.

Unless otherwise required by context, the use herein of the singular is to be read to include the plural and vice versa. The term "a" or "an" used in relation to an entity is to be read to refer to one or more of that entity. As such, the terms "a" (or "an"), "one or more," and "at least one" are used interchangeably herein. The terms "approximately" and "about" refer to a quantity, level, value, or amount that varies by as much as 30%, or in another embodiment by as much as 20%, and in a third embodiment by as much as 10% to a reference quantity, level, value or amount.

As used herein, the term "comprise," or variations thereof such as "comprises" or "comprising," are to be read to indicate the inclusion of any recited integer (e.g., a feature, element, characteristic, property, method/process step or limitation) or group of integers (e.g., features, element, characteristics, properties, method/process steps or limitations) but not the exclusion of any other integer or group of integers. Thus, as used herein the term "comprising" is inclusive or open-ended and does not exclude additional, unrecited integers or method/process steps.

SEQUENCE LISTING

SEQ ID NO. 1
DNA
GT3
Arabidopsis thaliana

GTTTGGCAAGTAAAAAACTGTTTACGTCATCTGTTACGTCTTCTGTTCACTTTATGTTTTACTCTCCACG
CATCTTATCCTTTATAAGCTCGCACAAATCTTAACCAAAACCAAAGCAAAACTTGAGAGTTTCTTACTAA
GGTTGAATCATGGTTTCCGAAACAACCAAATCTTCTCCACTTCACTTTGTTCTCTTCCCTTTCATGGCTC
AAGGCCACATGATTCCCATGGTTGATATTGCAAGGCTCTTGGCTCAGCGTGGTGTGATCATAACAATTGT
CACGACGCCTCACAATGCAGCGAGGTTCAAGAATGTCCTAAACCGTGCCATTGAGTCTGGCTTGCCCATC
AACTTAGTGCAAGTCAAGTTTCCATATCTAGAAGCTGGTTTGCAAGAAGGACAAGAGAATATCGATTCTC
TTGACACAATGGAGCGGATGATACCTTTCTTTAAAGCGGTTAACTTTCTCGAAGAACCAGTCCAGAAGCT
CATTGAAGAGATGAACCCTCGACCAAGCTGTCTAATTTCTGATTTTTGTTTGCCTTATACAAGCAAAATC
GCCAAGAAGTTCAATATCCCAAAGATCCTCTTCCATGGCATGGGTTGCTTTTGTCTTCTGTGTATGCATG
TTTTACGCAAGAACCGTGAGATCTTGGACAATTTAAAGTCAGATAAGGAGCTTTTCACTGTTCCTGATTT
TCCTGATAGAGTTGAATTCACAAGAACGCAAGTTCCGGTAGAAACATATGTTCCAGCTGGAGACTGGAAA
GATATCTTTGATGGTATGGTAGAAGCGAATGAGACATCTTATGGTGTGATCGTCAACTCATTTCAAGAGC
TCGAGCCTGCTTATGCCAAAGACTACAAGGAGGTAAGGTCCGGTAAAGCATGGACCATTGGACCCGTTTC
CTTGTGCAACAAGGTAGGAGCCGACAAAGCAGAGAGGGGAAACAAATCAGACATTGATCAAGATGAGTGC
CTTAAATGGCTCGATTCTAAGAAACATGGCTCGGTGCTTTACGTTTGTCTTGGAAGTATCTGTAATCTTC
CTTTGTCTCAACTCAAGGAGCTGGGACTAGGCCTAGAGGAATCCCAAAGACCTTTCATTTGGGTCATAAG
AGGTTGGGAGAAGTACAAAGAGTTAGTTGAGTGGTTCTCGGAAAGCGGCTTTGAAGATAGAATCCAAGAT
AGAGGACTTCTCATCAAAGGATGGTCCCCTCAAATGCTTATCCTTTCACATCCATCAGTTGGAGGGTTCC
TAACACACTGTGGTTGGAACTCGACTCTTGAGGGGATAACTGCTGGTCTACCGCTACTTACATGGCCGCT
ATTCGCAGACCAATTCTGCAATGAGAAATTGGTCGTTGAGGTACTAAAAGCCGGTGTAAGATCCGGGGTT
GAACAGCCTATGAAATGGGGAGAAGAGGAGAAAATAGGAGTGTTGGTGGATAAAGAAGGAGTGAAGAAGG
CAGTGGAAGAATTAATGGGTGAGAGTGATGATGCAAAAGAGAGAAGAAGAAGAGCCAAAGAGCTTGGAGA
TTCAGCTCACAAGGCTGTGGAAGAAGGAGGCTCTTCTCATTCTAACATCTCTTTCTTGCTACAAGACATA
ATGGAACTGGCAGAACCCAATAATTGAGTATACGTCATCTTTTTAAAGGAATTTAAAAATTAAATAGTTT
TGTTTTCTGTATTTGTGAAATTTAAAACAGAGTCTTAGTTACACATTTGTACGAATTAGCAGAAGACAAC
TAATTAGTGTAACTTTCACATACTCGAGTCATCTTTACTTTGAAAAAGTTGGTTTAAGGTTCGTTTAGAC
TTATCTAACATTAAGGCTGCGCCTAATTTTCCTAAACCAAGATCAATTGACATAACCTATTTATGTCCCT
TTTCAATGTCTCAAAACCATTTGAGGTGAAATTTCCACACAAATTTTGTTAAATGATTTTATTCCCCC

SEQ ID NO. 2
Amino Acid
GT3
Arabidopsis thaliana

MVSETTKSSPLHFVLFPFMAQGHMIPMVDIARLLAQRGVIITIVTTPHNAARFKNVLNRAIESGLPINLV
QVKFPYLEAGLQEGQENIDSLDTMERMIPFFKAVNFLEEPVQKLIEEMNPRPSCLISDFCLPYTSKIAKK
FNIPKILFHGMGCFCLLCMHVLRKNREILDNLKSDKELFTVPDFPDRVEFTRTQVPVETYVPAGDWKDIF
DGMVEANETSYGVIVNSFQELEPAYAKDYKEVRSGKAWTIGPVSLCNKVGADKAERGNKSDIDQDECLKW
LDSKKHGSVLYVCLGSICNLPLSQLKELGLGLEESQRPFIWVIRGWEKYKELVEWFSESGFEDRIQDRGL
LIKGWSPQMLILSHPSVGGFLTHCGWNSTLEGITAGLPLLTWPLFADQFCNEKLVVEVLKAGVRSGVEQP
MKWGEEEKIGVLVDKEGVKKAVEELMGESDDAKERRRRAKELGDSAHKAVEEGGSSHSNISFLLQDIMEL
AEPNN

SEQ ID NO. 3

DNA
GT4
Catharanthus roseus

CACTGTACCTCAATTCCATCTTCAATTTTTCCATTCATTTCATCATTTTGAGGTAGAAGAAGAAGAAGAG
GCCATTAATATGGTTAATCAGCTCCATATTTTCAACTTCCCATTCATGGCACAGGGCCATATGTTACCCG
CCTTAGACATGGCCAATCTATTCACTTCTCGTGGAGTCAAAGTAACATTAATCACAACCCATCAACATGT
TCCCATGTTTACAAAATCCATAGAAAGGAGCAGAAATTCTGGATTTGATATATCCATTCAATCCATCAAA
TTCCCAGCTTCAGAAGTTGGTTTACCTGAAGGAATCGAAAGTCTAGATCAAGTTTCAGGGGACGACGAAA
TGCTTCCTAAGTTCATGAGAGGAGTTAATTTACTCCAACAACCTCTCGAACAACTATTGCAAGAATCTCG
TCCTCATTGTCTTCTTTCTGATATGTTCTTCCCTTGGACTACTGAATCTGCTGCTAAATTTGGTATTCCC
AGATTGCTTTTTCATGGGTCCTGTTCCTTTGCCCTCTCTGCAGCTGAAAGTGTGAGAAGAAATAAACCTT
TCGAGAATGTTTCCACAGACACAGAGGAATTTGTTGTGCCTGATCTTCCCCACCAAATTAAATTAACCAG
AACACAAATTTCAACATACGAAAGGGAAAATATTGAGTCAGATTTTACCAAAATGCTGAAGAAAGTTAGG
GATTCAGAATCCACATCTTACGGAGTTGTAGTCAATAGTTTCTATGAACTTGAACCAGATTATGCCGATT
ATTACATCAACGTTTTGGGAAGAAAAGCATGGCATATAGGGCCTTTTTTGCTTTGTAACAAATTACAAGC
TGAAGATAAAGCCCAAAGGGGGAAGAAATCAGCAATTGATGCAGACGAATGTTTAAATTGGCTTGATTCG
AAACAACCAAATTCCGTAATTTATCTCTGTTTCGGAAGTATGGCCAATTTAAATTCTGCCCAATTACACG
AAATTGCAACAGCCCTTGAATCCTCCGGCCAAAATTTCATCTGGGTTGTTAGAAAATGTGTGGACGAAGA
AAACAGTTCAAAATGGTTTCCAGAAGGATTCGAAGAAAGAACAAAAGAAAAAGGGCTAATTATAAAGGGA
TGGGCACCACAAACCCTAATTCTTGAACACGAATCAGTAGGAGCATTTGTTACCCATTGTGGTTGGAATT
CAACTCTTGAAGGAATCTGCGCAGGGGTTCCTCTGGTGACTTGGCCTTTCTTTGCTGAGCAATTTTTCAA
TGAGAAATTGATTACAGAGGTACTGAAAACGGGATACGGAGTTGGGGCTCGGCAATGGAGTAGAGTTTCA
ACAGAGATTATAAAAGGAGAAGCCATAGCTAATGCTATTAATCGAGTAATGGTGGGTGATGAAGCTGTTG
AGATGAGAAACAGAGCAAAAGATTTGAAGGAAAAGGCAAGAAAAGCTTTGGAAGAAGATGGATCTTCTTA
TCGTGATCTTACTGCTCTTATTGAAGAATTGGGGGCATATCGTTCTCAAGTTGAAAGAAAGCAACAAGAC
TAGGAAATAAGAAGAGAGAGAGAATTCTCATTTATATTATTTTGTCCATTTAATCAAATTACTATATTCT
GTAAAATTTGCCATTTTTATTATGGAATCTAGAAGGCAGTCTATATATCTTCTTGCCAGGCACTAAAGAA
ATATTTGTCATTTCAAGTTTAAAAAAAAAAAAAAAAAAAA

SEQ ID NO. 4
Amino Acid
GT4
Catharanthus roseus

MVNQLHIFNFPFMAQGHMLPALDMANLFTSRGVKVTLITTHQHVPMFTKSIERSRNSGFDISIQSIKFPA
SEVGLPEGIESLDQVSGDDEMLPKFMRGVNLLQQPLEQLLQESRPHCLLSDMFFPWTTESAAKFGIPRLL
FHGSCSFALSAAESVRRNKPFENVSTDTEEFVVPDLPHQIKLTRTQISTYERENIESDFTKMLKKVRDSE
STSYGVVVNSFYELEPDYADYYINVLGRKAWHIGPFLLCNKLQAEDKAQRGKKSAIDADECLNWLDSKQP
NSVIYLCFGSMANLNSAQLHEIATALESSGQNFIWVVRKCVDEENSSKWFPEGFEERTKEKGLIIKGWAP
QTLILEHESVGAFVTHCGWNSTLEGICAGVPLVTWPFFAEQFFNEKLITEVLKTGYGVGARQWSRVSTEI
IKGEAIANAINRVMVGDEAVEMRNRAKDLKEKARKALEEDGSSYRDLTALIEELGAYRSQVERKQQD

SEQ ID NO. 5
DNA
GT5
Fragaria ananassa

ATGGCGATGGAAACCAAAAGCTGCCAGCAGCTGCATATTTTTTTTCTGCCGTTTATGGCGCGCGGCCATA
GCATTCCGCTGACCGATATTGCGAAACTGTTTAGCAGCCATGGCGCGCGCTGCACCATTGTGACCACCCC
GCTGAACGCGCCGCTGTTTAGCAAAGCGACCCAGCGCGGCGAAATTGAACTGGTGCTGATTAAATTTCCG
AGCGCGGAAGCGGGCCTGCCGCAGGATTGCGAAAGCGCGGATCTGATTACCACCCAGGATATGCTGGGCA
AATTTGTGAAAGCGACCTTTCTGATTGAACCGCATTTTGAAAAAATTCTGGATGAACATCGCCCGCATTG
CCTGGTGGCGGATGCGTTTTTTACCTGGGCGACCGATGTGGCGGCGAAATTTCGCATTCCGCGCCTGTAT
TTTCATGGCACCGGCTTTTTTGCGCTGTGCGCGAGCCTGAGCGTGATGATGTATCAGCCGCATAGCAACC
TGAGCAGCGATAGCGAAAGCTTTGTGATTCCGAACCTGCCGGATGAAATTAAAATGACCCGCAGCCAGCT
GCCGGTGTTTCCGGATGAAAGCGAATTTATGAAAATGCTGAAAGCGAGCATTGAAATTGAAGAACGCAGC
TATGGCGTGATTGTGAACAGCTTTTATGAACTGGAACCGGCGTATGCGAACCATTATCGCAAAGTGTTTG
GCCGCAAAGCGTGGCATATTGGCCCGGTGAGCTTTTGCAACAAAGCGATTGAAGATAAAGCGGAACGCGG
CAGCATTAAAAGCAGCACCGCGGAAAAACATGAATGCCTGAAATGGCTGGATAGCAAAAAACCGCGCAGC
GTGGTGTATGTGAGCTTTGGCAGCATGGTGCGCTTTGCGGATAGCCAGCTGCTGGAAATTGCGACCGGCC
TGGAAGCGAGCGGCCAGGATTTTATTTGGGTGGTGAAAAAAGAAAAAAAAGAAGTGGAAGAATGGCTGCC
GGAAGGCTTTGAAAAACGCATGGAAGGCAAAGGCCTGATTATTCGCGATTGGGCGCCGCAGGTGCTGATT
CTGGAACATGAAGCGATTGGCGCGTTTGTGACCCATTGCGGCTGGAACAGCATTCTGGAAGCGGTGAGCG
CGGGCGTGCCGATGATTACCTGGCCGGTGTTTGGCGAACAGTTTTATAACGAAAAACTGGTGACCGAAAT
TCATCGCATTGGCGTGCCGGTGGGCAGCGAAAAATGGGCGCTGAGCTTTGTGGATGTGAACGCGGAAACC
GAAGGCCGCGTGCGCCGCGAAGCGATTGAAGAAGCGGTGACCCGCATTATGGTGGGCGATGAAGCGGTGG
AAACCCGCAGCCGCGTGAAAGAACTGGGCGAAAACGCGCGCCGCGCGGTGGAAGAAGGCGGCAGCAGCTT
TCTGGATCTGAGCGCGCTGGTGGGCGAACTGAACGATCTGGCGTTTGGCGGCCTGGTGGAA

SEQ ID NO. 6

DNA

GT5

Fragaria ananassa

MAMETKSCQQLHIFFLPFMARGHSIPLTDIAKLFSSHGARCTIVTTPLNAPLFSKATQRGEIELVLIKFP
SAEAGLPQDCESADLITTQDMLGKFVKATFLIEPHFEKILDEHRPHCLVADAFFTWATDVAAKFRIPRLY
FHGTGFFALCASLSVMMYQPHSNLSSDSESFVIPNLPDEIKMTRSQLPVFPDESEFMKMLKASIEIEERS
YGVIVNSFYELEPAYANHYRKVFGRKAWHIGPVSFCNKAIEDKAERGSIKSSTAEKHECLKWLDSKKPRS
VVYVSFGSMVRFADSQLLEIATGLEASGQDFIWVVKKEKKEVEEWLPEGFEKRMEGKGLIIRDWAPQVLI
LEHEAIGAFVTHCGWNSILEAVSAGVPMITWPVFGEQFYNEKLVTEIHRIGVPVGSEKWALSFVDVNAET
EGRVRREAIEEAVTRIMVGDEAVETRSRVKELGENARRAVEEGGSSFLDLSALVGELNDLAFGGLVE

SEQ ID NO. 7

DNA

GT6

ATGGGCAGCGAGGGCAGGCAGCTCCACATCTTCATGTTCCCGTTCATGGCCCACGGCCACATGATCCCGATCGTGGACATGGCCAA
GCTCTTCGCCAGCAGGGGCATCAAGATCACCATCGTGACCACCCCGCTCAACAGCATCAGCATCAGCAAGAGCCTCCACAACTGCA
GCCCGAACAGCCTCATCCAGCTCCTCATCCTCAAGTTCCCGGCCGCCGAGGCCGGCCTCCCGGACGGCTGCGAGAACGCCGACAGC
ATCCCGAGCATGGACCTCCTCCCGAAGTTCTTCGAGGCCGTGAGCCTCCTCCAGCCGCCGTTCGAGGAGGCCCTCCACAACAACAG
GCCGGACTGCCTCATCAGCGACATGTTCTTCCCGTGGACCAACGACGTGGCCGACAGGGTGGGCATCCCGAGGCTCATCTTCCACG
GCACCAGCTGCTTCAGCCTCTGCAGCAGCGAGTTCATGAGGCTCCACAAGCCGTACCAGCACGTGAGCAGCGACACCGAGCCGTTC
ACCATCCCGTACCTCCCGGGCGACATCAAGCTCACCAAGATGAAGCTCCCGATCTTCGTGAGGGAGAACAGCGAGAACGAGTTC
AGCAAGTTCATCACCAAGGTGAAGGAGAGCGAGAGCTTCTGCTACGGCGTGGTGGTGAACAGCTTCTACGAGCTCGAGGCCGAGTA
CGTGGACTGCTACAAGGACGTGCTCGGCAGGAAGACCTGGACCATCGGCCCGCTCAGCCTCACCAACACCAAGACCCAGGAGATCA
CCCTCAGGGGCAGGGAGAGCGCCATCGACGAGCACGAGTGCCTCAAGTGGCTCGACAGCCAGAAGCCGAACAGCGTGGTGTACGTG
TGCTTCGGCAGCCTCGCCAAGTTCAACAGCGCCCAGCTCAAGGAGATCGCCATCGGCCTCGAGGCCAGCGGCAAGAAGTTCATCTG
GGTGGTGAGGAAGGGCAAGGGCGAGGAGGAGGAGGAGGAGCAGAACTGGCTCCCGGAGGGCTACGAGGAGAGGATGGAGGGCACCG

GCCTCATCATCAGGGGCTGGGCCCCGCAGGTGCTCATCCTCGACCACCCGAGCGTGGGCGGCTTCGTGACCCACTGCGGCTGGAAC
AGCACCCTCGAGGGCGTGGCCGCCGGCGTGCCGATGGTGACCTGGCCGGTGGGCGCCGAGCAGTTCTACAACGAGAAGCTCGTGAC
CGAGGTGCTCAAGACCGGCGTGGGCGTGGGCGTGCAGAAGTGGGCCCCGGGCGTGGGCGACTTCATCGAGAGCGAGGCCGTGGAGA
AGGCCATCAGGAGGATCATGGAGAAGGAGGGCGAGGAGATGAGGAACAGGGCCATCGAGCTCGGCAAGAAGGCCAAGTGGGCCGTG
GGCGAGGAGGGCAGCAGCTACAGCAACCTCGACGCCCTCATCGAGGAGCTCAAGAGCCTCGCCTTC

SEQ ID NO. 8
Amino Acid
GT6

MGSEGRQLHIFMFPFMAHGHMIPIVDMAKLFASRGIKITIVTTPLNSISISKSLHNCSPNSLIQLLILKF
PAAEAGLPDGCENADSIPSMDLLPKFFEAVSLLQPPFEEALHNNRPDCLISDMFFPWTNDVADRVGIPRL
IFHGTSCFSLCSSEFMRLHKPYQHVSSDTEPFTIPYLPGDIKLTKMKLPIFVRENSENEFSKFITKVKES
ESFCYGVVVNSFYELEAEYVDCYKDVLGRKTWTIGPLSLTNTKTQEITLRGRESAIDEHECLKWLDSQKP
NSVVYVCFGSLAKFNSAQLKEIAIGLEASGKKFIWVVRKGKGEEEEEEQNWLPEGYEERMEGTGLIIRGW
APQVLILDHPSVGGFVTHCGWNSTLEGVAAGVPMVTWPVGAEQFYNEKLVTEVLKTGVGVGVQKWAPGVG
DFIESEAVEKAIRRIMEKEGEEMRNRAIELGKKAKWAVGEEGSSYSNLDALIEELKSLAF

SEQ ID NO. 9
DNA
GT8
Fragaria vesca subsp. vesca scopoletin

ATGTCTTCTGATCCTCATAGAAAGCTTCATGTTGTTTTTTTTCCTTTTATGGCTTATGGACATATGATTC
CTACTCTTGATATGGCTAAGCTTTTTTCTTCTAGAGGAGCTAAGTCTACTATTCTTACTACTCCTCTTAA
TTCTAAGATTTTTCAAAAGCCTATTGAAAGATTTAAGAATCTTAATCCTTCTTTTGAAATTGATATTCAA
ATTTTTGATTTTCCTTGTGTTGATCTTGGACTTCCTGAAGGATGTGAAAATGTTGATTTTTTTACTTCTA
ATAATAATGATGATAGACAATATCTTACTCTTAAGTTTTTTAAGTCTACTAGATTTTTTAAGGATCAACT
TGAAAAGCTTCTTGAAACTACTAGACCTGATTGTCTTATTGCTGATATGTTTTTTCCTTGGGCTACTGAA
GCTGCTGAAAAGTTTAATGTTCCTAGACTTGTTTTTCATGGAACTGGATATTTTTCTCTTTGTTCTGAAT
ATTGTATTAGAGTTCATAATCCTCAAAATATTGTTGCTTCTAGATATGAACCTTTTGTTATTCCTGATCT
TCCTGGAAATATTGTTATTACTCAAGAACAAATTGCTGATAGAGATGAAGAATCTGAAATGGGAAAGTTT
ATGATTGAAGTTAAGGAATCTGATGTTAAGTCTTCTGGAGTTATTGTTAATTCTTTTTATGAACTTGAAC
CTGATTATGCTGATTTTTATAAGTCTGTTGTTCTTAAGAGAGCTTGGCATATTGGACCTCTTTCTGTTTA
TAATAGAGGATTTGAAGAAAAGGCTGAAAGAGGAAAGAAGGCTTCTATTAATGAAGTTGAATGTCTTAAG
TGGCTTGATTCTAAGAAGCCTGATTCTGTTATTTATATTTCTTTTGGATCTGTTGCTTGTTTTAAGAATG
AACAACTTTTTGAAATTGCTGCTGGACTTGAAACTTCTGGAGCTAATTTTATTTGGGTTGTTAGAAAGAA
TATTGGAATTGAAAAGGAAGAATGGCTTCCTGAAGGATTTGAAGAAAGAGTTAAGGGAAAGGGAATGATT
ATTAGAGGATGGGCTCCTCAAGTTCTTATTCTTGATCATCAAGCTACTTGTGGATTTGTTACTCATTGTG
GATGGAATTCTCTTCTTGAAGGAGTTGCTGCTGGACTTCCTATGGTTACTTGGCCTGTTGCTGCTGAACA
ATTTTATAATGAAAAGCTTGTTACTCAAGTTCTTAGAACTGGAGTTTCTGTTGGAGCTAAGAAGAATGTT
AGAACTACTGGAGATTTTATTTCTAGAGAAAAGGTTGTTAAGGCTGTTAGAGAAGTTCTTGTTGGAGAAG
AAGCTGATGAAAGAAGAGAAAGAGCTAAGAAGCTTGCTGAAATGGCTAAGGCTGCTGTTGAAGGAGGATC
TTCTTTTAATGATCTTAATTCTTTTATTGAAGAATTTACTTCT

SEQ ID NO. 10
Amino Acid
GT8
Fragaria vesca subsp. vesca scopoletin

MSSDPHRKLHVVFFPFMAYGHMIPTLDMAKLFSSRGAKSTILTTPLNSKIFQKPIERFKNLNPSFEIDIQ
IFDFPCVDLGLPEGCENVDFFTSNNNDDRQYLTLKFFKSTRFFKDQLEKLLETTRPDCLIADMFFPWATE
AAEKFNVPRLVFHGTGYFSLCSEYCIRVHNPQNIVASRYEPFVIPDLPGNIVITQEQIADRDEESEMGKF
MIEVKESDVKSSGVIVNSFYELEPDYADFYKSVVLKRAWHIGPLSVYNRGFEEKAERGKKASINEVECLK
WLDSKKPDSVIYISFGSVACFKNEQLFEIAAGLETSGANFIWVVRKNIGIEKEEWLPEGFEERVKGKGMI
IRGWAPQVLILDHQATCGFVTHCGWNSLLEGVAAGLPMVTWPVAAEQFYNEKLVTQVLRTGVSVGAKKNV
RTTGDFISREKVVKAVREVLVGEEADERRERAKKLAEMAKAAVEGGSSFNDLNSFIEEFTS

CLAIMS

What is claimed is:

1. A method of producing a cannabinoid glycoside comprising the step of:

- contacting a cannabinoid having at least one glycosylation site to a UDP-glucosyltransferase (UGT) peptide according to SEQ ID NO. 2, or a fragment thereof, or a UGT peptide having at least 90% sequence identity with SEQ ID NO. 2; and
- wherein said UGT catalyzes the bioconversion of the cannabinoid and one or more UDP-glucose substrates into a cannabinoid glycoside.

2. The method of claim 1, wherein said cannabinoid comprises cannabidiol (CBD).

3. The method of any of claims 1 and 2, wherein said cannabinoid glycoside comprises a CBD glycoside having between 1 and 4 UDP-glucose moieties.

4. The method of claim 1, wherein said step of contacting comprises the step of introducing selected from the group consisting of:

- contacting the cannabinoid with the UGT peptide in an *in vitro* system;
- contacting the cannabinoid with the UGT peptide in an in an *ex vivo* system; and
- contacting the cannabinoid with the UGT peptide in an in an *in vivo* system.

5. The method of claim 4, wherein said *ex vivo* system comprises a bioreactor system.

6. The method of claim 4, wherein said *in vitro* system comprises a synthetic cannabinoid synthesis system.

7. The method of claim 4, wherein said *in vivo* system comprises an *in vivo* system selected from the group consisting of: a *Cannabis* plant or part thereof, and a cell culture.

8. The method of claim 7, wherein said cell culture is selected from the group consisting of: a yeast cell culture, a bacterial cell culture, an algal cell culture, a fungi cell culture, and a plant cell culture.

9. A system for producing a cannabinoid glycoside comprising:

- an *in vivo*, *ex vivo* or *in vitro* system comprising:
  - a quantity of a cannabinoid having at least one glycosylation site;
  - a UDP-glucosyltransferase (UGT) peptide according to SEQ ID NO. 2, or a fragment thereof, or a UGT peptide having at least 90% sequence identity with SEQ ID NO. 2; and
  - wherein said UGT catalyzes the bioconversion of the cannabinoid and one or more UDP-glucose substrates into a cannabinoid glycoside in the *in vivo*, *ex vivo* or *in vitro* system.

10. The system of claim 9, wherein said cannabinoid comprises cannabidiol (CBD).

11. The system of any of claims 9 and 10, wherein said cannabinoid glycoside comprises a CBD glycoside.

12. The system of claim 9, wherein said CBD glycoside comprises a CBD having between 1 and 4 UDP-glucose moieties.

13. The system of claim 9, wherein said *ex vivo* system comprises a bioreactor system.

14. The system of claim 9, wherein said *in vitro* system comprises a synthetic cannabinoid synthesis system.

15. The system of claim 9, wherein said *in vivo* system comprises an *in vivo* system selected from the group consisting of: a *Cannabis* plant or part thereof, and a cell culture.

16. The system of claim 15, wherein said cell culture is selected from the group consisting of: a yeast cell culture, a bacterial cell culture, an algal cell culture, a fungi cell culture, and a plant cell culture.

17. An isolated nucleotide sequence encoding a UDP-glucosyltransferase (UGT) according to SEQ ID NO. 1, or a fragment thereof, or a UGT peptide having at least 90% sequence identity with SEQ ID NO. 1.

18. An expression vector encoding the nucleotide sequence according to claim 17, wherein said nucleotide sequence is operably linked to a promoter.

19. A cell transformed by the expression vector of claim 18.

20. The cell of claim 19, wherein the cell is selected from: a yeast cell, a bacterial cell, an algal cell, a fungi cell, and a plant cell.

21. A cell culture generated from the transformed call of claim 19 or 20.

22. A method of producing a cannabinoid glycoside comprising introducing a quantity of a cannabinoid to the cell culture of claim 21 and wherein the heterologously expressed UGT catalyzes the bioconversion of the cannabinoid and one or more UDP-glucose substrates into a cannabinoid glycoside.

23. A method of producing a cannabinoid glycoside comprising the step of:

- contacting a cannabinoid having at least one glycosylation site to a UDP-glucosyltransferase (UGT) peptide selected from SEQ ID NO's. 4, 6, 8, and 10, or a fragment thereof, or a UGT peptide having at least 90% sequence identity with SEQ ID NO's. 4, 6, 8, and 10; and
- wherein said UGT catalyzes the bioconversion of the cannabinoid and one or more UDP-glucose substrates into a cannabinoid glycoside.

24. The method of claim 23, wherein said cannabinoid comprises cannabinol (CBN).

25. The method of any of claims 23 and 24, wherein said cannabinoid glycoside comprises a CBN glycoside.

26. The method of claim 23, wherein said CBN glycoside comprises a CBN having at least one UDP-glucose moiety.

27. The method of claim 23, wherein said step of contacting comprises the step of introducing selected from the group consisting of:

- contacting the cannabinoid with the UGT peptide in an *in vitro* system;
- contacting the cannabinoid with the UGT peptide in an in an *ex vivo* system; and
- contacting the cannabinoid with the UGT peptide in an in an *in vivo* system.

28. The method of claim 27, wherein said *ex vivo* system comprises a bioreactor system.

29. The method of claim 27, wherein said *in vitro* system comprises a synthetic cannabinoid synthesis system.

30. The method of claim 27, wherein said *in vivo* system comprises an *in vivo* system selected from the group consisting of: a *Cannabis* plant or part thereof, and a cell culture.

31. The method of claim 30, wherein said cell culture is selected from the group consisting of: a yeast cell culture, a bacterial cell culture, an algal cell culture, a fungi cell culture, and a plant cell culture.

32. A system for producing a cannabinoid glycoside comprising:

- an *in vivo*, *ex vivo* or *in vitro* system comprising:
  - a quantity of a cannabinoid having at least one glycosylation site;
  - a UDP-glucosyltransferase (UGT) peptide selected from SEQ ID NO. 4, 6, 8, and 10, or a fragment thereof, or a UGT peptide having at least 90% sequence identity with SEQ ID NO. 4, 6, 8, and 10; and
  - wherein said UGT catalyzes the bioconversion of the cannabinoid and one or more UDP-glucose substrates into a cannabinoid glycoside in the *in vivo*, *ex vivo* or *in vitro* system.

33. The system of claim 32, wherein said cannabinoid comprises cannabinol (CBN).

34. The system of any of claims 32 and 33, wherein said cannabinoid glycoside comprises a CBN glycoside.

35. The system of claim 32, wherein said CBN glycoside comprises a CBD having at least one UDP-glucose moiety.

36. The system of claim 32, wherein said *ex vivo* system comprises a bioreactor system.

37. The system of claim 32, wherein said *in vitro* system comprises a synthetic cannabinoid synthesis system.

38. The system of claim 32, wherein said *in vivo* system comprises an *in vivo* system selected from the group consisting of: a *Cannabis* plant or part thereof, and a cell culture.

39. The system of claim 38, wherein said cell culture is selected from the group consisting of: a yeast cell culture, a bacterial cell culture, an algal cell culture, a fungi cell culture, and a plant cell culture.

40. An isolated nucleotide sequence encoding a UDP-glucosyltransferase (UGT) selected from SEQ ID NO. 3, 5, 7 and 9, or a fragment thereof, or a UGT peptide having at least 90% sequence identity with SEQ ID NO. 3, 5, 7 and 9.

41. An expression vector encoding the nucleotide sequence according to claim 40, wherein said nucleotide sequence is operably linked to a promoter.

42. A cell transformed by the expression vector of claim 41.

43. The cell of claim 42, wherein the cell is selected from: a yeast cell, a bacterial cell, an algal cell, a fungi cell, and a plant cell.

44. A cell culture generated from the transformed call of claim 42 or 43.

45. A method of producing a cannabinoid glycoside comprising introducing a quantity of a cannabinoid to the cell culture of claim 44 and wherein the heterologously expressed UGT catalyzes the bioconversion of the cannabinoid and one or more UDP-glucose substrates into a cannabinoid glycoside.

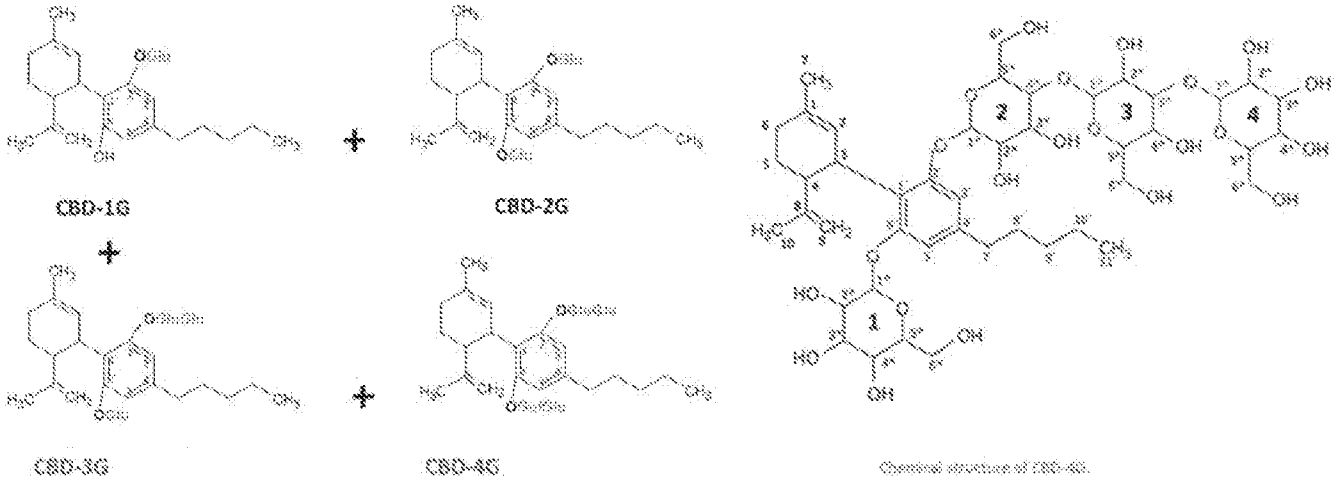


FIG. 1

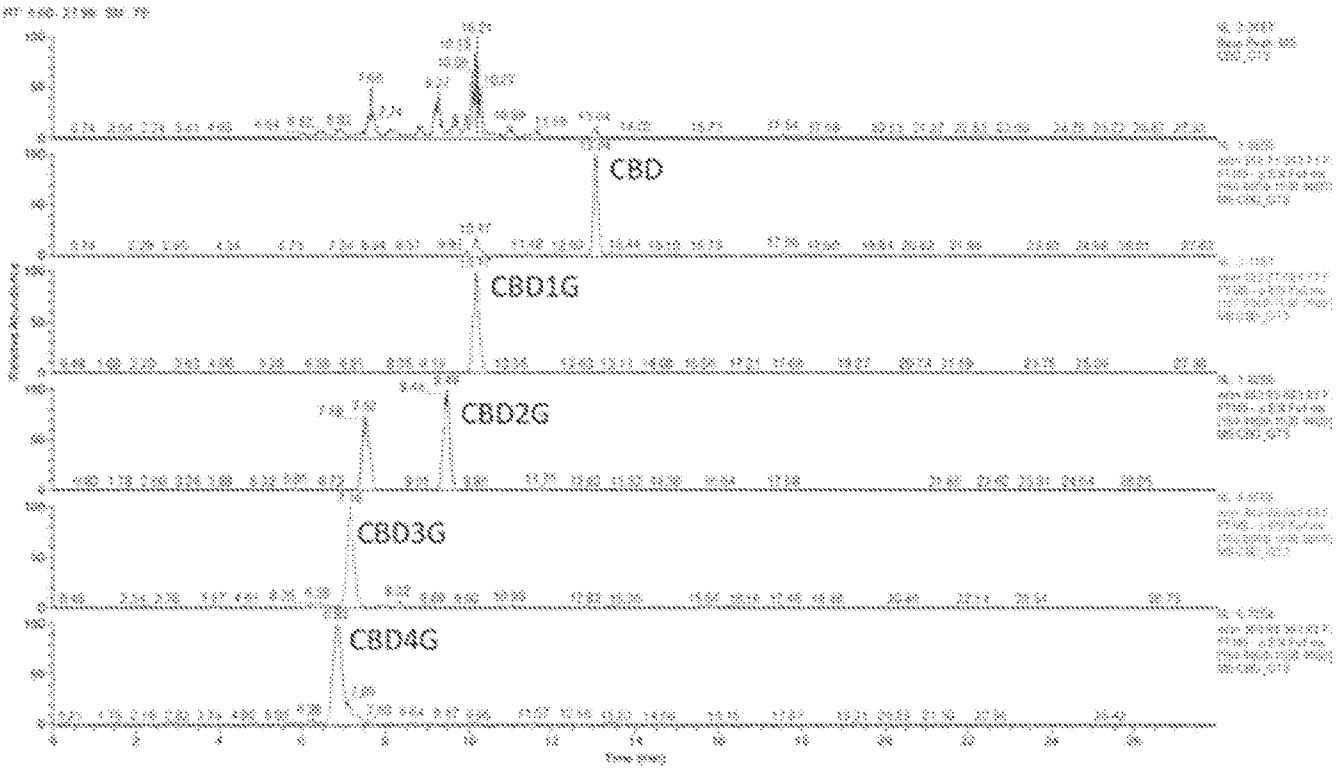


FIG. 2

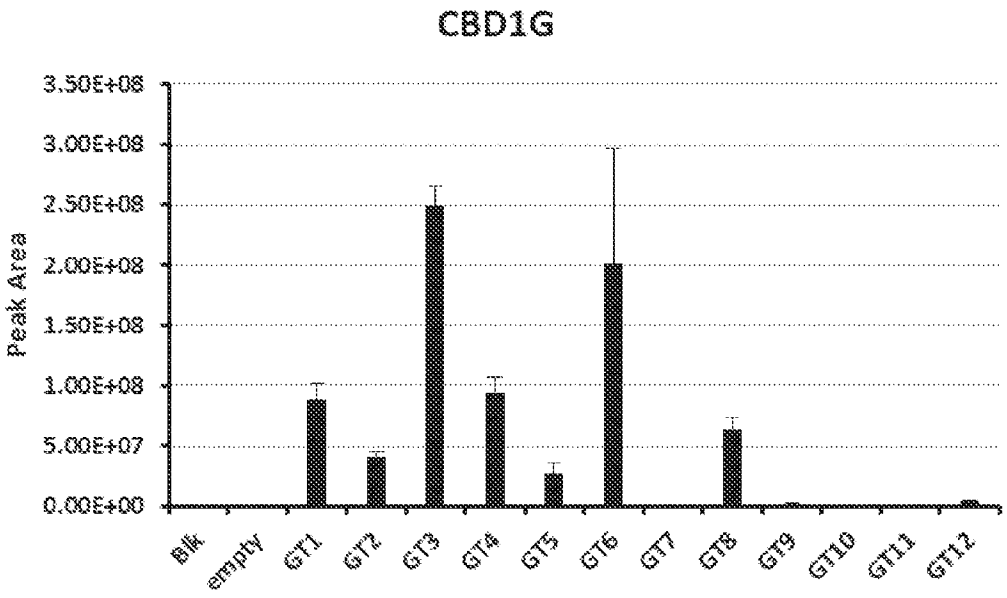

CBD1G
Peak Area
3.50E+08
3.00E+08
2.50E+08
2.00E+08
1.50E+08
1.00E+08
5.00E+07
0.00E+00
Blk
empty
GT1
GT2
GT3
GT4
GT5
GT6
GT7
GT8
GT9
GT10
GT11
GT12

FIG. 3

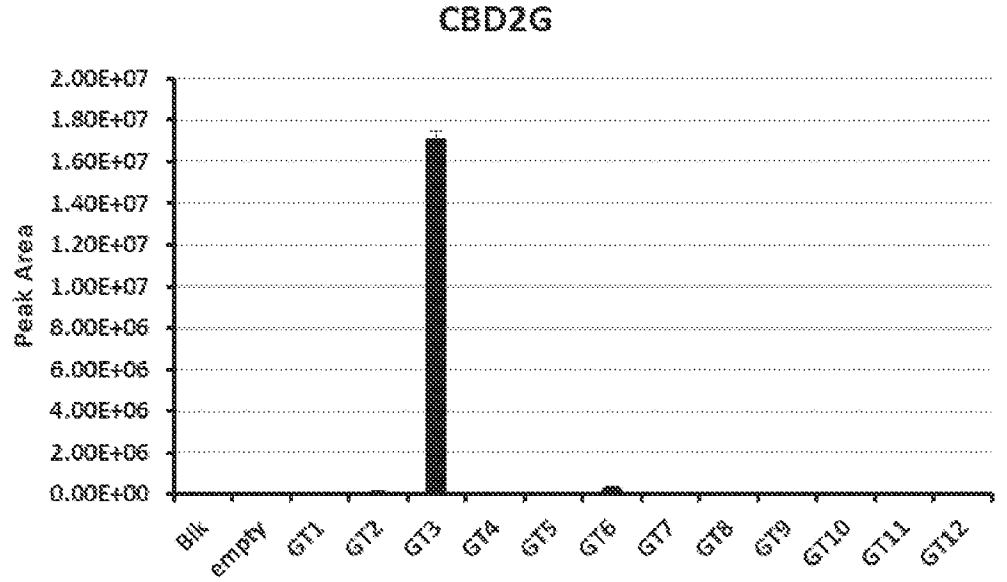

CBD2G
Peak Area
2.00E+07
1.80E+07
1.60E+07
1.40E+07
1.20E+07
1.00E+07
8.00E+06
6.00E+06
4.00E+06
2.00E+06
0.00E+00
Blk
empty
GT1
GT2
GT3
GT4
GT5
GT6
GT7
GT8
GT9
GT10
GT11
GT12

FIG. 4

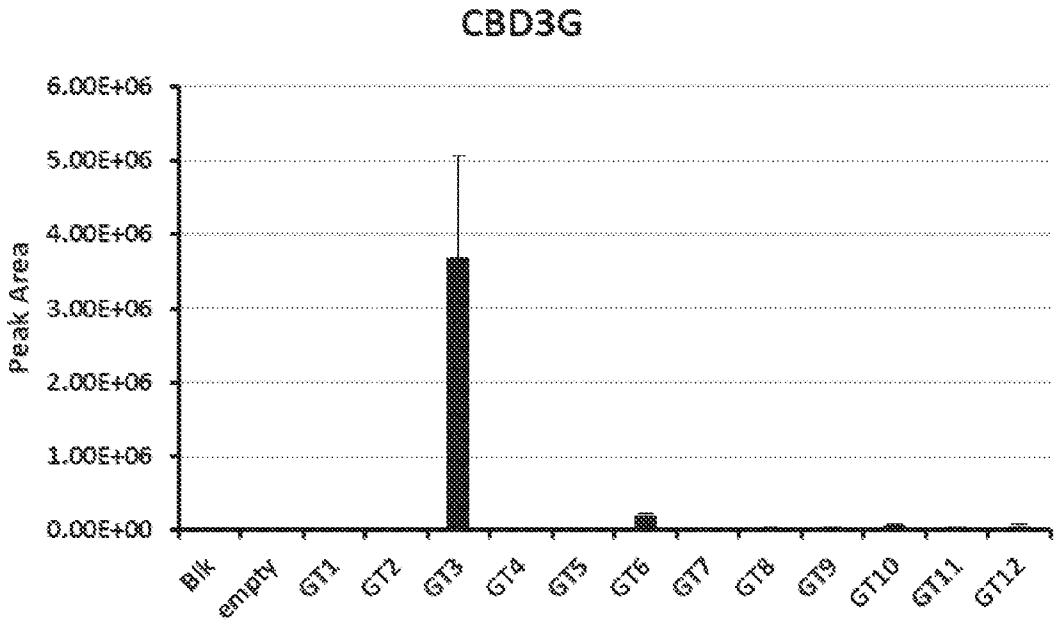

CBD3G
Peak Area
6.00E+06
5.00E+06
4.00E+06
3.00E+06
2.00E+06
1.00E+06
0.00E+00
Blk
empty
GT1
GT2
GT3
GT4
GT5
GT6
GT7
GT8
GT9
GT10
GT11
GT12

FIG. 5

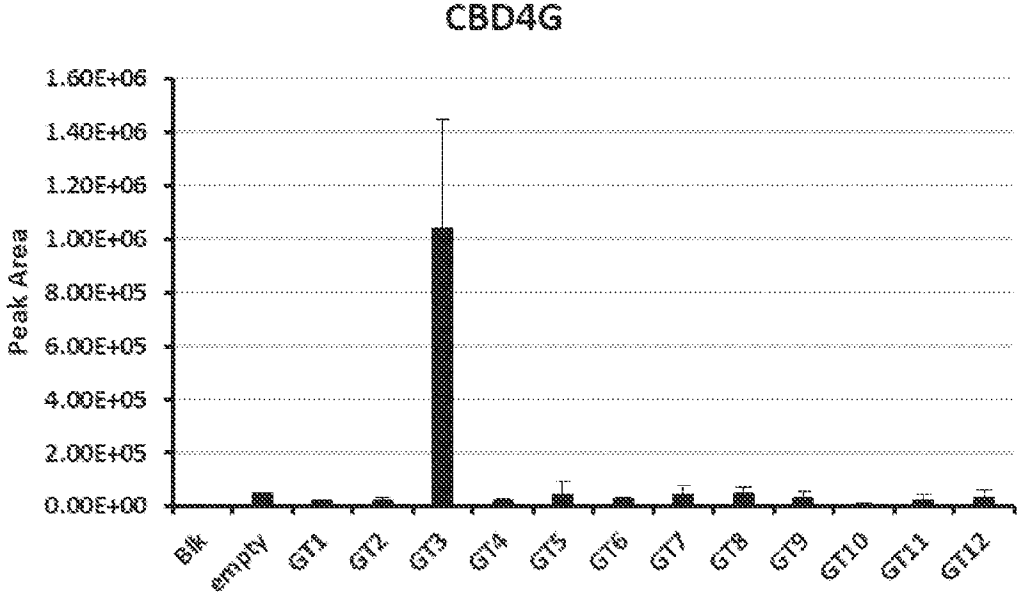

CBD4G
Peak Area
1.60E+06
1.40E+06
1.20E+06
1.00E+06
8.00E+05
6.00E+05
4.00E+05
2.00E+05
0.00E+00
Blk
empty
GT1
GT2
GT3
GT4
GT5
GT6
GT7
GT8
GT9
GT10
GT11
GT12

FIG. 6

CBN GT4

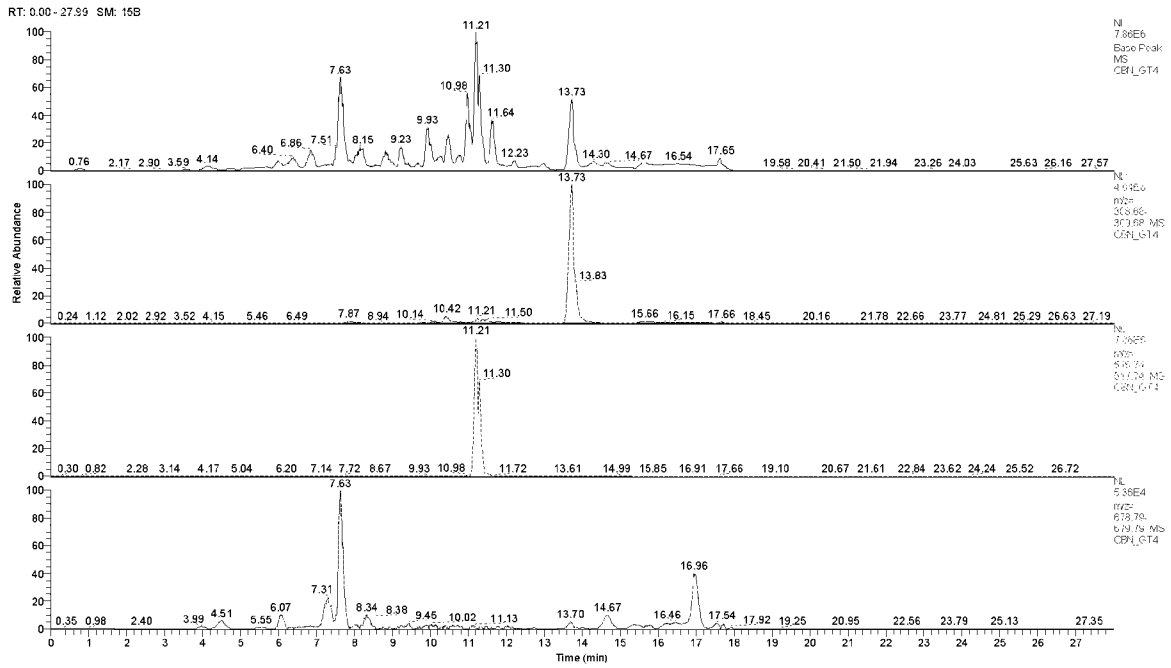


FIG. 7A

CBN GT5

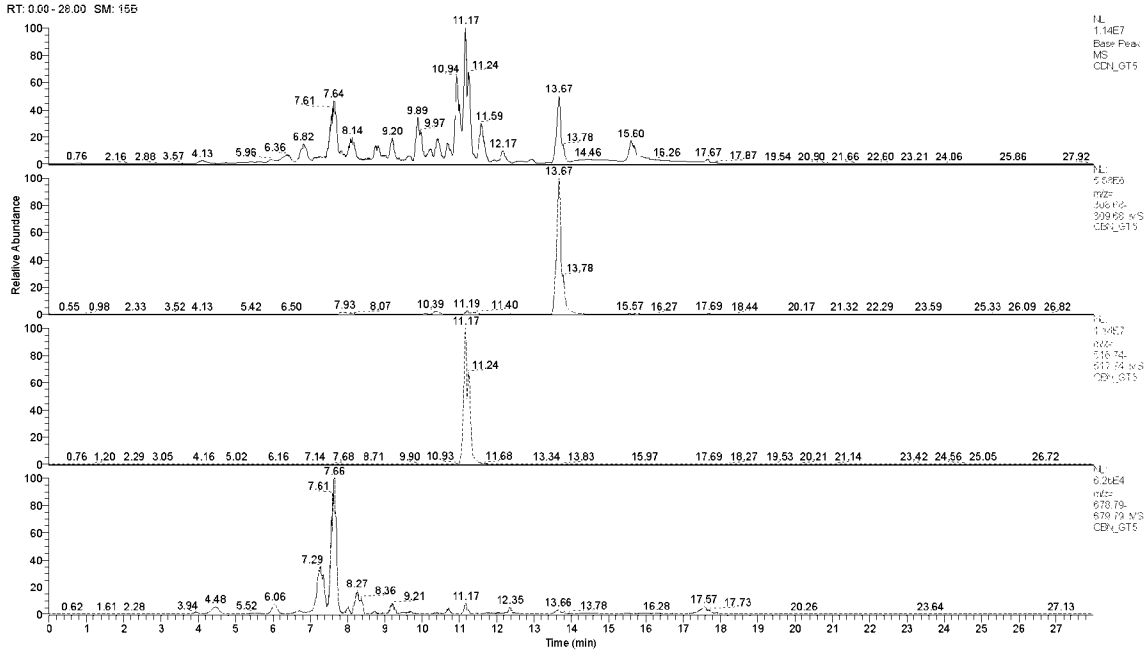


FIG. 7B

CBN GT6

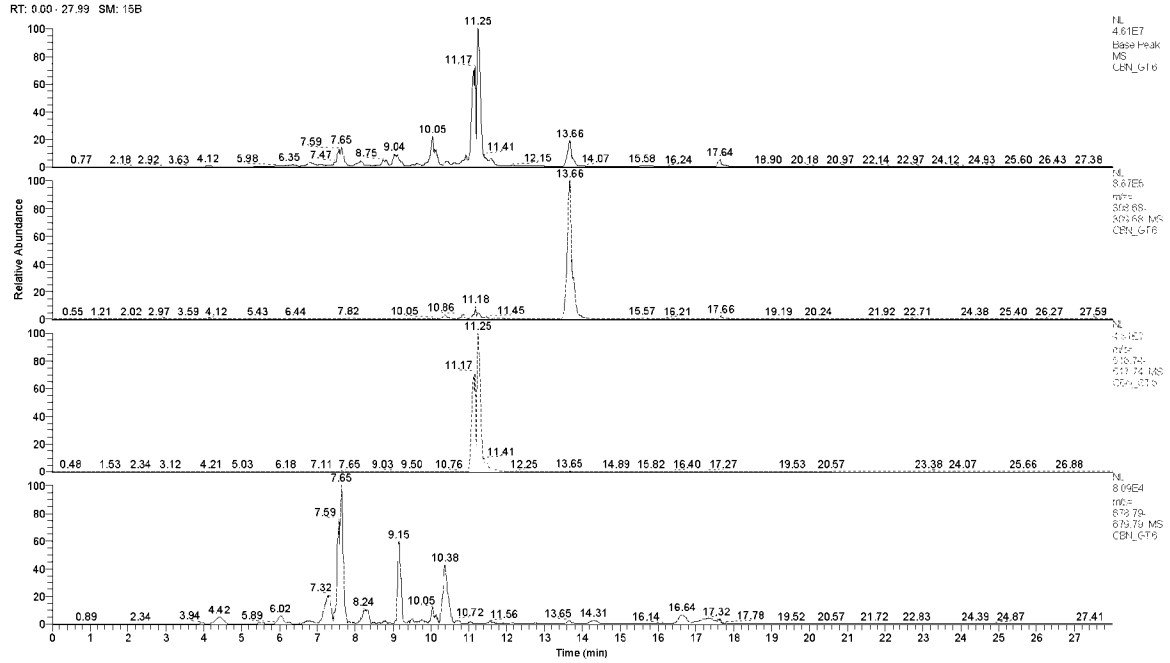


FIG. 7C

CBN GT8

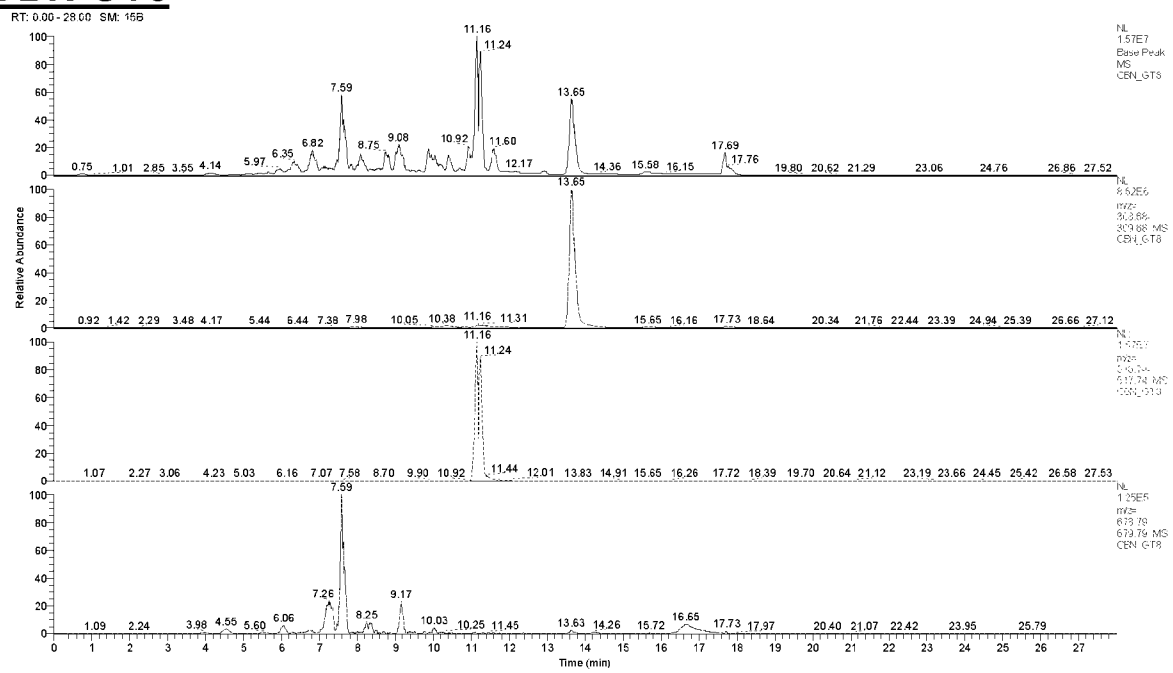


FIG. 7D

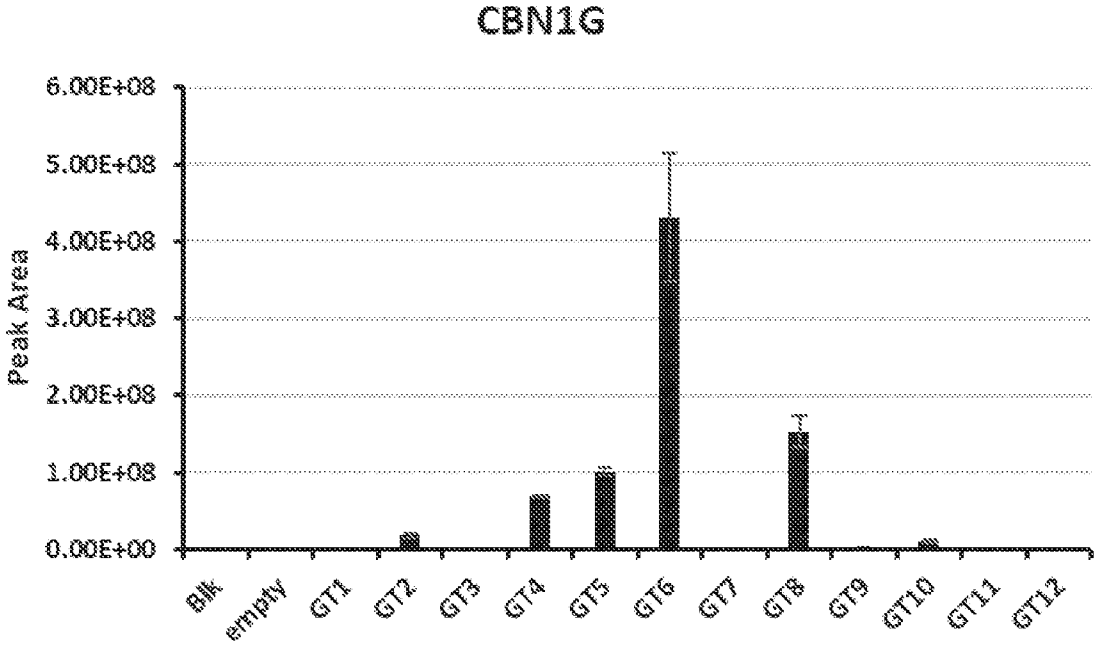

CBN1G
6.00E+08
5.00E+08
4.00E+08
3.00E+08
2.00E+08
1.00E+08
0.00E+00
Peak Area
Blk
empty
GT1
GT2
GT3
GT4
GT5
GT6
GT7
GT8
GT9
GT10
GT11
GT12

FIG. 7E

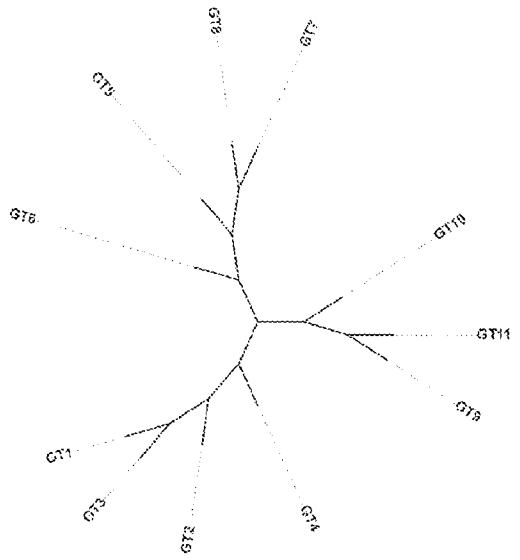

GT12
GT7
GT9
GT6
GT10
GT11
GT8
GT1
GT3
GT2
GT4

FIG. 8

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2024/033912

A. CLASSIFICATION OF SUBJECT MATTER

IPC: ***C12N 9/10*** (2024.01); ***C12N 5/04*** (2024.01); ***A01H 6/28*** (2024.01); *C12N 15/52* (2024.01)

CPC: ***C12N 9/1051; A01H 6/28; C12N 5/04;*** *C12N 15/52; C12Y 204/01*

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

See Search History Document

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

See Search History Document

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

See Search History Document

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 2021/174092 A1 (TRAIT BIOSCIENCES INC.) 02 September 2021 (02.09.2021) entire document	1-16
Y	WO 2018/204483 A2 (SENOMYX INC.) 08 November 2018 (08.11.2018) entire document	1-16
A	WO 2002/010210 A2 (BAYER AKTIENGESELLSCHAFT) 07 February 2002 (07.02.2002) entire document	1-16
A	US 7,847,156 B2 (INZE et al.) 07 December 2010 (07.12.2010) entire document	1-16

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

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

"D" document cited by the applicant in the international application

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

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

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

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

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
04 October 2024 (04.10.2024)	**22 November 2024 (22.11.2024)**

Name and mailing address of the ISA/US	Authorized officer
COMMISSIONER FOR PATENTS **MAIL STOP PCT, ATTN: ISA/US** **P.O. Box 1450** **Alexandria, VA 22313-1450** **UNITED STATES OF AMERICA**	**TAINA MATOS**
Facsimile No. **571-273-8300**	Telephone No. **571-272-4300**

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2024/033912

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

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:
   a. ☐ forming part of the international application as filed.
   b. ☑ furnished subsequent to the international filing date for the purposes of international search (Rule 13*ter*.1(a)),
      ☑ accompanied by a statement to the effect that the sequence listing does not go beyond the disclosure in the international application as filed.

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

3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2024/033912

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

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

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees need to be paid.

Group I: claims 1-16 are drawn to SEQ ID NO. 2, methods of producing a cannabinoid glycoside comprising the same, and systems for producing a cannabinoid glycoside comprising the same.

Group II: claims 17-22 are drawn to isolated nucleotide sequence encoding a UDP-glucosyltransferase (UGT) according to SEQ ID NO. 1, expression vectors comprising the same, cells transformed by the same, cell cultures generated by the same, method of producing a cannabinoid glycoside comprising the same.

Group III: claims 23-39 are drawn to a UDP-glucosyltransferase (UGT) peptide selected from SEQ ID NO's. 4, 6, 8, and 10, or a fragment thereof, methods of producing a cannabinoid glycoside comprising the same, and systems for producing a cannabinoid glycoside comprising the same.

Group IV: claims 40-45 are drawn to an isolated nucleotide sequence encoding a UDP-glucosyltransferase (UGT) selected from SEQ ID NO. 3, 5, 7 and 9, expression vectors comprising the same, cells transformed by the same, cell cultures generated by the same, method of producing a cannabinoid glycoside comprising the same.

The inventions listed in Groups I-IV do not relate to a single general inventive concept under PCT Rule 13.1, because under PCT Rule 13.2 they lack the same or corresponding special technical features for the following reasons:

The special technical features of Group I, SEQ ID NO. 2, methods of producing a cannabinoid glycoside comprising the same, and systems for producing a cannabinoid glycoside comprising the same, are not present in Groups II-IV; the special technical features of Group II, isolated nucleotide sequence encoding a UDP-glucosyltransferase (UGT) according to SEQ ID NO. 1, expression vectors comprising the same, cells transformed by the same, cell cultures generated by the same, method of producing a cannabinoid glycoside comprising the same, are not present in Groups I, III, and IV; the special technical features of Group III, a UDP-glucosyltransferase (UGT) peptide selected from SEQ ID NO's. 4, 6, 8, and 10, or a fragment thereof, methods of producing a cannabinoid glycoside comprising the same, and systems for producing a cannabinoid glycoside comprising the same, are not present in Groups I, II, and IV; and the special technical features of Group IV, an isolated nucleotide sequence encoding a UDP-glucosyltransferase (UGT) selected from SEQ ID NO. 3, 5, 7 and 9, expression vectors comprising the same, cells transformed by the same, cell cultures generated by the same, method of producing a cannabinoid glycoside comprising the same, are not present in Groups I-III.

Additionally, even if Groups I-IV were considered to share the technical features of a method of producing a cannabinoid glycoside comprising the step of: - contacting a cannabinoid having at least one glycosylation site to a UDP-glucosyltransferase (UGT) peptide; and - wherein said UGT catalyzes the bioconversion of the cannabinoid and one or more UDP-glucose substrates into a cannabinoid glycoside; a system for producing a cannabinoid glycoside comprising: - an in vivo, ex vivo or in vitro system comprising: - a quantity of a cannabinoid having at least one glycosylation site; - a UDP-glucosyltransferase (UGT) peptide; and - wherein said UGT catalyzes the bioconversion of the cannabinoid and one or more UDP-glucose substrates into a cannabinoid glycoside in the in vivo, ex vivo or in vitro system; An isolated nucleotide sequence encoding a UDP-glucosyltransferase (UGT); An expression vector encoding the nucleotide sequence, wherein said nucleotide sequence is operably linked to a promoter; a cell transformed by the expression vector; a cell culture generated from the transformed cell, these shared technical features do not represent a contribution over the prior art.

Specifically, WO 2021/174092 A1 to Trait Biosciences Inc. teaches of a method of producing a cannabinoid glycoside comprising the step of: - contacting a cannabinoid having at least one glycosylation site to a UDP-glucosyltransferase (UGT) peptide; and - wherein said UGT catalyzes the bioconversion of the cannabinoid and one or more UDP-glucose substrates into a cannabinoid glycoside (One aspect of the present invention further relates generally to novel methods of generating water-soluble cannabinoid glycoside compounds, and preferably THC-glycosides and CBD- glycosides, comprising the step of introducing one or more cannabinoids

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

to a UGT enzyme having specific activity towards one or more cannabinoid compounds according to the amino acid sequences identified as belonging to the Gram + of class of UGTs as described herein, in an in vitro, ex vivo, and in vivo systems, Pg. 5, Lns. 12-17; an exogenous or endogenous UDP-glucose/UDP-galactose transporter may be expressed in a trichome producing plant, such as Cannabis plant, Pg. 35, Lns. 18-19); a system for producing a cannabinoid glycoside comprising: - an in vivo, ex vivo or in vitro system comprising: - a quantity of a cannabinoid having at least one glycosylation site; - a UDP-glucosyltransferase (UGT) peptide; and - wherein said UGT catalyzes the bioconversion of the cannabinoid and one or more UDP-glucose substrates into a cannabinoid glycoside in the in vivo, ex vivo or in vitro system (One aspect of the present invention further relates generally to novel methods of generating water-soluble cannabinoid glycoside compounds, and preferably THC-glycosides and CBD- glycosides, comprising the step of introducing one or more cannabinoids to a UGT enzyme having specific activity towards one or more cannabinoid compounds according to the amino acid sequences identified as belonging to the Gram + of class of UGTs as described herein, in an in vitro, ex vivo, and in vivo systems, Pg. 5, Lns. 12-17; an exogenous or endogenous UDP-glucose/UDP- galactose transporter may be expressed in a trichome producing plant, such as Cannabis plant, Pg. 35, Lns. 18-19); An isolated nucleotide sequence encoding a UDP-glucosyltransferase (UGT) (An isolated nucleotide sequence operably linked to a promoter sequencer encoding at least one UDP-glucosyltransferases (UGT) having glycosylation activity towards at least one cannabinoid, Pg. 19, Lns. 15-17); An expression vector encoding the nucleotide sequence, wherein said nucleotide sequence is operably linked to a promoter (an expression vector having a nucleotide sequence encoding one or more UGTs that have glycosylation activity towards one or more cannabinoid compounds… operably linked to a promoter, Pg. 23, Lns. 24-28); a cell transformed by the expression vector (Another aspect of the current inventive technology includes one or more organisms, such as a plant, plant cell, bacteria, algae, fungi, or yeast cell, transformed by an expression vector, Pg. 23, Lns. 29-30); and a cell culture generated from the transformed cell (an in vivo system may include a whole organism system, such as a… cell culture, Pg. 23, Lns. 1-2).

The inventions listed in Groups I-IV therefore lack unity under Rule 13 because they do not share a same or corresponding special technical features.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☑ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.: **1-16**

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.