US 20230340429A1

(19) United States

(12) Patent Application Publication (10) Pub. No.: US 2023/0340429 A1

ONO et al. (43) Pub. Date: Oct. 26, 2023

(54) **STEVIOL GLYCOSIDE HEXOSE TRANSFERASE AND GENE CODING FOR SAME**

(71) Applicant: **SUNTORY HOLDINGS LIMITED**, Osaka (JP)

(72) Inventors: **Eiichiro ONO**, Kyoto (JP); **Misa OCHIAI**, Kyoto (JP); **Kazunari IWAKI**, Kyoto (JP); **Tadayoshi HIRAI**, Kyoto (JP)

(73) Assignee: **SUNTORY HOLDINGS LIMITED**, Osaka (JP)

(21) Appl. No.: **18/310,225**

(22) Filed: **May 1, 2023**

Related U.S. Application Data

(62) Division of application No. 16/473,819, filed on Jun. 26, 2019, now Pat. No. 11,702,641, filed as application No. PCT/JP2017/046804 on Dec. 26, 2017.

(30) **Foreign Application Priority Data**

Dec. 27, 2016 (JP) 2016-252643

Publication Classification

(51) **Int. Cl.**

C12N 9/10	(2006.01)
C12N 15/70	(2006.01)
C12N 15/81	(2006.01)
C12N 15/82	(2006.01)
C12P 19/56	(2006.01)

(52) **U.S. Cl.**

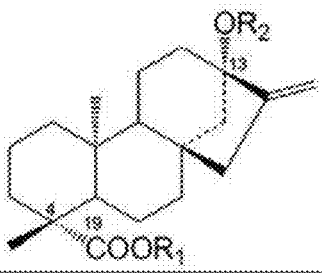
CPC ***C12N 9/1048*** (2013.01); ***C12N 15/70*** (2013.01); ***C12N 15/81*** (2013.01); ***C12N 15/8245*** (2013.01); ***C12P 19/56*** (2013.01)

(57) **ABSTRACT**

The purpose of the present invention is to provide a steviol glycoside hexose transferase, and a method for producing a steviol glycoside that contains glucose and/or rhamnose using said enzyme. The present invention provides a steviol glycoside hexose transferase, and a method for producing a steviol glycoside that contains glucose and/or rhamnose using said enzyme. The present invention also provides a transformant into which a steviol glycoside hexose transferase gene has been introduced, and a method for preparing said transformant.

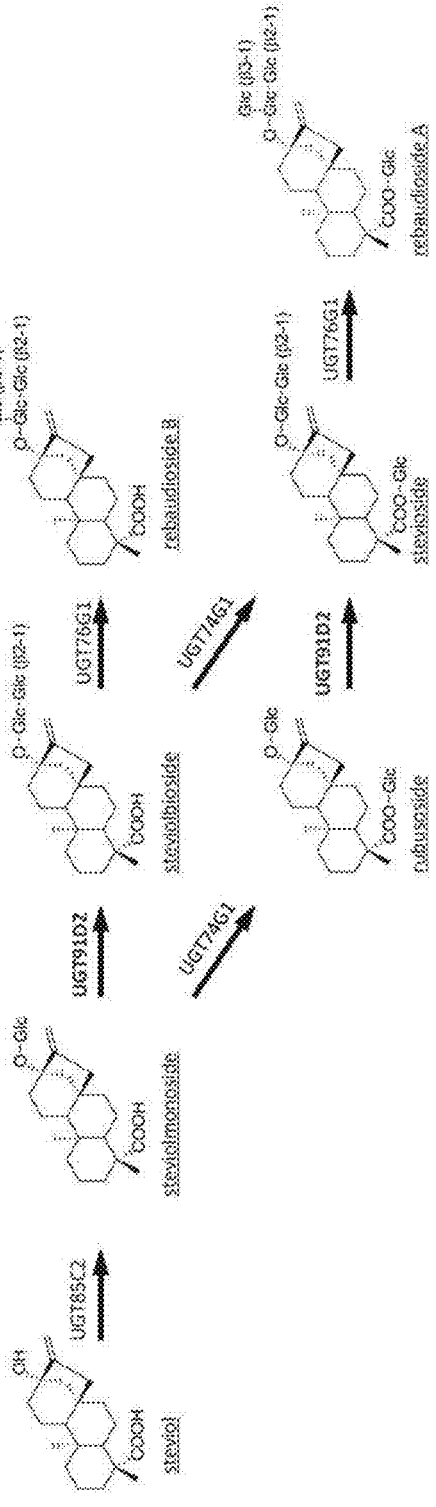
Specification includes a Sequence Listing.

[Figure 1]

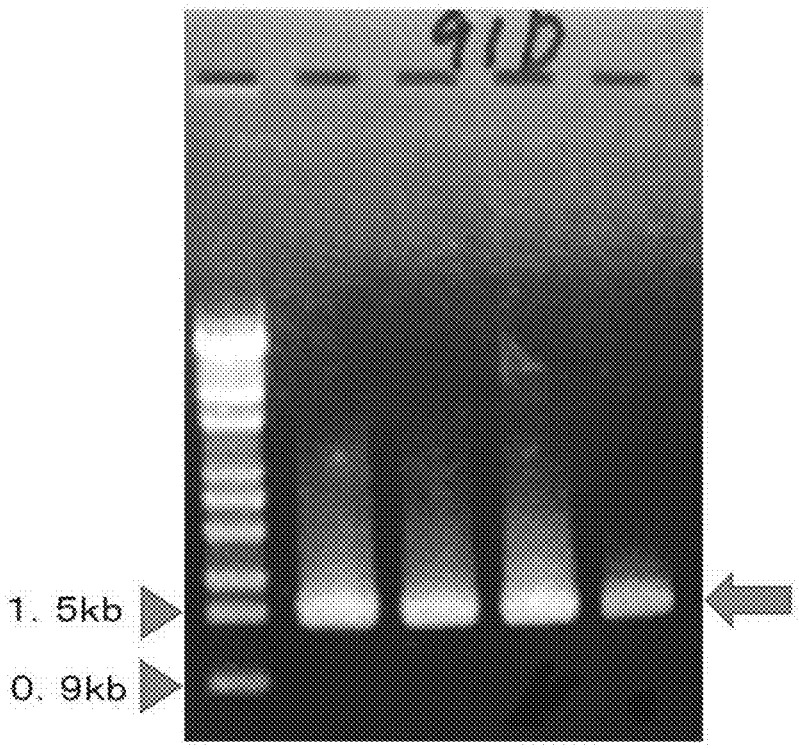


Name	R_1	R_2
Steviol	H	H
Steviolmonoside	H	Glc
Steviolbioside	H	Glc-Glc(β2→1)
Dulcoside A	Glc	Glc-Rha(β2→1)
Rubusoside	Glc	Glc
Stevioside	Glc	Glc-Glc(β2→1)
Rebaudioside A	Glc	Glc-Glc(β2→1) | Glc(β3→1)
Rebaudioside B	H	Glc-Glc(β2→1) | Glc(β3→1)
Rebaudioside C (Dulcoside B)	Glc	Glc-Rha(β2→1) | Glc(β3→1)
Rebaudioside D	Glc-Glc(β2→1)	Glc-Glc(β2→1) | Glc(β3→1)
Rebaudioside E	Glc-Glc(β2→1)	Glc-Glc(β2→1)
Rebaudioside F	Glc	Glc-Xyl(β2→1) | Glc(β3→1)

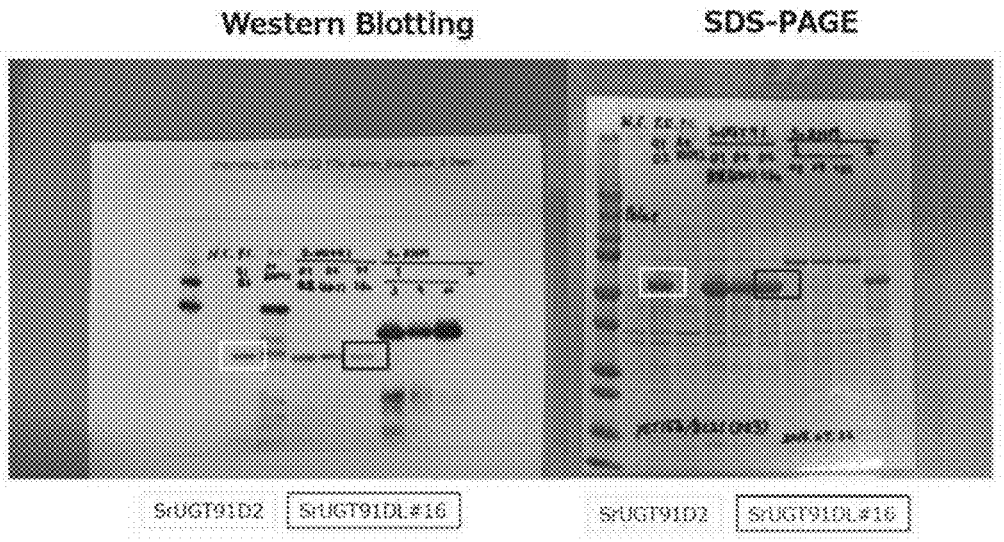
[Figure 2]


steviol
UGT85C2
steviolmonoside
UGT91D2
UGT74G1
steviobioside
UGT76G1
UGT74G1
rebaudioside B
UGT91D2
UGT74G1
rubusoside
UGT76G1
stevioside
UGT76G1
rebaudioside A

[Figure 3]



[Figure 4]


Western Blotting
SDS-PAGE
StUGT91D2
StUGT91DL#16
StUGT91D2
StUGT91DL#16

[Figure 5-1]

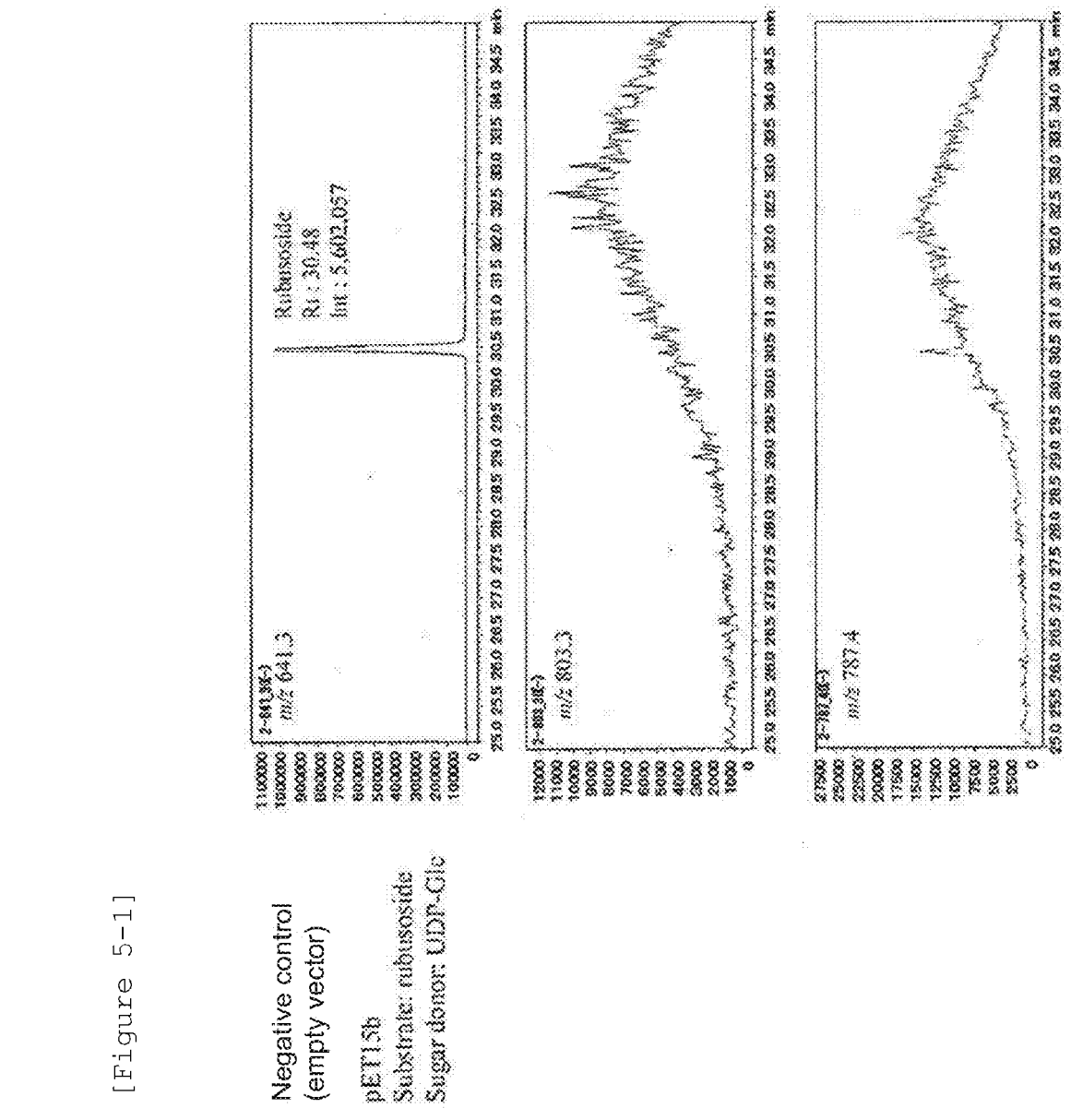
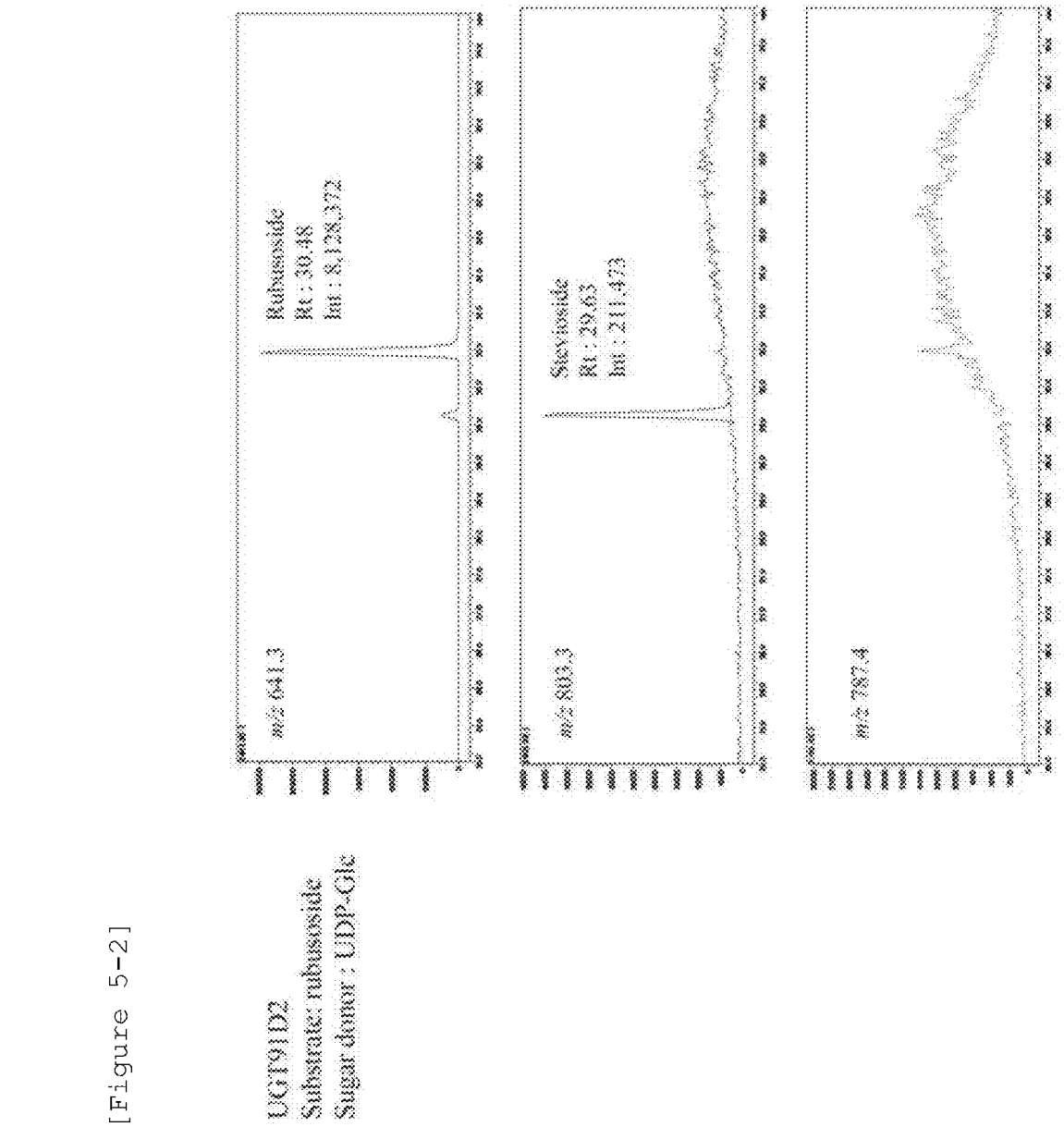
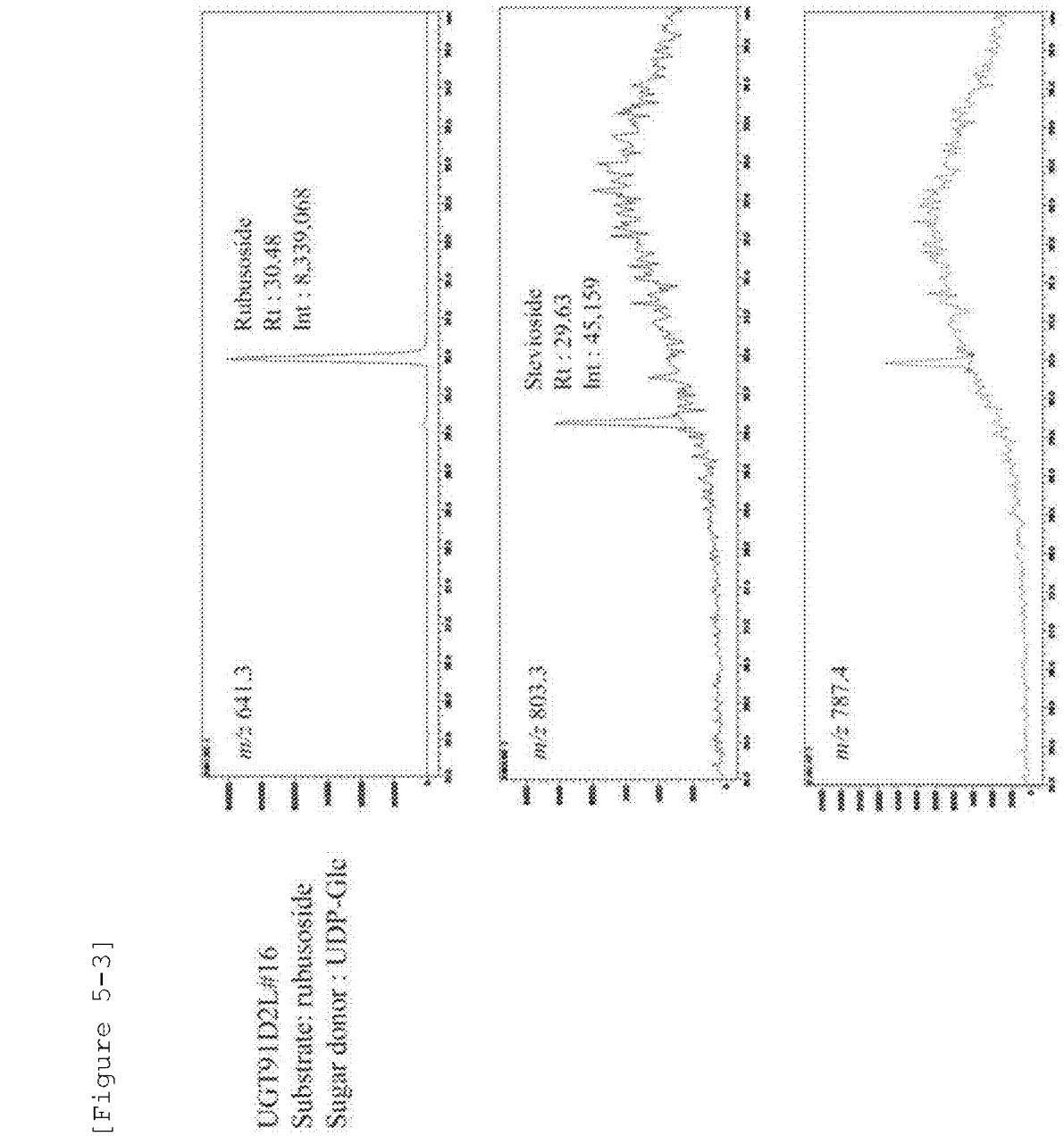


Figure 5-1: Enzymatic activity of recombinant protein

[Figure 5-2]


UGT91D2
Substrate: rubusoside
Sugar donor : UDP-Glc
m/z 641.3
Rubusoside
Rt : 30.48
Int : 8,128,372
m/z 803.3
Stevioside
Rt : 29.63
Int : 211,473
m/z 787.4

[Figure 5-3]


UGT91D2L#16
Substrate: rubusoside
Sugar donor : UDP-Glc
m/z 641.3
Rubusoside
Rt : 30.48
Int : 8,339,068
m/z 803.3
Stevioside
Rt : 29.63
Int : 45,159
m/z 787.4

[Figure 6-1]

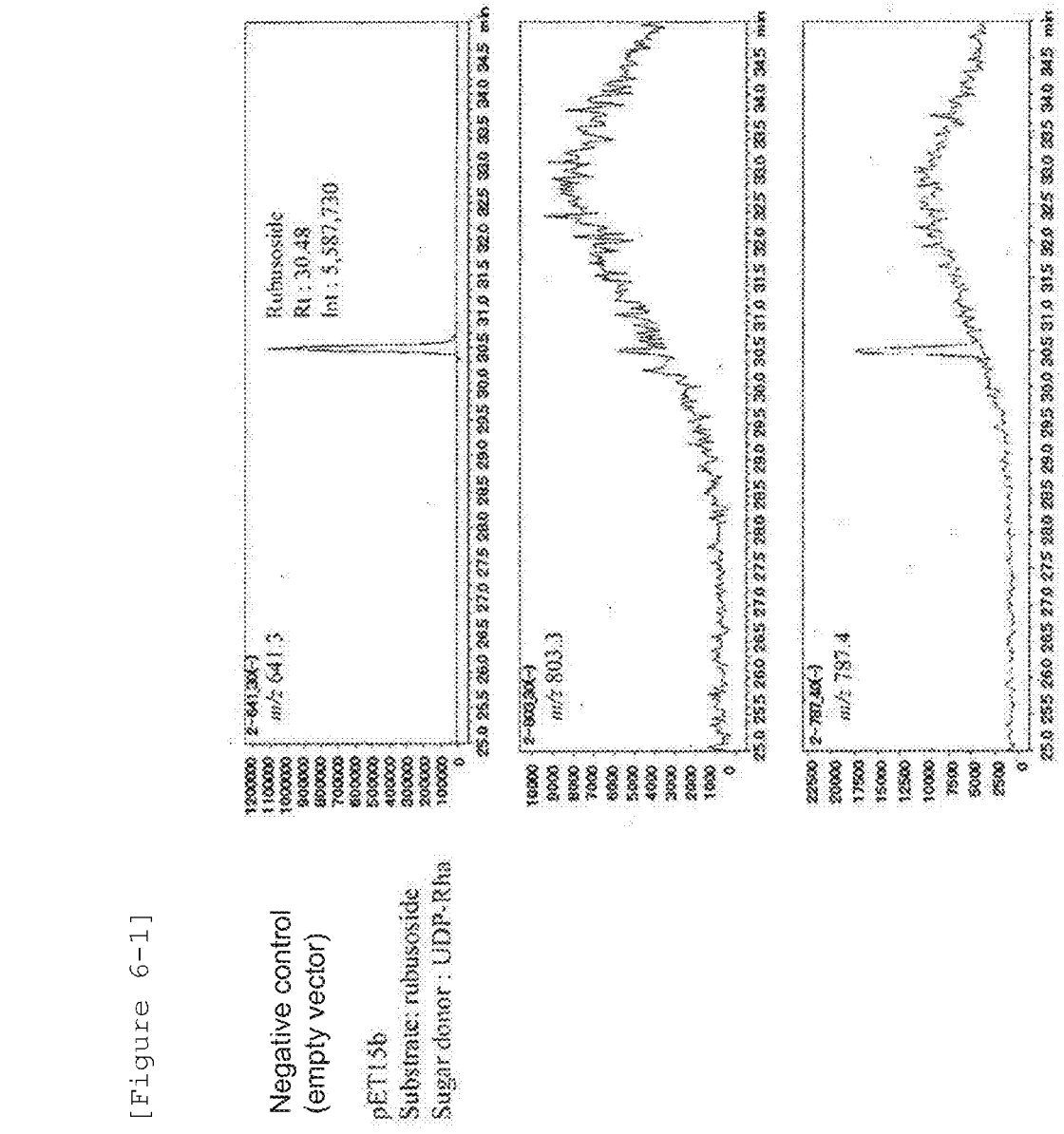
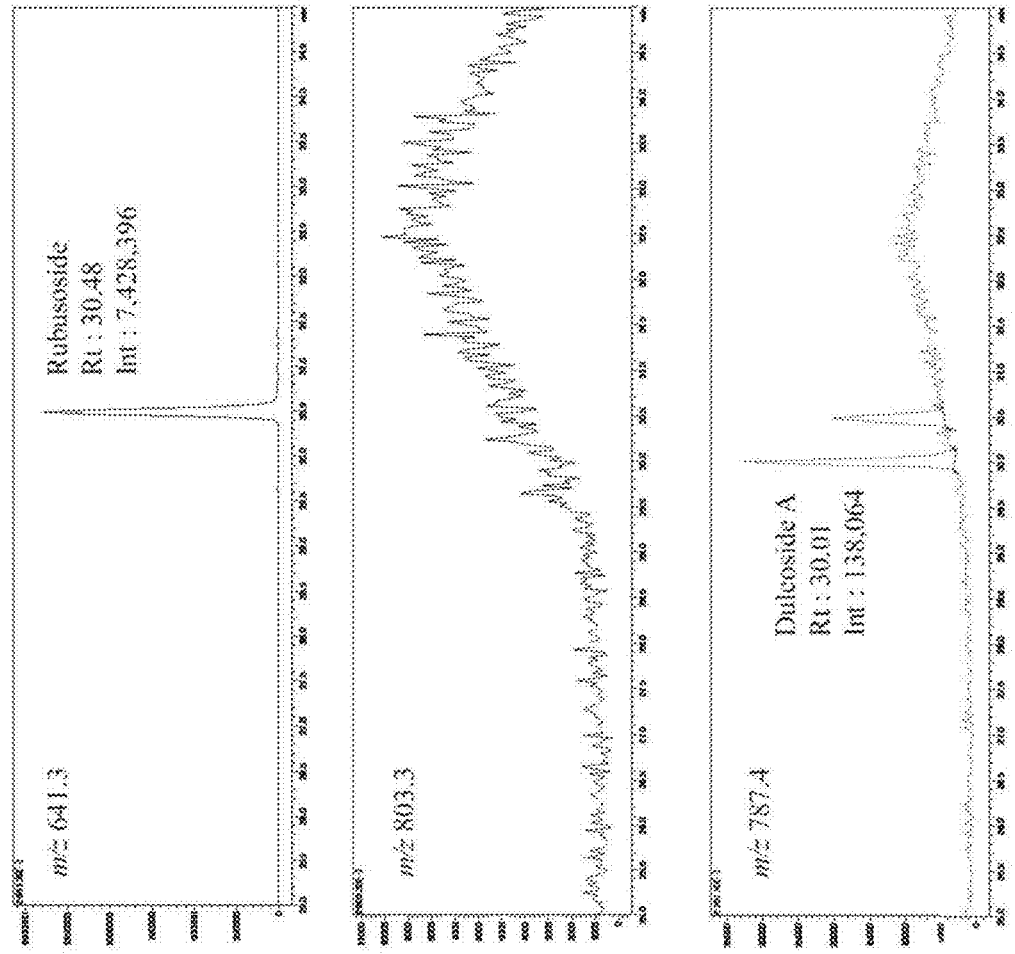
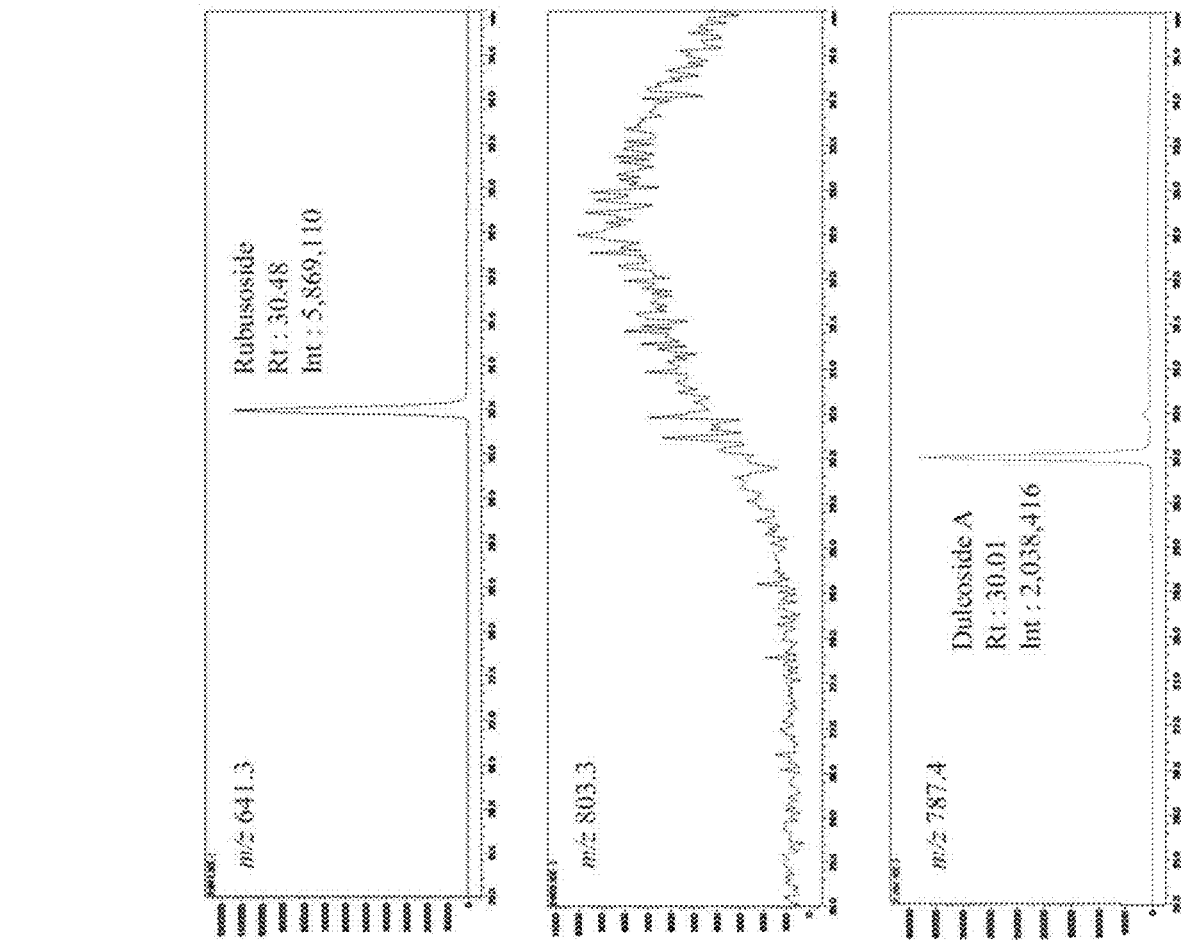


Figure 6-1: Enzymatic activity of recombinant protein

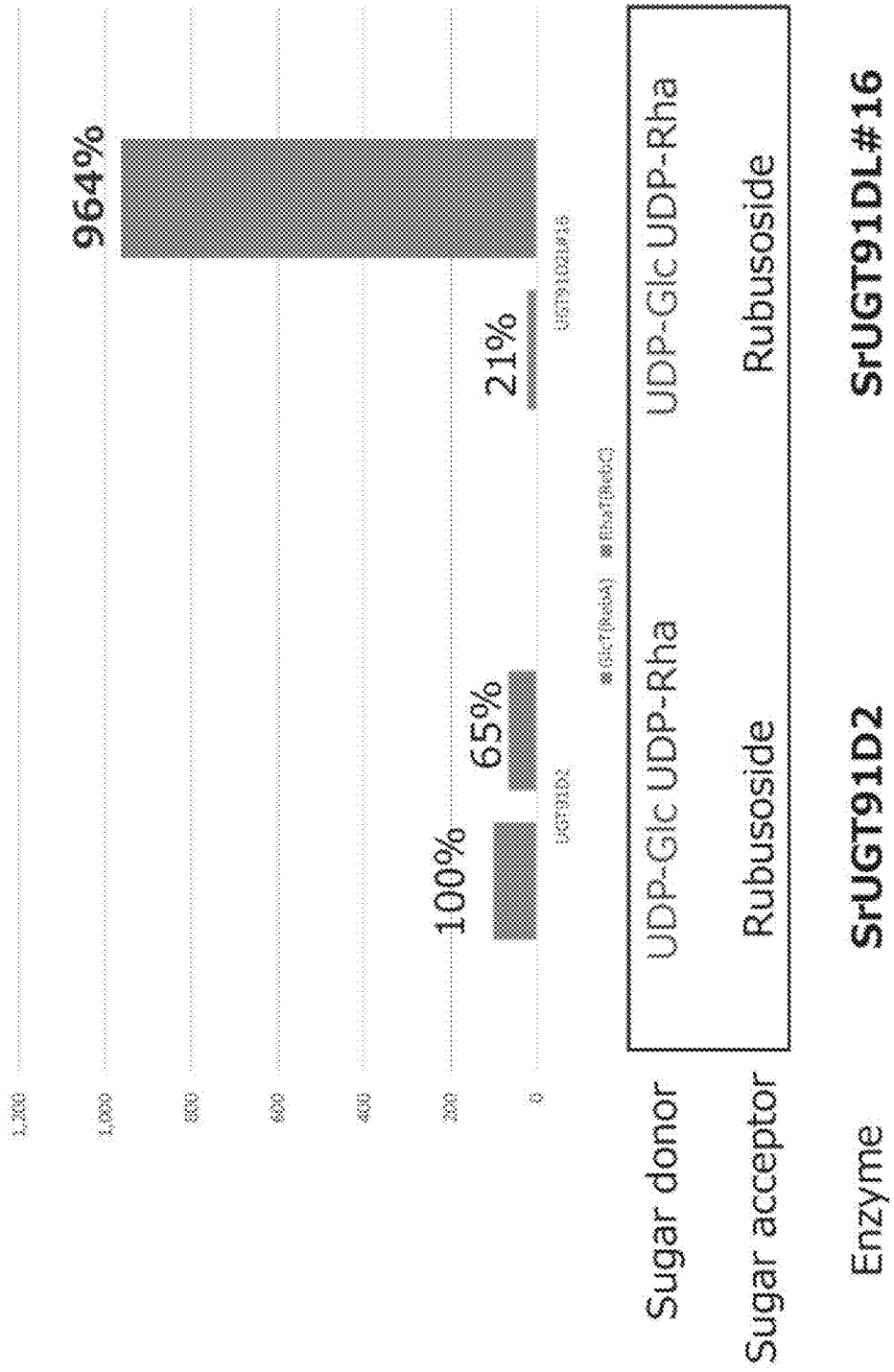
[Figure 6-2]


UGT91D2
Substrate: rubusoside
Suga donor : UDP-Rha
m/z 641.3
Rubusoside
Rt : 30.48
Int : 7,428,396
m/z 803.3
m/z 787.4
Dulcoside A
Rt : 30.01
Int : 138,064

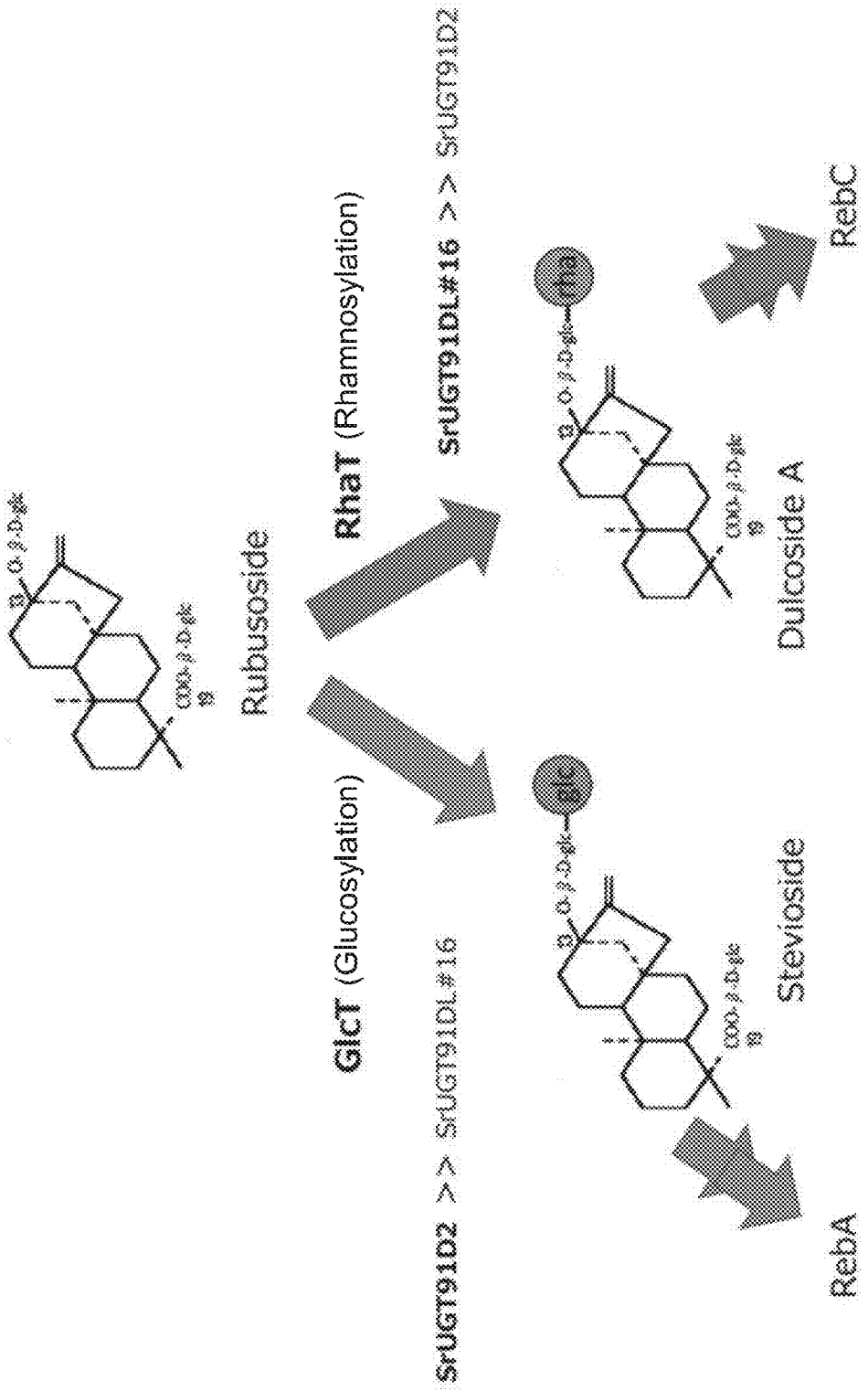
[Figure 6-3]


UGT91D2L#16
Substrate: rubusoside
Sugar donor : UDP-Rha
m/z 641.3
Rubusoside
Rt : 30.48
Int : 5,869,110
m/z 803.3
m/z 787.4
Dulcoside A
Rt : 30.01
Int : 2,038,416

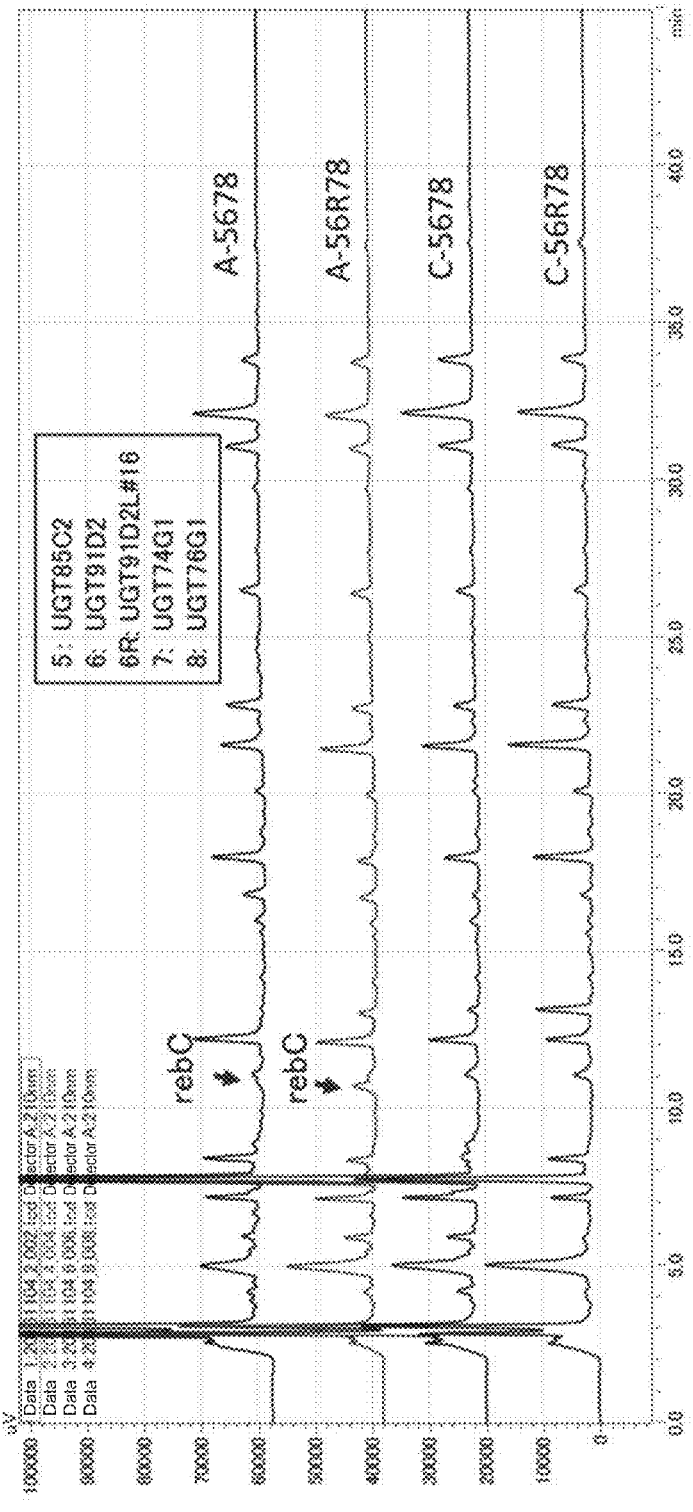
[Figure 7]


Relative activity of UGTs
1,200
1,000
800
600
400
200
0
100%
65%
964%
21%
UGT91D2
UGT91D2L#16
Sugar donor
UDP-Glc UDP-Rha
UDP-Glc UDP-Rha
Sugar acceptor
Rubusoside
Rubusoside
Enzyme
SrUGT91D2
SrUGT91DL#16

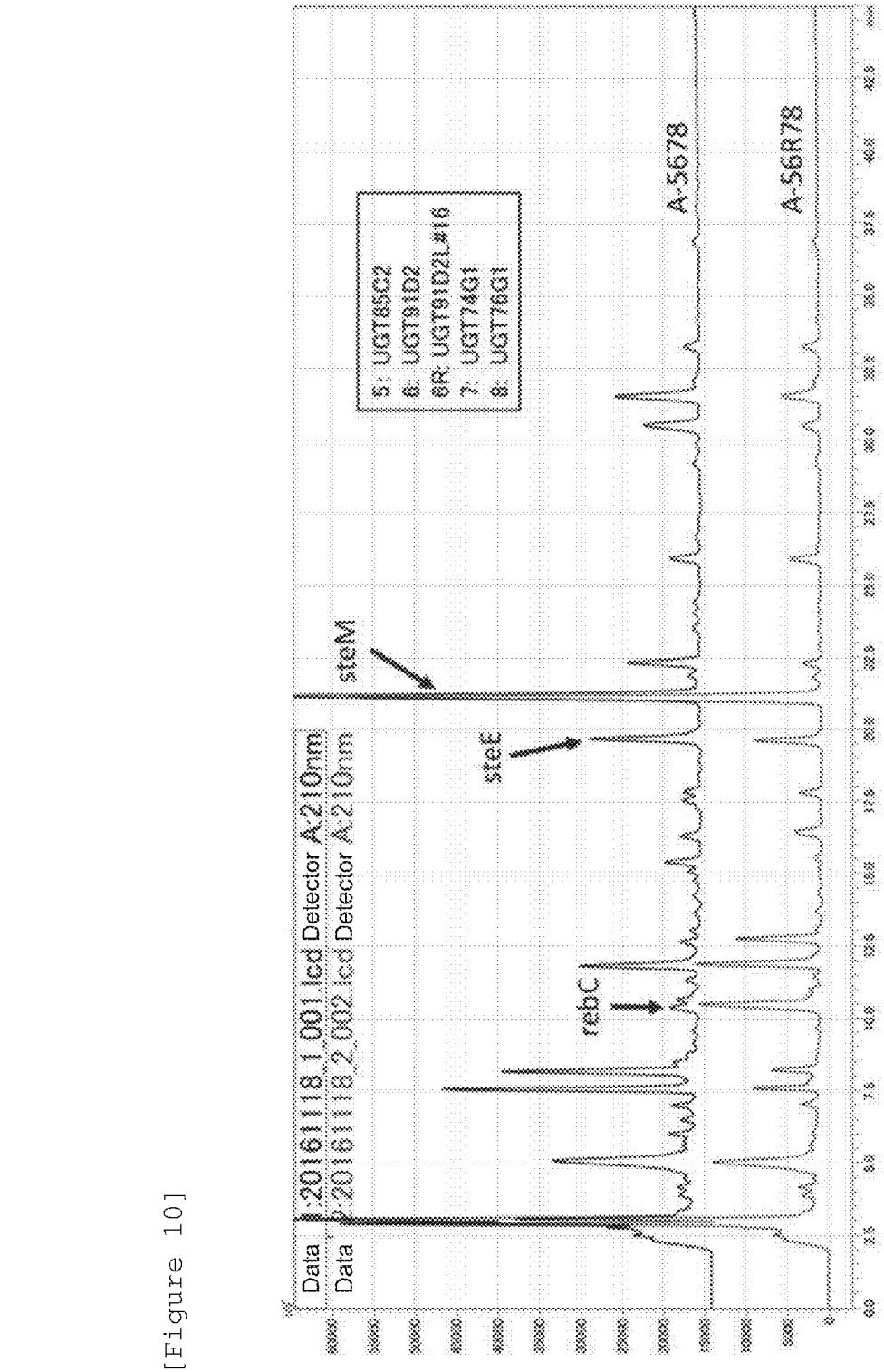
[Figure 8]


Rubusoside
RhaT (Rhamnosylation)
SrUGT91DL#16 >> SrUGT91D2
Dulcoside A
RebC
GlcT (Glucosylation)
SrUGT91D2 >> SrUGT91DL#16
Stevioside
RebA

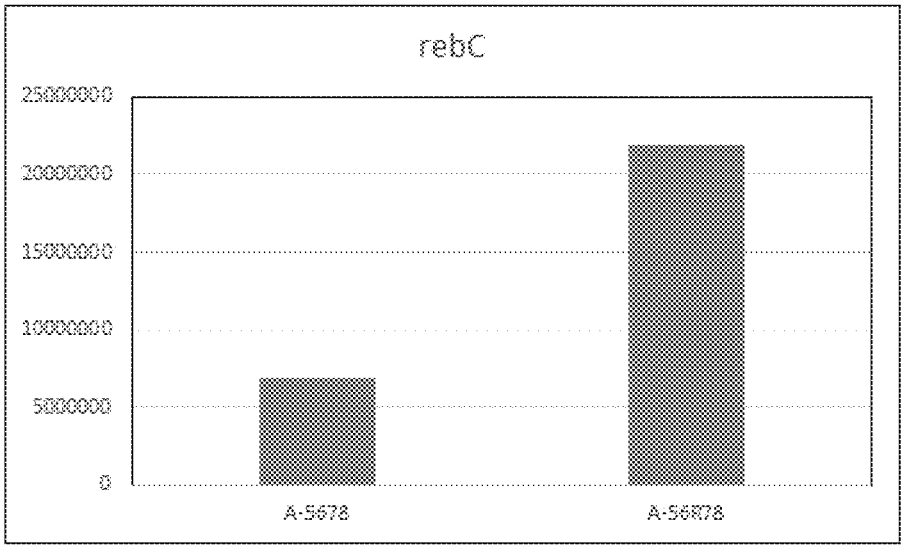
[Figure 9]


5: UGT85C2
6: UGT91D2
6R: UGT91D2L#16
7: UGT74G1
8: UGT76G1
rebC
rebC
A-5678
A-56R78
C-5678
C-56R78

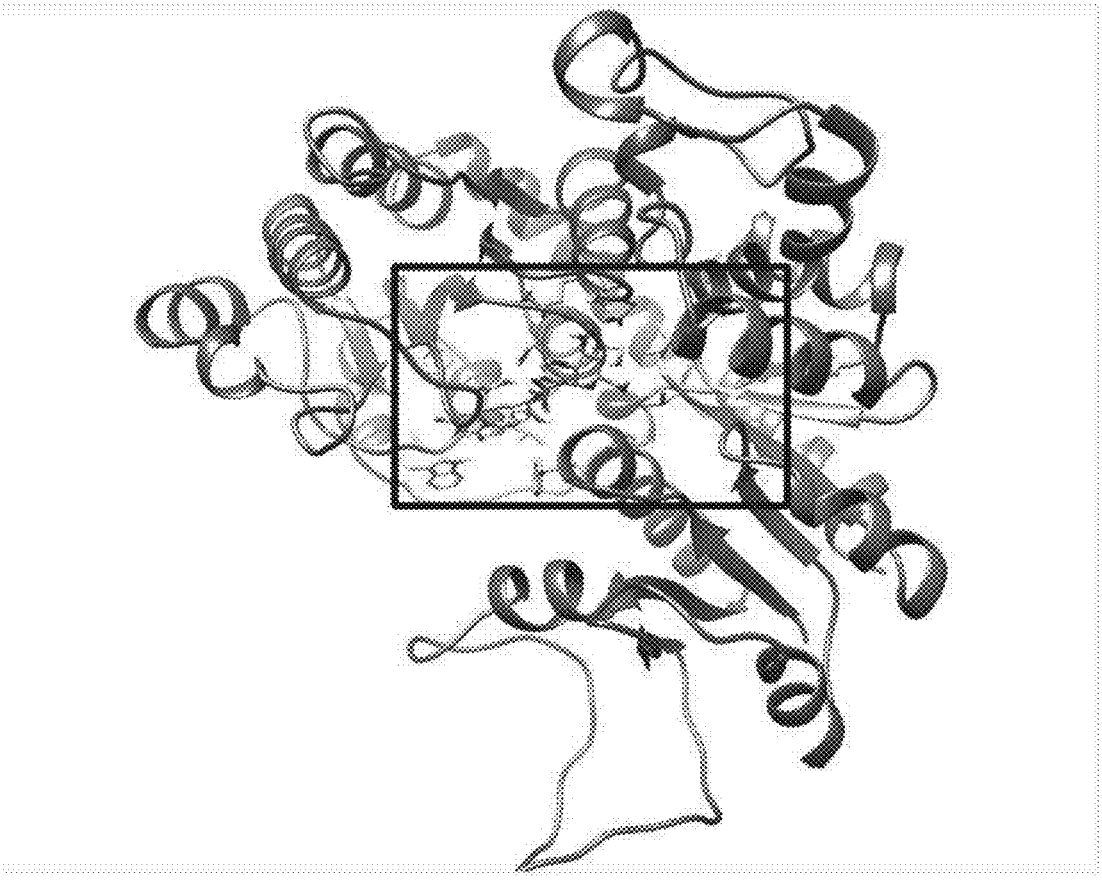
[Figure 10]


Data 1:20161118_1_001.lcd Detector A:210nm
Data 2:20161118_2_002.lcd Detector A:210nm
steM
steE
rebC
A-5678
A-56R78
5: UGT85C2
6: UGT91D2
6R: UGT91D2L#16
7: UGT74G1
8: UGT76G1

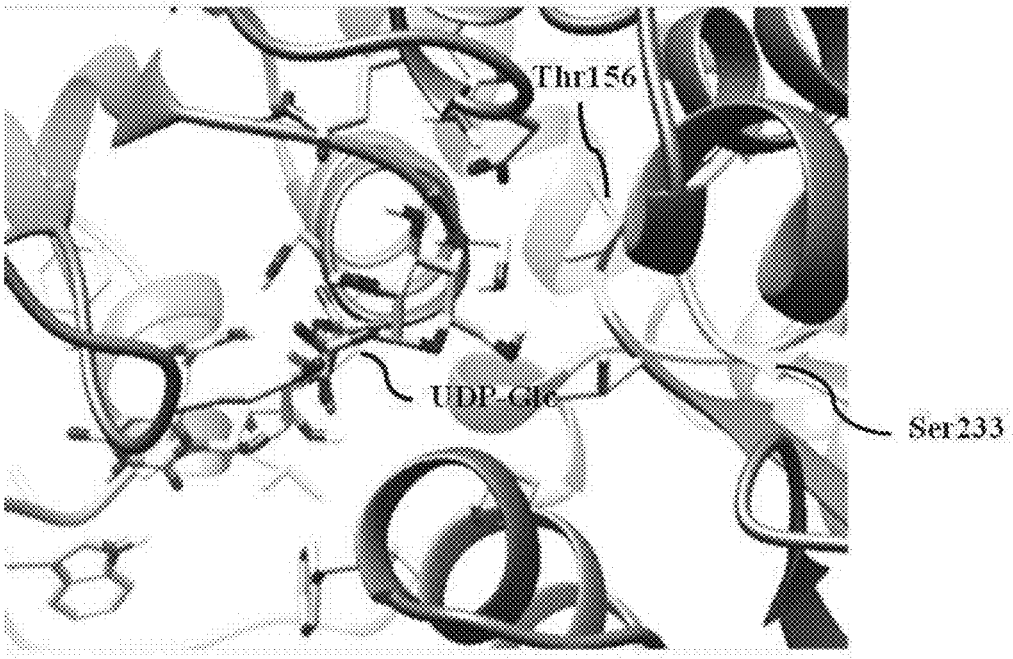
[Figure 11]


rebC
25000000
20000000
15000000
10000000
5000000
0
A-S678
A-S6R78

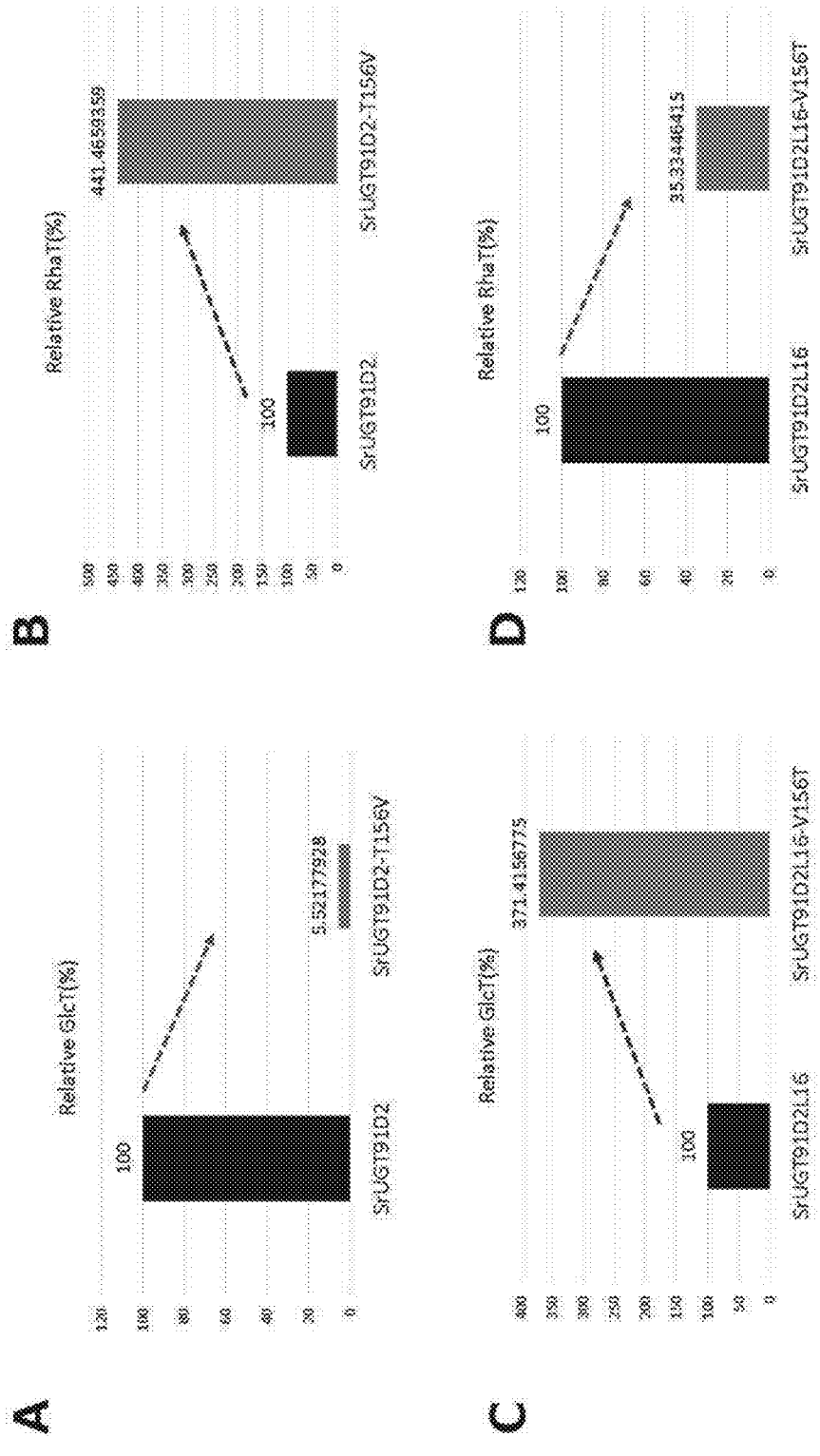
[Figure 12]



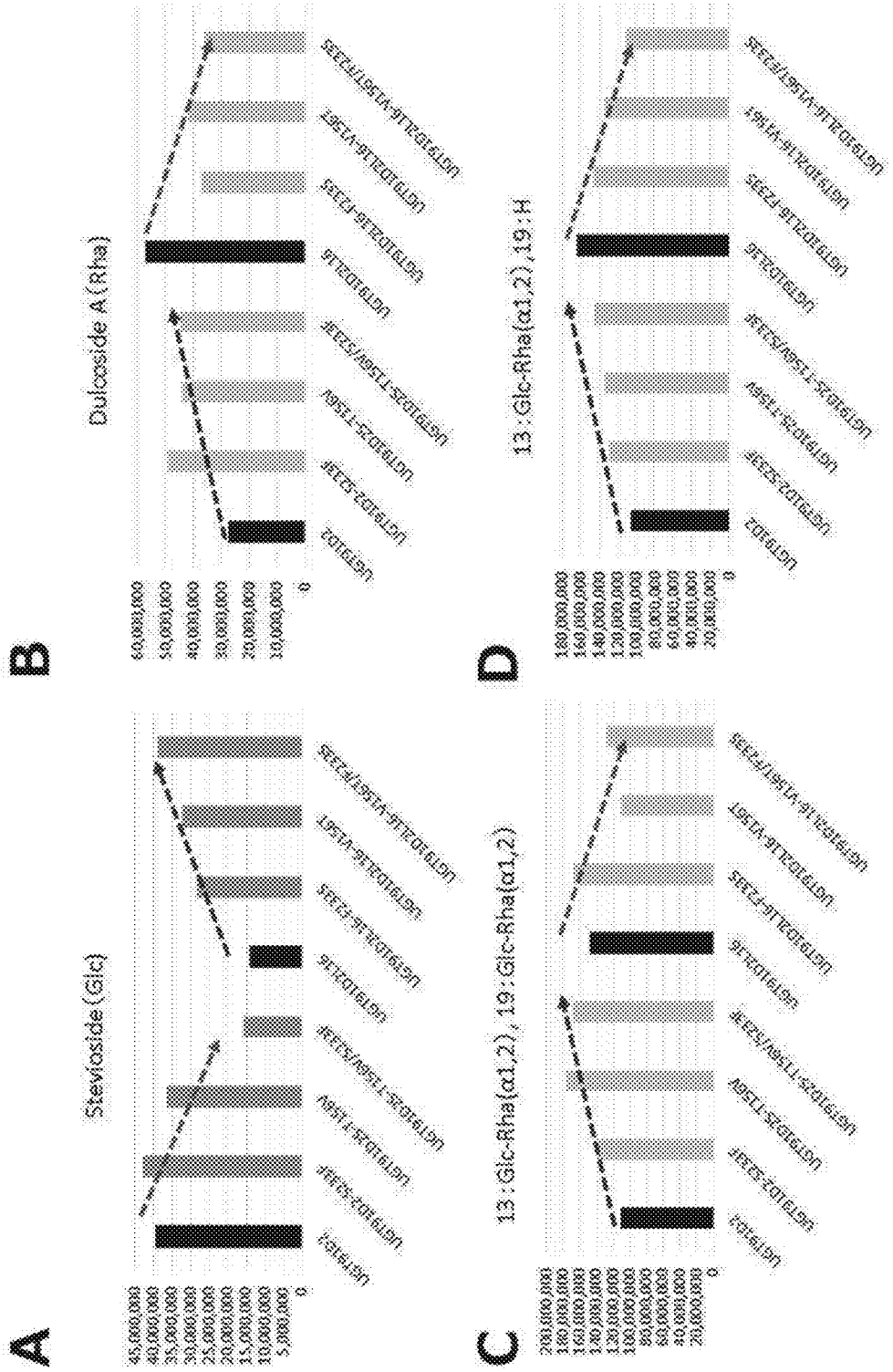
[Figure 13]



[Figure 14]


A
Relative GlcT(%)
100
5.52177928
SrUGT91D2
SrUGT91D2-T156V
B
Relative RhaT(%)
441.4659359
100
SrUGT91D2
SrUGT91D2-T156V
C
Relative GlcT(%)
371.4158775
100
SrUGT91D2116
SrUGT91D2116-V156T
D
Relative RhaT(%)
353.3446415
100
SrUGT91D2116
SrUGT91D2116-V156T

[Figure 15]


A
Stevioside (Glc)
B
Dulcoside A (Rha)
C
13:Glc-Rha(α1,2), 19:Glc-Rha(α1,2)
D
13:Glc-Rha(α1,2),19:H

STEVIOL GLYCOSIDE HEXOSE TRANSFERASE AND GENE CODING FOR SAME

PRIORITY

[0001] The present application is a divisional of U.S. application Ser. No. 16/473,819, which is the U.S. National Stage of PCT/JP2017/046804, filed Dec. 26, 2017, which application claims priority to JP App. No. 2016-252643, filed Dec. 27, 2016. The disclosures of each of these applications are incorporated by reference herein in their entireties.

SEQUENCE LISTING

[0002] The instant application contains a Sequence Listing XML which has been filed electronically in XML format and is hereby incorporated by reference in its entirety. Said XML copy, created on May 1, 2023, is named G1647US-DIV1.xml and is 65,400 bytes in size.

TECHNICAL FIELD

[0003] The present invention relates to a protein having the activity to transfer hexose to a steviol glycoside and a polynucleotide encoding this, a method of producing a steviol glycoside using the protein, a transformant that highly expresses a steviol glycosyltransferase, and a steviol glycoside made by the method and use thereof.

BACKGROUND ART

[0004] The leaves of *Stevia rebaudiana* of the family Asteraceae contain a secondary metabolite called "steviol" which is a kind of diterpenoid. Some steviol glycosides have sweetness about 300 times higher than that of sucrose and are therefore used as non-caloric sweeteners in the food industry. Obesity is becoming more of a serious social issue on an international scale, and non-caloric sweeteners are increasingly demanded from the viewpoint of promotion of health and reduction of medical cost. Currently, aspartame and acesulfame potassium, which are artificially-synthesized amino acid derivatives, are used as artificial sweeteners. However, naturally-occurring non-caloric sweeteners such as steviol glycosides are expected to be safer and gain more public acceptance.

[0005] The main steviol glycosides in *Stevia rebaudiana* are modified with sugar finally into a glycoside called Rebaudioside A (Reb.A) having 4 sugars attached (FIG. **1**). Its precursor stevioside, which is a trisaccharide glycoside, is most abundant and these 2 are central substances of the sweetness of *Stevia rebaudiana*. It is known that the stevioside content is highest in the leaves of *Stevia rebaudiana* and it exhibits sweetness about 250 to 300 times higher than that of sucrose. Reb.A is very sweet (350 to 450 times of that of sucrose) and is a tetrasaccharide glycoside, which is said to have good quality of taste. These have attracted attention as non-caloric sweeteners. In addition to these glycosides, glycosides considered to be reaction intermediates and analogs differing in the type of sugar are known to exist. For example, while the 4 glycoside sugars in Reb.A are all glucose, Rebaudioside C (Reb.C), in which rhamnose instead of glucose is added at position 2 of glucose at position 13, and Rebaudioside F (Reb.F), in which xylose was added at the same position, are known.

[0006] The genes of enzymes for bio-synthesis of Reb.A have been isolated by the Expressed Sequence Tag (EST) analysis of *Stevia rebaudiana* (Non-Patent Literature 1 and 2, Patent Literature 1). Ent-kaurenoic acid, which is a precursor of gibberellin, a plant hormone diterpenoid, is hydroxylated at position 13 by ent-kaurenoic acid 13-hydroxylase (EK13H) to produce steviol (FIG. **2**) (Patent Literature 1). Steviol is first glucosylated at the hydroxyl group at position 13 by UGT85C2, a UDP sugar-dependent glycosyltransferase (UGT) in *Stevia rebaudiana* to produce steviolmonoside (Non-Patent Literature 1, 2). Steviolmonoside is further glucosylated at position 2 of glucose at position 13 to produce steviolbioside or glucosylated at the carboxyl group at position 19 to produce a disaccharide glycoside of steviol called rubusoside. As an enzyme that glucosylates steviolmonoside or rubusoside at position 2 of glucose at position 13, UGT91D2 has been reported (Non-Patent Literature 2: in the report, UGT91D2 is referred to as UGT91D-like 3). Meanwhile, position 3 of glucose at position 13 and carboxylic acid at position 19 have been reported to be glucosylated by UGT76G1 and UGT74G1, respectively (Non-Patent Literature 2). As described above, the genes of enzymes responsible for glycosylation to Reb.A have been identified and the industrial use of *Stevia rebaudiana* enzymes is in progress, as seen in the report of ectopic expression of biosynthetic enzymes for steviol glycosides in yeast and production of the steviol glycosides in culture (Patent Literature 3). Meanwhile, the UGT enzymes responsible for the biosynthesis of Reb.C and Reb.F, which comprise a glycoside sugar other than glucose, are not elucidated.

CITATION LIST

Patent Literature

[0007] Patent Literature 1: EP 1 897 951 B1
[0008] Patent Literature 2: WO2013/137487
[0009] Patent Literature 3: WO2014 122328

Non-Patent Literature

[0010] Non-Patent Literature 1: Brandle and Telmer (2007) Phytochemistry 68, 1855-1863
[0011] Non-Patent Literature 2: Richman et al (2005) Plant J. 41, 56-67

SUMMARY OF INVENTION

Technical Problem

[0012] An object of the present invention is to provide a steviol glycoside hexose transferase and a method of producing a steviol glycoside comprising glucose and/or rhamnose using the enzyme.

Solution to Problem

[0013] The present inventors have succeeded, as a result of diligent studies for achieving the aforementioned object, in identifying the enzyme UGT91D2L #16, which catalyzes the reaction of adding rhamnose to rubusoside, a steviol glycoside, at position 2 of glucose at position 13 in *Stevia rebaudiana*, and a genetic sequence encoding the protein. It was also found that the enzyme UGT91D2L #16 catalyzes the transfer of glucose.

[0014] Moreover, the present inventors have identified the amino acid residues that are actually involved in the selection of the sugar donor among the 7 amino acid mutations between the glucose transferase (UGT91D2) and the rhamnose transferase (UGT91D2L #16) for steviol glycosides in *Stevia rebaudiana* through the homology model analysis and the evaluation of mutant proteins.

[0015] The present invention is based on the findings.

Advantageous Effects of Invention

[0016] By using the enzyme according to the present invention, it is possible to provide a steviol glycoside hexose transferase and a method of producing a steviol glycoside comprising glucose and/or rhamnose using the enzyme. Moreover, by promoting or inhibiting the function of the enzyme according to the present invention, the kind of steviol glycoside that is expressed in the plant of *Stevia rebaudiana* can be controlled. Furthermore, it is possible to select individuals containing a particular steviol glycoside and to produce steviol glycosides containing glucose and/or rhamnose by metabolic engineering, based on the sequence and the expression level of the gene of the enzyme.

[0017] Moreover, more highly glycosylated steviol glycosides (for example, Rebaudioside C, N, and O) can be produced by coexpressing the protein according to the present invention and a polynucleotide encoding the protein with another steviol glycoside hexose transferase or a polynucleotide encoding the enzyme in the same host cells.

[0018] Furthermore, by using the present invention, it is possible to efficiently design the sugar donor-selectivity of UGT and estimate the activity of a natural UGT and the glycoside sugar of an accumulated metabolite. Moreover, by specifically substituting one or more amino acid residues involved in the selection of sugar donor identified in the present invention by a genome editing technique, etc. it can be applied to controlling not only rhamnose-containing steviol glycosides such as Reb.C but also glycoside sugars of plant secondary metabolites.

BRIEF DESCRIPTION OF DRAWINGS

[0019] FIG. **1** illustrates the names and structures of the steviol glycosides. In FIG. **1**, "Glc-Glc" (β2→1) indicates that the binding of "Glc-Glc" is a β2,1 glycosidic linkage and "Glc-Glc" (β 3→1) indicates that the binding of "Glc-Glc" is a β3,1 glycosidic linkage.

[0020] FIG. **2** illustrates the biosynthetic pathway of steviol glycosides.

[0021] FIG. **3** illustrates a result of electrophoresis of PCR products in a 0.8% agarose gel and ethidium bromide staining.

[0022] FIG. **4** illustrates results of Western Blotting and SDS-PAGE of the *Stevia rebaudiana* UGT protein expressed in *Escherichia coli.*

[0023] FIG. **5-1** illustrates the activity to glucosylate rubusoside to generate stevioside in a negative control section.

[0024] FIG. **5-2** illustrates the activity to glucosylate rubusoside to generate stevioside in SrUGT91D2.

[0025] FIG. **5-3** illustrates the activity to glucosylate rubusoside to generate stevioside in SrUGT91D2L #16.

[0026] FIG. **6-1** illustrates the activity to rhamnosylate rubusoside to produce dulcoside A in a negative control section.

[0027] FIG. **6-2** illustrates the activity to rhamnosylate rubusoside to produce dulcoside A in SrUGT91D2.

[0028] FIG. **6-3** illustrates the activity to rhamnosylate rubusoside to produce dulcoside A in SrUGT91D2L #16.

[0029] FIG. **7** illustrates the specific activities of UGT91D2 and UGT91D2L #16 for glucosylation and rhamnosylation of rubusoside.

[0030] FIG. **8** illustrates the glycosylation pathway of steviol glycosides using UGT91D2 and UGT91D2L #16.

[0031] FIG. **9** illustrates results of the HPLC analysis of steviol glycosides obtained by the production of steviol glycosides in yeast (the amount of steviol added: 0.5 μg/ml).

[0032] FIG. **10** illustrates results of the HPLC analysis of steviol glycosides obtained by the production of steviol glycosides in yeast (the amount of steviol added: 2 μg/ml). In FIG. **10**, RebC represents Rebaudioside C, SteE represents steviol monoglucosyl ester, and SteM represents steviol monoglucoside.

[0033] FIG. **11** illustrates LC-MS results indicating the amount of Reb.C produced in the strains A-5678 and A-56R78 at an amount of steviol added of 2 μg/ml.

[0034] FIG. **12** illustrates a homology model of UGT91D2 bound with UDP-glucose.

[0035] FIG. **13** is an enlarged view of the part in the frame in FIG. **12**. The figure illustrates the positional relation between UDP-glucose (UDP-Glc) and the 156th amino acid (Thr156) and the 233rd amino acid (Ser233) adjacent thereto.

[0036] FIG. **14** is a graph illustrating a comparison of in vitro glycosylation activity (A) and rhamnosylation activity (B) of wild-type UGT91D2 and Mutant 1 expressed in *Escherichia coli* and a comparison of in vitro glucosylation activity (C) and rhamnosylation activity (D) of wild-type UGT91D2L #16 and Mutant 2. The numerical values indicate the relative activity of Mutant 1 or 2 with that of wild-type UGT91D2 or wild-type UGT91D2L #16 being 100%.

[0037] FIG. **15** is a graph illustrating a comparison of glucosylation activity (A) and rhamnosylation activity (B to D) of wild-type UGT91D2, wild-type UGT91D2L #16, and Mutants 1 to 6 in the yeast in which various UGTs are introduced. The numerical values indicate the signal intensity of the products as measured by LC-MS.

DESCRIPTION OF EMBODIMENTS

[0038] The present invention will be described in detail below. The following embodiments are illustrations for describing the present invention and it is not intended to limit the present invention to only these embodiments. The present invention can be embodied in various forms that do not deviates from the spirit of the present invention. All literature and patent literature such as unexamined patent publications and patent publications cited herein are incorporated herein by reference.

[0039] The present inventors have for the first time elucidated that the enzyme proteins are UGT91D2 and UGT91D2L #16 that are responsible for glucose- and/or rhamnose-addition reaction at position 2 of glucose at position 13 in steviol glycosides. The CDSs and the estimated amino acid sequences of UGT91D2 and UGT91D2L #16 are SEQ ID NOs: 1 to 4, respectively. These polynucleotides and enzymes can be obtained by the techniques described in Examples below, known genetic engineering techniques, known synthetic methods, and the like.

1. Steviol Glycoside Hexose Transferase

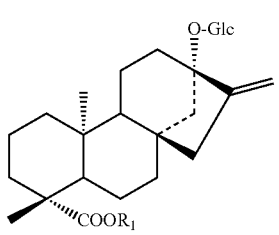
[0040] The present invention provides a protein (hereinafter, referred to as the "protein according to the present invention") according to any one selected from the group consisting of the following (a) to (c):

[0041] (a) a protein consisting of the amino acid sequence of SEQ ID NO: 2;

[0042] (b) a protein consisting of an amino acid sequence wherein 1 to 48 amino acids other than the amino acid residue X_7 are deleted, substituted, inserted, and/or added in the amino acid sequence of SEQ ID NO: 2 and having an activity to add hexose at position 2 of glucose at position 13 in a compound represented by the following formula (I);

[0043] (c) a protein having an amino acid sequence having a sequence identity of 90% or more to the amino acid sequence of SEQ ID NO: 2 wherein an amino acid corresponding to the amino acid residue X_7 is the same and having an activity to add hexose at position 2 of glucose at position 13 in a compound represented by the following formula (I);

[Formula 1]



(I)

[0044] wherein R_1 represents H, a C_1-C_{20} alkyl group, a C_2-C_{20} alkenyl group, a C_2-C_{20} alkynyl group, a C_4-C_{20} alkyldienyl group, a C_6-C_{18} aryl group, a C_6-C_{20} alkylaryl group, a C_6-C_{20} arylalkyl group, a C_4-C_{20} cycloalkyl group, a C_4-C_{20} cycloalkenyl group, a (C_3-C_{10} cycloalkyl)C_1-C_{10} alkyl group or a sugar residue.

[0045] According to another embodiment, the protein according to the present invention is

[0046] (b') a protein consisting of an amino acid sequence wherein 1 to 48 amino acids other than the amino acid residues X_1 to X_7 are deleted, substituted, inserted, and/or added in the amino acid sequence of SEQ ID NO: 2 and having an activity to add hexose at position 2 of glucose at position 13 in a compound represented by formula (I) above; or

[0047] (c') a protein having an amino acid sequence having a sequence identity of 90% or more to the amino acid sequence of SEQ ID NO: 2 wherein amino acids corresponding to the amino acid residues X_1 to X_7 are the same and having an activity to add hexose at position 2 of glucose at position 13 in a compound represented by formula (I) above.

[0048] According to another embodiment, the protein according to the present invention is (b") a protein consisting of an amino acid sequence wherein 1 to 48 amino acids other than the amino acid residues X_2 and/or X_3 are deleted, substituted, inserted, and/or added in the amino acid sequence of SEQ ID NO: 2 and having an activity to add hexose at position 2 of glucose at position 13 in a compound represented by formula (I) above; or

[0049] (c") a protein having an amino acid sequence having a sequence identity of 90% or more to the amino acid sequence of SEQ ID NO: 2 wherein amino acids corresponding to the amino acid residues X_2 and/or X_3 are the same and having an activity to add hexose at position 2 of glucose at position 13 in a compound represented by formula (I) above.

[0050] According to another embodiment, the protein according to the present invention is

[0051] (d) a protein consisting of an amino acid sequence wherein the 156th amino acid residue is Val and/or the 233rd amino acid residue is Phe in the amino acid sequence of SEQ ID NO: 4; or

[0052] (e) a protein consisting of an amino acid sequence wherein 1 to 48 amino acids are deleted, substituted, inserted, and/or added in the protein (d) above and an amino acid residue corresponding to the 156th amino acid residue of SEQ ID NO: 4 is Val and/or an amino acid residue corresponding to the 233rd amino acid residue of SEQ ID NO: 4 is Phe and having an activity to add hexose at position 2 of glucose at position 13 in a compound represented by formula (I) above; or

[0053] (f) a protein having an amino acid sequence having a sequence identity of 90% or more to the amino acid sequence of the protein (d) above wherein an amino acid residue corresponding to the 156th amino acid residue is Val and/or an amino acid residue corresponding to the 233rd amino acid residue is Phe in the amino acid sequence of SEQ ID NO: 4 and having an activity to add hexose at position 2 of glucose at position 13 in a compound represented by formula (I) above.

[0054] According to another embodiment, the protein according to the present invention is

[0055] (g) a protein wherein the amino acid residue X_2 is Thr and/or the amino acid residue X_3 is Ser in the amino acid sequence of SEQ ID NO: 2; or

[0056] (h) a protein consisting of an amino acid sequence wherein 1 to 48 amino acids are deleted, substituted, inserted, and/or added and an amino acid residue corresponding to the amino acid residue X_2 is Thr and/or an amino acid residue corresponding to the amino acid residue X_3 is Ser in the protein (g) above and having an activity to add hexose at position 2 of glucose at position 13 in a compound represented by formula (I) above; or

[0057] (i) a protein having an amino acid sequence having a sequence identity of 90% or more to the amino acid sequence of the protein (g) above wherein an amino acid residue corresponding to the amino acid residue X_2 is Thr and/or an amino acid residue corresponding to the amino acid residue X_3 is Ser and having an activity to add hexose at position 2 of glucose at position 13 in a compound represented by formula (I) above.

[0058] The proteins described in above (b), (b'), (b"), (c), (c'), (c"), and (d) to (i) are typically naturally occurring mutants of polypeptides of SEQ ID NO: 2 or 4, but include those that can be obtained artificially by using site-directed mutagenesis described in, for example, "Sambrook & Russell, Molecular Cloning: A Laboratory Manual Vol. 3, Cold

Spring Harbor Laboratory Press, 2001", "Ausubel, Current Protocols in Molecular Biology, John Wiley & Sons, 1987-1997", "Nuc. Acids. Res., 10, 6487 (1982)", "Proc. Natl. Acad. Sci. USA, 79, 6409 (1982)", "Gene, 34, 315 (1985)", "Nuc. Acids. Res., 13, 4431 (1985)", "Proc. Natl. Acad. Sci. USA, 82, 488 (1985)", or the like.

[0059] As used herein, the "protein consisting of an amino acid sequence wherein 1 to 48 amino acids other than the amino acid residue X_7 are deleted, substituted, inserted, and/or added in the amino acid sequence of SEQ ID NO: 2 and having an activity to add hexose at position 2 of glucose at position 13 in a compound represented by formula (I)" includes a protein consisting of an amino acid sequence wherein, for example, 1 to 48, 1 to 47, 1 to 46, 1 to 45, 1 to 44, 1 to 43, 1 to 42, 1 to 41, 1 to 40, 1 to 39, 1 to 38, 1 to 37, 1 to 36, 1 to 35, 1 to 34, 1 to 33, 1 to 32, 1 to 31, 1 to 30, 1 to 29, 1 to 28, 1 to 27, 1 to 26, 1 to 25, 1 to 24, 1 to 23, 1 to 22, 1 to 21, 1 to 20, 1 to 19, 1 to 18, 1 to 17, 1 to 16, 1 to 15, 1 to 14, 1 to 13, 1 to 12, 1 to 11, 1 to 10, 1 to 9 (1 to several), 1 to 8, 1 to 7, 1 to 6, 1 to 5, 1 to 4, 1 to 3, 1 to 2, or 1 amino acid residue(s) is/are deleted, substituted, inserted, and/or added, besides the amino acid residue X_7, in the amino acid sequence set forth in SEQ ID NO: 2 and having an activity to add hexose at position 2 of glucose at position 13 in a compound represented by formula (I). In general, the number of amino acid residues in the aforementioned deletion, substitution, insertion, and/or addition is preferably smaller.

[0060] Moreover, such proteins include proteins having an amino acid sequence having a sequence identity of 90% or more, 91% or more, 92% or more, 93% or more, 94% or more, 95% or more, 96% or more, 97% or more, 98% or more, 99% or more, 99.1% or more, 99.2% or more, 99.3% or more, 99.4% or more, 99.5% or more, 99.6% or more, 99.7% or more, 99.8% or more, or 99.9% or more to the amino acid sequence of SEQ ID NO: 2 wherein an amino acid corresponding to the amino acid residue X_7 is the same and having an activity to add hexose at position 2 of glucose at position 13 in a compound represented by formula (I). In general, the numerical value of the aforementioned sequence identity is preferably greater.

[0061] Similarly, the "protein consisting of an amino acid sequence wherein 1 to 48 amino acids other than the amino acid residues X_1 to X_7 are deleted, substituted, inserted, and/or added in the amino acid sequence of SEQ ID NO: 2 and having an activity to add hexose at position 2 of glucose at position 13 in a compound represented by formula (I)", as used herein, includes a protein consisting of an amino acid sequence wherein, for example, 1 to 48, 1 to 47, 1 to 46, 1 to 45, 1 to 44, 1 to 43, 1 to 42, 1 to 41, 1 to 40, 1 to 39, 1 to 38, 1 to 37, 1 to 36, 1 to 35, 1 to 34, 1 to 33, 1 to 32, 1 to 31, 1 to 30, 1 to 29, 1 to 28, 1 to 27, 1 to 26, 1 to 25, 1 to 24, 1 to 23, 1 to 22, 1 to 21, 1 to 20, 1 to 19, 1 to 18, 1 to 17, 1 to 16, 1 to 15, 1 to 14, 1 to 13, 1 to 12, 1 to 11, 1 to 10, 1 to 9 (1 to several), 1 to 8, 1 to 7, 1 to 6, 1 to 5, 1 to 4, 1 to 3, 1 to 2, or 1 amino acid residue(s) is/are deleted, substituted, inserted, and/or added, besides the amino acid residues X_1 to X_7, in the amino acid sequence set forth in SEQ ID NO: 2 and having an activity to add hexose at position 2 of glucose at position 13 in a compound represented by formula (I). In general, the number of amino acid residues in the aforementioned deletion, substitution, insertion, and/or addition is preferably smaller.

[0062] Moreover, such proteins include proteins having an amino acid sequence having a sequence identity of 90% or more, 91% or more, 92% or more, 93% or more, 94% or more, 95% or more, 96% or more, 97% or more, 98% or more, 99% or more, 99.1% or more, 99.2% or more, 99.3% or more, 99.4% or more, 99.5% or more, 99.6% or more, 99.7% or more, 99.8% or more, or 99.9% or more to the amino acid sequence of SEQ ID NO: 2 wherein amino acids corresponding to the amino acid residues X_1 to X_7 are the same and having an activity to add hexose at position 2 of glucose at position 13 in a compound represented by formula (I). In general, the numerical value of the aforementioned sequence identity is preferably greater.

[0063] As used herein, the "protein consisting of an amino acid sequence wherein 1 to 48 amino acids other than the amino acid residues X_2 and/or X_3 are deleted, substituted, inserted, and/or added in the amino acid sequence of SEQ ID NO: 2 and having an activity to add hexose at position 2 of glucose at position 13 in a compound represented by formula (I) above" includes a protein consisting of an amino acid sequence wherein, for example, 1 to 48, 1 to 47, 1 to 46, 1 to 45, 1 to 44, 1 to 43, 1 to 42, 1 to 41, 1 to 40, 1 to 39, 1 to 38, 1 to 37, 1 to 36, 1 to 35, 1 to 34, 1 to 33, 1 to 32, 1 to 31, 1 to 30, 1 to 29, 1 to 28, 1 to 27, 1 to 26, 1 to 25, 1 to 24, 1 to 23, 1 to 22, 1 to 21, 1 to 20, 1 to 19, 1 to 18, 1 to 17, 1 to 16, 1 to 15, 1 to 14, 1 to 13, 1 to 12, 1 to 11, 1 to 10, 1 to 9 (1 to several), 1 to 8, 1 to 7, 1 to 6, 1 to 5, 1 to 4, 1 to 3, 1 to 2, or 1 amino acid residue(s) is/are deleted, substituted, inserted, and/or added, besides the amino acid residues X_2 and/or X_3, in the amino acid sequence set forth in SEQ ID NO: 2 and having an activity to add hexose at position 2 of glucose at position 13 in a compound represented by formula (I). In general, the number of amino acid residues in the aforementioned deletion, substitution, insertion, and/or addition is preferably smaller.

[0064] Moreover, such proteins include proteins having an amino acid sequence having a sequence identity of 90% or more, 91% or more, 92% or more, 93% or more, 94% or more, 95% or more, 96% or more, 97% or more, 98% or more, 99% or more, 99.1% or more, 99.2% or more, 99.3% or more, 99.4% or more, 99.5% or more, 99.6% or more, 99.7% or more, 99.8% or more, or 99.9% or more to the amino acid sequence of SEQ ID NO: 2 wherein an amino acid corresponding to the amino acid residues X_2 and/or X_3 are the same and having an activity to add hexose at position 2 of glucose at position 13 in a compound represented by formula (I). In general, the numerical value of the aforementioned sequence identity is preferably greater.

[0065] As used herein, the "protein consisting of an amino acid sequence wherein 1 to 48 amino acids are deleted, substituted, inserted, and/or added in the protein (d) and an amino acid residue corresponding to the 156th amino acid residue is Val and/or an amino acid residue corresponding to the 233rd amino acid residue is Phe in the amino acid sequence of SEQ ID NO: 4 and having an activity to add hexose at position 2 of glucose at position 13 in a compound represented by formula (I) above" includes a protein consisting of an amino acid sequence wherein 1 to 48, 1 to 47, 1 to 46, 1 to 45, 1 to 44, 1 to 43, 1 to 42, 1 to 41, 1 to 40, 1 to 39, 1 to 38, 1 to 37, 1 to 36, 1 to 35, 1 to 34, 1 to 33, 1 to 32, 1 to 31, 1 to 30, 1 to 29, 1 to 28, 1 to 27, 1 to 26, 1 to 25, 1 to 24, 1 to 23, 1 to 22, 1 to 21, 1 to 20, 1 to 19, 1 to 18, 1 to 17, 1 to 16, 1 to 15, 1 to 14, 1 to 13, 1 to 12, 1 to 11, 1 to 10, 1 to 9 (1 to several), 1 to 8, 1 to 7, 1 to 6,

1 to 5, 1 to 4, 1 to 3, 1 to 2, or 1 amino acid residues is/are deleted, substituted, inserted, and/or added in the protein (d) and an amino acid residue corresponding to the 156th amino acid residue is Val and/or an amino acid residue corresponding to the 233rd amino acid residue is Phe in the amino acid sequence of SEQ ID NO: 4 and having an activity to add hexose at position 2 of glucose at position 13 in a compound represented by formula (I). In general, the number of amino acid residues in the aforementioned deletion, substitution, insertion, and/or addition is preferably smaller.

[0066] Moreover, such proteins include proteins having a sequence identity of 90% or more, 91% or more, 92% or more, 93% or more, 94% or more, 95% or more, 96% or more, 97% or more, 98% or more, 99% or more, 99.1% or more, 99.2% or more, 99.3% or more, 99.4% or more, 99.5% or more, 99.6% or more, 99.7% or more, 99.8% or more, or 99.9% or more to the amino acid sequence of the protein (d), having an amino acid sequence wherein an amino acid residue corresponding to the 156th amino acid residue is Val and/or an amino acid residue corresponding to the 233rd amino acid residue is Phe in the amino acid sequence of SEQ ID NO: 4 and having an activity to add hexose at position 2 of glucose at position 13 in a compound represented by formula (I). In general, the numerical value of the aforementioned sequence identity is preferably greater.

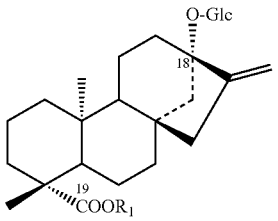
[0067] As used herein, the "protein consisting of an amino acid sequence wherein 1 to 48 amino acids are deleted, substituted, inserted, and/or added and an amino acid residue corresponding to the amino acid residue X_2 is Thr and/or the amino acid residue corresponding to the amino acid residue X_3 is Ser in the protein (g) and having an activity to add hexose at position 2 of glucose at position 13 in a compound represented by formula (I) above" includes a protein consisting of an amino acid sequence wherein, for example, 1 to 48, 1 to 47, 1 to 46, 1 to 45, 1 to 44, 1 to 43, 1 to 42, 1 to 41, 1 to 40, 1 to 39, 1 to 38, 1 to 37, 1 to 36, 1 to 35, 1 to 34, 1 to 33, 1 to 32, 1 to 31, 1 to 30, 1 to 29, 1 to 28, 1 to 27, 1 to 26, 1 to 25, 1 to 24, 1 to 23, 1 to 22, 1 to 21, 1 to 20, 1 to 19, 1 to 18, 1 to 17, 1 to 16, 1 to 15, 1 to 14, 1 to 13, 1 to 12, 1 to 11, 1 to 10, 1 to 9 (1 to several), 1 to 8, 1 to 7, 1 to 6, 1 to 5, 1 to 4, 1 to 3, 1 to 2, or 1 amino acid(s) other than the amino acid residues X_1 to X_7 is/are deleted, substituted, inserted, and/or added and an amino acid residue corresponding to the amino acid residue X_2 is Thr and/or an amino acid residue corresponding to the amino acid residue X_3 is Ser in the protein (g) and having an activity to add hexose at position 2 of glucose at position 13 in a compound represented by formula (I). In general, the number of amino acid residues in the aforementioned deletion, substitution, insertion, and/or addition is preferably smaller.

[0068] Moreover, such proteins include proteins having a sequence identity of 90% or more, 91% or more, 92% or more, 93% or more, 94% or more, 95% or more, 96% or more, 97% or more, 98% or more, 99% or more, 99.1% or more, 99.2% or more, 99.3% or more, 99.4% or more, 99.5% or more, 99.6% or more, 99.7% or more, 99.8% or more, or 99.9% or more to the amino acid sequence of the protein (g), having an amino acid sequence wherein the amino acid residue corresponding to the amino acid residue X_2 is Thr and/or the amino acid residue corresponding to the amino acid residue X_3 is Ser and having an activity to add hexose at position 2 of glucose at position 13 in a compound represented by formula (I). In general, the numerical value of the aforementioned sequence identity is preferably greater.

[0069] Here, the "activity to add hexose at position 2 of glucose at position 13 in a compound represented by formula (I)" means an activity to add hexose at position 2 of the glucose group at position 13 in a compound represented by the following formula (I).

[Formula 2]

(I)



[0070] In formula (I), Glc represents a glucose residue. Moreover, in formula (I), R_1 represents H, a C_1-C_{20} alkyl group, a C_2-C_{20} alkenyl group, a C_2-C_{20} alkynyl group, a C_4-C_{20} alkyldienyl group, a C_6-C_{18} aryl group, a C_6-C_{20} alkylaryl group, a C_6-C_{20} arylalkyl group, a C_4-C_{20} cycloalkyl group, a C_4-C_{20} cycloalkenyl group, a (C_3-C_{10} cycloalkyl)C_1-C_{10} alkyl group or a sugar residue.

[0071] As used herein, the "C_1-C_{20} alkyl group" is preferably a C_1-C_{10} alkyl group and more preferably a C_1-C_6 alkyl group, Examples of the alkyl group include, but are not limited to, methyl, ethyl, propyl, isopropyl, n-butyl, sec-butyl, tert-butyl, pentyl, hexyl, dodecanyl, and the like.

[0072] As used herein, the "C_2-C_{20} alkenyl group" is preferably a C_2-C_{10} alkenyl group and more preferably a C_2-C_6 alkenyl group. Examples of the alkenyl group include, but are not limited to, vinyl, allyl, propenyl, isopropenyl, 2-methyl-1-propenyl, 2-methylallyl, 2-butenyl, and the like.

[0073] As used herein, the "C_2-C_{20} alkynyl group" is preferably a C_2-C_{10} alkynyl group and more preferably a C_2-C_6 alkynyl group. Examples of the alkynyl group include, but are not limited to, ethynyl, 2-propynyl, 2-butynyl, and the like.

[0074] As used herein, the "C_4-C_{20} alkyldienyl group" is preferably a C_4-C_{10} alkyldienyl group and more preferably a C_4-C_6 alkyldienyl group. Examples of the alkyldienyl group include, but are not limited to, 1,3-butadienyl, and the like.

[0075] As used herein, the "C_6-C_{18} aryl group" is preferably a C_6-C_{10} aryl group. Examples of the aryl group include, but are not limited to, phenyl, 1-naphthyl, 2-naphthyl, indenyl, biphenylyl, anthryl, phenanthryl, and the like.

[0076] As used herein, the "C_6-C_{20} alkylaryl group" is preferably a C_6-C_{12} alkylaryl group. Examples of the alkylaryl group include, but are not limited to, o-tolyl, m-tolyl, p-tolyl, 2,3-xylyl, 2,4-xylyl, 2,5-xylyl, o-cumenyl, m-cumenyl, p-cumenyl, mesityl, and the like.

[0077] As used herein, the "C_6-C_{20} arylalkyl group" is preferably a C_6-C_{12} arylalkyl group. Examples of the arylalkyl group include, but are not limited to, benzyl, phenethyl, diphenylmethyl, triphenylmethyl, 1-naphthylmethyl, 2-naphthylmethyl, 2,2-diphenylethyl, 3-phenylpropyl, 4-phenylbutyl, 5-phenylpentyl, and the like.

[0078] As used herein, the "C_4-C_{20} cycloalkyl group" is preferably a C_4-C_{10} cycloalkyl group. Examples of the cycloalkyl group include, but are not limited to, cyclopropyl, cyclobutyl, cyclopentyl, cyclohexyl, and the like.

[0079] As used herein, the "C_4-C_{20} cycloalkenyl group" is preferably a C_4-C_{10} cycloalkenyl group. Examples of the cycloalkenyl group include, but are not limited to, cyclopropenyl, cyclobutenyl, 2-cyclopenten-1-yl, 2-cyclohexen-1-yl, 3-cyclohexen-1-yl, and the like.

[0080] Examples of the "C_3-C_{10} cycloalkyl)C_1-C_{10} alkyl group", as used herein, include methylcyclopropyl, ethylcyclopropyl, methylcyclobutyl, ethylcyclopentyl, methylcyclohexyl, and the like.

[0081] The "sugar residue", as used herein, is not particularly limited, but may be a sugar residue consisting of one or more pentose, hexose (including deoxyhexose), or a combination thereof.

[0082] Examples of the pentose include ribose, arabinose, and lyxose and examples of the hexose include allose, altrose, glucose, mannose, gulose, idose, galactose, talose, and rhamnose.

[0083] Preferably, the "sugar residue" is a sugar residue consisting of one or more hexose units and, more preferably, a sugar residue of a glucose monomer (-Glc) or a glucose dimer (-Glc-Glc). In the sugar residue of a glucose dimer, the glucose is preferably bound to each other by the β2,1-glycosidic linkage.

[0084] Preferably, the compound of formula (I) is steviolmonoside or rubusoside.

[0085] By the protein according to the present invention, hexose to be added to the compound represented by formula (I) at position 2 of glucose at position 13 in the compound is not particularly limited, but preferably a hexose selected from the group consisting of glucose, rhamnose, mannose, and galactose and, more preferably, glucose or rhamnose. Most preferably, the hexose is rhamnose. In the proteins according to (a), (b'), (b"), (c'), (c"), and (d) to (f) above, rhamnose can particularly advantageously be used as the hexose.

[0086] Moreover, in the proteins according to (g) to (i) above, glucose can particularly advantageously be used as the hexose.

[0087] The activity to add hexose at position 2 of glucose at position 13 in a compound represented by formula (I) can be validated by incubating a buffer solution (for example, a sodium phosphate buffer or a potassium phosphate buffer), at a pH in a neutral region of pH 6.0 to 8.0, containing a test protein, a 1 to 1000 μM (preferably 100 to 700 μM and most preferably 500 PM) UDP sugar (for example, UDP-glucose), and a 1 to 500 μM (preferably 100 to 500 μM and most preferably 250 μM) substrate compound (a compound of formula (I)) at a temperature of 20 to 40° C. for from 10 minutes to 2 hours, then purifying the aforementioned substrate compound, and analyzing the purified monoterpene by a known technique such as the liquid chromatography-mass spectrometry (LC-MS) analysis.

[0088] If a compound in which hexose is added at position 2 of glucose at position 13 in the compound represented by formula (I) is detected, as a result of the LC-MS analysis, then the test protein is considered to be that having an activity to add hexose at position 2 in glucose at position 13 in a compound represented by formula (I).

[0089] The aforementioned hexose addition reaction is usually completed in around 1 minute to 12 hours.

[0090] The one or more amino acid residues deleted, substituted, inserted, and/or added in the amino acid sequence of the protein according to the present invention means that there is deletion, substitution, insertion, and/or addition of one or more amino acid residues at the position (s) of any one or more amino acid sequences in the same sequence. Two or more of the deletion, substitution, insertion, and addition may occur simultaneously.

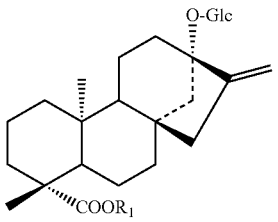
[0091] Examples of amino acid residues that may be substituted with each other are illustrated below. The amino acid residues included in a group may be substituted with each other. Group A: leucine, isoleucine, norleucine, valine, norvaline, alanine, 2-aminobutanoic acid, methionine, o-methylserine, t-butylglycine, t-butylalanine, cyclohexylalanine; Group B: aspartic acid, glutamic acid, isoaspartic acid, isoglutamic acid, 2-aminoadipic acid, 2-aminosuberic acid; Group C: asparagine, glutamine; Group D: lysine, arginine, ornithine, 2,4-diaminobutanoic acid, 2,3-diaminopropionic acid; Group E: proline, 3-hydroxyproline, 4-hydroxyproline; Group F: serine, threonine, homoserine; Group G: phenylalanine, tyrosine.

[0092] The protein according to the present invention can be obtained by expressing a polynucleotide (see the "Polynucleotide of the present invention" below) encoding this in an appropriate host cell, but it can be produced by a chemical synthetic method such as the fluorenylmethyl oxycarbonyl (Fmoc) method, the t-butyloxycarbonyl (tBoc) method, or the like. Moreover, the protein according to the present invention can be chemically synthesized with a peptide synthesizer such as that manufactured by Advanced Automation Peptide Protein Technologies, Inc., PerkinElmer, Inc., Protein Technologies Ltd., PerSeptive Biosystems, Inc., Applied Biosystems, or SHIMADZU CORPORATION.

2. Method of Producing Steviol Glycoside Steviol glycosides can easily be produced in large amounts by making use of the activity to add hexose at position 2 of glucose at position 13 in a compound represented by formula (I) that the protein according to the present invention has.

[0093] Therefore, in another embodiment, the present invention provides a method of producing a steviol glycoside, comprising reacting the protein according to the present invention, a UDP-sugar, and a compound represented by the following formula (I) to add hexose at position 2 of glucose at position 13 in the compound represented by formula (I).

[Formula 3]



(I)

[0094] The definition of Glc and R_1 in formula (I) is as described above. Preferably, the compound of formula (I) is steviolmonoside or rubusoside.

[0095] As used herein, the "UDP-sugar" is uridine diphosphate (UDP)-bound sugar. Preferred examples of the sugar moiety in the UDP-sugar include sugar consisting of one or more hexose. Examples of the hexose are as described above. Preferably, the UDP-sugar is UDP-hexose and more preferably, the sugar is a hexose selected from the group consisting of glucose, rhamnose, mannose, and galactose. Most preferably, the aforementioned UDP-sugar is UDP-rhamnose. When a protein described in any of the aforementioned (a), (b'), (b"), (c'), (c"), and (d) to (f) is used as the protein according to the present invention, UDP-rhamnose can be used particularly advantageously as the UDP-sugar. Moreover, when a protein described in any of the aforementioned (g) to (i) is used as the protein according to the present invention, UDP-glucose can be used particularly advantageously as the UDP-sugar.

[0096] The first method of producing a steviol glycoside according to the present invention comprises reacting the protein according to the present invention, a UDP-sugar, and a compound represented by formula (I) to add hexose at position 2 of glucose at position 13 in the compound represented by formula (I). The first method of production according to the present invention may further comprise purifying the steviol glycoside produced in the aforementioned step. Moreover, the first method of production according to the present invention may comprise further adding sugar to the steviol glycoside produced in the aforementioned step.

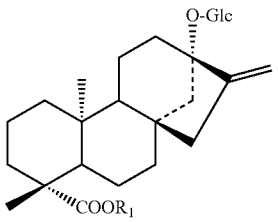
[0097] Examples of the steviol glycoside that is produced by the first method of production include, but are not limited to, steviolbioside, stevioside, dulcoside A, Reb.E, or Reb.C, or a combination thereof.

[0098] The produced steviol glycoside can be purified by a known method such as extraction with an appropriate solvent (an aqueous solvent such as water or an organic solvent such as alcohol, ether, and acetone), the gradient of an organic solvent such as ethyl acetate:water, high performance liquid chromatography (HPLC), gas chromatography, time-of-flight mass spectrometry (TOF-MS), or ultra (high) performance liquid chromatography (UPLC).

3. Non-Human Transformant Highly Containing Steviol Glycoside

[0099] The steviol glycoside can be produced using the protein according to the present invention in cells of a microorganism (*Escherichia coli*, yeast, or the like), a plant, an insect, a mammal other than humans. This is because the protein according to the present invention is an enzyme from *Stevia rebaudiana* or a variant thereof and therefore expected to have high activity in the intracellular environment. In this case, a steviol glycoside can be produced by introducing a polynucleotide encoding the protein according to the present invention (see the "polynucleotide of the present invention" described below) into host cells from a microorganism, a plant, an insect, or a mammal other than humans to express the protein according to the present invention and reacting the protein according to the present invention, UDP-sugar present in the aforementioned cells, and a compound represented by formula (I).

[Formula 4]



(I)

[0100] Accordingly, the present invention provides a non-human transformant (hereinafter, referred to as the "transformant of the present invention") in which a polynucleotide described in any one selected from the group consisting of the following (a) to (e) is introduced (hereinafter, referred to as the "polynucleotide of the present invention").

[0101] (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO: 1;

[0102] (b) a polynucleotide encoding a protein consisting of the amino acid sequences of SEQ ID NO: 2;

[0103] (c) a polynucleotide encoding a protein consisting of an amino acid sequence wherein 1 to 48 amino acids other than the amino acid residue X_7 are deleted, substituted, inserted, and/or added in the amino acid sequence of SEQ ID NO: 2 and having an activity to add hexose at position 2 of glucose at position 13 in a compound represented by formula (I) above;

[0104] (d) a polynucleotide encoding a protein having an amino acid sequence having a sequence identity of 90% or more to the amino acid sequence of SEQ ID NO: 2 wherein an amino acid corresponding to the amino acid residue X_7 is the same and having an activity to add hexose at position 2 of glucose at position 13 in a compound represented by formula (I) above:

[0105] Moreover, according to one aspect of the present invention,

[0106] (c') a polynucleotide encoding a protein consisting of an amino acid sequence wherein 1 to 48 amino acids other than the amino acid residues X_1 to X_7 are deleted, substituted, inserted, and/or added in the amino acid sequence of SEQ ID NO: 2 and having an activity to add hexose at position 2 of glucose at position 13 in a compound represented by formula (I) above; or

[0107] (d') a polynucleotide encoding a protein having an amino acid sequence having a sequence identity of 90% or more to the amino acid sequence of SEQ ID NO: 2 wherein amino acids corresponding to the amino acid residues X_1 to X_7 are the same and having an activity to add hexose at position 2 of glucose at position 13 in a compound represented by formula (I) above is provided.

[0108] According to another aspect of the present invention,

[0109] (c") a polynucleotide encoding a protein consisting of an amino acid sequence wherein 1 to 48 amino acids other than the amino acid residues X_2 and/or X_3 are deleted, substituted, inserted, and/or added in the amino acid sequence of SEQ ID NO: 2 and having an activity to add hexose at position 2 of glucose at position 13 in a compound represented by formula (I) above; or

[0110] (d") a polynucleotide encoding a protein having an amino acid sequence having a sequence identity of 90% or more to the amino acid sequence of SEQ ID NO: 2 wherein amino acids corresponding to the amino acid residues X_2 and/or X_3 are the same and having an activity to add hexose at position 2 of glucose at position 13 in a compound represented by formula (I) above is provided.

[0111] According to another aspect of the present invention,

[0112] (e) a polynucleotide encoding a protein wherein the 156th amino acid residue is Val and/or the 233rd amino acid residue is Phe in the amino acid sequence of SEQ ID NO: 4; or

[0113] (f) a polynucleotide encoding a protein consisting of an amino acid sequence wherein 1 to 48 amino acids are deleted, substituted, inserted, and/or added in the protein according to (e) above and an amino acid residue corresponding to the 156th amino acid residue of SEQ ID NO: 4 is Val and/or an amino acid residue corresponding to the amino acid residue of the 233rd of SEQ ID NO: 4 is Phe and having an activity to add hexose at position 2 of glucose at position 13 in a compound represented by formula (I) above; or

[0114] (g) a polynucleotide encoding a protein having an amino acid sequence having a sequence identity of 90% or more to the amino acid sequence of the protein according to (e) above wherein an amino acid residue corresponding to the 156th amino acid residue of SEQ ID NO: 4 is Val and/or an amino acid residue corresponding to the 233rd amino acid residue of SEQ ID NO: 4 is Phe and having an activity to add hexose at position 2 of glucose at position 13 in a compound represented by formula (I) above is provided.

[0115] According to another aspect of the present invention,

[0116] (h) a polynucleotide encoding a protein wherein the amino acid residue X_2 is Thr and/or the amino acid residue X_3 is Ser in the amino acid sequence of SEQ ID NO: 2; or

[0117] (i) a polynucleotide encoding a protein consisting of an amino acid sequence wherein 1 to 48 amino acids are deleted, substituted, inserted, and/or added in the protein according to (h) above and an amino acid residue corresponding to the amino acid residue X_2 is Thr and/or an amino acid residue corresponding to the amino acid residue X_3 is Ser and having an activity to add hexose at position 2 of glucose at position 13 in a compound represented by formula (I) above; or

[0118] (j) a polynucleotide encoding a protein having an amino acid sequence having a sequence identity of 90% or more to the amino acid sequence of the protein according to (h) above wherein an amino acid residue corresponding to the amino acid residue X_2 is Thr and/or an amino acid residue corresponding to the amino acid residue X_3 is Ser and having an activity to add hexose at position 2 of glucose at position 13 in a compound represented by formula (I) above is provided.

[0119] The definition and specific examples of formula (I) are as described above and the definition and specific examples of the hexose to be added at position 2 of glucose at position 13 in the compound represented by formula (I) are also as described above.

[0120] As used herein, the "polynucleotide" means DNA or RNA.

[0121] As used herein, the "polynucleotide that hybridizes under highly stringent conditions" refers to a polynucleotide obtained by performing colony hybridization, plaque hybridization, Southern hybridization, or the like, using for example, all or a part of a polynucleotide consisting of a nucleotide sequence complementary to the nucleotide sequence set forth in SEQ ID NO: 1 or a polynucleotide consisting of a nucleotide sequence encoding the amino acid sequence set forth in SEQ ID NO: 2 as a probe. Examples of available methods of hybridization include methods described in "Sambrook & Russell, Molecular Cloning A Laboratory Manual Vol. 3, 2001 Cold Spring Harbor, Laboratory Press", "Ausubel, Current Protocols in Molecular Biology, John Wiley & Sons 1987-1997", and the like.

[0122] Examples of the "highly stringent conditions" as used herein are, but not limited to, conditions of (1) 5×SSC, 5× den Hald solution, 0.5% SDS, 50% formamide, 50° C.; (2) 0.2×SSC, 0.1% SDS, 60° C., (3) 0.2×SSC, 0.1% SDS, 62° C.; (4) 0.2×SSC, 0.1% SDS, 65° C.; or (5) 0.1×SSC, 0.1% SDS, 65° C. Under these conditions, it can be expected that DNA having a high sequence identity is obtained more efficiently at higher temperatures. Meanwhile, it is considered that there are plural factors that have effects on the stringency of hybridization, such as temperature, the probe concentration, the probe length, the ionic strength, time, and the salt concentration and a person skilled in the art can attain similar stringency by selecting these factors as appropriate.

[0123] When using a commercially available kit for hybridization, for example, Alkphos Direct Labelling and Detection System (GE Healthcare) may be used. In this case, hybridized DNA can be detected after incubating a membrane with a labelled probe overnight in accordance with the protocol attached to the kit and then washing the membrane with a primary washing buffer containing 0.1% (w/v) SDS under conditions at 55 to 60° C. Alternatively, hybridization can be detected by using the DIG nucleic acid detection kit (Roche Diagnostics), when the probe is labeled with digoxigenin (DIG) by using a commercially available reagent (for example, PCR labeling mixture (Roche Diagnostics)), in the production of a probe based on a nucleotide sequence complementary to the nucleotide sequence set forth in SEQ ID NO: 1 or a sequence complementary to all or a part of a nucleotide sequence encoding the amino acid sequence set forth in SEQ ID NO: 2.

[0124] Examples of hybridizable polynucleotide other than those described above include DNA having a sequence identity of 80% or more, 81% or more, 82% or more, 83% or more, 84% or more, 85% or more, 86% or more, 87% or more, 88% or more, 89% or more, 90% or more, 91% or more, 92% or more, 93% or more, 94% or more, 95% or more, 96% or more, 97% or more, 98% or more, 99% or more, 99.1% or more, 99.2% or more, 99.3% or more, 99.4% or more, 99.5% or more, 99.6% or more, 99.7% or more, 99.8% or more, or 99.9% or more to the DNA of SEQ ID NO: 1 or DNA encoding the amino acid sequence set forth in SEQ ID NO: 2 as calculated by homology search software such as FASTA or BLAST using default parameters.

[0125] The sequence identity of an amino acid sequence or a nucleotide sequence can be determined using FASTA (Science 227 (4693): 1435-1441 (1985)) or the algorithm by

Karlin and Altschul BLAST (Basic Local Alignment Search Tool) (Proc. Natl. Acad. Sci. USA 872264-2268, 1990; Proc Natl Acad Sci USA 90: 5873, 1993). The programs called blastn, blastx, blastp, tblastn, and tblastx based on the algorithm of BLAST have been developed (Altschul S F, et al: J Mol Biol 215:403, 1990). When analyzing a nucleotide sequence using blastn, parameters are set at, for example, score=100, wordlength=12. Moreover, when analyzing an amino acid sequence using blastp, parameters are, for example, score=50, wordlength=3. When using BLAST and Gapped BLAST programs, the default parameters of each program are used.

[0126] The polynucleotide of the present invention described above can be obtained by a known genetic engineering technique or a known synthetic method.

[0127] The polynucleotide of the present invention is preferably introduced into a host in a state inserted into an appropriate expression vector.

[0128] An appropriate expression vector is usually configured to comprise:

[0129] (i) a promoter that allows the transcription in host cells,

[0130] (ii) the polynucleotide of the present invention connected to the promoter; and

[0131] (iii) an expression cassette comprising signals that function in host cells as a component for the termination of transcription of RNA molecules and the polyadenylation.

[0132] Examples of the method of producing the expression vector include, but are not particularly limited to, methods using a plasmid, a bacteriophage, a cosmid or the like.

[0133] The specific kind of vector is not particularly limited, but a vector expressible in host cells can be selected as appropriate. More specifically, a vector obtained by selecting a promoter sequence that ensures the expression of the polynucleotide of the present invention, as appropriate, depending on the kind of host cells and incorporating the promoter and the polynucleotide of the present invention into a certain plasmid may be used as an expression vector.

[0134] The expression vector according to the present invention contains an expression regulatory region (for example, a promoter, a terminator, and/or a replication origin) depending on the kind of the host in which the expression vector is to be introduced. A conventional promoter (for example, trc promoter, tac promoter, lac promoter) is used as the promoter of the expression vector for bacteria, examples of a promoter for yeast include the GAL1 promoter, the GAL10 promoter, the glyceraldehyde-3-phosphate dehydrogenase promoter, the PH05 promoter, and the like, and examples of a promoter for filamentous fungi include those for amylase and trpC, and the like. Moreover, examples of a promoter for expressing a gene of interest in plant cells include the 35S RNA promoter from cauliflower mosaic virus, the rd29A gene promoter, the rbcS promoter, the mac-1 promoter, which is a promoter obtained by adding an enhancer sequence of the aforementioned 35S RNA promoter from cauliflower mosaic virus on 5' of the mannopine synthase promoter sequence derived from *Agrobacterium*, and the like. Examples of a promoter for an animal cell host include a viral promoter (for example, the SV40 early promoter, the SV40 late promoter, or the like).

[0135] The expression vector preferably comprises at least one selection marker. Such markers that are available include auxotrophic markers (ura5, niaD, TRP1, URA3, HIS3, LEU2), drug resistance markers (hygromycin, Zeocin), Geneticin resistance genes (G418r), copper resistance genes (CUP1) (Marin et al., Proc. Natl. Acad. Sci. USA, vol. 81, p. 337, 1984), and cerulenin resistance genes (fas2m, PDR4) (Inokoshi, Junji, et al., Biochemistry, vol. 64, p. 660, 1992; and Hussain et al., Gene, vol. 101, p. 149, 1991, respectively).

[0136] The method of producing the transformant according to the present invention (method of production) is not particularly limited, but examples thereof include a method involving introducing an expression vector containing the polynucleotide of the present invention into a host to transform the host.

[0137] The transformant according to the present invention is expected to produce a steviol glycoside at high efficiency. The host cells to be used in transformation are not particularly limited, but various known cells may suitably be used. For example, examples of the host cells include bacteria such as *Escherichia coli*, yeast (the budding yeast *Saccharomyces cerevisiae*, the fission yeast *Schizosaccharomyces pombe*), plant cells, animal cells other than human cells, and the like. The host cells may be xenogeneic cells (cells from an organism other than *Stevia rebaudiana*) or allogeneic cells (cells from *Stevia rebaudiana*). When using allogeneic cells as a transformant, the transformant according to the present invention has the enzymatic reaction promoted in comparison with wild type *Stevia rebaudiana* cells since the protein according to the present invention introduced exogenously is expressed in addition to the protein expressed endogenously.

[0138] The host cells are preferably host cells that can produce a compound represented by formula (I). The host cells are not limited to those that can produce a compound represented by formula (I) in natural states, but may be, for example, those recombinantly engineered with a known gene so as to be capable of producing a compound represented by formula (I).

[0139] Examples of the known gene encoding an enzyme that contributes to the synthesis of the compound represented by formula (I) include, but are not limited to, EK13H, UGT74G1, UGT76G1 (Non-Patent Literature 2), and the like.

[0140] When the host cells are cells that cannot produce a compound represented by formula (I), a steviol glycoside can be produced by adding a compound of formula (I) or a plant extract containing the compound as a substrate to a culture system of the transformant obtained by introducing the gene of the present invention into a host cell, without introducing a gene encoding an enzyme that contributes to the synthesis of a compound represented by formula (I).

[0141] Furthermore, a more glycosylated steviol glycoside (for example, steviolbioside, Rebaudioside A, stevioside, and Rebaudioside B) can be produced by expressing the polynucleotide of the present invention in the host cells using the host cells in which a gene encoding a glycosyltransferase involved in a series of glycoside synthesis from steviol to Rebaudioside A is introduced. Examples of the glycosyltransferase involved in a series of glycoside synthesis from steviol to Rebaudioside A and a gene thereof include UGT85C2 (CDS sequence: SEQ ID NO: 7, amino acid sequence: SEQ ID NO: 8), UGT74G1 (CDS sequence: SEQ ID NO: 9, amino acid sequence: SEQ ID NO: 10),

UGT76G1 (CDS sequence: SEQ ID NO: 11, amino acid sequence: SEQ ID NO: 12), and the like.

[0142] The aforementioned appropriate culture media and conditions for host cells are well-known in the art. Moreover, the organism to be transformed is not particularly limited, but examples thereof include the microorganisms, plants, or animals other than humans illustrated for the aforementioned host cells.

[0143] For other general techniques in molecular biology, see Sambrook & Russell, "Molecular Cloning: A Laboratory Manual" Vol. 3, Cold Spring Harbor Laboratory Press 2001; "Methods in Yeast Genetics, A laboratory manual" (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY).

[0144] By culturing the transformant obtained in this way, it is possible to have the transformant produce a steviol glycoside. As described above, the production of the steviol glycoside can be promoted by adding the compound of formula (I) or a plant extract containing the compound as a substrate to a culture system of the transformant. The steviol glycoside of interest can be obtained by extracting and purifying the accumulated steviol glycoside.

[0145] Accordingly, the present invention provides a second method of producing a steviol glycoside characterized by using the transformant according to the present invention. Appropriate culture media and conditions are well-known in the art. The method of extracting and purifying a steviol glycoside is as described above.

[0146] The steviol glycoside is not particularly limited, but may preferably be one selected from the group consisting of steviolbioside, stevioside, dulcoside A, Reb.E, or Reb.C or a combination thereof.

[0147] In one aspect of the present invention, the host cells for transformation to be used may be any yeast. Specifically, the host cells include, but are not limited to, yeast such as those in the genus *Saccharomyces*.

[0148] Examples of available methods of transforming yeast include known methods that are generally used. The transformation can be conducted by methods such as, but not limited to, those described in Meth. Enzym., 194, p 182 (1990) (electroporation); Proc. Natl. Acad. Sci. USA, 75 p 1929(1978) (the spheroplast method); J.Bacteriology, 153, p. 163 (1983) (the lithium acetate method); Proc. Natl. Acad. Sci. USA, 75 p. 1929 (1978); Methods in yeast genetics, 2000 Edition: A Cold Spring Harbor Laboratory Course Manual; and the like. The transformant strain is obtained by selecting a strain that grows in a medium with a selective pressure for the selection marker used (for example, a medium containing an antibiotic or a medium lacking a nutrient).

[0149] In one aspect of the present invention, the transformant may be a plant transformant. The plant transformant according to the present embodiment is obtained by introducing a recombinant vector containing the polynucleotide according to the present invention into a plant such that the polypeptide encoded by the polynucleotide can be expressed. Alternatively, a new plant having the gene can be obtained by using the transformant according to the present invention as a crossing parent since the gene of the present invention is inherited by offspring.

[0150] When using a recombination expression vector, the recombination expression vector used in transformation of plant bodies is not particularly limited as long as it is a vector that can express the polynucleotide according to the present invention in the plant. Examples of such a vector include a vector having a promoter that allows the constitutive expression of a polynucleotide in plant cells or a vector having a promoter that is inductively activated by an external stimulus.

[0151] Examples of the promoter that allows the constitutive expression of a polynucleotide in plant cells include the 35S RNA promoter from cauliflower mosaic virus, the rd29A gene promoter, the rbcS promoter, the mac-1 promoter, and the like.

[0152] Examples of the promoter inducibly activated by an external stimulus include the mouse mammary tumor virus (MMTV) promoter, the tetracycline responsiveness promoter, the metallothionein promoter, the heat shock protein promoter, and the like.

[0153] The plant to be transformed in the present invention means the whole plant body, a plant organ (for example, a leaf, a petal, a stem, a root, a seed, or the like), plant tissue (for example, epidermis, phloem, parenchyma, xylem, vascular bundle, palisade tissue, spongy tissue, or the like) or cultured plant cells, various forms of plant cells (for example, suspension cultured cells), protoplasts, a leaf section, callus, or the like. The plant used in transformation is not particularly limited, but may be any of plants belonging to Monocotyledoneae or Dicotyledoneae.

[0154] The introduction of a gene into a plant is conducted by a method of transformation known to those skilled in the art (for example, the *Agrobacterium* method, the gene gun method, the PEG method, electroporation, particle bombardment, or the like).

[0155] The cells or plant tissue in which a gene has been introduced is first selected for drug resistance such as the hygromycin resistance and then reproduced into a plant by a conventional method. The reproduction of a plant from transformed cells may be conducted by a method known to those skilled in the art depending on the kind of the plant cells.

[0156] Whether the polynucleotide of the present invention has been introduced into a plant or not can be confirmed by PCR, Southern hybridization, Northern hybridization, or the like.

[0157] Once the transformed plant in which the polynucleotide of the present invention is incorporated in the genome is obtained, offspring can be obtained by sexual reproduction or asexual reproduction of the plant. Moreover, it is possible to obtain, for example, a seed, a fruit, cutting, a tuber, a tuberous root, a strain, a callus, a protoplast, or the like from the plant or offspring thereof or a clone thereof and produce the plant therefrom at a large quantity. Accordingly, the present invention also includes a plant in which the polynucleotide according to the present invention has been expressibly introduced, or offspring of the plant having the same traits as the plant, or tissue derived therefrom.

[0158] Moreover, methods for transforming various plants have been already reported. The transformant plant according to the present invention is not limited, but particularly preferable examples desirable to be used include plants known to bio-synthesize various glycosides using steviol as aglycones and examples of such a plant include *Stevia rebaudiana* and *Rubus suavissimus*.

[0159] The plant transformed with the polynucleotide of the present invention (hereinafter, the "plant of the present invention" or "plant body of the present invention") can produce a steviol glycoside more than the wild type when it

has an appropriate substrate or when an appropriate substrate is added from the outside.

[0160] From the plant of the present invention, a complete plant body can easily be obtained by growing a seed, a cutting, a bulb, or the like of the plant of the present invention.

[0161] Accordingly, the plant of the present invention includes a whole plant body, a plant organ (for example, a leaf, a petal, a stem, a root, a seed, a bulb), plant tissue (for example, epidermis, phloem, parenchyma, xylem, vascular bundle, palisade tissue, spongy tissue) or cultured plant cells, or various forms of plant cells (for example, suspension cultured cells), a protoplast, a section of leaf, a callus, and the like.

4. Extract of Transformant and Use Thereof

[0162] In another embodiment, the present invention also provides an extract of the aforementioned transformant. Since the transformant according to the present invention contains steviol glycosides at a higher content than the wild type when it has an appropriate substrate or when an appropriate substrate is added from the outside, extracts thereof are considered to contain steviol glycosides at high concentrations.

[0163] The extract of the transformant according to the present invention has a higher ratio of hexose, in particular rhamnose and/or glucose, to the aglycones in the whole steviol glycosides than extracts of the wild type *Stevia rebaudiana*, even when the transformant is cells from *Stevia rebaudiana*. Since the introduced protein according to the present invention is expressed and adds hexose to steviol glycosides in the transformant according to the present invention, the number of hexose residues added to a unit quantity (for example, a unit number of molecules) of steviol aglycones, when collectively taking the whole steviol glycosides, is of course increased in comparison with cells expressing no such protein (for example, wild type *Stevia rebaudiana* cells). The change in ratio of hexose to the aglycones in the whole steviol glycosides has effects on the properties of extract, for example, the sensory properties, for example, sweetness and the like.

[0164] The extract of the transformant according to the present invention can be obtained by homogenizing the transformant using glass beads, a homogenizer, or a sonicator, centrifuging the resultant homogenate, and collecting the supernatant. Furthermore, a further extraction step by the methods of extracting steviol glycosides described above may be conducted.

[0165] The extract of the transformant according to the present invention can be used, according to a conventional method, for a purpose such as the production of a food, a pharmaceutical product, an industrial raw material, or the like.

[0166] In another embodiment, the present invention also provides a food, a medicament, an industrial raw material (a raw material for a food or the like) comprising an extract of the transformant according to the present invention. The food, medicament, or industrial raw material containing an extract of the transformant according to the present invention is prepared according to a conventional method. As seen above, the food, medicament, or industrial raw material containing an extract of the transformant according to the present invention contains a steviol glycoside produced using the transformant according to the present invention. The aforementioned food, medicament, or industrial raw material may comprise an unnatural ingredient. Examples of the unnatural ingredient include a compound that does not occur naturally, for example, a synthetic additive such as a synthetic flavorant and a synthetic preservative, a fermentation product, and the like.

[0167] Examples of the food of the present invention include a dietary supplement, a health food, a functional food, a food for infants, a food for the elderly, and the like. As used herein, the food is a solid, a fluid, and a liquid, and a mixture thereof and is a generic name for edibles.

[0168] The dietary supplement refers to a food enriched with a particular nutrition ingredient. The health food refers to a food that is healthy or considered to be good for health and includes a dietary supplement, a natural food, a diet food, and the like. The functional food refers to a food for supplying a nutrition ingredient which serves a function in regulating physical conditions and is synonymous with a food for a specified health use. The food for infants refers to a food for feeding a child to up to about 6 years old. The food for the elderly refers to a food treated so as to be more easily digested and absorbed than a food with no treatment.

[0169] The food of the present invention uses a calorieless steviol glycoside as a sweetener. Therefore, the food of the present invention is low-calorie and has a merit of contributing to health promotion or health maintenance.

[0170] Examples of the forms of these foods may be agricultural foods such as bread, noodles, pasta, rice, confectionery (a cake, ice cream, popsicles, doughnuts, baked confectionery, candy, chewing gum, gummy candy, tablets, and Japanese sweets such as a dumpling and a steamed bun), tofu and processed products thereof; fermented foods such as refined sake, alcoholic drinks with medical properties, sweet sake, vinegar, soy sauce, and miso; livestock foods such as yogurt, ham, bacon, and sausage; sea foods such as kamaboko, fried fish paste, and cakes of ground fish; fruit juice beverages, refreshing beverages, sports beverages, alcoholic beverages; tea, and the like or flavoring agents. Examples of the forms of further foods include low calorie beverage, non-sugar beverage, canned fruits, milk beverage, beverage powder, yogurt, Jelly, dressing, noodle soup, pickle, a food boiled down in soy, soy sauce, miso, fish guts pickled in salt, Vermont vinegar, sweet pickled scallions, sweet and sour ginger, and a pickled lotus root, as well as a pickle, sauces for tempura and kabayaki, sauce for grilled meats, sources, gum, candy, toothpaste, a deep-fried patty of fish paste, rolled omelet, chow mein source, sauce for cold Chinese noodles, cut mackerel sprinkled with salt and then pickled in vinegar, ice cream, sherbet, soft ice cream, fish paste, snack food, rice confectionery, a corn cup, seasoned laver, bits of tempura batter that have fallen into the hot oil and been deep-fried, flaked seasoning for sprinkling over rice, and the like.

[0171] The food of the present invention encompasses any processed food. In one aspect, the processed food is made from natural raw materials, but different from natural products in the properties (for example, physical properties such as elasticity, viscosity, hardness, and the like; and sensory properties such as taste, smell, and texture). Many existing processed foods belong to this class. In another aspect, the processed food comprises an unnatural ingredient.

[0172] Moreover, the food of the present invention encompasses any beverage. In one aspect, the beverage is made from natural raw materials, but different from natural prod-

ucts in the properties (for example, physical properties such as viscosity and cohesive power; and sensory properties such as taste, smell, and texture). Examples of such a beverage include fermentation beverages (for example, lactic fermenting beverages, alcoholic beverages such as refined sake, wine, beer, alcoholic drinks with medical properties, half fermented tea such as oolong tea, total fermented tea such as black tea, and post-heating fermented tea such as Pu-erh tea), smoothie, milk shake, and the like. In another aspect, the beverage comprises an unnatural ingredient.

[0173] The dosage form of the pharmaceutical product (composition) according to the present invention is not particularly limited, and may be any dosage form, such as a solution, paste, gel, solid, or powder. Moreover, the pharmaceutical product (pharmaceutical composition) according to the present invention can be used in a skin external preparations such as an oil, a lotion, a cream, an emulsion, a gel, a shampoo, a hair rinse, a hair conditioner, enamel, a foundation, a lipstick, face powder, a pack, ointment, powder, toothpaste, aerosol, or a cleansing foam, as well as a bath preparation, hair tonic, skin essence, a sunscreen, or the like.

[0174] The pharmaceutical composition according to the present invention may further comprise another pharmaceutically active ingredients (for example, an anti-inflammatory ingredient) or an auxiliary ingredient (for example, a lubricant ingredient, a carrier ingredient) as needed. The pharmaceutically active ingredient or auxiliary ingredient may be a natural ingredient or an unnatural ingredient.

5. Method of Screening for Plant with High Content of Steviol Glycoside Having Rhamnose Group

[0175] The present invention provides a method of screening for a plant with a high content of steviol glycoside having rhamnose groups. Specifically, the aforementioned method comprises the following steps (1) to (3).

[0176] (1) extracting mRNA from a test plant;

[0177] (2) hybridizing the aforementioned mRNA or a cDNA prepared from the aforementioned mRNA with a polynucleotide that hybridizes with a polynucleotide consisting of a nucleotide sequence complementary to the polynucleotide of the present invention under highly stringent conditions;

[0178] (3) detecting the aforementioned hybridization.

[0179] The aforementioned step (1) can be carried out by extracting mRNA from a test plant. The part in the test plant from which mRNA is extracted is not particularly limited, but preferably a petal. When mRNA is extracted, cDNA may be prepared from mRNA by reverse transcription.

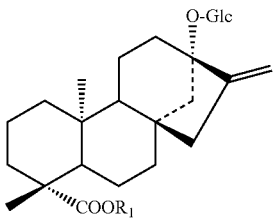
[0180] The step (2) can be carried out by hybridizing a polynucleotide or oligonucleotide consisting of nucleotide sequence complementary to the polynucleotide of the present invention as a probe or a primer with the mRNA extracted as described above under highly stringent conditions. The highly stringent conditions are as described above. The polynucleotide or oligonucleotide is preferably 5 to 500 bp, more preferably 10 to 200 bp, and further preferably 10 to 100 bp in length. The polynucleotide or oligonucleotide can easily be synthesized using various automatic synthesizers (for example, AKTA oligopilot plus 10/100 (GE Healthcare)) or the synthesis can be outsourced to a third party (for example, Promega Corporation or Takara Bio Inc.).

[0181] When using a polynucleotide consisting of a nucleotide sequence complementary to the polynucleotide of the present invention as a probe in step (2), the step (3) can be carried out by a method of hybridization detection such as usual Southern blotting, Northern blotting (Sambrook, Fritsch and Maniatis, "Molecular Cloning: A Laboratory Manual" 2nd Edition (1989), Cold Spring Harbor Laboratory Press), Microarray (see Affymetrix Inc., U.S. Pat. Nos. 6,045,996, 5,925,525, and 5,858,659), TaqMan PCR (Sambrook, Fritsch and Maniatis, "Molecular Cloning: A Laboratory Manual" 2nd Edition (1989), Cold Spring Harbor Laboratory Press), or Fluorescent In Situ Hybridization (FISH) (Sieben V. J. et al., (2007-06).IET Nanobiotechnology 1 (3): 27-35). Meanwhile, when using a polynucleotide consisting of a nucleotide sequence complementary to the polynucleotide of the present invention as a primer in the step (2), hybridization can be detected in the step (3) by performing the PCR amplification reaction and analyzing the obtained amplified product by electrophoresis or sequencing, or the like (Sambrook, Fritsch and Maniatis, "Molecular Cloning: A Laboratory Manual" 2nd Edition (1989), Cold Spring Harbor Laboratory Press).

[0182] The plant for which more hybridization has been detected is expected to have a high content of steviol glycosides having rhamnose groups in comparison with other plant bodies since it is considered to have higher expression of a protein having an activity to add hexose, particularly rhamnose, at position 2 of glucose at position 13 in a compound represented by the following formula (I).

[Formula 5]

(I)



6. Method of Producing UGT with Changed Sugar Donor Selectivity

[0183] In another side, the present invention provides a method of producing UGT with changed sugar donor selectivity. In one aspect, the aforementioned method comprises the following steps (1) to (2).

[0184] (1) identifying an amino acid residue corresponding to an amino acid residue selected from the amino acid residue X_2 in SEQ ID NO: 2, the amino acid residue X_3 in SEQ ID NO: 2, the 156th amino acid residue in SEQ ID NO: 4, and/or the 233rd amino acid residue in SEQ ID NO: 4 in the amino acid sequence of UGT; and

[0185] (2) substituting the amino acid residue corresponding to the amino acid residue X_2 in SEQ ID NO: 2 or the 156th amino acid residue in SEQ ID NO: 4, if the amino acid residue is Thr or Ser, preferably Thr, with Val, Leu, Ile, Ala, or Met, preferably Val, and, if the amino acid is Val, Leu, Ile, Ala, or Met, preferably Val, with Thr or Ser, preferably Thr, and/or substituting the amino acid residue corresponding to the amino acid

residue X_3 in SEQ ID NO: 2 or the 233rd amino acid residue in SEQ ID NO: 4, if the amino acid residue is Ser or Thr, preferably Ser, with Phe or Tyr, preferably Phe, and, if the amino acid is Phe or Tyr, preferably Phe, with Ser or Thr, preferably Ser.

[0186] In a particular aspect, the present invention provides a method of producing UGT whose sugar donor selectivity is shifted from UDP-glucose to UDP-rhamnose. In one aspect, the aforementioned method comprises the following steps (1) to (2).

[0187] (1) identifying UGT having an amino acid sequence in which the amino acid residue corresponding to the amino acid residue X_2 in SEQ ID NO: 2 or the 156th amino acid residue in SEQ ID NO: 4 is Thr or Ser, preferably Thr and/or the amino acid residue corresponding to the amino acid residue X_3 in SEQ ID NO: 2 or the 233rd amino acid residue in SEQ ID NO: 4 is Ser or Thr, preferably Ser;

[0188] (2) substituting the amino acid residue corresponding to the amino acid residue X_2 in SEQ ID NO: 2 or the 156th amino acid residue in SEQ ID NO: 4 in the amino acid sequence of UGT identified in (1), if the amino acid residue is Thr or Ser, preferably Thr, with Val, Leu, Ile, Ala, or Met, preferably Val, and/or substituting the amino acid residue corresponding to the amino acid residue X_3 in SEQ ID NO: 2 or the 233rd amino acid residue in SEQ ID NO: 4, if the amino acid is Ser or Thr, preferably Ser, with Phe or Tyr, preferably Phe.

[0189] In another aspect, the aforementioned method comprises the following steps (1) to (3).

[0190] (1) identifying UGT having a sugar donor selectivity for UDP-glucose,

[0191] (2) identifying an amino acid residue corresponding to an amino acid residue selected from the amino acid residue X_2 in SEQ ID NO: 2, the amino acid residue X_3 in SEQ ID NO: 2, the 156th amino acid residue in SEQ ID NO: 4, and/or the 233rd amino acid residue in SEQ ID NO: 4 in the amino acid sequence of UGT identified in (1);

[0192] (3) substituting the amino acid residue corresponding to the amino acid residue X_2 in SEQ ID NO: 2 or the 156th amino acid residue in SEQ ID NO: 4, if the amino acid residue is Thr or Ser, preferably Thr, with Val, Leu, Ile, Ala, or Met, preferably Val, and/or substituting the amino acid residue corresponding to the amino acid residue X_3 in SEQ ID NO: 2 or the 233rd amino acid residue in SEQ ID NO: 4, if the amino acid is Ser or Thr, preferably Ser, with Phe or Tyr, preferably Phe.

[0193] In another particular aspect, the present invention provides a method of producing UGT whose sugar donor selectivity is shifted from UDP-rhamnose to UDP-glucose.

[0194] In one aspect, the aforementioned method comprises the following steps (1) to (2).

[0195] (1) identifying UGT having an amino acid sequence in which the amino acid residue corresponding to the amino acid residue X_2 in SEQ ID NO: 2 or the 156th amino acid residue in SEQ ID NO: 4 is Val, Leu, Ile, Ala, or Met, preferably Val, and/or the amino acid residue corresponding to the amino acid residue X_3 in SEQ ID NO: 2 or the 233rd amino acid residue in SEQ ID NO: 4 is Phe or Tyr, preferably Phe;

[0196] (2) substituting the amino acid residue corresponding to the amino acid residue X_2 in SEQ ID NO: 2 or the 156th amino acid residue in SEQ ID NO: 4, if the amino acid residue is Val, Leu, Ile, Ala, or Met, preferably Val, with Thr or Ser, preferably Thr, and/or substituting the amino acid residue corresponding to the amino acid residue X_3 in SEQ ID NO: 2 or the 233rd amino acid residue in SEQ ID NO: 4, if the amino acid residue is Phe or Tyr, preferably Phe, with Ser or Thr, preferably Ser in the amino acid sequence of UGT identified in (1).

[0197] In another aspect, the aforementioned method comprises the following steps (1) to (3).

[0198] (1) identifying UGT having a sugar donor selectivity for UDP-rhamnose;

[0199] (2) identifying an amino acid residue corresponding to an amino acid residue selected from the amino acid residue X_2 in SEQ ID NO: 2, the amino acid residue X_3 in SEQ ID NO: 2, the 156th amino acid residue in SEQ ID NO: 4, and/or the 233rd amino acid residue in SEQ ID NO: 4 in the amino acid sequence of UGT identified in (1);

[0200] (3) substituting the amino acid residue corresponding to the amino acid residue X_2 in SEQ ID NO: 2 or the 156th amino acid residue in SEQ ID NO: 4, if the amino acid residue is Val, Leu, Ile, Ala, or Met, preferably Val, with Thr or Ser, preferably Thr, and/or substituting the amino acid residue corresponding to the amino acid residue X_3 in SEQ ID NO: 2 or the 233rd amino acid residue in SEQ ID NO: 4, if the amino acid residue is Phe or Tyr, preferably Phe, with Ser or Thr, preferably Ser.

[0201] In each method of producing UGT with changed sugar donor selectivity described above, the identification of the amino acid residue corresponding to the predetermined amino acid residue in the amino acid sequence of the UGT of interest can be made by any known method, for example, the alignment of the amino acid sequence of the UGT of interest and a standard amino acid sequence (SEQ ID NO: 2 or 4), the homology modeling using the protein consisting of the amino acid sequence of SEQ ID NO: 2 or 4 as a template, or the like.

[0202] The substitution of amino acid in the UGT of interest can be made, for example, by the aforementioned site-directed mutagenesis or the like. Moreover, the polynucleotide encoding the aforementioned amino acid related to the sugar donor selectivity may be modified by genome editing or the like to express a desired amino acid in the genome of the plant having the UGT of interest.

[0203] The identification of UGT having a sugar donor selectivity for UDP-glucose or UDP-rhamnose can be made based on literature or experimentally. To identify the aforementioned UGT experimentally, for example, the comparison can be made between the result of the reaction of UGT of interest, UDP-glucose or UDP-rhamnose, and a sugar receptor substrate (for example, the compound represented by formula (I)) under conditions suitable for the glycosidation reaction and the result of a reaction under the same conditions except that a compound different from one used in the aforementioned reaction is used as a sugar donor. When the reaction using UDP-glucose or UDP-rhamnose is more efficient than the reaction using another sugar donor, the UGT can be determined to have a sugar donor selectivity for UDP-glucose or UDP-rhamnose.

[0204] By any of the aforementioned methods, the UGT having a sugar donor selectivity for UDP-glucose or UDP-rhamnose can easily be designed or produced. Moreover, the sugar donor selectivity of UGT expressed in the plant can be changed to control the kind of glycoside sugar accumulated in the plant by the substitution of an amino acid residue of UGT at the genomic level.

7. Method of Estimating Sugar Donor Selectivity of UGT

[0205] In another side, the present invention provides a method of estimating the sugar donor selectivity of UGT. In one aspect, the aforementioned method comprises the following steps (1) to (2).

[0206] (1) identifying an amino acid residue corresponding to an amino acid residue selected from the amino acid residue X_2 in SEQ ID NO: 2, the amino acid residue X_3 in SEQ ID NO: 2, the 156th amino acid residue in SEQ ID NO: 4, and/or the 233rd amino acid residue in SEQ ID NO: 4 in the amino acid sequence of UGT; and

[0207] (2) estimating that the aforementioned UGT has a sugar donor selectivity for UDP-glucose if the amino acid residue corresponding to the amino acid residue X_2 in SEQ ID NO: 2 or the 156th amino acid residue in SEQ ID NO: 4 is Thr or Ser, preferably Thr, and/or the amino acid residue corresponding to the amino acid residue X_3 in SEQ ID NO: 2 or the 233rd amino acid residue in SEQ ID NO: 4 is Ser or Thr, preferably Ser; and estimating that the aforementioned UGT has a sugar donor selectivity for UDP-rhamnose if the amino acid residue corresponding to the amino acid residue X_2 in SEQ ID NO: 2 or the 156th amino acid residue in SEQ ID NO: 4 is Val, Leu, Ile, Ala, or Met, preferably Val, and/or the amino acid residue corresponding to the amino acid residue X_3 in SEQ ID NO: 2 or the 233rd amino acid residue in SEQ ID NO: 4 is Phe or Tyr, preferably Phe.

[0208] In the step (1) of the aforementioned method, the identification of the amino acid residue corresponding to the predetermined amino acid residue in the amino acid sequence of UGT of interest may be made as in the method of producing UGT with changed sugar donor selectivity.

[0209] By estimating the sugar donor selectivity of UGT by this method, it is possible to carry out designing of experiments for the characterization of UGT reasonably and emit unnecessary experiments.

8. Method of Estimating Kind of Sugar Attached to Position 2 of Glucose at Position 13 in Steviol Glycoside Accumulated in Plant

[0210] In another side, the present invention provides a method of estimating the kind of sugar attached to position 2 of glucose at position 13 in a steviol glycoside accumulated in a plant. In one aspect, the aforementioned method comprises the following steps (1) to (3).

[0211] (1) determining the amino acid sequence of UGT expressed in a plant;

[0212] (2) identifying an amino acid residue corresponding to an amino acid residue selected from the amino acid residue X_2 in SEQ ID NO: 2, the amino acid residue X_3 in SEQ ID NO: 2, the 156th amino acid residue in SEQ ID NO: 4, and/or the 233rd amino acid residue in SEQ ID NO: 4 in the amino acid sequence of UGT identified in (1);

[0213] (3) estimating that a proportion of glucose in the sugar attached to position 2 of glucose at position 13 in the steviol glycoside accumulated in the plant is high if the amino acid residue corresponding to the amino acid residue X_2 in SEQ ID NO: 2 or the 156th amino acid residue in SEQ ID NO: 4 is Thr or Ser, preferably Thr, and/or the amino acid residue corresponding to the amino acid residue X_3 in SEQ ID NO: 2 or the 233rd amino acid residue in SEQ ID NO: 4 is Ser or Thr, preferably Ser, and estimating that a proportion of rhamnose in the sugar attached to position 2 of glucose at position 13 in the steviol glycoside accumulated in the plant is high if the amino acid residue corresponding to the amino acid residue X_2 in SEQ ID NO: 2 or the 156th amino acid residue in SEQ ID NO: 4 is Val, Leu, Ile, Ala, or Met, preferably Val, and/or the amino acid residue corresponding to the amino acid residue X_3 in SEQ ID NO: 2 or the 233rd amino acid residue in SEQ ID NO: 4 is Phe or Tyr, preferably Phe.

[0214] The determination of the amino acid sequence of the UGT expressed in the plant in the step (1) of the aforementioned method can be made philologically or by any known experimental technique. Examples of the experimental identification technique of the amino acid sequence of the UGT include a technique involving extracting mRNA from a plant of interest, obtaining a gene having a similar nucleotide sequence using a probe or primer, etc. produced based on a known UGT nucleotide sequence, and identifying the amino acid sequence based on the gene, and the like.

[0215] In the step (2) of the aforementioned method, the identification of the amino acid residue corresponding to the predetermined amino acid residue in the amino acid sequence of the UGT of interest can be made similarly to the method of producing UGT with changed sugar donor selectivity.

[0216] In the step (3) of the aforementioned method, the term "a proportion of glucose in the sugar attached to position 2 of glucose at position 13 in the steviol glycoside accumulated in the plant is high" means that the proportion of glucose in the sugar attached to position 2 of glucose at position 13 in the steviol glycoside accumulated in the plant is higher when the amino acid residue corresponding to the amino acid residue X_2 in SEQ ID NO: 2 or the 156th amino acid residue in SEQ ID NO: 4 is Thr or Ser, preferably Thr, and/or the amino acid residue corresponding to the amino acid residue X_3 in SEQ ID NO: 2 or the 233rd amino acid residue in SEQ ID NO: 4 is Ser or Thr, preferably Ser, in the amino acid sequence of the UGT expressed in the aforementioned plant, in comparison with a different plant, for example, a plant that expresses UGT having an amino acid sequence in which the amino acid residue corresponding to the amino acid residue X_2 in SEQ ID NO: 2 or the 156th amino acid residue in SEQ ID NO: 4 is Val and the amino acid residue corresponding to the amino acid residue X_3 in SEQ ID NO: 2 or the 233rd amino acid residue in SEQ ID NO: 4 is Phe. The aforementioned proportion of glucose may be higher than a plant that expresses, for example, UGT having an amino acid sequence in which the amino acid residue corresponding to the amino acid residue X_2 in SEQ ID NO: 2 or the 156th amino acid residue in SEQ ID NO: 4 is Val and the amino acid residue corresponding to the

amino acid residue X_3 in SEQ ID NO: 2 or the 233rd amino acid residue in SEQ ID NO: 4 is Phe by about 10% or more, about 20% or more, about 30% or more, about 40% or more, about 50% or more, about 60% or more, about 70% or more, about 80% or more, about 90% or more, or about 100% or more.

[0217] In the step (3) of the aforementioned method, the term "a proportion of rhamnose in the sugar attached to position 2 of glucose at position 13 in the steviol glycoside accumulated in the plant is high" means that the proportion of rhamnose in the sugar attached to position 2 of glucose at position 13 in the steviol glycoside accumulated in the plant is higher when the amino acid residue corresponding to the amino acid residue X_2 in SEQ ID NO: 2 or the 156th amino acid residue in SEQ ID NO: 4 is Val, Leu, Ile, Ala, or Met, preferably Val, and/or the amino acid residue corresponding to the amino acid residue X_3 in SEQ ID NO: 2 or the 233rd amino acid residue in SEQ ID NO: 4 is Phe or Tyr, preferably Phe, in the amino acid sequence of the UGT expressed in the aforementioned plant, in comparison with a different plant, for example, a plant that expresses UGT having an amino acid sequence in which the amino acid residue corresponding to the amino acid residue X_2 in SEQ ID NO: 2 or the 156th amino acid residue in SEQ ID NO: 4 is Thr and the amino acid residue corresponding to the amino acid residue X_3 in SEQ ID NO: 2 or the 233rd amino acid residue in SEQ ID NO: 4 is Ser. The aforementioned proportion of rhamnose may be higher than a plant that expresses, for example, UGT having an amino acid sequence in which the amino acid residue corresponding to the amino acid residue X_2 in SEQ ID NO: 2 or the 156th amino acid residue in SEQ ID NO: 4 is Thr and the amino acid residue corresponding to the amino acid residue X_3 in SEQ ID NO: 2 or the 233rd amino acid residue in SEQ ID NO: 4 is Ser by about 10% or more, about 20% or more, about 30% or more, about 40% or more, about 50% or more, about 60% or more, about 70% or more, about 80% or more, about 90% or more, or about 100% or more.

[0218] By estimating the glycoside sugar accumulated in the plant by this method, it is possible to carry out designing of experiments for the characterization of the glycoside sugar reasonably and emit unnecessary experiments and to efficiently carry out the screening for a plant having a desired glycoside sugar profile.

9. Method of Screening for Plant Accumulating Glycoside Comprising Glucose or Rhamnose as Steviol Glycoside

[0219] In another side, the present invention provides a method of screening for a plant accumulating glycoside comprising glucose or rhamnose at position 2 of glucose at position 13 as a steviol glycoside. In one aspect, the aforementioned method comprises the following steps (1) to (3).

[0220] (1) determining the amino acid sequence of UGT expressed in a plant;

[0221] (2) identifying an amino acid residue corresponding to an amino acid residue selected from the amino acid residue X_2 in SEQ ID NO: 2, the amino acid residue X_3 in SEQ ID NO: 2, the 156th amino acid residue in SEQ ID NO: 4, and/or the 233rd amino acid residue in SEQ ID NO: 4 in the amino acid sequence of UGT identified in (1); (3) determining that the plant may have accumulated glycoside comprising glucose at position 2 of glucose at position 13 as a steviol glycoside, if the amino acid residue corresponding to the amino acid residue X_2 in SEQ ID NO: 2 or the 156th amino acid residue in SEQ ID NO: 4 is Thr or Ser, preferably Thr, and/or the amino acid residue corresponding to the amino acid residue X_3 in SEQ ID NO: 2 or the 233rd amino acid residue in SEQ ID NO: 4 is Ser or Thr, preferably Ser, and determining that the plant may have accumulated glycoside comprising rhamnose at position 2 of glucose at position 13 as a steviol glycoside, if the amino acid residue corresponding to the amino acid residue X_2 in SEQ ID NO: 2 or the 156th amino acid residue in SEQ ID NO: 4 is Val, Leu, Ile, Ala, or Met, preferably Val, and/or the amino acid residue corresponding to the amino acid residue X_3 in SEQ ID NO: 2 or the 233rd amino acid residue in SEQ ID NO: 4 is Phe or Tyr, preferably Phe.

[0222] The step (1) and (2) of the aforementioned method can be carried out similarly to the aforementioned "method of estimating the kind of sugar attached to position 2 of glucose at position 13 in steviol glycoside accumulated in a plant". The UGT to be investigated in the same plant may be one plant or plural plants. When there are plural UGTs to be investigated, there may be a case in which it is determined for a certain UGT that the plant may have accumulated glycoside comprising glucose at position 2 of glucose at position 13 as a steviol glycoside and for another UGT that the plant may have accumulated glycoside comprising rhamnose at position 2 of glucose at position 13 as a steviol glycoside. In such a case, it is possible to determine that the plant has accumulated both glycosides comprising glucose or rhamnose at position 2 of glucose at position 13 as a steviol glycoside.

[0223] By this method, it is possible to screen for a plant accumulating a desired glycoside without directly detecting the glycoside contained in the plant and to increase the efficiency of selection for a food raw material.

EXAMPLES

[0224] The present invention will be described more specifically by Examples below, but the scope of the present invention is not limited to these Examples.

[Example 1] Isolation of Candidate Gene for Steviol Glycoside Hexose Transferase

[0225] The molecular biological techniques used in this Example were according to methods described in Molecular Cloning (Sambrook et al., Cold Spring Harbor Laboratory Press, 2001) otherwise specified in detail.

[0226] cDNA from *Stevia rebaudiana* leaves was obtained by extracting total RNA from *Stevia rebaudiana* leaves using RNeasy Plant Mini kit (QIAGEN) and reverse-transcribing 0.5 μg the total RNA with Random Oligo-dT primers (RT).

[0227] The PCR reaction solution (50 μl) was prepared to have the composition of 1 μl of cDNA from *Stevia rebaudiana* leaves, 1×ExTaq buffer (Takara Bio), 0.2 mM dNTPs, 0.4 pmol/μl each of primers, 2.5 U of ExTaq polymerase. The following primers were used. Primer set for amplification of UGT91D2 and UGT91D2 #16 gene

Primer set for amplification of UGT91D2 and UGT91D2#16 gene

SrUGT91D2-pET15b-FW (SEQ ID NO: 37)

5 -TGCCGCGCGGCAGCCATATGTACAACGTTACTTATCATC-3

SrUGT91D2-pET15b-RV (SEQ ID NO: 38)

5 -GTTAGCAGCCGGATCCTTAACTCTCATGATCGATGGCAA-3

[0228] The PCR reaction included the reaction at 94° C. for 3 minutes and subsequent amplification with total 30 cycles of the reaction at 94° C. for 1 minute, at 50° C. for 1 minute, and at 72° C. for 2 minutes. Electrophoresis of the PCR product on a 0.8% agarose gel and staining with ethidium bromide have resulted in an amplified band at a size of about 1.4 kb estimated from each template DNA (FIG. **3**).

[0229] This PCR product was subcloned into pENTR-TOPO Directional vector (Invitrogen) in a way recommended by the manufacturer. The sequencing was carried out by primer walking with the synthesized oligonucleotide primers with DNA Sequencer model 3100 (Applied Biosystems). As a result, the presence of 2 genes was revealed.

[0230] One of them was a known gene UGT91D2 (CDS sequence: SEQ ID NO: 3, amino acid sequence: SEQ ID NO: 4) and the other one was a novel *Stevia rebaudiana* UGT gene (CDS sequence: SEQ ID NO: 1, amino acid sequence: SEQ ID NO: 2) highly homologous with UGT91D2. This homologous gene (UGT91D2L #16) exhibited a sequence identity of 99% (9 nucleotides are different) at the DNA level and 98% (7 residues are different) at the amino acid level with UGT91D2.

[Example 2] Construction of Expression Vector

[0231] An *Escherichia coli* expression vector for this enzyme gene was obtained by cutting out the UGT91D2 and UGT91D2L #16 ORF fragments with a size of about 1.4 kb using the NdeI and BamHI restriction enzyme sites (underlined parts in SEQ ID NOs: 5 and 6) added to the primers and ligating the fragments at the NdeI and BamHI sites in the *Escherichia coli* expression vector pET15b (Novagen). The expression vector was designed such that the His tag upstream of the NdeI site in this vector is in frame with the opening reading frame of the inserted gene and a chimeric protein in which UGT and the His tag are fused is expressed.

[Example 3] Expression and Purification of Recombination Protein

[0232] To elucidate the biochemical function of this enzyme, this enzyme was expressed in *Escherichia coli*. The *Escherichia coli* strain BL21 (DE3) was transformed with the plasmids for *Escherichia coli* expression of the two UGT91D2 genes obtained as described above according to a conventional method. The obtained transformant was cultured with shaking at 37° C. overnight in 4 ml of LB medium (10 g/l tryptone pepton, 5 g/l yeast extract, 1 g/l NaCl) containing 50 μg/ml ampicillin. 80 ml of the medium of the same composition was inoculated with 4 ml of the culture liquid reached to the stationary phase and the resultant culture was cultured with shaking at 37° C. IPTG was added at a final concentration of 0.5 mM when the bacterial turbidity (OD600) reached approximately 0.5 and the shaking culture was continued for 20 hr at 18° C.

[0233] All following operations were carried out at 4° C. The cultured transformant was collected by centrifugation (5,000×g, 10 min) and suspended by adding 1 ml/g cell of Buffer S [20 mM HEPES buffer (pH 7.5), 20 mM imidazole, 14 mM β-mercaptoethanol]. Subsequently, sonication (15 sec×8 times) and centrifugation (15,000×g, 15 min) were conducted. The obtained supernatant was collected as a crude enzyme liquid. The crude enzyme liquid was loaded onto equilibrated His SpinTrap (GE Healthcare) with Buffer S and centrifuged (70× g, 30 sec). After washing with the buffer, proteins bound to the column were eluted stepwise with 5 ml each of Buffer S containing 100 mM and 500 mM imidazole. Each elution fraction was subjected to buffer exchange into 20 mM HEPES buffer (pH 7.5), 14 mM β-mercaptoethanol using Microcon YM-30 (Amicon) (dialysis against approximately 500 volumes).

[0234] As a result of CBB staining and Western blot analysis using an anti-HisTag antibody after the SDS-PAGE separation of the prepared enzyme, protein was confirmed in the vicinity of the estimated molecular weight of about 50 kDa for chimeric HisTag-fused UGT91D2 and UGT91D2L #16 proteins in the 200 mM imidazole elution fraction. Therefore, this fraction was used for the enzymatic analysis (FIG. **4**).

[Example 4] Enzymatic Activity Measurement of UGT91D2 and UGT91D2L #16

[0235] The standard enzymatic reaction conditions are as follows. A reaction solution (2 mM UDP-glucose, 0.1 mM sugar receptor substrate, 100 mM potassium phosphate buffer (pH 7.5), 25 μl of purified enzyme solution) was prepared to 50 μl with distilled water and incubated at 30° C. for 1 hour to react. The LC-MS analysis of 5 μl of the enzymatic reaction solution was carried out under the following conditions.

LC Conditions

[0236] Column: Imtakt SM-C18 3.0 um 4.6 mm I.D.× 250 mm

[0237] Mobile phase: A: MilliQ Water (+0.2% acetic acid),

[0238] B: Methanol

[0239] Gradient: 0 to 5 min (B conc 10% constant), 5 to 20 min (B conc 10%→70%), 20 to 25 min (B conc 70%→100%), 25 to 35 min (B conc 100% constant), 35 to 36 min (B conc 100%→10%), 45 min end of analysis

[0240] Flow rate: 0.4 mL/min

[0241] Column oven: 40° C.

MS Conditions

[0242] ESI (negative mode)

[0243] Selected ion monitoring: m/z 641.3, 787.3, 803.3, 935.4, 949.4, 965.4, 1111.4, 1127.4, 1259.5, 1273.5, 1289.5, 1435.5

[0244] As described in prior literature (Patent Literature 2), the reaction of the recombinant SrUGT91D2 protein and UDP-glucose with rubusoside resulted in the production of stevioside in which position 2 of glucose at position 13 is glucosylated (FIG. **5-2**). Under similar conditions, SrUGT91D2L #16 also had the activity to glucosylate rubusoside to produce stevioside, but its activity was weaker than SrUGT91D2 (FIG. **5-3**). Meanwhile, transfer of hexose to rubusoside was not found in the negative control (FIG. **5-1**).

[0245] The reaction of the recombinant SrUGT91D2 protein and UDP-rhamnose with rubusoside then resulted in the production of dulcoside A in which position 2 of glucose at position 13 was rhamnosylated (FIG. **6-2**). Under similar conditions, SrUGT91D2L #16 has the activity to rhamnosylate rubusoside to produce dulcoside A and the activity was even markedly higher than SrUGT91D2 (FIG. **6-3**). Meanwhile, transfer of hexose to rubusoside was not found in the negative control (FIG. **6-1**). The specific activities of UGT91D2 and UGT91D2L #16 based on the above results are summarized in FIG. **7**. Moreover, the glycosylation pathway of steviol glycosides using UGT91D2 and UGT91D2L #16 is summarized in FIG. **8**.

[0246] Based on the foregoing results, it was revealed that SrUGT91D2 and SrUGT91D2L #16 have the activity to rhamnosylate position 2 of glucose at position 13 in rubusoside to produce dulcoside A. In particular, SrUGT91D2L #16, which has 7-amino acid residue difference from SrUGT91D2, has a striking rhamnose transfer activity and is considered to contribute to the synthesis of dulcoside A and Reb.C, a derivative thereof, in *Stevia rebaudiana*.

[0247] [Example 5] RebC synthesis by fermentative production using yeast Various steviol glycosides were produced from steviol in yeast. First, yeast that simultaneously expresses the 4 glycosylation enzyme genes UGT85C2 (SEQ ID NO: 7), UGT91D2 (SEQ ID NO: 3), UGT74G1 (SEQ ID NO: 9), and UGT76G1 (SEQ ID NO: 11) or the 4 genes UGT85C2, UGT91D2L #16, UGT74G1, and UGT76G1 from *Stevia rebaudiana* and the UDP-rhamnose synthase gene AtRHM2 (SEQ ID NO: 13) from *Arabidopsis thaliana* has been created.

[0248] Cloning of glycosylation enzyme gene and rhamnose synthase gene cDNA The cDNA cloning of UGT91D2 and UGT91D2L #16 from *Stevia rebaudiana* was carried out as described above. To clone other UGT genes from *Stevia rebaudiana*, the following primer sets were prepared.

Primer Set for Amplification of UGT85C2 Gene

[0249] CACC-NdeI-SrUGT85C2-Fw (underlined part is NdeI recognition site):

14)
5′ -CACCCATATGGATGCAATGGCTACAACTGAGAA-3′ (SEQ ID NO:

[0250] BglII-SrUGT85C2-Rv (underlined part is BglII recognition site):

15)
5′ -AGATCTCTAGTTTCTTGCTAGCACGGTGATTT-3′ (SEQ ID NO:

Primer set for amplification of UGT74G1 CACC-NdeI-SrUGT74G1-Fw (underlined part is NdeI recognition site): GP-29,DNA

5′ -CACCCATATGGCGGAACAACAAAAGATCAAGAAAT-3′ (SEQ ID NO: 16)

[0251] BamHI-SrUGT74G1-Rv (underlined part is BamHI recognition site):

5′ -GGATCCTTAAGCCTTAATTAGCTCACTTACAAATT-3′ (SEQ ID NO: 17)

Primer Set for Amplification of UGT76G1

[0252] CACC-NdeI-SrUGT76G1-Fw (underlined part is NdeI recognition site):

(SEQ ID NO: 18)
5′ -CACCCATATGGAAAATAAAACGGAGACCA-3′

[0253] BamHI-SrUGT76G1-Rv (underlined part is BamHI recognition site):

19)
5′ -GGATCCTTACAACGATGAAATGTAAGAAACTA-3′ (SEQ ID NO:

[0254] The PCR reaction solution (50 μl) was prepared to have the composition of 1 μl of cDNA from *Stevia rebaudiana* leaves, 1×KOD plus buffer (TOYOBO), 0.2 mM dNTPs, 0.4 pmol/μl primers, 1 mM MgSO4, and 1 U heat-resistant KOD plus polymerase. The PCR reaction included the reaction at 95° C. for 5 minutes and subsequent amplification with total 30 cycles of the reaction at 94° C. for 0.5 minutes, at 50° C. for 0.5 minutes, and at 68° C. for 2 minutes. Electrophoresis of the PCR products on a 0.8% agarose gel and staining with ethidium bromide resulted in an amplified band at a size of about 1.4 kb estimated from each template DNA.

[0255] This PCR product was subcloned into pENTR-TOPO Directional vector (Invitrogen) in a way recommended by the manufacturer. The sequencing was carried out by primer walking with the synthesized oligonucleotide primers with DNA Sequencer model 3100 (Applied Biosystems) to confirm the cloning of all the intended UGT genes, that is to say, UGT85C2, UGT74G1, and UGT76G1.

Construction of Expression Vector for Yeast

[0256] To incorporate these UGT and UDP-rhamnose synthase genes and the UDP-rhamnose synthase gene AtRHM2 from *Arabidopsis thaliana* (J Biol Chem 2007 Oka et. al) into a yeast expression vector, the following primer sets were designed.

SrUGT85C2 Set

[0257] Bgl2-UGT85C2-F (underlined part is BglII recognition site):

20)
5′ -ACAGATCTATGGATGCAATGGCTACAACTGAGA-3′
(SEQ ID NO:

[0258] Sal-UGT85C2-R (underlined part is SalI recognition site):

21)
5′ -TAGTCGACTAGTTTCTTGCTAGCACGGTGATTTC-3′
(SEQ ID NO:

SrUGT91D2 Set

[0259] NotI-UGT91DIL3-F (underlined part is NotI recognition site):

(SEQ ID NO: 22)
5′ -AAGCGGCCGCATGTACAACGTTACTTATCATCAAAATTCAAA-3′

[0260] Pac-UGT91D1L3-R (underlined part is PacI recognition site):

23)
5′ -CGTTAATTAACTCTCATGATCGATGGCAACC-3′ (SEQ ID NO:

SrUGT74G1 Set

[0261] Not-UGT74G1-F (underlined part is NotI recognition site):

24)
5′ -AAGCGGCCGCATGGCGGAACAACAAAAGATCAAG-3′
(SEQ ID NO:

[0262] Pac-UGT74G1-R (underlined part is PacI recognition site):

5′ -CGTTAATTAAGCCTTAATTAGCTCACTTACAAATTCG-3′
(SEQ ID
NO: 25)

SrUGT76G1 Set

[0263] Bam-UGT76G1-F (underlined part is BamHI recognition site):

26)
5′ -AAGGATCCATGGAAAATAAAACGGAGACCACCG-3′
(SEQ ID NO:

[0264] Sal-UGT76G1-R (underlined part is SalI recognition site):

(SEQ ID NO: 27)
5′ -GCGTCGACTTACAACGATGAAATGTAAGAAACTAGAGACTCTAA-3′

AtRHM2 Set

[0265] Bam-AtRHM2-F (underlined part is BamHI recognition site):

28)
5′ -GGATCCATGGATGATACTACGTATAAGCCAAAG-3′
(SEQ ID NO:

[0266] Xho-AtRHM2-R (underlined part is XhoI recognition site):

29)
5′ -CTCGAGTTAGGTTCTCTTGTTTGGTTCAAAGA-3′
(SEQ ID NO:

[0267] Using the combinations of template and primers, UGT85C2 as template and the SrUGT85C2 set, UGT91D2

or UGT91D2L #16 as template and the SrUGT91D2 set, UGT74G1 as template and the SrUGT74G1 set, UGT76G1 as template and the SrUGT76G1 set, and AtRHM2 as template and the AtRHM2 set, and heat-resistant KOD DNA polymerase (Toyobo), PCR amplification was conducted to add restriction enzyme sites to the both ends of each ORF. The obtained DNA fragments were subcloned using the zero Blunt-TOPO PCR cloning kit (Invitrogen) and the sequencing was carried out by primer walking with the synthesized oligonucleotide primers with DNA Sequencer model 3100 (Applied Biosystems) to confirm that each of the intended UGT genes was cloned.

[0268] To express the genes in yeast, the following expression vectors were constructed using the pESC yeast expression system (Stratagene).

[0269] (1) Construction of Plasmid pESC-URA-UGT56 or pESC-URA-UGT56R

[0270] The plasmid pESC-URA-UGT-1 was obtained by cutting out UGT85C2 with the restriction enzyme BglII and the restriction enzyme SalI and ligating the UGT85C2 into the vector pESC-URA (Stratagene) cut with the restriction enzyme BamHI and the restriction enzyme SalI. This plasmid pESC-URA-UGT-1 cut with the restriction enzyme NotI and the restriction enzyme PacI and UGT91D2 or UGT91D2L #16 cut with the restriction enzyme NotI and the restriction enzyme PacI were ligated to obtain pESC-URA-UGT56 or pESC-URA-UGT56R.

[0271] (2) Construction of Plasmid pESC-HIS-UGT78

[0272] The plasmid pESC-HIS-UGT-8 was obtained by cutting out UGT76G1 with the restriction enzyme BamHI and the restriction enzyme SalI and ligating the UGT76G1 with the vector pESC-HIS (Stratagene) cut with the same restriction enzymes. This plasmid pESC-HIS-UGT-8 cut with the restriction enzyme NotI and the restriction enzyme PacI and UGT74G1 cut with NotI and PacI were ligated to obtain pESC-HIS-UGT78.

[0273] (3) Construction of Plasmid pESC-TRP-AtRHM2

[0274] The plasmid pESC-TRP-AtRHM2 was obtained by cutting out AtRHM2 with the restriction enzyme BamHI and the restriction enzyme XhoI and ligating the AtRHM2 with the vector pESC-TRP (Stratagene) cut with the same restriction enzymes.

Transformation of Yeast

[0275] The plasmids set forth in Table 1 were introduced into the *Saccharomyces cerevisiae* strain YPH499 (ura3-52 lys2-801^{amber} ade2-101^{ochre} trp1-Δ63 his3-Δ200 leu2-Δ1 a) as a host by the lithium acetate method. Those that grew on SC-Trp&Ura&His agar medium (6.7 g of Yeast nitrogen base without amino acids, 20 g of glucose, 1.3 g of amino acid mixture powder Trp&Ura&His, 20 g of Bacto agar, per 1 L) were selected as transformant strains.

TABLE 1

Transformant strain	Introduced plasmid	Introduced gene
A-5678	pESC-URA-UGT56	SrUGT85C2, SrUGT91D2
	pESC-HIS-UGT78	SrUGT74G1, SrUGT76G1
	pESC-TRP-AtRHM2	AtPHM2
A-56R78	pESC-URA-UGT56R	SrUGT85C2, UGT91D2L#16
	pESC-HIS-UGT78	SrUGT74G1, SrUGT76G1
	pESC-TRP-AtARM2	AtRHM2

TABLE 1-continued

Transformant strain	Introduced plasmid	Introduced gene
C-5678	pESC-URA-UGT56	SrUGT85C2, SrUGT91D2
	pESC-HIS-UGT78	SrUGT74G1, SrUGT76G1
	pESC-TRP	
C-56R78	pESC-URA-UGT56R	SrUGT85C2, UGT91D2L#16
	pESC-HIS-UGT78	SrUGT74G1, SrUGT76G1
	pESC-TRP	

[0276] The amino acid mixture powder Trp&Ura&His was prepared by mixing 2.5 g of adenine sulfate, 1.2 g of L-arginine hydrochloride, 6.0 g of L-aspartic acid, 6.0 g of L-glutamic acid, 3.6 g of L-leucine, 1.8 g of L-lysine, 1.2 g of L-methionine, 3.0 g of L-phenylalanine, 22.5 g of L-serine, 12 g of L-threonine, 1.8 g of L-tyrosine, and 9.0 g of L-valine.

Induction and Analysis of Expression of Transgene

[0277] The obtained transformant strain was cultured as follows.

[0278] First, 10 ml of SC-Trp&Ura&His liquid medium (SC-Trp&Ura&His agar medium without Bacto agar) was inoculated each transformant strain as a preculture and cultured with shaking at 30° C. for 1 day. Then, 10 ml of SG-Trp&Ura&His liquid medium (6.7 g of Yeast nitrogen base without amino acids, 20 g of galactose, 1.3 g of amino acid mixture powder Trp&Ura&His, per 1 L) was inoculated with 1 ml of the liquid preculture as a main culture and cultured with shaking at 30° C. for 2 days.

[0279] To confirm whether the gene introduced in the transformant strain is expressed, bacterial cells were collected from the liquid culture and total RNA was purified with RNeasy Mini Kit.

[0280] cDNA was synthesized by taking 1 μg of the total RNA and using Super script II reverse transcriptase (Thermo Fisher Scientific) and random hexamers as primers.

[0281] To confirm the expression of the transgenes, the following primers were prepared.

For confirmation of expression of UGT85C2 UGT85C2-r1:

(SEQ ID NO: 30)
5′ -CAAGTCCCCAACCAAATTCCGT-3′

For confirmation of expression of UGT91D2 and UGT91D2L3 #16

[0282] UGT91D1L3-r1:

(SEQ ID NO: 31)
5′ -CACGAACCCGTCTGGCAACTC-3′

For confirmation of expression of UGT74G1

[0283] UGT74G1-r1:

(SEQ ID NO: 32)
5′ -CCCGTGTGATTTCTTCCACTTGTTC-3′

For confirmation of expression of UGT76G1

[0284] UGT76G1-r1:

(SEQ ID NO: 33)
5′ -CAAGAACCCATCTGGCAACGG-3′

For confirmation of expression of AtRHM2

[0285] AtRHM2-r1

(SEQ ID NO: 34)
5′ -GCTTTGTCACCAGAATCACCATT-3′

GAL10p region (Promoter region)

[0286] PGAL10-f3:

(SEQ ID NO: 35)
5′ -GATTATTAAACTTCTTTGCGTCCATCCA-3′

GAL1p region (Promoter region

[0287] PGAL1-f3:

36)
5′ -CCTCTATACTTTAACGTCAAGGAGAAAAAACC-3′ (SEQ ID NO:

[0288] Expression of each transgene was confirmed by performing PCR using the following combinations of primers, the previously synthesized cDNA as template, and ExTaq (Takara Bio) and agarose gel electrophoresis of the products.

UGT85C2:UGT85C2-r1 (SEQ ID NO: 30) and PGAL1-f3 (SEQ ID NO: 36)

UGT91D2 or UGT91D2L3:UGT91D1L3-r1 (SEQ ID NO: 31) and PGAL10-f3 (SEQ ID NO: 35)

UGT74G1:UGT74G1-r1 (SEQ ID NO: 32) and PGAL1-f3 (SEQ ID NO: 36)

UGT76G1:UGT76G1-r1 (SEQ ID NO: 33) and PGAL10-f3 (SEQ ID NO: 35)

AtRHM2:AtRHM2-r1 (SEQ ID NO: 34) and PGAL10-f3 (SEQ ID NO: 35)

[0289] This confirmed that the introduced genes were expressed in the transformant strains.

Production of Reb.C

[0290] The culture was conducted under the same conditions as Example 3 above, except that 0.5 μg or 2 μg of steviol (ChromaDex Inc.) per 1 ml of medium was added to the liquid medium for the main culture. After completing the culture, the supernatant and bacterial cells were separated by centrifugation of the liquid culture. The culture supernatant was washed with acetonitrile and then loaded on the Sep-Pak C18 column equilibrated with water, washed with 20% acetonitrile and then eluted with 80% acetonitrile, dried, and then dissolved in a little amount of 80% acetonitrile to prepare a glycoside sample. This glycoside sample was subjected to the following analyses.

Analysis by HPLC

[0291] The obtained glycoside sample was analyzed with high performance liquid chromatography (HPLC). The HPLC conditions are as follows.

[0292] Column: COSMOSIL $5C_{18}$-AR-II 4.6 mm I.D.× 250 mm (Nacalai Tesque, Inc.)

[0293] Mobile phase: A; Acetonitrile, B; 10 mM sodium phosphate buffer (pH 2.6), Gradient: 40 min, B conc 70%→30%, linear gradient

[0294] Flow rate: 1 ml/min

[0295] Temperature: 40° C.

[0296] Detection: UV 210 nm

Analysis by LC-MS

[0297] The analysis by LC-MS was conducted under the following conditions.

LC conditions

[0298] Column: Imtakt SM-C18 3.0 um 4.6 mm I.D.× 250 mm

[0299] Mobile phase: A: MilliQ Water (+0.2% acetic acid),

B: Methanol

[0300] Gradient: 0 to 5 min (B conc 10% constant), 5 to 20 min (B conc 10%→70%), 20 to 25 min (B conc 70%→100%), 25 to 35 min (B conc 100% constant), 35 to 36 min (B conc 100%→10%), 45 min end of analysis

[0301] Flow rate: 0.4 mL/min

[0302] Column oven: 40° C.

MS conditions

[0303] ESI (negative mode)

[0304] Selected ion monitoring: m/z 641.3, 787.3, 803.3, 935.4, 949.4, 965.4, 1111.4, 1127.4, 1259.5, 1273.5, 1289.5, 1435.5

[0305] The results are shown in FIGS. **9** to **11**. Reb.C was produced in the strains A-5678 and A-56R78 coexpressing the UDP-rhamnose synthase gene and the steviol glycosylation enzyme gene. The amount of Reb.C produced in the strain A-56R78 expressing UGT91D2L #16 was higher than that in the strain A-5678 expressing UGT91D2.

[Example 6] Analysis of SrUGT91D2L #16

Homology Structure Analysis

[0306] The amino acid residue that contributes to a striking rhamnose transfer activity in UGT91D2L #16 was estimated. Homology structure analysis was performed to estimate which residue of the 7 residues different between UGT91D2L #16 and UGT91D2 is related to the activity. A homology model of UGT91D2 in an existing method (Noguchi et al (2009) Plant Cell. 21 (5): 1556-152) was produced by using the structural information (RCSB Protein Data Bank (PDB), PDB ID:2PQ6) of UGT85H2 from *Medicago* in the family Fabaceae, of which crystal structure has been already solved, as the protein structure template, with the sugar donor UDP-glucose (PDB ID:$2C_1Z$) docked as a substrate.

[0307] The determination of the locations of the 7 residues that are different from UGT91D2L #16 in a constructed UGT91D2 homology model revealed that the 156th Thr residue is in the vicinity of UDP-glucose, suggesting its contribution to the selectivity of UDP-sugar. The 233rd Ser residue was also found in the vicinity of the Thr residue, suggesting its contribution to the UDP-sugar selectivity (FIGS. **12** to **13**).

Generation of Mutants

[0308] The mutants in which the amino acids estimated by the homology model are substituted in UGT91D2 and UGT91D2L #16 were generated. Mutants 1 and 2 were used in the in vitro assay with the *Escherichia coli*-expressed protein and Mutants 1 to 6 were used in the in vivo assay in yeast.

Mutant 1: UGT91D2-T156V (with substitution of the 156th Thr residue in UGT91D2 with a Val residue)

Mutant 2: UGT91D2L #16-V156T (with substitution of the 156th Val residue in UGT91D2L #16 with a Thr residue)

Mutant 3: UGT91D2-S233F (with substitution of the 233rd Ser residue in UGT91D2 with a Phe residue)

Mutant 4: UGT91D2L #16-F233S (with substitution of the 233rd Phe residue in UGT91D2L #16 with a Ser residue)

Mutant 5: UGT91D2-T156V/S233F (with substitution of the 156th Thr residue and the 233rd Ser residue in UGT91D2 with a Val residue and a Phe residue, respectively)

Mutant 6: UGT91D2L #16-V156T/F233 (with substitution of the 156th Val residue and the 233rd Phe residue in UGT91D2L #16 with a Thr residue and a Ser residue, respectively)

[0309] The plasmids containing DNA encoding the mutants were prepared as follows. For Mutant 1 as an example, an amplified fragment in which two fragments were ligated was obtained by mixing a DNA fragment obtained by the PCR with cDNA of UGT91D2 as a template and a primer set of the following SrUGT91D2-T156V-FW and SrUGT91D2-pET15b-RV (SEQ ID NO: 39 and SEQ ID NO: 38) and a DNA fragment obtained by the PCR with cDNA of UGT91D2 as a template and a primer set of SrUGT91D2-T156V-RV and SrUGT91D2-pET15b-FW (SEQ ID NO: 40 and SEQ ID NO: 37) and performing PCR again with a primer set of SrUGT91D2-pET15b-FW and SrUGT91D2-pET15b-RV (SEQ ID NO: 37 and SEQ ID NO: 38). By inserting this ligated fragment into the *Escherichia coli* expression vector pET15b (Novagen) digested with NdeI and BamHI using GeneArt Seamless Cloning and Assembly (Thermo Fisher Scientific Inc.), a plasmid for expression in *Escherichia coli* containing DNA encoding the UGT91D2-T156V mutant was obtained. The DNA encoding the UGT91D2-T156V mutant was sequenced with DNA Sequencer model 3100 (Applied Biosystems) by primer walking with synthesized oligonucleotide primers to confirm that there was no mutation other than the intended mutation. The plasmids containing DNA encoding other mutants were similarly prepared.

SrUGT91D2-pET15b-FW:

[0310]

(SEQ ID NO: 37)
5′ -TGCCGCGCGGCAGCCATATGTACAACGTTACTTATCATC-3′

SrUGT91D2-pET15b-RV:

[0311]

(SEQ ID No: 38)
5′ -GTTAGCAGCCGGATCCTTAACTCTCATGATCGATGGCAA-3′

SrUGT91D2-T156V-FW:

[0312]

(SEQ ID NO: 39)
5′ -CACTTCTCCGTCGTCACTCCATG-3′

SrUGT91D2-T156V-RV:

[0313]

(SEQ ID NO: 40)
5′ -CATGGAGTGACGACGGAGAAGTG-3′

SrUGT91D2-Like-V156T-FW

[0314]

(SEQ ID NO: 41)
5′ -CACTTCTCCGTCACCACTCCATG-3′

SrUGT91D2-Like-V156T-RV

[0315]

(SEQ ID NO: 42)
5′ -CATGGAGTGGTGACGGAGAAGTG-3′

SrUGT91D2-S233F-FW

[0316]

(SEQ ID NO: 43)
5′ -CTGATTGTTTGCTTTTCAAATGTTACCATGAG-3′

SrUGT91D2-S233F-RV

[0317]

(SEQ ID NO: 44)
5′ -CTCATGGTAACATTTGAAAAGCAAACAATCAG-3′

SrUGT91D2L-F233S-FW

[0318]

(SEQ ID NO: 45)
5′ -CTGATTGTTTGCTTTCCAAATGTTACCATGAG-3′

SrUGT91D2L-F233S-RV

[0319]

(SEQ ID NO: 46)
5′ -CTCATGGTAACATTTGGAAAGCAAACAATCAG-3′

Evaluation in *Escherichia coli*

[0320] Any of the 6 mutant proteins expressed in *Escherichia coli* using the same method as Examples 3 and 4 was

confirmed by CBB staining and Western blot analysis to have a band in the vicinity of the estimated size of 55 kDa.

[0321] Evaluation of Mutants 1 and 2 was then made using the protein expressed in *Escherichia coli* in the same manner as in Example 4. The glucosylation activity (GlcT) and the rhamnosylation activity (RhaT) were evaluated using UDP-glucose and UDP-rhamnose, respectively, as the sugar donor and using rubusoside as the substrate, which is a sugar receptor. Stevioside and dulcoside A were produced as the reaction products in the former and in the latter, respectively.

[0322] In Mutant 1, the relative GlcT activity decreased to 5.5% based on the GlcT activity of the wildtype UGT91D2 of 100% (FIG. **14**A) and the relative RhaT activity was a high activity of about 4.4 times based on the RhaT activity of the wildtype UGT91D2 of 100% (FIG. **14**B). Meanwhile, in Mutant 2, the relative GlcT activity increased to a high activity of 3.7 times based on the GlcT activity of the wildtype UGT91D2L #16 of 100% (FIG. **14**C) and the relative RhaT activity was a low activity of 35% based on the RhaT activity of the wildtype UGT91D2L #16 of 100% (FIG. **14**D). From the foregoing results, it was found that the sugar donor selectivity mainly shifts to UDP-glucose when the amino acid at the 156th position in UGT91D2 is a Thr residue and to UDP-rhamnose when the position is a Val residue.

Evaluation in Yeast

[0323] To confirm the specificity of the mutants in yeast, the enzyme activities of Mutants 1 to 6 were then compared with UGT91D2 and UGT91D2L #16 and evaluated. The genes to be introduced and expressed in the yeast transformants were 4 genes: any one of the UGT91 genes, UGT85C2, UGT74G1, and AtRHM2.

(1) Construction of plasmid pESC-URA-UGT56M

[0324] The mutant vector for yeast expression was constructed as follows. DNA was amplified by PCR using a plasmid containing DNA encoding one of the mutants produced as described above as a template and the SrUGT91D2 set of primers (SEQ ID NO: 22, 23). The plasmids pESC-URA-UGT56M1 to pESC-URA-UGT56M6, which can coexpress UGT85C2 and one of the UGT91D2 mutants, were constructed by incorporating a DNA fragment cut with restriction enzymes into the plasmid pESC-URA-UGT-1 in a similar way to that in Example 5.

(2) Construction of plasmid pESC-HIS-UGT7

[0325] The plasmid pESC-HIS-UGT7 was obtained by cutting out UGT74G1 with the restriction enzymes NotI and PacI and ligating the UGT74G1 with the vector pESC-HIS cut with the same restriction enzymes.

[0326] pESC-URA-UGT56, pESC-URA-UGT56R, or pESC-TRP-AtRHM2 used in Example 5 was used as the vector for UGT91D2, UGT91D2L #16, or AtRHM2.

Transformation of Yeast

[0327] A combination of the plasmids pESC-TRP-AtRHM2 and pESC-HIS-UGT7 and any one of ESC-URA-UGT56M1 to pESC-URA-UGT56M6, pESC-URA-UGT56, and pESC-URA-UGT56R was introduced into the *Saccharomyces cerevisiae* strain YPH499 as a host by the lithium acetate method and transformant strains were selected in a similar way to that in Example 5. In this way, 8 transformant strains, different in UGT91D2, were obtained. It was confirmed in a similar way to Example 5 that the transgenes are expressed in the obtained transformant strains. The analysis of steviol glycosides in the culture and culture supernatant for each transformant strain was carried out in a similar way to that in Example 5, except that 2 pg steviol per 1 ml medium was added.

[0328] The results of analysis of steviol glycosides in the culture supernatant of yeast indicated a decreasing tendency of the amount of the GlcT activity product stevioside produced in the yeast in which UGT91D2-based mutants (Mutants 1, 3, 5) were expressed. For example, in the yeast in which Mutant 5 was expressed, the amount of stevioside was decreased to 37.5% with the amount of stevioside produced in the yeast in which wildtype UGT91D2 was expressed being 100%, and the activity was equivalent to the wildtype UGT91D2L #16 (FIG. **15**A). Meanwhile, the GlcT activity in any of the UGT91D2L #16-based mutants (Mutants 2, 4, 6) was higher than that in the wildtype UGT91D2L #16. The activity in Mutant 6 was increased to about 2.7 times of the wildtype UGT91D2L #16 and equivalent to that in the wildtype UGT91D2 (FIG. **15**A).

[0329] The amount of the RhaT activity product dulcoside A produced in the yeast in which Mutants 1, 3, and 5 was expressed was found to be increased to about 1.5 times or more in comparison with the amount of dulcoside A produced in the yeast in which the wildtype UGT91D2 was expressed (FIG. **15**B). Meanwhile, the RhaT activity in the UGT91D2L #16-based mutants was found to be decreased by 30% or more in any of the mutants in comparison with that in the wildtype UGT91D2L #16.

[0330] Furthermore, the product behaviors of the two steviol rhamnosides (steviol position 13-Glc-Rha(α1,2), position 19-Glc-Rha(α1,2) and steviol position 13-Glc-Rha (α1,2), position 19-H) other than dulcoside A similar to those of dulcoside A were found (FIGS. **15**C to D).

[0331] The forgoing results confirmed that the sugar in the steviol glycoside shifts to rhamnose when the 156th residue of UGT91D2 is Val and to glucose when the residue is Thr in de novo expression in yeast, similar to the results of the *Escherichia coli* recombinant protein. Furthermore, it was also indicated that the sugar shifts to rhamnose when the 233rd residue is Phe and to glucose when the residue is Ser.

INDUSTRIAL AVAILABILITY

[0332] According to the present invention, it is possible to rhamnosylate position 2 of glucose at position 13 in steviol and control the sweetness and taste quality of the steviol glycoside using the SrUGT91D2L #16 gene. Using these genes as markers, it is possible to select plant bodies different in the Reb.C content. Moreover, the present invention provides a molecule tool for producing rhamnose-containing steviol glycosides represented by Reb.C not only in plants but also in microorganisms.

Sequence Listing Free Text

[0333] In SEQ ID NO: 2, the amino acid residue X_1 represents the 42nd amino acid residue Phe, the amino acid residue X_2 represents the 156th amino acid residue Val, the amino acid residue X_3 represents the 233rd amino acid residue Phe, the amino acid residue X_4 represents the 298th amino acid residue Ala, the amino acid residue X_5 represents the 394th amino acid residue Leu, the amino acid residue X_6 represents the 439th amino acid residue Asn, and the amino acid residue X_7 represents the 471st amino acid residue Phe.

SEQUENCE LISTING

Sequence total quantity: 46

SEQ ID NO: 1 moltype = DNA length = 1458
FEATURE Location/Qualifiers
source 1..1458
mol_type = other DNA
organism = Stevia rebaudiana
CDS 1..1458
SEQUENCE: 1
atgtacaacg ttacttatca tcaaaattca aaagcaatgg ctaccagtga ctccatagtt 60
gacgaccgta agcagcttca tgttgcgacg ttcccatggc ttgctttcgg tcacatcctc 120
cctttccttc agctttcgaa attgatagct gaaaagggtc acaaagtctc gtttctttct 180
accaccagaa acattcaacg tctctcttct catatctcgc cactcataaa tgttgttcaa 240
ctcacacttc cacgtgtcca agagctgccg gaggatgcag aggcgaccac tgacgtccac 300
cctgaagata ttccatatct caagaaggct tctgatggtc ttcaaccgga ggtcacccgg 360
tttctagaac aacactctcc ggactggatt atttatgatt atactcacta ctggttgcca 420
tccatcgcgg ctagcctcgg tatctcacga gcccacttct ccgtcgtcac tccatgggcc 480
attgcttata tgggaccctc agctgacgcc atgataaatg gttcagatgg tcgaaccacg 540
gttgaggatc tcacgacacc gcccaagtgg tttccctttc cgaccaaagt atgctggcgg 600
aagcatgatc ttgcccgact ggtgccttac aaagctccgg ggatatctga tggataccgt 660
atggggctgg ttcttaaggg atctgattgt ttgcttttca aatgttacca tgagtttgga 720
actcaatggc tacctctttt ggagacacta caccaagtac cggtggttcc ggtgggatta 780
ctgccaccgg aaatacccgg agacgagaaa gatgaaacat gggtgtcaat caagaaatgg 840
ctcgatggta aacaaaaagg cagtgtggtg tacgttgcat taggaagcga ggctttggtg 900
agccaaaccg aggttgttga gttagcattg ggtctcgagc tttctgggtt gccatttgtt 960
tgggcttata gaaaaccaaa aggtcccgcg aagtcagact cggtggagtt gccagacggg 1020
ttcgtggaac gaactcgtga ccgtgggttg gtctggacga gttgggcacc tcagttacga 1080
atactgagcc atgagtcggt ttgtggtttc ttgactcatt gtggttctgg atcaattgtg 1140
gaagggctaa tgtttggtca ccctctaatc atgctaccgc tttttgggga ccaacctctg 1200
aatgctcgat tactggagga caaacaggtg ggaatcgaga taccaagaaa tgaggaagat 1260
ggttgcttga ccaaggagtc ggttgctaga tcactgaggt ccgttgttgt ggaaaacgaa 1320
ggggagatct acaaggcgaa cgcgagggag ctgagtaaaa tctataacga cactaaggtg 1380
gaaaaagaat atgtaagcca attcgtagac tttttggaaa agaatgcgcg tgcggttgcc 1440
atcgatcatg agagttaa 1458

SEQ ID NO: 2 moltype = AA length = 485
FEATURE Location/Qualifiers
source 1..485
mol_type = protein
organism = Stevia rebaudiana
SEQUENCE: 2
MYNVTYHQNS KAMATSDSIV DDRKQLHVAT FPWLAFGHIL PFLQLSKLIA EKGHKVSFLS 60
TTRNIQRLSS HISPLINVVQ LTLPRVQELP EDAEATTDVH PEDIPYLKKA SDGLQPEVTR 120
FLEQHSPDWI IYDYTHYWLP SIAASLGISR AHFSVVTPWA IAYMGPSADA MINGSDGRTT 180
VEDLTTPPKW FPFPTKVCWR KHDLARLVPY KAPGISDGYR MGLVLKGSDC LLFKCYHEFG 240
TQWLPLLETL HQVPVVPVGL LPPEIPGDEK DETWVSIKKW LDGKQKGSVV YVALGSEALV 300
SQTEVVELAL GLELSGLPFV WAYRKPKGPA KSDSVELPDG FVERTRDRGL VWTSWAPQLR 360
ILSHESVCGF LTHCGSGSIV EGLMPGHPLI MLPLFGDQPL NARLLEDKQV GIEIPRNEED 420
GCLTKESVAR SLRSVVVENE GEIYKANARE LSKIYNDTKV EKEYVSQFVD FLEKNARAVA 480
IDHES 485

SEQ ID NO: 3 moltype = DNA length = 1458
FEATURE Location/Qualifiers
source 1..1458
mol_type = other DNA
organism = Stevia rebaudiana
CDS 1..1458
SEQUENCE: 3
atgtacaacg ttacttatca tcaaaattca aaagcaatgg ctaccagtga ctccatagtt 60
gacgaccgta agcagcttca tgttgcgacg ttcccatggc ttgctttcgg tcacatcctc 120
ccttaccttc agctttcgaa attgatagct gaaaagggtc acaaagtctc gtttctttct 180
accaccagaa acattcaacg tctctcttct catatctcgc cactcataaa tgttgttcaa 240
ctcacacttc cacgtgtcca agagctgccg gaggatgcag aggcgaccac tgacgtccac 300
cctgaagata ttccatatct caagaaggct tctgatggtc ttcaaccgga ggtcacccgg 360
tttctagaac aacactctcc ggactggatt atttatgatt atactcacta ctggttgcca 420
tccatcgcgg ctagcctcgg tatctcacga gcccacttct ccgtcaccac tccatgggcc 480
attgcttata tgggaccctc agctgacgcc atgataaatg gttcagatgg tcgaaccacg 540
gttgaggatc tcacgacacc gcccaagtgg tttccctttc cgaccaaagt atgctggcgg 600
aagcatgatc ttgcccgact ggtgccttac aaagctccgg ggatatctga tggataccgt 660
atggggctgg ttcttaaggg atctgattgt ttgctttcca aatgttacca tgagtttgga 720
actcaatggc tacctctttt ggagacacta caccaagtac cggtggttcc ggtgggatta 780
ctgccaccgg aaatacccgg agacgagaaa gatgaaacat gggtgtcaat caagaaatgg 840
ctcgatggta aacaaaaagg cagtgtggtg tacgttgcat taggaagcga ggttttggtg 900
agccaaaccg aggttgttga gttagcattg ggtctcgagc tttctgggtt gccatttgtt 960
tgggcttata gaaaaccaaa aggtcccgcg aagtcagact cggtggagtt gccagacggg 1020
ttcgtggaac gaactcgtga ccgtgggttg gtctggacga gttgggcacc tcagttacga 1080

-continued

```
atactgagcc atgagtcggt ttgtggtttc ttgactcatt gtggttctgg atcaattgtg  1140
gaagggctaa tgtttggtca ccctctaatc atgctaccga tttttgggga ccaacctctg  1200
aatgctcgat tactggagga caaacaggtg ggaatcgaga taccaagaaa tgaggaagat  1260
ggttgcttga ccaaggagtc ggttgctaga tcactgaggt ccgttgttgt ggaaaaagaa  1320
ggggagatct acaaggcgaa cgcgagggag ctgagtaaaa tctataacga cactaaggtt  1380
gaaaaagaat atgtaagcca attcgtagac tatttggaaa agaatgcgcg tgcggttgcc  1440
atcgatcatg agagttaa                                               1458

SEQ ID NO: 4           moltype = AA  length = 485
FEATURE                Location/Qualifiers
source                 1..485
                       mol_type = protein
                       organism = Stevia rebaudiana
SEQUENCE: 4
MYNVTYHQNS KAMATSDSIV DDRKQLHVAT FPWLAFGHIL PYLQLSKLIA EKGHKVSFLS  60
TTRNIQRLSS HISPLINVVQ LTLPRVQELP EDAEATTDVH PEDIPYLKKA SDGLQPEVTR  120
FLEQHSPDWI IYDYTHYWLP SIAASLGISR AHFSVTTPWA IAYMGPSADA MINGSDGRTT  180
VEDLTTPPKW FPFPTKVCWR KHDLARLVPY KAPGISDGYR MGLVLKGSDC LLSKCYHEFG  240
TQWLPLLETL HQVPVVPVGL LPPEIPGDEK DETWVSIKKW LDGKQKGSVV YVALGSEVLV  300
SQTEVVELAL GLELSGLPFV WAYRKPKGPA KSDSVELPDG FVERTRDRGL VWTSWAPQLR  360
ILSHESVCGF LTHCGSGSIV EGLMFGHPLI MLPIFGDQPL NARLLEDKQV GIEIPRNEED  420
GCLTKESVAR SLRSVVVEKE GEIYKANARE LSKIYNDTKV EKEYVSQFVD YLEKNARAVA  480
IDHES                                                             485

SEQ ID NO: 5           moltype = DNA  length = 39
FEATURE                Location/Qualifiers
misc_feature           1..39
                       note = Artificial Nucleic Acid
source                 1..39
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 5
tgccgcgcgg cagccatatg tacaacgtta cttatcatc                        39

SEQ ID NO: 6           moltype = DNA  length = 39
FEATURE                Location/Qualifiers
misc_feature           1..39
                       note = Artificial Nucleic Acid
source                 1..39
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 6
gttagcagcc ggatccttaa ctctcatgat cgatggcaa                        39

SEQ ID NO: 7           moltype = DNA  length = 1446
FEATURE                Location/Qualifiers
source                 1..1446
                       mol_type = other DNA
                       organism = Stevia rebaudiana
CDS                    1..1446
SEQUENCE: 7
atggatgcaa tggctacaac tgagaagaaa ccacacgtca tcttcatacc atttccagca  60
caaagccaca ttaaagccat gctcaaacta gcacaacttc tccaccacaa aggactccag  120
ataaccttcg tcaacaccga cttcatccac aaccagtttc ttgaatcatc gggcccacat  180
tgtttggacg gttcaccggg tttccggttc gaaaccatcc cggatggtgt ttctcacagt  240
ccggaagcga gcatcccaat cagagaatca ctcttgagat ccattgaaac caacttcttg  300
gatcgtttca ttgatcttgt aaccaaactt ccggatcctc cgacttgtat tatctcagat  360
gggttcttgt cggttttcac aattgacgct gcaaaaaagc ttggaattcc ggtcatgatg  420
tattggacac ttgctgcctg tgggttcatg ggtttttacc atattcattc tctcattgag  480
aaaggatttg caccacttaa agatgcaagt tacttgacaa atgggtattt ggacaccgtc  540
attgattggg ttccgggaat ggaaggcatc cgtctcaagg atttcccgct ggactggagc  600
actgacctca atgacaaagt tttgatgttc actacagaag ctcctcaaag gtcacacaag  660
gtttcacatc atattttcca cacgttcgat gagttggagc ctagtattat aaaaactttg  720
tcattgaggt ataatcacat ttacaccatc ggcccactgc aattacttct tgatcaaata  780
cccgaagaga aaaagcaaac tggaattacg agtctccatg gatacagttt agtaaaagaa  840
gaaccagagt gtttccagtg gcttcagtct aaagaaccaa attccgtcgt ttatgtaaat  900
tttggaagta ctacagtaat gtctttagaa gacatgacgg aatttggttg gggacttgct  960
aatagcaacc attatttcct ttggatcatc cgatcaaact tggtgatagg ggaaaatgca  1020
gttttgcccc ctgaacttga ggaacatata aagaaaagag gctttattgc tagctggtgt  1080
tcacaagaaa aggtcttgaa gcacccttcg gttggagggt tcttgactca ttgtgggtgg  1140
ggatcgacca tcgagagctt gtctgctggg gtgccaatga tatgctggcc ttattcgtgg  1200
gaccagctga ccaactgtag gtatatatgc aaagaatggg aggttgggct cgagatggga  1260
accaaagtga aacgagatga agtcaagagg cttgtacaag agttgatggg agaaggaggt  1320
cacaaaatga ggaacaaggc taaagattgg aaagaaaagg ctcgcattgc aatagctcct  1380
aacggttcat cttctttgaa catagacaaa atggtcaagg aaatcaccgt gctagcaaga  1440
aactag                                                            1446
```

-continued

```
SEQ ID NO: 8           moltype = AA  length = 481
FEATURE                Location/Qualifiers
source                 1..481
                       mol_type = protein
                       organism = Stevia rebaudiana
SEQUENCE: 8
MDAMATTEKK PHVIFIPFPA QSHIKAMLKL AQLLHHKGLQ ITFVNTDFIH NQFLESSGPH  60
CLDGSPGFRF ETIPDGVSHS PEASIPIRES LLRSIETNFL DRFIDLVTKL PDPPTCIISD  120
GFLSVFTIDA AKKLGIPVMM YWTLAACGFM GFYHIHSLIE KGFAPLKDAS YLTNGYLDTV  180
IDWVPGMEGI RLKDFPLDWS TDLNDKVLMF TTEAPQRSHK VSHHIFHTFD ELEPSIIKTL  240
SLRYNHIYTI GPLQLLLDQI PEEKKQTGIT SLHGYSLVKE EPECFQWLQS KEPNSVVYVN  300
FGSTTVMSLE DMTEFGWGLA NSNHYFLWII RSNLVIGENA VLPPELEEHI KKRGFIASWC  360
SQEKVLKHPS VGGFLTHCGW GSTIESLSAG VPMICWPYSW DQLTNCRYIC KEWEVGLEMG  420
TKVKRDEVKR LVQELMGEGG HKMRNKAKDW KEKARIAIAP NGSSSLNIDK MVKEITVLAR  480
N                                                                  481

SEQ ID NO: 9           moltype = DNA  length = 1383
FEATURE                Location/Qualifiers
source                 1..1383
                       mol_type = other DNA
                       organism = Stevia rebaudiana
CDS                    1..1383
SEQUENCE: 9
atggcggaac aacaaaagat caagaaatca ccacacgttc tactcatccc attcccttta  60
caaggccata taaacccttt catccagttt ggcaaacgat taatctccaa aggtgtcaaa  120
acaacacttg ttaccaccat ccacacctta aactcaaccc taaaccacag taacaccacc  180
accacctcca tcgaaatcca agcaatttcc gatggttgtg atgaaggcgg ttttatgagt  240
gcaggagaat catatttgga aacattcaaa caagttgggt ctaaatcact agctgactta  300
atcaagaagc ttcaaagtga aggaaccaca attgatgcaa tcatttatga ttctatgact  360
gaatgggttt tagatgttgc aattgagttt ggaatcgatg gtggttcgtt tttcactcaa  420
gcttgtgttg taaacagctt atattatcat gttcataagg gtttgatttc tttgccattg  480
ggtgaaactg tttcggttcc tggatttcca gagcttcaac ggtgggagac accgttaatt  540
ttgcagaatc atgagcaaat acagagccct tggtctcaga tgttgtttgg tcagtttgct  600
aatattgatc aagcacgttg ggtcttcaca aatagttttt acaagctcga ggaagaggta  660
atagagtgga cgagaaagat atggaacttg aaggtaatcg ggccaacact tccatccatg  720
taccttgaca aacgacttga tgatgataaa gataacggat ttaatctcta caaagcaaac  780
catcatgagt gcatgaactg gttagacgat aagccaaagg aatcagttgt ttacgtagca  840
tttggtagcc tggtgaaaca tggacccgaa caagtggaag aaatcacacg ggctttaata  900
gatagtgatg tcaacttctt gtgggttatc aaacataaag aagagggaaa gctcccagaa  960
aatctttcgg aagtaataaa aaccggaaag ggtttgattg tagcatggtg caaacaattg  1020
gatgtgttag cacacgaatc agtaggatgc tttgttacac attgtgggtt caactcaact  1080
cttgaagcaa taagtcttgg agtccccgtt gttgcaatgc ctcaattttc ggatcaaact  1140
acaaatgcca agcttctaga tgaaattttg ggtgttggag ttagagttaa ggctgatgag  1200
aatgggatag tgagaagagg aaatcttgcg tcatgtatta agatgattat ggaggaggaa  1260
agaggagtaa taatccgaaa gaatgcggta aaatggaagg atttggctaa agtagccgtt  1320
catgaaggtg gtagctcaga caatgatatt gtcgaatttg taagtgagct aattaaggct  1380
taa                                                                1383

SEQ ID NO: 10          moltype = AA  length = 460
FEATURE                Location/Qualifiers
source                 1..460
                       mol_type = protein
                       organism = Stevia rebaudiana
SEQUENCE: 10
MAEQQKIKKS PHVLLIPFPL QGHINPFIQF GKRLISKGVK TTLVTTIHTL NSTLNHSNTT  60
TTSIEIQAIS DGCDEGGFMS AGESYLETFK QVGSKSLADL IKKLQSEGTT IDAIIYDSMT  120
EWVLDVAIEF GIDGGSFFTQ ACVVNSLYYH VHKGLISLPL GETVSVPGFP ELQRWETPLI  180
LQNHEQIQSP WSQMLFGQFA NIDQARWVFT NSFYKLEEEV IEWTRKIWNL KVIGPTLPSM  240
YLDKRLDDDK DNGFNLYKAN HHECMNWLDD KPKESVVYVA FGSLVKHGPE QVEEITRALI  300
DSDVNFLWVI KHKEEGKLPE NLSEVIKTGK GLIVAWCKQL DVLAHESVGC FVTHCGFNST  360
LEAISLGVPV VAMPQFSDQT TNAKLLDEIL GVGVRVKADE NGIVRRGNLA SCIKMIMEEE  420
RGVIIRKNAV KWKDLAKVAV HEGGSSDNDI VEFVSELIKA                        460

SEQ ID NO: 11          moltype = DNA  length = 1377
FEATURE                Location/Qualifiers
source                 1..1377
                       mol_type = other DNA
                       organism = Stevia rebaudiana
CDS                    1..1377
SEQUENCE: 11
atggaaaata aaacggagac caccgttcgc cggcgccgga gaataatatt attcccggta  60
ccatttcaag gccacattaa cccaattctt cagctagcca atgtgttgta ctctaaagga  120
ttcagtatca ccatctttca caccaacttc aacaaaccca aaacatctaa ttaccctcac  180
ttcactttca gattcatcct cgacaacgac ccacaagacg aacgcatttc caatctaccg  240
actcatggtc cgctcgctgg tatgcggatt ccgattatca acgaacacgg agctgacgaa  300
```

-continued

```
ttacgacgcg aactggaact gttgatgtta gcttctgaag aagatgaaga ggtatcgtgt  360
ttaatcacgg atgctctttg gtacttcgcg caatctgttg ctgacagtct taacctccga  420
ccgcttgttt tgatgacaag cagcttgttt aattttcatg cacatgtttc acttcctcag  480
tttgatgagc ttggttacct cgatcctgat gacaaaaccc gtttggaaga acaagcgagt  540
gggtttccta tgctaaaagt gaaagacatc aagtctgcgt attcgaactg gcaaatactc  600
aaagagatat tagggaagat gataaaacaa acaaaagcat cttcaggagt catctggaac  660
tcatttaagg aactcgaaga gtctgagctc gaaactgtta tccgtgagat cccggctcca  720
agtttcttga taccactccc caagcatttg acagcctctt ccagcagctt actagaccac  780
gatcgaaccg tttttcaatg gttagaccaa caaccgccaa gttcggtact gtatgttagt  840
tttggtagta ctagtgaagt ggatgagaaa gatttcttgg aaatagctcg tgggttggtt  900
gatagcaagc agtcgttttt atgggtggtt cgacctgggt ttgtcaaggg ttcgacgtgg  960
gtcgaaccgt tgccagatgg gttcttgggt gaaagaggac gtattgtgaa atgggttcca 1020
cagcaagaag tgctagctca tggagcaata ggcgcattct ggactcatag cggatggaac 1080
tctacgttgg aaagcgtttg tgaaggtgtt cctatgattt tctcggattt tgggctcgat 1140
caaccgttga atgctagata catgagtgat gttttgaagg taggggtgta tttggaaaat 1200
gggtgggaaa gaggagagat agcaaatgca ataagaagag ttatggtgga tgaagaagga 1260
gaatacatta gacagaatgc aagagttttg aaacaaaagg cagatgtttc tttgatgaag 1320
ggtggttcgt cttacgaatc attagagtct ctagtttctt acatttcatc gttgtaa     1377

SEQ ID NO: 12          moltype = AA  length = 458
FEATURE                Location/Qualifiers
source                 1..458
                       mol_type = protein
                       organism = Stevia rebaudiana
SEQUENCE: 12
MENKTETTVR RRRRIILFPV PFQGHINPIL QLANVLYSKG FSITIFHTNF NKPKTSNYPH  60
FTFRFILDND PQDERISNLP THGPLAGMRI PIINEHGADE LRRELELLML ASEEDEEVSC  120
LITDALWYFA QSVADSLNLR PLVLMTSSLF NFHAHVSLPQ FDELGYLDPD DKTRLEEQAS  180
GFPMLKVKDI KSAYSNWQIL KEILGKMIKQ TKASSGVIWN SFKELEESEL ETVIREIPAP  240
SFLIPLPKHL TASSSSLLDH DRTVFQWLDQ QPPSSVLYVS FGSTSEVDEK DFLEIARGLV  300
DSKQSFLWVV RPGFVKGSTW VEPLPDGFLG ERGRIVKWVP QQEVLAHGAI GAFWTHSGWN  360
STLESVCEGV PMIFSDFGLD QPLNARYMSD VLKVGVYLEN GWERGEIANA IRRVMVDEEG  420
EYIRQNARVL KQKADVSLMK GGSSYESLES LVSYISSL                          458

SEQ ID NO: 13          moltype = DNA  length = 2004
FEATURE                Location/Qualifiers
source                 1..2004
                       mol_type = other DNA
                       organism = Arabidopsis thaliana
SEQUENCE: 13
atggatgata ctacgtataa gccaaagaac attctcatta ctggagctgc tggatttatt  60
gcttctcatg ttgccaacag attaatccgt aactatcctg attacaagat cgttgttctt  120
gacaagcttg attactgttc agatctgaag aatcttgatc cttctttttc ttcaccaaat  180
ttcaagtttg tcaaaggaga tatcgcgagt gatgatctcg ttaactacct tctcatcact  240
gaaaacattg atacgataat gcattttgct gctcaaactc atgttgataa ctcttttggt  300
aatagctttg agtttaccaa gaacaatatt tatggtactc atgttctttt ggaagcctgt  360
aaagttacag gacagatcag gaggtttatc catgtgagta ccgatgaagt ctatggagaa  420
accgatgagg atgctgctgt aggaaaccat gaagcttctc agctgttacc gacgaatcct  480
tactctgcaa ctaaggctgg tgctgagatg cttgtgatgg cttatggtag atcatatgga  540
ttgcctgtta ttacgactcg cgggaacaat gtttatgggc ctaaccagtt tcctgaaaaa  600
atgattccta agttcatctt gttggctatg agtgggaagc cgcttcccat ccatggagat  660
ggatctaatg tccggagtta cttgtactgc gaagacgttg ctgaggcttt tgaggttgtt  720
cttcacaaag gagaaatcgg tcatgtctac aatgtcggca caaaaagaga aaggagagtg  780
atcgatgtgg ctagagacat ctgcaaactt ttcgggaaag accctgagtc aagcattcag  840
tttgtggaga accggccctt taatgatcaa aggtacttcc ttgatgatca gaagctgaag  900
aaattggggt ggcaagagcg aacaaattgg gaagatggat tgaagaagac aatggactgg  960
tacactcaga atcctgagtg gtggggtgat gtttctggag ctttgcttcc tcatccgaga 1020
atgcttatga tgcccggtgg aagactttct gatggatcta gtgagaagaa agacgtttca 1080
agcaacacgg tccagacatt tacggttgta acacctaaga atggtgattc tggtgacaaa 1140
gcttcgttga agtttttgat ctatggtaag actggttggc ttggtggtct tctagggaaa 1200
ctatgtgaga agcaagggat tacatatgag tatgggaaag gacgtctgga ggatagagct 1260
tctcttgtgg cggatattcg tagcatcaaa cctactcatg tgtttaatgc tgctggttta 1320
actggcagac ccaacgttga ctggtgtgaa tctcacaaac cagagaccat tcgtgtaaat 1380
gtcgcaggta ctttgactct agctgatgtt tgcagagaga atgatctctt gatgatgaac 1440
ttcgccaccg gttgcatctt tgagtatgac gctacacatc ctgagggttc gggtataggt 1500
ttcaaggaag aagacaagcc aaatttcttt ggttctttct actcgaaaac caaagccatg 1560
gttgaggagc tcttgagaga atttgacaat gtatgtacct tgagagtccg gatgccaatc 1620
tcctcagacc taaacaaccc gagaaacttc atcacgaaga tctcgcgcta caacaaagtg 1680
gtggacatcc cgaacagcat gaccgtacta gacgagcttc tcccaatctc tatcgagatg 1740
gcgaagagaa acctaagagg catatggaat ttcaccaacc caggggtggt gagccacaac 1800
gagatattgg agatgtacaa gaattacatc gagccaggtt ttaaatggtc caacttcaca 1860
gtggaagaac aagcaaaggt cattgttgct gctcgaagca acaacgaaat ggatggatct 1920
aaactaagca aggagttccc agagatgctc tccatcaaag agtcactgct caaatacgtc 1980
tttgaaccaa acaagagaac ctaa                                        2004

SEQ ID NO: 14          moltype = DNA  length = 33
```

-continued

```
FEATURE                 Location/Qualifiers
misc_feature            1..33
                        note = Artificial Nucleic Acid
source                  1..33
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 14
cacccatatg gatgcaatgg ctacaactga gaa                                33

SEQ ID NO: 15           moltype = DNA  length = 32
FEATURE                 Location/Qualifiers
misc_feature            1..32
                        note = Artificial Nucleic Acid
source                  1..32
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 15
agatctctag tttcttgcta gcacggtgat tt                                 32

SEQ ID NO: 16           moltype = DNA  length = 35
FEATURE                 Location/Qualifiers
misc_feature            1..35
                        note = Artificial Nucleic Acid
source                  1..35
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 16
cacccatatg gcggaacaac aaaagatcaa gaaat                               35

SEQ ID NO: 17           moltype = DNA  length = 35
FEATURE                 Location/Qualifiers
misc_feature            1..35
                        note = Artificial Nucleic Acid
source                  1..35
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 17
ggatccttaa gccttaatta gctcacttac aaatt                               35

SEQ ID NO: 18           moltype = DNA  length = 29
FEATURE                 Location/Qualifiers
misc_feature            1..29
                        note = Artificial Nucleic Acid
source                  1..29
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 18
cacccatatg gaaaataaaa cggagacca                                     29

SEQ ID NO: 19           moltype = DNA  length = 32
FEATURE                 Location/Qualifiers
misc_feature            1..32
                        note = Artificial Nucleic Acid
source                  1..32
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 19
ggatccttac aacgatgaaa tgtaagaaac ta                                 32

SEQ ID NO: 20           moltype = DNA  length = 33
FEATURE                 Location/Qualifiers
misc_feature            1..33
                        note = Artificial Nucleic Acid
source                  1..33
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 20
acagatctat ggatgcaatg gctacaactg aga                                33

SEQ ID NO: 21           moltype = DNA  length = 34
FEATURE                 Location/Qualifiers
misc_feature            1..34
                        note = Artificial Nucleic Acid
source                  1..34
                        mol_type = other DNA
                        organism = synthetic construct
```

-continued

SEQUENCE: 21
tagtcgacta gtttcttgct agcacggtga tttc 34

SEQ ID NO: 22 moltype = DNA length = 42
FEATURE Location/Qualifiers
misc_feature 1..42
note = Artificial Nucleic Acid
source 1..42
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 22
aagcggccgc atgtacaacg ttacttatca tcaaaattca aa 42

SEQ ID NO: 23 moltype = DNA length = 31
FEATURE Location/Qualifiers
misc_feature 1..31
note = Artificial Nucleic Acid
source 1..31
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 23
cgttaattaa ctctcatgat cgatggcaac c 31

SEQ ID NO: 24 moltype = DNA length = 34
FEATURE Location/Qualifiers
misc_feature 1..34
note = Artificial Nucleic Acid
source 1..34
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 24
aagcggccgc atggcggaac aacaaaagat caag 34

SEQ ID NO: 25 moltype = DNA length = 37
FEATURE Location/Qualifiers
misc_feature 1..37
note = Artificial Nucleic Acid
source 1..37
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 25
cgttaattaa gccttaatta gctcacttac aaattcg 37

SEQ ID NO: 26 moltype = DNA length = 33
FEATURE Location/Qualifiers
misc_feature 1..33
note = Artificial Nucleic Acid
source 1..33
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 26
aaggatccat ggaaaataaa acggagacca ccg 33

SEQ ID NO: 27 moltype = DNA length = 44
FEATURE Location/Qualifiers
misc_feature 1..44
note = Artificial Nucleic Acid
source 1..44
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 27
gcgtcgactt acaacgatga aatgtaagaa actagagact ctaa 44

SEQ ID NO: 28 moltype = DNA length = 33
FEATURE Location/Qualifiers
misc_feature 1..33
note = Artificial Nucleic Acid
source 1..33
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 28
ggatccatgg atgatactac gtataagcca aag 33

SEQ ID NO: 29 moltype = DNA length = 32
FEATURE Location/Qualifiers
misc_feature 1..32

-continued

```
                     note = Artificial Nucleic Acid
source               1..32
                     mol_type = other DNA
                     organism = synthetic construct
SEQUENCE: 29
ctcgagttag gttctcttgt ttggttcaaa ga                                  32

SEQ ID NO: 30        moltype = DNA  length = 22
FEATURE              Location/Qualifiers
misc_feature         1..22
                     note = Artificial Nucleic Acid
source               1..22
                     mol_type = other DNA
                     organism = synthetic construct
SEQUENCE: 30
caagtcccca accaaattcc gt                                             22

SEQ ID NO: 31        moltype = DNA  length = 21
FEATURE              Location/Qualifiers
misc_feature         1..21
                     note = Artificial Nucleic Acid
source               1..21
                     mol_type = other DNA
                     organism = synthetic construct
SEQUENCE: 31
cacgaacccg tctggcaact c                                              21

SEQ ID NO: 32        moltype = DNA  length = 25
FEATURE              Location/Qualifiers
misc_feature         1..25
                     note = Artificial Nucleic Acid
source               1..25
                     mol_type = other DNA
                     organism = synthetic construct
SEQUENCE: 32
cccgtgtgat ttcttccact tgttc                                          25

SEQ ID NO: 33        moltype = DNA  length = 21
FEATURE              Location/Qualifiers
misc_feature         1..21
                     note = Artificial Nucleic Acid
source               1..21
                     mol_type = other DNA
                     organism = synthetic construct
SEQUENCE: 33
caagaaccca tctggcaacg g                                              21

SEQ ID NO: 34        moltype = DNA  length = 23
FEATURE              Location/Qualifiers
misc_feature         1..23
                     note = Artificial Nucleic Acid
source               1..23
                     mol_type = other DNA
                     organism = synthetic construct
SEQUENCE: 34
gctttgtcac cagaatcacc att                                            23

SEQ ID NO: 35        moltype = DNA  length = 28
FEATURE              Location/Qualifiers
misc_feature         1..28
                     note = Artificial Nucleic Acid
source               1..28
                     mol_type = other DNA
                     organism = synthetic construct
SEQUENCE: 35
gattattaaa cttctttgcg tccatcca                                       28

SEQ ID NO: 36        moltype = DNA  length = 32
FEATURE              Location/Qualifiers
misc_feature         1..32
                     note = Artificial Nucleic Acid
source               1..32
                     mol_type = other DNA
                     organism = synthetic construct
SEQUENCE: 36
cctctatact ttaacgtcaa ggagaaaaaa cc                                  32
```

-continued

```
SEQ ID NO: 37          moltype = DNA  length = 39
FEATURE                Location/Qualifiers
misc_feature           1..39
                       note = Artificial Nucleic Acid
source                 1..39
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 37
tgccgcgcgg cagccatatg tacaacgtta cttatcatc                          39

SEQ ID NO: 38          moltype = DNA  length = 39
FEATURE                Location/Qualifiers
misc_feature           1..39
                       note = Artificial Nucleic Acid
source                 1..39
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 38
gttagcagcc ggatccttaa ctctcatgat cgatggcaa                          39

SEQ ID NO: 39          moltype = DNA  length = 23
FEATURE                Location/Qualifiers
misc_feature           1..23
                       note = Artificial Nucleic Acid
source                 1..23
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 39
cacttctccg tcgtcactcc atg                                           23

SEQ ID NO: 40          moltype = DNA  length = 23
FEATURE                Location/Qualifiers
misc_feature           1..23
                       note = Artificial Nucleic Acid
source                 1..23
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 40
catggagtga cgacggagaa gtg                                           23

SEQ ID NO: 41          moltype = DNA  length = 23
FEATURE                Location/Qualifiers
misc_feature           1..23
                       note = Artificial Nucleic Acid
source                 1..23
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 41
cacttctccg tcaccactcc atg                                           23

SEQ ID NO: 42          moltype = DNA  length = 23
FEATURE                Location/Qualifiers
misc_feature           1..23
                       note = Artificial Nucleic Acid
source                 1..23
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 42
catggagtgg tgacggagaa gtg                                           23

SEQ ID NO: 43          moltype = DNA  length = 32
FEATURE                Location/Qualifiers
misc_feature           1..32
                       note = Artificial Nucleic Acid
source                 1..32
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 43
ctgattgttt gcttttcaaa tgttaccatg ag                                 32

SEQ ID NO: 44          moltype = DNA  length = 32
FEATURE                Location/Qualifiers
misc_feature           1..32
                       note = Artificial Nucleic Acid
source                 1..32
```

-continued

	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 44	
ctcatggtaa catttgaaaa gcaaacaatc ag	32
SEQ ID NO: 45	moltype = DNA length = 32
FEATURE	Location/Qualifiers
misc_feature	1..32
	note = Artificial Nucleic Acid
source	1..32
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 45	
ctgattgttt gctttccaaa tgttaccatg ag	32
SEQ ID NO: 46	moltype = DNA length = 32
FEATURE	Location/Qualifiers
misc_feature	1..32
	note = Artificial Nucleic Acid
source	1..32
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 46	
ctcatggtaa catttggaaa gcaaacaatc ag	32

What is claimed is:

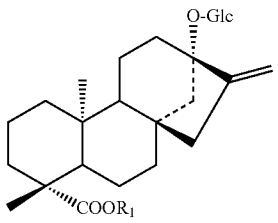
1. A protein according to any one selected from the group consisting of the following (a) to (c):

(a) a protein consisting of the amino acid sequence of SEQ ID NO: 2;

(b) a protein consisting of an amino acid sequence wherein 1 to 48 amino acids other than the amino acid residue X_7 are deleted, substituted, inserted, and/or added in the amino acid sequence of SEQ ID NO: 2 and having an activity to add hexose at position 2 of glucose at position 13 in a compound represented by the following formula (I);

(c) a protein having an amino acid sequence having a sequence identity of 90% or more to the amino acid sequence of SEQ ID NO: 2 wherein an amino acid corresponding to the amino acid residue X_7 is the same and having an activity to add hexose at position 2 of glucose at position 13 in a compound represented by the following formula (I);

(I)



wherein R_1 represents H, a C_1-C_{20} alkyl group, a C_2-C_{20} alkenyl group, a C_2-C_{20} alkynyl group, a C_4-C_{20} alkyldienyl group, a C_6-C_{18} aryl group, a C_6-C_{20} alkylaryl group, a C_6-C_{20} arylalkyl group, a C_4-C_{20} cycloalkyl group, a C_4-C_{20} cycloalkenyl group, a (C_3-C_{10} cycloalkyl)C_1-C_{10} alkyl group or a sugar residue.

2. The protein according to claim **1**, wherein the protein is

(b') a protein consisting of an amino acid sequence wherein 1 to 48 amino acids other than the amino acid residues X_1 to X_7 are deleted, substituted, inserted, and/or added in the amino acid sequence of SEQ ID NO: 2 and having an activity to add hexose at position 2 of glucose at position 13 in a compound represented by the formula (I); or

(c') a protein having an amino acid sequence having a sequence identity of 90% or more to the amino acid sequence of SEQ ID NO: 2 wherein amino acids corresponding to the amino acid residues X_1 to X_7 are the same and having an activity to add hexose at position 2 of glucose at position 13 in a compound represented by the formula (I).

3. The protein according to claim **1**, wherein the hexose is selected from the group consisting of glucose or rhamnose.

4. The protein according to claim **1**, wherein the R_1 is H or a sugar residue of a glucose monomer or a glucose dimer.

5. The protein according to claim **1**, wherein the compound is steviolmonoside or rubusoside.

6. A polynucleotide selected from the group consisting of the following (a) to (d):

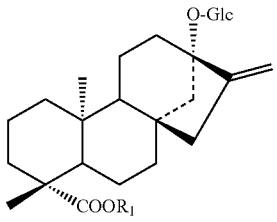
(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO: 1;

(b) a polynucleotide encoding a protein consisting of the amino acid sequences of SEQ ID NO: 2;

(c) a polynucleotide encoding a protein consisting of an amino acid sequence wherein 1 to 48 amino acids other than the amino acid residue X_7 are deleted, substituted, inserted, and/or added in the amino acid sequence of SEQ ID NO: 2 and having an activity to add hexose at position 2 of glucose at position 13 in a compound represented by the following formula (I);

(d) a polynucleotide encoding a protein having an amino acid sequence having a sequence identity of 90% or more to the amino acid sequence of SEQ ID NO: 2 wherein an amino acid corresponding to the amino acid residue X_7 is the same and having an activity to add hexose at position 2 of glucose at position 13 in a compound represented by formula (I):

(I)



wherein R_1 represents H, a C_1-C_{20} alkyl group, a C_2-C_{20} alkenyl group, a C_2-C_{20} alkynyl group, a C_4-C_{20} alkyldienyl group, a C_6-C_{18} aryl group, a C_6-C_{20} alkylaryl group, a C_6-C_{20} arylalkyl group, a C_4-C_{20} cycloalkyl group, a C_4-C_{20} cycloalkenyl group, a (C_3-C_{10} cycloalkyl)C_1-C_{10} alkyl group or a sugar residue.

7. The polynucleotide according to claim **6**, wherein the polynucleotide is

(c') a polynucleotide encoding a protein consisting of an amino acid sequence wherein 1 to 48 amino acids other than the amino acid residues X_1 to X_7 are deleted, substituted, inserted, and/or added in the amino acid sequence of SEQ ID NO: 2 and having an activity to add hexose at position 2 of glucose at position 13 in a compound represented by the formula (I); or

(d') a polynucleotide encoding a protein having an amino acid sequence having a sequence identity of 90% or more to the amino acid sequence of SEQ ID NO: 2 wherein amino acids corresponding to the amino acid residues X_1 to X_7 are the same and having an activity to add hexose at position 2 of glucose at position 13 in a compound represented by the formula (I).

8. The polynucleotide according to claim **6**, wherein the hexose is selected from the group consisting of glucose or rhamnose.

9. The polynucleotide according to claim **6**, wherein the R_1 is H or a sugar residue of a glucose monomer or a glucose dimer.

10. The polynucleotide according to claim **6**, wherein the compound is steviolmonoside or rubusoside.

11. A non-human transformant in which the polynucleotide according to claim **6** is introduced.

12. The transformant according to claim **11**, wherein the polynucleotide is inserted into an expression vector.

13. The transformant according to claim **11**, wherein the transformant is a plant.

14. An extract of the transformant according to claim **11**.

15. A food or drink, a pharmaceutical product, or an industrial raw material comprising the extract according to claim **14**.

* * * * *