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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
19 October 2017 (19.10.2017)

WIPO | PCT

(10) International Publication Number
WO 2017/178632 A1

(51) **International Patent Classification**:
A23L 27/30 (2016.01) *C12N 15/63* (2006.01)
A23L 2/60 (2006.01)

(21) **International Application Number**:
PCT/EP2017/059028

(22) **International Filing Date**:
13 April 2017 (13.04.2017)

(25) **Filing Language**: English

(26) **Publication Language**: English

(30) **Priority Data**:
62/321,850 13 April 2016 (13.04.2016) US

(71) **Applicant**: **EVOLVA SA** [CH/CH]; Duggingerstrasse 23, CH-4153 Reinach (CH).

(72) **Inventors**: **DOUCHIN, Veronique**; Aurikelvej 2-4th, 2000 Frederiksberg (DK). **HALLWYL, Swee Chuang Lim**; Amalieparken 87, 4-4, 2665 Vallensbaek Strand (DK). **OLSSON, Kim**; Stenhuggervej 1, 1th, 2400 Copenhagen (DK).

(74) **Agent**: **REES, Kerry**; WP Thompson, 138 Fetter Lane, London EC4A 1BT (GB).

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

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

Published:

— *with international search report (Art. 21(3))*

— *with sequence listing part of description (Rule 5.2(a))*

WO 2017/178632 A1

(54) **Title**: PRODUCTION OF STEVIOL GLYCOSIDES IN RECOMBINANT HOSTS

(57) **Abstract**: The invention relates to recombinant microorganisms and methods for producing steviol glycosides and steviol glycoside precursors.

PRODUCTION OF STEVIOL GLYCOSIDES IN RECOMBINANT HOSTS

BACKGROUND OF THE INVENTION

Field of the Invention

[0001] This disclosure relates to recombinant production of steviol glycosides and steviol glycoside precursors in recombinant hosts. In particular, this disclosure relates to production of steviol glycosides comprising steviol-13-*O*-glucoside (13-SMG), rubusoside, rebaudioside B (RebB), rebaudioside A (RebA), rebaudioside D (RebD), and rebaudioside M (RebM) in recombinant hosts comprising genes involved in uridine diphosphate (UDP)-glucose formation.

Description of Related Art

[0002] Sweeteners are well known as ingredients used most commonly in the food, beverage, or confectionary industries. The sweetener can either be incorporated into a final food product during production or for stand-alone use, when appropriately diluted, as a tabletop sweetener or an at-home replacement for sugars in baking. Sweeteners include natural sweeteners such as sucrose, high fructose corn syrup, molasses, maple syrup, and honey and artificial sweeteners such as aspartame, saccharine, and sucralose. Stevia extract is a natural sweetener that can be isolated and extracted from a perennial shrub, *Stevia rebaudiana*. Stevia is commonly grown in South America and Asia for commercial production of stevia extract. Stevia extract, purified to various degrees, is used commercially as a high intensity sweetener in foods and in blends or alone as a tabletop sweetener. Extracts of the Stevia plant generally comprise steviol glycosides that contribute to the sweet flavor, although the amount of each steviol glycoside often varies, *inter alia*, among different production batches.

[0003] Chemical structures for several steviol glycosides are shown in Figure 2, including the diterpene steviol and various steviol glycosides. Extracts of the Stevia plant generally comprise steviol glycosides that contribute to the sweet flavor, although the amount of each steviol glycoside often varies, *inter alia*, among different production batches.

[0004] As recovery and purification of steviol glycosides from the Stevia plant have proven to be labor intensive and inefficient, there remains a need for a recombinant production system that can accumulate high yields of desired steviol glycosides, such as RebM. There also remains a need for improved production of steviol glycosides in recombinant hosts for

commercial uses. As well, there remains a need for increasing UDP-glucose formation in recombinant hosts in order to produce higher yields of steviol glycosides, including RebM.

SUMMARY OF THE INVENTION

[0005] It is against the above background that the present invention provides certain advantages and advancements over the prior art.

[0006] Although this invention as disclosed herein is not limited to specific advantages or functionalities, the invention provides a recombinant host cell capable of producing one or more steviol glycosides or a steviol glycoside composition in a cell culture, comprising:

(a) a recombinant gene encoding a polypeptide capable of synthesizing uridine 5'-triphosphate (UTP) from uridine diphosphate (UDP);

(b) a recombinant gene encoding a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate; and/or

(c) a recombinant gene encoding a polypeptide capable of synthesizing uridine diphosphate glucose (UDP-glucose) from UTP and glucose-1-phosphate.

[0007] In one aspect of the recombinant host cell disclosed herein:

(a) the polypeptide capable of synthesizing UTP from UDP comprises a polypeptide having at least 60% sequence identity to the amino acid sequence set forth in SEQ ID NO:123;

(b) the polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate comprises a polypeptide having at least 60% sequence identity to the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:119, SEQ ID NO:143 or a polypeptide having at least 55% sequence identity to the amino acid sequence set forth in SEQ ID NO:141, SEQ ID NO:145, or SEQ ID NO:147; and/or

(c) the polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate comprises a polypeptide having at least 60% sequence identity to the amino acid sequence set forth in SEQ ID NO:121, SEQ ID NO:127, a polypeptide having at least 55% sequence identity to the amino acid sequence set forth in SEQ ID NO:125, SEQ ID NO:129, SEQ ID NO:133, SEQ ID NO:135, SEQ ID

NO:137, or SEQ ID NO:139 or a polypeptide having at least 70% sequence identity to the amino acid sequence set forth in SEQ ID NO:131.

[0008] In one aspect, the recombinant host cell disclosed herein further comprises:

(a) a gene encoding a polypeptide capable of glycosylating steviol or a steviol glycoside at its C-13 hydroxyl group thereof;

(b) a gene encoding a polypeptide capable of beta 1,3 glycosylation of the C3' of the 13-*O*-glucose, 19-*O*-glucose, or both 13-*O*-glucose and 19-*O*-glucose of a steviol glycoside;

(c) a gene encoding a polypeptide capable of glycosylating steviol or a steviol glycoside at its C-19 carboxyl group thereof; and/or

(d) a gene encoding a polypeptide capable of beta 1,2 glycosylation of the C2' of the 13-*O*-glucose, 19-*O*-glucose, or both 13-*O*-glucose and 19-*O*-glucose of a steviol glycoside.

[0009] In one aspect, the recombinant host cell disclosued herein further comprises:

(e) a gene encoding a polypeptide capable of synthesizing geranylgeranyl pyrophosphate (GGPP) from farnesyl diphosphate (FPP) and isopentenyl diphosphate (IPP);

(f) a gene encoding a polypeptide capable of synthesizing *ent*-copalyl diphosphate from GGPP;

(g) a gene encoding an a polypeptide capable of synthesizing *ent*-kaurene from *ent*-copalyl diphosphate;

(h) a gene encoding a polypeptide capable of synthesizing *ent*-kaurenoic acid from *ent*-kaurene;

(i) a gene encoding a polypeptide capable of reducing cytochrome P450 complex; and/or

(j) a gene encoding a polypeptide capable of synthesizing steviol from *ent*-kaurenoic acid.

[0010] In one aspect of the recombinant host cell disclosed herein:

(a) the polypeptide capable of glycosylating steviol or a steviol glycoside at its C-13 hydroxyl group thereof comprises a polypeptide having at least 55% sequence identity to the amino acid sequence set forth in SEQ ID NO:7;

(b) the polypeptide capable of beta 1,3 glycosylation of the C3' of the 13-O-glucose, 19-O-glucose, or both 13-O-glucose and 19-O-glucose of a steviol glycoside comprises a polypeptide having at least 50% sequence identity to the amino acid sequence set forth in SEQ ID NO:9;

(c) the polypeptide capable of glycosylating steviol or a steviol glycoside at its C-19 carboxyl group thereof comprises a polypeptide having at least 55% sequence identity to the amino acid sequence set forth in SEQ ID NO:4;

(d) the polypeptide capable of beta 1,2 glycosylation of the C2' of the 13-O-glucose, 19-O-glucose, or both 13-O-glucose and 19-O-glucose of a steviol glycoside comprises a polypeptide having 80% or greater identity to the amino acid sequence set forth in SEQ ID NO:11; a polypeptide having 80% or greater identity to the amino acid sequence set forth in SEQ ID NO:13; or a polypeptide having at least 65% sequence identity to the amino acid sequence set forth in SEQ ID NO:16;

(e) the polypeptide capable of synthesizing GGPP comprises a polypeptide having at least 70% sequence identity to the amino acid sequence set forth in SEQ ID NO:20, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, or SEQ ID NO:116;

(f) the polypeptide capable of synthesizing ent-copalyl diphosphate comprises a polypeptide having at least 70% sequence identity to the amino acid sequence set forth in SEQ ID NO:34, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:40, SEQ ID NO:42, or SEQ ID NO:120;

(g) the polypeptide capable of synthesizing ent-kaurene comprises a polypeptide having at least 70% sequence identity to the amino acid sequence set forth in SEQ ID NO:44, SEQ ID NO:46, SEQ ID NO:48, SEQ ID NO:50, or SEQ ID NO:52;

(h) the polypeptide capable of synthesizing ent-kaurenoic acid comprises a polypeptide having at least 70% sequence identity to the amino acid sequence set forth in SEQ ID NO:60, SEQ ID NO:62, SEQ ID NO:117, SEQ ID NO:66, SEQ ID NO:68, SEQ ID NO:70, SEQ ID NO:72, SEQ ID NO:74, or SEQ ID NO:76;

(i) the polypeptide capable of reducing cytochrome P450 complex comprises a polypeptide having at least 70% sequence identity to the amino acid sequence

set forth in SEQ ID NO:78, SEQ ID NO:80, SEQ ID NO:82, SEQ ID NO:84, SEQ ID NO:86, SEQ ID NO:88, SEQ ID NO:90, SEQ ID NO:92; and/or

(k) the polypeptide capable of synthesizing steviol comprises a polypeptide having at least 70% sequence identity to the amino acid sequence set forth in SEQ ID NO:94, SEQ ID NO:97, SEQ ID NO:100, SEQ ID NO:101, SEQ ID NO:102, SEQ ID NO:103, SEQ ID NO:104, SEQ ID NO:106, SEQ ID NO:108, SEQ ID NO:110, SEQ ID NO:112, or SEQ ID NO:114.

[0011] In one aspect, the recombinant host cell disclosued herein comprises:

(a) a gene encoding a polypeptide capable of synthesizing uridine 5'-triphosphate (UTP) from uridine diphosphate (UDP) having at least 60% sequence identity to the amino acid sequence set forth in SEQ ID NO:123;

(b) one or more genes encoding a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate, each having at least 60% sequence identity to the amino acid sequence set forth in SEQ ID NO:2 and/or SEQ ID NO:119; and

(c) a gene encoding a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate having at least 60% sequence identity to the amino acid sequence set forth in SEQ ID NO:121.

[0012] In one aspect, the recombinant host cell disclosued herein comprises:

(a) a gene encoding a polypeptide capable of synthesizing uridine 5'-triphosphate (UTP) from uridine diphosphate (UDP);

(b) a gene encoding a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate;

(c) a gene encoding a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate having at least 60% sequence identity to the amino acid sequence set forth in SEQ ID NO:121;

(d) a gene encoding a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate having at least 55% sequence identity to the amino acid sequence set forth in SEQ ID NO:125, SEQ ID NO:129, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, or SEQ ID NO:139; at least 60% sequence identity to the amino acid sequence set forth in SEQ ID NO:127; or at least 70% sequence identity to the amino acid sequence set forth in SEQ ID NO:131; and

one or more of:

(e) a gene encoding a polypeptide capable of glycosylating steviol or a steviol glycoside at its C-13 hydroxyl group thereof having at least 55% sequence identity to the amino acid sequence set forth in SEQ ID NO:7;

(b) a gene encoding a polypeptide capable of beta 1,3 glycosylation of the C3' of the 13-O-glucose, 19-O-glucose, or both 13-O-glucose and 19-O-glucose of a steviol glycoside having at least 50% sequence identity to the amino acid sequence set forth in SEQ ID NO:9;

(c) a gene encoding a polypeptide capable of glycosylating steviol or a steviol glycoside at its C-19 carboxyl group thereof having at least 55% sequence identity to the amino acid sequence set forth in SEQ ID NO:4;

(d) a gene encoding a polypeptide capable of beta 1,2 glycosylation of the C2' of the 13-O-glucose, 19-O-glucose, or both 13-O-glucose and 19-O-glucose of a steviol glycoside comprises a polypeptide having 80% or greater identity to the amino acid sequence set forth in SEQ ID NO:11; a polypeptide having 80% or greater identity to the amino acid sequence set forth in SEQ ID NO:13; or a polypeptide having at least 65% sequence identity to the amino acid sequence set forth in SEQ ID NO:16.

[0013] In one aspect, the recombinant host cell disclosued herein comprises:

(a) a recombinant gene encoding a polypeptide capable of synthesizing uridine 5'-triphosphate (UTP) from uridine diphosphate (UDP) having at least 60% sequence identity to the amino acid sequence set forth in SEQ ID NO:123;

(b) one or more recombinant genes encoding a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate, each having at least 60% sequence identity to the amino acid sequence set forth in SEQ ID NO:2 and/or SEQ ID NO:119; and/or

(c) a recombinant gene encoding a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate having at least 60% sequence identity to the amino acid sequence set forth in SEQ ID NO:121;

wherein the gene encoding a polypeptide capable of synthesizing uridine 5'-triphosphate (UTP) from uridine diphosphate (UDP), the one or more genes encoding a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate, and/or the gene encoding a polypeptide capable of synthesizing UDP-glucose from UTP and

glucose-1-phosphate are overexpressed relative to a corresponding host cell lacking the one or more recombinant genes.

[0014] In one aspect of the recombinant host cell disclosed herein, the gene encoding a polypeptide capable of synthesizing uridine 5'-triphosphate (UTP) from uridine diphosphate (UDP), the one or more genes encoding a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate, and/or the gene encoding a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate are overexpressed by at least 10%, or at least 15%, or at least 20%, or at least 30%, or at least 40%, or at least 50%, or at least 60%, or at least 70%, or at least 80%, or at least 90%, or at least 100%, or at least 125%, or at least 150%, or at least 175%, or at least 200% relative to a corresponding host cell lacking the one or more recombinant genes.

[0015] In one aspect of the recombinant host cell disclosed herein, expression of the one or more recombinant genes increase the amount of UDP-glucose accumulated by the cell relative to a corresponding host lacking the one or more recombinant genes.

[0016] In one aspect of the recombinant host cell disclosed herein, expression of the one or more recombinant genes increases the amount of UDP-glucose accumulated by the cell by at least about 10%, at least about 25%, or at least about 50%, at least about 100%, at least about 150%, at least about 200%, or at least about 250% relative to a corresponding host lacking the one or more recombinant genes.

[0017] In one aspect of the recombinant host cell disclosed herein, expression of the one or more recombinant genes increases an amount of the one or more steviol glycosides or the steviol glycoside composition produced by the cell relative to a corresponding host lacking the one or more recombinant genes.

[0018] In one aspect of the recombinant host cell disclosed herein, expression of the one or more recombinant genes increases the amount of the one or more steviol glycosides produced by the cell by at least about 5%, at least about 10%, at least about 25%, at least about 50%, at least about 75%, or at least about 100% relative to a corresponding host lacking the one or more recombinant genes.

[0019] In one aspect of the recombinant host cell disclosed herein, expression of the one or more recombinant genes increases the amount of RebA, RebB, Reb D, and/or RebM produced by the cell relative to a corresponding host lacking the one or more recombinant genes.

[0020] In one aspect of the recombinant host cell disclosed herein, expression of the one or more recombinant genes decreases the one of one or more steviol glycosides or the steviol glycoside composition accumulated by the cell relative to a corresponding host lacking the one or more recombinant genes.

[0021] In one aspect of the recombinant host cell disclosed herein, expression of the one or more recombinant genes decreases the amount of the one or more steviol glycosides accumulated by the cell by at least about 5%, at least about 10%, at least about 25%, or at least about 50% relative to a corresponding host lacking the one or more recombinant genes.

[0022] In one aspect of the recombinant host cell disclosed herein, expression of the one or more recombinant genes decreases the amount of RebB, RebD, and/or 13-SMG accumulated by the cell relative to a corresponding host lacking the one or more recombinant genes.

[0023] In one aspect of the recombinant host cell disclosed herein, expression of the one or more recombinant genes increases or decreases the amount of total steviol glycosides produced by the cell by less than 5%, less than 2.5%, or less than 1% relative to a corresponding host lacking the one or more recombinant genes.

[0024] In one aspect of the recombinant host cell disclosed herein, expression of the one or more recombinant genes increases the amount of total steviol glycosides produced by the cell by at least about 5%, at least about 10%, or at least about 25% relative to a corresponding host lacking the one or more recombinant genes.

[0025] In one aspect of the recombinant host cell disclosed herein, the one or more steviol glycosides is, or the steviol glycoside composition comprises, steviol-13-*O*-glucoside (13-SMG), steviol-1,2-Bioside, steviol-1,3-Bioside, steviol-19-*O*-glucoside (19-SMG), 1,2-Stevioside, 1,3-stevioside (RebG), rubusoside, rebaudioside A (RebA), rebaudioside B (RebB), rebaudioside C (RebC), rebaudioside D (RebD), rebaudioside E (RebE), rebaudioside F (RebF), rebaudioside M (RebM), rebaudioside Q (RebQ), rebaudioside I (RebI), dulcoside A, and/or an isomer thereof.

[0026] In one aspect of the recombinant host cell disclosed herein, the recombinant host cell is a plant cell, a mammalian cell, an insect cell, a fungal cell, an algal cell or a bacterial cell.

[0027] The invention also provides method of producing one or more steviol glycosides or a steviol glycoside composition in a cell culture, comprising culturing the recombinant host cell disclosed herein, under conditions in which the genes are expressed, and wherein the one or

more steviol glycosides or the steviol glycoside composition is produced by the recombinant host cell.

[0028] In one aspect of the methods disclosed herein, the genes are constitutively expressed and/or expression of the genes is induced.

[0029] In one aspect of the methods disclosed herein, the amount of UDP-glucose accumulated by the cell is increased by at least by at least about 10% relative to a corresponding host lacking the one or more recombinant genes.

[0030] In one aspect of the methods disclosed herein, the amount of RebA, RebB, RebD, and/or RebM produced by the cell is increased by at least about 5% relative to a corresponding host lacking the one or more recombinant genes.

[0031] In one aspect of the methods disclosed herein, the amount of RebB, RebD, and/or 13-SMG accumulated by the cell is decreased by at least about 5% relative to a corresponding host lacking the one or more recombinant genes.

[0032] In one aspect of the methods disclosed herein, the amount of total steviol glycosides produced by the cell is increased or decreased by less than about 5% relative to a corresponding host lacking the one or more recombinant genes.

[0033] In one aspect of the methods disclosed herein, the amount of total steviol glycosides produced by the cell is increased by at least about 5% relative to a corresponding host lacking the one or more recombinant genes.

[0034] In one aspect of the methods disclosed herein, the recombinant host cell is grown in a fermentor at a temperature for a period of time, wherein the temperature and period of time facilitate the production of the one or more steviol glycosides or the steviol glycoside composition.

[0035] In one aspect of the methods disclosed herein, the amount of UDP-glucose present in the cell culture is increased by at least about 10%, at least about 25%, or at least about 50%, at least about 100%, at least about 150%, at least about 200%, or at least about 250% at any point throughout the period of time.

[0036] In one aspect, the methods disclosed herein further comprise isolating the produced one or more steviol glycosides or the steviol glycoside composition from the cell culture.

[0037] In one aspect of the methods disclosed herein, the isolating step comprises:

(a) providing the cell culture comprising the one or more steviol glycosides or the steviol glycoside composition;

(b) separating a liquid phase of the cell culture from a solid phase of the cell culture to obtain a supernatant comprising the produced one or more steviol glycosides or the steviol glycoside composition;

(c) providing one or more adsorbent resins, comprising providing the adsorbent resins in a packed column; and

(d) contacting the supernatant of step (b) with the one or more adsorbent resins in order to obtain at least a portion of the produced one or more steviol glycosides or the steviol glycoside composition, thereby isolating the produced one or more steviol glycosides or the steviol glycoside composition;

or

(a) providing the cell culture comprising the one or more steviol glycosides or the steviol glycoside composition;

(b) separating a liquid phase of the cell culture from a solid phase of the cell culture to obtain a supernatant comprising the produced one or more steviol glycosides or the steviol glycoside composition;

(c) providing one or more ion exchange or ion exchange or reversed-phase chromatography columns; and

(d) contacting the supernatant of step (b) with the one or more ion exchange or ion exchange or reversed-phase chromatography columns in order to obtain at least a portion of the produced one or more steviol glycosides or the steviol glycoside composition, thereby isolating the produced one or more steviol glycosides or the steviol glycoside composition;

or

(a) providing the cell culture comprising the one or more steviol glycosides or the steviol glycoside composition;

(b) separating a liquid phase of the cell culture from a solid phase of the cell culture to obtain a supernatant comprising the produced one or more steviol glycosides or the steviol glycoside composition;

(c) crystallizing or extracting the produced one or more steviol glycosides or the steviol glycoside composition, thereby isolating the produced one or more steviol glycosides or the steviol glycoside composition.

[0038] In one aspect, the methods disclosed herein further comprise recovering the one or more steviol glycosides or the steviol glycoside composition from the cell culture.

[0039] In one aspect of the methods disclosed herein, the recovered one or more steviol glycosides or the steviol glycoside composition has a reduced level of Stevia plant-derived components relative to a plant-derived Stevia extract.

[0040] The invention also provides a method for producing one or more steviol glycosides or a steviol glycoside composition, comprising whole-cell bioconversion of plant-derived or synthetic steviol and/or steviol glycosides in a cell culture medium of a recombinant host cell using:

(a) a polypeptide capable of synthesizing UTP from UDP having at least 60% sequence identity to the amino acid sequence set forth in SEQ ID NO:123;

(b) a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate having at least 60% sequence identity to the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:119, or SEQ ID NO:143; at least 55% sequence identity to the amino acid sequence set forth in SEQ ID NO:141, SEQ ID NO:145, or SEQ ID NO:147; and/or

(c) a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate having at least 60% sequence identity to the amino acid sequence set forth in SEQ ID NO:121, SEQ ID NO:127; at least 55% sequence identity to the amino acid sequence set forth in SEQ ID NO:125, SEQ ID NO:129, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, or SEQ ID NO:139; or at least 70% sequence identity to the amino acid sequence set forth in SEQ ID NO:131; and

one or more of:

(d) a polypeptide capable of glycosylating steviol or a steviol glycoside at its C-13 hydroxyl group thereof;

(e) a polypeptide capable of beta 1,3 glycosylation of the C3' of the 13-O-glucose, 19-O-glucose, or both 13-O-glucose and 19-O-glucose of a steviol glycoside;

(f) a polypeptide capable of glycosylating steviol or a steviol glycoside at its C-19 carboxyl group thereof; and/or

(g) a polypeptide capable of beta 1,2 glycosylation of the C2' of the 13-O-glucose, 19-O-glucose, or both 13-O-glucose and 19-O-glucose of a steviol glycoside;

wherein at least one of the polypeptides is a recombinant polypeptide expressed in the recombinant host cell; and producing the one or more steviol glycosides or the steviol glycoside composition thereby.

[0041] In one aspect of the methods disclosed herein:

(d) the polypeptide capable of glycosylating steviol or a steviol glycoside at its C-13 hydroxyl group thereof comprises a polypeptide having at least 55% sequence identity to the amino acid sequence set forth in SEQ ID NO:7;

(e) the polypeptide capable of beta 1,3 glycosylation of the C3' of the 13-O-glucose, 19-O-glucose, or both 13-O-glucose and 19-O-glucose of a steviol glycoside comprises a polypeptide having at least 50% sequence identity to the amino acid sequence set forth in SEQ ID NO:9;

(f) the polypeptide capable of glycosylating steviol or a steviol glycoside at its C-19 carboxyl group thereof comprises a polypeptide having at least 55% sequence identity to the amino acid sequence set forth in SEQ ID NO:4;

(g) the polypeptide capable of beta 1,2 glycosylation of the C2' of the 13-O-glucose, 19-O-glucose, or both 13-O-glucose and 19-O-glucose of a steviol glycoside comprises a polypeptide having 80% or greater identity to the amino acid sequence set forth in SEQ ID NO:11; a polypeptide having 80% or greater identity to the amino acid sequence set forth in SEQ ID NO:13; or a polypeptide having at least 65% sequence identity to the amino acid sequence set forth in SEQ ID NO:16.

[0042] In one aspect of the methods disclosed herein, the recombinant host cell is a plant cell, a mammalian cell, an insect cell, a fungal cell, an algal cell or a bacterial cell.

[0043] In one aspect of the methods disclosed herein, the one or more steviol glycosides is, or the steviol glycoside composition comprises, steviol-13-*O*-glucoside (13-SMG), steviol-1,2-Bioside, steviol-1,3-Bioside, steviol-19-*O*-glucoside (19-SMG), 1,2-stevioside, 1,3-stevioside (RebG), rubusoside, rebaudioside A (RebA), rebaudioside B (RebB), rebaudioside C (RebC), rebaudioside D (RebD), rebaudioside E (RebE), rebaudioside F (RebF), rebaudioside M (RebM), rebaudioside Q (RebQ), rebaudioside I (RebI), dulcoside A, and/or an isomer thereof.

[0044] The invention also provides a cell culture, comprising the recombinant host cell disclosed herein, the cell culture further comprising:

(a) the one or more steviol glycosides or the steviol glycoside composition produced by the recombinant host cell;

(b) glucose, fructose, sucrose, xylose, rhamnose, UDP-glucose, UDP-rhamnose, UDP-xylose, and/or N-acetyl-glucosamine; and

(c) supplemental nutrients comprising trace metals, vitamins, salts, YNB, and/or amino acids;

wherein the one or more steviol glycosides or the steviol glycoside composition is present at a concentration of at least 1 mg/liter of the cell culture;

wherein the cell culture is enriched for the one or more steviol glycosides or the steviol glycoside composition relative to a steviol glycoside composition from a Stevia plant and has a reduced level of Stevia plant-derived components relative to a plant-derived Stevia extract.

[0045] The invention also provides a cell culture, comprising the recombinant host cell disclosed herein, the cell culture further comprising:

(a) the one or more steviol glycosides or the steviol glycoside composition produced by the recombinant host cell;

(b) glucose, fructose, sucrose, xylose, rhamnose, UDP-glucose, UDP-rhamnose, UDP-xylose, and/or N-acetyl-glucosamine; and

(c) supplemental nutrients comprising trace metals, vitamins, salts, YNB, and/or amino acids;

wherein UDP-glucose is present in the cell culture at a concentration of at least 100 μM;

wherein the cell culture is enriched for UGP-glucose relative to a steviol glycoside composition from a Stevia plant and has a reduced level of Stevia plant-derived components relative to a plant-derived Stevia extract.

[0046] The invention also provides cell lysate from the recombinant host cell disclosed herein grown in the cell culture, comprising:

(a) the one or more steviol glycosides or the steviol glycoside composition produced by the recombinant host cell;

(b) glucose, fructose, sucrose, xylose, rhamnose, UDP-glucose, UDP-rhamnose, UDP-xylose, and/or N-acetyl-glucosamine; and/or

(c) supplemental nutrients comprising trace metals, vitamins, salts, yeast nitrogen base, YNB, and/or amino acids;

wherein the one or more steviol glycosides or the steviol glycoside composition produced by the recombinant host cell is present at a concentration of at least 1 mg/liter of the cell culture.

[0047] The invention also provides one or more steviol glycosides produced by the recombinant host cell disclosed herein;

wherein the one or more steviol glycosides produced by the recombinant host cell are present in relative amounts that are different from a steviol glycoside composition from a Stevia plant and have a reduced level of Stevia plant-derived components relative to a plant-derived Stevia extract.

[0048] The invention also provides one or more steviol glycosides produced by the method disclosed herein;

wherein the one or more steviol glycosides produced by the recombinant host cell are present in relative amounts that are different from a steviol glycoside composition from a Stevia plant and have a reduced level of Stevia plant-derived components relative to a plant-derived Stevia extract.

[0049] The invention also provides a sweetener composition, comprising the one or more steviol glycosides disclosed herein.

[0050] The invention also provides a food product comprising, the sweetener composition disclosed herein.

[0051] The invention also provides a beverage or a beverage concentrate, comprising the sweetener composition disclosed herein.

[0052] These and other features and advantages of the present invention will be more fully understood from the following detailed description taken together with the accompanying claims. It is noted that the scope of the claims is defined by the recitations therein and not by the specific discussion of features and advantages set forth in the present description.

BRIEF DESCRIPTION OF THE DRAWINGS

[0053] The following detailed description of the embodiments of the present invention can be best understood when read in conjunction with the following drawings, where like structure is indicated with like reference numerals and in which:

[0054] Figure 1 shows the biochemical pathway for producing steviol from geranylgeranyl diphosphate using geranylgeranyl diphosphate synthase (GGPPS), ent-copalyl diphosphate synthase (CDPS), ent-kaurene synthase (KS), ent-kaurene oxidase (KO), and ent-kaurenoic acid hydroxylase (KAH) polypeptides.

[0055] Figure 2 shows representative primary steviol glycoside glycosylation reactions catalyzed by suitable UGT enzymes and chemical structures for several of the compounds found in Stevia extracts.

[0056] Figure 3 shows representative reactions catalyzed by enzymes involved in the UDP-glucose biosynthetic pathway, including uracil permease (FUR4), uracil phosphoribosyltransferase (FUR1), orotate phosphoribosyltransferase 1 (URA5), orotate phosphoribosyltransferase 2 (URA10), orotidine 5'-phosphate decarboxylase (URA3), uridylate kinase (URA6), nucleoside diphosphate kinase (YNK1), phosphoglucomutase-1 (PGM1), phosphoglucomutase-2 (PGM2), and UTP-glucose-1-phosphate uridylyltransferase (UGP1). *See*, *e.g.*, Daran *et al.*, 1995, *Eur J Biochem.* 233(2):520-30.

[0057] Skilled artisans will appreciate that elements in the Figures are illustrated for simplicity and clarity and have not necessarily been drawn to scale. For example, the dimensions of some of the elements in the Figures can be exaggerated relative to other elements to help improve understanding of the embodiment(s) of the present invention.

DETAILED DESCRIPTION OF THE INVENTION

[0058] All publications, patents and patent applications cited herein are hereby expressly incorporated by reference for all purposes.

[0059] Before describing the present invention in detail, a number of terms will be defined. As used herein, the singular forms "a," "an," and "the" include plural referents unless the context

clearly dictates otherwise. For example, reference to a "nucleic acid" means one or more nucleic acids.

[0060] It is noted that terms like "preferably," "commonly," and "typically" are not utilized herein to limit the scope of the claimed invention or to imply that certain features are critical, essential, or even important to the structure or function of the claimed invention. Rather, these terms are merely intended to highlight alternative or additional features that can or cannot be utilized in a particular embodiment of the present invention.

[0061] For the purposes of describing and defining the present invention it is noted that the term "substantially" is utilized herein to represent the inherent degree of uncertainty that can be attributed to any quantitative comparison, value, measurement, or other representation. The term "substantially" is also utilized herein to represent the degree by which a quantitative representation can vary from a stated reference without resulting in a change in the basic function of the subject matter at issue.

[0062] Methods well known to those skilled in the art can be used to construct genetic expression constructs and recombinant cells according to this invention. These methods include *in vitro* recombinant DNA techniques, synthetic techniques, *in vivo* recombination techniques, and polymerase chain reaction (PCR) techniques. *See*, for example, techniques as described in Green & Sambrook, 2012, MOLECULAR CLONING: A LABORATORY MANUAL, Fourth Edition, Cold Spring Harbor Laboratory, New York; Ausubel *et al.*, 1989, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, Greene Publishing Associates and Wiley Interscience, New York, and PCR Protocols: A Guide to Methods and Applications (Innis *et al.*, 1990, Academic Press, San Diego, CA).

[0063] As used herein, the terms "polynucleotide," "nucleotide," "oligonucleotide," and "nucleic acid" can be used interchangeably to refer to nucleic acid comprising DNA, RNA, derivatives thereof, or combinations thereof, in either single-stranded or double-stranded embodiments depending on context as understood by the skilled worker.

[0064] As used herein, the terms "microorganism," "microorganism host," "microorganism host cell," "recombinant host," and "recombinant host cell" can be used interchangeably. As used herein, the term "recombinant host" is intended to refer to a host, the genome of which has been augmented by at least one DNA sequence. Such DNA sequences include but are not limited to genes that are not naturally present, DNA sequences that are not normally transcribed into RNA or translated into a protein ("expressed"), and other genes or DNA sequences which

one desires to introduce into a host. It will be appreciated that typically the genome of a recombinant host described herein is augmented through stable introduction of one or more recombinant genes. Generally, introduced DNA is not originally resident in the host that is the recipient of the DNA, but it is within the scope of this disclosure to isolate a DNA segment from a given host, and to subsequently introduce one or more additional copies of that DNA into the same host, *e.g.*, to enhance production of the product of a gene or alter the expression pattern of a gene. In some instances, the introduced DNA will modify or even replace an endogenous gene or DNA sequence by, *e.g.*, homologous recombination or site-directed mutagenesis. Suitable recombinant hosts include microorganisms.

[0065] As used herein, the term "recombinant gene" refers to a gene or DNA sequence that is introduced into a recipient host, regardless of whether the same or a similar gene or DNA sequence may already be present in such a host. "Introduced," or "augmented" in this context, is known in the art to mean introduced or augmented by the hand of man. Thus, a recombinant gene can be a DNA sequence from another species or can be a DNA sequence that originated from or is present in the same species but has been incorporated into a host by recombinant methods to form a recombinant host. It will be appreciated that a recombinant gene that is introduced into a host can be identical to a DNA sequence that is normally present in the host being transformed, and is introduced to provide one or more additional copies of the DNA to thereby permit overexpression or modified expression of the gene product of that DNA. In some aspects, said recombinant genes are encoded by cDNA. In other embodiments, recombinant genes are synthetic and/or codon-optimized for expression in *S. cerevisiae*.

[0066] As used herein, the term "engineered biosynthetic pathway" refers to a biosynthetic pathway that occurs in a recombinant host, as described herein. In some aspects, one or more steps of the biosynthetic pathway do not naturally occur in an unmodified host. In some embodiments, a heterologous version of a gene is introduced into a host that comprises an endogenous version of the gene.

[0067] As used herein, the term "endogenous" gene refers to a gene that originates from and is produced or synthesized within a particular organism, tissue, or cell. In some embodiments, the endogenous gene is a yeast gene. In some embodiments, the gene is endogenous to *S. cerevisiae*, including, but not limited to *S. cerevisiae* strain S288C. In some embodiments, an endogenous yeast gene is overexpressed. As used herein, the term "overexpress" is used to refer to the expression of a gene in an organism at levels higher than

the level of gene expression in a wild type organism. *See*, *e.g.*, Prelich, 2012, *Genetics* 190:841-54. *See*, *e.g.*, Giaever & Nislow, 2014, *Genetics* 197(2):451-65. In some aspects, overexpression can be performed by integration using the USER cloning system; *see*, *e.g.*, Nour-Eldin *et al.*, 2010, Methods Mol Biol. 643:185-200. As used herein, the terms "deletion," "deleted," "knockout," and "knocked out" can be used interchangeably to refer to an endogenous gene that has been manipulated to no longer be expressed in an organism, including, but not limited to, *S. cerevisiae*.

[0068] As used herein, the terms "heterologous sequence" and "heterologous coding sequence" are used to describe a sequence derived from a species other than the recombinant host. In some embodiments, the recombinant host is an *S. cerevisiae* cell, and a heterologous sequence is derived from an organism other than *S. cerevisiae*. A heterologous coding sequence, for example, can be from a prokaryotic microorganism, a eukaryotic microorganism, a plant, an animal, an insect, or a fungus different than the recombinant host expressing the heterologous sequence. In some embodiments, a coding sequence is a sequence that is native to the host.

[0069] A "selectable marker" can be one of any number of genes that complement host cell auxotrophy, provide antibiotic resistance, or result in a color change. Linearized DNA fragments of the gene replacement vector then are introduced into the cells using methods well known in the art (*see* below). Integration of the linear fragments into the genome and the disruption of the gene can be determined based on the selection marker and can be verified by, for example, PCR or Southern blot analysis. Subsequent to its use in selection, a selectable marker can be removed from the genome of the host cell by, *e.g.*, Cre-LoxP systems (*see*, *e.g.*, Gossen *et al.*, 2002, *Ann. Rev. Genetics* 36:153-173 and U.S. 2006/0014264). Alternatively, a gene replacement vector can be constructed in such a way as to include a portion of the gene to be disrupted, where the portion is devoid of any endogenous gene promoter sequence and encodes none, or an inactive fragment of, the coding sequence of the gene.

[0070] As used herein, the terms "variant" and "mutant" are used to describe a protein sequence that has been modified at one or more amino acids, compared to the wild-type sequence of a particular protein.

[0071] As used herein, the term "inactive fragment" is a fragment of the gene that encodes a protein having, *e.g.*, less than about 10% (*e.g.*, less than about 9%, less than about 8%, less than about 7%, less than about 6%, less than about 5%, less than about 4%, less than about

3%, less than about 2%, less than about 1%, or 0%) of the activity of the protein produced from the full-length coding sequence of the gene. Such a portion of a gene is inserted in a vector in such a way that no known promoter sequence is operably linked to the gene sequence, but that a stop codon and a transcription termination sequence are operably linked to the portion of the gene sequence. This vector can be subsequently linearized in the portion of the gene sequence and transformed into a cell. By way of single homologous recombination, this linearized vector is then integrated in the endogenous counterpart of the gene with inactivation thereof.

[0072] As used herein, the term "steviol glycoside" refers to rebaudioside A (RebA) (CAS # 58543-16-1), rebaudioside B (RebB) (CAS # 58543-17-2), rebaudioside C (RebC) (CAS # 63550-99-2), rebaudioside D (RebD) (CAS # 63279-13-0), rebaudioside E (RebE) (CAS # 63279-14-1), rebaudioside F (RebF) (CAS # 438045-89-7), rebaudioside M (RebM) (CAS # 1220616-44-3), Rubusoside (CAS # 63849-39-4), Dulcoside A (CAS # 64432-06-0), rebaudioside I (RebI) (MassBank Record: FU000332), rebaudioside Q (RebQ), 1,2-Stevioside (CAS # 57817-89-7), 1,3-Stevioside (RebG), Steviol-1,2-Bioside (MassBank Record: FU000299), Steviol-1,3-Bioside, Steviol-13-*O*-glucoside (13-SMG), Steviol-19-*O*-glucoside (19-SMG), a tri-glycosylated steviol glycoside, a tetra-glycosylated steviol glycoside, a penta-glycosylated steviol glycoside, a hexa-glycosylated steviol glycoside, a hepta-glycosylated steviol glycoside, and isomers thereof. *See* Figure 2; *see also*, Steviol Glycosides Chemical and Technical Assessment 69th JECFA, 2007, prepared by Harriet Wallin, Food Agric. Org.

[0073] As used herein, the terms "steviol glycoside precursor" and "steviol glycoside precursor compound" are used to refer to intermediate compounds in the steviol glycoside biosynthetic pathway. Steviol glycoside precursors include, but are not limited to, geranylgeranyl diphosphate (GGPP), *ent*-copalyl-diphosphate, *ent*-kaurene, *ent*-kaurenol, *ent*-kaurenal, *ent*-kaurenoic acid, and steviol. *See* Figure 1. In some embodiments, steviol glycoside precursors are themselves steviol glycoside compounds. For example, 19-SMG, rubusoside, 1,2-stevioside, and RebE are steviol glycoside precursors of RebM. *See* Figure 2.

[0074] Also as used herein, the terms "steviol precursor" and "steviol precursor compound" are used to refer to intermediate compounds in the steviol biosynthetic pathway. Steviol precursors may also be steviol glycoside precursors, and include, but are not limited to, geranylgeranyl diphosphate (GGPP), *ent*-copalyl-diphosphate, *ent*-kaurene, *ent*-kaurenol, *ent*-kaurenal, and *ent*-kaurenoic acid. Steviol glycosides and/or steviol glycoside precursors can be produced *in vivo* (*i.e.*, in a recombinant host), *in vitro* (*i.e.*, enzymatically), or by whole cell

bioconversion. As used herein, the terms "produce" and "accumulate" can be used interchangeably to describe synthesis of steviol glycosides and steviol glycoside precursors *in vivo*, *in vitro*, or by whole cell bioconversion.

[0075] As used herein, the terms "culture broth," "culture medium," and "growth medium" can be used interchangeably to refer to a liquid or solid that supports growth of a cell. A culture broth can comprise glucose, fructose, sucrose, trace metals, vitamins, salts, yeast nitrogen base (YNB), and/or amino acids. The trace metals can be divalent cations, including, but not limited to, Mn^{2+} and/or Mg^{2+}. In some embodiments, Mn^{2+} can be in the form of $MnCl_2$ dihydrate and range from approximately 0.01 g/L to 100 g/L. In some embodiments, Mg^{2+} can be in the form of $MgSO_4$ heptahydrate and range from approximately 0.01 g/L to 100 g/L. For example, a culture broth can comprise i) approximately 0.02-0.03 g/L $MnCl_2$ dihydrate and approximately 0.5-3.8 g/L $MgSO_4$ heptahydrate, ii) approximately 0.03-0.06 g/L $MnCl_2$ dihydrate and approximately 0.5-3.8 g/L $MgSO_4$ heptahydrate, and/or iii) approximately 0.03-0.17 g/L $MnCl_2$ dihydrate and approximately 0.5-7.3 g/L $MgSO_4$ heptahydrate. Additionally, a culture broth can comprise one or more steviol glycosides produced by a recombinant host, as described herein.

[0076] Recombinant steviol glycoside-producing *Saccharomyces cerevisiae* (*S. cerevisiae*) strains are described in WO 2011/153378, WO 2013/022989, WO 2014/122227, and WO 2014/122328, each of which is incorporated by reference in their entirety. Methods of producing steviol glycosides in recombinant hosts, by whole cell bio-conversion, and *in vitro* are also described in WO 2011/153378, WO 2013/022989, WO 2014/122227, and WO 2014/122328.

[0077] In some embodiments, a recombinant host comprising a gene encoding a polypeptide capable of synthesizing geranylgeranyl pyrophosphate (GGPP) from farnesyl diphosphate (FPP) and isopentenyl diphosphate (IPP) (*e.g.*, geranylgeranyl diphosphate synthase (GGPPS)); a gene encoding a polypeptide capable of synthesizing *ent*-copalyl diphosphate from GGPP (*e.g.*, ent-copalyl diphosphate synthase (CDPS)); a gene encoding a polypeptide capable of synthesizing *ent*-kaurene from *ent*-copalyl diphosphate (*e.g.*, kaurene synthase (KS)); a gene encoding a polypeptide capable of synthesizing ent-kaurenoic acid, ent-kaurenol, and/or ent-kaurenal from *ent*-kaurene (*e.g.*, kaurene oxidase (KO)); a gene encoding a polypeptide capable of reducing cytochrome P450 complex (*e.g.*, cytochrome P450 reductase (CPR) or P450 oxidoreductase (POR); for example, but not limited to a polypeptide capable of electron transfer from NADPH to cytochrome P450 complex during conversion of NADPH to $NADP^+$, which is utilized as a cofactor for terpenoid biosynthesis); a gene encoding a

polypeptide capable of synthesizing steviol from *ent*-kaurenoic acid (*e.g.*, steviol synthase (KAH)); and/or a gene encoding a bifunctional polypeptide capable of synthesizing *ent*-copalyl diphosphate from GGPP and synthesizing *ent*-kaurene from *ent*-copalyl diphosphate (*e.g.*, an ent-copalyl diphosphate synthase (CDPS) – *ent*-kaurene synthase (KS) polypeptide) can produce steviol *in vivo*. *See*, *e.g.*, Figure 1. The skilled worker will appreciate that one or more of these genes can be endogenous to the host provided that at least one (and in some embodiments, all) of these genes is a recombinant gene introduced into the recombinant host.

[0078] In some embodiments, a recombinant host comprising a gene encoding a polypeptide capable of glycosylating steviol or a steviol glycoside at its C-13 hydroxyl group (*e.g.*, UGT85C2 polypeptide); a gene encoding a polypeptide capable of beta 1,3 glycosylation of the C3' of the 13-*O*-glucose, 19-*O*-glucose, or both 13-*O*-glucose and 19-*O*-glucose of a steviol glycoside (*e.g.*, UGT76G1 polypeptide); a gene encoding a polypeptide capable of glycosylating steviol or a steviol glycoside at its C-19 carboxyl group (*e.g.*, UGT74G1 polypeptide); and/or a gene encoding a polypeptide capable of beta 1,2 glycosylation of the C2' of the 13-*O*-glucose, 19-*O*-glucose, or both 13-*O*-glucose and 19-*O*-glucose of a steviol glycoside (*e.g.*, UGT91D2 and EUGT11 polypeptide) can produce a steviol glycoside *in vivo*. The skilled worker will appreciate that one or more of these genes can be endogenous to the host provided that at least one (and in some embodiments, all) of these genes is a recombinant gene introduced into the recombinant host.

[0079] In some embodiments, steviol glycosides and/or steviol glycoside precursors are produced *in vivo* through expression of one or more enzymes involved in the steviol glycoside biosynthetic pathway in a recombinant host. For example, a recombinant host comprising a gene encoding a polypeptide capable of synthesizing geranylgeranyl pyrophosphate (GGPP) from farnesyl diphosphate (FPP) and isopentenyl diphosphate (IPP); a gene encoding a polypeptide capable of synthesizing *ent*-copalyl diphosphate from GGPP; a gene encoding a polypeptide capable of synthesizing *ent*-kaurene from *ent*-copalyl diphosphate; a gene encoding a polypeptide capable of synthesizing ent-kaurenoic acid, ent-kaurenol, and/or ent-kaurenal from *ent*-kaurene; a gene encoding a polypeptide capable of reducing cytochrome P450 complex; a gene encoding a bifunctional polypeptide capable of synthesizing *ent*-copalyl diphosphate from GGPP and synthesizing *ent*-kaurene from *ent*-copalyl diphosphate; a gene encoding a polypeptide capable of glycosylating steviol or a steviol glycoside at its C-13 hydroxyl group (*e.g.*, UGT85C2 polypeptide); a gene encoding a polypeptide capable of beta 1,3 glycosylation of the C3' of the 13-*O*-glucose, 19-*O*-glucose, or both 13-*O*-glucose and 19-*O*-

glucose of a steviol glycoside (*e.g.*, UGT76G1 polypeptide); a gene encoding a polypeptide capable of glycosylating steviol or a steviol glycoside at its C-19 carboxyl group (*e.g.*, UGT74G1 polypeptide); and/or a gene encoding a polypeptide capable of beta 1,2 glycosylation of the C2' of the 13-*O*-glucose, 19-*O*-glucose, or both 13-*O*-glucose and 19-*O*-glucose of a steviol glycoside (*e.g.*, UGT91D2 and EUGT11 polypeptide) can produce a steviol glycoside and/or steviol glycoside precursors *in vivo*. *See*, *e.g.*, Figures 1 and 2. The skilled worker will appreciate that one or more of these genes can be endogenous to the host provided that at least one (and in some embodiments, all) of these genes is a recombinant gene introduced into the recombinant host.

[0080] In some embodiments, a steviol-producing recombinant microorganism comprises heterologous nucleic acids encoding a polypeptide capable of glycosylating steviol or a steviol glycoside at its C-13 hydroxyl group; a polypeptide capable of beta 1,3 glycosylation of the C3' of the 13-*O*-glucose, 19-*O*-glucose, or both 13-*O*-glucose and 19-*O*-glucose of a steviol glycoside; a polypeptide capable of glycosylating steviol or a steviol glycoside at its C-19 carboxyl group; and a polypeptide capable of beta 1,2 glycosylation of the C2' of the 13-*O*-glucose, 19-O-glucose, or both 13-*O*-glucose and 19-*O*-glucose of a steviol glycoside.

[0081] In some embodiments, a steviol-producing recombinant microorganism comprises heterologous nucleic acids encoding a polypeptide capable of glycosylating steviol or a steviol glycoside at its C-13 hydroxyl group, a polypeptide capable of beta 1,3 glycosylation of the C3' of the 13-*O*-glucose, 19-*O*-glucose, or both 13-*O*-glucose and 19-*O*-glucose of a steviol glycoside, and a polypeptide capable of beta 1,2 glycosylation of the C2' of the 13-*O*-glucose, 19-*O*-glucose, or both 13-*O*-glucose and 19-*O*-glucose of a steviol glycoside polypeptides.

[0082] In some aspects, a polypeptide capable of glycosylating steviol or a steviol glycoside at its C-13 hydroxyl group, a polypeptide capable of beta 1,3 glycosylation of the C3' of the 13-*O*-glucose, 19-*O*-glucose, or both 13-*O*-glucose and 19-*O*-glucose of a steviol glycoside, a polypeptide capable of glycosylating steviol or a steviol glycoside at its C-19 carboxyl group, and/or a polypeptide capable of beta 1,2 glycosylation of the C2' of the 13-*O*-glucose, 19-*O*-glucose, or both 13-*O*-glucose and 19-*O*-glucose of a steviol glycoside, transfers a glucose molecule from uridine diphosphate glucose (UDP-glucose) to steviol and/or a steviol glycoside.

[0083] In some aspects, UDP-glucose is produced *in vivo* through expression of one or more enzymes involved in the UDP-glucose biosynthetic pathway in a recombinant host. For example, a recombinant host comprising a gene encoding a polypeptide capable of transporting

uracil into the host cell (*e.g.*, uracil permease (FUR4)); a gene encoding a polypeptide capable of synthesizing uridine monophosphate (UMP) from uracil (*e.g.*, uracil phosphoribosyltransferase (FUR1)); a gene encoding a polypeptide capable of synthesizing orotidine monophosphate (OMP) from orotate or orotic acid (*e.g.*, orotate phosphoribosyltransferase 1 (URA5) and orotate phosphoribosyltransferase 2 (URA10)); a gene encoding a polypeptide capable of synthesizing UMP from OMP (*e.g.*, orotidine 5'-phosphate decarboxylase (URA3)); a gene encoding a polypeptide capable of synthesizing uridine diphosphate (UDP) from UMP (*e.g.*, uridylate kinase (URA6)); a gene encoding a polypeptide capable of synthesizing uridine 5'-triphosphate (UTP) from UDP (*i.e.*, a polypeptide capable of catalyzing the transfer of gamma phosphates from nucleoside triphosphates, *e.g.*, nucleoside diphosphate kinase (YNK1)); a gene encoding a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate (*e.g.*, phosphoglucomutase-1 (PGM1) and phosphoglucomutase-2 (PGM2)); and/or a gene encoding a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate (*e.g.*, UTP-glucose-1-phosphate uridylyltransferase (UGP1) can produce UDP-glucose *in vivo*. *See*, *e.g.*, Figure 3. The skilled worker will appreciate that one or more of these genes may be endogenous to the host.

[0084] In some embodiments, a recombinant host comprises a gene encoding a polypeptide capable of synthesizing UTP from UDP. In some aspects, the gene encoding a polypeptide capable of synthesizing UTP from UDP is a recombinant gene. In some aspects, the recombinant gene comprises a nucleotide sequence native to the host. In other aspects, the recombinant gene comprises a heterologous nucleotide sequence. In some aspects, the recombinant gene is operably linked to a promoter. In some aspects, the recombinant gene is operably linked to a terminator, for example but not limited to, tCYC1 (SEQ ID NO:154) or tADH1 (SEQ ID NO:155). In some aspects, the promoter and terminator drive high expression of the recombinant gene. In some aspects, the recombinant gene is operably linked to a strong promoter, for example but not limited to, pTEF1 (SEQ ID NO:148), pPGK1 (SEQ ID NO:149), pTDH3 (SEQ ID NO:150), pTEF2 (SEQ ID NO:151), pTPI1 (SEQ ID NO:152), or pPDC1 (SEQ ID NO:153). In some aspects, the recombinant gene comprises a nucleotide sequence that originated from or is present in the same species as the recombinant host. In some aspects, expression of a recombinant gene encoding a polypeptide capable of synthesizing UTP from UDP results in a total expression level of genes encoding a polypeptide capable of synthesizing UTP from UDP that is higher than the expression level of endogenous genes encoding a

polypeptide capable of synthesizing UTP from UDP, *i.e.*, an overexpression of a polypeptide capable of synthesizing UTP from UDP.

[0085] In some aspects, the gene encoding the polypeptide capable of synthesizing UTP from UDP is a gene present in the same species as the recombinant host, *i.e.*, an endogenous gene. In some embodiments, the wild-type promoter of an endogenous gene encoding the polypeptide capable of synthesizing UTP from UDP can be exchanged for a strong promoter. In some aspects, the strong promoter drives high expression of the endogenous gene (*i.e.*, overexpression of the gene). In other embodiments, the wild-type enhancer of an endogenous gene encoding a polypeptide capable of synthesizing UTP from UDP can be exchanged for a strong enhancer. In some embodiments, the strong enhancer drives high expression of the endogenous gene (*i.e.*, overexpression of the gene). In some embodiments, both the wild-type enhancer (*i.e.*, operably linked to the promoter) and the wild-type promoter (*i.e.*, operably linked to the endogenous gene) of the endogenous gene can be exchanged for a strong enhancer and strong promoter, respectively, resulting in overexpression of a polypeptide capable of synthesizing UTP from UDP (*i.e.*, relative to the expression level of endogenous genes operably linked to wild-type enhancers and/or promoters). The endogenous gene operably linked to the strong enhancer and/or promoter may be located at the native loci, and/or may be located elsewhere in the genome.

[0086] For example, in some embodiments, a recombinant host comprising an endogenous gene encoding a polypeptide capable of synthesizing UTP from UDP, operably linked to a wild-type promoter, further comprises a recombinant gene encoding a polypeptide capable of synthesizing UTP from UDP, comprising a nucleotide sequence native to the host, operably linked to, *e.g.*, a wild-type promoter, a promoter native to the host, or a heterologous promoter. In another example, in some embodiments, a recombinant host comprising an endogenous gene encoding a polypeptide capable of synthesizing UTP from UDP, operably linked to a wild-type promoter, further comprises a recombinant gene encoding a polypeptide capable of synthesizing UTP from UDP, comprising a heterologous nucleotide sequence, operably linked to, *e.g.*, a wild-type promoter, a promoter native to the host, or a heterologous promoter. In yet another example, in some embodiments, a recombinant host comprises an endogenous gene encoding a polpeptide capable of synthesizing UTP from UDP, operably linked to, *e.g.*, a strong promoter native to the host, or a heterologous promoter.

[0087] The person of ordinary skill in the art will appreciate that, *e.g.*, expression of a recombinant gene encoding a polypeptide capable of synthesizing UTP from UDP; expression of a recombinant gene and an endogenous gene encoding a polypeptide capable of synthesizing UTP from UDP, and expression of an endogenous gene encoding a polypeptide capable of synthesizing UTP from UDP, wherein the wild-type promoter and/or enhancer of the endogenous gene are exchanged for a strong promoter and/or enhancer, each result in overexpression of a polypeptide capable of synthesizing UTP from UDP relative to a corresponding host not expressing a recombinant gene encoding a polypeptide capable of synthesizing UTP from UDP and/or a corresponding host expressing only a native gene encoding a polypeptide capable of synthesizing UTP from UDP, operably linked to the wild-type promoter and enhancer—*i.e.*, as used herein, the term "expression" may include "overexpression."

[0088] In some embodiments, a polypeptide capable of synthesizing UTP from UDP is overexpressed such that the total expression level of genes encoding the polypeptide capable of synthesizing UTP from UDP is at least 5% higher than the expression level of endogenous genes encoding a polypeptide capable of synthesizing UTP from UDP. In some embodiments, the total expression level of genes encoding a polypeptide capable of synthesizing UTP from UDP is at least 10%, or at least 15%, or at least 20%, or at least 30%, or at least 40%, or at least 50%, or at least 60%, or at least 70%, or at least 80%, or at least 90%, or at least 100%, or at least 125%, or at least 150%, or at least 175%, or at least 200% higher than the expression level of endogenous genes encoding a polypeptide capable of synthesizing UTP from UDP.

[0089] In some embodiments, a recombinant host comprises a gene encoding a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate. In some aspects, the gene encoding a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate is a recombinant gene. In some aspects, the recombinant gene comprises a nucleotide sequence native to the host. In other aspects, the recombinant gene comprises a heterologous nucleotide sequence. In some aspects, the recombinant gene is operably linked to a promoter. In some aspects, the recombinant gene is operably linked to a terminator, for example but not limited to, tCYC1 (SEQ ID NO:154) or tADH1 (SEQ ID NO:155). In some aspects, the promoter and terminator drive high expression of the recombinant gene. In some aspects, the recombinant gene is operably linked to a strong promoter, for example but not limited to, pTEF1 (SEQ ID NO:148), pPGK1 (SEQ ID NO:149), pTDH3 (SEQ ID NO:150), pTEF2 (SEQ ID NO:151), pTPI1 (SEQ ID NO:152), or pPDC1 (SEQ ID NO:153). In some aspects, the recombinant gene

comprises a nucleotide sequence that originated from or is present in the same species as the recombinant host. In some aspects, expression of a recombinant gene encoding a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate results in a total expression level of genes encoding a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate that is higher than the expression level of endogenous genes encoding a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate, *i.e.*, an overexpression of a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate.

[0090] In some aspects, the gene encoding the polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate is a gene present in the same species as the recombinant host, *i.e.*, an endogenous gene. In some embodiments, the wild-type promoter of an endogenous gene encoding the polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate can be exchanged for a strong promoter. In some aspects, the strong promoter drives high expression of the endogenous gene (*i.e.*, overexpression of the gene). In other embodiments, the wild-type enhancer of an endogenous gene encoding a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate can be exchanged for a strong enhancer. In some embodiments, the strong enhancer drives high expression of the endogenous gene (*i.e.*, overexpression of the gene). In some embodiments, both the wild-type enhancer (*i.e.*, operably linked to the promoter) and the wild-type promoter (*i.e.*, operably linked to the endogenous gene) of the endogenous gene can be exchanged for a strong enhancer and strong promoter, respectively, resulting in overexpression of a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate (*i.e.*, relative to the expression level of endogenous genes operably linked to wild-type enhancers and/or promoters). The endogenous gene operably linked to the strong enhancer and/or promoter may be located at the native loci, and/or may be located elsewhere in the genome.

[0091] For example, in some embodiments, a recombinant host comprising an endogenous gene encoding a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate, operably linked to a wild-type promoter, further comprises a recombinant gene encoding a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate, comprising a nucleotide sequence native to the host, operably linked to, *e.g.*, a wild-type promoter, a promoter native to the host, or a heterologous promoter. In another example, in some embodiments, a recombinant host comprising an endogenous gene encoding a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate, operably linked to a wild-type promoter, further comprises a recombinant gene encoding a polypeptide capable

of converting glucose-6-phosphate to glucose-1-phosphate, comprising a heterologous nucleotide sequence, operably linked to, *e.g.*, a wild-type promoter, a promoter native to the host, or a heterologous promoter. In yet another example, in some embodiments, a recombinant host comprises an endogenous gene encoding a polpeptide capable of converting glucose-6-phosphate to glucose-1-phosphate, operably linked to, *e.g.*, a strong promoter native to the host, or a heterologous promoter.

[0092] In some embodiments, a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate is overexpressed such that the total expression level of genes encoding the polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate is at least 5% higher than the expression level of endogenous genes encoding a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate. In some embodiments, the total expression level of genes encoding a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate is at least 10%, or at least 15%, or at least 20%, or at least 30%, or at least 40%, or at least 50%, or at least 60%, or at least 70%, or at least 80%, or at least 90%, or at least 100%, or at least 125%, or at least 150%, or at least 175%, or at least 200% higher than the expression level of endogenous genes encoding a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate.

[0093] In some embodiments, a recombinant host comprises a gene encoding a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate. In some aspects, the gene encoding a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate is a recombinant gene. In some aspects, the recombinant gene comprises a nucleotide sequence native to the host. In other aspects, the recombinant gene comprises a heterologous nucleotide sequence. In some aspects, the recombinant gene is operably linked to a promoter. In some aspects, the recombinant gene is operably linked to a terminator, for example but not limited to, tCYC1 (SEQ ID NO:154) or tADH1 (SEQ ID NO:155). In some aspects, the promoter and terminator drive high expression of the recombinant gene. In some aspects, the recombinant gene is operably linked to a strong promoter, for example but not limited to, pTEF1 (SEQ ID NO:148), pPGK1 (SEQ ID NO:149), pTDH3 (SEQ ID NO:150), pTEF2 (SEQ ID NO:151), pTPI1 (SEQ ID NO:152), or pPDC1 (SEQ ID NO:153). In some aspects, the recombinant gene comprises a nucleotide sequence that originated from or is present in the same species as the recombinant host. In some aspects, expression of a recombinant gene encoding a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate results in a total expression level of genes encoding a polypeptide capable

of synthesizing UDP-glucose from UTP and glucose-1-phosphate that is higher than the expression level of endogenous genes encoding a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate, *i.e.*, an overexpression of a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate.

[0094] In some aspects, the gene encoding the polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate is a gene present in the same species as the recombinant host, *i.e.*, an endogenous gene. In some embodiments, the wild-type promoter of an endogenous gene encoding the polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate can be exchanged for a strong promoter. In some aspects, the strong promoter drives high expression of the endogenous gene (*i.e.*, overexpression of the gene). In other embodiments, the wild-type enhancer of an endogenous gene encoding a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate can be exchanged for a strong enhancer. In some embodiments, the strong enhancer drives high expression of the endogenous gene (*i.e.*, overexpression of the gene). In some embodiments, both the wild-type enhancer (*i.e.*, operably linked to the promoter) and the wild-type promoter (*i.e.*, operably linked to the endogenous gene) of the endogenous gene can be exchanged for a strong enhancer and strong promoter, respectively, resulting in overexpression of a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate (*i.e.*, relative to the expression level of endogenous genes operably linked to wild-type enhancers and/or promoters). The endogenous gene operably linked to the strong enhancer and/or promoter may be located at the native loci, and/or may be located elsewhere in the genome.

[0095] For example, in some embodiments, a recombinant host comprising an endogenous gene encoding a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate, operably linked to a wild-type promoter, further comprises a recombinant gene encoding a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate, comprising a nucleotide sequence native to the host, operably linked to, *e.g.*, a wild-type promoter, a promoter native to the host, or a heterologous promoter. In another example, in some embodiments, a recombinant host comprising an endogenous gene encoding a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate, operably linked to a wild-type promoter, further comprises a recombinant gene encoding a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate, comprising a heterologous nucleotide sequence, operably linked to, *e.g.*, a wild-type promoter, a promoter native to the host, or a heterologous promoter. In yet another example, in some embodiments,

a recombinant host comprises an endogenous gene encoding a polpeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate, operably linked to, *e.g.*, a strong promoter native to the host, or a heterologous promoter.

[0096] In some embodiments, a recombinant host comprising a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate is overexpressed such that the total expression level of genes encoding the polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate is at least 5% higher than the expression level of endogenous genes encoding a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate. In some embodiments, the total expression level of genes encoding a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate is at least 10%, or at least 15%, or at least 20%, or at least 30%, or at least 40%, or at least 50%, or at least 60%, or at least 70%, or at least 80%, or at least 90%, or at least 100%, or at least 125%, or at least 150%, or at least 175%, or at least 200% higher than the expression level of endogenous genes encoding a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate.

[0097] In some aspects, a recombinant host comprising one or more genes encoding one or more polypeptides capable of synthesizing UTP from UDP, one or more genes encoding one or more polypeptides capable of converting glucose-6-phosphate to glucose-1-phosphate, and/or one or more genes encoding one or more polypeptides capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate may further comprise a recombinant gene encoding a polypeptide capable of transporting uracil into the host cell; a recombinant gene encoding a polypeptide capable of synthesizing uridine monophosphate (UMP) from uracil; a recombinant gene encoding a polypeptide capable of synthesizing orotidine monophosphate (OMP) from orotate or orotic acid; a recombinant gene encoding a polypeptide capable of synthesizing UMP from OMP; and/or a recombinant gene encoding a polypeptide capable of synthesizing uridine diphosphate (UDP) from UMP. In some embodiments, a recombinant host comprising one or more genes encoding one or more polypeptides capable of synthesizing UTP from UDP, one or more genes encoding one or more polypeptides capable of converting glucose-6-phosphate to glucose-1-phosphate, and/or one or more genes encoding one or more polypeptides capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate may overexpress a gene encoding a polypeptide capable of transporting uracil into the host cell; a gene encoding a polypeptide capable of synthesizing uridine monophosphate (UMP) from uracil; a gene encoding a polypeptide capable of synthesizing orotidine monophosphate (OMP) from orotate

or orotic acid; a gene encoding a polypeptide capable of synthesizing UMP from OMP; and/or a gene encoding a polypeptide capable of synthesizing uridine diphosphate (UDP) from UMP.

[0098] In some aspects, the polypeptide capable of synthesizing UTP from UDP comprises a polypeptide having the amino acid sequence set forth in SEQ ID NO:123 (which can be encoded by the nucleotide sequence set forth in SEQ ID NO:122).

[0099] In some aspects, the polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate comprises a polypeptide having the amino acid sequence set forth in SEQ ID NO:2 (which can be encoded by the nucleotide sequence set forth in SEQ ID NO:1), SEQ ID NO:119 (encoded by the nucleotide sequence set forth in SEQ ID NO:118), SEQ ID NO:141 (encoded by the nucleotide sequence set forth in SEQ ID NO:140), SEQ ID NO:143 (encoded by the nucleotide sequence set forth in SEQ ID NO:142), SEQ ID NO:145 (encoded by the nucleotide sequence set forth in SEQ ID NO:144), or SEQ ID NO:147 (encoded by the nucleotide sequence set forth in SEQ ID NO:146).

[00100] In some aspects, the polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate comprises a polypeptide having the amino acid sequence set forth in SEQ ID NO:121 (which can be encoded by the nucleotide sequence set forth in SEQ ID NO:120), SEQ ID NO:125 (encoded by the nucleotide sequence set forth in SEQ ID NO:124), SEQ ID NO:127 (encoded by the nucleotide sequence set forth in SEQ ID NO:126), SEQ ID NO:129 (encoded by the nucleotide sequence set forth in SEQ ID NO:128), SEQ ID NO:131 (encoded by the nucleotide sequence set forth in SEQ ID NO:130), SEQ ID NO:133 (encoded by the nucleotide sequence set forth in SEQ ID NO:132), SEQ ID NO:135 (encoded by the nucleotide sequence set forth in SEQ ID NO:134), SEQ ID NO:137 (encoded by the nucleotide sequence set forth in SEQ ID NO:136), or SEQ ID NO:139 (encoded by the nucleotide sequence set forth in SEQ ID NO:138).

[00101] In some embodiments, a recombinant host comprises a recombinant gene encoding a polypeptide capable of synthesizing UTP from UDP and a recombinant gene encoding a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate. In some embodiments, a recombinant host comprises a recombinant gene encoding a polypeptide capable of synthesizing UTP from UDP and a recombinant gene encoding a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate. In some embodiments, a recombinant host comprises a recombinant gene encoding a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate and a recombinant gene

encoding a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate. In some embodiments, a recombinant host comprises a recombinant gene encoding a polypeptide capable of synthesizing UTP from UDP, a recombinant gene encoding a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate, and a recombinant gene encoding a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate.

[00102] In some embodiments, a recombinant host comprises two or more recombinant genes encoding a polypeptide involved in the UDP-glucose biosynthetic pathway, *e.g.*, a gene encoding a polypeptide capable of converting glucose-6-phosphate having a first amino acid sequence and a gene encoding a polypeptide capable of converting glucose-6-phosphate having a second amino acid sequence distinct from the first amino acid sequence. For example, in some embodiments, a recombinant host comprises a gene encoding a polypeptide having the amino acid sequence of PGM1 (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:2) and a gene encoding a polypeptide having the amino acid sequence of PGM2 (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:119, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, or SEQ ID NO:147). In certain such embodiments, the two or more genes encoding a polypeptide involved in the UDP-glucose biosynthetic pathway comprise nucleotide sequences native to the recombinant host cell (*e.g.*, a recombinant *S. cerevisiae* host cell comprising a gene encoding a polypeptide having the amino acid sequence set forth in SEQ ID NO:2 and a gene encoding a polypeptide having the amino acid sequence set forth in SEQ ID NO:119). In other such embodiments, one of the two or more genes encoding a polypeptide involved in the UDP-glucose biosynthetic pathway comprises a nucleotide sequence native to the recombinant host cell, while one or more of the two or more genes encoding a polypeptide involved in the UDP-glucose biosynthetic pathway comprises a heterologous nucleotide sequence. For example, in some embodiments, a recombinant *S. cerevisiae* host cell expressing a recombinant gene encoding a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate having the amino acid sequence set forth in SEQ ID NO:121 (*i.e.*, a recombinant host overexpressing the polypeptide) further expresses a recombinant gene encoding a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate having the amino acid sequence set forth in, *e.g.*, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, or SEQ ID NO:139. In another example, in some embodiments, a recombinant *S. cerevisiae* host cell expressing a recombinant gene encoding a

polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate having the amino acid sequence set forth in SEQ ID NO:119 (*i.e.*, a recombinant host overexpressing the polypeptide) further expresses a recombinant gene encoding a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate having the amino acid sequence set forth in, *e.g.*, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, or SEQ ID NO:147. Accordingly, as used herein, the term "a recombinant gene" may include "one or more recombinant genes."

[00103] In some embodiments, a recombinant host comprises two or more copies of a recombinant gene encoding a polypeptide involved in the UDP-glucose biosynthetic pathway or the steviol glycoside biosynthetic pathway. In some embodiments, a recombinant host is preferably transformed with, *e.g.*, two copies, three copies, four copies, or five copies of a recombinant gene encoding a polypeptide involved in the UDP-glucose biosynthetic pathway or the steviol glycoside biosynthetic pathway. For example, in some embodiments, a recombinant host is transformed with two copies of a recombinant gene encoding a polypeptide capable of synthesizing UTP from UDP (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:123). The person of ordinary skill in the art will appreciate that, in some embodiments, recombinant genes may be replicated in a host cell independently of cell replication; accordingly, a recombinant host cell may comprise, *e.g.*, more copies of a recombinant gene than the number of copies the cell was transformed with. Accordingly, as used herein, the term "a recombinant gene" may include "one or more copies of a recombinant gene."

[00104] In some aspects, expression of a polypeptide capable of synthesizing UTP from UDP, a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate, and/or a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate in a recombinant host cell increases the amount of UDP-glucose produced by the cell. In some aspects, expression of a polypeptide capable of synthesizing UTP from UDP, a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate, and/or a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate in a recombinant host cell maintains, or even increases, the pool of UDP-glucose available for, *e.g.*, glycosylation of steviol or a steviol glycoside. In some aspects, expression of a polypeptide capable of synthesizing UTP from UDP, a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate, and/or a polypeptide capable sunthesizing UDP-glucose from UTP and glucose-1-phosphate in a recombinant host cell increases the speed which which UDP-glucose

is regenerated, thus maintaining, or even increasing, the UDP-glucose pool, which can be used to synthesize one or more steviol glycosides.

[00105] In some embodiments, expression of a recombinant gene encoding a polypeptide capable of synthesizing UTP from UDP (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:123), a recombinant gene encoding a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate (*e.g.* a polypeptide having the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:119, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, or SEQ ID NO:147), and a recombinant gene encoding a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:121, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, or SEQ ID NO:139) in a recombinant host cell increases the amount of UDP-glucose produced by the cell by at least about 10%, *e.g.*, at least about 25%, or at least about 50%, or at least about 75%, or at least about 100%, or at least about 125%, or at least about 150%, or at least about 175%, or at least about 200%, or at least about 225%, or at least about 250%, or at least about 275%, or at least about 300%, calculated as an increase in intracellular UDP-glucose concentration relative to a corresponding host lacking the recombinant genes.

[00106] In certain such embodiments, one or more of the recombinant gene encoding a polypeptide capable of synthesizing UTP from UDP, the recombinant gene encoding a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate, and the recombinant gene encoding a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate comprise a nucleotide sequence native to the host cell. For example, in some embodiments, expression of a recombinant gene encoding a polypeptide capable of synthesizing UTP from UDP having the amino acid sequence set forth in SEQ ID NO:123, a recombinant gene encoding a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate having the amino acid sequence set forth in SEQ ID NO:2 and/or SEQ ID NO:119, and a recombinant gene encoding a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate having the amino acid sequence set forth in SEQ ID NO:121 in a steviol glycoside-producing *S. cerevisiae* host cell (*i.e.*, providing a recombinant host overexpressing the polypeptides) increases the amount of UDP-glucose produced by the cell by at least about 10%, *e.g.*, at least about 25%, or at least about 50%, or at least about 75%, or at least about 100%, or at least about 125%, or at least about 150%, or at least about 175%, or at least about 200%, or at least about 225%, or at least about 250%, or at least about

275%, or at least about 300%, calculated as an increase in intracellular UDP-glucose concentration relative to a corresponding host lacking the recombinant genes.

[00107] In some aspects, expression of a polypeptide capable of synthesizing UTP from UDP, a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate, and/or a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate in a steviol-glycoside producing recombinant host cell further expressing a gene encoding a polypeptide capable of glycosylating steviol or a steviol glycoside at its C-13 hydroxyl group; a gene encoding a polypeptide capable of beta 1,3 glycosylation of the C3' of the 13-*O*-glucose, 19-*O*-glucose, or both 13-*O*-glucose and 19-*O*-glucose of a steviol glycoside; a gene encoding a polypeptide capable of glycosylating steviol or a steviol glycoside at its C-19 carboxyl group; and/or a gene encoding a polypeptide capable of beta 1,2 glycosylation of the C2' of the 13-*O*-glucose, 19-*O*-glucose, or both 13-*O*-glucose and 19-*O*-glucose of a steviol glycoside, increases the amount of one or more steviol glycosides produced by the cell, and/or decreases the amount of one or more steviol glycosides produced by the cell. In some embodiments, the steviol glycoside-producing host further expresses a gene encoding a polypeptide capable of synthesizing GGPP from FPP and IPP; a gene encoding a polypeptide capable of synthesizing *ent*-copalyl diphosphate from GGPP; a gene encoding a polypeptide capable of synthesizing *ent*-kaurene from *ent*-copalyl diphosphate; a gene encoding a polypeptide capable of synthesizing ent-kaurenoic acid, ent-kaurenol, and/or ent-kaurenal from *ent*-kaurene; a gene encoding a polypeptide capable of reducing cytochrome P450 complex; and a gene encoding a polypeptide capable of synthesizing steviol from *ent*-kaurenoic acid; and/or a gene encoding a bifunctional polypeptide capable of synthesizing *ent*-copalyl diphosphate from GGPP and synthesizing *ent*-kaurene from *ent*-copalyl diphosphate.

[00108] In some aspects, the polypeptide capable of synthesizing geranylgeranyl pyrophosphate (GGPP) from farnesyl diphosphate (FPP) and isopentenyl diphosphate (IPP) comprises a polypeptide having an amino acid sequence set forth in SEQ ID NO:20 (which can be encoded by the nucleotide sequence set forth in SEQ ID NO:19), SEQ ID NO:22 (encoded by the nucleotide sequence set forth in SEQ ID NO:21), SEQ ID NO:24 (encoded by the nucleotide sequence set forth in SEQ ID NO:23), SEQ ID NO:26 (encoded by the nucleotide sequence set forth in SEQ ID NO:25), SEQ ID NO:28 (encoded by the nucleotide sequence set forth in SEQ ID NO:27), SEQ ID NO:30 (encoded by the nucleotide sequence set forth in SEQ ID NO:29), SEQ ID NO:32 (encoded by the nucleotide sequence set forth in SEQ ID NO:31), or SEQ ID NO:116 (encoded by the nucleotide sequence set forth in SEQ ID NO:115). In some

embodiments, a recombinant host comprising a gene encoding a polypeptide capable of synthesizing geranylgeranyl pyrophosphate (GGPP) from farnesyl diphosphate (FPP) and isopentenyl diphosphate (IPP) further comprises one or more genes encoding one or more polypeptides capable of synthesizing UTP from UDP (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:123), one or more genes encoding one or more polypeptides capable of converting glucose-6-phosphate to glucose-1-phosphate (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:119, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, and/or SEQ ID NO:147), and/or one or more genes encoding one or more polypeptides capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:121, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, and/or SEQ ID NO:139). In some embodiments, the recombinant host is an *S. cerevisiae* host cell overexpressing one or more genes encoding one or more polypeptides involved in the UDP-glucose biosynthetic pathway (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:119, SEQ ID NO:121, and/or SEQ ID NO:123).

[00109] In some aspects, the polypeptide capable of synthesizing *ent*-copalyl diphosphate from GGPP comprises a polypeptide having an amino acid sequence set forth in SEQ ID NO:34 (which can be encoded by the nucleotide sequence set forth in SEQ ID NO:33), SEQ ID NO:36 (encoded by the nucleotide sequence set forth in SEQ ID NO:35), SEQ ID NO:38 (encoded by the nucleotide sequence set forth in SEQ ID NO:37), SEQ ID NO:40 (encoded by the nucleotide sequence set forth in SEQ ID NO:39), or SEQ ID NO:42 (encoded by the nucleotide sequence set forth in SEQ ID NO:41). In some embodiments, the polypeptide capable of synthesizing *ent*-copalyl diphosphate from GGPP lacks a chloroplast transit peptide. In some embodiments, a recombinant host comprising a gene encoding a polypeptide capable of synthesizing *ent*-copalyl diphosphate from GGPP further comprises one or more genes encoding one or more polypeptides capable of synthesizing UTP from UDP (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:123), one or more genes encoding one or more polypeptides capable of converting glucose-6-phosphate to glucose-1-phosphate (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:119, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, and/or SEQ ID NO:147), and/or one or more genes encoding one or more polypeptides capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:121, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135,

SEQ ID NO:137, and/or SEQ ID NO:139). In some embodiments, the recombinant host is an *S. cerevisiae* host cell overexpressing one or more genes encoding one or more polypeptides involved in the UDP-glucose biosynthetic pathway (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:119, SEQ ID NO:121, and/or SEQ ID NO:123).

[00110] In some aspects, the polypeptide capable of synthesizing *ent*-kaurene from *ent*-copalyl diphosphate comprises a polypeptide having an amino acid sequence set forth in SEQ ID NO:44 (which can be encoded by the nucleotide sequence set forth in SEQ ID NO:43), SEQ ID NO:46 (encoded by the nucleotide sequence set forth in SEQ ID NO:45), SEQ ID NO:48 (encoded by the nucleotide sequence set forth in SEQ ID NO:47), SEQ ID NO:50 (encoded by the nucleotide sequence set forth in SEQ ID NO:49), or SEQ ID NO:52 (encoded by the nucleotide sequence set forth in SEQ ID NO:51). In some embodiments, a recombinant host comprising a gene encoding a polypeptide capable of synthesizing *ent*-kaurene from *ent*-copalyl diphosphate further comprises one or more genes encoding one or more polypeptides capable of synthesizing UTP from UDP (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:123), one or more genes encoding one or more polypeptides capable of converting glucose-6-phosphate to glucose-1-phosphate (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:119, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, and/or SEQ ID NO:147), and/or one or more genes encoding one or more polypeptides capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:121, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, and/or SEQ ID NO:139). In some embodiments, the recombinant host is an *S. cerevisiae* host cell overexpressing one or more genes encoding one or more polypeptides involved in the UDP-glucose biosynthetic pathway (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:119, SEQ ID NO:121, and/or SEQ ID NO:123).

[00111] In some embodiments, a recombinant host comprises a gene encoding a bifunctional polypeptide capable of synthesizing *ent*-copalyl diphosphate from GGPP and synthesizing *ent*-kaurene from *ent*-copalyl diphosphate. In some aspects, the bifunctional polypeptide comprises a polypeptide having an amino acid sequence set forth in SEQ ID NO:54 (which can be encoded by the nucleotide sequence set forth in SEQ ID NO:53), SEQ ID NO:56 (encoded by the nucleotide sequence set forth in SEQ ID NO:55), or SEQ ID NO:58 (encoded by the nucleotide sequence set forth in SEQ ID NO:57). In some embodiments, a recombinant host comprising a gene encoding a bifunctional polypeptide capable of synthesizing *ent*-copalyl

diphosphate from GGPP and synthesizing *ent*-kaurene from *ent*-copalyl diphosphate further comprises one or more genes encoding one or more polypeptides capable of synthesizing UTP from UDP (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:123), one or more genes encoding one or more polypeptides capable of converting glucose-6-phosphate to glucose-1-phosphate (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:119, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, and/or SEQ ID NO:147), and/or one or more genes encoding one or more polypeptides capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:121, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, and/or SEQ ID NO:139). In some embodiments, the recombinant host is an *S. cerevisiae* host cell overexpressing one or more genes encoding one or more polypeptides involved in the UDP-glucose biosynthetic pathway (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:119, SEQ ID NO:121, and/or SEQ ID NO:123).

[00112] In some aspects, the polypeptide capable of synthesizing ent-kaurenoic acid, ent-kaurenol, and/or ent-kaurenal from *ent*-kaurene comprises a polypeptide having an amino acid sequence set forth in SEQ ID NO:60 (which can be encoded by the nucleotide sequence set forth in SEQ ID NO:59), SEQ ID NO:62 (encoded by the nucleotide sequence set forth in SEQ ID NO:61), SEQ ID NO:117 (encoded by the nucleotide sequence set forth in SEQ ID NO:63 or SEQ ID NO:64), SEQ ID NO:66 (encoded by the nucleotide sequence set forth in SEQ ID NO:65), SEQ ID NO:68 (encoded by the nucleotide sequence set forth in SEQ ID NO:67), SEQ ID NO:70 (encoded by the nucleotide sequence set forth in SEQ ID NO:69), SEQ ID NO:72 (encoded by the nucleotide sequence set forth in SEQ ID NO:71), SEQ ID NO:74 (encoded by the nucleotide sequence set forth in SEQ ID NO:73), or SEQ ID NO:76 (encoded by the nucleotide sequence set forth in SEQ ID NO:75). In some embodiments, a recombinant host comprising a gene encoding a polypeptide capable of synthesizing ent-kaurenoic acid, ent-kaurenol, and/or ent-kaurenal from *ent*-kaurene further comprises one or more genes encoding one or more polypeptides capable of synthesizing UTP from UDP (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:123), one or more genes encoding one or more polypeptides capable of converting glucose-6-phosphate to glucose-1-phosphate (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:119, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, and/or SEQ ID NO:147), and/or one or more genes encoding one or more polypeptides capable of synthesizing UDP-glucose from UTP and

glucose-1-phosphate (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:121, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, and/or SEQ ID NO:139). In some embodiments, the recombinant host is an *S. cerevisiae* host cell overexpressing one or more genes encoding one or more polypeptides involved in the UDP-glucose biosynthetic pathway (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:119, SEQ ID NO:121, and/or SEQ ID NO:123).

[00113] In some aspects, the polypeptide capable of reducing cytochrome P450 complex comprises a polypeptide having an amino acid sequence set forth in SEQ ID NO:78 (which can be encoded by the nucleotide sequence set forth in SEQ ID NO:77), SEQ ID NO:80 (encoded by the nucleotide sequence set forth in SEQ ID NO:79), SEQ ID NO:82 (encoded by the nucleotide sequence set forth in SEQ ID NO:81), SEQ ID NO:84 (encoded by the nucleotide sequence set forth in SEQ ID NO:83), SEQ ID NO:86 (encoded by the nucleotide sequence set forth in SEQ ID NO:85), SEQ ID NO:88 (encoded by the nucleotide sequence set forth in SEQ ID NO:87), SEQ ID NO:90 (encoded by the nucleotide sequence set forth in SEQ ID NO:89), or SEQ ID NO:92 (encoded by the nucleotide sequence set forth in SEQ ID NO:91). In some embodiments, a recombinant host comprising a gene encoding a polypeptide capable of reducing cytochrome P450 complex further comprises one or more genes encoding one or more polypeptides capable of synthesizing UTP from UDP (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:123), one or more genes encoding one or more polypeptides capable of converting glucose-6-phosphate to glucose-1-phosphate (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:119, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, and/or SEQ ID NO:147), and/or one or more genes encoding one or more polypeptides capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:121, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, and/or SEQ ID NO:139). In some embodiments, the recombinant host is an *S. cerevisiae* host cell overexpressing one or more genes encoding one or more polypeptides involved in the UDP-glucose biosynthetic pathway (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:119, SEQ ID NO:121, and/or SEQ ID NO:123).

[00114] In some aspects, the polypeptide capable of synthesizing steviol from *ent*-kaurenoic acid comprises a polypeptide having an amino acid sequence set forth in SEQ ID NO:94 (which

can be encoded by the nucleotide sequence set forth in SEQ ID NO:93), SEQ ID NO:97 (encoded by the nucleotide sequence set forth in SEQ ID NO:95 or SEQ ID NO:96), SEQ ID NO:100 (encoded by the nucleotide sequence set forth in SEQ ID NO:98 or SEQ ID NO:99), SEQ ID NO:101, SEQ ID NO:102, SEQ ID NO:103, SEQ ID NO:104, SEQ ID NO:106 (encoded by the nucleotide sequence set forth in SEQ ID NO:105), SEQ ID NO:108 (encoded by the nucleotide sequence set forth in SEQ ID NO:107), SEQ ID NO:110 (encoded by the nucleotide sequence set forth in SEQ ID NO:109), SEQ ID NO:112 (encoded by the nucleotide sequence set forth in SEQ ID NO:111), or SEQ ID NO:114 (encoded by the nucleotide sequence set forth in SEQ ID NO:113). In some embodiments, a recombinant host comprising a gene encoding a polypeptide capable of synthesizing steviol from *ent*-kaurenoic acid further comprises one or more genes encoding one or more polypeptides capable of synthesizing UTP from UDP (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:123), one or more genes encoding one or more polypeptides capable of converting glucose-6-phosphate to glucose-1-phosphate (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:119, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, and/or SEQ ID NO:147), and/or one or more genes encoding one or more polypeptides capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:121, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, and/or SEQ ID NO:139). In some embodiments, the recombinant host is an *S. cerevisiae* host cell overexpressing one or more genes encoding one or more polypeptides involved in the UDP-glucose biosynthetic pathway (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:119, SEQ ID NO:121, and/or SEQ ID NO:123).

[00115] In some embodiments, a recombinant host comprises a nucleic acid encoding a polypeptide capable of glycosylating steviol or a steviol glycoside at its C-13 hydroxyl group (*e.g.*, UGT85C2 polypeptide) (SEQ ID NO:7), a nucleic acid encoding a polypeptide capable of beta 1,3 glycosylation of the C3' of the 13-*O*-glucose, 19-*O*-glucose, or both 13-*O*-glucose and 19-*O*-glucose of a steviol glycoside (*e.g.*, UGT76G1 polypeptide) (SEQ ID NO:9), a nucleic acid encoding a polypeptide capable of glycosylating steviol or a steviol glycoside at its C-19 carboxyl group (*e.g.*, UGT74G1 polypeptide) (SEQ ID NO:4), a nucleic acid encoding a polypeptide capable of beta 1,2 glycosylation of the C2' of the 13-*O*-glucose, 19-*O*-glucose, or both 13-*O*-glucose and 19-*O*-glucose of a steviol glycoside (*e.g.*, EUGT11 polypeptide) (SEQ ID NO:16). In some aspects, the polypeptide capable of beta 1,2 glycosylation of the C2' of the 13-

O-glucose, 19-*O*-glucose, or both 13-*O*-glucose and 19-*O*-glucose of a steviol glycoside (*e.g.*, UGT91D2 polypeptide) can be a UGT91D2e polypeptide (SEQ ID NO:11) or a UGT91D2e-b polypeptide (SEQ ID NO:13). In some embodiments, a recombinant host comprising a gene encoding a polypeptide capable of glycosylating steviol or a steviol glycoside further comprises one or more genes encoding one or more polypeptides capable of synthesizing UTP from UDP (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:123), one or more genes encoding one or more polypeptides capable of converting glucose-6-phosphate to glucose-1-phosphate (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:119, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, and/or SEQ ID NO:147), and/or one or more genes encoding one or more polypeptides capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:121, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, and/or SEQ ID NO:139). In some embodiments, the recombinant host is an *S. cerevisiae* host cell overexpressing one or more genes encoding one or more polypeptides involved in the UDP-glucose biosynthetic pathway (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:119, SEQ ID NO:121, and/or SEQ ID NO:123).

[00116] In some aspects, the polypeptide capable of glycosylating steviol or a steviol glycoside at its C-13 hydroxyl group is encoded by the nucleotide sequence set forth in SEQ ID NO:5 or SEQ ID NO:6, the polypeptide capable of beta 1,3 glycosylation of the C3' of the 13-*O*-glucose, 19-*O*-glucose, or both 13-*O*-glucose and 19-*O*-glucose of a steviol glycoside is encoded by the nucleotide sequence set forth in SEQ ID NO:8, the polypeptide capable of glycosylating steviol or a steviol glycoside at its C-19 carboxyl group is encoded by the nucleotide sequence set forth in SEQ ID NO:3, the polypeptide capable of beta 1,2 glycosylation of the C2' of the 13-*O*-glucose, 19-*O*-glucose, or both 13-*O*-glucose and 19-*O*-glucose of a steviol glycoside is encoded by the nucleotide sequence set forth in SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, or SEQ ID NO:15. The skilled worker will appreciate that expression of these genes may be necessary to produce a particular steviol glycoside but that one or more of these genes can be endogenous to the host provided that at least one (and in some embodiments, all) of these genes is a recombinant gene introduced into the recombinant host.

[00117] In some embodiments, expression of a recombinant gene encoding a polypeptide capable of synthesizing UTP from UDP, a recombinant gene encoding a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate, and a recombinant gene encoding a

polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate in a steviol glycoside-producing recombinant host increases the amount of one or more steviol glycosides, *e.g.*, rubusoside, RebB, RebA, RebD, and RebM, produced by the cell by at least about 5%, *e.g.*, at least about 10%, or at least about 15%, or at least about 20%, or at least about 25%, or at least about 30%, or at least about 35%, or at least about 40%, or at least about 45%, or at least about 50%, or at least about 60%, or at least about 70%, or at least about 80%, or at least about 90%, or at least about 100%, calculated as an increase in intracellular steviol glycoside concentration relative to a corresponding steviol glycoside-producing host lacking the recombinant genes.

[00118] For example, in some embodiments, expression of a recombinant gene encoding a polypeptide capable of synthesizing UTP from UDP (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:123), a recombinant gene encoding a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate (*e.g.* a polypeptide having the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:119, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, or SEQ ID NO:147), and a recombinant gene encoding a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:121, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, or SEQ ID NO:139) in a steviol glycoside-producing host increases the amount of one or more steviol glycosides, *e.g.*, rubusoside, RebB, RebA, RebD, and RebM, produced by the cell by at least about 5%, *e.g.*, at least about 10%, or at least about 15%, or at least about 20%, or at least about 25%, or at least about 30%, or at least about 35%, or at least about 40%, or at least about 45%, or at least about 50%, or at least about 60%, or at least about 70%, or at least about 80%, or at least about 90%, or at least about 100%, calculated as an increase in intracellular glycoside concentration relative to a corresponding steviol glycoside-producing host lacking the recombinant genes.

[00119] In some embodiments, expression of a recombinant gene encoding a polypeptide capable of synthesizing UTP from UDP, a recombinant gene encoding a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate, and a recombinant gene encoding a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate in a steviol glycoside-producing recombinant host decreases the amount of one or more steviol glycosides, *e.g.*, 13-SMG and RebD, produced by the cell by at least about 5%, *e.g.*, at least about 10%, or at least about 15%, or at least about 20%, or at least about 25%, or at least about

30%, or at least about 35%, or at least about 40%, or at least about 45%, or at least about 50%, calculated as a decrease in intracellular steviol glycoside concentration relative to a corresponding steviol glycoside-producing host lacking the recombinant genes.

[00120] For example, in some embodiments, expression of a recombinant gene encoding a polypeptide capable of synthesizing UTP from UDP having the amino acid sequence set forth in SEQ ID NO:123, a recombinant gene encoding a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate having the amino acid sequence set forth in SEQ ID NO:2, a recombinant gene encoding a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate having the amino acid sequence set forth in SEQ ID NO:119, a recombinant gene encoding a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate having the amino acid sequence set forth in SEQ ID NO:121, and further expression of a recombinant gene encoding a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate having the amino acid sequence set forth in, *e.g.*, SEQ ID NO:127, SEQ ID NO:133, SEQ ID NO:129, SEQ ID NO:125, SEQ ID NO:139, or SEQ ID NO:135, in a steviol glycoside-producing recombinant host decreases the amount of 13-SMG produced by the cell by at least about 5%, *e.g.*, at least about 7.5%, or at least about 10%, or at least about 15%, or at least about 20%, or at least about 25%, or at least about 30%, or at least about 35%.

[00121] In some embodiments, expression of a recombinant gene encoding a polypeptide capable of synthesizing UTP from UDP, a recombinant gene encoding a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate, and a recombinant gene encoding a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate in a steviol glycoside-producing recombinant host increases the total amount of steviol glycosides (*i.e.*, the total amount of mono-, di-, tri-, tetra- penta-, hexa-, and hepta-glycosylated steviol compounds) by at least about 5%, *e.g.*, at least about 7.5%, or at least about 10%, or at least about 12.5%, or at least about 15%, or at least about 17.5%, or at least about 20%, or at least about 25%, or at least about 27.5%, or at least about 30%, or at least about 35%, calculated as an increase in intracellular steviol glycoside concentration relative to a corresponding steviol glycoside-producing host lacking the recombinant genes.

[00122] For example, in some embodiments, expression of a recombinant gene encoding a polypeptide capable of synthesizing UTP from UDP having the amino acid sequence set forth in SEQ ID NO:123, a recombinant gene encoding a polypeptide capable of converting glucose-6-

phosphate to glucose-1-phosphate having the amino acid sequence set forth in SEQ ID NO:2, a recombinant gene encoding a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate having the amino acid sequence set forth in SEQ ID NO:119, a recombinant gene encoding a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate having the amino acid sequence set forth in SEQ ID NO:121, and further expression of a recombinant gene encoding a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate having the amino acid sequence set forth in, *e.g.*, SEQ ID NO:133, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:125, SEQ ID NO:139, or SEQ ID NO:135, in a steviol glycoside-producing recombinant host increases the total amount of steviol glycosides (*i.e.*, the total amount of mono-, di-, tri-, tetra- penta-, hexa-, and hepta-glycosylated steviol compounds) by at least about 5%, *e.g.*, at least about 7.5%, or at least about 10%, or at least about 12.5%, or at least about 15%, or at least about 17.5%, or at least about 20%, or at least about 25%, or at least about 27.5%, or at least about 30%, or at least about 35%, calculated as an increase in intracellular steviol glycoside concentration relative to a corresponding steviol glycoside-producing host lacking the recombinant genes.

[00123] In some other embodiments, the total amount of steviol glycosides produced by a steviol glycoside-producing recombinant host cell is unchanged (*i.e.*, increased or decreased by less than about 5%, or less than about 4%, or less than about 3%, or less than about 2%, or less than about 1%) by expression in the host of a recombinant gene encoding a polypeptide capable of synthesizing UTP from UDP, a recombinant gene encoding a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate, and/or a recombinant gene encoding a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate. For example, in some embodiments, expression of a recombinant gene encoding a polypeptide capable of synthesizing UTP from UDP having the amino acid sequence set forth in SEQ ID NO:123, a recombinant gene encoding a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate having the amino acid sequence set forth in SEQ ID NO:2, a recombinant gene encoding a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate having the amino acid sequence set forth in SEQ ID NO:119, a recombinant gene encoding a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate having the amino acid sequence set forth in SEQ ID NO:121 in a steviol glycoside-producing recombinant host increases the total amount of steviol glycosides produced by the host by less than about 5%, *e.g.*, less than about 4%, or less than about 3%, or less than about 2%.

[00124] The person of ordinary skill in the art will appreciate that, in such embodiments, expression of one or more genes encoding a polypeptide involved in the involved in the UDP-glucose biosynthetic pathway may affect the relative levels of steviol glycosides produced by the recombinant host, *e.g.*, by increasing the level of UDP-glucose available as a substrate for a polypeptide capable of glycosylating steviol or a steviol glycoside. For example, in some embodiments, expression of a recombinant gene encoding a polypeptide capable of synthesizing UTP from UDP having the amino acid sequence set forth in SEQ ID NO:123, a recombinant gene encoding a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate having the amino acid sequence set forth in SEQ ID NO:2, a recombinant gene encoding a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate having the amino acid sequence set forth in SEQ ID NO:119, a recombinant gene encoding a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate having the amino acid sequence set forth in SEQ ID NO:121 in a steviol glycoside-producing recombinant host increases the total amount of steviol glycosides produced by the host by less than about 5%, *e.g.*, less than about 4%, or less than about 3%, or less than about 2%, increases the amount of RebM produced by the host by at least about 50%, *e.g.*, at least about 60%, or at least about 70%, or at least about 80%, or at least about 90%, and decreases the amount of RebD produced by the host by at least about 10%, *e.g.*, at least about 20%, or at least about 30%, or at least about 40%.

[00125] In some embodiments, a recombinant host cell comprises one or more genes encoding one or more polypeptides capable of synthesizing UTP from UDP (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:123), one or more genes encoding one or more polypeptides capable of converting glucose-6-phosphate to glucose-1-phosphate (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:119, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, and/or SEQ ID NO:147), and/or one or more genes encoding one or more polypeptides capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:121, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, and/or SEQ ID NO:139).

[00126] In certain embodiments, a recombinant host comprises one or more recombinant genes having a nucleotide sequence native to the host that encode one or more polypeptides capable of synthesizing UTP from UDP, one or more polypeptides capable of converting glucose-6-phosphate to glucose-1-phosphate, and/or one or more polypeptides capable of

synthesizing UDP-glucose from UTP and glucose-1-phosphate, *i.e.*, a recombinant host overexpresses one or more polypeptides capable of synthesizing UTP from UDP, one or more polypeptides capable of converting glucose-6-phosphate to glucose-1-phosphate, and/or one or more polypeptides capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate.

[00127] In certain such embodiments, a recombinant host cell overexpresses one or more genes encoding one or more polypeptides capable of synthesizing UTP from UDP (*e.g.*, an *S. cerevisiae* host cell expressing a recombinant gene encoding a polypeptide having the amino acid sequence set forth in SEQ ID NO:123), one or more genes encoding one or more polypeptides capable of converting glucose-6-phosphate to glucose-1-phosphate (*e.g.*, an *S. cerevisiae* host cell expressing a recombinant gene encoding a polypeptide having the amino acid sequence set forth in SEQ ID NO:2, and/or SEQ ID NO:119), and/or one or more genes encoding one or more polypeptides capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate (*e.g.*, an *S. cerevisiae* host cell expressing a recombinant gene encoding a polypeptide having the amino acid sequence set forth in SEQ ID NO:121). In one example, a recombinant *S. cerevisiae* host cell overexpresses a gene encoding a polypeptide having the amino acid sequence set forth in SEQ ID NO:123, a gene encoding a polypeptide having the amino acid sequence set forth in SEQ ID NO:2, a gene encoding a polypeptide having the amino acid sequence set forth in SEQ ID NO:119, and a gene encoding a polypeptide having the amino acid sequence set forth in SEQ ID NO:121.

[00128] In certain embodiments, a recombinant host cell comprising one or more genes encoding one or more polypeptides capable of synthesizing UTP from UDP (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:123), one or more genes encoding one or more polypeptides capable of converting glucose-6-phosphate to glucose-1-phosphate (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:119, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, and/or SEQ ID NO:147), and/or one or more genes encoding one or more polypeptides capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:121, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, and/or SEQ ID NO:139), further comprises a gene encoding a polypeptide capable of glycosylating steviol or a steviol glycoside at its C-13 hydroxyl group (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:7); a gene encoding a polypeptide capable of beta 1,3 glycosylation of the C3' of the 13-*O*-glucose, 19-*O*-glucose, or both 13-*O*-glucose and 19-*O*-glucose of a steviol glycoside (*e.g.*, a polypeptide

having the amino acid sequence set forth in SEQ ID NO:9); a gene encoding a polypeptide capable of glycosylating steviol or a steviol glycoside at its C-19 carboxyl group (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:4); and/or a gene encoding a polypeptide capable of beta 1,2 glycosylation of the C2' of the 13-*O*-glucose, 19-*O*-glucose, or both 13-*O*-glucose and 19-*O*-glucose of a steviol glycoside (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:11, SEQ ID NO:13, or SEQ ID NO:16). In certain such embodiments, the recombinant host cell further comprises a gene encoding a polypeptide capable of synthesizing GGPP from FPP and IPP (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:20); a gene encoding a polypeptide capable of synthesizing *ent*-copalyl diphosphate from GGPP (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:40); a gene encoding a polypeptide capable of synthesizing *ent*-kaurene from *ent*-copalyl diphosphate (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:52); a gene encoding a polypeptide capable of synthesizing ent-kaurenoic acid, ent-kaurenol, and/or ent-kaurenal from *ent*-kaurene (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:60 or SEQ ID NO:117); a gene encoding a polypeptide capable of reducing cytochrome P450 complex (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:78, SEQ ID NO:86, or SEQ ID NO:92); and/or a gene encoding a polypeptide capable of synthesizing steviol from *ent*-kaurenoic acid (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:94).

[00129] In some embodiments, a recombinant host comprises two or more genes encoding two or more polypeptides capable of converting glucose-6-phosphate to glucose-1-phosphate (*e.g.*, two or more polypeptides having the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:119, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, and/or SEQ ID NO:147), and/or two or more genes encoding two or more polypeptides capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate (*e.g.*, two or more polypeptides having the amino acid sequence set forth in SEQ ID NO:121, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, and/or SEQ ID NO:139).

[00130] In certain such embodiments, a recombinant host comprises two or more genes encoding two or more polypeptides capable of converting glucose-6-phosphate to glucose-1-phosphate, *e.g.*, two or more genes encoding two or more polypeptides having the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:119, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, and/or SEQ ID NO:147. In one example, a recombinant host comprises a gene encoding a polypeptide having the amino acid sequence set forth in SEQ ID NO:2 and a

polypeptide having the amino acid sequence set forth in SEQ ID NO:119. In another example, a recombinant host comprises a gene encoding a polypeptide having the amino acid sequence set forth in SEQ ID NO:2, a polypeptide having the amino acid sequence set forth in SEQ ID NO:119, and a polypeptide having the amino acid sequence set forth in SEQ ID NO:145. In some embodiments, the recombinant host further comprises a gene encoding a polypeptide capable of synthesizing UTP from UDP (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:123) and/or one or more genes encoding one or more polypeptides capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:121, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, and/or SEQ ID NO:139).

[00131] In certain such embodiments, a recombinant host comprises two or more genes encoding two or more polypeptides capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate, *e.g.*, two or more genes encoding two or more polypeptides having the amino acid sequence set forth in SEQ ID NO:121, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, and/or SEQ ID NO:139. In one example, a recombinant host comprises a gene encoding a polypeptide having the amino acid sequence set forth in SEQ ID NO:121 and a polypeptide having the amino acid sequence set forth in SEQ ID NO:125. In another example, a recombinant host comprises a gene encoding a polypeptide having the amino acid sequence set forth in SEQ ID NO:121 and a polypeptide having the amino acid sequence set forth in SEQ ID NO:127. In another example, a recombinant host comprises a gene encoding a polypeptide having the amino acid sequence set forth in SEQ ID NO:121 and a polypeptide having the amino acid sequence set forth in SEQ ID NO:129. In another example, a recombinant host comprises a gene encoding a polypeptide having the amino acid sequence set forth in SEQ ID NO:121 and a polypeptide having the amino acid sequence set forth in SEQ ID NO:131. In another example, a recombinant host comprises a gene encoding a polypeptide having the amino acid sequence set forth in SEQ ID NO:121 and a gene encoding a gene encoding a polypeptide having the amino acid sequence set forth in SEQ ID NO:133. In another example, a recombinant host comprises a gene encoding a polypeptide having the amino acid sequence set forth in SEQ ID NO:121 and a gene encoding a polypeptide having the amino acid sequence set forth in SEQ ID NO:135. In another example, a recombinant host comprises a gene encoding a polypeptide having the amino acid sequence set forth in SEQ ID NO:121 and a gene encoding a polypeptide having the

amino acid sequence set forth in SEQ ID NO:137. In another example, a recombinant host comprises a gene encoding a polypeptide having the amino acid sequence set forth in SEQ ID NO:121 and a gene encoding a polypeptide having the amino acid sequence set forth in SEQ ID NO:139. In some embodiments, the recombinant host further comprises a gene encoding a polypeptide capable of synthesizing UTP from UDP (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:123) and/or one or more genes encoding one or more polypeptides capable of converting glucose-6-phosphate to glucose-1-phosphate (*e.g.*, one or more polypeptides having the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:119, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, and/or SEQ ID NO:147).

[00132] In certain such embodiments, a recombinant host comprising two or more genes encoding two or more polypeptides capable of converting glucose-6-phosphate to glucose-1-phosphate (*e.g.*, two or more polypeptides having the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:119, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, and/or SEQ ID NO:147), and/or two or more genes encoding two or more polypeptides capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate (*e.g.*, two or more polypeptides having the amino acid sequence set forth in SEQ ID NO:121, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, and/or SEQ ID NO:139) is a host cell overexpressing one or more genes encoding one or more polypeptides involved in the UDP-glucose biosynthetic pathway (*e.g.*, an *S. cerevisiae* host cell expressing one or more genes encoding one or more polypeptides having the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:119, SEQ ID NO:121, and/or SEQ ID NO:123).

[00133] In certain embodiments, a recombinant host cell comprising two or more genes encoding two or more polypeptides capable of converting glucose-6-phosphate to glucose-1-phosphate (*e.g.*, two or more polypeptides having the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:119, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, and/or SEQ ID NO:147), and/or two or more genes encoding two or more polypeptides capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate (*e.g.*, two or more polypeptides having the amino acid sequence set forth in SEQ ID NO:121, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, and/or SEQ ID NO:139), further comprises a gene encoding polypeptide capable of synthesizing UTP from UDP (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:123), a gene encoding a polypeptide capable of glycosylating steviol or a steviol glycoside at its C-13 hydroxyl group (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:7);

a gene encoding a polypeptide capable of beta 1,3 glycosylation of the C3' of the 13-*O*-glucose, 19-*O*-glucose, or both 13-*O*-glucose and 19-*O*-glucose of a steviol glycoside (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:9); a gene encoding a polypeptide capable of glycosylating steviol or a steviol glycoside at its C-19 carboxyl group (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:4); and/or a gene encoding a polypeptide capable of beta 1,2 glycosylation of the C2' of the 13-*O*-glucose, 19-*O*-glucose, or both 13-*O*-glucose and 19-*O*-glucose of a steviol glycoside (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:11, SEQ ID NO:13, or SEQ ID NO:16). In certain such embodiments, the recombinant host cell further comprises a gene encoding a polypeptide capable of synthesizing GGPP from FPP and IPP (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:20); a gene encoding a polypeptide capable of synthesizing *ent*-copalyl diphosphate from GGPP (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:40); a gene encoding a polypeptide capable of synthesizing *ent*-kaurene from *ent*-copalyl diphosphate (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:52); a gene encoding a polypeptide capable of synthesizing ent-kaurenoic acid, ent-kaurenol, and/or ent-kaurenal from *ent*-kaurene (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:60 or SEQ ID NO:117); a gene encoding a polypeptide capable of reducing cytochrome P450 complex (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:78, SEQ ID NO:86, or SEQ ID NO:92); and/or a gene encoding a polypeptide capable of synthesizing steviol from *ent*-kaurenoic acid (*e.g.*, a polypeptide having the amino acid sequence set forth in SEQ ID NO:94).

[00134] In some embodiments, a steviol glycoside or steviol glycoside precursor is produced by whole cell bioconversion. For whole cell bioconversion to occur, a host cell expressing one or more enzymes involved in the steviol glycoside pathway takes up and modifies a steviol glycoside precursor in the cell; following modification *in vivo*, a steviol glycoside remains in the cell and/or is excreted into the culture medium. For example, a host cell expressing a gene encoding a polypeptide capable of synthesizing UTP from UDP, a gene encoding a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate, and/or a gene encoding a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate; and further expressing a gene encoding a polypeptide capable of glycosylating steviol or a steviol glycoside at its C-13 hydroxyl group; a gene encoding a polypeptide capable of beta 1,3 glycosylation of the C3' of the 13-*O*-glucose, 19-*O*-glucose, or both 13-*O*-glucose and 19-*O*-glucose of a steviol glycoside; a gene encoding a polypeptide capable of glycosylating steviol or

a steviol glycoside at its C-19 carboxyl group; and/or a gene encoding a polypeptide capable of beta 1,2 glycosylation of the C2' of the 13-*O*-glucose, 19-*O*-glucose, or both 13-*O*-glucose and 19-*O*-glucose of a steviol glycoside can take up steviol and glycosylate steviol in the cell; following glycosylation *in vivo*, a steviol glycoside can be excreted into the culture medium. In certain such embodiments, the host cell may further express a gene encoding a polypeptide capable of synthesizing GGPP from FPP and IPP; a gene encoding a polypeptide capable of synthesizing *ent*-copalyl diphosphate from GGPP; a gene encoding a polypeptide capable of synthesizing *ent*-kaurene from *ent*-copalyl diphosphate; a gene encoding a polypeptide capable of synthesizing ent-kaurenoic acid, ent-kaurenol, and/or ent-kaurenal from *ent*-kaurene; a gene encoding a polypeptide capable of reducing cytochrome P450 complex; a gene encoding a polypeptide capable of synthesizing steviol from *ent*-kaurenoic acid; and/or a gene encoding a bifunctional polypeptide capable of synthesizing *ent*-copalyl diphosphate from GGPP and synthesizing *ent*-kaurene from *ent*-copalyl diphosphate.

[00135] In some embodiments, the method for producing one or more steviol glycosides or a steviol glycoside composition disclosed herein comprises whole-cell bioconversion of plant-derived or synthetic steviol and/or steviol glycosides in a cell culture medium of a recombinant host cell using: (a) a polypeptide capable of synthesizing UTP from UDP; (b) a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate; and/or (c) a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate, andone or more of: (d) a polypeptide capable of glycosylating steviol or a steviol glycoside at its C-13 hydroxyl group thereof; (e) a polypeptide capable of beta 1,3 glycosylation of the C3' of the 13-O-glucose, 19-O-glucose, or both 13-O-glucose and 19-O-glucose of a steviol glycoside; (f) a polypeptide capable of glycosylating steviol or a steviol glycoside at its C-19 carboxyl group thereof; and/or (g) a polypeptide capable of beta 1,2 glycosylation of the C2' of the 13-O-glucose, 19-O-glucose, or both 13-O-glucose and 19-O-glucose of a steviol glycoside; wherein at least one of the polypeptides is a recombinant polypeptide expressed in the recombinant host cell; and producing the one or more steviol glycosides or the steviol glycoside composition thereby.

[00136] In some embodiments of the methods for producing one or more steviol glycosides or a steviol glycoside composition disclosed herein comprises whole-cell bioconversion of plant-derived or synthetic steviol and/or steviol glycosides in a cell culture medium of a recombinant host cell disclosed herein, the polypeptide capable of synthesizing UTP from UDP comprises a polypeptide having at least 60% sequence identity to the amino acid sequence set forth in SEQ

ID NO:123; the polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate comprises a polypeptide having at least 60% sequence identity to the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:119, or SEQ ID NO:143; or at least 55% sequence identity to the amino acid sequence set forth in SEQ ID NO:141, SEQ ID NO:145, or SEQ ID NO:147; and/or the polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate comprises a polypeptide having at least 60% sequence identity to the amino acid sequence set forth in SEQ ID NO:121, SEQ ID NO:127; at least 55% sequence identity to the amino acid sequence set forth in SEQ ID NO:125, SEQ ID NO:129, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, or SEQ ID NO:139; or at least 70% sequence identity to the amino acid sequence set forth in SEQ ID NO:131.

[00137] In some embodiments, a polypeptide capable of synthesizing UTP from UDP, a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate, and/or a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate can be displayed on the surface of the recombinant host cells disclosed herein by fusing it with the anchoring motifs.

[00138] In some embodiments, the cell is permeabilized to take up a substrate to be modified or to excrete a modified product. In some embodiments, a permeabilizing agent can be added to aid the feedstock entering into the host and product getting out. In some embodiments, the cells are permeabilized with a solvent such as toluene, or with a detergent such as Triton-X or Tween. In some embodiments, the cells are permeabilized with a surfactant, for example a cationic surfactant such as cetyltrimethylammonium bromide (CTAB). In some embodiments, the cells are permeabilized with periodic mechanical shock such as electroporation or a slight osmotic shock. For example, a crude lysate of the cultured microorganism can be centrifuged to obtain a supernatant. The resulting supernatant can then be applied to a chromatography column, *e.g.*, a C18 column, and washed with water to remove hydrophilic compounds, followed by elution of the compound(s) of interest with a solvent such as methanol. The compound(s) can then be further purified by preparative HPLC. *See also*, WO 2009/140394.

[00139] In some embodiments, steviol, one or more steviol glycoside precursors, and/or one or more steviol glycosides are produced by co-culturing of two or more hosts. In some embodiments, one or more hosts, each expressing one or more enzymes involved in the steviol glycoside pathway, produce steviol, one or more steviol glycoside precursors, and/or one or more steviol glycosides. For example, a host expressing a gene encoding a polypeptide

capable of synthesizing GGPP from FPP and IPP; a gene encoding a polypeptide capable of synthesizing *ent*-copalyl diphosphate from GGPP; a gene encoding a polypeptide capable of synthesizing *ent*-kaurene from *ent*-copalyl diphosphate; a gene encoding a polypeptide capable of synthesizing ent-kaurenoic acid, ent-kaurenol, and/or ent-kaurenal from *ent*-kaurene; a gene encoding a polypeptide capable of reducing cytochrome P450 complex; a gene encoding a polypeptide capable of synthesizing steviol from *ent*-kaurenoic acid; and/or a gene encoding a bifunctional polypeptide capable of synthesizing *ent*-copalyl diphosphate from GGPP and synthesizing *ent*-kaurene from *ent*-copalyl diphosphate and a host expressing a gene encoding a polypeptide capable of synthesizing UTP from UDP, a gene encoding a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate, and/or a gene encoding a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate; and further expressing a gene encoding a polypeptide capable of glycosylating steviol or a steviol glycoside at its C-13 hydroxyl group; a gene encoding a polypeptide capable of beta 1,3 glycosylation of the C3' of the 13-*O*-glucose, 19-*O*-glucose, or both 13-*O*-glucose and 19-*O*-glucose of a steviol glycoside; a gene encoding a polypeptide capable of glycosylating steviol or a steviol glycoside at its C-19 carboxyl group; and/or a gene encoding a polypeptide capable of beta 1,2 glycosylation of the C2' of the 13-*O*-glucose, 19-*O*-glucose, or both 13-*O*-glucose and 19-*O*-glucose of a steviol glycoside, produce one or more steviol glycosides.

[00140] In some embodiments, the steviol glycoside comprises, for example, but not limited to, 13-SMG, steviol-1,2-bioside, steviol-1,3-bioside, 19-SMG, 1,2-stevioside, 1,3-stevioside (RebG), rubusoside, RebA, RebB, RebC, RebD, RebE, RebF, RebM, RebQ, RebI, dulcoside A, di-glycosylated steviol, tri-glycosylated steviol, tetra-glycosylated steviol, penta-glycosylated steviol, hexa-glycosylated steviol, hepta-glycosylated steviol, or isomers thereof.

[00141] In some embodiments, a steviol glycoside or steviol glycoside precursor composition produced *in vivo, in vitro*, or by whole cell bioconversion does not comprise or comprises a reduced amount or reduced level of plant-derived components than a Stevia extract from, *inter alia*, a Stevia plant. Plant-derived components can contribute to off-flavors and include pigments, lipids, proteins, phenolics, saccharides, spathulenol and other sesquiterpenes, labdane diterpenes, monoterpenes, decanoic acid, 8,11,14-eicosatrienoic acid, 2-methyloctadecane, pentacosane, octacosane, tetracosane, octadecanol, stigmasterol, β-sitosterol, α- and β-amyrin, lupeol, β-amryin acetate, pentacyclic triterpenes, centauredin, quercitin, epi-alpha-cadinol, carophyllenes and derivatives, beta-pinene, beta-sitosterol, and

gibberellin. In some embodiments, the plant-derived components referred to herein are non-glycoside compounds.

[00142] As used herein, the terms "detectable amount," "detectable concentration," "measurable amount," and "measurable concentration" refer to a level of steviol glycosides measured in AUC, $\mu M/OD_{600}$, mg/L, μM, or mM. Steviol glycoside production (*i.e.*, total, supernatant, and/or intracellular steviol glycoside levels) can be detected and/or analyzed by techniques generally available to one skilled in the art, for example, but not limited to, liquid chromatography-mass spectrometry (LC-MS), thin layer chromatography (TLC), high-performance liquid chromatography (HPLC), ultraviolet-visible spectroscopy/ spectrophotometry (UV-Vis), mass spectrometry (MS), and nuclear magnetic resonance spectroscopy (NMR).

[00143] As used herein, the term "undetectable concentration" refers to a level of a compound that is too low to be measured and/or analyzed by techniques such as TLC, HPLC, UV-Vis, MS, or NMR. In some embodiments, a compound of an "undetectable concentration" is not present in a steviol glycoside or steviol glycoside precursor composition.

[00144] After the recombinant microorganism has been grown in culture for the period of time, wherein the temperature and period of time facilitate the production of a steviol glycoside, steviol and/or one or more steviol glycosides can then be recovered from the culture using various techniques known in the art. Steviol glycosides can be isolated using a method described herein. For example, following fermentation, a culture broth can be centrifuged for 30 min at 7000 rpm at 4°C to remove cells, or cells can be removed by filtration. The cell-free lysate can be obtained, for example, by mechanical disruption or enzymatic disruption of the host cells and additional centrifugation to remove cell debris. Mechanical disruption of the dried broth materials can also be performed, such as by sonication. The dissolved or suspended broth materials can be filtered using a micron or sub-micron prior to further purification, such as by preparative chromatography. The fermentation media or cell-free lysate can optionally be treated to remove low molecular weight compounds such as salt; and can optionally be dried prior to purification and re-dissolved in a mixture of water and solvent.

[00145] The supernatant or cell-free lysate can be purified as follows: a column can be filled with, for example, HP20 Diaion resin (aromatic type Synthetic Adsorbent; Supelco) or other suitable non-polar adsorbent or reversed-phase chromatography resin, and an aliquot of supernatant or cell-free lysate can be loaded on to the column and washed with water to remove the hydrophilic components. The steviol glycoside product can be eluted by stepwise

incremental increases in the solvent concentration in water or a gradient from, *e. g.*, 0% → 100% methanol). The levels of steviol glycosides, glycosylated *ent*-kaurenol, and/or glycosylated *ent*-kaurenoic acid in each fraction, including the flow-through, can then be analyzed by LC-MS. Fractions can then be combined and reduced in volume using a vacuum evaporator. Additional purification steps can be utilized, if desired, such as additional chromatography steps and crystallization. For example, steviol glycosides can be isolated by methods not limited to ion exchange chromatography, reversed-phase chromatography (*i.e.*, using a C18 column), extraction, crystallization, and carbon columns and/or decoloring steps.

[00146] As used herein, the terms "or" and "and/or" is utilized to describe multiple components in combination or exclusive of one another. For example, "x, y, and/or z" can refer to "x" alone, "y" alone, "z" alone, "x, y, and z," "(x and y) or z," "x or (y and z)," or "x or y or z." In some embodiments, "and/or" is used to refer to the exogenous nucleic acids that a recombinant cell comprises, wherein a recombinant cell comprises one or more exogenous nucleic acids selected from a group. In some embodiments, "and/or" is used to refer to production of steviol glycosides and/or steviol glycoside precursors. In some embodiments, "and/or" is used to refer to production of steviol glycosides, wherein one or more steviol glycosides are produced. In some embodiments, "and/or" is used to refer to production of steviol glycosides, wherein one or more steviol glycosides are produced through one or more of the following steps: culturing a recombinant microorganism, synthesizing one or more steviol glycosides in a recombinant microorganism, and/or isolating one or more steviol glycosides.

Functional Homologs

[00147] Functional homologs of the polypeptides described above are also suitable for use in producing steviol glycosides in a recombinant host. A functional homolog is a polypeptide that has sequence similarity to a reference polypeptide, and that carries out one or more of the biochemical or physiological function(s) of the reference polypeptide. A functional homolog and the reference polypeptide can be a natural occurring polypeptide, and the sequence similarity can be due to convergent or divergent evolutionary events. As such, functional homologs are sometimes designated in the literature as homologs, or orthologs, or paralogs. Variants of a naturally occurring functional homolog, such as polypeptides encoded by mutants of a wild type coding sequence, can themselves be functional homologs. Functional homologs can also be

created via site-directed mutagenesis of the coding sequence for a polypeptide, or by combining domains from the coding sequences for different naturally-occurring polypeptides ("domain swapping"). Techniques for modifying genes encoding functional polypeptides described herein are known and include, *inter alia*, directed evolution techniques, site-directed mutagenesis techniques and random mutagenesis techniques, and can be useful to increase specific activity of a polypeptide, alter substrate specificity, alter expression levels, alter subcellular location, or modify polypeptide-polypeptide interactions in a desired manner. Such modified polypeptides are considered functional homologs. The term "functional homolog" is sometimes applied to the nucleic acid that encodes a functionally homologous polypeptide.

[00148] Functional homologs can be identified by analysis of nucleotide and polypeptide sequence alignments. For example, performing a query on a database of nucleotide or polypeptide sequences can identify homologs of steviol glycoside biosynthesis polypeptides. Sequence analysis can involve BLAST, Reciprocal BLAST, or PSI-BLAST analysis of non-redundant databases using a UGT amino acid sequence as the reference sequence. Amino acid sequence is, in some instances, deduced from the nucleotide sequence. Those polypeptides in the database that have greater than 40% sequence identity are candidates for further evaluation for suitability as a steviol glycoside biosynthesis polypeptide. Amino acid sequence similarity allows for conservative amino acid substitutions, such as substitution of one hydrophobic residue for another or substitution of one polar residue for another. If desired, manual inspection of such candidates can be carried out in order to narrow the number of candidates to be further evaluated. Manual inspection can be performed by selecting those candidates that appear to have domains present in steviol glycoside biosynthesis polypeptides, *e.g.*, conserved functional domains. In some embodiments, nucleic acids and polypeptides are identified from transcriptome data based on expression levels rather than by using BLAST analysis.

[00149] Conserved regions can be identified by locating a region within the primary amino acid sequence of a steviol glycoside biosynthesis polypeptide that is a repeated sequence, forms some secondary structure (*e.g.*, helices and beta sheets), establishes positively or negatively charged domains, or represents a protein motif or domain. *See, e.g.*, the Pfam web site describing consensus sequences for a variety of protein motifs and domains on the World Wide Web at sanger.ac.uk/Software/Pfam/ and pfam.janelia.org/. The information included at the Pfam database is described in Sonnhammer *et al.*, *Nucl. Acids Res.*, 26:320-322 (1998); Sonnhammer *et al.*, Proteins, 28:405-420 (1997); and Bateman *et al.*, *Nucl. Acids Res.*, 27:260-

262 (1999). Conserved regions also can be determined by aligning sequences of the same or related polypeptides from closely related species. Closely related species preferably are from the same family. In some embodiments, alignment of sequences from two different species is adequate to identify such homologs.

[00150] Typically, polypeptides that exhibit at least about 40% amino acid sequence identity are useful to identify conserved regions. Conserved regions of related polypeptides exhibit at least 45% amino acid sequence identity (*e.g.*, at least 50%, at least 60%, at least 70%, at least 80%, or at least 90% amino acid sequence identity). In some embodiments, a conserved region exhibits at least 92%, 94%, 96%, 98%, or 99% amino acid sequence identity.

[00151] For example, polypeptides suitable for producing steviol in a recombinant host include functional homologs of UGTs.

[00152] Methods to modify the substrate specificity of, for example, a UGT, are known to those skilled in the art, and include without limitation site-directed/rational mutagenesis approaches, random directed evolution approaches and combinations in which random mutagenesis/saturation techniques are performed near the active site of the enzyme. For example see Osmani *et al.*, 2009, *Phytochemistry* 70: 325–347.

[00153] A candidate sequence typically has a length that is from 80% to 200% of the length of the reference sequence, *e.g.*, 82, 85, 87, 89, 90, 93, 95, 97, 99, 100, 105, 110, 115, 120, 130, 140, 150, 160, 170, 180, 190, or 200% of the length of the reference sequence. A functional homolog polypeptide typically has a length that is from 95% to 105% of the length of the reference sequence, *e.g.*, 90, 93, 95, 97, 99, 100, 105, 110, 115, or 120% of the length of the reference sequence, or any range between. A % identity for any candidate nucleic acid or polypeptide relative to a reference nucleic acid or polypeptide can be determined as follows. A reference sequence (*e.g.*, a nucleic acid sequence or an amino acid sequence described herein) is aligned to one or more candidate sequences using the computer program Clustal Omega (version 1.2.1, default parameters), which allows alignments of nucleic acid or polypeptide sequences to be carried out across their entire length (global alignment). Chenna *et al.*, 2003, *Nucleic Acids Res.* 31(13):3497-500.

[00154] ClustalW calculates the best match between a reference and one or more candidate sequences, and aligns them so that identities, similarities and differences can be determined. Gaps of one or more residues can be inserted into a reference sequence, a candidate sequence, or both, to maximize sequence alignments. For fast pairwise alignment of nucleic

acid sequences, the following default parameters are used: word size: 2; window size: 4; scoring method: % age; number of top diagonals: 4; and gap penalty: 5. For multiple alignment of nucleic acid sequences, the following parameters are used: gap opening penalty: 10.0; gap extension penalty: 5.0; and weight transitions: yes. For fast pairwise alignment of protein sequences, the following parameters are used: word size: 1; window size: 5; scoring method:% age; number of top diagonals: 5; gap penalty: 3. For multiple alignment of protein sequences, the following parameters are used: weight matrix: blosum; gap opening penalty: 10.0; gap extension penalty: 0.05; hydrophilic gaps: on; hydrophilic residues: Gly, Pro, Ser, Asn, Asp, Gln, Glu, Arg, and Lys; residue-specific gap penalties: on. The ClustalW output is a sequence alignment that reflects the relationship between sequences. ClustalW can be run, for example, at the Baylor College of Medicine Search Launcher site on the World Wide Web (searchlauncher.bcm.tmc.edu/multi-align/multi-align.html) and at the European Bioinformatics Institute site on the World Wide Web (ebi.ac.uk/clustalw).

[00155] To determine a % identity of a candidate nucleic acid or amino acid sequence to a reference sequence, the sequences are aligned using Clustal Omega, the number of identical matches in the alignment is divided by the length of the reference sequence, and the result is multiplied by 100. It is noted that the% identity value can be rounded to the nearest tenth. For example, 78.11, 78.12, 78.13, and 78.14 are rounded down to 78.1, while 78.15, 78.16, 78.17, 78.18, and 78.19 are rounded up to 78.2.

[00156] It will be appreciated that functional UGT proteins (*e.g.*, a polypeptide capable of glycosylating steviol or a steviol glycoside at its C-19 carboxyl group) can include additional amino acids that are not involved in the enzymatic activities carried out by the enzymes. In some embodiments, UGT proteins are fusion proteins. The terms “chimera,” “fusion polypeptide,” “fusion protein,” “fusion enzyme,” “fusion construct,” “chimeric protein,” “chimeric polypeptide,” “chimeric construct,” and “chimeric enzyme” can be used interchangeably herein to refer to proteins engineered through the joining of two or more genes that code for different proteins. In some embodiments, a nucleic acid sequence encoding a UGT polypeptide (*e.g.*, a polypeptide capable of glycosylating steviol or a steviol glycoside at its C-19 carboxyl group) can include a tag sequence that encodes a “tag” designed to facilitate subsequent manipulation (*e.g.*, to facilitate purification or detection), secretion, or localization of the encoded polypeptide. Tag sequences can be inserted in the nucleic acid sequence encoding the polypeptide such that the encoded tag is located at either the carboxyl or amino terminus of the polypeptide. Non-limiting examples of encoded tags include green fluorescent protein (GFP), human influenza

hemagglutinin (HA), glutathione S transferase (GST), polyhistidine-tag (HIS tag), and Flag™ tag (Kodak, New Haven, CT). Other examples of tags include a chloroplast transit peptide, a mitochondrial transit peptide, an amyloplast peptide, signal peptide, or a secretion tag.

[00157] In some embodiments, a fusion protein is a protein altered by domain swapping. As used herein, the term "domain swapping" is used to describe the process of replacing a domain of a first protein with a domain of a second protein. In some embodiments, the domain of the first protein and the domain of the second protein are functionally identical or functionally similar. In some embodiments, the structure and/or sequence of the domain of the second protein differs from the structure and/or sequence of the domain of the first protein. In some embodiments, a UGT polypeptide (*e.g.*, a polypeptide capable of glycosylating steviol or a steviol glycoside at its C-19 carboxyl group) is altered by domain swapping.

[00158] In some embodiments, a fusion protein is a protein altered by circular permutation, which consists in the covalent attachment of the ends of a protein that would be opened elsewhere afterwards. Thus, the order of the sequence is altered without causing changes in the amino acids of the protein. In some embodiments, a targeted circular permutation can be produced, for example but not limited to, by designing a spacer to join the ends of the original protein. Once the spacer has been defined, there are several possibilities to generate permutations through generally accepted molecular biology techniques, for example but not limited to, by producing concatemers by means of PCR and subsequent amplification of specific permutations inside the concatemer or by amplifying discrete fragments of the protein to exchange to join them in a different order. The step of generating permutations can be followed by creating a circular gene by binding the fragment ends and cutting back at random, thus forming collections of permutations from a unique construct. In some embodiments, DAP1 polypeptide is altered by circular permutation.

Steviol and Steviol Glycoside Biosynthesis Nucleic Acids

[00159] A recombinant gene encoding a polypeptide described herein comprises the coding sequence for that polypeptide, operably linked in sense orientation to one or more regulatory regions suitable for expressing the polypeptide. Because many microorganisms are capable of expressing multiple gene products from a polycistronic mRNA, multiple polypeptides can be expressed under the control of a single regulatory region for those microorganisms, if desired.

A coding sequence and a regulatory region are considered to be operably linked when the regulatory region and coding sequence are positioned so that the regulatory region is effective for regulating transcription or translation of the sequence. Typically, the translation initiation site of the translational reading frame of the coding sequence is positioned between one and about fifty nucleotides downstream of the regulatory region for a monocistronic gene.

[00160] In many cases, the coding sequence for a polypeptide described herein is identified in a species other than the recombinant host, *i.e.*, is a heterologous nucleic acid. Thus, if the recombinant host is a microorganism, the coding sequence can be from other prokaryotic or eukaryotic microorganisms, from plants or from animals. In some case, however, the coding sequence is a sequence that is native to the host and is being reintroduced into that organism. A native sequence can often be distinguished from the naturally occurring sequence by the presence of non-natural sequences linked to the exogenous nucleic acid, *e.g.*, non-native regulatory sequences flanking a native sequence in a recombinant nucleic acid construct. In addition, stably transformed exogenous nucleic acids typically are integrated at positions other than the position where the native sequence is found. "Regulatory region" refers to a nucleic acid having nucleotide sequences that influence transcription or translation initiation and rate, and stability and/or mobility of a transcription or translation product. Regulatory regions include, without limitation, promoter sequences, enhancer sequences, response elements, protein recognition sites, inducible elements, protein binding sequences, 5′ and 3′ untranslated regions (UTRs), transcriptional start sites, termination sequences, polyadenylation sequences, introns, and combinations thereof. A regulatory region typically comprises at least a core (basal) promoter. A regulatory region also may include at least one control element, such as an enhancer sequence, an upstream element or an upstream activation region (UAR). A regulatory region is operably linked to a coding sequence by positioning the regulatory region and the coding sequence so that the regulatory region is effective for regulating transcription or translation of the sequence. For example, to operably link a coding sequence and a promoter sequence, the translation initiation site of the translational reading frame of the coding sequence is typically positioned between one and about fifty nucleotides downstream of the promoter. A regulatory region can, however, be positioned as much as about 5,000 nucleotides upstream of the translation initiation site, or about 2,000 nucleotides upstream of the transcription start site.

[00161] The choice of regulatory regions to be included depends upon several factors, including, but not limited to, efficiency, selectability, inducibility, desired expression level, and preferential expression during certain culture stages. It is a routine matter for one of skill in the

art to modulate the expression of a coding sequence by appropriately selecting and positioning regulatory regions relative to the coding sequence. It will be understood that more than one regulatory region may be present, *e.g.*, introns, enhancers, upstream activation regions, transcription terminators, and inducible elements.

[00162] One or more genes can be combined in a recombinant nucleic acid construct in "modules" useful for a discrete aspect of steviol and/or steviol glycoside production. Combining a plurality of genes in a module, particularly a polycistronic module, facilitates the use of the module in a variety of species. For example, a steviol biosynthesis gene cluster, or a UGT gene cluster, can be combined in a polycistronic module such that, after insertion of a suitable regulatory region, the module can be introduced into a wide variety of species. As another example, a UGT gene cluster can be combined such that each UGT coding sequence is operably linked to a separate regulatory region, to form a UGT module. Such a module can be used in those species for which monocistronic expression is necessary or desirable. In addition to genes useful for steviol or steviol glycoside production, a recombinant construct typically also contains an origin of replication, and one or more selectable markers for maintenance of the construct in appropriate species.

[00163] It will be appreciated that because of the degeneracy of the genetic code, a number of nucleic acids can encode a particular polypeptide; *i.e.*, for many amino acids, there is more than one nucleotide triplet that serves as the codon for the amino acid. Thus, codons in the coding sequence for a given polypeptide can be modified such that optimal expression in a particular host is obtained, using appropriate codon bias tables for that host (*e.g.*, microorganism). As isolated nucleic acids, these modified sequences can exist as purified molecules and can be incorporated into a vector or a virus for use in constructing modules for recombinant nucleic acid constructs.

[00164] In some cases, it is desirable to inhibit one or more functions of an endogenous polypeptide in order to divert metabolic intermediates towards steviol or steviol glycoside biosynthesis. For example, it may be desirable to downregulate synthesis of sterols in a yeast strain in order to further increase steviol or steviol glycoside production, e.g., by downregulating squalene epoxidase. As another example, it may be desirable to inhibit degradative functions of certain endogenous gene products, *e.g.*, glycohydrolases that remove glucose moieties from secondary metabolites or phosphatases as discussed herein. In such cases, a nucleic acid that overexpresses the polypeptide or gene product may be included in a recombinant construct that

is transformed into the strain. Alternatively, mutagenesis can be used to generate mutants in genes for which it is desired to increase or enhance function.

Host Microorganisms

[00165] Recombinant hosts can be used to express polypeptides for the producing steviol glycosides, including mammalian, insect, plant, and algal cells. A number of prokaryotes and eukaryotes are also suitable for use in constructing the recombinant microorganisms described herein, *e.g.*, gram-negative bacteria, yeast, and fungi. A species and strain selected for use as a steviol glycoside production strain is first analyzed to determine which production genes are endogenous to the strain and which genes are not present. Genes for which an endogenous counterpart is not present in the strain are advantageously assembled in one or more recombinant constructs, which are then transformed into the strain in order to supply the missing function(s).

[00166] Typically, the recombinant microorganism is grown in a fermenter at a temperature(s) for a period of time, wherein the temperature and period of time facilitate the production of a steviol glycoside. The constructed and genetically engineered microorganisms provided by the invention can be cultivated using conventional fermentation processes, including, *inter alia*, chemostat, batch, fed-batch cultivations, semi-continuous fermentations such as draw and fill, continuous perfusion fermentation, and continuous perfusion cell culture. Depending on the particular microorganism used in the method, other recombinant genes such as isopentenyl biosynthesis genes and terpene synthase and cyclase genes may also be present and expressed. Levels of substrates and intermediates, *e.g.*, isopentenyl diphosphate, dimethylallyl diphosphate, GGPP, *ent*-kaurene and *ent*-kaurenoic acid, can be determined by extracting samples from culture media for analysis according to published methods.

[00167] Carbon sources of use in the instant method include any molecule that can be metabolized by the recombinant host cell to facilitate growth and/or production of the steviol glycosides. Examples of suitable carbon sources include, but are not limited to, sucrose (*e.g.*, as found in molasses), fructose, xylose, ethanol, glycerol, glucose, cellulose, starch, cellobiose or other glucose-comprising polymer. In embodiments employing yeast as a host, for example, carbons sources such as sucrose, fructose, xylose, ethanol, glycerol, and glucose are suitable. The carbon source can be provided to the host organism throughout the cultivation period or

alternatively, the organism can be grown for a period of time in the presence of another energy source, *e.g.*, protein, and then provided with a source of carbon only during the fed-batch phase.

[00168] It will be appreciated that the various genes and modules discussed herein can be present in two or more recombinant hosts rather than a single host. When a plurality of recombinant hosts is used, they can be grown in a mixed culture to accumulate steviol and/or steviol glycosides.

[00169] Alternatively, the two or more hosts each can be grown in a separate culture medium and the product of the first culture medium, *e.g.*, steviol, can be introduced into second culture medium to be converted into a subsequent intermediate, or into an end product such as, for example, RebA. The product produced by the second, or final host is then recovered. It will also be appreciated that in some embodiments, a recombinant host is grown using nutrient sources other than a culture medium and utilizing a system other than a fermenter.

[00170] Exemplary prokaryotic and eukaryotic species are described in more detail below. However, it will be appreciated that other species can be suitable. For example, suitable species can be in a genus such as *Agaricus*, *Aspergillus*, *Bacillus*, *Candida*, *Corynebacterium*, *Eremothecium*, *Escherichia*, *Fusarium/Gibberella*, *Kluyveromyces*, *Laetiporus*, *Lentinus*, *Phaffia*, *Phanerochaete*, *Pichia*, *Physcomitrella*, *Rhodoturula*, *Saccharomyces*, *Schizosaccharomyces*, *Sphaceloma*, *Xanthophyllomyces* or *Yarrowia*. Exemplary species from such genera include *Lentinus tigrinus*, *Laetiporus sulphureus*, *Phanerochaete chrysosporium*, *Pichia pastoris*, *Cyberlindnera jadinii*, *Physcomitrella patens*, *Rhodoturula glutinis*, *Rhodoturula mucilaginosa*, *Phaffia rhodozyma*, *Xanthophyllomyces dendrorhous*, *Fusarium fujikuroi/Gibberella fujikuroi*, *Candida utilis*, *Candida glabrata*, *Candida albicans*, and *Yarrowia lipolytica*.

[00171] In some embodiments, a microorganism can be a prokaryote such as *Escherichia* bacteria cells, for example, *Escherichia coli* cells; *Lactobacillus* bacteria cells; *Lactococcus* bacteria cells; *Comebacterium* bacteria cells; *Acetobacter* bacteria cells; *Acinetobacter* bacteria cells; or *Pseudomonas* bacterial cells.

[00172] In some embodiments, a microorganism can be an Ascomycete such as *Gibberella fujikuroi*, *Kluyveromyces lactis*, *Schizosaccharomyces pombe*, *Aspergillus niger*, *Yarrowia lipolytica*, *Ashbya gossypii*, or *S. cerevisiae*.

[00173] In some embodiments, a microorganism can be an algal cell such as *Blakeslea trispora, Dunaliella salina, Haematococcus pluvialis, Chlorella sp., Undaria pinnatifida, Sargassum, Laminaria japonica, Scenedesmus almeriensis* species.

[00174] In some embodiments, a microorganism can be a cyanobacterial cell such as *Blakeslea trispora, Dunaliella salina, Haematococcus pluvialis, Chlorella sp., Undaria pinnatifida, Sargassum, Laminaria japonica, Scenedesmus almeriensis*.

Saccharomyces spp.

[00175] S*accharomyces* is a widely used chassis organism in synthetic biology, and can be used as the recombinant microorganism platform. For example, there are libraries of mutants, plasmids, detailed computer models of metabolism and other information available for *S. cerevisiae*, allowing for rational design of various modules to enhance product yield. Methods are known for making recombinant microorganisms.

Aspergillus spp.

[00176] *Aspergillus* species such as *A. oryzae*, *A. niger* and *A. sojae* are widely used microorganisms in food production and can also be used as the recombinant microorganism platform. Nucleotide sequences are available for genomes of *A. nidulans*, *A. fumigatus*, *A. oryzae*, *A. clavatus*, *A. flavus*, *A. niger*, and *A. terreus*, allowing rational design and modification of endogenous pathways to enhance flux and increase product yield. Metabolic models have been developed for *Aspergillus*, as well as transcriptomic studies and proteomics studies. *A. niger* is cultured for the industrial production of a number of food ingredients such as citric acid and gluconic acid, and thus species such as *A. niger* are generally suitable for producing steviol glycosides.

E. coli

[00177] *E. coli*, another widely used platform organism in synthetic biology, can also be used as the recombinant microorganism platform. Similar to *Saccharomyces*, there are libraries of mutants, plasmids, detailed computer models of metabolism and other information available for *E. coli*, allowing for rational design of various modules to enhance product yield. Methods similar to those described above for *Saccharomyces* can be used to make recombinant *E. coli* microorganisms.

Agaricus, Gibberella, and Phanerochaete spp.

[00178] *Agaricus*, *Gibberella*, and *Phanerochaete* spp. can be useful because they are known to produce large amounts of isoprenoids in culture. Thus, the terpene precursors for producing large amounts of steviol glycosides are already produced by endogenous genes. Thus, modules comprising recombinant genes for steviol glycoside biosynthesis polypeptides can be introduced into species from such genera without the necessity of introducing mevalonate or MEP pathway genes.

Arxula adeninivorans (Blastobotrys adeninivorans)

[00179] *Arxula adeninivorans* is dimorphic yeast (it grows as budding yeast like the baker's yeast up to a temperature of 42°C, above this threshold it grows in a filamentous form) with unusual biochemical characteristics. It can grow on a wide range of substrates and can assimilate nitrate. It has successfully been applied to the generation of strains that can produce natural plastics or the development of a biosensor for estrogens in environmental samples.

Yarrowia lipolytica

[00180] *Yarrowia lipolytica* is dimorphic yeast (*see Arxula adeninivorans*) and belongs to the family Hemiascomycetes. The entire genome of *Yarrowia lipolytica* is known. *Yarrowia* species is aerobic and considered to be non-pathogenic. *Yarrowia* is efficient in using hydrophobic substrates (*e.g.,* alkanes, fatty acids, oils) and can grow on sugars. It has a high potential for industrial applications and is an oleaginous microorgamism. *Yarrowia lipolyptica* can accumulate lipid content to approximately 40% of its dry cell weight and is a model organism for lipid accumulation and remobilization. *See e.g.,* Nicaud, 2012, *Yeast* 29(10):409-18; Beopoulos *et al.,* 2009, *Biochimie* 91(6):692-6; Bankar *et al.*, 2009, *Appl Microbiol Biotechnol.* 84(5):847-65.

Rhodotorula sp.

[00181] *Rhodotorula* is unicellular, pigmented yeast. The oleaginous red yeast, *Rhodotorula glutinis*, has been shown to produce lipids and carotenoids from crude glycerol (Saenge *et al.*, 2011, *Process Biochemistry* 46(1):210-8). *Rhodotorula toruloides* strains have been shown to be an efficient fed-batch fermentation system for improved biomass and lipid productivity (Li *et al.,* 2007, *Enzyme and Microbial Technology* 41:312-7).

Rhodosporidium toruloides

[00182] *Rhodosporidium toruloides* is oleaginous yeast and useful for engineering lipid-production pathways (*See e.g.* Zhu *et al.*, 2013, *Nature Commun.* 3:1112; Ageitos *et al.*, 2011, *Applied Microbiology and Biotechnology* 90(4):1219-27).

Candida boidinii

[00183] *Candida boidinii* is methylotrophic yeast (it can grow on methanol). Like other methylotrophic species such as *Hansenula polymorpha* and *Pichia pastoris*, it provides an excellent platform for producing heterologous proteins. Yields in a multigram range of a secreted foreign protein have been reported. A computational method, IPRO, recently predicted mutations that experimentally switched the cofactor specificity of *Candida boidinii* xylose reductase from NADPH to NADH. *See*, *e.g.*, Mattanovich *et al.*, 2012, *Methods Mol Biol.* 824:329-58; Khoury *et al.*, 2009, *Protein Sci.* 18(10):2125-38.

Hansenula polymorpha (Pichia angusta)

[00184] *Hansenula polymorpha* is methylotrophic yeast (*see Candida boidinii*). It can furthermore grow on a wide range of other substrates; it is thermo-tolerant and can assimilate nitrate (see also *Kluyveromyces lactis*). It has been applied to producing hepatitis B vaccines, insulin and interferon alpha-2a for the treatment of hepatitis C, furthermore to a range of technical enzymes. *See*, *e.g.*, Xu *et al.*, 2014, *Virol Sin.* 29(6):403-9.

Kluyveromyces lactis

[00185] *Kluyveromyces lactis* is yeast regularly applied to the production of kefir. It can grow on several sugars, most importantly on lactose which is present in milk and whey. It has successfully been applied among others for producing chymosin (an enzyme that is usually present in the stomach of calves) for producing cheese. Production takes place in fermenters on a 40,000 L scale. *See*, *e.g.*, van Ooyen *et al.*, 2006, *FEMS Yeast Res.* 6(3):381-92.

Pichia pastoris

[00186] *Pichia pastoris* is methylotrophic yeast (*see Candida boidinii* and *Hansenula polymorpha*). It provides an efficient platform for producing foreign proteins. Platform elements are available as a kit and it is worldwide used in academia for producing proteins. Strains have been engineered that can produce complex human N-glycan (yeast glycans are similar but not identical to those found in humans). *See*, *e.g.*, Piirainen *et al.*, 2014, *N Biotechnol.* 31(6):532-7.

Physcomitrella spp.

[00187] *Physcomitrella mosses*, when grown in suspension culture, have characteristics similar to yeast or other fungal cultures. This genera can be used for producing plant secondary metabolites, which can be difficult to produce in other types of cells.

[00188] It can be appreciated that the recombinant host cell disclosed herein can comprise a plant cell, comprising a plant cell that is grown in a plant, a mammalian cell, an insect cell, a fungal cell, comprising a yeast cell, wherein the yeast cell is a cell from *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Yarrowia lipolytica*, *Candida glabrata*, *Ashbya gossypii*, *Cyberlindnera jadinii*, *Pichia pastoris*, *Kluyveromyces lactis*, *Hansenula polymorpha*, *Candida boidinii*, *Arxula adeninivorans*, *Xanthophyllomyces dendrorhous*, or *Candida albicans species* or is a Saccharomycete or is a Saccharomyces cerevisiae cell, an algal cell or a bacterial cell, comprising *Escherichia* cells, *Lactobacillus* cells, *Lactococcus* cells, *Cornebacterium* cells, *Acetobacter* cells, *Acinetobacter* cells, or *Pseudomonas* cells.

Steviol Glycoside Compositions

[00189] Steviol glycosides do not necessarily have equivalent performance in different food systems. It is therefore desirable to have the ability to direct the synthesis to steviol glycoside compositions of choice. Recombinant hosts described herein can produce compositions that are selectively enriched for specific steviol glycosides (*e.g.*, RebD or RebM) and have a consistent taste profile. As used herein, the term “enriched” is used to describe a steviol glycoside composition with an increased proportion of a particular steviol glycoside, compared to a steviol glycoside composition (extract) from a stevia plant. Thus, the recombinant hosts described herein can facilitate the production of compositions that are tailored to meet the sweetening profile desired for a given food product and that have a proportion of each steviol glycoside that is consistent from batch to batch. In some embodiments, hosts described herein do not produce or produce a reduced amount of undesired plant by-products found in Stevia extracts. Thus, steviol glycoside compositions produced by the recombinant hosts described herein are distinguishable from compositions derived from *Stevia* plants.

[00190] The amount of an individual steviol glycoside (*e.g.*, RebA, RebB, RebD, or RebM) accumulated can be from about 1 to about 7,000 mg/L, *e.g.*, about 1 to about 10 mg/L, about 3 to about 10 mg/L, about 5 to about 20 mg/L, about 10 to about 50 mg/L, about 10 to about 100 mg/L, about 25 to about 500 mg/L, about 100 to about 1,500 mg/L, or about 200 to about 1,000

mg/L, at least about 1,000 mg/L, at least about 1,200 mg/L, at least about at least 1,400 mg/L, at least about 1,600 mg/L, at least about 1,800 mg/L, at least about 2,800 mg/L, or at least about 7,000 mg/L. In some aspects, the amount of an individual steviol glycoside can exceed 7,000 mg/L. The amount of a combination of steviol glycosides (*e.g.*, RebA, RebB, RebD, or RebM) accumulated can be from about 1 mg/L to about 7,000 mg/L, *e.g.*, about 200 to about 1,500, at least about 2,000 mg/L, at least about 3,000 mg/L, at least about 4,000 mg/L, at least about 5,000 mg/L, at least about 6,000 mg/L, or at least about 7,000 mg/L. In some aspects, the amount of a combination of steviol glycosides can exceed 7,000 mg/L. In general, longer culture times will lead to greater amounts of product. Thus, the recombinant microorganism can be cultured for from 1 day to 7 days, from 1 day to 5 days, from 3 days to 5 days, about 3 days, about 4 days, or about 5 days.

[00191] It will be appreciated that the various genes and modules discussed herein can be present in two or more recombinant microorganisms rather than a single microorganism. When a plurality of recombinant microorganisms is used, they can be grown in a mixed culture to produce steviol and/or steviol glycosides. For example, a first microorganism can comprise one or more biosynthesis genes for producing a steviol glycoside precursor, while a second microorganism comprises steviol glycoside biosynthesis genes. The product produced by the second, or final microorganism is then recovered. It will also be appreciated that in some embodiments, a recombinant microorganism is grown using nutrient sources other than a culture medium and utilizing a system other than a fermenter.

[00192] Alternatively, the two or more microorganisms each can be grown in a separate culture medium and the product of the first culture medium, *e.g.*, steviol, can be introduced into second culture medium to be converted into a subsequent intermediate, or into an end product such as RebA. The product produced by the second, or final microorganism is then recovered. It will also be appreciated that in some embodiments, a recombinant microorganism is grown using nutrient sources other than a culture medium and utilizing a system other than a fermenter.

[00193] Steviol glycosides and compositions obtained by the methods disclosed herein can be used to make food products, dietary supplements and sweetener compositions. *See*, *e.g.*, WO 2011/153378, WO 2013/022989, WO 2014/122227, and WO 2014/122328.

[00194] For example, substantially pure steviol or steviol glycoside such as RebM or RebD can be included in food products such as ice cream, carbonated 2s, fruit juices, yogurts, baked

goods, chewing gums, hard and soft candies, and sauces. Substantially pure steviol or steviol glycoside can also be included in non-food products such as pharmaceutical products, medicinal products, dietary supplements and nutritional supplements. Substantially pure steviol or steviol glycosides may also be included in animal feed products for both the agriculture industry and the companion animal industry. Alternatively, a mixture of steviol and/or steviol glycosides can be made by culturing recombinant microorganisms separately, each producing a specific steviol or steviol glycoside, recovering the steviol or steviol glycoside in substantially pure form from each microorganism and then combining the compounds to obtain a mixture comprising each compound in the desired proportion. The recombinant microorganisms described herein permit more precise and consistent mixtures to be obtained compared to current Stevia products.

[00195] In another alternative, a substantially pure steviol or steviol glycoside can be incorporated into a food product along with other sweeteners, *e.g.,* saccharin, dextrose, sucrose, fructose, erythritol, aspartame, sucralose, monatin, or acesulfame potassium. The weight ratio of steviol or steviol glycoside relative to other sweeteners can be varied as desired to achieve a satisfactory taste in the final food product. See, *e.g.*, U.S. 2007/0128311. In some embodiments, the steviol or steviol glycoside may be provided with a flavor (*e.g.*, citrus) as a flavor modulator.

[00196] Compositions produced by a recombinant microorganism described herein can be incorporated into food products. For example, a steviol glycoside composition produced by a recombinant microorganism can be incorporated into a food product in an amount ranging from about 20 mg steviol glycoside/kg food product to about 1800 mg steviol glycoside/kg food product on a dry weight basis, depending on the type of steviol glycoside and food product. For example, a steviol glycoside composition produced by a recombinant microorganism can be incorporated into a dessert, cold confectionary (*e.g.*, ice cream), dairy product (*e.g.*, yogurt), or beverage (*e.g.*, a carbonated beverage) such that the food product has a maximum of 500 mg steviol glycoside/kg food on a dry weight basis. A steviol glycoside composition produced by a recombinant microorganism can be incorporated into a baked good (*e.g.*, a biscuit) such that the food product has a maximum of 300 mg steviol glycoside/kg food on a dry weight basis. A steviol glycoside composition produced by a recombinant microorganism can be incorporated into a sauce (*e.g.*, chocolate syrup) or vegetable product (*e.g.*, pickles) such that the food product has a maximum of 1000 mg steviol glycoside/kg food on a dry weight basis. A steviol glycoside composition produced by a recombinant microorganism can be incorporated into

bread such that the food product has a maximum of 160 mg steviol glycoside/kg food on a dry weight basis. A steviol glycoside composition produced by a recombinant microorganism, plant, or plant cell can be incorporated into a hard or soft candy such that the food product has a maximum of 1600 mg steviol glycoside/kg food on a dry weight basis. A steviol glycoside composition produced by a recombinant microorganism, plant, or plant cell can be incorporated into a processed fruit product (*e.g.*, fruit juices, fruit filling, jams, and jellies) such that the food product has a maximum of 1000 mg steviol glycoside/kg food on a dry weight basis. In some embodiments, a steviol glycoside composition produced herein is a component of a pharmaceutical composition. *See*, *e.g.*, Steviol Glycosides Chemical and Technical Assessment 69th JECFA, 2007, prepared by Harriet Wallin, Food Agric. Org.; EFSA Panel on Food Additives and Nutrient Sources added to Food (ANS), "Scientific Opinion on the safety of steviol glycosides for the proposed uses as a food additive," 2010, *EFSA Journal* 8(4):1537; U.S. Food and Drug Administration GRAS Notice 323; U.S Food and Drug Administration GRAS Notice 329; WO 2011/037959; WO 2010/146463; WO 2011/046423; and WO 2011/056834.

[00197] For example, such a steviol glycoside composition can have from 90-99 weight % RebA and an undetectable amount of stevia plant-derived contaminants, and be incorporated into a food product at from 25-1600 mg/kg, *e.g.*, 100-500 mg/kg, 25-100 mg/kg, 250-1000 mg/kg, 50-500 mg/kg or 500-1000 mg/kg on a dry weight basis.

[00198] Such a steviol glycoside composition can be a RebB-enriched composition having greater than 3 weight % RebB and be incorporated into the food product such that the amount of RebB in the product is from 25-1600 mg/kg, *e.g.*, 100-500 mg/kg, 25-100 mg/kg, 250-1000 mg/kg, 50-500 mg/kg or 500-1000 mg/kg on a dry weight basis. Typically, the RebB-enriched composition has an undetectable amount of stevia plant-derived contaminants.

[00199] Such a steviol glycoside composition can be a RebD-enriched composition having greater than 3 weight % RebD and be incorporated into the food product such that the amount of RebD in the product is from 25-1600 mg/kg, *e.g.*, 100-500 mg/kg, 25-100 mg/kg, 250-1000 mg/kg, 50-500 mg/kg or 500-1000 mg/kg on a dry weight basis. Typically, the RebD-enriched composition has an undetectable amount of stevia plant-derived contaminants.

[00200] Such a steviol glycoside composition can be a RebE-enriched composition having greater than 3 weight % RebE and be incorporated into the food product such that the amount of RebE in the product is from 25-1600 mg/kg, *e.g.*, 100-500 mg/kg, 25-100 mg/kg, 250-1000

mg/kg, 50-500 mg/kg or 500-1000 mg/kg on a dry weight basis. Typically, the RebE-enriched composition has an undetectable amount of stevia plant-derived contaminants.

[00201] Such a steviol glycoside composition can be a RebM-enriched composition having greater than 3 weight % RebM and be incorporated into the food product such that the amount of RebM in the product is from 25-1600 mg/kg, *e.g.*, 100-500 mg/kg, 25-100 mg/kg, 250-1000 mg/kg, 50-500 mg/kg or 500-1000 mg/kg on a dry weight basis. Typically, the RebM-enriched composition has an undetectable amount of stevia plant-derived contaminants.

[00202] In some embodiments, a substantially pure steviol or steviol glycoside is incorporated into a tabletop sweetener or "cup-for-cup" product. Such products typically are diluted to the appropriate sweetness level with one or more bulking agents, *e.g.*, maltodextrins, known to those skilled in the art. Steviol glycoside compositions enriched for RebA, RebB, RebD, RebE, or RebM, can be package in a sachet, for example, at from 10,000 to 30,000 mg steviol glycoside/kg product on a dry weight basis, for tabletop use. In some embodiments, a steviol glycoside produced *in vitro*, *in vivo*, or by whole cell bioconversion

[00203] The invention will be further described in the following examples, which do not limit the scope of the invention described in the claims.

EXAMPLES

[00204] The Examples that follow are illustrative of specific embodiments of the invention, and various uses thereof. They are set forth for explanatory purposes only, and are not to be taken as limiting the invention.

Example 1: Strain Engineering

[00205] Steviol glycoside-producing *S. cerevisiae* strains were constructed as described in WO 2011/153378, WO 2013/022989, WO 2014/122227, and WO 2014/122328, each of which is incorporated by reference in its entirety. For example, yeast strains comprising and expressing a native gene encoding a YNK1 polypeptide (SEQ ID NO:122, SEQ ID NO:123), a native gene encoding a PGM1 polypeptide (SEQ ID NO:1, SEQ ID NO:2), a native gene encoding a PGM2 polypeptide (SEQ ID NO:118, SEQ ID NO:119), a native gene encoding a UGP1 polypeptide (SEQ ID NO:120, SEQ ID NO:121), a recombinant gene encoding a GGPPS polypeptide (SEQ ID NO:19, SEQ ID NO:20), a recombinant gene encoding a truncated CDPS

polypeptide (SEQ ID NO:39, SEQ ID NO:40), a recombinant gene encoding a KS polypeptide (SEQ ID NO:51, SEQ ID NO:52), a recombinant gene encoding a KO polypeptide (SEQ ID NO:59, SEQ ID NO:60), a recombinant gene encoding a KO polypeptide (SEQ ID NO:63, SEQ ID NO:64), a recombinant gene encoding an ATR2 polypeptide (SEQ ID NO:91, SEQ ID NO:92), a recombinant gene encoding a KAHe1 polypeptide (SEQ ID NO:93, SEQ ID NO:94), a recombinant gene encoding a CPR8 polypeptide (SEQ ID NO:85, SEQ ID NO:86), a recombinant gene encoding a CPR1 polypeptide (SEQ ID NO:77, SEQ ID NO:78), a recombinant gene encoding a UGT76G1 polypeptide (SEQ ID NO:8, SEQ ID NO:9), a recombinant gene encoding a UGT85C2 polypeptide (SEQ ID NO:5/SEQ ID NO:6, SEQ ID NO:7), a recombinant gene encoding a UGT74G1 polypeptide (SEQ ID NO:3, SEQ ID NO:4), a recombinant gene encoding a UGT91d2e-b polypeptide (SEQ ID NO:12, SEQ ID NO:13) and a recombinant gene encoding an EUGT11 polypeptide (SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16) were engineered to accumulate steviol glycosides.

Example 2: Overexpression of PGM1, PGM2, UGP1, and YNK1

[00206] A steviol glycoside-producing *S. cerevisiae* strain as described in Example 1, further engineered to comprise and express a recombinant gene encoding a KAH polypeptide (SEQ ID NO:96, SEQ ID NO:97) and a recombinant gene encoding a KO polypeptide (SEQ ID NO:117, SEQ ID NO:64), was transformed with vectors comprising an additional copy of the gene encoding a YNK1 polypeptide (SEQ ID NO:122, SEQ ID NO:123), operably linked to a pTEF1 promoter (SEQ ID NO:148) and a CYC1 terminator (SEQ ID NO:154), an additional copy of the gene encoding a PGM1 polypeptide (SEQ ID NO:1, SEQ ID NO:2), operably linked to a pTEF1 promoter (SEQ ID NO:148) and a CYC1 terminator (SEQ ID NO:154), an additional copy of the gene encoding a PGM2 polypeptide (SEQ ID NO:118, SEQ ID NO:119), operably linked to a pPGK1 promoter (SEQ ID NO:149) and a tADH1 terminator (SEQ ID NO:155), and an additional copy of the gene encoding a UGP1 polypeptide (SEQ ID NO:120, SEQ ID NO:121), operably linked to a pPGK1 promoter (SEQ ID NO:149) and a tADH1 terminator (SEQ ID NO:155).

[00207] Fed-batch fermentation with cultures of the transformed *S. cerevisiae* strain and a control *S. cerevisiae* strain (a steviol glycoside-producing *S. cerevisiae* strain as described in Example 2, further engineered to comprise and express a recombinant gene encoding a KAH

polypeptide and a recombinant gene encoding a KO polypeptide) was carried out aerobically in 2L fermenters at 30ºC with an approximate 16 h growth phase in minimal medium comprising glucose, ammonium sulfate, trace metals, vitamins, salts, and buffer followed by an approximate 100 h feeding phase with a glucose-comprising defined feed medium. A pH near 6.0 and glucose-limiting conditions were maintained. Extractions of whole culture samples (without cell removal) were performed and extracts were analyzed by LC-UV to determine levels of steviol glycosides.

[00208] LC-UV was conducted with an Agilent 1290 instrument comprising a variable wavelength detector (VWD), a thermostatted column compartment (TCC), an autosampler, an autosampler cooling unit, and a binary pump, using SB-C18 rapid resolution high definition (RRHD) 2.1 mm x 300 mm, 1.8 µm analytical columns (two 150 mm columns in series; column temperature of 65°C). Steviol glycosides were separated by a reversed-phase C18 column followed by detection by UV absorbance at 210 mm. Quantification of steviol glycosides was done by comparing the peak area of each analyte to standards of RebA and applying a correction factor for species with differing molar absorptivities. For LC-UV, 0.5 mL cultures were spun down, the supernatant was removed, and the wet weight of the pellets was calculated. The LC-UV results were normalized by pellet wet weight. Total steviol glycoside values of the fed-batch fermentation were calculated based upon the measured levels of steviol glycosides calculated as a sum (in g/L RebD equivalents) of measured RebA, RebB, RebD, RebE, RebM, 13-SMG, rubusoside, steviol-1,2-bioside, di-glycosylated steviol, tri-glycosylated steviol, tetra-glycosylated steviol, penta-glycosylated steviol, hexa-glycosylated steviol, and hepta-glycosylated steviol. Results are shown in Table 1.

Table 1: Steviol Glycoside accumulation by transformed *S. cerevisiae* strain and *S. cerevisiae* control strain.

	Transformed Strain		Control Strain	
	Accumulation (g/L RebD Equiv.)	Std. Error (g/L RebD Equiv.)	Accumulation (g/L RebD Equiv.)	Std. Error (g/L RebD Equiv.)
13-SMG	2.40	0.14	4.2	0.02
RebA	0.59	0.007	0.45	0.07
RebD	1.21	0.16	2.16	0.12
RebM	6.31	0.22	3.22	0.06
Total SG	11.90	0.33	11.76	0.34

[00209] A decrease in 13-SMG and RebD accumulation, and an increase in RebA and RebM accumulation were observed for the *S. cerevisiae* strain overexpressing UGP1, YNK1, PGM1, and PGM2, relative to the control strain. Furthermore, RebD + RebM accumulation levels increased upon overexpression of UGP1, YNK1, PGM1, and PGM2, while the total steviol glycosides produced by the experimental strain increased negligibly. In addition, RebD / RebM ratios of 0.2 and below were observed for the *S. cerevisiae* strain overexpressing UGP1, YNK1, PGM1, and PGM2, relative to the control strain.

Example 3: UGP1, PGM2 Activity Assay

[00210] Fed-batch fermentation with cultures of a *S. cerevisiae* strain overexpressing PGM1, PGM2, UGP1, and YNK1, as described in Example 2, and a control *S. cerevisiae* strain (a steviol glycoside-producing *S. cerevisiae* strain as described in Example 1) was carried out aerobically in 2L fermenters at 30ºC with an approximate 16 h growth phase in minimal medium comprising glucose, ammonium sulfate, trace metals, vitamins, salts, and buffer followed by an approximate 100 h feeding phase with a glucose-comprising defined feed medium. A pH near 6.0 and glucose-limiting conditions were maintained. Whole culture samples (without cell removal) were analyzed to determine the activity levels of PGM and UGP.

[00211] For both assays, frozen fermentation cell pellets were resuspended in CelLytic™ Y Cell Lysis Reagent (Sigma) to an OD_{600} of 44. Samples were shaken 30 min at 25°C and then centrifuged at 13,000 rpm for 10 min. The supernatant was recovered and stored on ice.

[00212] The PGM enzyme assay relies on a coupled activity assay wherein supplied glucose-1-phosphate is first converted to glucose-6-phosphate by a PGM polypeptide/PGM polypeptide containing cell lysate, followed by glucose-6-phosphate conversion by a glucose-6-phosphate dehydrogenase (added to the assay as a purified enzyme in excess) to phosphogluconolactone under β-$NADP^+$ consumption. The kinetics of the concomitant β-NAPDH released are recorded by monitoring the absorbance at 340 nm.

[00213] 180 mM glycylglycine, pH 7.4 (adjusted with NaOH/HCl); 5.0 mM glucose-1-phosphate; 3.00 mM β-$NADP^+$; 0.4 mM G1,6-bisphosphate; 30 mM $MgCl_2$, 43 mM L-cysteine; 0.65 U/ml G6P-DH, and previously stored cell lysate were mixed together at 30°C at different cell-lysate/buffer concentrations (0.5% (v/v), 1%(v/v), 2%(v/v), and 3%(v/v)). The kinetics for the release of β-NAPDH were followed over a maximum of 1000 sec. for each concentration of

supernatant added. PGM activity for each cell-lysate/buffer concentration was defined by the maximum slope of the curve of OD_{340} versus time. Cell-lysate/buffer concentration corrected PGM activity was defined as the slope of the curve of OD340/sec as a function of Cell-lysate/buffer concentrations. The value obtained in this way for a certain strain can be compared to the values from other strains and differences in PGM activity can be pointed out. The increase in activity of the cell-lysate of the *S. cerevisiae* strain overexpressing PGM1, PGM2, UGP1, and YNK1 is shown in Table 3, below, relative to that of the control strain.

[00214] The UGP assay relies on a coupled activity assay of the yeast UDP-glucose pyrophosphorylase wherein supplied glucose-1-phosphate is first converted to UDP-glucose by a UGP polypeptide/UGP polypeptide-containing cell-lysate under UTP consumption, followed by UDP-glucose convertion to UDP-Glucuronate and β-NADH by UDP-glucose dehydrogenase (added to the assay as a purified enzyme in excess) under β-NAD^+ consumption. The kinetics for the release of β-NADH are followed by monitoring the change in absorbance at 340 nm. Alternative UGP assays using, for example but not limited to, hydrophilic interaction liquid chromatography coupled with tandem mass spectrometry for the quantification of UDP-glucose (*see*Warth *et al.*, Journal of Chromatography A, 1423, pp. 183–189 (2016)) may be used as well.

[00215] 100 mM Tris/HCl, pH 8.5; 10 mM MgCl2; 100 mM NaCl; 5.0 mM β-NAD^+; 2 mM UTP; 2 mM ATP; 0.12 mg/ml UDPG-DH; 5 mM; and previously stored cell lysate were mixed together at 30°C at different supernatant/buffer concentrations (0.5% (v/v), 1%(v/v), 1.5%(v/v), and 2%(v/v)). The kinetics for the release of β-NADH were followed over a maximum of 1000 sec. for each supernatant/buffer concentration. UGP activity for each cell-lysate/buffer concentration was defined by the maximum slope of the curve of OD_{340} versus time. Cell-lysate/buffer concentration corrected UGP activity was defined as the slope of the curve of OD340/sec as a function of Cell-lysate/buffer concentrations. The value obtained in this way for a certain strain can be compared to the values from other strains and differences in UGP activity can be pointed out. The increase in activity of the lysate of the *S. cerevisiae* strain overexpressing PGM1, PGM2, UGP1, and YNK1 is shown in Table 2, below, relative to that of the control strain.

Table 2. Relative UGP and PGM activity

	Transformed Strain	**Control Strain**
UGP Activity relative to control strain	250%	100%

PGM Activity relative to control stain	160%	100%

[00216] Individual and total steviol glycoside values of the fed-batch fermentation were calculated according to Example 2. Results are shown in Table 3.

Table 3: Steviol Glycoside accumulation by transformed *S. cerevisiae* strain and *S. cerevisiae* control strain.

	Transformed Strain	**Control Strain**
	Accumulation (g/L RebD Equiv.)	Accumulation (g/L RebD Equiv.)
RebD	2.19	1.21
RebM	5.71	5.12
Total SG	12.10	9.43

[00217] An increase in both UGP and PGM activity was observed for the *S. cerevisiae* strain overexpressing UGP1, YNK1, PGM1, and PGM2, relative to the control strain. As shown in Table 3, RebD and total steviol glycoside accumulation increased upon overexpression of UGP1, YNK1, PGM1, and PGM2. Without being bound to a particular theory, the results suggest that increasing UGP and PGM activity (*i.e.*, by expressing genes encoding polypeptides involved in the UDP-glucose biosynthetic pathway) allows for conversion of partially glycosylated steviol glycosides to higher moleculae weight steviol glycosides, including, *e.g.*, RebD.

Example 4: LC-MS Analytical Procedures (UDP-glucose Analysis)

[00218] LC-MS analyses were performed on a Thermo Scientific Accela UPLC (Ultra Performance Liquid Chromatography system; Thermo Scientific) with a Thermo Scientific PAL autosampler system (Thermo Scientific) SeQuant ZIC-cHILIC column (2.1 mm x 150 mm, 3.0 µm analytical column, 100 Å pore size) coupled to a Thermo Scientific Exactive Orbitrap mass spectrometer with electrospray ionization (ESI) operated in negative ionization mode. Compound separation was achieved using a gradient of the two mobile phases: A (water with 0.1% ammonium acetate) and B (MeCN). Separation was achieved by using a gradient from time 0 min with 15% A holding until 0.5 min and increasing to 50% A at time 15.50 min, holding until time 17.50 min, and reducing to 15% A at time 17.60 min, equilibrating at 15% A until 25.50

min. The flow rate was 0.3 mL/min, and the column was maintained at room temperature. UDP-glucose was monitored by full-scan analysis in the mass range 130-1400 m/z. EIC (Extracted ion chromatogram) of 565.04492-565.05058 corresponding to UDP-glucose was extracted and quantified by comparing against authentic standards. See Table 4 for m/z trace and retention time values of UDP-glucose.

Table 4: LC-MS Analytical Data for UDP-glucose

Compound	MS Trace	RT (mins)
UDP-glucose	565.04775	8.4

[00219] To determine the intracellular concentration of UDP-Glucose, full fermentation broth was sampled (via syringe) at desired time points during different stages of fermentation. Biomass (cells) was quickly separated by centrifugation and supernatant was removed. Cell pellets were quenched and extracted using a mixture of methanol, chloroform and an aqueous buffer solution. The final intracellular extracts were stored at -80°C prior to LC-MS analysis.

Example 5: UDP-glucose Accumulation Quantification

[00220] Fed-batch fermentation with cultures of a *S. cerevisiae* strain overexpressing PGM1, PGM2, UGP1, and YNK1, as described in Example 2, and a control *S. cerevisiae* strain (a *S. cerevisiae* strain comprising and expressing a native gene encoding a YNK1 polypeptide (SEQ ID NO:122, SEQ ID NO:123), a native gene encoding a PGM1 polypeptide (SEQ ID NO:1, SEQ ID NO:2), a native gene encoding a PGM2 polypeptide (SEQ ID NO:118, SEQ ID NO:119), a native gene encoding a UGP1 polypeptide (SEQ ID NO:120, SEQ ID NO:121), a recombinant gene encoding a GGPPS polypeptide (SEQ ID NO:19, SEQ ID NO:20), a recombinant gene encoding a truncated CDPS polypeptide (SEQ ID NO:39, SEQ ID NO:40), a recombinant gene encoding a KS polypeptide (SEQ ID NO:51, SEQ ID NO:52), a recombinant gene encoding a KO polypeptide (SEQ ID NO:59, SEQ ID NO:60), a recombinant gene encoding a KAHe1 polypeptide (SEQ ID NO:93, SEQ ID NO:94), a recombinant gene encoding a CPR8 polypeptide (SEQ ID NO:85, SEQ ID NO:86), a recombinant gene encoding a CPR1 polypeptide (SEQ ID NO:77, SEQ ID NO:78), a recombinant gene encoding an ATR2 polypeptide (SEQ ID NO:91, SEQ ID NO:92), a recombinant gene encoding a UGT85C2

polypeptide (SEQ ID NO:5/SEQ ID NO:6, SEQ ID NO:7), and a recombinant gene encoding a UGT74G1 polypeptide (SEQ ID NO:3, SEQ ID NO:4)) was carried out aerobically in 2L fermenters at 30°C with an approximate 16 h growth phase in minimal medium comprising glucose, ammonium sulfate, trace metals, vitamins, salts, and buffer followed by an approximate 100 h feeding phase with a glucose-comprising defined feed medium. A pH near 6.0 and glucose-limiting conditions were maintained. Whole culture samples (without cell removal) were analyzed by LC-UV to determine the levels of steviol glycosides, according to Example 2, and by LC-MS to analyze the intracellular level of UDP-glucose, according to Example 4. Results are shown in Tables 5-6.

Table 5: Steviol Glycoside accumulation by transformed *S. cerevisiae* strain and *S. cerevisiae* control strain.

	Transformed Strain	**Control Strain**
	Accumulation (g/L RebD Equiv.)	Accumulation (g/L RebD Equiv.)
RebD	1.05	1.92
RebM	5.75	2.23
Total SG	10.18	7.40

Table 6: UDP-glucose accumulation by transformed *S. cerevisiae* strain and *S. cerevisiae* control strain.

	Transformed Strain		**Control Strain**	
Time (h)	UDP-glucose Accumulation (μM)	Std. Deviation (μM)	UDP-glucose Accumulation (μM)	Std. Deviation (μM)
22	450.52	54.96	306.50	51.75
30	495.66	10.83	198.88	36.95
46	518.26	26.13	241.30	45.69
55	425.39	70.01	221.35	64.36
72	398.08	41.85	206.26	19.54
76	299.16	33.57	159.96	5.06
96	270.53	82.67	160.74	9.19
104	310.97	24.57	132.08	21.17
120	359.92	24.30	119.32	37.39

[00221] An increase in UDP-glucose accumulation, by up to 300%, was observed for the *S. cerevisiae* strain overexpressing UGP1, YNK1, PGM1, and PGM2, relative to the control strain. RebD + RebM accumulation levels increased upon overexpression of UGP1, YNK1, PGM1, and PGM2; this result further demonstrates a beneficial effect of expression of UDP-glucose biosynthetic pathway genes on the production of higher molecular weight steviol glycosides such as RebD or RebM.

[00222] One of skill in the art would appreciate a disctinction between improving the total amount of UDP-glucose as compared to the recycling of UDP-glucose. As shown in Table 6 above, taking the highest and lowest number over fermentation time, the worst decrase in parental strain is 2.5 while the worst decrease in UDP-glucose boosted strain (*i.e.*, the *S. cerevisiae* strain overexpressing UGP1, YNK1, PGM1, and PGM2) is 1.9 times. This demonstrates that overexpressing UGP1, YNK1, PGM1, and PGM2 increases the UDP-glucose pool and UDP-glucose. In fact, the net increase (consumption/formation) is higher is the UDP-glucose boosted strain.

[00223] Without being bound to a particular theory, the results observed in Examples 2-5 suggest that increasing UDP-glucose levels (*i.e.*, by expressing genes encoding polypeptides involved in the UDP-glucose biosynthetic pathway) allows for conversion of 13-SMG and other partially glycosylated steviol glycosides to higher molecular weight steviol glycosides, including, *e.g.*, RebM. Furthermore, the difference between the magnitude of the increase in accumulation levels of, *e.g.*, RebM and/or RebD and that of the increase in accumulation levels of the total steviol glycosides suggests that maintaining and/or increasing UDP-glucose levels allows for more efficient production of higher molecular weight steviol glycosides, including, *e.g.*, RebM (*i.e.*, by shifting the profile of produced steviol glycosides away from lower molecular weight steviol glycosides).

Example 6: Expression of Heterologous UGP1 and PGM2

[00224] A steviol glycoside-producing *S. cerevisiae* strain overexpressing UGP1, YNK1, PGM1, and PGM2, as described in Example 2, was transformed with vectors comprising a gene encoding a UGP1 polypeptide (SEQ ID NO:132, SEQ ID NO:133) operably linked to a pPDC1 promoter (SEQ ID NO:153) and a tCYC1 terminator (SEQ ID NO:154) and a gene encoding a

PGM2 polypeptide (SEQ ID NO:144, SEQ ID NO:145), operably linked to a pTPI1 promoter (SEQ ID NO:152) and an tADH1 terminator (SEQ ID NO:155).

[00225] Fed-batch fermentation with cultures of the transformed *S. cerevisiae* strain and a control *S. cerevisiae* strain (a steviol glycoside-producing *S. cerevisiae* strain as described in Example 2, further engineered to comprise and express a recombinant gene encoding a ~~Stevia KAH polypeptide,~~ KAH polypeptide and a recombinant gene encoding a KO polypeptide) was carried out aerobically in 2L fermenters at 30ºC with an approximate 16 h growth phase in minimal medium comprising glucose, ammonium sulfate, trace metals, vitamins, salts, and buffer followed by an approximate 100 h feeding phase with a glucose-comprising defined feed medium. A pH near 6.0 and glucose-limiting conditions were maintained. Whole culture samples (without cell removal) were analyzed by LC-UV to determine levels of steviol glycosides, as described in Example 2. Results are shown in Table 7.

Table 7: Steviol Glycoside accumulation by transformed *S. cerevisiae* strain and *S. cerevisiae* control strain.

	Transformed Strain	**Control Strain**
	Accumulation (g/L RebD Equiv.)	Accumulation (g/L RebD Equiv.)
RebD	2.27	1.80
RebM	5.33	4.50
Total SG	14.27	12.39

[00226] An increase in RebD and RebM accumulation were observed for the *S. cerevisiae* strain expressing PGM2 and UGP1, relative to the control strain. Furthermore, total steviol glycosides produced by the experimental strain also increased. Without being bound to a particular theory, the results observed in Table 7 suggest that increasing UDP-glucose levels (*i.e.*, by expressing genes encoding polypeptides involved in the UDP-glucose biosynthetic pathway) allows for conversion of 13-SMG and other partially glycosylated steviol glycosides to higher molecular weight steviol glycosides, including, *e.g.*, RebM.

Example 7: LC-MS Analytical Procedures (Steviol Glycoside Analysis)

[00227] LC-MS analyses were performed on a Waters ACQUITY UPLC (Ultra Performance Liquid Chromatography system; Waters Corporation) with a Waters ACQUITY UPLC (Ultra Performance Liquid Chromatography system; Waters Corporation) BEH C18 column (2.1 x 50 mm, 1.7 µm particles, 130 Å pore size) equipped with a pre-column (2.1 x 5 mm, 1.7 µm particles, 130 Å pore size) coupled to a Waters ACQUITY TQD triple quadropole mass spectrometer with electrospray ionization (ESI) operated in negative ionization mode. Compound separation was achieved using a gradient of the two mobile phases, A (water with 0.1% formic acid) and B (MeCN with 0.1% formic acid), by increasing from 20% to 50 % B between 0.3 to 2.0 min, increasing to 100% B at 2.01 min and holding 100% B for 0.6 min, and re-equilibrating for 0.6 min. The flow rate was 0.6 mL/min, and the column temperature was set at 55°C. Steviol glycosides were monitored using SIM (Single Ion Monitoring) and quantified by comparing against authentic standards. *See* Table 1 for m/z trace and retention time values of steviol glycosides and glycosides of steviol precursors detected.

Table 8: LC-MS Analytical Data for Steviol and Glycosides of Steviol and Steviol Precursors

Compound	MS Trace	RT (mins)
steviol+5Glc (#22) [also referred to as compound 5.22]	1127.48	0.85
steviol+6Glc (isomer 1) [also referred to as compound 6.1]	1289.53	0.87
steviol+7Glc (isomer 2) [also referred to as compound 7.2]	1451.581	0.94
steviol+6Glc (#23) [also referred to as compound 6.23]	1289.53	0.97

Compound	MS Trace	RT (mins)
RebE	965.42	1.06
RebD	1127.48	1.08
RebM	1289.53	1.15
steviol+7Glc (isomer 5) [also referred to as compound 7.5]	1451.581	1.09
steviol+4Glc (#26) [also referred to as compound 4.26]	965.42	1.21
steviol+5Glc (#24) [also referred to as compound 5.24]	1127.48	1.18
steviol+4Glc (#25) [also referred to as compound 5.25]	1127.48	1.40
RebA	965.42	1.43
1,2-Stevioside	803.37	1.43
steviol+4Glc (#33) [also referred to as compound 4.33]	965.42	1.49
steviol+3Glc (#1) [also referred to as compound 3.1]	803.37	1.52
steviol+2Glc (#57) [also referred to as compound 2.57]	641.32	1.57
RebQ	965.42	1.59
1,3-Stevioside (RebG)	803.37	1.60
Rubusoside	641.32	1.67
RebB	803.37	1.76
Steviol-1,2-Bioside	641.32	1.80
Steviol-1,3-Bioside	641.32	1.95
19-SMG	525.27	1.98
13-SMG	479.26	2.04
ent-kaurenoic acid+3Glc (isomer 1) [also referred to as compound KA3.1]	787.37	2.16

Compound	MS Trace	RT (mins)
ent-kaurenoic acid+3Glc (isomer 2) [also referred to as compound KA3.2]	787.37	2.28
ent-kaurenol+3Glc (isomer 1) co-eluted with *ent*-kaurenol+3Glc (#6) [also referred to as compounds KL3.1 and KL3.6]	773.4	2.36
ent-kaurenoic acid+2Glc (#7) [also referred to as compound KA2.7]	625.32	2.35
ent-kaurenol+2Glc (#8) [also referred to as compound KL2.8]	611.34	2.38
Steviol	317.21	2.39

[00228] Steviol glycosides can be isolated using a method described herein. For example, following fermentation, a culture broth can be centrifuged for 30 min at 7000 rpm at 4°C to remove cells, or cells can be removed by filtration. The cell-free lysate can be obtained, for example, by mechanical disruption or enzymatic disruption of the host cells and additional centrifugation to remove cell debris. Mechanical disruption of the dried broth materials can also be performed, such as by sonication. The dissolved or suspended broth materials can be filtered using a micron or sub-micron filter prior to further purification, such as by preparative chromatography. The fermentation media or cell-free lysate can optionally be treated to remove low molecular weight compounds such as salt, and can optionally be dried prior to purification and re-dissolved in a mixture of water and solvent. The supernatant or cell-free lysate can be purified as follows: a column can be filled with, for example, HP20 Diaion resin (aromatic-type Synthetic Adsorbent; Supelco) or another suitable non-polar adsorbent or reverse phase chromatography resin, and an aliquot of supernatant or cell-free lysate can be loaded on to the column and washed with water to remove the hydrophilic components. The steviol glycoside product can be eluted by stepwise incremental increases in the solvent concentration in water or a gradient from, *e.g.*, 0% → 100% methanol. The levels of steviol glycosides, glycosylated *ent*-kaurenol, and/or glycosylated *ent*-kaurenoic acid in each fraction, including the flow-through, can then be analyzed by LC-MS. Fractions can then be combined and reduced in volume using

a vacuum evaporator. Additional purification steps can be utilized, if desired, such as additional chromatography steps and crystallization.

Example 8: Expression of Heterologous UGP1

[00229] A steviol glycoside-producing *S. cerevisiae* strain overexpressing UGP1, YNK1, PGM1, and PGM2, as described in Example 2, was transformed with a vector comprising a codon-optimized nucleotide sequence encoding a UGP1 polypeptide (SEQ ID NO:132, SEQ ID NO:133) operably linked to a pTDH3 promoter (SEQ ID NO:150) and a tCYC1 terminator (SEQ ID NO:154), as summarized in Table 9, below.

Table 9. UGP1 Polypeptides Expressed

Strain	SEQ ID
1	126, 127
2	132, 133
3	128, 129
4	130, 131
5	124, 125
6	138, 139
7	136, 137
8	134, 135

[00230] Single colonies of the transformed strains provided in Table 9, and a control strain, transformed with a blank vector, were grown in 500 µL of Delft medium in a 96-well plate for 2 days at 30°C, shaking at 280 rpm. 50 µL of the cell culture of each strain was then transferred to a second 96-well plate and grown in 450 µL Feed-in-Time medium (m2p-labs GmbH, Baesweiler, Germany) for 4 days at 30°C, shaking at 280 rpm. Samples for LC-MS analysis were prepared by extracting 100 µL of cell solution with 100 µL of DMSO, vortexing until mixed, and incubating at 80°C for 10 minutes. The resultant extract was clarified by centrifugation at 15,000 *g* for 10 min. 20 µL of the supernatant was diluted with 140 µL of 50% (v/v) DMSO for LC-MS injection. LC-MS data was normalized to the OD_{600} of a mixture of 100 µL of the cell solution and 100 µL of water, measured on an ENVISION® Multilabel Reader (PerkinElmer, Waltham, MA).

[00231] LC-MS analysis was performed according to Example 7. Whole culture accumulation of compounds in $\mu M/OD_{600}$ was quantified by LC-MS against a known standard. Results are shown in Table 10, below. Each value is an average of 6 independent clones.

Table 10. Concentration of Steviol Glycosides

	Accumulated Concentration ($\mu M/OD_{600}$)						
Strain	**13-SMG**	**Rubu.**	**RebB**	**RebA**	**RebD**	**RebM**	**Total**
Control	9.96 ± 2.19	0.05 ± 0.08	0.67 ± 0.14	1.95 ± 0.79	3.89 ± 0.60	20.73 ± 4.48	37.38 ± 6.71
1	6.15 ± 1.83	0.26 ± 0.04	0.59 ± 0.09	2.37 ± 0.65	1.49 ± 0.36	25.91 ± 1.35	37.38 ± 3.03
2	7.06 ± 2.48	0.23 ± 0.12	0.76 ± 0.30	2.03 ± 0.37	1.34 ± 0.24	27.99 ± 3.17	39.43 ± 5.88
3	8.73 ± 3.20	0.25 ± 0.08	0.69 ± 0.24	2.50 + 0.81	1.69 ± 0.43	29.41 ± 6.19	43.34 ± 9.22
4	13.02 ± 2.39	0.14 ± 0.08	0.99 ± 0.23	2.88 ± 0.51	4.89 ± 0.75	30.41 ± 5.90	52.50 ± 9.51
5	7.91 ± 2.30	0.28 ± 0.08	0.62 ± 0.14	2.55 ± 0.96	1.42 ± 0.33	29.54 ± 4.23	42.37 ± 5.98
6	8.89 ± 2.94	0.28 ± 0.04	0.68 ± 0.18	2.36 ± 0.66	1.43 ± 0.49	27.64 ± 3.49	41.32 ± 5.08
7	5.68 ± 2.05	0.23 ± 0.09	0.51 ± 0.19	2.04 ± 0.50	1.26 ± 0.28	23.63 ± 2.27	33.38 ± 4.98
8	6.59 ± 2.65	0.22 ± 0.12	0.63 ± 0.17	2.28 ± 1.03	1.49 ± 0.59	26.64 ± 6.51	37.90 ± 10.21

[00232] Increases in steviol glycoside accumulation, by up to about 600%, was observed for the *S. cereivisiae* strain overexpressing UGP1, YNK1, PGM1, and PGM2, and further expressing heterologous UGP1, relative to the control strain. RebD + RebM accumulation levels increased upon expression of heterologous UGP1, further demonstrating a beneficial effect of expression of heterologous UDP-glucose biosynthetic pathway genes on the production of higher molecular weight steviol glycosides such as RebD or RebM.

[00233] Having described the invention in detail and by reference to specific embodiments thereof, it will be apparent that modifications and variations are possible without departing from the scope of the invention defined in the appended claims. More specifically, although some aspects of the present invention are identified herein as particularly advantageous, it is

contemplated that the present invention is not necessarily limited to these particular aspects of the invention.

Table 11. Sequences disclosed herein.

SEQ ID NO:1
S. cerevisiae

```
atgtcacttc taatagattc tgtaccaaca gttgcttata aggaccaaaa accgggtact        60
tcaggtttac gtaagaagac caaggttttc atggatgagc ctcattatac tgagaacttc       120
attcaagcaa caatgcaatc tatccctaat ggctcagagg gaaccacttt agttgttgga       180
ggagatggtc gtttctacaa cgatgttatc atgaacaaga ttgccgcagt aggtgctgca       240
aacggtgtca gaaagttagt cattggtcaa ggcggtttac tttcaacacc agctgcttct       300
catataatta gaacatacga ggaaaagtgt accggtggtg gtatcatatt aactgcctca       360
cacaacccag gcggtccaga gaatgattta ggtatcaagt ataatttacc taatggtggg       420
ccagctccag agagtgtcac taacgctatc tgggaagcgt ctaaaaaatt aactcactat       480
aaaattataa agaacttccc caagttgaat ttgaacaagc ttggtaaaaa ccaaaaatat       540
ggcccattgt tagtggacat aattgatcct gccaaagcat acgttcaatt tctgaaggaa       600
atttttgatt ttgacttaat taaaagcttc ttagcgaaac agcgcaaaga caaagggtgg       660
aagttgttgt ttgactcctt aaatggtatt acaggaccat atggtaaggc tatatttgtt       720
gatgaatttg gtttaccggc agaggaagtt cttcaaaatt ggcacccttt acctgatttc       780
ggcggtttac atcccgatcc gaatctaacc tatgcacgaa ctcttgttga cagggttgac       840
cgcgaaaaaa ttgcctttgg agcagcctcc gatggtgatg gtgataggaa tatgatttac       900
ggttatggcc ctgctttcgt ttcgccaggt gattctgttg ccattattgc cgaatatgca       960
cccgaaattc catacttcgc caaacaaggt atttatggct tggcacgttc atttcctaca      1020
tcctcagcca ttgatcgtgt tgcagcaaaa aagggattaa gatgttacga agttccaacc      1080
ggctggaaat tcttctgtgc cttatttgat gctaaaaagc tatcaatctg tggtgaagaa      1140
tccttcggta caggttccaa tcatatcaga gaaaaggacg gtctatgggc cattattgct      1200
tggttaaata tcttggctat ctaccatagg cgtaaccctg aaaaggaagc ttcgatcaaa      1260
actattcagg acgaattttg gaacgagtat ggccgtactt tcttcacaag atacgattac      1320
gaacatatcg aatgcgagca ggccgaaaaa gttgtagctc ttttgagtga atttgtatca      1380
aggccaaacg tttgtggctc ccacttccca gctgatgagt ctttaaccgt tatcgattgt      1440
ggtgattttt cgtatagaga tctagatggc tccatctctg aaaatcaagg ccttttcgta      1500
aagttttcga atgggactaa atttgttttg aggttatccg gcacaggcag ttctggtgca      1560
acaataagat tatacgtaga aaagtatact gataaaaagg agaactatgg ccaaacagct      1620
gacgtcttct tgaaacccgt catcaactcc attgtaaaat tcttaagatt taaagaaatt      1680
ttaggaacag acgaaccaac agtccgcaca tag                                   1713
```

SEQ ID NO:2
S. cerevisiae

```
MSLLIDSVPT VAYKDQKPGT SGLRKKTKVF MDEPHYTENF IQATMQSIPN GSEGTTLVVG        60
GDGRFYNDVI MNKIAAVGAA NGVRKLVIGQ GGLLSTPAAS HIIRTYEEKC TGGGIILTAS       120
HNPGGPENDL GIKYNLPNGG PAPESVTNAI WEASKKLTHY KIIKNFPKLN LNKLGKNQKY       180
GPLLVDIIDP AKAYVQFLKE IFDFDLIKSF LAKQRKDKGW KLLFDSLNGI TGPYGKAIFV       240
DEFGLPAEEV LQNWHPLPDF GGLHPDPNLT YARTLVDRVD REKIAFGAAS DGDGDRNMIY       300
GYGPAFVSPG DSVAIIAEYA PEIPYFAKQG IYGLARSFPT SSAIDRVAAK KGLRCYEVPT       360
GWKFFCALFD AKKLSICGEE SFGTGSNHIR EKDGLWAIIA WLNILAIYHR RNPEKEASIK       420
TIQDEFWNEY GRTFFTRYDY EHIECEQAEK VVALLSEFVS RPNVCGSHFP ADESLTVIDC       480
GDFSYRDLDG SISENQGLFV KFSNGTKFVL RLSGTGSSGA TIRLYVEKYT DKKENYGQTA       540
DVFLKPVINS IVKFLRFKEI LGTDEPTVRT                                        570
```

SEQ ID NO:3
S. rebaudiana

```
atggcagagc aacaaaagat caaaaagtca cctcacgtct tacttattcc atttcctctg        60
caaggacata tcaacccatt catacaattt gggaaaagat tgattagtaa gggtgtaaag       120
acaacactgg taaccactat ccacactttg aattctactc tgaaccactc aaatactact       180
actacaagta tagaaattca agctatatca gacggatgcg atgagggtgg ctttatgtct       240
```

```
gccggtgaat cttacttgga aacattcaag caagtgggat ccaagtctct ggccgatcta      300
atcaaaaagt tacagagtga aggcaccaca attgacgcca taatctacga ttctatgaca      360
gagtgggttt tagacgttgc tatcgaattt ggtattgatg gaggttcctt tttcacacaa      420
gcatgtgttg tgaattctct atactaccat gtgcataaag ggttaatctc tttaccattg      480
ggtgaaactg tttcagttcc aggttttcca gtgttacaac gttgggaaac cccattgatc      540
ttacaaaatc atgaacaaat acaatcacct tggtcccaga tgttgtttgg tcaattcgct      600
aacatcgatc aagcaagatg ggtctttact aattcattct ataagttaga ggaagaggta      660
attgaatgga ctaggaagat ctggaatttg aaagtcattg gtccaacatt gccatcaatg      720
tatttggaca aaagacttga tgatgataaa gataatggtt tcaatttgta caaggctaat      780
catcacgaat gtatgaattg gctggatgac aaaccaaagg aatcagttgt atatgttgct      840
ttcggctctc ttgttaaaca tggtccagaa caagttgagg agattacaag agcacttata      900
gactctgacg taaacttttt gtgggtcatt aagcacaaag aggaggggaa actgccagaa      960
aacctttctg aagtgataaa gaccggaaaa ggtctaatcg ttgcttggtg taaacaattg     1020
gatgttttag ctcatgaatc tgtaggctgt tttgtaacac attgcggatt caactctaca     1080
ctagaagcca tttccttagg cgtacctgtc gttgcaatgc ctcagttctc cgatcagaca     1140
accaacgcta aacttttgga cgaaatacta ggggtgggtg tcagagttaa agcagacgag     1200
aatggtatcg tcagaagagg gaacctagct tcatgtatca aaatgatcat ggaagaggaa     1260
agaggagtta tcataaggaa aaacgcagtt aagtggaagg atcttgcaaa ggttgccgtc     1320
catgaaggcg gctcttcaga taatgatatt gttgaatttg tgtccgaact aatcaaagcc     1380
taa                                                                  1383
```

SEQ ID NO:4

S. rebaudiana

```
MAEQQKIKKS PHVLLIPFPL QGHINPFIQF GKRLISKGVK TTLVTTIHTL NSTLNHSNTT       60
TTSIEIQAIS DGCDEGGFMS AGESYLETFK QVGSKSLADL IKKLQSEGTT IDAIIYDSMT      120
EWVLDVAIEF GIDGGSFFTQ ACVVNSLYYH VHKGLISLPL GETVSVPGFP VLQRWETPLI      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:5

S. rebaudiana

```
atggatgcaa tggctacaac tgagaagaaa ccacacgtca tcttcatacc atttccagca       60
caaagccaca ttaaagccat gctcaaacta gcacaacttc tccaccacaa aggactccag      120
ataaccttcg tcaacaccga cttcatccac aaccagtttc ttgaatcatc gggcccacat      180
tgtctagacg gtgcaccggg tttccggttc gaaaccattc 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 actacggaag 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
aactagttac aaagttgttt cacattgtgc tttctattta agatgtaact ttgttctaat      1500
ttaatattgt ctagatgtat tgaaccataa gtttagttgg tctcaggaat tgatttttaa      1560
tgaaataatg gtcattaggg gtgagt                                           1586
```

SEQ ID NO:6

S. rebaudiana

```
atggatgcaa tggcaactac tgagaaaaag cctcatgtga tcttcattcc atttcctgca        60
caatctcaca taaaggcaat gctaaagtta gcacaactat tacaccataa gggattacag       120
ataactttcg tgaataccga cttcatccat aatcaatttc tggaatctag tggccctcat       180
tgtttggacg gagccccagg gtttagattc gaaacaattc ctgacggtgt ttcacattcc       240
ccagaggcct ccatcccaat aagagagagt ttactgaggt caatagaaac caacttttg        300
gatcgtttca ttgacttggt cacaaaactt ccagacccac caacttgcat aatctctgat       360
ggctttctgt cagtgtttac tatcgacgct gccaaaaagt tgggtatccc agttatgatg       420
tactggactc ttgctgcatg cggtttcatg ggtttctatc acatccattc tcttatcgaa       480
aagggttttg ctccactgaa agatgcatca tacttaacca acggctacct ggatactgtt       540
attgactggg taccaggtat ggaaggtata agacttaaag attttccttt ggattggtct       600
acagacctta atgataaagt attgatgttt actacagaag ctccacaaag atctcataag       660
gtttcacatc atatctttca cacctttgat gaattggaac catcaatcat caaaaccttg       720
tctctaagat acaatcatat ctacactatt ggtccattac aattacttct agatcaaatt       780
cctgaagaga aaaagcaaac tggtattaca tccttacacg gctactcttt agtgaaagag       840
gaaccagaat gttttcaatg gctacaaagt aaagagccta attctgtggt ctacgtcaac       900
ttcggaagta caacagtcat gtccttggaa gatatgactg aatttggttg gggccttgct       960
aattcaaatc attactttct atggattatc aggtccaatt tggtaatagg ggaaaacgcc      1020
gtattacctc cagaattgga ggaacacatc aaaaagagag gtttcattgc ttcctggtgt      1080
tctcaggaaa aggtattgaa acatccttct gttggtggtt tccttactca ttgcggttgg      1140
ggctctacaa tcgaatcact aagtgcagga gttccaatga tttgttggcc atattcatgg      1200
gaccaactta caaattgtag gtatatctgt aaagagtggg aagttggatt agaaatggga      1260
acaaaggtta aacgtgatga agtgaaaaga ttggttcagg agttgatggg ggaaggtggc      1320
cacaagatga gaaacaaggc caaagattgg aaggaaaaag ccagaattgc tattgctcct      1380
aacgggtcat cctctctaaa cattgataag atggtcaaag agattacagt cttagccaga      1440
aactaa                                                                 1446
```

SEQ ID NO:7

S. rebaudiana

```
MDAMATTEKK PHVIFIPFPA QSHIKAMLKL AQLLHHKGLQ ITFVNTDFIH NQFLESSGPH        60
CLDGAPGFRF 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:8

S. rebaudiana

```
atggaaaaca agaccgaaac aacagttaga cgtaggcgta gaatcattct gtttccagta        60
ccttttcaag ggcacatcaa tccaatacta caactagcca acgttttgta ctctaaaggt       120
ttttctatta caatctttca caccaatttc aacaaaccaa aaacatccaa ttacccacat       180
```

```
ttcacattca gattcatact tgataatgat ccacaagatg aacgtatttc aaacttacct       240
acccacggtc ctttagctgg aatgagaatt ccaatcatca atgaacatgg tgccgatgag       300
cttagaagag aattagagtt acttatgttg gcatccgaag aggacgagga agtctcttgt       360
ctgattactg acgctctatg gtactttgcc caatctgtgg ctgatagttt gaatttgagg       420
agattggtac taatgacatc cagtctgttt aactttcacg ctcatgttag tttaccacaa       480
tttgacgaat tgggatactt ggaccctgat gacaagacta ggttagagga acaggcctct       540
ggttttccta tgttgaaagt caaagatatc aagtctgcct attctaattg gcaaatcttg       600
aaagagatct taggaaagat gatcaaacag acaaaggctt catctggagt gatttggaac       660
agtttcaaag agttagaaga gtctgaattg gagactgtaa tcagagaaat tccagcacct       720
tcattcctga taccattacc aaaacatttg actgcttcct cttcctcttt gttggatcat       780
gacagaacag tttttcaatg gttggaccaa caaccaccta gttctgtttt gtacgtgtca       840
tttggtagta cttctgaagt cgatgaaaag gacttccttg aaatcgcaag aggcttagtc       900
gatagtaagc agtcattcct ttgggtcgtg cgtccaggtt tcgtgaaagg ctcaacatgg       960
gtcgaaccac ttccagatgg ttttctaggc gaaagaggta gaatagtcaa atgggttcct      1020
caacaggaag ttttagctca tggcgctatt ggggcattct ggactcattc cggatggaat      1080
tcaactttag aatcagtatg cgaaggggta cctatgatct tttcagattt tggtcttgat      1140
caaccactga acgcaagata catgtctgat gttttgaaag tgggtgtata tctagaaaat      1200
ggctgggaaa ggggtgaaat agctaatgca ataagacgtg ttatggttga tgaagagggg      1260
gagtatatca gacaaaacgc aagagtgctg aagcaaaagg ccgacgtttc tctaatgaag      1320
ggaggctctt catacgaatc cttagaatct cttgtttcct acatttcatc actgtaa         1377
```

SEQ ID NO:9

S. rebaudiana

```
MENKTETTVR RRRRIILFPV PFQGHINPIL QLANVLYSKG FSITIFHTNF NKPKTSNYPH        60
FTFRFILDND PQDERISNLP THGPLAGMRI PIINEHGADE LRRELELLML ASEEDEEVSC       120
LITDALWYFA QSVADSLNLR RLVLMTSSLF 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:10

```
atggctacat ctgattctat tgttgatgac aggaagcagt tgcatgtggc tactttccct        60
tggcttgctt tcggtcatat actgccttac ctacaactat caaaactgat agctgaaaaa       120
ggacataaag tgtcattcct ttcaacaact agaaacattc aaagattatc ttcccacata       180
tcaccattga ttaacgtcgt tcaattgaca cttccaagag tacaggaatt accagaagat       240
gctgaagcta caacagatgt gcatcctgaa gatatccctt acttgaaaaa ggcatccgat       300
ggattacagc ctgaggtcac tagattcctt gagcaacaca gtccagattg gatcatatac       360
gactacactc actattggtt gccttcaatt gcagcatcac taggcatttc tagggcacat       420
ttcagtgtaa ccacaccttg ggccattgct tacatgggtc catccgctga tgctatgatt       480
aacggcagtg atggtagaac taccgttgaa gatttgacaa ccccaccaaa gtggtttcca       540
tttccaacta aagtctgttg gagaaaacac gacttagcaa gactggttcc atacaaggca       600
ccaggaatct cagacggcta tagaatgggt ttagtcctta aagggtctga ctgcctattg       660
tctaagtgtt accatgagtt tgggacacaa tggctaccac ttttggaaac attacaccaa       720
gttcctgtcg taccagttgg tctattacct ccagaaatcc ctggtgatga gaaggacgag       780
acttgggttt caatcaaaaa gtggttagac gggaagcaaa aaggctcagt ggtatatgtg       840
gcactgggtt ccgaagtttt agtatctcaa acagaagttg tggaacttgc cttaggtttg       900
gaactatctg gattgccatt tgtctgggcc tacagaaaac caaaaggccc tgcaaagtcc       960
gattcagttg aattgccaga cggctttgtc gagagaacta gagatagagg gttggtatgg      1020
acttcatggg ctccacaatt gagaatcctg agtcacgaat ctgtgtgcgg tttcctaaca      1080
cattgtggtt ctggttctat agttgaagga ctgatgtttg gtcatccact tatcatgttg      1140
ccaatctttg gtgaccagcc tttgaatgca cgtctgttag aagataaaca agttggaatt      1200
gaaatcccac gtaatgagga agatggatgt ttaaccaagg agtctgtggc cagatcatta      1260
```

```
cgttccgttg tcgttgaaaa ggaaggcgaa atctacaagg ccaatgcccg tgaactttca      1320
aagatctaca atgacacaaa agtagagaag gaatatgttt ctcaatttgt agattaccta      1380
gagaaaaacg ctagagccgt agctattgat catgaatcct aa                         1422
```

SEQ ID NO:11

```
MATSDSIVDD RKQLHVATFP WLAFGHILPY LQLSKLIAEK GHKVSFLSTT RNIQRLSSHI        60
SPLINVVQLT LPRVQELPED AEATTDVHPE DIPYLKKASD GLQPEVTRFL EQHSPDWIIY       120
DYTHYWLPSI AASLGISRAH FSVTTPWAIA YMGPSADAMI NGSDGRTTVE DLTTPPKWFP       180
FPTKVCWRKH DLARLVPYKA PGISDGYRMG LVLKGSDCLL SKCYHEFGTQ WLPLLETLHQ       240
VPVVPVGLLP PEIPGDEKDE TWVSIKKWLD GKQKGSVVYV ALGSEVLVSQ TEVVELALGL       300
ELSGLPFVWA YRKPKGPAKS DSVELPDGFV ERTRDRGLVW TSWAPQLRIL SHESVCGFLT       360
HCGSGSIVEG LMFGHPLIML PIFGDQPLNA RLLEDKQVGI EIPRNEEDGC LTKESVARSL       420
RSVVVEKEGE IYKANARELS KIYNDTKVEK EYVSQFVDYL EKNARAVAID HES              473
```

SEQ ID NO:12

```
atggctactt ctgattccat cgttgacgat agaaagcaat tgcatgttgc tacttttcca        60
tggttggctt tcggtcatat tttgccatac ttgcaattgt ccaagttgat tgctgaaaag       120
ggtcacaagg tttcattctt gtctaccacc agaaacatcc aaagattgtc ctctcatatc       180
tccccattga tcaacgttgt tcaattgact ttgccaagag tccaagaatt gccagaagat       240
gctgaagcta ctactgatgt tcatccagaa gatatccctt acttgaaaaa ggcttccgat       300
ggtttacaac cagaagttac tagattcttg gaacaacatt ccccagattg gatcatctac       360
gattatactc attactggtt gccatccatt gctgcttcat tgggtatttc tagagcccat       420
ttctctgtta ctactccatg ggctattgct tatatgggtc catctgctga tgctatgatt       480
aacggttctg atggtagaac taccgttgaa gatttgacta ctccaccaaa gtggtttcca       540
tttccaacaa aagtctgttg gagaaaacac gatttggcta gattggttcc atacaaagct       600
ccaggtattt ctgatggtta cagaatgggt atggttttga aaggttccga ttgcttgttg       660
tctaagtgct atcatgaatt cggtactcaa tggttgcctt tgttggaaac attgcatcaa       720
gttccagttg ttccagtagg tttgttgcca ccagaaattc caggtgacga aaaagacgaa       780
acttgggttt ccatcaaaaa gtggttggat ggtaagcaaa agggttctgt tgtttatgtt       840
gctttgggtt ccgaagcttt ggtttctcaa accgaagttg ttgaattggc tttgggtttg       900
gaattgtctg gtttgccatt tgtttgggct tacagaaaac ctaaaggtcc agctaagtct       960
gattctgttg aattgccaga tggtttcgtt gaaagaacta gagatagagg tttggtttgg      1020
acttcttggg ctccacaatt gagaattttg tctcatgaat ccgtctgtgg tttcttgact      1080
cattgtggtt ctggttctat cgttgaaggt ttgatgtttg gtcacccatt gattatgttg      1140
ccaatctttg gtgaccaacc attgaacgct agattattgg aagataagca agtcggtatc      1200
gaaatcccaa gaaatgaaga agatggttgc ttgaccaaag aatctgttgc tagatctttg      1260
agatccgttg tcgttgaaaa agaaggtgaa atctacaagg ctaacgctag agaattgtcc      1320
aagatctaca acgataccaa ggtcgaaaaa gaatacgttt cccaattcgt tgactacttg      1380
gaaaagaatg ctagagctgt tgccattgat catgaatctt ga                         1422
```

SEQ ID NO:13

```
MATSDSIVDD RKQLHVATFP WLAFGHILPY LQLSKLIAEK GHKVSFLSTT RNIQRLSSHI        60
SPLINVVQLT LPRVQELPED AEATTDVHPE DIPYLKKASD GLQPEVTRFL EQHSPDWIIY       120
DYTHYWLPSI AASLGISRAH FSVTTPWAIA YMGPSADAMI NGSDGRTTVE DLTTPPKWFP       180
FPTKVCWRKH DLARLVPYKA PGISDGYRMG MVLKGSDCLL SKCYHEFGTQ WLPLLETLHQ       240
VPVVPVGLLP PEIPGDEKDE TWVSIKKWLD GKQKGSVVYV ALGSEALVSQ TEVVELALGL       300
ELSGLPFVWA YRKPKGPAKS DSVELPDGFV ERTRDRGLVW TSWAPQLRIL SHESVCGFLT       360
HCGSGSIVEG LMFGHPLIML PIFGDQPLNA RLLEDKQVGI EIPRNEEDGC LTKESVARSL       420
RSVVVEKEGE IYKANARELS KIYNDTKVEK EYVSQFVDYL EKNARAVAID HES              473
```

SEQ ID NO:14

O. sativa

```
atggactccg gctactcctc ctcctacgcc gccgccgccg ggatgcacgt cgtgatctgc        60
```

```
ccgtggctcg ccttcggcca cctgctcccg tgcctcgacc tcgcccagcg cctcgcgtcg       120
cggggccacc gcgtgtcgtt cgtctccacg ccgcggaaca tatcccgcct cccgccggtg       180
cgccccgcgc tcgcgccgct cgtcgccttc gtggcgctgc cgctcccgcg cgtcgagggg       240
ctccccgacg gcgccgagtc caccaacgac gtcccccacg acaggccgga catggtcgag       300
ctccaccgga gggccttcga cgggctcgcc gcgcccttct cggagttctt gggcaccgcg       360
tgcgccgact gggtcatcgt cgacgtcttc caccactggg ccgcagccgc cgctctcgag       420
cacaaggtgc catgtgcaat gatgttgttg ggctctgcac atatgatcgc ttccatagca       480
gacagacggc tcgagcgcgc ggagacagag tcgcctgcgg ctgccgggca gggacgccca       540
gcggcggcgc caacgttcga ggtggcgagg atgaagttga tacgaaccaa aggctcatcg       600
ggaatgtccc tcgccgagcg cttctccttg acgctctcga ggagcagcct cgtcgtcggg       660
cggagctgcg tggagttcga gccggagacc gtcccgctcc tgtcgacgct ccgcggtaag       720
cctattacct tccttggcct tatgccgccg ttgcatgaag gccgccgcga ggacggcgag       780
gatgccaccg tccgctggct cgacgcgcag ccggccaagt ccgtcgtgta cgtcgcgcta       840
ggcagcgagg tgccactggg agtggagaag gtccacgagc tcgcgctcgg gctggagctc       900
gccgggacgc gcttcctctg ggctcttagg aagcccactg gcgtctccga cgccgacctc       960
ctccccgccg gcttcgagga gcgcacgcgc ggccgcggcg tcgtggcgac gagatgggtt      1020
cctcagatga gcatactggc gcacgccgcc gtgggcgcgt tcctgaccca ctgcggctgg      1080
aactcgacca tcgaggggct catgttcggc cacccgctta tcatgctgcc gatcttcggc      1140
gaccagggac cgaacgcgcg gctaatcgag gcgaagaacg ccggattgca ggtggcaaga      1200
aacgacggcg atggatcgtt cgaccgagaa ggcgtcgcgg cggcgattcg tgcagtcgcg      1260
gtggaggaag aaagcagcaa agtgtttcaa gccaaagcca agaagctgca ggagatcgtc      1320
gcggacatgg cctgccatga gaggtacatc gacggattca ttcagcaatt gagatcttac      1380
aaggattga                                                             1389
```

SEQ ID NO:15

O. sativa

```
atggatagtg gctactcctc atcttatgct gctgccgctg gtatgcacgt tgtgatctgc        60
ccttggttgg cctttggtca cctgttacca tgtctggatt tagcccaaag actggcctca       120
agaggccata gagtatcatt tgtgtctact cctagaaata tctctcgttt accaccagtc       180
agacctgctc tagctcctct agttgcattc gttgctcttc cacttccaag agtagaagga       240
ttgccagacg gcgctgaatc tactaatgac gtaccacatg atagacctga catggtcgaa       300
ttgcatagaa gagcctttga tggattggca gctccatttt ctgagttcct gggcacagca       360
tgtgcagact gggttatagt cgatgtattt catcactggg ctgctgcagc cgcattggaa       420
cataaggtgc cttgtgctat gatgttgtta gggtcagcac acatgatcgc atccatagct       480
gatagaagat tggaaagagc tgaaacagaa tccccagccg cagcaggaca aggtaggcca       540
gctgccgccc caacctttga agtggctaga atgaaattga ttcgtactaa aggtagttca       600
gggatgagtc ttgctgaaag gttttctctg acattatcta gatcatcatt agttgtaggt       660
agatcctgcg tcgagttcga acctgaaaca gtacctttac tatctacttt gagaggcaaa       720
cctattactt tccttggtct aatgcctcca ttacatgaag gaaggagaga agatggtgaa       780
gatgctactg ttaggtggtt agatgcccaa cctgctaagt ctgttgttta cgttgcattg       840
ggttctgagg taccactagg ggtggaaaag gtgcatgaat tagcattagg acttgagctg       900
gccggaacaa gattcctttg ggctttgaga aaaccaaccg gtgtttctga cgccgacttg       960
ctaccagctg ggttcgaaga gagaacaaga ggccgtggtg tcgttgctac tagatgggtc      1020
ccacaaatga gtattctagc tcatgcagct gtaggggcct ttctaaccca ttgcggttgg      1080
aactcaacaa tagaaggact gatgtttggt catccactta ttatgttacc aatctttggc      1140
gatcagggac ctaacgcaag attgattgag gcaaagaacg caggtctgca ggttgcacgt      1200
aatgatggtg atggttcctt tgatagagaa ggcgttgcag ctgccatcag agcagtcgcc      1260
gttgaggaag agtcatctaa agttttccaa gctaaggcca aaaaattaca agagattgtg      1320
gctgacatgg cttgtcacga aagatacatc gatggtttca tccaacaatt gagaagttat      1380
aaagactaa                                                             1389
```

SEQ ID NO:16

O. sativa

```
MDSGYSSSYA AAAGMHVVIC PWLAFGHLLP CLDLAQRLAS RGHRVSFVST PRNISRLPPV        60
```

```
RPALAPLVAF VALPLPRVEG LPDGAESTND VPHDRPDMVE LHRRAFDGLA APFSEFLGTA       120
CADWVIVDVF HHWAAAAALE HKVPCAMMLL GSAHMIASIA DRRLERAETE SPAAAGQGRP       180
AAAPTFEVAR MKLIRTKGSS GMSLAERFSL TLSRSSLVVG RSCVEFEPET VPLLSTLRGK       240
PITFLGLMPP LHEGRREDGE DATVRWLDAQ PAKSVVYVAL GSEVPLGVEK VHELALGLEL       300
AGTRFLWALR KPTGVSDADL LPAGFEERTR GRGVVATRWV PQMSILAHAA VGAFLTHCGW       360
NSTIEGLMFG HPLIMLPIFG DQGPNARLIE AKNAGLQVAR NDGDGSFDRE GVAAAIRAVA       420
VEEESSKVFQ AKAKKLQEIV ADMACHERYI DGFIQQLRSY KD                          462
```

SEQ ID NO:17

```
MDSGYSSSYA AAAGMHVVIC PWLAFGHLLP CLDLAQRLAS RGHRVSFVST PRNISRLPPV        60
RPALAPLVAF VALPLPRVEG LPDGAESTND VPHDRPDMVE LHRRAFDGLA APFSEFLGTA       120
CADWVIVDVF HHWAAAAALE HKVPCAMMLL GSAHMIASIA DRRLERAETE SPAAAGQGRP       180
AAAPTFEVAR MKLIRTKGSS GMSLAERFSL TLSRSSLVVG RSCVEFEPET VPLLSTLRGK       240
PITFLGLLPP EIPGDEKDET WVSIKKWLDG KQKGSVVYVA LGSEALVSQT EVVELALGLE       300
LSGLPFVWAY RKPKGPAKSD SVELPDGFVE RTRDRGLVWT SWAPQLRILS HESVCGFLTH       360
CGSGSIVEGL MFGHPLIMLP IFGDQPLNAR LLEDKQVGIE IARNDGDGSF DREGVAAAIR       420
AVAVEEEESSK VFQAKAKKLQ EIVADMACHE RYIDGFIQQL RSYKD                      465
```

SEQ ID NO:18

```
MATSDSIVDD RKQLHVATFP WLAFGHILPY LQLSKLIAEK GHKVSFLSTT RNIQRLSSHI        60
SPLINVVQLT LPRVQELPED AEATTDVHPE DIPYLKKASD GLQPEVTRFL EQHSPDWIIY       120
DYTHYWLPSI AASLGISRAH FSVTTPWAIA YMGPSADAMI NGSDGRTTVE DLTTPPKWFP       180
FPTKVCWRKH DLARLVPYKA PGISDGYRMG MVLKGSDCLL SKCYHEFGTQ WLPLLETLHQ       240
VPVVPVGLMP PLHEGRREDG EDATVRWLDA QPAKSVVYVA LGSEVPLGVE KVHELALGLE       300
LAGTRFLWAL RKPTGVSDAD LLPAGFEERT RGRGVVATRW VPQMSILAHA AVGAFLTHCG       360
WNSTIEGLMF GHPLIMLPIF GDQGPNARLI EAKNAGLQVP RNEEDGCLTK ESVARSLRSV       420
VVEKEGEIYK ANARELSKIY NDTKVEKEYV SQFVDYLEKN ARAVAIDHES                  470
```

SEQ ID NO:19

Synechococcus sp.

```
atggctttgg taaacccaac cgctcttttc tatggtacct ctatcagaac aagacctaca        60
aacttactaa atccaactca aaagctaaga ccagtttcat catcttcctt accttctttc       120
tcatcagtta gtgcgattct tactgaaaaa catcaatcta atccttctga gaacaacaat       180
ttgcaaactc atctagaaac tcctttcaac tttgatagtt atatgttgga aaaagtcaac       240
atggttaacg aggcgcttga tgcatctgtc ccactaaaag acccaatcaa aatccatgaa       300
tccatgagat actctttatt ggcaggcggt aagagaatca gaccaatgat gtgtattgca       360
gcctgcgaaa tagtcggagg taatatcctt aacgccatgc cagccgcatg tgccgtggaa       420
atgattcata ctatgtcttt ggtgcatgac gatcttccat gtatggataa tgatgacttc       480
agaagaggta aacctatttc acacaaggtc tacggggagg aaatggcagt attgaccggc       540
gatgctttac taagtttatc tttcgaacat atagctactg ctacaaaggg tgtatcaaag       600
gatagaatcg tcagagctat aggggagttg gcccgttcag ttggctccga aggtttagtg       660
gctggacaag ttgtagatat cttgtcagag ggtgctgatg ttggattaga tcacctagaa       720
tacattcaca tccacaaaac agcaatgttg cttgagtcct cagtagttat tggcgctatc       780
atgggaggag gatctgatca gcagatcgaa aagttgagaa aattcgctag atctattggt       840
ctactattcc aagttgtgga tgacattttg gatgttacaa aatctaccga agagttgggg       900
aaaacagctg gtaaggattt gttgacagat aagacaactt acccaaagtt gttaggtata       960
gaaaagtcca gagaatttgc cgaaaaactt aacaaggaag cacaagagca attaagtggc      1020
tttgatagac gtaaggcagc tcctttgatc gcgttagcca actacaatgc gtaccgtcaa      1080
aattga                                                                 1086
```

SEQ ID NO:20

Synechococcus sp.

```
MALVNPTALF YGTSIRTRPT NLLNPTQKLR PVSSSSLPSF SSVSAILTEK HQSNPSENNN        60
```

```
LQTHLETPFN FDSYMLEKVN MVNEALDASV PLKDPIKIHE SMRYSLLAGG KRIRPMMCIA       120
ACEIVGGNIL NAMPAACAVE MIHTMSLVHD DLPCMDNDDF RRGKPISHKV YGEEMAVLTG       180
DALLSLSFEH IATATKGVSK DRIVRAIGEL ARSVGSEGLV AGQVVDILSE GADVGLDHLE       240
YIHIHKTAML LESSVVIGAI MGGGSDQQIE KLRKFARSIG LLFQVVDDIL DVTKSTEELG       300
KTAGKDLLTD KTTYPKLLGI EKSREFAEKL NKEAQEQLSG FDRRKAAPLI ALANYNAYRQ       360
N                                                                       361
```

SEQ ID NO:21

```
atggctgagc aacaaatatc taacttgctg tctatgtttg atgcttcaca tgctagtcag        60
aaattagaaa ttactgtcca aatgatggac acataccatt acagagaaac gcctccagat       120
tcctcatctt ctgaaggcgg ttcattgtct agatacgacg agagaagagt ctctttgcct       180
ctcagtcata atgctgcctc tccagatatt gtatcacaac tatgtttttc cactgcaatg       240
tcttcagagt tgaatcacag atggaaatct caaagattaa aggtggccga ttctccttac       300
aactatatcc taacattacc atcaaaagga attagaggtg cctttatcga ttccctgaac       360
gtatggttgg aggttccaga ggatgaaaca tcagtcatca aggaagttat tggtatgctc       420
cacaactctt cattaatcat tgatgacttc caagataatt ctccacttag aagaggaaag       480
ccatctaccc atacagtctt cggccctgcc caggctatca atactgctac ttacgttata       540
gttaaagcaa tcgaaaagat acaagacata gtgggacacg atgcattggc agatgttacg       600
ggtactatta caactatttt ccaaggtcag gccatggact tgtggtggac agcaaatgca       660
atcgttccat caatacagga atacttactt atggtaaacg ataaaaccgg tgctctcttt       720
agactgagtt tggagttgtt agctctgaat tccgaagcca gtatttctga ctctgcttta       780
gaaagtttat ctagtgctgt ttccttgcta ggtcaatact tccaaatcag agacgactat       840
atgaacttga tcgataacaa gtatacagat cagaaaggct tctgcgaaga tcttgatgaa       900
ggcaagtact cactaacact tattcatgcc ctccaaactg attcatccga tctactgacc       960
aacatccttt caatgagaag agtgcaagga aagttaacgg cacaaaagag atgttggttc      1020
tggaaatga                                                              1029
```

SEQ ID NO:22

```
MAEQQISNLL SMFDASHASQ KLEITVQMMD TYHYRETPPD SSSSEGGSLS RYDERRVSLP        60
LSHNAASPDI VSQLCFSTAM SSELNHRWKS QRLKVADSPY NYILTLPSKG IRGAFIDSLN       120
VWLEVPEDET SVIKEVIGML HNSSLIIDDF QDNSPLRRGK PSTHTVFGPA QAINTATYVI       180
VKAIEKIQDI VGHDALADVT GTITTIFQGQ AMDLWWTANA IVPSIQEYLL MVNDKTGALF       240
RLSLELLALN SEASISDSAL ESLSSAVSLL GQYFQIRDDY MNLIDNKYTD QKGFCEDLDE       300
GKYSLTLIHA LQTDSSDLLT NILSMRRVQG KLTAQKRCWF WK                          342
```

SEQ ID NO:23

```
atggaaaaga ctaaggagaa agcagaacgt atcttgctgg agccatacag atacttatta        60
caactaccag gaaagcaagt ccgttctaaa ctatcacaag cgttcaatca ctggttaaaa       120
gttcctgaag ataagttaca aatcattatt gaagtcacag aaatgctaca caatgcttct       180
ttactgatcg atgatataga ggattcttcc aaactgagaa gaggttttcc tgtcgctcat       240
tccatatacg gggtaccaag tgtaatcaac tcagctaatt acgtctactt cttgggattg       300
gaaaaagtat tgacattaga tcatccagac gctgtaaagc tattcaccag acaacttctt       360
gaattgcatc aaggtcaagg tttggatatc tattggagag acacttatac ttgcccaaca       420
gaagaggagt acaaagcaat ggttctacaa aagactggcg gtttgttcgg acttgccgtt       480
ggtctgatgc aacttttctc tgattacaag gaggacttaa agcctctgtt ggataccttg       540
ggcttgtttt tccagattag agatgactac gctaacttac attcaaagga atattcagaa       600
aacaaatcat tctgtgaaga tttgactgaa gggaagttta gttttccaac aatccacgcc       660
atttggtcaa gaccagaatc tactcaagtg caaaacattc tgcgtcagag aacagagaat       720
attgacatca aaaagtattg tgttcagtac ttggaagatg ttggttcttt tgcttacaca       780
agacatacac ttagagaatt agaggcaaaa gcatacaagc aaatagaagc ctgtggaggc       840
aatccttctc tagtggcatt ggttaaacat ttgtccaaaa tgttcaccga ggaaaacaag       900
taa                                                                    903
```

SEQ ID NO:24

```
MEKTKEKAER ILLEPYRYLL QLPGKQVRSK LSQAFNHWLK VPEDKLQIII EVTEMLHNAS        60
LLIDDIEDSS KLRRGFPVAH SIYGVPSVIN SANYVYFLGL EKVLTLDHPD AVKLFTRQLL       120
ELHQGQGLDI YWRDTYTCPT EEEYKAMVLQ KTGGLFGLAV GLMQLFSDYK EDLKPLLDTL       180
GLFFQIRDDY ANLHSKEYSE NKSFCEDLTE GKFSFPTIHA IWSRPESTQV QNILRQRTEN       240
IDIKKYCVQY LEDVGSFAYT RHTLRELEAK AYKQIEACGG NPSLVALVKH LSKMFTEENK       300
```

SEQ ID NO:25

```
atggcaagat tctattttct taacgcacta ttgatggtta tctcattaca atcaactaca        60
gccttcactc cagctaaact tgcttatcca acaacaacaa cagctctaaa tgtcgcctcc       120
gccgaaactt ctttcagtct agatgaatac ttggcctcta agataggacc tatagagtct       180
gccttggaag catcagtcaa atccagaatt ccacagaccg ataagatctg cgaatctatg       240
gcctactctt tgatggcagg aggcaagaga attagaccag tgttgtgtat cgctgcatgt       300
gagatgttcg gtggatccca agatgtcgct atgcctactg ctgtggcatt agaaatgata       360
cacacaatgt ctttgattca tgatgatttg ccatccatgg ataacgatga cttgagaaga       420
ggtaaaccaa caaaccatgt cgttttcggc gaagatgtag ctattcttgc aggtgactct       480
ttattgtcaa cttccttcga gcacgtcgct agagaaacaa aaggagtgtc agcagaaaag       540
atcgtggatg ttatcgctag attaggcaaa tctgttggtg ccgagggcct tgctggcggt       600
caagttatgg acttagaatg tgaagctaaa ccaggtacca cattagacga cttgaaatgg       660
attcatatcc ataaaaccgc tacattgtta caagttgctg tagcttctgg tgcagttcta       720
ggtggtgcaa ctcctgaaga ggttgctgca tgcgagttgt ttgctatgaa tataggtctt       780
gcctttcaag ttgccgacga tatccttgat gtaaccgctt catcagaaga tttgggtaaa       840
actgcaggca aagatgaagc tactgataag acaacttacc caaagttatt aggattagaa       900
gagagtaagg catacgcaag acaactaatc gatgaagcca aggaaagttt ggctcctttt       960
ggagatagag ctgccccttt attggccatt gcagatttca ttattgatag aaagaattga      1020
```

SEQ ID NO:26

```
MARFYFLNAL LMVISLQSTT AFTPAKLAYP TTTTALNVAS AETSFSLDEY LASKIGPIES        60
ALEASVKSRI PQTDKICESM AYSLMAGGKR IRPVLCIAAC EMFGGSQDVA MPTAVALEMI       120
HTMSLIHDDL PSMDNDDLRR GKPTNHVVFG EDVAILAGDS LLSTSFEHVA RETKGVSAEK       180
IVDVIARLGK SVGAEGLAGG QVMDLECEAK PGTTLDDLKW IHIHKTATLL QVAVASGAVL       240
GGATPEEVAA CELFAMNIGL AFQVADDILD VTASSEDLGK TAGKDEATDK TTYPKLLGLE       300
ESKAYARQLI DEAKESLAPF GDRAAPLLAI ADFIIDRKN                              339
```

SEQ ID NO:27

```
atgcacttag caccacgtag agtccctaga ggtagaagat caccacctga cagagttcct        60
gaaagacaag gtgccttggg tagaagacgt ggagctggct ctactggctg tgcccgtgct       120
gctgctggtg ttcaccgtag aagaggagga ggcgaggctg atccatcagc tgctgtgcat       180
agaggctggc aagccggtgg tggcaccggt ttgcctgatg aggtggtgtc taccgcagcc       240
gccttagaaa tgtttcatgc ttttgcttta atccatgatg atatcatgga tgatagtgca       300
actagaagag gctccccaac tgttcacaga gccctagctg atcgtttagg cgctgctctg       360
gacccagatc aggccggtca actaggagtt tctactgcta tcttggttgg agatctggct       420
ttgacatggt ccgatgaatt gttatacgct ccattgactc cacatagact ggcagcagta       480
ctaccattgg taacagctat gagagctgaa accgttcatg gccaatatct tgatataact       540
agtgctagaa gacctgggac cgatacttct cttgcattga gaatagccag atataagaca       600
gcagcttaca caatggaacg tccactgcac attggtgcag ccctggctgg ggcaagacca       660
gaactattag cagggctttc agcatacgcc ttgccagctg gagaagcctt ccaattggca       720
gatgacctgc taggcgtctt cggtgatcca agacgtacag ggaaacctga cctagatgat       780
cttagaggtg gaaagcatac tgtcttagtc gccttggcaa gagaacatgc cactccagaa       840
cagagacaca cattggatac attattgggt acaccaggtc ttgatagaca aggcgcttca       900
agactaagat gcgtattggt agcaactggt gcaagagccg aagccgaaag acttattaca       960
gagagaagag atcaagcatt aactgcattg aacgcattaa cactgccacc tcctttagct      1020
gaggcattag caagattgac attagggtct acagctcatc ctgcctaa                   1068
```

SEQ ID NO:28

```
MHLAPRRVPR GRRSPPDRVP ERQGALGRRR GAGSTGCARA AAGVHRRRGG GEADPSAAVH        60
RGWQAGGGTG LPDEVVSTAA ALEMFHAFAL IHDDIMDDSA TRRGSPTVHR ALADRLGAAL       120
DPDQAGQLGV STAILVGDLA LTWSDELLYA PLTPHRLAAV LPLVTAMRAE TVHGQYLDIT       180
SARRPGTDTS LALRIARYKT AAYTMERPLH IGAALAGARP ELLAGLSAYA LPAGEAFQLA       240
DDLLGVFGDP RRTGKPDLDD LRGGKHTVLV ALAREHATPE QRHTLDTLLG TPGLDRQGAS       300
RLRCVLVATG ARAEAERLIT ERRDQALTAL NALTLPPPLA EALARLTLGS TAHPA             355
```

SEQ ID NO:29

```
atgtcatatt tcgataacta cttcaatgag atagttaatt ccgtgaacga catcattaag        60
tcttacatct ctggcgacgt accaaaacta tacgaagcct cctaccattt gtttacatca       120
ggaggaaaga gactaagacc attgatcctt acaatttctt ctgatctttt cggtggacag       180
agagaaagag catactatgc tggcgcagca atcgaagttt tgcacacatt cactttggtt       240
cacgatgata tcatggatca agataacatt cgtagaggtc ttcctactgt acatgtcaag       300
tatggcctac ctttggccat tttagctggt gacttattgc atgcaaaagc ctttcaattg       360
ttgactcagg cattgagagg tctaccatct gaaactatca tcaaggcgtt tgatatcttt       420
acaagatcta tcattatcat atcagaaggt caagctgtcg atatggaatt cgaagataga       480
attgatatca aggaacaaga gtatttggat atgatatctc gtaaaaccgc tgccttattc       540
tcagcttctt cttccattgg ggcgttgata gctggagcta atgataacga tgtgagatta       600
atgtccgatt tcggtacaaa tcttgggatc gcatttcaaa ttgtagatga tatacttggt       660
ttaacagctg atgaaaaaga gctaggaaaa cctgttttca gtgatatcag agaaggtaaa       720
aagaccatat tagtcattaa gactttagaa ttgtgtaagg aagacgagaa aaagattgtg       780
ttaaaagcgc taggcaacaa gtcagcatca aaggaagagt tgatgagttc tgctgacata       840
atcaaaaagt actcattgga ttacgcctac aacttagctg agaaatacta caaaaacgcc       900
atcgattctc taaatcaagt ttcaagtaaa agtgatattc cagggaaggc attgaaatat       960
cttgctgaat tcaccatcag aagacgtaag taa                                    993
```

SEQ ID NO:30

```
MSYFDNYFNE IVNSVNDIIK SYISGDVPKL YEASYHLFTS GGKRLRPLIL TISSDLFGGQ        60
RERAYYAGAA IEVLHTFTLV HDDIMDQDNI RRGLPTVHVK YGLPLAILAG DLLHAKAFQL       120
LTQALRGLPS ETIIKAFDIF TRSIIIISEG QAVDMEFEDR IDIKEQEYLD MISRKTAALF       180
SASSSIGALI AGANDNDVRL MSDFGTNLGI AFQIVDDILG LTADEKELGK PVFSDIREGK       240
KTILVIKTLE LCKEDEKKIV LKALGNKSAS KEELMSSADI IKKYSLDYAY NLAEKYYKNA       300
IDSLNQVSSK SDIPGKALKY LAEFTIRRRK                                        330
```

SEQ ID NO:31

```
atggtcgcac aaactttcaa cctggatacc tacttatccc aaagacaaca acaagttgaa        60
gaggccctaa gtgctgctct tgtgccagct tatcctgaga gaatatacga agctatgaga       120
tactccctcc tggcaggtgg caaaagatta agacctatct tatgtttagc tgcttgcgaa       180
ttggcaggtg gttctgttga acaagccatg ccaactgcgt gtgcacttga aatgatccat       240
acaatgtcac taattcatga tgacctgcca gccatggata acgatgattt cagaagagga       300
aagccaacta atcacaaggt gttcggggaa gatatagcca tcttagcggg tgatgcgctt       360
ttagcttacg cttttgaaca tattgcttct caaacaagag gagtaccacc tcaattggtg       420
ctacaagtta ttgctagaat cggacacgcc gttgctgcaa caggcctcgt tggaggccaa       480
gtcgtagacc ttgaatctga aggtaaagct atttccttag aaacattgga gtatattcac       540
tcacataaga ctggagcctt gctggaagca tcagttgtct caggcggtat tctcgcaggg       600
gcagatgaag agcttttggc cagattgtct cattacgcta gagatatagg cttggctttt       660
caaatcgtcg atgatatcct ggatgttact gctacatctg aacagttggg gaaaaccgct       720
ggtaaagacc aggcagccgc aaaggcaact tatccaagtc tattgggttt agaagcctct       780
agacagaaag cggaagagtt gattcaatct gctaaggaag ccttaagacc ttacggttca       840
caagcagagc cactcctagc gctggcagac ttcatcacac gtcgtcagca ttaa              894
```

SEQ ID NO:32

```
MVAQTFNLDT YLSQRQQQVE EALSAALVPA YPERIYEAMR YSLLAGGKRL RPILCLAACE        60
LAGGSVEQAM PTACALEMIH TMSLIHDDLP AMDNDDFRRG KPTNHKVFGE DIAILAGDAL       120
LAYAFEHIAS QTRGVPPQLV LQVIARIGHA VAATGLVGGQ VVDLESEGKA ISLETLEYIH       180
SHKTGALLEA SVVSGGILAG ADEELLARLS HYARDIGLAF QIVDDILDVT ATSEQLGKTA       240
GKDQAAAKAT YPSLLGLEAS RQKAEELIQS AKEALRPYGS QAEPLLALAD FITRRQH          297
```

SEQ ID NO:33

```
atgaaaaccg ggtttatctc accagcaaca gtatttcatc acagaatctc accagcgacc        60
actttcagac atcacttatc acctgctact acaaactcta caggcattgt cgccttaaga       120
gacatcaact tcagatgtaa agcagtttct aaagagtact ctgatctgtt gcagaaagat       180
gaggcttctt tcacaaaatg ggacgatgac aaggtgaaag atcatcttga taccaacaaa       240
aacttatacc caaatgatga gattaaggaa tttgttgaat cagtaaaggc tatgttcggt       300
agtatgaatg acggggagat aaacgtctct gcatacgata ctgcatgggt tgctttggtt       360
caagatgtcg atggatcagg tagtcctcag ttcccttctt ctttagaatg gattgccaac       420
aatcaattgt cagatggatc atggggagat catttgctgt tctcagctca cgatagaatc       480
atcaacacat tagcatgcgt tattgcactt acaagttgga atgttcatcc ttctaagtgt       540
gaaaaaggtt tgaattttct gagagaaaac atttgcaaat tagaagatga aaacgcagaa       600
catatgccaa ttggttttga agtaacattc ccatcactaa ttgatatcgc gaaaaagttg       660
aacattgaag tacctgagga tactccagca cttaaagaga tctacgcacg tagagatatc       720
aagttaacta agatcccaat ggaagttctt cacaaggtac ctactacttt gttacattct       780
ttggaaggaa tgcctgattt ggagtgggaa aaactgttaa agctacaatg taaagatggt       840
agtttcttgt tttccccatc tagtaccgca ttcgccctaa tgcaaacaaa agatgagaaa       900
tgcttacagt atctaacaaa tatcgtcact aagttcaacg gtggcgtgcc taatgtgtac       960
ccagtcgatt tgtttgaaca tatttgggtt gttgatagac tgcagagatt ggggattgcc      1020
agatacttca aatcagagat aaaagattgt gtagagtata tcaataagta ctggaccaaa      1080
aatggaattt gttgggctag aaatactcac gttcaagata tcgatgatac agccatggga      1140
ttcagagtgt tgagagcgca cggttatgac gtcactccag atgtttttag acaatttgaa      1200
aaagatggta aattcgtttg ctttgcaggg caatcaacac aagccgtgac aggaatgttt      1260
aacgtttaca gagcctctca aatgttgttc ccaggggaga gaattttgga agatgccaaa      1320
aagttctctt acaattactt aaaggaaaag caaagtacca acgaattgct ggataaatgg      1380
ataatcgcta aagatctacc tggtgaagtt ggttatgctc tggatatccc atggtatgct      1440
tccttaccaa gattggaaac tcgttattac cttgaacaat acggcggtga agatgatgtc      1500
tggataggca agacattata cagaatgggt tacgtgtcca ataacacata tctagaaatg      1560
gcaaagctgg attacaataa ctatgttgca gtccttcaat tagaatggta cacaatacaa      1620
caatggtacg tcgatattgg tatagagaag ttcgaatctg acaacatcaa gtcagtcctg      1680
```

SEQ ID NO:34

```
MKTGFISPAT VFHHRISPAT TFRHHLSPAT TNSTGIVALR DINFRCKAVS KEYSDLLQKD        60
EASFTKWDDD KVKDHLDTNK NLYPNDEIKE FVESVKAMFG SMNDGEINVS AYDTAWVALV       120
QDVDGSGSPQ FPSSLEWIAN NQLSDGSWGD HLLFSAHDRI INTLACVIAL TSWNVHPSKC       180
EKGLNFLREN ICKLEDENAE HMPIGFEVTF PSLIDIAKKL NIEVPEDTPA LKEIYARRDI       240
KLTKIPMEVL HKVPTTLLHS LEGMPDLEWE KLLKLQCKDG SFLFSPSSTA FALMQTKDEK       300
CLQYLTNIVT KFNGGVPNVY PVDLFEHIWV VDRLQRLGIA RYFKSEIKDC VEYINKYWTK       360
NGICWARNTH VQDIDDTAMG FRVLRAHGYD VTPDVFRQFE KDGKFVCFAG QSTQAVTGMF       420
NVYRASQMLF PGERILEDAK KFSYNYLKEK QSTNELLDKW IIAKDLPGEV GYALDIPWYA       480
SLPRLETRYY LEQYGGEDDV WIGKTLYRMG YVSNNTYLEM AKLDYNNYVA VLQLEWYTIQ       540
QWYVDIGIEK FESDNIKSVL VSYYLAAASI FEPERSKERI AWAKTTILVD KITSIFDSSQ       600
SSKEDITAFI DKFRNKSSSK KHSINGEPWH EVMVALKKTL HGFALDALMT HSQDIHPQLH       660
QAWEMWLTKL QDGVDVTAEL MVQMINMTAG RWVSKELLTH PQYQRLSTVT NSVCHDITKL       720
HNFKENSTTV DSKVQELVQL VFSDTPDDLD QDMKQTFLTV MKTFYYKAWC DPNTINDHIS       780
KVFEIVI                                                                 787
```

SEQ ID NO:35

```
atgcctgatg cacacgatgc tccacctcca caaataagac agagaacact agtagatgag        60
gctacccaac tgctaactga gtccgcagaa gatgcatggg gtgaagtcag tgtgtcagaa       120
tacgaaacag caaggctagt tgcccatgct acatggttag gtggacacgc cacaagagtg       180
gccttccttc tggagagaca acacgaagac gggtcatggg gtccaccagg tggatatagg       240
ttagtcccta cattatctgc tgttcacgca ttattgacat gtcttgcctc tcctgctcag       300
gatcatggcg ttccacatga tagactttta agagctgttg acgcaggctt gactgccttg       360
agaagattgg ggacatctga ctccccacct gatactatag cagttgagct ggttatccca       420
tctttgctag agggcattca acacttactg gaccctgctc atcctcatag tagaccagcc       480
ttctctcaac atagaggctc tcttgtttgt cctggtggac tagatgggag aactctagga       540
gctttgagat cacacgccgc agcaggtaca ccagtaccag gaaaagtctg gcacgcttcc       600
gagactttgg gcttgagtac cgaagctgct tctcacttgc aaccagccca aggtataatc       660
ggtggctctg ctgctgccac agcaacatgg ctaaccaggg ttgcaccatc tcaacagtca       720
gattctgcca gaagatacct tgaggaatta caacacagat actctggccc agttccttcc       780
attaccccta tcacatactt cgaaagagca tggttattga acaattttgc agcagccggt       840
gttccttgtg aggctccagc tgctttgttg gattccttag aagcagcact tacaccacaa       900
ggtgctcctg ctggagcagg attgcctcca gatgctgatg atacagccgc tgtgttgctt       960
gcattggcaa cacatgggag aggtagaaga ccagaagtac tgatggatta caggactgac      1020
gggtatttcc aatgctttat tggggaaagg actccatcaa tttcaacaaa cgctcacgta      1080
ttggaaacat tagggcatca tgtggcccaa catccacaag atagagccag atacggatca      1140
gccatggata ccgcatcagc ttggctgctg gcagctcaaa agcaagatgg ctcttggtta      1200
gataaatggc atgcctcacc atactacgct actgtttgtt gcacacaagc cctagccgct      1260
catgcaagtc ctgcaactgc accagctaga cagagagctg tcagatgggt tttagccaca      1320
caaagatccg atggcggttg gggtctatgg cattcaactg ttgaagagac tgcttatgcc      1380
ttacagatct tggccccacc ttctggtggt ggcaatatcc cagtccaaca agcacttact      1440
agaggcagag caagattgtg tggagccttg ccactgactc ctttatggca tgataaggat      1500
ttgtatactc cagtaagagt agtcagagct gccagagctg ctgctctgta cactaccaga      1560
gatctattgt taccaccatt gtaa                                             1584
```

SEQ ID NO:36

```
MPDAHDAPPP QIRQRTLVDE ATQLLTESAE DAWGEVSVSE YETARLVAHA TWLGGHATRV        60
AFLLERQHED GSWGPPGGYR LVPTLSAVHA LLTCLASPAQ DHGVPHDRLL RAVDAGLTAL       120
RRLGTSDSPP DTIAVELVIP SLLEGIQHLL DPAHPHSRPA FSQHRGSLVC PGGLDGRTLG       180
ALRSHAAAGT PVPGKVWHAS ETLGLSTEAA SHLQPAQGII GGSAAATATW LTRVAPSQQS       240
DSARRYLEEL QHRYSGPVPS ITPITYFERA WLLNNFAAAG VPCEAPAALL DSLEAALTPQ       300
GAPAGAGLPP DADDTAAVLL ALATHGRGRR PEVLMDYRTD GYFQCFIGER TPSISTNAHV       360
LETLGHHVAQ HPQDRARYGS AMDTASAWLL AAQKQDGSWL DKWHASPYYA TVCCTQALAA       420
HASPATAPAR QRAVRWVLAT QRSDGGWGLW HSTVEETAYA LQILAPPSGG GNIPVQQALT       480
RGRARLCGAL PLTPLWHDKD LYTPVRVVRA ARAAALYTTR DLLLPPL                     527
```

SEQ ID NO:37

```
atgaacgccc tatccgaaca cattttgtct gaattgagaa gattattgtc tgaaatgagt        60
gatggcggat ctgttggtcc atctgtgtat gatacggccc aggccctaag attccacggt       120
aacgtaacag gtagacaaga tgcatatgct tggttgatcg cccagcaaca agcagatgga       180
ggttggggct ctgccgactt tccactcttt agacatgctc caacatgggc tgcacttctc       240
gcattacaaa gagctgatcc acttcctggc gcagcagacg cagttcagac cgcaacaaga       300
ttcttgcaaa gacaaccaga tccatacgct catgccgttc ctgaggatgc ccctattggt       360
gctgaactga tcttgcctca gttttgtgga gaggctgctt ggttgttggg aggtgtggcc       420
ttccctagac acccagccct attaccatta agacaggctt gtttagtcaa actgggtgca       480
gtcgccatgt tgccttcagg acacccattg ctccactcct gggaggcatg gggtacttct       540
ccaacaacag cctgtccaga cgatgatggt tctataggta tctcaccagc agctacagcc       600
gcctggagag cccaggctgt gaccagaggc tcaactcctc aagtgggcag agctgacgca       660
tacttacaaa tggcttcaag agcaacgaga tcaggcatag aaggagtctt ccctaatgtt       720
tggcctataa acgtattcga accatgctgg tcactgtaca ctctccatct tgccggtctg       780
```

```
ttcgcccatc cagcactggc tgaggctgta agagttatcg ttgctcaact tgaagcaaga       840
ttgggagtgc atggcctcgg accagcttta cattttgctg ccgacgctga tgatactgca       900
gttgccttat gcgttctgca tttggctggc agagatcctg cagttgacgc attgagacat       960
tttgaaattg gtgagctctt tgttacattc ccaggagaga gaaatgctag tgtctctacg      1020
aacattcacg ctcttcatgc tttgagattg ttaggtaaac cagctgccgg agcaagtgca      1080
tacgtcgaag caaatagaaa tccacatggt ttgtgggaca acgaaaaatg gcacgtttca      1140
tggctttatc caactgcaca cgccgttgca gctctagctc aaggcaagcc tcaatggaga      1200
gatgaaagag cactagccgc tctactacaa gctcaaagag atgatggtgg ttggggagct      1260
ggtagaggat ccactttcga ggaaaccgcc tacgctcttt tcgctttaca cgttatggac      1320
ggatctgagg aagccacagg cagaagaaga atcgctcaag tcgtcgcaag agccttagaa      1380
tggatgctag ctagacatgc cgcacatgga ttaccacaaa caccactctg gattggtaag      1440
gaattgtact gtcctactag agtcgtaaga gtagctgagc tagctggcct gtggttagca      1500
ttaagatggg gtagaagagt attagctgaa ggtgctggtg ctgcacctta a               1551
```

SEQ ID NO:38

```
MNALSEHILS ELRRLLSEMS DGGSVGPSVY DTAQALRFHG NVTGRQDAYA WLIAQQQADG        60
GWGSADFPLF RHAPTWAALL ALQRADPLPG AADAVQTATR FLQRQPDPYA HAVPEDAPIG       120
AELILPQFCG EAAWLLGGVA FPRHPALLPL RQACLVKLGA VAMLPSGHPL LHSWEAWGTS       180
PTTACPDDDG SIGISPAATA AWRAQAVTRG STPQVGRADA YLQMASRATR SGIEGVFPNV       240
WPINVFEPCW SLYTLHLAGL FAHPALAEAV RVIVAQLEAR LGVHGLGPAL HFAADADDTA       300
VALCVLHLAG RDPAVDALRH FEIGELFVTF PGERNASVST NIHALHALRL LGKPAAGASA       360
YVEANRNPHG LWDNEKWHVS WLYPTAHAVA ALAQGKPQWR DERALAALLQ AQRDDGGWGA       420
GRGSTFEETA YALFALHVMD GSEEATGRRR IAQVVARALE WMLARHAAHG LPQTPLWIGK       480
ELYCPTRVVR VAELAGLWLA LRWGRRVLAE GAGAAP                                 516
```

SEQ ID NO:39

Z. mays

```
atggttttgt cttcttcttg tactacagta ccacacttat cttcattagc tgtcgtgcaa        60
cttggtcctt ggagcagtag gattaaaaag aaaaccgata ctgttgcagt accagccgct       120
gcaggaaggt ggagaagggc cttggctaga gcacagcaca catcagaatc cgcagctgtc       180
gcaaagggca gcagtttgac ccctatagtg agaactgacg ctgagtcaag gagaacaaga       240
tggccaaccg atgacgatga cgccgaacct ttagtggatg agatcagggc aatgcttact       300
tccatgtctg atggtgacat ttccgtgagc gcatacgata cagcctgggt cggattggtt       360
ccaagattag acggcggtga aggtcctcaa tttccagcag ctgtgagatg gataagaaat       420
aaccagttgc ctgacggaag ttggggcgat gccgcattat tctctgccta tgacaggctt       480
atcaataccc ttgcctgcgt tgtaactttg acaaggtggt ccctagaacc agagatgaga       540
ggtagaggac tatctttttt gggtaggaac atgtggaaat tagcaactga agatgaagag       600
tcaatgccta ttggcttcga attagcattt ccatctttga tagagcttgc taagagccta       660
ggtgtccatg acttccctta tgatcaccag gccctacaag gaatctactc ttcaagagag       720
atcaaaatga agaggattcc aaaagaagtg atgcataccg ttccaacatc aatattgcac       780
agtttggagg gtatgcctgg cctagattgg gctaaactac ttaaactaca gagcagcgac       840
ggaagttttt tgttctcacc agctgccact gcatatgctt taatgaatac cggagatgac       900
aggtgtttta gctacatcga tagaacagta aagaaattca acggcggcgt ccctaatgtt       960
tatccagtgg atctatttga acatatttgg gccgttgata gacttgaaag attaggaatc      1020
tccaggtact tccaaaagga gatcgaacaa tgcatggatt atgtaaacag gcattggact      1080
gaggacggta tttgttgggc aaggaactct gatgtcaaag aggtggacga cacagctatg      1140
gcctttagac ttcttaggtt gcacggctac agcgtcagtc ctgatgtgtt taaaaacttc      1200
gaaaaggacg gtgaattttt cgcatttgtc ggacagtcta atcaagctgt taccggtatg      1260
tacaacttaa acagagcaag ccagatatcc ttcccaggcg aggatgtgct tcatagagct      1320
ggtgccttct catatgagtt cttgaggaga aaagaagcag agggagcttt gagggacaag      1380
tggatcattt ctaaagatct acctggtgaa gttgtgtata ctttggattt tccatggtac      1440
ggcaacttac ctagagtcga ggccagagac tacctagagc aatacggagg tggtgatgac      1500
gtttggattg gcaagacatt gtataggatg ccacttgtaa acaatgatgt atatttggaa      1560
ttggcaagaa tggatttcaa ccactgccag gctttgcatc agttagagtg gcaaggacta      1620
```

```
aaaagatggt atactgaaaa taggttgatg gactttggtg tcgcccaaga agatgccctt      1680
agagcttatt ttcttgcagc cgcatctgtt tacgagcctt gtagagctgc cgagaggctt      1740
gcatgggcta gagccgcaat actagctaac gccgtgagca cccacttaag aaatagccca      1800
tcattcagag aaaggttaga gcattctctt aggtgtagac ctagtgaaga gacagatggc      1860
tcctggttta actcctcaag tggctctgat gcagttttag taaaggctgt cttaagactt      1920
actgattcat tagccaggga agcacagcca atccatggag gtgacccaga agatattata      1980
cacaagttgt taagatctgc ttgggccgag tgggttaggg aaaaggcaga cgctgccgat      2040
agcgtgtgca atggtagttc tgcagtagaa caagagggat caagaatggt ccatgataaa      2100
cagacctgtc tattattggc tagaatgatc gaaatttctg ccggtagggc agctggtgaa      2160
gcagccagtg aggacggcga tagaagaata attcaattaa caggctccat ctgcgacagt      2220
cttaagcaaa aaatgctagt ttcacaggac cctgaaaaaa atgaagagat gatgtctcac      2280
gtggatgacg aattgaagtt gaggattaga gagttcgttc aatatttgct tagactaggt      2340
gaaaaaaaga ctggatctag cgaaaccagg caaacatttt taagtatagt gaaatcatgt      2400
tactatgctg ctcattgccc acctcatgtc gttgatagac acattagtag agtgattttc      2460
gagccagtaa gtgccgcaaa gtaaccgcgg                                       2490
```

SEQ ID NO:40

Z. mays

```
MVLSSSCTTV PHLSSLAVVQ LGPWSSRIKK KTDTVAVPAA AGRWRRALAR AQHTSESAAV        60
AKGSSLTPIV RTDAESRRTR WPTDDDDAEP LVDEIRAMLT SMSDGDISVS AYDTAWVGLV       120
PRLDGGEGPQ FPAAVRWIRN NQLPDGSWGD AALFSAYDRL INTLACVVTL TRWSLEPEMR       180
GRGLSFLGRN MWKLATEDEE SMPIGFELAF PSLIELAKSL GVHDFPYDHQ ALQGIYSSRE       240
IKMKRIPKEV MHTVPTSILH SLEGMPGLDW AKLLKLQSSD GSFLFSPAAT AYALMNTGDD       300
RCFSYIDRTV KKFNGGVPNV YPVDLFEHIW AVDRLERLGI SRYFQKEIEQ CMDYVNRHWT       360
EDGICWARNS DVKEVDDTAM AFRLLRLHGY SVSPDVFKNF EKDGEFFAFV GQSNQAVTGM       420
YNLNRASQIS FPGEDVLHRA GAFSYEFLRR KEAEGALRDK WIISKDLPGE VVYTLDFPWY       480
GNLPRVEARD YLEQYGGGDD VWIGKTLYRM PLVNNDVYLE LARMDFNHCQ ALHQLEWQGL       540
KRWYTENRLM DFGVAQEDAL RAYFLAAASV YEPCRAAERL AWARAAILAN AVSTHLRNSP       600
SFRERLEHSL RCRPSEETDG SWFNSSSGSD AVLVKAVLRL TDSLAREAQP IHGGDPEDII       660
HKLLRSAWAE WVREKADAAD SVCNGSSAVE QEGSRMVHDK QTCLLLARMI EISAGRAAGE       720
AASEDGDRRI IQLTGSICDS LKQKMLVSQD PEKNEEMMSH VDDELKLRIR EFVQYLLRLG       780
EKKTGSSETR QTFLSIVKSC YYAAHCPPHV VDRHISRVIF EPVSAAK                     827
```

SEQ ID NO:41

```
cttcttcact aaatacttag acagagaaaa cagagctttt taaagccatg tctcttcagt        60
atcatgttct aaactccatt ccaagtacaa cctttctcag ttctactaaa acaacaatat       120
cttcttcttt ccttaccatc tcaggatctc ctctcaatgt cgctagagac aaatccagaa       180
gcggttccat acattgttca aagcttcgaa ctcaagaata cattaattct caagaggttc       240
aacatgattt gcctctaata catgagtggc aacagcttca aggagaagat gctcctcaga       300
ttagtgttgg aagtaatagt aatgcattca aagaagcagt gaagagtgtg aaaacgatct       360
tgagaaacct aacggacggg gaaattacga tatcggctta cgatacagct tgggttgcat       420
tgatcgatgc cggagataaa actccggcgt ttccctccgc cgtgaaatgg atcgccgaga       480
accaactttc cgatggttct tggggagatg cgtatctctt ctcttatcat gatcgtctca       540
tcaataccct tgcatgcgtc gttgctctaa gatcatggaa tctctttcct catcaatgca       600
acaaaggaat cacgtttttc cgggaaaata ttgggaagct agaagacgaa aatgatgagc       660
atatgccaat cggattcgaa gtagcattcc catcgttgct tgagatagct cgaggaataa       720
acattgatgt accgtacgat tctccggtct taaaagatat atacgccaag aaagagctaa       780
agcttacaag gataccaaaa gagataatgc acaagatacc aacaacattg ttgcatagtt       840
tggaggggat gcgtgattta gattgggaaa agctcttgaa acttcaatct caagacggat       900
ctttcctctt ctctccttcc tctaccgctt ttgcattcat gcagacccga gacagtaact       960
gcctcgagta tttgcgaaat gccgtcaaac gtttcaatgg aggagttccc aatgtctttc      1020
ccgtggatct tttcgagcac atatggatag tggatcggtt acaacgttta gggatatcga      1080
gatactttga agaagagatt aaagagtgtc ttgactatgt ccacagatat tggaccgaca      1140
atggcatatg ttgggctaga tgttcccatg tccaagacat cgatgataca gccatggcat      1200
```

```
ttaggctctt aagacaacat ggataccaag tgtccgcaga tgtattcaag aactttgaga      1260
aagagggaga gtttttctgc tttgtggggc aatcaaacca agcagtaacc ggtatgttca      1320
acctataccg ggcatcacaa ttggcgtttc caagggaaga gatattgaaa aacgccaaag      1380
agttttctta taattatctg ctagaaaaac gggagagaga ggagttgatt gataagtgga      1440
ttataatgaa agacttacct ggcgagattg ggtttgcgtt agagattcca tggtacgcaa      1500
gcttgcctcg agtagagacg agattctata ttgatcaata tggtggagaa aacgacgttt      1560
ggattggcaa gactctttat aggatgccat acgtgaacaa taatggatat ctggaattag      1620
caaaacaaga ttacaacaat tgccaagctc agcatcagct cgaatgggac atattccaaa      1680
agtggtatga agaaaatagg ttaagtgagt ggggtgtgcg cagaagtgag cttctcgagt      1740
gttactactt agcggctgca actatatttg aatcagaaag gtcacatgag agaatggttt      1800
gggctaagtc aagtgtattg gttaaagcca tttcttcttc ttttggggaa tcctctgact      1860
ccagaagaag cttctccgat cagtttcatg aatacattgc caatgctcga cgaagtgatc      1920
atcactttaa tgacaggaac atgagattgg accgaccagg atcggttcag gccagtcggc      1980
ttgccggagt gttaatcggg actttgaatc aaatgtcttt tgacctttc atgtctcatg      2040
gccgtgacgt taacaatctc ctctatctat cgtggggaga ttggatggaa aaatggaaac      2100
tatatggaga tgaaggagaa ggagagctca tggtgaagat gataattcta atgaagaaca      2160
atgacctaac taacttcttc acccacactc acttcgttcg tctcgcggaa atcatcaatc      2220
gaatctgtct tcctcgccaa tacttaaagg caaggagaaa cgatgagaag gagaagacaa      2280
taaagagtat ggagaaggag atggggaaaa tggttgagtt agcattgtcg gagagtgaca      2340
catttcgtga cgtcagcatc acgtttcttg atgtagcaaa agcattttac tactttgctt      2400
tatgtggcga tcatctccaa actcacatct ccaaagtctt gtttcaaaaa gtctagtaac      2460
ctcatcatca tcatcgatcc attaacaatc agtggatcga tgtatccata gatgcgtgaa      2520
taatatttca tgtagagaag gagaacaaat tagatcatgt agggttatca                 2570
```

SEQ ID NO:42

```
MSLQYHVLNS IPSTTFLSST KTTISSSFLT ISGSPLNVAR DKSRSGSIHC SKLRTQEYIN        60
SQEVQHDLPL IHEWQQLQGE DAPQISVGSN SNAFKEAVKS VKTILRNLTD GEITISAYDT       120
AWVALIDAGD KTPAFPSAVK WIAENQLSDG SWGDAYLFSY HDRLINTLAC VVALRSWNLF       180
PHQCNKGITF FRENIGKLED ENDEHMPIGF EVAFPSLLEI ARGINIDVPY DSPVLKDIYA       240
KKELKLTRIP KEIMHKIPTT LLHSLEGMRD LDWEKLLKLQ SQDGSFLFSP SSTAFAFMQT       300
RDSNCLEYLR NAVKRFNGGV PNVFPVDLFE HIWIVDRLQR LGISRYFEEE IKECLDYVHR       360
YWTDNGICWA RCSHVQDIDD TAMAFRLLRQ HGYQVSADVF KNFEKEGEFF CFVGQSNQAV       420
TGMFNLYRAS QLAFPREEIL KNAKEFSYNY LLEKREREEL IDKWIIMKDL PGEIGFALEI       480
PWYASLPRVE TRFYIDQYGG ENDVWIGKTL YRMPYVNNNG YLELAKQDYN NCQAQHQLEW       540
DIFQKWYEEN RLSEWGVRRS ELLECYYLAA ATIFESERSH ERMVWAKSSV LVKAISSSFG       600
ESSDSRRSFS DQFHEYIANA RRSDHHFNDR NMRLDRPGSV QASRLAGVLI GTLNQMSFDL       660
FMSHGRDVNN LLYLSWGDWM EKWKLYGDEG EGELMVKMII LMKNNDLTNF FTHTHFVRLA       720
EIINRICLPR QYLKARRNDE KEKTIKSMEK EMGKMVELAL SESDTFRDVS ITFLDVAKAF       780
YYFALCGDHL QTHISKVLFQ KV                                                802
```

SEQ ID NO:43

```
atgaatttga gtttgtgtat agcatctcca ctattgacca aatctaatag accagctgct        60
ttatcagcaa ttcatacagc tagtacatcc catggtggcc aaaccaaccc tacgaatctg       120
ataatcgata cgaccaagga gagaatacaa aaacaattca aaaatgttga aatttcagtt       180
tcttcttatg atactgcgtg ggttgccatg gttccatcac ctaattctcc aaagtctcca       240
tgtttcccag aatgtttgaa ttggctgatt aacaaccagt tgaatgatgg atcttggggt       300
ttagtcaatc acacgcacaa tcacaaccat ccacttttga aagattcttt atcctcaact       360
ttggcttgca tcgtggccct aaagagatgg aacgtaggtg aggatcagat taacaagggg       420
cttagtttca ttgaatctaa cttggcttcc gcgactgaaa aatctcaacc atctccaata       480
ggattcgata tcatctttcc aggtctgtta gagtacgcca aaaatctaga tatcaactta       540
ctgtctaagc aaactgattt ctcactaatg ttacacaaga gagaattaga acaaaagaga       600
tgtcattcaa acgaaatgga tggttaccta gcttatatct ctgaaggtct tggtaatctt       660
tacgattgga atatggtgaa aaagtaccag atgaaaaatg gctcagtttt caattcccct       720
tctgcaactg cggcagcatt cattaaccat caaaatccag gatgcctgaa ctatttgaat       780
```

```
tcactactag acaaattcgg caacgcagtt ccaactgtat accctcacga tttgtttatc       840
agattgagta tggtggatac aattgaaaga cttggtatat cccaccactt tagagtcgag       900
atcaaaaatg ttttggatga gacataccgt tgttgggtgg agagagatga acaaatcttt       960
atggatgttg tgacgtgcgc gttggccttt agattgttgc gtattaacgg ttacgaagtt      1020
agtccagatc cacttgccga aattacaaac gaattagctt taaaggatga atacgccgct      1080
cttgaaacat atcatgcgtc acatatcctt taccaagagg acttatcatc tggaaaacaa      1140
attcttaaat ctgctgattt cctgaaggaa atcatatcca ctgatagtaa tagactgtcc      1200
aaactgatcc ataaagaggt tgaaaatgca cttaagttcc ctattaacac cggcttagaa      1260
cgtattaaca caagacgtaa catccagctt tacaacgtag acaatactag aatcttgaaa      1320
accacttacc attcttccaa catatcaaac actgattacc taagattagc tgttgaagat      1380
ttctacacat gtcagtctat ctatagagaa gagctgaaag gattagagag atgggtcgtt      1440
gagaataagc tagatcaatt gaaatttgcc agacaaaaga cagcttattg ttacttctca      1500
gttgccgcca ctttatcaag tccagaattg tcagatgcac gtatttcttg ggctaaaaac      1560
ggaattttga caactgttgt tgatgatttc tttgatattg gcgggacaat cgacgaattg      1620
acaaacctga ttcaatgcgt tgaaaagtgg aatgtcgatg tcgataaaga ctgttgctca      1680
gaacatgtta gaatactgtt cttggctctg aaagatgcta tctgttggat cggggatgag      1740
gctttcaaat ggcaagctag agatgtgacg tctcacgtca ttcaaacctg gctagaactg      1800
atgaactcta tgttgagaga agcaatttgg actagagatg catacgttcc tacattaaac      1860
gagtatatgg aaaacgctta tgtctccttt gctttgggtc ctatcgttaa gcctgccata      1920
tactttgtag gaccaaagct atccgaggaa atcgtcgaat catcagaata ccataacttg      1980
ttcaagttaa tgtccacaca aggcagatta cttaatgata ttcattcttt caaaagagag      2040
tttaaggaag gaaagttaaa tgctgttgct ctgcatcttt ctaatggcga aagtggtaaa      2100
gtcgaagagg aagtagttga ggaaatgatg atgatgatca aaaacaagag aaaggagttg      2160
atgaaactaa tcttcgaaga gaacggttca attgttccta gagcatgtaa ggatgcattt      2220
tggaacatgt gtcatgtgct aaacttttc tacgcaaacg acgatggttt tactgggaac       2280
acaatactag atacagtaaa agacatcata tacaaccctt tggtcttagt aaacgaaaac      2340
gaggagcaaa gataa                                                       2355
```

SEQ ID NO:44

```
MNLSLCIASP LLTKSNRPAA LSAIHTASTS HGGQTNPTNL IIDTTKERIQ KQFKNVEISV        60
SSYDTAWVAM VPSPNSPKSP CFPECLNWLI NNQLNDGSWG LVNHTHNHNH PLLKDSLSST       120
LACIVALKRW NVGEDQINKG LSFIESNLAS ATEKSQPSPI GFDIIFPGLL EYAKNLDINL       180
LSKQTDFSLM LHKRELEQKR CHSNEMDGYL AYISEGLGNL YDWNMVKKYQ MKNGSVFNSP       240
SATAAAFINH QNPGCLNYLN SLLDKFGNAV PTVYPHDLFI RLSMVDTIER LGISHHFRVE       300
IKNVLDETYR CWVERDEQIF MDVVTCALAF RLLRINGYEV SPDPLAEITN ELALKDEYAA       360
LETYHASHIL YQEDLSSGKQ ILKSADFLKE IISTDSNRLS KLIHKEVENA LKFPINTGLE       420
RINTRRNIQL YNVDNTRILK TTYHSSNISN TDYLRLAVED FYTCQSIYRE ELKGLERWVV       480
ENKLDQLKFA RQKTAYCYFS VAATLSSPEL SDARISWAKN GILTTVVDDF FDIGGTIDEL       540
TNLIQCVEKW NVDVDKDCCS EHVRILFLAL KDAICWIGDE AFKWQARDVT SHVIQTWLEL       600
MNSMLREAIW TRDAYVPTLN EYMENAYVSF ALGPIVKPAI YFVGPKLSEE IVESSEYHNL       660
FKLMSTQGRL LNDIHSFKRE FKEGKLNAVA LHLSNGESGK VEEEVVEEMM MMIKNKRKEL       720
MKLIFEENGS IVPRACKDAF WNMCHVLNFF YANDDGFTGN TILDTVKDII YNPLVLVNEN       780
EEQR                                                                    784
```

SEQ ID NO:45

```
atgaatctgt ccctttgtat agctagtcca ctgttgacaa aatcttctag accaactgct        60
ctttctgcaa ttcatactgc cagtactagt catggaggtc aaacaaaccc aacaaatttg       120
ataatcgata ctactaagga gagaatccaa aagctattca aaaatgttga aatctcagta       180
tcatcttatg acaccgcatg ggttgcaatg gtgccatcac ctaattcccc aaaaagtcca       240
tgttttccag agtgcttgaa ttggttaatc aataatcagt taaacgatgg ttcttggggt       300
ttagtcaacc acactcataa ccacaatcat ccattattga aggactcttt atcatcaaca       360
ttagcctgta ttgttgcatt gaaaagatgg aatgtaggtg aagatcaaat caacaagggt       420
ttatcattca tagaatccaa tctagcttct gctaccgaca aatcacaacc atctccaatc       480
gggttcgaca taatcttccc tggtttgctg gagtatgcca aaaaccttga tatcaactta       540
```

```
ctgtctaaac aaacagattt ctctttgatg ctacacaaaa gagagttaga gcagaaaaga       600
tgccattcta acgaaattga cgggtactta gcatatatct cagaaggttt gggtaatttg       660
tatgactgga acatggtcaa aaagtatcag atgaaaaatg gatccgtatt caattctcct       720
tctgcaactg ccgcagcatt cattaatcat caaaaccctg ggtgtcttaa ctacttgaac       780
tcactattag ataagtttgg aaatgcagtt ccaacagtct atcctttgga cttgtacatc       840
agattatcta tggttgacac tatagagaga ttaggtattt ctcatcattt cagagttgag       900
atcaaaaatg ttttggacga gacatacaga tgttgggtcg aaagagatga gcaaatcttt       960
atggatgtcg tgacctgcgc tctggctttt agattgctaa ggatacacgg atacaaagta      1020
tctcctgatc aactggctga gattacaaac gaactggctt tcaaagacga atacgccgca      1080
ttagaaacat accatgcatc ccaaatactt taccaggaag acctaagttc aggaaaacaa      1140
atcttgaagt ctgcagattt cctgaaaggc attctgtcta cagatagtaa taggttgtct      1200
aaattgatac acaaggaagt agaaaacgca ctaaagtttc ctattaacac tggtttagag      1260
agaatcaata ctaggagaaa cattcagctg tacaacgtag ataatacaag gattcttaag      1320
accacctacc atagttcaaa catttccaac acctattact taagattagc tgtcgaagac      1380
ttttacactt gtcaatcaat ctacagagag gagttaaagg gcctagaaag atgggtagtt      1440
caaaacaagt tggatcaact gaagtttgct agacagaaga cagcatactg ttatttctct      1500
gttgctgcta ccctttcatc cccagaattg tctgatgcca gaataagttg ggccaaaaat      1560
ggtattctta caactgtagt cgatgatttc tttgatattg gaggtactat tgatgaactg      1620
acaaatctta ttcaatgtgt tgaaaagtgg aacgtggatg tagataagga ttgctgcagt      1680
gaacatgtga gaatactttt cctggctcta aaagatgcaa tatgttggat tggcgacgag      1740
gccttcaagt ggcaagctag agatgttaca tctcatgtca tccaaacttg gcttgaactg      1800
atgaactcaa tgctaagaga agcaatctgg acaagagatg catacgttcc aacattgaac      1860
gaatacatgg aaaacgctta cgtctcattt gccttgggtc ctattgttaa gccagccata      1920
tactttgttg ggccaaagtt atccgaagag attgttgagt cttccgaata tcataaccta      1980
ttcaagttaa tgtcaacaca aggcagactt ctgaacgata tccactcctt caaaagagaa      2040
ttcaaggaag gtaagctaaa cgctgttgct ttgcacttgt ctaatggtga atctggcaaa      2100
gtggaagagg aagtcgttga ggaaatgatg atgatgatca aaaacaagag aaaggaattg      2160
atgaaattga ttttcgagga aaatggttca atcgtaccta gagcttgtaa agatgctttt      2220
tggaatatgt gccatgttct taacttcttt tacgctaatg atgatggctt cactggaaat      2280
acaatattgg atacagttaa agatatcatc tacaacccac ttgttttggt caatgagaac      2340
gaggaacaaa gataa                                                       2355
```

SEQ ID NO:46

```
MNLSLCIASP LLTKSSRPTA LSAIHTASTS HGGQTNPTNL IIDTTKERIQ KLFKNVEISV        60
SSYDTAWVAM VPSPNSPKSP CFPECLNWLI NNQLNDGSWG LVNHTHNHNH PLLKDSLSST       120
LACIVALKRW NVGEDQINKG LSFIESNLAS ATDKSQPSPI GFDIIFPGLL EYAKNLDINL       180
LSKQTDFSLM LHKRELEQKR CHSNEIDGYL AYISEGLGNL YDWNMVKKYQ MKNGSVFNSP       240
SATAAAFINH QNPGCLNYLN SLLDKFGNAV PTVYPLDLYI RLSMVDTIER LGISHHFRVE       300
IKNVLDETYR CWVERDEQIF MDVVTCALAF RLLRIHGYKV SPDQLAEITN ELAFKDEYAA       360
LETYHASQIL YQEDLSSGKQ ILKSADFLKG ILSTDSNRLS KLIHKEVENA LKFPINTGLE       420
RINTRRNIQL YNVDNTRILK TTYHSSNISN TYYLRLAVED FYTCQSIYRE ELKGLERWVV       480
QNKLDQLKFA RQKTAYCYFS VAATLSSPEL SDARISWAKN GILTTVVDDF FDIGGTIDEL       540
TNLIQCVEKW NVDVDKDCCS EHVRILFLAL KDAICWIGDE AFKWQARDVT SHVIQTWLEL       600
MNSMLREAIW TRDAYVPTLN EYMENAYVSF ALGPIVKPAI YFVGPKLSEE IVESSEYHNL       660
FKLMSTQGRL LNDIHSFKRE FKEGKLNAVA LHLSNGESGK VEEEVVEEMM MMIKNKRKEL       720
MKLIFEENGS IVPRACKDAF WNMCHVLNFF YANDDGFTGN TILDTVKDII YNPLVLVNEN       780
EEQR                                                                    784
```

SEQ ID NO:47

```
atggctatgc cagtgaagct aacacctgcg tcattatcct taaaagctgt gtgctgcaga        60
ttctcatccg gtggccatgc tttgagattc gggagtagtc tgccatgttg gagaaggacc       120
cctacccaaa gatctacttc ttcctctact actagaccag ctgccgaagt gtcatcaggt       180
aagagtaaac aacatgatca ggaagctagt gaagcgacta tcagacaaca attacaactt       240
gtggatgtcc tggagaatat gggaatatcc agacattttg ctgcagagat aaagtgcata       300
```

```
ctagacagaa cttacagatc ttggttacaa agacacgagg aaatcatgct ggacactatg       360
acatgtgcta tggcttttag aatcctaaga ttgaacggat acaacgtttc atcagatgaa       420
ctataccacg ttgtagaggc atctggtctg cataattctt tgggtgggta tcttaacgat       480
accagaacac tacttgaatt acacaaggct tcaacagtta gtatctctga ggatgaatct       540
atcttagatt caattggctc tagatccaga acattgctta gagaacaatt ggagtctggt       600
ggcgcactga gaaagccttc tttattcaaa gaggttgaac atgcactgga tggacctttt       660
tacaccacac ttgatagact tcatcatagg tggaatattg aaaacttcaa cattattgag       720
caacacatgt tggagactcc atacttatct aaccagcata catcaaggga tatcctagca       780
ttgtcaatta gagatttttc ctcctcacaa ttcacttatc aacaagagct acagcatctg       840
gagagttggg ttaaggaatg tagattagat caactacagt tcgcaagaca gaaattagcg       900
tacttttacc tatcagccgc aggcaccatg ttttctcctg agctttctga tgcgagaaca       960
ttatgggcca aaaacggggt gttgacaact attgttgatg atttctttga tgttgccggt      1020
tctaaagagg aattggaaaa cttagtcatg ctggtcgaaa tgtgggatga acatcacaaa      1080
gttgaattct attctgagca ggtcgaaatc atcttctctt ccatctacga ttctgtcaac      1140
caattgggtg agaaggcctc tttggttcaa gacagatcaa ttacaaaaca ccttgttgaa      1200
atatggttag acttgttaaa gtccatgatg acggaagttg aatggagact gtcaaaatac      1260
gtgcctacag aaaaggaata catgattaat gcctctctta tcttcggcct aggtccaatc      1320
gttttaccag ctttgtattt cgttggtcca aagatttcag aaagtatagt aaaggaccca      1380
gaatatgatg aattgttcaa actaatgtca acatgtggta gattgttgaa tgacgtgcaa      1440
acgttcgaaa gagaatacaa tgagggtaaa ctgaattctg tcagtctatt ggttcttcac      1500
ggaggcccaa tgtctatttc agacgcaaag aggaaattac aaaagcctat tgatacgtgt      1560
agaagagatc ttctttcttt ggtccttaga gaagagtctg tagtaccaag accatgtaag      1620
gaactattct ggaaaatgtg taaagtgtgc tatttctttt actcaacaac tgatgggttt      1680
tctagtcaag tcgaaagagc aaaagaggta gacgctgtca taaatgagcc actgaagttg      1740
caaggttctc atacactggt atctgatgtt taa                                  1773
```

SEQ ID NO:48

```
MAMPVKLTPA SLSLKAVCCR FSSGGHALRF GSSLPCWRRT PTQRSTSSST TRPAAEVSSG        60
KSKQHDQEAS EATIRQQLQL VDVLENMGIS RHFAAEIKCI LDRTYRSWLQ RHEEIMLDTM       120
TCAMAFRILR LNGYNVSSDE LYHVVEASGL HNSLGGYLND TRTLLELHKA STVSISEDES       180
ILDSIGSRSR TLLREQLESG GALRKPSLFK EVEHALDGPF YTTLDRLHHR WNIENFNIIE       240
QHMLETPYLS NQHTSRDILA LSIRDFSSSQ FTYQQELQHL ESWVKECRLD QLQFARQKLA       300
YFYLSAAGTM FSPELSDART LWAKNGVLTT IVDDFFDVAG SKEELENLVM LVEMWDEHHK       360
VEFYSEQVEI IFSSIYDSVN QLGEKASLVQ DRSITKHLVE IWLDLLKSMM TEVEWRLSKY       420
VPTEKEYMIN ASLIFGLGPI VLPALYFVGP KISESIVKDP EYDELFKLMS TCGRLLNDVQ       480
TFEREYNEGK LNSVSLLVLH GGPMSISDAK RKLQKPIDTC RRDLLSLVLR EESVVPRPCK       540
ELFWKMCKVC YFFYSTTDGF SSQVERAKEV DAVINEPLKL QGSHTLVSDV                  590
```

SEQ ID NO:49

```
atgcagaact tccatggtac aaaggaaagg atcaaaaaga tgtttgacaa gattgaattg        60
tccgtttctt cttatgatac agcctgggtt gcaatggtcc catcccctga ttgcccagaa       120
acaccttgtt ttccagaatg tactaaatgg atcctagaaa atcagttggg tgatggtagt       180
tggtcacttc ctcatggcaa tccacttcta gttaaagatg cattatcttc cactcttgct       240
tgtattctgg ctcttaaaag atggggaatc ggtgaggaac agattaacaa aggactgaga       300
ttcatagaac tcaactctgc tagtgtaacc gataacgaac aacacaaacc aattggattt       360
gacattatct ttccaggtat gattgaatac gctatagact tagacctgaa tctaccacta       420
aaaccaactg acattaactc catgttgcat cgtagagccc ttgaattgac atcaggtgga       480
ggcaaaaatc tagaaggtag aagagcttac ttggcctacg tctctgaagg aatcggtaag       540
ctgcaagatt gggaaatggc tatgaaatac caacgtaaaa acggatctct gttcaatagt       600
ccatcaacaa ctgcagctgc attcatccat atacaagatg ctgaatgcct ccactatatt       660
cgttctcttc tccagaaatt tggaaacgca gtccctacaa tataccctct cgatatctat       720
gccagacttt caatggtaga tgccctggaa cgtcttggta ttgatagaca tttcagaaag       780
gagagaaagt tcgttctgga tgaaacatac agattttggt tgcaaggaga agaggagatt       840
ttctccgata acgcaacctg tgctttggcc ttcagaatat tgagacttaa tggttacgat       900
```

```
gtctctcttg aagatcactt ctctaactct ctgggcggtt acttaaagga ctcaggagca       960
gctttagaac tgtacagagc cctccaattg tcttacccag acgagtccct cctggaaaag      1020
caaaattcta gaacttctta cttcttaaaa caaggtttat ccaatgtctc cctctgtggt      1080
gacagattgc gtaaaaacat aattggagag gtgcatgatg ctttaaactt ttccgaccac      1140
gctaacttac aaagattagc tattcgtaga aggattaagc attacgctac tgacgataca      1200
aggattctaa aaacttccta cagatgctca acaatcggta accaagattt tctaaaactt      1260
gcagtggaag atttcaatat ctgtcaatca atacaaagag aggaattcaa gcatattgaa      1320
agatgggtcg ttgaaagacg tctagacaag ttaaagttcg ctagacaaaa agaggcctat      1380
tgctatttct cagccgcagc aacattgttt gcccctgaat tgtctgatgc tagaatgtct      1440
tgggccaaaa atggtgtatt gacaactgtg gttgatgatt tcttcgatgt cggaggctct      1500
gaagaggaat tagttaactt gatagaattg atcgagcgtt gggatgtgaa tggcagtgca      1560
gatttttgta gtgaggaagt tgagattatc tattctgcta tccactcaac tatctctgaa      1620
ataggtgata agtcatttgg ctggcaaggt agagatgtaa agtctcaagt tatcaagatc      1680
tggctggact tattgaaatc aatgttaact gaagctcaat ggtcttcaaa caagtctgtt      1740
cctaccctag atgagtatat gacaaccgcc catgtttcat tcgcacttgg tccaattgta      1800
cttccagcct tatacttcgt tggcccaaag ttgtcagaag aggttgcagg tcatcctgaa      1860
ctactaaacc tctacaaagt cacatctact tgtggcagac tactgaatga ttggagaagt      1920
tttaagagag aatccgagga aggtaagctc aacgctatta gtttatacat gatccactcc      1980
ggtggtgctt ctacagaaga ggaaacaatc gaacatttca aaggtttgat tgattctcag      2040
agaaggcaac tgttacaatt ggtgttgcaa gagaaggata gtatcatacc tagaccatgt      2100
aaagatctat tttggaatat gattaagtta ttacacactt tctacatgaa agatgatggc      2160
ttcacctcaa atgagatgag gaatgtagtt aaggcaatca ttaacgaacc aatctcactg      2220
gatgaattat ga                                                         2232
```

SEQ ID NO:50

```
MSCIRPWFCP SSISATLTDP ASKLVTGEFK TTSLNFHGTK ERIKKMFDKI ELSVSSYDTA        60
WVAMVPSPDC PETPCFPECT KWILENQLGD GSWSLPHGNP LLVKDALSST LACILALKRW       120
GIGEEQINKG LRFIELNSAS VTDNEQHKPI GFDIIFPGMI EYAKDLDLNL PLKPTDINSM       180
LHRRALELTS GGGKNLEGRR AYLAYVSEGI GKLQDWEMAM KYQRKNGSLF NSPSTTAAAF       240
IHIQDAECLH YIRSLLQKFG NAVPTIYPLD IYARLSMVDA LERLGIDRHF RKERKFVLDE       300
TYRFWLQGEE EIFSDNATCA LAFRILRLNG YDVSLEDHFS NSLGGYLKDS GAALELYRAL       360
QLSYPDESLL EKQNSRTSYF LKQGLSNVSL CGDRLRKNII GEVHDALNFP DHANLQRLAI       420
RRRIKHYATD DTRILKTSYR CSTIGNQDFL KLAVEDFNIC QSIQREEFKH IERWVVERRL       480
DKLKFARQKE AYCYFSAAAT LFAPELSDAR MSWAKNGVLT TVVDDFFDVG GSEEELVNLI       540
ELIERWDVNG SADFCSEEVE IIYSAIHSTI SEIGDKSFGW QGRDVKSHVI KIWLDLLKSM       600
LTEAQWSSNK SVPTLDEYMT TAHVSFALGP IVLPALYFVG PKLSEEVAGH PELLNLYKVM       660
STCGRLLNDW RSFKRESEEG KLNAISLYMI HSGGASTEEE TIEHFKGLID SQRRQLLQLV       720
LQEKDSIIPR PCKDLFWNMI KLLHTFYMKD DGFTSNEMRN VVKAIINEPI SLDEL            775
```

SEQ ID NO:51

A. thaliana

```
atgtctatca accttcgctc ctccggttgt tcgtctccga tctcagctac tttggaacga        60
ggattggact cagaagtaca gacaagagct aacaatgtga gctttgagca aacaaaggag       120
aagattagga agatgttgga gaaagtggag ctttctgttt cggcctacga tactagttgg       180
gtagcaatgg ttccatcacc gagctcccaa aatgctccac ttttcccaca gtgtgtgaaa       240
tggttattgg ataatcaaca tgaagatgga tcttggggac ttgataacca tgaccatcaa       300
tctcttaaga aggatgtgtt atcatctaca ctggctagta tcctcgcgtt aaagaagtgg       360
ggaattggtg aaagacaaat aaacaagggt ctccagttta ttgagctgaa ttctgcatta       420
gtcactgatg aaaccataca gaaaccaaca gggtttgata ttatatttcc tgggatgatt       480
aaatatgcta gagatttgaa tctgacgatt ccattgggct cagaagtggt ggatgacatg       540
atacgaaaaa gagatctgga tcttaaatgt gatagtgaaa agttttcaaa gggaagagaa       600
gcatatctgg cctatgtttt agaggggaca agaaacctaa aagattggga tttgatagtc       660
aaatatcaaa ggaaaaatgg gtcactgttt gattctccag ccacaacagc agctgctttt       720
actcagtttg ggaatgatgg ttgtctccgt tatctctgtt ctctccttca gaaattcgag       780
```

```
gctgcagttc cttcagttta tccatttgat caatatgcac gccttagtat aattgtcact       840
cttgaaagct taggaattga tagagatttc aaaaccgaaa tcaaaagcat attggatgaa       900
acctatagat attggcttcg tggggatgaa gaaatatgtt tggacttggc cacttgtgct       960
ttggctttcc gattattgct tgctcatggc tatgatgtgt cttacgatcc gctaaaacca      1020
tttgcagaag aatctggttt ctctgatact ttggaaggat atgttaagaa tacgttttct      1080
gtgttagaat tatttaaggc tgctcaaagt tatccacatg aatcagcttt gaagaagcag      1140
tgttgttgga ctaaacaata tctggagatg gaattgtcca gctgggttaa gacctctgtt      1200
cgagataaat acctcaagaa agaggtcgag gatgctcttg cttttccctc ctatgcaagc      1260
ctagaaagat cagatcacag gagaaaaata ctcaatggtt ctgctgtgga aaacaccaga      1320
gttacaaaaa cctcatatcg tttgcacaat atttgcacct ctgatatcct gaagttagct      1380
gtggatgact tcaatttctg ccagtccata caccgtgaag aaatggaacg tcttgatagg      1440
tggattgtgg agaatagatt gcaggaactg aaatttgcca gacagaagct ggcttactgt      1500
tatttctctg gggctgcaac tttattttct ccagaactat ctgatgctcg tatatcgtgg      1560
gccaaaggtg gagtacttac aacggttgta gacgacttct ttgatgttgg agggtccaaa      1620
gaagaactgg aaaacctcat acacttggtc gaaaagtggg atttgaacgg tgttcctgag      1680
tacagctcag aacatgttga gatcatattc tcagttctaa gggacaccat tctcgaaaca      1740
ggagacaaag cattcaccta tcaaggacgc aatgtgacac accacattgt gaaaatttgg      1800
ttggatctgc tcaagtctat gttgagagaa gccgagtggt ccagtgacaa gtcaacacca      1860
agcttggagg attacatgga aaatgcgtac atatcatttg cattaggacc aattgtcctc      1920
ccagctacct atctgatcgg acctccactt ccagagaaga cagtcgatag ccaccaatat      1980
aatcagctct acaagctcgt gagcactatg ggtcgtcttc taaatgacat acaaggtttt      2040
aagagagaaa gcgcggaagg gaagctgaat gcggtttcat tgcacatgaa acacgagaga      2100
gacaatcgca gcaaagaagt gatcatagaa tcgatgaaag gtttagcaga gagaaagagg      2160
gaagaattgc ataagctagt tttggaggag aaaggaagtg tggttccaag ggaatgcaaa      2220
gaagcgttct tgaaaatgag caaagtgttg aacttatttt acaggaagga cgatggattc      2280
acatcaaatg atctgatgag tcttgttaaa tcagtgatct acgagcctgt tagcttacag      2340
aaagaatctt taacttga                                                    2358
```

SEQ ID NO:52

A. thaliana

```
MSINLRSSGC SSPISATLER GLDSEVQTRA NNVSFEQTKE KIRKMLEKVE LSVSAYDTSW        60
VAMVPSPSSQ NAPLFPQCVK WLLDNQHEDG SWGLDNHDHQ SLKKDVLSST LASILALKKW       120
GIGERQINKG LQFIELNSAL VTDETIQKPT GFDIIFPGMI KYARDLNLTI PLGSEVVDDM       180
IRKRDLDLKC DSEKFSKGRE AYLAYVLEGT RNLKDWDLIV KYQRKNGSLF DSPATTAAAF       240
TQFGNDGCLR YLCSLLQKFE AAVPSVYPFD QYARLSIIVT LESLGIDRDF KTEIKSILDE       300
TYRYWLRGDE EICLDLATCA LAFRLLLAHG YDVSYDPLKP FAEESGFSDT LEGYVKNTFS       360
VLELFKAAQS YPHESALKKQ CCWTKQYLEM ELSSWVKTSV RDKYLKKEVE DALAFPSYAS       420
LERSDHRRKI LNGSAVENTR VTKTSYRLHN ICTSDILKLA VDDFNFCQSI HREEMERLDR       480
WIVENRLQEL KFARQKLAYC YFSGAATLFS PELSDARISW AKGGVLTTVV DDFFDVGGSK       540
EELENLIHLV EKWDLNGVPE YSSEHVEIIF SVLRDTILET GDKAFTYQGR NVTHHIVKIW       600
LDLLKSMLRE AEWSSDKSTP SLEDYMENAY ISFALGPIVL PATYLIGPPL PEKTVDSHQY       660
NQLYKLVSTM GRLLNDIQGF KRESAEGKLN AVSLHMKHER DNRSKEVIIE SMKGLAERKR       720
EELHKLVLEE KGSVVPRECK EAFLKMSKVL NLFYRKDDGF TSNDLMSLVK SVIYEPVSLQ       780
KESLT                                                                   785
```

SEQ ID NO:53

```
atggaatttg atgaaccatt ggttgacgaa gcaagatctt tagtgcagcg tactttacaa        60
gattatgatg acagatacgg cttcggtact atgtcatgtg ctgcttatga tacagcctgg       120
gtgtctttag ttacaaaaac agtcgatggg agaaaacaat ggcttttccc agagtgtttt       180
gaatttctac tagaaacaca atctgatgcc ggaggatggg aaatcgggaa ttcagcacca       240
atcgacggta tattgaatac agctgcatcc ttacttgctc taaaacgtca cgttcaaact       300
gagcaaatca tccaacctca acatgaccat aaggatctag caggtagagc tgaacgtgcc       360
gctgcatctt tgagagcaca attggctgca ttggatgtgt ctacaactga acacgtcggt       420
tttgagataa ttgttcctgc aatgctagac ccattagaag ccgaagatcc atctctagtt       480
```

```
ttcgattttc cagctaggaa acctttgatg aagattcatg atgctaagat gagtagattc       540
aggccagaat acttgtatgg caaacaacca atgaccgcct tacattcatt agaggctttc       600
ataggcaaaa tcgacttcga taaggtaaga caccaccgta cccatgggtc tatgatgggt       660
tctccttcat ctaccgcagc ctacttaatg cacgcttcac aatgggatgg tgactcagag       720
gcttacctta gacacgtgat taaacacgca gcagggcagg gaactggtgc tgtaccatct       780
gctttcccat caacacattt tgagtcatct tggattctta ccacattgtt tagagctgga       840
ttttcagctt ctcatcttgc ctgtgatgag ttgaacaagt tggtcgagat acttgagggc       900
tcattcgaga aggaaggtgg ggcaatcggt tacgctccag ggtttcaagc agatgttgat       960
gatactgcta aaacaataag tacattagca gtccttggaa gagatgctac accaagacaa      1020
atgatcaagg tatttgaagc taatacacat tttagaacat accctggtga aagagatcct      1080
tctttgacag ctaattgtaa tgctctatca gccttactac accaaccaga tgcagcaatg      1140
tatggatctc aaattcaaaa gattaccaaa tttgtctgtg actattggtg gaagtctgat      1200
ggtaagatta aagataagtg gaacacttgc tacttgtacc catctgtctt attagttgag      1260
gttttggttg atcttgttag tttattggag cagggtaaat tgcctgatgt tttggatcaa      1320
gagcttcaat acagagtcgc catcacattg ttccaagcat gtttaaggcc attactagac      1380
caagatgccg aaggatcatg gaacaagtct atcgaagcca cagcctacgg catccttatc      1440
ctaactgaag ctaggagagt ttgtttcttc gacagattgt ctgagccatt gaatgaggca      1500
atccgtagag gtatcgcttt cgccgactct atgtctggaa ctgaagctca gttgaactac      1560
atttggatcg aaaaggttag ttacgcacct gcattattga ctaaatccta tttgttagca      1620
gcaagatggg ctgctaagtc tcctttaggc gcttccgtag gctcttcttt gtggactcca      1680
ccaagagaag gattggataa gcatgtcaga ttattccatc aagctgagtt attcagatcc      1740
cttccagaat gggaattaag agcctccatg attgaagcag ctttgttcac accacttcta      1800
agagcacata gactagacgt tttccctaga caagatgtag gtgaagacaa atatcttgat      1860
gtagttccat tcttttggac tgccgctaac aacagagata gaacttacgc ttccactcta      1920
ttcctttacg atatgtgttt tatcgcaatg ttaaacttcc agttagacga attcatggag      1980
gccacagccg gtatcttatt cagagatcat atggatgatt tgaggcaatt gattcatgat      2040
cttttggcag agaaaacttc cccaaagagt tctggtagaa gtagtcaggg cacaaaagat      2100
gctgactcag gtatagagga agacgtgtca atgtccgatt cagcttcaga ttcccaggat      2160
agaagtccag aatacgactt ggttttcagt gcattgagta cctttacaaa acatgtcttg      2220
caacacccat ctatacaaag tgcctctgta tgggatagaa aactacttgc tagagagatg      2280
aaggcttact tacttgctca tatccaacaa gcagaagatt caactccatt gtctgaattg      2340
aaagatgtgc ctcaaaagac tgatgtaaca agagtttcta catctactac taccttcttt      2400
aactgggtta gaacaacttc cgcagaccat atatcctgcc catactcctt ccactttgta      2460
gcatgccatc taggcgcagc attgtcacct aaagggtcta acggtgattg ctatccttca      2520
gctggtgaga agttcttggc agctgcagtc tgcagacatt tggccaccat gtgtagaatg      2580
tacaacgatc ttggatcagc tgaacgtgat tctgatgaag gtaatttgaa ctccttggac      2640
ttccctgaat tcgccgattc cgcaggaaac ggagggatag aaattcagaa ggccgctcta      2700
ttaaggttag ctgagtttga gagagattca tacttagagg ccttccgtcg tttacaagat      2760
gaatccaata gagttcacgg tccagccggt ggtgatgaag ccagattgtc cagaaggaga      2820
atggcaatcc ttgaattctt cgcccagcag gtagatttgt acggtcaagt atacgtcatt      2880
agggatattt ccgctcgtat tcctaaaaac gaggttgaga aaaagagaaa attggatgat      2940
gctttcaatt ga                                                          2952
```

SEQ ID NO:54

```
MEFDEPLVDE ARSLVQRTLQ DYDDRYGFGT MSCAAYDTAW VSLVTKTVDG RKQWLFPECF        60
EFLLETQSDA GGWEIGNSAP IDGILNTAAS LLALKRHVQT EQIIQPQHDH KDLAGRAERA       120
AASLRAQLAA LDVSTTEHVG FEIIVPAMLD PLEAEDPSLV FDFPARKPLM KIHDAKMSRF       180
RPEYLYGKQP MTALHSLEAF IGKIDFDKVR HHRTHGSMMG SPSSTAAYLM HASQWDGDSE       240
AYLRHVIKHA AGQGTGAVPS AFPSTHFESS WILTTLFRAG FSASHLACDE LNKLVEILEG       300
SFEKEGGAIG YAPGFQADVD DTAKTISTLA VLGRDATPRQ MIKVFEANTH FRTYPGERDP       360
SLTANCNALS ALLHQPDAAM YGSQIQKITK FVCDYWWKSD GKIKDKWNTC YLYPSVLLVE       420
VLVDLVSLLE QGKLPDVLDQ ELQYRVAITL FQACLRPLLD QDAEGSWNKS IEATAYGILI       480
LTEARRVCFF DRLSEPLNEA IRRGIAFADS MSGTEAQLNY IWIEKVSYAP ALLTKSYLLA       540
ARWAAKSPLG ASVGSSLWTP PREGLDKHVR LFHQAELFRS LPEWELRASM IEAALFTPLL       600
RAHRLDVFPR QDVGEDKYLD VVPFFWTAAN NRDRTYASTL FLYDMCFIAM LNFQLDEFME       660
```

ATAGILFRDH MDDLRQLIHD LLAEKTSPKS SGRSSQGTKD ADSGIEEDVS MSDSASDSQD 720
RSPEYDLVFS ALSTFTKHVL QHPSIQSASV WDRKLLAREM KAYLLAHIQQ AEDSTPLSEL 780
KDVPQKTDVT RVSTSTTTFF NWVRTTSADH ISCPYSFHFV ACHLGAALSP KGSNGDCYPS 840
AGEKFLAAAV CRHLATMCRM YNDLGSAERD SDEGNLNSLD FPEFADSAGN GGIEIQKAAL 900
LRLAEFERDS YLEAFRRLQD ESNRVHGPAG GDEARLSRRR MAILEFFAQQ VDLYGQVYVI 960
RDISARIPKN EVEKKRKLDD AFN 983

SEQ ID NO:55

atggcttcta gtacacttat ccaaaacaga tcatgtggcg tcacatcatc tatgtcaagt 60
tttcaaatct tcagaggtca accactaaga tttcctggca ctagaacccc agctgcagtt 120
caatgcttga aaaagaggag atgccttagg ccaaccgaat ccgtactaga atcatctcct 180
ggctctggtt catatagaat agtaactggc ccttctggaa ttaaccctag ttctaacggg 240
cacttgcaag agggttcctt gactcacagg ttaccaatac caatggaaaa atctatcgat 300
aacttccaat ctactctata tgtgtcagat atttggtctg aaacactaca gagaactgaa 360
tgtttgctac aagtaactga aaacgtccag atgaatgagt ggattgagga aattagaatg 420
tactttagaa atatgacttt aggtgaaatt tccatgtccc cttacgacac tgcttgggtg 480
gctagagttc cagcgttgga cggttctcat gggcctcaat tccacagatc tttgcaatgg 540
attatcgaca accaattacc agatggggac tggggcgaac cttctctttt cttgggttac 600
gatagagttt gtaatacttt agcctgtgtg attgcgttga aaacatgggg tgttggggca 660
caaaacgttg aaagaggaat tcagttccta caatctaaca tatacaagat ggaggaagat 720
gacgctaatc atatgccaat aggattcgaa atcgtattcc ctgctatgat ggaagatgcc 780
aaagcattag gtttggattt gccatacgat gctactattt tgcaacagat ttcagccgaa 840
agagagaaaa agatgaaaaa gatcccaatg gcaatggtgt acaaataccc aaccacttta 900
cttcactcct tagaaggctt gcatagagaa gttgattgga ataagttgtt acaattacaa 960
tctgaaaatg gtagttttct ttattcacct gcttcaaccg catgcgcctt aatgtacact 1020
aaggacgtta aatgttttga ttacttaaac cagttgttga tcaagttcga ccacgcatgc 1080
ccaaatgtat atccagtcga tctattcgaa agattatgga tggttgacag attgcagaga 1140
ttagggatct ccagatactt tgaaagagag attagagatt gtttacaata cgtctacaga 1200
tattggaaag attgtggaat cggatgggct tctaactctt ccgtacaaga tgttgatgat 1260
acagccatgg cgtttagact tttaaggact catggtttcg acgtaaagga agattgcttt 1320
agacagtttt tcaaggacgg agaattcttc tgcttcgcag gccaatcatc tcaagcagtt 1380
acaggcatgt ttaatctttc aagagccagt caaacattgt ttccaggaga atctttattg 1440
aaaaaggcta gaaccttctc tagaaacttc ttgagaacaa agcatgagaa caacgaatgt 1500
ttcgataaat ggatcattac taaagatttg gctggtgaag tcgagtataa cttgaccttc 1560
ccatggtatg cctctttgcc tagattagaa cataggacat acttagatca atatggaatc 1620
gatgatatct ggataggcaa atctttatac aaaatgcctg ctgttaccaa cgaagttttc 1680
ctaaagttgg caaaggcaga ctttaacatg tgtcaagctc tacacaaaaa ggaattggaa 1740
caagtgataa agtggaacgc gtcctgtcaa ttcagagatc ttgaattcgc cagacaaaaa 1800
tcagtagaat gctattttgc tggtgcagcc acaatgttcg aaccagaaat ggttcaagct 1860
agattagtct gggcaagatg ttgtgtattg acaactgtct tagacgatta ctttgaccac 1920
gggacacctg ttgaggaact tagagtgttt gttcaagctg tcagaacatg gaatccagag 1980
ttgatcaacg gtttgccaga gcaagctaaa atcttgttta tgggcttata caaaacagtt 2040
aacacaattg cagaggaagc attcatggca cagaaaagag acgtccatca tcatttgaaa 2100
cactattggg acaagttgat aacaagtgcc ctaaaggagg ccgaatgggc agagtcaggt 2160
tacgtcccaa catttgatga atacatggaa gtagctgaaa tttctgttgc tctagaacca 2220
attgtctgta gtaccttgtt ctttgcgggt catagactag atgaggatgt tctagatagt 2280
tacgattacc atctagttat gcatttggta aacagagtcg gtagaatctt gaatgatata 2340
caaggcatga agagggaggc ttcacaaggt aagatctcat cagttcaaat ctacatggag 2400
gaacatccat ctgttccatc tgaggccatg gcgatcgctc atcttcaaga gttagttgat 2460
aattcaatgc agcaattgac atacgaagtt cttaggttca ctgcggttcc aaaaagttgt 2520
aagagaatcc acttgaatat ggctaaaatc atgcatgcct tctacaagga tactgatgga 2580
ttctcatccc ttactgcaat gacaggattc gtcaaaaagg ttcttttcga acctgtgcct 2640
gagtaa 2646

SEQ ID NO:56

```
MASSTLIQNR SCGVTSSMSS FQIFRGQPLR FPGTRTPAAV QCLKKRRCLR PTESVLESSP        60
GSGSYRIVTG PSGINPSSNG HLQEGSLTHR LPIPMEKSID NFQSTLYVSD IWSETLQRTE       120
CLLQVTENVQ MNEWIEEIRM YFRNMTLGEI SMSPYDTAWV ARVPALDGSH GPQFHRSLQW       180
IIDNQLPDGD WGEPSLFLGY DRVCNTLACV IALKTWGVGA QNVERGIQFL QSNIYKMEED       240
DANHMPIGFE IVFPAMMEDA KALGLDLPYD ATILQQISAE REKKMKKIPM AMVYKYPTTL       300
LHSLEGLHRE VDWNKLLQLQ SENGSFLYSP ASTACALMYT KDVKCFDYLN QLLIKFDHAC       360
PNVYPVDLFE RLWMVDRLQR LGISRYFERE IRDCLQYVYR YWKDCGIGWA SNSSVQDVDD       420
TAMAFRLLRT HGFDVKEDCF RQFFKDGEFF CFAGQSSQAV TGMFNLSRAS QTLFPGESLL       480
KKARTFSRNF LRTKHENNEC FDKWIITKDL AGEVEYNLTF PWYASLPRLE HRTYLDQYGI       540
DDIWIGKSLY KMPAVTNEVF LKLAKADFNM CQALHKKELE QVIKWNASCQ FRDLEFARQK       600
SVECYFAGAA TMFEPEMVQA RLVWARCCVL TTVLDDYFDH GTPVEELRVF VQAVRTWNPE       660
LINGLPEQAK ILFMGLYKTV NTIAEEAFMA QKRDVHHHLK HYWDKLITSA LKEAEWAESG       720
YVPTFDEYME VAEISVALEP IVCSTLFFAG HRLDEDVLDS YDYHLVMHLV NRVGRILNDI       780
QGMKREASQG KISSVQIYME EHPSVPSEAM AIAHLQELVD NSMQQLTYEV LRFTAVPKSC       840
KRIHLNMAKI MHAFYKDTDG FSSLTAMTGF VKKVLFEPVP E                           881
```

SEQ ID NO:57

```
atgcctggta aaattgaaaa tggtacccca aaggacctca agactggaaa tgattttgtt        60
tctgctgcta agagtttact agatcgagct ttcaaaagtc atcattccta ctacggatta       120
tgctcaactt catgtcaagt ttatgataca gcttgggttg caatgattcc aaaaacaaga       180
gataatgtaa aacagtggtt gtttccagaa tgtttccatt acctcttaaa aacacaagcc       240
gcagatggct catggggttc attgcctaca acacagacag cgggtatcct agatacagcc       300
tcagctgtgc tggcattatt gtgccacgca caagagcctt tacaaatatt ggatgtatct       360
ccagatgaaa tggggttgag aatagaacac ggtgtcacat ccttgaaacg tcaattagca       420
gtttggaatg atgtggagga caccaaccat attggcgtcg agtttatcat accagcctta       480
ctttccatgc tagaaaagga attagatgtt ccatcttttg aatttccatg taggtccatc       540
ttagagagaa tgcacgggga gaaattaggt catttcgacc tggaacaagt ttacggcaag       600
ccaagctcat tgttgcactc attggaagca tttctcggta agctagattt tgatcgacta       660
tcacatcacc tataccacgg cagtatgatg gcatctccat cttcaacggc tgcttatctt       720
attggggcta caaaatggga tgacgaagcc gaagattacc taagacatgt aatgcgtaat       780
ggtgcaggac atgggaatgg aggtatttct ggtacatttc caactactca tttcgaatgt       840
agctggatta tagcaacgtt gttaaaggtt ggctttactt tgaagcaaat tgacggcgat       900
ggcttaagag gtttatcaac catcttactt gaggcgcttc gtgatgagaa tggtgtcata       960
ggctttgccc ctagaacagc agatgtagat gacacagcca aagctctatt ggccttgtca      1020
ttggtaaacc agccagtgtc acctgatatc atgattaagg tctttgaggg caaagaccat      1080
tttaccactt ttggttcaga aagagatcca tcattgactt ccaacctgca cgtcctttta      1140
tctttactta aacaatctaa cttgtctcaa taccatcctc aaatcctcaa aacaacatta      1200
ttcacttgta gatggtggtg gggttccgat cattgtgtca aagacaaatg gaatttgagt      1260
cacctatatc caactatgtt gttggttgaa gccttcactg aagtgctcca tctcattgac      1320
ggtggtgaat tgtctagtct gtttgatgaa tcctttaagt gtaagattgg tcttagcatc      1380
tttcaagcgg tacttagaat aatcctcacc caagacaacg acggctcttg gagaggatac      1440
agagaacaga cgtgttacgc aatattggct ttagttcaag cgagacatgt atgctttttc      1500
actcacatgg ttgacagact gcaatcatgt gttgatcgag gtttctcatg gttgaaatct      1560
tgctcttttc attctcaaga cctgacttgg acctctaaaa cagcttatga agtgggtttc      1620
gtagctgaag catataaact agctgcttta caatctgctt ccctggaggt tcctgctgcc      1680
accattggac attctgtcac gtctgccgtt ccatcaagtg atcttgaaaa atacatgaga      1740
ttggtgagaa aaactgcgtt attctctcca ctggatgagt ggggtctaat ggcttctatc      1800
atcgaatctt catttttcgt accattactg caggcacaaa gagttgaaat ataccctaga      1860
gataatatca aggtggacga agataagtac ttgtctatta tcccattcac atgggtcgga      1920
tgcaataata ggtctagaac tttcgcaagt aacagatggc tatacgatat gatgtacctt      1980
tcattactcg gctatcaaac cgacgagtac atggaagctg tagctgggcc agtgtttggg      2040
gatgtttcct tgttacatca aacaattgat aaggtgattg ataatacaat gggtaacctt      2100
gcgagagcca atggaacagt acacagtggt aatggacatc agcacgaatc tcctaatata      2160
ggtcaagtcg aggacacctt gactcgtttc acaaattcag tcttgaatca caaagacgtc      2220
```

```
cttaactcta gctcatctga tcaagatact ttgagaagag agtttagaac attcatgcac      2280
gctcatataa cacaaatcga agataactca cgattcagta agcaagcctc atccgatgcg      2340
ttttcctctc ctgaacaatc ttactttcaa tgggtgaact caactggtgg ctcacatgtc      2400
gcttgcgcct attcatttgc cttctctaat tgcctcatgt ctgcaaattt gttgcagggt      2460
aaagacgcat ttccaagcgg aacgcaaaag tacttaatct cctctgttat gagacatgcc      2520
acaaacatgt gtagaatgta taacgacttt ggctctattg ccagagacaa cgctgagaga      2580
aatgttaata gtattcattt tcctgagttt actctctgta acggaacttc tcaaaaccta      2640
gatgaaagga aggaaagact tctgaaaatc gcaacttacg aacaagggta tttggataga      2700
gcactagagg ccttggaaag acagagtaga gatgatgccg gagacagagc tggatctaaa      2760
gatatgagaa agttgaaaat cgttaagtta ttctgtgatg ttacggactt atacgatcag      2820
ctctacgtta tcaaagattt gtcatcctct atgaagtaa                              2859
```

SEQ ID NO:58

```
MPGKIENGTP KDLKTGNDFV SAAKSLLDRA FKSHHSYYGL CSTSCQVYDT AWVAMIPKTR        60
DNVKQWLFPE CFHYLLKTQA ADGSWGSLPT TQTAGILDTA SAVLALLCHA QEPLQILDVS       120
PDEMGLRIEH GVTSLKRQLA VWNDVEDTNH IGVEFIIPAL LSMLEKELDV PSFEFPCRSI       180
LERMHGEKLG HFDLEQVYGK PSSLLHSLEA FLGKLDFDRL SHHLYHGSMM ASPSSTAAYL       240
IGATKWDDEA EDYLRHVMRN GAGHGNGGIS GTFPTTHFEC SWIIATLLKV GFTLKQIDGD       300
GLRGLSTILL EALRDENGVI GFAPRTADVD DTAKALLALS LVNQPVSPDI MIKVFEGKDH       360
FTTFGSERDP SLTSNLHVLL SLLKQSNLSQ YHPQILKTTL FTCRWWWGSD HCVKDKWNLS       420
HLYPTMLLVE AFTEVLHLID GGELSSLFDE SFKCKIGLSI FQAVLRIILT QDNDGSWRGY       480
REQTCYAILA LVQARHVCFF THMVDRLQSC VDRGFSWLKS CSFHSQDLTW TSKTAYEVGF       540
VAEAYKLAAL QSASLEVPAA TIGHSVTSAV PSSDLEKYMR LVRKTALFSP LDEWGLMASI       600
IESSFFVPLL QAQRVEIYPR DNIKVDEDKY LSIIPFTWVG CNNRSRTFAS NRWLYDMMYL       660
SLLGYQTDEY MEAVAGPVFG DVSLLHQTID KVIDNTMGNL ARANGTVHSG NGHQHESPNI       720
GQVEDTLTRF TNSVLNHKDV LNSSSSDQDT LRREFRTFMH AHITQIEDNS RFSKQASSDA       780
FSSPEQSYFQ WVNSTGGSHV ACAYSFAFSN CLMSANLLQG KDAFPSGTQK YLISSVMRHA       840
TNMCRMYNDF GSIARDNAER NVNSIHFPEF TLCNGTSQNL DERKERLLKI ATYEQGYLDR       900
ALEALERQSR DDAGDRAGSK DMRKLKIVKL FCDVTDLYDQ LYVIKDLSSS MK               952
```

SEQ ID NO:59

S. rebaudiana

```
atggatgctg tgacgggttt gttaactgtc ccagcaaccg ctataactat tggtggaact        60
gctgtagcat tggcggtagc gctaatcttt tggtacctga aatcctacac atcagctaga       120
agatcccaat caaatcatct tccaagagtg cctgaagtcc caggtgttcc attgttagga       180
aatctgttac aattgaagga gaaaaagcca tacatgactt ttacgagatg ggcagcgaca       240
tatggaccta tctatagtat caaaactggg gctacaagta tggttgtggt atcatctaat       300
gagatagcca aggaggcatt ggtgaccaga ttccaatcca tatctacaag gaacttatct       360
aaagccctga aagtacttac agcagataag acaatggtcg caatgtcaga ttatgatgat       420
tatcataaaa cagttaagag acacatactg accgccgtct tgggtcctaa tgcacagaaa       480
aagcatagaa ttcacagaga tatcatgatg gataacatat ctactcaact tcatgaattc       540
gtgaaaaaca acccagaaca ggaagaggta gaccttagaa aaatctttca atctgagtta       600
ttcggcttag ctatgagaca agccttagga aaggatgttg aaagtttgta cgttgaagac       660
ctgaaaatca ctatgaatag agacgaaatc tttcaagtcc ttgttgttga tccaatgatg       720
ggagcaatcg atgttgattg gagagacttc tttccatacc taaagtgggt cccaaacaaa       780
aagttcgaaa atactattca acaaatgtac atcagaagag aagctgttat gaaatcttta       840
atcaaagagc acaaaaagag aatagcgtca ggcgaaaagc taaatagtta tatcgattac       900
cttttatctg aagctcaaac tttaaccgat cagcaactat tgatgtcctt gtgggaacca       960
atcattgaat cttcagatac aacaatggtc acaacagaat gggcaatgta cgaattagct      1020
aaaaacccta aattgcaaga taggttgtac agagacatta agtccgtctg tggatctgaa      1080
aagataaccg aagagcatct atcacagctg ccttacatta cagctatttt ccacgaaaca      1140
ctgagaagac actcaccagt tcctatcatt cctctaagac atgtacatga agataccgtt      1200
ctaggcggct accatgttcc tgctggcaca gaacttgccg ttaacatcta cggttgcaac      1260
atggacaaaa acgtttggga aaatccagag gaatggaacc cagaaagatt catgaaagag      1320
```

```
aatgagacaa ttgattttca aaagacgatg gccttcggtg gtggtaagag agtttgtgct      1380
ggttccttgc aagccctttt aactgcatct attgggattg ggagaatggt tcaagagttc      1440
gaatggaaac tgaaggatat gactcaagag gaagtgaaca cgataggcct aactacacaa      1500
atgttaagac cattgagagc tattatcaaa cctaggatct aa                         1542
```

SEQ ID NO:60

S. rebaudiana

```
MDAVTGLLTV PATAITIGGT AVALAVALIF WYLKSYTSAR RSQSNHLPRV PEVPGVPLLG        60
NLLQLKEKKP YMTFTRWAAT YGPIYSIKTG ATSMVVVSSN EIAKEALVTR FQSISTRNLS       120
KALKVLTADK TMVAMSDYDD YHKTVKRHIL TAVLGPNAQK KHRIHRDIMM DNISTQLHEF       180
VKNNPEQEEV DLRKIFQSEL FGLAMRQALG KDVESLYVED LKITMNRDEI FQVLVVDPMM       240
GAIDVDWRDF FPYLKWVPNK KFENTIQQMY IRREAVMKSL IKEHKKRIAS GEKLNSYIDY       300
LLSEAQTLTD QQLLMSLWEP IIESSDTTMV TTEWAMYELA KNPKLQDRLY RDIKSVCGSE       360
KITEEHLSQL PYITAIFHET LRRHSPVPII PLRHVHEDTV LGGYHVPAGT ELAVNIYGCN       420
MDKNVWENPE EWNPERFMKE NETIDFQKTM AFGGGKRVCA GSLQALLTAS IGIGRMVQEF       480
EWKLKDMTQE EVNTIGLTTQ MLRPLRAIIK PRI                                    513
```

SEQ ID NO:61

```
aagcttacta gtaaaatgga cggtgtcatc gatatgcaaa ccattccatt gagaaccgct         60
attgctattg gtggtactgc tgttgctttg gttgttgcat tatacttttg gttcttgaga        120
tcctacgctt ccccatctca tcattctaat catttgccac cagtacctga agttccaggt        180
gttccagttt tgggtaattt gttgcaattg aaagaaaaaa agccttacat gaccttcacc        240
aagtgggctg aaatgtatgg tccaatctac tctattagaa ctggtgctac ttccatggtt        300
gttgtctctt ctaacgaaat cgccaaagaa gttgttgtta ccagattccc atctatctct        360
accagaaaat tgtcttacgc cttgaaggtt ttgaccgaag ataagtctat ggttgccatg        420
tctgattatc acgattacca taagaccgtc aagagacata ttttgactgc tgttttgggt        480
ccaaacgccc aaaaaaagtt tagagcacat agagacacca tgatggaaaa cgtttccaat        540
gaattgcatg ccttcttcga aaagaaccca aatcaagaag tcaacttgag aaagatcttc        600
caatcccaat tattcggttt ggctatgaag caagccttgg gtaaagatgt tgaatccatc        660
tacgttaagg atttggaaac caccatgaag agagaagaaa tcttcgaagt tttggttgtc        720
gatccaatga tgggtgctat tgaagttgat tggagagact ttttcccata cttgaaatgg        780
gttccaaaca agtccttcga aaacatcatc catagaatgt acactagaag agaagctgtt        840
atgaaggcct tgatccaaga acacaagaaa agaattgcct ccggtgaaaa cttgaactcc        900
tacattgatt acttgttgtc tgaagcccaa accttgaccg ataagcaatt attgatgtct        960
ttgtgggaac ctattatcga atcttctgat accactatgg ttactactga atgggctatg       1020
tacgaattgg ctaagaatcc aaacatgcaa gacagattat acgaagaaat ccaatccgtt       1080
tgcggttccg aaaagattac tgaagaaaac ttgtcccaat tgccatactt gtacgctgtt       1140
ttccaagaaa ctttgagaaa gcactgtcca gttcctatta tgccattgag atatgttcac       1200
gaaaacaccg ttttgggtgg ttatcatgtt ccagctggta ctgaagttgc tattaacatc       1260
tacggttgca acatggataa gaaggtctgg gaaaatccag aagaatggaa tccagaaaga       1320
ttcttgtccg aaaaagaatc catggacttg tacaaaacta tggcttttgg tggtggtaaa       1380
agagtttgcg ctggttcttt acaagccatg gttatttctt gcattggtat cggtagattg       1440
gtccaagatt ttgaatggaa gttgaaggat gatgccgaag aagatgttaa cactttgggt       1500
ttgactaccc aaaagttgca tccattattg gccttgatta acccaagaaa gtaactcgag       1560
ccgcgg                                                                  1566
```

SEQ ID NO:62

```
MDGVIDMQTI PLRTAIAIGG TAVALVVALY FWFLRSYASP SHHSNHLPPV PEVPGVPVLG        60
NLLQLKEKKP YMTFTKWAEM YGPIYSIRTG ATSMVVVSSN EIAKEVVVTR FPSISTRKLS       120
YALKVLTEDK SMVAMSDYHD YHKTVKRHIL TAVLGPNAQK KFRAHRDTMM ENVSNELHAF       180
FEKNPNQEVN LRKIFQSQLF GLAMKQALGK DVESIYVKDL ETTMKREEIF EVLVVDPMMG       240
AIEVDWRDFF PYLKWVPNKS FENIIHRMYT RREAVMKALI QEHKKRIASG ENLNSYIDYL       300
LSEAQTLTDK QLLMSLWEPI IESSDTTMVT TEWAMYELAK NPNMQDRLYE EIQSVCGSEK       360
```

```
ITEENLSQLP YLYAVFQETL RKHCPVPIMP LRYVHENTVL GGYHVPAGTE VAINIYGCNM       420
DKKVWENPEE WNPERFLSEK ESMDLYKTMA FGGGKRVCAG SLQAMVISCI GIGRLVQDFE       480
WKLKDDAEED VNTLGLTTQK LHPLLALINP RK                                     512
```

SEQ ID NO:63

R. suavissimus

```
atggccaccc tccttgagca tttccaagct atgccctttg ccatccctat tgcactggct        60
gctctgtctt ggctgttcct cttttacatc aaagtttcat tcttttccaa caagagtgct       120
caggctaagc tccctcctgt gccagtggtt cctgggctgc cggtgattgg gaatttactg       180
caactcaagg agaagaaacc ctaccagact tttacaaggt gggctgagga gtatggacca       240
atctattcta tcaggactgg tgcttccacc atggtcgttc tcaataccac ccaagttgca       300
aaagaggcca tggtgaccag atatttatcc atctcaacca gaaagctatc aaacgcacta       360
aagattctta ctgctgataa atgtatggtt gcaataagtg actacaacga ttttcacaag       420
atgataaagc gatacatact ctcaaatgtt cttggaccta gtgctcagaa gcgtcaccgg       480
agcaacagag ataccttgag agctaatgtc tgcagccgat tgcattctca agtaaagaac       540
tctcctcgag aagctgtgaa tttcagaaga gtttttgagt gggaactctt tggaattgca       600
ttgaagcaag cctttggaaa ggacatagaa aagcccattt atgtggagga acttggcact       660
acactgtcaa gagatgagat ctttaaggtt ctagtgcttg acataatgga gggtgcaatt       720
gaggttgatt ggagagattt cttcccttac ctgagatgga ttccgaatac gcgcatggaa       780
acaaaaattc agcgactcta tttccgcagg aaagcagtga tgactgccct gatcaacgag       840
cagaagaagc gaattgcttc aggagaggaa atcaactgtt atatcgactt cttgcttaag       900
gaagggaaga cactgacaat ggaccaaata agtatgttgc tttgggagac ggttattgaa       960
acagcagata ctacaatggt aacgacagaa tgggctatgt atgaagttgc taaagactca      1020
aagcgtcagg atcgtctcta tcaggaaatc caaaaggttt gtggatcgga gatggttaca      1080
gaggaatact tgtcccaact gccgtacctg aatgcagttt tccatgaaac gctaaggaag      1140
cacagtccgg ctgcgttagt tcctttaaga tatgcacatg aagataccca actaggaggt      1200
tactacattc cagctggaac tgagattgct ataaacatat acgggtgtaa catggacaag      1260
catcaatggg aaagccctga ggaatggaaa ccggagagat ttttggaccc gaaatttgat      1320
cctatggatt tgtacaagac catggctttt ggggctggaa agagggtatg tgctggttct      1380
cttcaggcaa tgttaatagc gtgcccgacg attggtaggc tggtgcagga gtttgagtgg      1440
aagctgagag atggagaaga agaaaatgta gatactgttg ggctcaccac tcacaaacgc      1500
tatccaatgc atgcaatcct gaagccaaga agtta                                 1535
```

SEQ ID NO:64

R. suavissimus

```
atggctacct tgttggaaca ttttcaagct atgccattcg ctattccaat tgctttggct        60
gctttgtctt ggttgttttt gttctacatc aaggtttctt tcttctccaa caaatccgct       120
caagctaaat tgccaccagt tccagttgtt ccaggtttgc cagttattgg taatttgttg       180
caattgaaag aaaagaagcc ataccaaacc ttcactagat gggctgaaga atatggtcca       240
atctactcta ttagaactgg tgcttctact atggttgtct tgaacactac tcaagttgcc       300
aaagaagcta tggttaccag atacttgtct atctctacca gaaagttgtc caacgccttg       360
aaaatttga ccgctgataa gtgcatggtt gccatttctg attacaacga tttccacaag       420
atgatcaaga gatatatctt gtctaacgtt ttgggtccat ctgcccaaaa aagacataga       480
tctaacagag ataccttgag agccaacgtt tgttctagat tgcattccca agttaagaac       540
tctccaagag aagctgtcaa ctttagaaga gttttcgaat gggaattatt cggtatcgct       600
ttgaaacaag ccttcggtaa ggatattgaa aagccaatct acgtcgaaga attgggtact       660
actttgtcca gagatgaaat cttcaaggtt ttggtcttgg acattatgga aggtgccatt       720
gaagttgatt ggagagattt tttcccatac ttgcgttgga ttccaaacac cagaatggaa       780
actaagatcc aaagattata ctttagaaga aaggccgtta tgaccgcctt gattaacgaa       840
caaaagaaaa gaattgcctc cggtgaagaa atcaactgct acatcgattt cttgttgaaa       900
gaaggtaaga ccttgaccat ggaccaaatc tctatgttgt tgtgggaaac cgttattgaa       960
actgctgata ccacaatggt tactactgaa tgggctatgt acgaagttgc taaggattct      1020
aaaagacaag acagattata ccaagaaatc caaaaggtct gcggttctga aatggttaca      1080
gaagaatact tgtcccaatt gccatacttg aatgctgttt tccacgaaac tttgagaaaa      1140
```

```
cattctccag ctgctttggt tccattgaga tatgctcatg aagatactca attgggtggt      1200
tattacattc cagccggtac tgaaattgcc attaacatct acggttgcaa catggacaaa      1260
caccaatggg aatctccaga agaatggaag ccagaaagat ttttggatcc taagtttgac      1320
ccaatggact tgtacaaaac tatggctttt ggtgctggta aaagagtttg cgctggttct      1380
ttacaagcta tgttgattgc ttgtccaacc atcggtagat tggttcaaga atttgaatgg      1440
aagttgagag atggtgaaga agaaaacgtt gatactgttg gtttgaccac ccataagaga      1500
tatccaatgc atgctatttt gaagccaaga tcttaa                                1536
```

SEQ ID NO:65

```
aagcttacta gtaaaatggc ctccatcacc catttcttac aagattttca agctactcca        60
ttcgctactg cttttgctgt tggtggtgtt tctttgttga tattcttctt cttcatccgt       120
ggtttccact ctactaagaa aaacgaatat tacaagttgc caccagttcc agttgttcca       180
ggtttgccag ttgttggtaa tttgttgcaa ttgaaagaaa agaagccata caagactttc       240
ttgagatggg ctgaaattca tggtccaatc tactctatta gaactggtgc ttctaccatg       300
gttgttgtta actctactca tgttgccaaa gaagctatgg ttaccagatt ctcttcaatc       360
tctaccagaa agttgtccaa ggctttggaa ttattgacct ccaacaaatc tatggttgcc       420
acctctgatt acaacgaatt tcacaagatg gtcaagaagt acatcttggc cgaattattg       480
ggtgctaatg ctcaaaagag acacagaatt catagagaca ccttgatcga aaacgtcttg       540
aacaaattgc atgcccatac caagaattct ccattgcaag ctgttaactt cagaaagatc       600
ttcgaatctg aattattcgg tttggctatg aagcaagcct tgggttatga tgttgattcc       660
ttgttcgttg aagaattggg tactaccttg tccagagaag aaatctacaa cgttttggtc       720
agtgacatgt tgaagggtgc tattgaagtt gattggagag actttttccc atacttgaaa       780
tggatcccaa acaagtcctt cgaaatgaag attcaaagat tggcctctag aagacaagcc       840
gttatgaact ctattgtcaa agaacaaaag aagtccattg cctctggtaa gggtgaaaac       900
tgttacttga attacttgtt gtccgaagct aagactttga ccgaaaagca aatttccatt       960
ttggcctggg aaaccattat tgaaactgct gatacaactg ttgttaccac tgaatgggct      1020
atgtacgaat tggctaaaaa cccaaagcaa caagacagat tatacaacga aatccaaaac      1080
gtctgcggta ctgataagat taccgaagaa catttgtcca agttgcctta cttgtctgct      1140
gtttttcacg aaaccttgag aaagtattct ccatctccat tggttccatt gagatacgct      1200
catgaagata ctcaattggg tggttattat gttccagccg gtactgaaat tgctgttaat      1260
atctacggtt gcaacatgga caagaatcaa tgggaaactc cagaagaatg gaagccagaa      1320
agatttttgg acgaaaagta cgatccaatg gacatgtaca agactatgtc ttttggttcc      1380
ggtaaaagag tttgcgctgg ttctttacaa gctagtttga ttgcttgtac ctccatcggt      1440
agattggttc aagaatttga atggagattg aaagacggtg aagttgaaaa cgttgatacc      1500
ttgggtttga ctacccataa gttgtatcca atgcaagcta tcttgcaacc tagaaactga      1560
ctcgagccgc gg                                                          1572
```

SEQ ID NO:66

```
MASITHFLQD FQATPFATAF AVGGVSLLIF FFFIRGFHST KKNEYYKLPP VPVVPGLPVV        60
GNLLQLKEKK PYKTFLRWAE IHGPIYSIRT GASTMVVVNS THVAKEAMVT RFSSISTRKL       120
SKALELLTSN KSMVATSDYN EFHKMVKKYI LAELLGANAQ KRHRIHRDTL IENVLNKLHA       180
HTKNSPLQAV NFRKIFESEL FGLAMKQALG YDVDSLFVEE LGTTLSREEI YNVLVSDMLK       240
GAIEVDWRDF FPYLKWIPNK SFEMKIQRLA SRRQAVMNSI VKEQKKSIAS GKGENCYLNY       300
LLSEAKTLTE KQISILAWET IIETADTTVV TTEWAMYELA KNPKQQDRLY NEIQNVCGTD       360
KITEEHLSKL PYLSAVFHET LRKYSPSPLV PLRYAHEDTQ LGGYYVPAGT EIAVNIYGCN       420
MDKNQWETPE EWKPERFLDE KYDPMDMYKT MSFGSGKRVC AGSLQASLIA CTSIGRLVQE       480
FEWRLKDGEV ENVDTLGLTT HKLYPMQAIL QPRN                                   514
```

SEQ ID NO:67

```
atgatttcct tgttgttggg ttttgttgtc tcctccttct tgtttatctt cttcttgaaa        60
aaattgttgt tcttcttcag tcgtcacaaa atgtccgaag tttctagatt gccatctgtt       120
ccagttccag gttttccatt gattggtaac ttgttgcaat tgaaagaaaa gaagccacac       180
aagactttca ccaagtggtc tgaattatat ggtccaatct actctatcaa gatgggttcc       240
```

```
tcttctttga tcgtcttgaa ctctattgaa accgccaaag aagctatggt cagtagattc       300
tcttcaatct ctaccagaaa gttgtctaac gctttgactg ttttgacctg caacaaatct       360
atggttgcta cctctgatta cgatgacttt cataagttcg tcaagagatg cttgttgaac       420
ggtttgttgg gtgctaatgc tcaagaaaga aaaagacatt acagagatgc cttgatcgaa       480
aacgttacct ctaaattgca tgcccatacc agaaatcatc cacaagaacc agttaacttc       540
agagccattt tcgaacacga attattcggt gttgctttga aacaagcctt cggtaaagat       600
gtcgaatcca tctatgtaaa agaattgggt gtcaccttgt ccagagatga aattttcaag       660
gttttggtcc acgacatgat ggaaggtgct attgatgttg attggagaga tttcttccca       720
tacttgaaat ggatcccaaa caactctttc gaagccagaa ttcaacaaaa gcacaagaga       780
agattggctg ttatgaacgc cttgatccaa gacagattga atcaaaacga ttccgaatcc       840
gatgatgact gctacttgaa tttcttgatg tctgaagcta agaccttgac catggaacaa       900
attgctattt tggtttggga aaccattatc gaaactgctg ataccacttt ggttactact       960
gaatgggcta tgtacgaatt ggccaaacat caatctgttc aagatagatt attcaaagaa      1020
atccaatccg tctgcggtgg tgaaaagatc aaagaagaac aattgccaag attgccttac      1080
gtcaatggtg tttttcacga aaccttgaga aagtattctc cagctccatt ggttccaatt      1140
agatacgctc atgaagatac ccaaattggt ggttatcata ttccagccgg ttctgaaatt      1200
gccattaaca tctacggttg caacatggat aagaagagat gggaaagacc tgaagaatgg      1260
tggccagaaa gatttttgga agatagatac gaatcctccg acttgcataa gactatggct      1320
tttggtgctg gtaaaagagt ttgtgctggt gctttacaag ctagtttgat ggctggtatt      1380
gctatcggta gattggttca agaattcgaa tggaagttga gagatggtga agaagaaaac      1440
gttgatactt acggtttgac ctcccaaaag ttgtatccat tgatggccat tatcaaccca      1500
agaagatctt aa                                                          1512
```

SEQ ID NO:68

```
MASMISLLLG FVVSSFLFIF FLKKLLFFFS RHKMSEVSRL PSVPVPGFPL IGNLLQLKEK        60
KPHKTFTKWS ELYGPIYSIK MGSSSLIVLN SIETAKEAMV SRFSSISTRK LSNALTVLTC       120
NKSMVATSDY DDFHKFVKRC LLNGLLGANA QERKRHYRDA LIENVTSKLH AHTRNHPQEP       180
VNFRAIFEHE LFGVALKQAF GKDVESIYVK ELGVTLSRDE IFKVLVHDMM EGAIDVDWRD       240
FFPYLKWIPN NSFEARIQQK HKRRLAVMNA LIQDRLNQND SESDDDCYLN FLMSEAKTLT       300
MEQIAILVWE TIIETADTTL VTTEWAMYEL AKHQSVQDRL FKEIQSVCGG EKIKEEQLPR       360
LPYVNGVFHE TLRKYSPAPL VPIRYAHEDT QIGGYHIPAG SEIAINIYGC NMDKKRWERP       420
EEWWPERFLE DRYESSDLHK TMAFGAGKRV CAGALQASLM AGIAIGRLVQ EFEWKLRDGE       480
EENVDTYGLT SQKLYPLMAI INPRRS                                            506
```

SEQ ID NO:69

```
aagcttacta gtaaaatgga catgatgggt attgaagctg ttccatttgc tactgctgtt        60
gttttgggtg gtatttcctt ggttgttttg atcttcatca gaagattcgt ttccaacaga       120
aagagatccg ttgaaggttt gccaccagtt ccagatattc caggtttacc attgattggt       180
aacttgttgc aattgaaaga aaagaagcca cataagacct ttgctagatg ggctgaaact       240
tacggtccaa ttttctctat tagaactggt gcttctacca tgatcgtctt gaattcttct       300
gaagttgcca aagaagctat ggtcactaga ttctcttcaa tctctaccag aaagttgtcc       360
aacgccttga agattttgac cttcgataag tgtatggttg ccacctctga ttacaacgat       420
tttcacaaaa tggtcaaggg tttcatcttg agaaacgttt taggtgctcc agcccaaaaa       480
agacatagat gtcatagaga taccttgatc gaaaacatct ctaagtactt gcatgcccat       540
gttaagactt ctccattgga accagttgtc ttgaagaaga ttttcgaatc cgaaattttc       600
ggtttggctt tgaaacaagc cttgggtaag gatatcgaat ccatctatgt tgaagaattg       660
ggtactacct tgtccagaga agaaattttt gccgttttgg ttgttgatcc aatggctggt       720
gctattgaag ttgattggag agatttttcc ccatacttgt cctggattcc aaacaagtct       780
atggaaatga agatccaaag aatggatttt agaagaggtg ctttgatgaa ggccttgatt       840
ggtgaacaaa agaaaagaat cggttccggt gaagaaaaga actcctacat tgatttcttg       900
ttgtctgaag ctaccacttt gaccgaaaag caaattgcta tgttgatctg ggaaaccatc       960
atcgaaattt ccgatacaac tttggttacc tctgaatggg ctatgtacga attggctaaa      1020
gacccaaata gacaagaaat cttgtacaga gaaatccaca aggtttgcgg ttctaacaag      1080
ttgactgaag aaaacttgtc caagttgcca tacttgaact ctgttttcca cgaaaccttg      1140
```

agaaagtatt ctccagctcc aatggttcca gttagatatg ctcatgaaga tactcaattg 1200
ggtggttacc atattccagc tggttctcaa attgccatta acatctacgg ttgcaacatg 1260
aacaaaaagc aatgggaaaa tcctgaagaa tggaagccag aaagattctt ggacgaaaag 1320
tatgacttga tggacttgca taagactatg gcttttggtg gtggtaaaag agtttgtgct 1380
ggtgctttac aagcaatgtt gattgcttgc acttccatcg gtagattcgt tcaagaattt 1440
gaatggaagt tgatgggtgg tgaagaagaa aacgttgata ctgttgcttt gacctcccaa 1500
aaattgcatc caatgcaagc cattattaag gccagagaat gactcgagcc gcgg 1554

SEQ ID NO:70

MDMMGIEAVP FATAVVLGGI SLVVLIFIRR FVSNRKRSVE GLPPVPDIPG LPLIGNLLQL 60
KEKKPHKTFA RWAETYGPIF SIRTGASTMI VLNSSEVAKE AMVTRFSSIS TRKLSNALKI 120
LTFDKCMVAT SDYNDFHKMV KGFILRNVLG APAQKRHRCH RDTLIENISK YLHAHVKTSP 180
LEPVVLKKIF ESEIFGLALK QALGKDIESI YVEELGTTLS REEIFAVLVV DPMAGAIEVD 240
WRDFFPYLSW IPNKSMEMKI QRMDFRRGAL MKALIGEQKK RIGSGEEKNS YIDFLLSEAT 300
TLTEKQIAML IWETIIEISD TTLVTSEWAM YELAKDPNRQ EILYREIHKV CGSNKLTEEN 360
LSKLPYLNSV FHETLRKYSP APMVPVRYAH EDTQLGGYHI PAGSQIAINI YGCNMNKKQW 420
ENPEEWKPER FLDEKYDLMD LHKTMAFGGG KRVCAGALQA MLIACTSIGR FVQEFEWKLM 480
GGEEENVDTV ALTSQKLHPM QAIIKARE 508

SEQ ID NO:71

aagcttaaaa tgagtaagtc taatagtatg aattctacat cacacgaaac ccttttcaa 60
caattggtct tgggtttgga ccgtatgcca ttgatggatg ttcactggtt gatctacgtt 120
gctttcggcg catggttatg ttcttatgtg atacatgttt tatcatcttc ctctacagta 180
aaagtgccag ttgttggata caggtctgta ttcgaaccta catggttgct tagacttaga 240
ttcgtctggg aaggtggctc tatcataggt caagggtaca ataagtttaa agactctatt 300
ttccaagtta ggaaattggg aactgatatt gtcattatac cacctaacta tattgatgaa 360
gtgagaaaat tgtcacagga caagactaga tcagttgaac ctttcattaa tgattttgca 420
ggtcaataca caagaggcat ggttttcttg caatctgact tacaaaaccg tgttatacaa 480
caaagactaa ctccaaaatt ggtttccttg accaaggtca tgaaggaaga gttggattat 540
gctttaacaa aagagatgcc tgatatgaaa aatgacgaat gggtagaagt agatatcagt 600
agtataatgg tgagattgat ttccaggatc tccgccagag tctttctagg gcctgaacac 660
tgtcgtaacc aggaatggtt gactactaca gcagaatatt cagaatcact tttcattaca 720
gggtttatct taagagttgt acctcatatc ttaagaccat tcatcgcccc tctattacct 780
tcatacagga ctctacttag aaacgtttca agtggtagaa gagtcatcgg tgacatcata 840
agatctcagc aaggggatgg taacgaagat atactttcct ggatgagaga tgctgccaca 900
ggagaggaaa agcaaatcga taacattgct cagagaatgt taattctttc tttagcatca 960
atccacacta ctgcgatgac catgacacat gccatgtacg atctatgtgc ttgccctgag 1020
tacattgaac cattaagaga tgaagttaaa tctgttgttg gggcttctgg ctgggacaag 1080
acagcgttaa acagatttca taagttggac tccttcctaa aagagtcaca aagattcaac 1140
ccagtattct tattgacatt caatagaatc taccatcaat ctatgacctt atcagatggc 1200
actaacattc catctggaac acgtattgct gttccatcac acgcaatgtt gcaagattct 1260
gcacatgtcc caggtccaac cccacctact gaatttgatg gattcagata tagtaagata 1320
cgttctgata gtaactacgc acaaaagtac ctattctcca tgaccgattc ttcaaacatg 1380
gctttcggat acggcaagta tgcttgtcca ggtagatttt acgcgtctaa tgagatgaaa 1440
ctaacattag ccattttgtt gctacaattt gagttcaaac taccagatgg taaaggtcgt 1500
cctagaaata tcactatcga ttctgatatg attccagacc caagagctag actttgcgtc 1560
agaaaaagat cacttagaga tgaatgaccg cgg 1593

SEQ ID NO:72

MSKSNSMNST SHETLFQQLV LGLDRMPLMD VHWLIYVAFG AWLCSYVIHV LSSSSTVKVP 60
VVGYRSVFEP TWLLRLRFVW EGGSIIGQGY NKFKDSIFQV RKLGTDIVII PPNYIDEVRK 120
LSQDKTRSVE PFINDFAGQY TRGMVFLQSD LQNRVIQQRL TPKLVSLTKV MKEELDYALT 180
KEMPDMKNDE WVEVDISSIM VRLISRISAR VFLGPEHCRN QEWLTTTAEY SESLFITGFI 240

```
LRVVPHILRP FIAPLLPSYR TLLRNVSSGR RVIGDIIRSQ QGDGNEDILS WMRDAATGEE       300
KQIDNIAQRM LILSLASIHT TAMTMTHAMY DLCACPEYIE PLRDEVKSVV GASGWDKTAL       360
NRFHKLDSFL KESQRFNPVF LLTFNRIYHQ SMTLSDGTNI PSGTRIAVPS HAMLQDSAHV       420
PGPTPPTEFD GFRYSKIRSD SNYAQKYLFS MTDSSNMAFG YGKYACPGRF YASNEMKLTL       480
AILLLQFEFK LPDGKGRPRN ITIDSDMIPD PRARLCVRKR SLRDE                       525
```

SEQ ID NO:73

```
aagcttaaaa tggaagatcc tactgtctta tatgcttgtc ttgccattgc agttgcaact        60
ttcgttgtta gatggtacag agatccattg agatccatcc caacagttgg tggttccgat       120
ttgcctattc tatcttacat cggcgcacta agatggacaa gacgtggcag agagatactt       180
caagagggat atgatggcta cagaggatct acattcaaaa tcgcgatgtt agaccgttgg       240
atcgtgatcg caaatggtcc taaactagct gatgaagtca gacgtagacc agatgaagag       300
ttaaacttta tggacggatt aggagcattc gtccaaacta agtacacctt aggtgaagct       360
attcataacg atccatacca tgtcgatatc ataagagaaa aactaacaag aggccttcca       420
gccgtgcttc ctgatgtcat tgaagagttg acacttgcgg ttagacagta cattccaaca       480
gaaggtgatg aatgggtgtc cgtaaactgt tcaaaggccg caagagatat tgttgctaga       540
gcttctaata gagtctttgt aggtttgcct gcttgcagaa accaaggtta cttagatttg       600
gcaatagact ttacattgtc tgttgtcaag gatagagcca tcatcaatat gtttccagaa       660
ttgttgaagc caatagttgg cagagttgta ggtaacgcca ccagaaatgt tcgtagagct       720
gttccttttg ttgctccatt ggtggaggaa agacgtagac ttatggaaga gtacggtgaa       780
gactggtctg aaaaacctaa tgatatgtta cagtggataa tggatgaagc tgcatccaga       840
gatagttcag tgaaggcaat cgcagagaga ttgttaatgg tgaacttcgc ggctattcat       900
acctcatcaa acactatcac tcatgctttg taccaccttg ccgaaatgcc tgaaactttg       960
caaccactta gagaagagat cgaaccatta gtcaaagagg agggctggac caaggctgct      1020
atgggaaaaa tgtggtggtt agattcattt ctaagagaat ctcaaagata caatggcatt      1080
aacatcgtat ctttaactag aatggctgac aaagatatta cattgagtga tggcacattt      1140
ttgccaaaag gtactctagt ggccgttcca gcgtattcta ctcatagaga tgatgctgtc      1200
tacgctgatg ccttagtatt cgatcctttc agattctcac gtatgagagc gagagaaggt      1260
gaaggtacaa agcaccagtt cgttaatact tcagtcgagt acgttccatt tggtcacgga      1320
aagcatgctt gtccaggaag attcttcgcc gcaaacgaat tgaaagcaat gttggcttac      1380
attgttctaa actatgatgt aaagttgcct ggtgacggta aacgtccatt gaacatgtat      1440
tggggtccaa cagttttgcc tgcaccagca ggccaagtat tgttcagaaa gagacaagtt      1500
agtctataac cgcgg                                                       1515
```

SEQ ID NO:74

```
MEDPTVLYAC LAIAVATFVV RWYRDPLRSI PTVGGSDLPI LSYIGALRWT RRGREILQEG        60
YDGYRGSTFK IAMLDRWIVI ANGPKLADEV RRRPDEELNF MDGLGAFVQT KYTLGEAIHN       120
DPYHVDIIRE KLTRGLPAVL PDVIEELTLA VRQYIPTEGD EWVSVNCSKA ARDIVARASN       180
RVFVGLPACR NQGYLDLAID FTLSVVKDRA IINMFPELLK PIVGRVVGNA TRNVRRAVPF       240
VAPLVEERRR LMEEYGEDWS EKPNDMLQWI MDEAASRDSS VKAIAERLLM VNFAAIHTSS       300
NTITHALYHL AEMPETLQPL REEIEPLVKE EGWTKAAMGK MWWLDSFLRE SQRYNGINIV       360
SLTRMADKDI TLSDGTFLPK GTLVAVPAYS THRDDAVYAD ALVFDPFRFS RMRAREGEGT       420
KHQFVNTSVE YVPFGHGKHA CPGRFFAANE LKAMLAYIVL NYDVKLPGDG KRPLNMYWGP       480
TVLPAPAGQV LFRKRQVSL                                                    499
```

SEQ ID NO:75

```
atggcatttt tctctatgat ttcaattttg ttgggatttg ttatttcttc tttcatcttc        60
atctttttct tcaaaaagtt acttagtttt agtaggaaaa acatgtcaga agtttctact       120
ttgccaagtg ttccagtagt gcctggtttt ccagttattg ggaatttgtt gcaactaaag       180
gagaaaaagc ctcataaaac tttcactaga tggtcagaga tatatggacc tatctactct       240
ataaagatgg gttcttcatc tcttattgta ttgaacagta cagaaactgc taaggaagca       300
atggtcacta gattttcatc aatatctacc agaaaattgt caaacgccct aacagttcta       360
acctgcgata agtctatggt cgccacttct gattatgatg acttccacaa attagttaag       420
```

```
agatgtttgc taaatggact tcttggtgct aatgctcaaa agagaaaaag acactacaga       480
gatgctttga ttgaaaatgt gagttccaag ctacatgcac acgctagaga tcatccacaa       540
gagccagtta actttagagc aattttcgaa cacgaattgt ttggtgtagc attaaagcaa       600
gccttcggta aagacgtaga atccatatac gtcaaggagt taggcgtaac attatcaaaa       660
gatgaaatct ttaaggtgct tgtacatgat atgatggagg gtgcaattga tgtagattgg       720
agagatttct tcccatattt gaaatggatc cctaataagt cttttgaagc taggatacaa       780
caaaagcaca agagaagact agctgttatg aacgcactta tacaggacag attgaagcaa       840
aatgggtctg aatcagatga tgattgttac cttaacttct taatgtctga ggctaaaaca       900
ttgactaagg aacagatcgc aatccttgtc tgggaaacaa tcattgaaac agcagatact       960
accttagtca caactgaatg ggccatatac gagctagcca aacatccatc tgtgcaagat      1020
aggttgtgta aggagatcca gaacgtgtgt ggtggagaga aattcaagga agagcagttg      1080
tcacaagttc cttaccttaa cggcgttttc catgaaacct tgagaaaata ctcacctgca      1140
ccattagttc ctattagata cgcccacgaa gatacacaaa tcggtggcta ccatgttcca      1200
gctgggtccg aaattgctat aaacatctac gggtgcaaca tggacaaaaa gagatgggaa      1260
agaccagaag attggtggcc agaaagattc ttagatgatg gcaaatatga aacatctgat      1320
ttgcataaaa caatggcttt cggagctggc aaaagagtgt gtgccggtgc tctacaagcc      1380
tccctaatgg ctggtatcgc tattggtaga ttggtccaag agttcgaatg gaaacttaga      1440
gatggtgaag aggaaaatgt cgatacttat gggttaacat ctcaaaagtt atacccacta      1500
atggcaatca tcaatcctag aagatcctaa                                       1530
```

SEQ ID NO:76

```
MAFFSMISIL LGFVISSFIF IFFFKKLLSF SRKNMSEVST LPSVPVVPGF PVIGNLLQLK        60
EKKPHKTFTR WSEIYGPIYS IKMGSSSLIV LNSTETAKEA MVTRFSSIST RKLSNALTVL       120
TCDKSMVATS DYDDFHKLVK RCLLNGLLGA NAQKRKRHYR DALIENVSSK LHAHARDHPQ       180
EPVNFRAIFE HELFGVALKQ AFGKDVESIY VKELGVTLSK DEIFKVLVHD MMEGAIDVDW       240
RDFFPYLKWI PNKSFEARIQ QKHKRRLAVM NALIQDRLKQ NGSESDDDCY LNFLMSEAKT       300
LTKEQIAILV WETIIETADT TLVTTEWAIY ELAKHPSVQD RLCKEIQNVC GGEKFKEEQL       360
SQVPYLNGVF HETLRKYSPA PLVPIRYAHE DTQIGGYHVP AGSEIAINIY GCNMDKKRWE       420
RPEDWWPERF LDDGKYETSD LHKTMAFGAG KRVCAGALQA SLMAGIAIGR LVQEFEWKLR       480
DGEEENVDTY GLTSQKLYPL MAIINPRRS                                         509
```

SEQ ID NO:77

S. rebaudiana

```
atgcaatcag attcagtcaa agtctctcca tttgatttgg tttccgctgc tatgaatggc        60
aaggcaatgg aaaagttgaa cgctagtgaa tctgaagatc caacaacatt gcctgcacta       120
aagatgctag ttgaaaatag agaattgttg acactgttca caacttcctt cgcagttctt       180
attgggtgtc ttgtatttct aatgtggaga cgttcatcct ctaaaaagct ggtacaagat       240
ccagttccac aagttatcgt tgtaaagaag aaagagaagg agtcagaggt tgatgacggg       300
aaaaagaaag tttctatttt ctacggcaca caaacaggaa ctgccgaagg ttttgctaaa       360
gcattagtcg aggaagcaaa agtgagatat gaaaagacct ctttcaaggt tatcgatcta       420
gatgactacg ctgcagatga tgatgaatat gaggaaaaac tgaaaaagga atccttagcc       480
ttcttcttct tggccacata cggtgatggt gaacctactg ataatgctgc taacttctac       540
aagtggttca cagaaggcga cgataaaggt gaatggctga aaaagttaca atacggagta       600
tttggtttag gtaacagaca atatgaacat ttcaacaaga tcgctattgt agttgatgat       660
aaacttactg aaatgggagc caaaagatta gtaccagtag gattagggga tgatgatcag       720
tgtatagaag atgacttcac cgcctggaag gaattggtat ggccagaatt ggatcaactt       780
ttaagggacg aagatgatac ttctgtgact accccataca ctgcagccgt attggagtac       840
agagtggttt accatgataa accagcagac tcatatgctg aagatcaaac ccatacaaac       900
ggtcatgttg ttcatgatgc acagcatcct tcaagatcta atgtggcttt caaaaaggaa       960
ctacacacct ctcaatcaga taggtcttgt actcacttag aattcgatat ttctcacaca      1020
ggactgtctt acgaaactgg cgatcacgtt ggcgtttatt ccgagaactt gtccgaagtt      1080
gtcgatgaag cactaaaact gttagggtta tcaccagaca catacttctc agtccatgct      1140
gataaggagg atgggacacc tatcggtggt gcttcactac caccaccttt tcctccttgc      1200
acattgagag acgctctaac cagatacgca gatgtcttat cctcacctaa aaaggtagct      1260
```

```
ttgctggcat tggctgctca tgctagtgat cctagtgaag ccgataggtt aaagttcctg      1320
gcttcaccag ccggaaaaga tgaatatgca caatggatcg tcgccaacca acgttctttg      1380
ctagaagtga tgcaaagttt tccatctgcc aagcctccat taggtgtgtt cttcgcagca      1440
gtagctccac gtttacaacc aagatactac tctatcagtt catctcctaa gatgtctcct      1500
aacagaatac atgttacatg tgctttggtg tacgagacta ctccagcagg cagaattcac      1560
agaggattgt gttcaacctg gatgaaaaat gctgtccctt taacagagtc acctgattgc      1620
tctcaagcat ccattttcgt tagaacatca aatttcagac ttccagtgga tccaaaagtt      1680
ccagtcatta tgataggacc aggcactggt cttgccccat tcaggggctt tcttcaagag      1740
agattggcct tgaaggaatc tggtacagaa ttgggttctt ctatcttttt ctttggttgc      1800
cgtaatagaa aagttgactt tatctacgag gacgagctta acaatttttgt tgagacagga      1860
gcattgtcag aattgatcgt cgcattttca agagaaggga ctgccaaaga gtacgttcag      1920
cacaagatga gtcaaaaagc ctccgatata tggaaacttc taagtgaagg tgcctatctt      1980
tatgtctgtg gcgatgcaaa gggcatggcc aaggatgtcc atagaactct gcatacaatt      2040
gttcaggaac aagggagtct ggattcttcc aaggctgaat tgtacgtcaa aaacttacag      2100
atgtctggaa gatacttaag agatgtttgg taa                                  2133
```

SEQ ID NO:78

S. rebaudiana

```
MQSDSVKVSP FDLVSAAMNG KAMEKLNASE SEDPTTLPAL KMLVENRELL TLFTTSFAVL        60
IGCLVFLMWR RSSSKKLVQD PVPQVIVVKK KEKESEVDDG KKKVSIFYGT QTGTAEGFAK       120
ALVEEAKVRY EKTSFKVIDL DDYAADDDEY EEKLKKESLA FFFLATYGDG EPTDNAANFY       180
KWFTEGDDKG EWLKKLQYGV FGLGNRQYEH FNKIAIVVDD KLTEMGAKRL VPVGLGDDDQ       240
CIEDDFTAWK ELVWPELDQL LRDEDDTSVT TPYTAAVLEY RVVYHDKPAD SYAEDQTHTN       300
GHVVHDAQHP SRSNVAFKKE LHTSQSDRSC THLEFDISHT GLSYETGDHV GVYSENLSEV       360
VDEALKLLGL SPDTYFSVHA DKEDGTPIGG ASLPPPFPPC TLRDALTRYA DVLSSPKKVA       420
LLALAAHASD PSEADRLKFL ASPAGKDEYA QWIVANQRSL LEVMQSFPSA KPPLGVFFAA       480
VAPRLQPRYY SISSSPKMSP NRIHVTCALV YETTPAGRIH RGLCSTWMKN AVPLTESPDC       540
SQASIFVRTS NFRLPVDPKV PVIMIGPGTG LAPFRGFLQE RLALKESGTE LGSSIFFFGC       600
RNRKVDFIYE DELNNFVETG ALSELIVAFS REGTAKEYVQ HKMSQKASDI WKLLSEGAYL       660
YVCGDAKGMA KDVHRTLHTI VQEQGSLDSS KAELYVKNLQ MSGRYLRDVW                  710
```

SEQ ID NO:79

```
atgaaggtca gtccattcga attcatgtcc gctattatca agggtagaat ggacccatct        60
aactcctcat ttgaatctac tggtgaagtt gcctccgtta tctttgaaaa cagagaattg       120
gttgccatct tgaccacttc tattgctgtt atgattggtt gcttcgttgt cttgatgtgg       180
agaagagctg gttctagaaa ggttaagaat gtcgaattgc caaagccatt gattgtccat       240
gaaccagaac ctgaagttga agatggtaag aagaaggttt ccatcttctt cggtactcaa       300
actggtactg ctgaaggttt tgctaaggct ttggctgatg aagctaaagc tagatacgaa       360
aaggctacct tcagagttgt tgatttggat gattatgctg ccgatgatga ccaatacgaa       420
gaaaaattga agaacgaatc cttcgccgtt ttcttgttgg ctacttatgg tgatggtgaa       480
cctactgata atgctgctag attttacaag tggttcgccg aaggtaaaga aagaggtgaa       540
tggttgcaaa acttgcacta tgctgttttt ggtttgggta acagacaata cgaacacttc       600
aacaagattg ctaaggttgc cgacgaatta ttggaagctc aaggtggtaa tagattggtt       660
aaggttggtt taggtgatga cgatcaatgc atcgaagatg atttttctgc ttggagagaa       720
tctttgtggc cagaattgga tatgttgttg agagatgaag atgatgctac tactgttact       780
actccatata ctgctgctgt cttggaatac agagttgtct ttcatgattc tgctgatgtt       840
gctgctgaag ataagtcttg gattaacgct aatggtcatg ctgttcatga tgctcaacat       900
ccattcagat ctaacgttgt cgtcagaaaa gaattgcata cttctgcctc tgatagatcc       960
tgttctcatt tggaattcaa catttccggt tccgctttga attacgaaac tggtgatcat      1020
gttggtgtct actgtgaaaa cttgactgaa actgttgatg aagccttgaa cttgttgggt      1080
ttgtctccag aaacttactt ctctatctac accgataacg aagatggtac tccattgggt      1140
ggttcttcat tgccaccacc atttccatca tgtactttga gaactgcttt gaccagatac      1200
gctgatttgt tgaactctcc aaaaaagtct gctttgttgg ctttagctgc tcatgcttct      1260
aatccagttg aagctgatag attgagatac ttggcttctc cagctggtaa agatgaatat      1320
```

```
gcccaatctg ttatcggttc ccaaaagtct ttgttggaag ttatggctga attcccatct      1380
gctaaaccac cattaggtgt ttttttgct gctgttgctc caagattgca acctagattc      1440
tactccattt catcctctcc aagaatggct ccatctagaa tccatgttac ttgtgctttg      1500
gtttacgata agatgccaac tggtagaatt cataagggtg tttgttctac ctggatgaag      1560
aattctgttc caatggaaaa gtcccatgaa tgttcttggg ctccaatttt cgttagacaa      1620
tccaatttta agttgccagc cgaatccaag gttccaatta tcatggttgg tccaggtact      1680
ggtttggctc cttttagagg ttttttacaa gaaagattgg ccttgaaaga atccggtgtt      1740
gaattgggtc catccatttt gtttttcggt tgcagaaaca gaagaatgga ttacatctac      1800
gaagatgaat tgaacaactt cgttgaaacc ggtgctttgt ccgaattggt tattgctttt      1860
tctagagaag gtcctaccaa agaatacgtc caacataaga tggctgaaaa ggcttctgat      1920
atctggaact tgatttctga aggtgcttac ttgtacgttt gtggtgatgc taaaggtatg      1980
gctaaggatg ttcatagaac cttgcatacc atcatgcaag aacaaggttc tttggattct      2040
tccaaagctg aatccatggt caagaacttg caaatgaatg gtagatactt aagagatgtt      2100
tggtaa                                                                2106
```

SEQ ID NO:80

```
MKVSPFEFMS AIIKGRMDPS NSSFESTGEV ASVIFENREL VAILTTSIAV MIGCFVVLMW        60
RRAGSRKVKN VELPKPLIVH EPEPEVEDGK KKVSIFFGTQ TGTAEGFAKA LADEAKARYE       120
KATFRVVDLD DYAADDDQYE EKLKNESFAV FLLATYGDGE PTDNAARFYK WFAEGKERGE       180
WLQNLHYAVF GLGNRQYEHF NKIAKVADEL LEAQGGNRLV KVGLGDDDQC IEDDFSAWRE       240
SLWPELDMLL RDEDDATTVT TPYTAAVLEY RVVFHDSADV AAEDKSWINA NGHAVHDAQH       300
PFRSNVVVRK ELHTSASDRS CSHLEFNISG SALNYETGDH VGVYCENLTE TVDEALNLLG       360
LSPETYFSIY TDNEDGTPLG GSSLPPPFPS CTLRTALTRY ADLLNSPKKS ALLALAAHAS       420
NPVEADRLRY LASPAGKDEY AQSVIGSQKS LLEVMAEFPS AKPPLGVFFA AVAPRLQPRF       480
YSISSSPRMA PSRIHVTCAL VYDKMPTGRI HKGVCSTWMK NSVPMEKSHE CSWAPIFVRQ       540
SNFKLPAESK VPIIMVGPGT GLAPFRGFLQ ERLALKESGV ELGPSILFFG CRNRRMDYIY       600
EDELNNFVET GALSELVIAF SREGPTKEYV QHKMAEKASD IWNLISEGAY LYVCGDAKGM       660
AKDVHRTLHT IMQEQGSLDS SKAESMVKNL QMNGRYLRDV W                          701
```

SEQ ID NO:81

```
atggcagaat tagatacact tgatatagta gtattaggtg ttatcttttt gggtactgtg        60
gcatacttta ctaagggtaa attgtggggt gttaccaagg atccatacgc taacggattc       120
gctgcaggtg gtgcttccaa gcctggcaga actagaaaca tcgtcgaagc tatggaggaa       180
tcaggtaaaa actgtgttgt tttctacggc agtcaaacag gtacagcgga ggattacgca       240
tcaagacttg caaaggaagg aaagtccaga ttcggtttga acactatgat cgccgatcta       300
gaagattatg acttcgataa cttagacact gttccatctg ataacatcgt tatgtttgta       360
ttggctactt acggtgaagg cgaaccaaca gataacgccg tggatttcta tgagttcatt       420
actggcgaag atgcctcttt caatgagggc aacgatcctc cactaggtaa cttgaattac       480
gttgcgttcg gtctgggcaa caatacctac gaacactaca actcaatggt caggaacgtt       540
aacaaggctc tagaaaagtt aggagctcat agaattggag aagcaggtga gggtgacgac       600
ggagctggaa ctatggaaga ggacttttta gcttggaaag atccaatgtg ggaagccttg       660
gctaaaaaga tgggcttgga ggaaagagaa gctgtatatg aacctatttt cgctatcaat       720
gagagagatg atttgacccc tgaagcgaat gaggtatact tgggagaacc taataagcta       780
cacttggaag gtacagcgaa aggtccattc aactcccaca acccatatat cgcaccaatt       840
gcagaatcat acgaactttt ctcagctaag gatagaaatt gtctgcatat ggaaattgat       900
atttctggta gtaatctaaa gtatgaaaca ggcgaccata tcgcgatctg gcctaccaac       960
ccaggtgaag aggtcaacaa atttcttgac attctagatc tgtctggtaa gcaacattcc      1020
gtcgtaacag tgaaagcctt agaacctaca gccaaagttc cttttccaaa tccaactacc      1080
tacgatgcta tattgagata ccatctggaa atatgcgctc cagtttctag acagtttgtc      1140
tcaactttag cagcattcgc ccctaatgat gatatcaaag ctgagatgaa ccgtttggga      1200
tcagacaaag attacttcca cgaaaagaca ggaccacatt actacaatat cgctagattt      1260
ttggcctcag tctctaaagg tgaaaaatgg acaaagatac cattttctgc tttcatagaa      1320
ggccttacaa aactacaacc aagatactat tctatctctt cctctagttt agttcagcct      1380
aaaaagatta gtattactgc tgttgtcgaa tctcagcaaa ttccaggtag agatgaccca      1440
```

```
ttcagaggtg tagcgactaa ctacttgttc gctttgaagc agaaacaaaa cggtgatcca      1500
aatccagctc cttttggcca atcatacgag ttgacaggac caaggaataa gtatgatggt      1560
atacatgttc cagtccatgt aagacattct aactttaagc taccatctga tccaggcaaa      1620
cctattatca tgatcggtcc aggtaccggt gttgcccctt ttagaggctt cgtccaagag      1680
agggcaaaac aagccagaga tggtgtagaa gttggtaaaa cactgctgtt ctttggatgt      1740
agaaagagta cagaagattt catgtatcaa aaagagtggc aagagtacaa ggaagctctt      1800
ggcgacaaat tcgaaatgat tacagctttt tcaagagaag gatctaaaaa ggtttatgtt      1860
caacacagac tgaaggaaag atcaaaggaa gtttctgatc ttctatccca aaaagcatac      1920
ttctacgttt gcggagacgc cgcacatatg gcacgtgaag tgaacactgt gttagcacag      1980
atcatagcag aaggccgtgg tgtatcagaa gccaagggtg aggaaattgt caaaaacatg      2040
agatcagcaa atcaatacca agtgtgttct gatttcgtaa ctttacactg taaagagaca      2100
acatacgcga attcagaatt gcaagaggat gtctggagtt aa                        2142
```

SEQ ID NO:82

```
MAELDTLDIV VLGVIFLGTV AYFTKGKLWG VTKDPYANGF AAGGASKPGR TRNIVEAMEE        60
SGKNCVVFYG SQTGTAEDYA SRLAKEGKSR FGLNTMIADL EDYDFDNLDT VPSDNIVMFV       120
LATYGEGEPT DNAVDFYEFI TGEDASFNEG NDPPLGNLNY VAFGLGNNTY EHYNSMVRNV       180
NKALEKLGAH RIGEAGEGDD GAGTMEEDFL AWKDPMWEAL AKKMGLEERE AVYEPIFAIN       240
ERDDLTPEAN EVYLGEPNKL HLEGTAKGPF NSHNPYIAPI AESYELFSAK DRNCLHMEID       300
ISGSNLKYET GDHIAIWPTN PGEEVNKFLD ILDLSGKQHS VVTVKALEPT AKVPFPNPTT       360
YDAILRYHLE ICAPVSRQFV STLAAFAPND DIKAEMNRLG SDKDYFHEKT GPHYYNIARF       420
LASVSKGEKW TKIPFSAFIE GLTKLQPRYY SISSSSLVQP KKISITAVVE SQQIPGRDDP       480
FRGVATNYLF ALKQKQNGDP NPAPFGQSYE LTGPRNKYDG IHVPVHVRHS NFKLPSDPGK       540
PIIMIGPGTG VAPFRGFVQE RAKQARDGVE VGKTLLFFGC RKSTEDFMYQ KEWQEYKEAL       600
GDKFEMITAF SREGSKKVYV QHRLKERSKE VSDLLSQKAY FYVCGDAAHM AREVNTVLAQ       660
IIAEGRGVSE AKGEEIVKNM RSANQYQVCS DFVTLHCKET TYANSELQED VWS               713
```

SEQ ID NO:83

```
atgcaatcgg aatccgttga agcatcgacg attgatttga tgactgctgt tttgaaggac        60
acagtgatcg atacagcgaa cgcatctgat aacggagact caaagatgcc gccggcgttg       120
gcgatgatgt tcgaaattcg tgatctgttg ctgattttga ctacgtcagt tgctgttttg       180
gtcggatgtt tcgttgtttt ggtgtggaag agatcgtccg ggaagaagtc cggcaaggaa       240
ttggagccgc cgaagatcgt tgtgccgaag aggcggctgg agcaggaggt tgatgatggt       300
aagaagaagg ttacgatttt cttcggaaca caaactggaa cggctgaagg tttcgctaag       360
gcacttttcg aagaagcgaa agcgcgatat gaaaaggcag cgtttaaagt gattgatttg       420
gatgattatg ctgctgattt ggatgagtat gcagagaagc tgaagaagga aacatatgct       480
ttcttcttct tggctacata tggagatggt gagccaactg ataatgctgc caaatttat        540
aaatggttta ctgagggaga cgagaaaggc gtttggcttc aaaaacttca atatggagta       600
tttggtcttg gcaacagaca atatgaacat ttcaacaaga ttggaatagt ggttgatgat       660
ggtctcaccg agcagggtgc aaaacgcatt gttcccgttg gtcttggaga cgacgatcaa       720
tcaattgaag acgatttttc ggcatggaaa gagttagtgt ggcccgaatt ggatctattg       780
cttcgcgatg aagatgacaa agctgctgca actccttaca cagctgcaat ccctgaatac       840
cgcgtcgtat ttcatgacaa acccgatgcg ttttctgatg atcatactca aaccaatggt       900
catgctgttc atgatgctca acatccatgc agatccaatg tggctgttaa aaaagagctt       960
catactcctg aatccgatcg ttcatgcaca catcttgaat ttgacatttc tcacactgga      1020
ttatcttatg aaactgggga tcatgttggt gtatactgtg aaaacctaat tgaagtagtg      1080
gaagaagctg ggaaattgtt aggattatca acagatactt atttctcgtt acatattgat      1140
aacgaagatg gttcaccact tggtggacct tcattacaac ctccttttcc tccttgtact      1200
ttaagaaaag cattgactaa ttatgcagat ctgttaagct ctcccaaaaa gtcaactttg      1260
cttgctctag ctgctcatgc ttccgatccc actgaagctg atcgtttaag atttcttgca      1320
tctcgcgagg gcaaggatga atatgctgaa tgggttgttg caaaccaaag aagtcttctt      1380
gaagtcatgg aagctttccc gtcagctaga ccgccacttg gtgttttctt tgcagcggtt      1440
gcaccgcgtt tacagcctcg ttactactct atttcttcct ccccaaagat ggaaccaaac      1500
aggattcatg ttacttgcgc gttggtttat gaaaaaactc ccgcaggtcg tatccacaaa      1560
```

```
ggaatctgct caacctggat gaagaacgct gtacctttga ccgaaagtca agattgcagt      1620
tgggcaccga tttttgttag aacatcaaac ttcagacttc caattgaccc gaaagtcccg      1680
gttatcatga ttggtcctgg aaccgggttg gctccattta ggggttttct tcaagaaaga      1740
ttggctctta aagaatccgg aaccgaactc gggtcatcta ttttattctt cggttgtaga      1800
aaccgcaaag tggattacat atatgagaat gaactcaaca actttgttga aaatggtgcg      1860
ctttctgagc ttgatgttgc tttctcccgc gatggcccga cgaaagaata cgtgcaacat      1920
aaaatgaccc aaaaggcttc tgaaatatgg aatatgcttt ctgagggagc atatttatat      1980
gtatgtggtg atgctaaagg catggctaaa gatgtacacc gtacacttca caccattgtg      2040
caagaacagg gaagtttgga ctcgtctaaa gcggagttgt atgtgaagaa tctacaaatg      2100
tcaggaagat acctccgtga tgtttggtaa                                      2130
```

SEQ ID NO:84

```
MQSESVEAST IDLMTAVLKD TVIDTANASD NGDSKMPPAL AMMFEIRDLL LILTTSVAVL        60
VGCFVVLVWK RSSGKKSGKE LEPPKIVVPK RRLEQEVDDG KKKVTIFFGT QTGTAEGFAK       120
ALFEEAKARY EKAAFKVIDL DDYAADLDEY AEKLKKETYA FFFLATYGDG EPTDNAAKFY       180
KWFTEGDEKG VWLQKLQYGV FGLGNRQYEH FNKIGIVVDD GLTEQGAKRI VPVGLGDDDQ       240
SIEDDFSAWK ELVWPELDLL LRDEDDKAAA TPYTAAIPEY RVVFHDKPDA FSDDHTQTNG       300
HAVHDAQHPC RSNVAVKKEL HTPESDRSCT HLEFDISHTG LSYETGDHVG VYCENLIEVV       360
EEAGKLLGLS TDTYFSLHID NEDGSPLGGP SLQPPFPPCT LRKALTNYAD LLSSPKKSTL       420
LALAAHASDP TEADRLRFLA SREGKDEYAE WVVANQRSLL EVMEAFPSAR PPLGVFFAAV       480
APRLQPRYYS ISSSPKMEPN RIHVTCALVY EKTPAGRIHK GICSTWMKNA VPLTESQDCS       540
WAPIFVRTSN FRLPIDPKVP VIMIGPGTGL APFRGFLQER LALKESGTEL GSSILFFGCR       600
NRKVDYIYEN ELNNFVENGA LSELDVAFSR DGPTKEYVQH KMTQKASEIW NMLSEGAYLY       660
VCGDAKGMAK DVHRTLHTIV QEQGSLDSSK AELYVKNLQM SGRYLRDVW                   709
```

SEQ ID NO:85

S. rebaudiana

```
atgcaatcta actccgtgaa gatttcgccg cttgatctgg taactgcgct gtttagcggc        60
aaggttttgg acacatcgaa cgcatcggaa tcgggagaat ctgctatgct gccgactata       120
gcgatgatta tggagaatcg tgagctgttg atgatactca caacgtcggt tgctgtattg       180
atcggatgcg ttgtcgtttt ggtgtggcgg agatcgtcta cgaagaagtc ggcgttggag       240
ccaccggtga ttgtggttcc gaagagagtg caagaggagg aagttgatga tggtaagaag       300
aaagttacgg ttttcttcgg cacccaaact ggaacagctg aaggcttcgc taaggcactt       360
gttgaggaag ctaaagctcg atatgaaaag gctgtcttta aagtaattga tttggatgat       420
tatgctgctg atgacgatga gtatgaggag aaactaaaga aagaatcttt ggccttttc        480
tttttggcta cgtatggaga tggtgagcca acagataatg ctgccagatt ttataaatgg       540
tttactgagg gagatgcgaa aggagaatgg cttaataagc ttcaatatgg agtatttggt       600
ttgggtaaca gacaatatga acatttaac aagatcgcaa aagtggttga tgatggtctt        660
gtagaacagg gtgcaaagcg tcttgttcct gttggacttg gagatgatga tcaatgtatt       720
gaagatgact tcaccgcatg gaaagagtta gtatggccgg agttggatca attacttcgt       780
gatgaggatg acacaactgt tgctactcca tacacagctg ctgttgcaga atatcgcgtt       840
gtttttcatg aaaaaccaga cgcgctttct gaagattata gttatacaaa tggccatgct       900
gttcatgatg ctcaacatcc atgcagatcc aacgtggctg tcaaaaagga acttcatagt       960
cctgaatctg accggtcttg cactcatctt gaatttgaca tctcgaacac cggactatca      1020
tatgaaactg gggaccatgt tggagtttac tgtgaaaact tgagtgaagt tgtgaatgat      1080
gctgaaagat tagtaggatt accaccagac acttactcct ccatccacac tgatagtgaa      1140
gacgggtcgc cacttggcgg agcctcattg ccgcctcctt tcccgccatg cactttaagg      1200
aaagcattga cgtgttatgc tgatgttttg agttctccca agaagtcggc tttgcttgca      1260
ctagctgctc atgccaccga tcccagtgaa gctgatagat tgaaatttct tgcatccccc      1320
gccggaaagg atgaatattc tcaatggata gttgcaagcc aaagaagtct ccttgaagtc      1380
atggaagcat tcccgtcagc taagccttca cttggtgttt tctttgcatc tgttgccccg      1440
cgcttacaac caagatacta ctctatttct tcctcaccca agatggcacc ggataggatt      1500
catgttacat gtgcattagt ctatgagaaa acacctgcag gccgcatcca caaaggagtt      1560
tgttcaactt ggatgaagaa cgcagtgcct atgaccgaga gtcaagattg cagttgggcc      1620
```

```
ccaatatacg tccgaacatc caatttcaga ctaccatctg accctaaggt cccggttatc      1680
atgattggac ctggcactgg tttggctcct tttagaggtt tccttcaaga gcggttagct      1740
ttaaaggaag ccggaactga cctcggttta tccattttat tcttcggatg taggaatcgc      1800
aaagtggatt tcatatatga aaacgagctt aacaactttg tggagactgg tgctctttct      1860
gagcttattg ttgctttctc ccgtgaaggc ccgactaagg aatatgtgca acacaagatg      1920
agtgagaagg cttcggatat ctggaacttg ctttctgaag gagcatattt atacgtatgt      1980
ggtgatgcca aaggcatggc caaagatgta catcgaaccc tccacacaat tgtgcaagaa      2040
cagggatctc ttgactcgtc aaaggcagaa ctctacgtga agaatctaca aatgtcagga      2100
agatacctcc gtgacgtttg gtaa                                            2124
```

SEQ ID NO:86

S. rebaudiana

```
MQSNSVKISP LDLVTALFSG KVLDTSNASE SGESAMLPTI AMIMENRELL MILTTSVAVL        60
IGCVVVLVWR RSSTKKSALE PPVIVVPKRV QEEEVDDGKK KVTVFFGTQT GTAEGFAKAL       120
VEEAKARYEK AVFKVIDLDD YAADDDEYEE KLKKESLAFF FLATYGDGEP TDNAARFYKW       180
FTEGDAKGEW LNKLQYGVFG LGNRQYEHFN KIAKVVDDGL VEQGAKRLVP VGLGDDDQCI       240
EDDFTAWKEL VWPELDQLLR DEDDTTVATP YTAAVAEYRV VFHEKPDALS EDYSYTNGHA       300
VHDAQHPCRS NVAVKKELHS PESDRSCTHL EFDISNTGLS YETGDHVGVY CENLSEVVND       360
AERLVGLPPD TYSSIHTDSE DGSPLGGASL PPPFPPCTLR KALTCYADVL SSPKKSALLA       420
LAAHATDPSE ADRLKFLASP AGKDEYSQWI VASQRSLLEV MEAFPSAKPS LGVFFASVAP       480
RLQPRYYSIS SSPKMAPDRI HVTCALVYEK TPAGRIHKGV CSTWMKNAVP MTESQDCSWA       540
PIYVRTSNFR LPSDPKVPVI MIGPGTGLAP FRGFLQERLA LKEAGTDLGL SILFFGCRNR       600
KVDFIYENEL NNFVETGALS ELIVAFSREG PTKEYVQHKM SEKASDIWNL LSEGAYLYVC       660
GDAKGMAKDV HRTLHTIVQE QGSLDSSKAE LYVKNLQMSG RYLRDVW                     707
```

SEQ ID NO:87

```
atgtcctcca actccgattt ggtcagaaga ttggaatctg ttttgggtgt ttctttcggt        60
ggttctgtta ctgattccgt tgttgttatt gctaccacct ctattgcttt ggttatcggt       120
gttttggttt tgttgtggag aagatcctct gacagatcta gagaagttaa gcaattggct       180
gttccaaagc cagttactat cgttgaagaa gaagatgaat tcgaagttgc ttctggtaag       240
accagagttt ctattttcta cggtactcaa actggtactg ctgaaggttt tgctaaggct       300
ttggctgaag aaatcaaagc cagatacgaa aaagctgccg ttaaggttat tgatttggat       360
gattacacag ccgaagatga caaatacggt gaaaagttga agaaagaaac tatggccttc       420
ttcatgttgg ctacttatgg tgatggtgaa cctactgata atgctgctag attttacaag       480
tggttcaccg aaggtactga tagaggtgtt tggttggaac atttgagata cggtgtattc       540
ggtttgggta acagacaata cgaacacttc aacaagattg ccaaggttgt tgatgatttg       600
ttggttgaac aaggtgccaa gagattggtt actgttggtt tgggtgatga tgatcaatgc       660
atcgaagatg atttctccgc ttggaaagaa gccttgtggc cagaattgga tcaattattg       720
caagatgata ccaacaccgt ttctactcca tacactgctg ttattccaga atacagagtt       780
gttatccacg atccatctgt tacctcttat gaagatccat actctaacat ggctaacggt       840
aatgcctctt acgatattca tcatccatgt agagctaacg ttgccgtcca aaaagaattg       900
cataagccag aatctgacag aagttgcatc catttggaat tcgatatttt cgctactggt       960
ttgacttacg aaaccggtga tcatgttggt gtttacgctg ataattgtga tgatactgta      1020
gaagaagccg ctaagttgtt gggtcaacca ttggatttgt tgttctccat tcataccgat      1080
aacaacgacg gtacttcttt gggttcttct ttgccaccac catttccagg tccatgtact      1140
ttgagaactg ctttggctag atatgccgat ttgttgaatc caccaaaaaa ggctgctttg      1200
attgctttag ctgctcatgc tgatgaacca tctgaagctg aaagattgaa gttcttgtca      1260
tctccacaag gtaaggacga atattctaaa tgggttgtcg gttcccaaag atccttggtt      1320
gaagttatgg ctgaatttcc atctgctaaa ccaccattgg gtgtattttt tgctgctgtt      1380
gttcctagat tgcaacctag atattactcc atctcttcca gtccaagatt tgctccacat      1440
agagttcatg ttacttgcgc tttggtttat ggtccaactc caactggtag aattcacaga      1500
ggtgtatgtt cattctggat gaagaatgtt gtcccattgg aaaagtctca aaactgttct      1560
tgggccccaa ttttcatcag acaatctaat ttcaagttgc cagccgatca ttctgttcca      1620
atagttatgg ttggtccagg tactggttta gctccttta gaggtttctt acaagaaaga       1680
```

```
ttggccttga aagaagaagg tgctcaagtt ggtcctgctt tgttgttttt tggttgcaga      1740
aacagacaaa tggacttcat ctacgaagtc gaattgaaca actttgtcga acaaggtgct      1800
ttgtccgaat tgatcgttgc tttttcaaga gaaggtccat ccaaagaata cgtccaacat      1860
aagatggttg aaaaggcagc ttacatgtgg aacttgattt ctcaaggtgg ttacttctac      1920
gtttgtggtg atgctaaagg tatggctaga gatgttcata gaacattgca taccatcgtc      1980
caacaagaag aaaaggttga ttctaccaag gccgaatcca tcgttaagaa attgcaaatg      2040
gacggtagat acttgagaga tgtttggtga                                      2070
```

SEQ ID NO:88

```
MSSNSDLVRR LESVLGVSFG GSVTDSVVVI ATTSIALVIG VLVLLWRRSS DRSREVKQLA        60
VPKPVTIVEE EDEFEVASGK TRVSIFYGTQ TGTAEGFAKA LAEEIKARYE KAAVKVIDLD       120
DYTAEDDKYG EKLKKETMAF FMLATYGDGE PTDNAARFYK WFTEGTDRGV WLEHLRYGVF       180
GLGNRQYEHF NKIAKVVDDL LVEQGAKRLV TVGLGDDDQC IEDDFSAWKE ALWPELDQLL       240
QDDTNTVSTP YTAVIPEYRV VIHDPSVTSY EDPYSNMANG NASYDIHHPC RANVAVQKEL       300
HKPESDRSCI HLEFDIFATG LTYETGDHVG VYADNCDDTV EEAAKLLGQP LDLLFSIHTD       360
NNDGTSLGSS LPPPFPGPCT LRTALARYAD LLNPPKKAAL IALAAHADEP SEAERLKFLS       420
SPQGKDEYSK WVVGSQRSLV EVMAEFPSAK PPLGVFFAAV VPRLQPRYYS ISSSPRFAPH       480
RVHVTCALVY GPTPTGRIHR GVCSFWMKNV VPLEKSQNCS WAPIFIRQSN FKLPADHSVP       540
IVMVGPGTGL APFRGFLQER LALKEEGAQV GPALLFFGCR NRQMDFIYEV ELNNFVEQGA       600
LSELIVAFSR EGPSKEYVQH KMVEKAAYMW NLISQGGYFY VCGDAKGMAR DVHRTLHTIV       660
QQEEKVDSTK AESIVKKLQM DGRYLRDVW                                         689
```

SEQ ID NO:89

```
atgacttctg cactttatgc ctccgatctt ttcaaacaat tgaaaagtat catgggaacg        60
gattctttgt ccgatgatgt tgtattagtt attgctacaa cttctctggc actggttgct       120
ggtttcgttg tcttattgtg gaaaaagacc acggcagatc gttccggcga gctaaagcca       180
ctaatgatcc ctaagtctct gatggcgaaa gatgaggatg atgacttaga tctaggttct       240
ggaaaaacga gagtctctat cttcttcggc acacaaaccg gaacagccga aggattcgct       300
aaagcacttt cagaagagat caaagcaaga tacgaaaagg cggctgtaaa agtaatcgat       360
ttggatgatt acgctgccga tgatgaccaa tatgaggaaa agttgaaaaa ggaaacattg       420
gctttctttt gtgtagccac gtatggtgat ggtgaaccaa ccgataacgc cgcaagattc       480
tacaagtggt ttactgaaga gaacgaaaga gatatcaagt tgcagcaact tgcttacggc       540
gtttttgcct taggtaacag acaatacgag cactttaaca agataggtat tgtcttagat       600
gaagagttat gcaaaaaggg tgcgaagaga ttgattgaag tcggtttagg agatgatgat       660
caatctatcg aggatgactt taatgcatgg aaggaatctt tgtggtctga attagataag       720
ttacttaagg acgaagatga taaatccgtt gccactccat acacagccgt cattccagaa       780
tatagagtag ttactcatga tccaagattc acaacacaga aatcaatgga aagtaatgtg       840
gctaatggta atactaccat cgatattcat catccatgta gagtagacgt tgcagttcaa       900
aaggaattgc acactcatga atcagacaga tcttgcatac atcttgaatt tgatatatca       960
cgtactggta tcacttacga aacaggtgat cacgtgggtg tctacgctga aaaccatgtt      1020
gaaattgtag aggaagctgg aaagttgttg ggccatagtt tagatcttgt tttctcaatt      1080
catgccgata aagaggatgg ctcaccacta gaaagtgcag tgcctccacc atttccagga      1140
ccatgcaccc taggtaccgg tttagctcgt tacgcggatc tgttaaatcc tccacgtaaa      1200
tcagctctag tggccttggc tgcgtacgcc acagaacctt ctgaggcaga aaaactgaaa      1260
catctaactt caccagatgg taaggatgaa tactcacaat ggatagtagc tagtcaacgt      1320
tctttactag aagttatggc tgctttccca tccgctaaac ctcctttggg tgttttcttc      1380
gccgcaatag cgcctagact gcaaccaaga tactattcaa tttcatcctc acctagactg      1440
gcaccatcaa gagttcatgt cacatccgct ttagtgtacg gtccaactcc tactggtaga      1500
atccataagg gcgtttgttc aacatggatg aaaaacgcgg ttccagcaga gaagtctcac      1560
gaatgttctg gtgctccaat ctttatcaga gcctccaact tcaaactgcc ttccaatcct      1620
tctactccta ttgtcatggt cggtcctggt acaggtcttg ctccattcag aggtttctta      1680
caagagagaa tggccttaaa ggaggatggt gaagagttgg gatcttcttt gttgtttttc      1740
ggctgtagaa acagacaaat ggatttcatc tacgaagatg aactgaataa ctttgtagat      1800
caaggagtta tttcagagtt gataatggct ttttctagag aaggtgctca gaaggagtac      1860
```

```
gtccaacaca aaatgatgga aaaggccgca caagtttggg acttaatcaa agaggaaggc      1920
tatctatatg tctgtggtga tgcaaagggt atggcaagag atgttcacag aacacttcat      1980
actatagtcc aggaacagga aggcgttagt tcttctgaag cggaagcaat tgtgaaaaag      2040
ttacaaacag agggaagata cttgagagat gtgtggtaa                             2079
```

SEQ ID NO:90

```
MTSALYASDL FKQLKSIMGT DSLSDDVVLV IATTSLALVA GFVVLLWKKT TADRSGELKP        60
LMIPKSLMAK DEDDDLDLGS GKTRVSIFFG TQTGTAEGFA KALSEEIKAR YEKAAVKVID       120
LDDYAADDDQ YEEKLKKETL AFFCVATYGD GEPTDNAARF YKWFTEENER DIKLQQLAYG       180
VFALGNRQYE HFNKIGIVLD EELCKKGAKR LIEVGLGDDD QSIEDDFNAW KESLWSELDK       240
LLKDEDDKSV ATPYTAVIPE YRVVTHDPRF TTQKSMESNV ANGNTTIDIH HPCRVDVAVQ       300
KELHTHESDR SCIHLEFDIS RTGITYETGD HVGVYAENHV EIVEEAGKLL GHSLDLVFSI       360
HADKEDGSPL ESAVPPPFPG PCTLGTGLAR YADLLNPPRK SALVALAAYA TEPSEAEKLK       420
HLTSPDGKDE YSQWIVASQR SLLEVMAAFP SAKPPLGVFF AAIAPRLQPR YYSISSSPRL       480
APSRVHVTSA LVYGPTPTGR IHKGVCSTWM KNAVPAEKSH ECSGAPIFIR ASNFKLPSNP       540
STPIVMVGPG TGLAPFRGFL QERMALKEDG EELGSSLLFF GCRNRQMDFI YEDELNNFVD       600
QGVISELIMA FSREGAQKEY VQHKMMEKAA QVWDLIKEEG YLYVCGDAKG MARDVHRTLH       660
TIVQEQEGVS SSEAEAIVKK LQTEGRYLRD VW                                     692
```

SEQ ID NO:91

A. thaliana

```
atgtcttcct cttcctcttc cagtacctct atgattgatt tgatggctgc tattattaaa        60
ggtgaaccag ttatcgtctc cgacccagca aatgcctctg cttatgaatc agttgctgca       120
gaattgtctt caatgttgat cgaaaacaga caattcgcca tgatcgtaac tacatcaatc       180
gctgttttga tcggttgtat tgtcatgttg gtatggagaa gatccggtag tggtaattct       240
aaaagagtcg aacctttgaa accattagta attaagccaa gagaagaaga aatagatgac       300
ggtagaaaga aagttacaat atttttcggt acccaaactg gtacagctga aggttttgca       360
aaagccttag gtgaagaagc taaggcaaga tacgaaaaga ctagattcaa gatagtcgat       420
ttggatgact atgccgctga tgacgatgaa tacgaagaaa agttgaagaa agaagatgtt       480
gcatttttct ttttggcaac ctatggtgac ggtgaaccaa ctgacaatgc agccagattc       540
tacaaatggt ttacagaggg taatgatcgt ggtgaatggt tgaaaaactt aaagtacggt       600
gttttcggtt tgggtaacag acaatacgaa catttcaaca aagttgcaaa ggttgtcgac       660
gatattttgg tcgaacaagg tgctcaaaga ttagtccaag taggtttggg tgacgatgac       720
caatgtatag aagatgactt tactgcctgg agagaagctt tgtggcctga attagacaca       780
atcttgagag aagaaggtga caccgccgtt gctaccccat atactgctgc agtattagaa       840
tacagagttt ccatccatga tagtgaagac gcaaagttta atgatatcac tttggccaat       900
ggtaacggtt atacagtttt cgatgcacaa cacccttaca aagctaacgt tgcagtcaag       960
agagaattac atacaccaga atccgacaga agttgtatac acttggaatt tgatatcgct      1020
ggttccggtt taaccatgaa gttgggtgac catgtaggtg ttttatgcga caatttgtct      1080
gaaactgttg atgaagcatt gagattgttg gatatgtccc ctgacactta ttttagtttg      1140
cacgctgaaa aagaagatgg tacaccaatt tccagttctt taccacctcc attccctcca      1200
tgtaacttaa gaacagcctt gaccagatac gcttgcttgt tatcatcccc taaaaagtcc      1260
gccttggttg ctttagccgc tcatgctagt gatcctactg aagcagaaag attgaaacac      1320
ttagcatctc cagccggtaa agatgaatat tcaaagtggg tagttgaatc tcaaagatca      1380
ttgttagaag ttatggcaga atttccatct gccaagcctc cattaggtgt cttctttgct      1440
ggtgtagcac ctagattgca accaagattc tactcaatca gttcttcacc taagatcgct      1500
gaaactagaa ttcatgttac atgtgcatta gtctacgaaa agatgccaac cggtagaatt      1560
cacaagggtg tatgctctac ttggatgaaa aatgctgttc cttacgaaaa atcagaaaag      1620
ttgttcttag gtagaccaat cttcgtaaga caatcaaact tcaagttgcc ttctgattca      1680
aaggttccaa taatcatgat aggtcctggt acaggtttag ccccattcag aggtttcttg      1740
caagaaagat tggctttagt tgaatctggt gtcgaattag gtccttcagt tttgttcttt      1800
ggttgtagaa acagaagaat ggatttcatc tatgaagaag aattgcaaag attcgtcgaa      1860
tctggtgcat tggccgaatt atctgtagct ttttcaagag aaggtccaac taaggaatac      1920
gttcaacata agatgatgga taaggcatcc gacatatgga acatgatcag tcaaggtgct      1980
```

```
tatttgtacg tttgcggtga cgcaaagggt atggccagag atgtccatag atctttgcac      2040
acaattgctc aagaacaagg ttccatggat agtaccaaag ctgaaggttt cgtaaagaac      2100
ttacaaactt ccggtagata cttgagagat gtctggtga                             2139
```

SEQ ID NO:92

A. thaliana

```
MSSSSSSSTS MIDLMAAIIK GEPVIVSDPA NASAYESVAA ELSSMLIENR QFAMIVTTSI        60
AVLIGCIVML VWRRSGSGNS KRVEPLKPLV IKPREEEIDD GRKKVTIFFG TQTGTAEGFA       120
KALGEEAKAR YEKTRFKIVD LDDYAADDDE YEEKLKKEDV AFFFLATYGD GEPTDNAARF       180
YKWFTEGNDR GEWLKNLKYG VFGLGNRQYE HFNKVAKVVD DILVEQGAQR LVQVGLGDDD       240
QCIEDDFTAW REALWPELDT ILREEGDTAV ATPYTAAVLE YRVSIHDSED AKFNDITLAN       300
GNGYTVFDAQ HPYKANVAVK RELHTPESDR SCIHLEFDIA GSGLTMKLGD HVGVLCDNLS       360
ETVDEALRLL DMSPDTYFSL HAEKEDGTPI SSSLPPPFPP CNLRTALTRY ACLLSSPKKS       420
ALVALAAHAS DPTEAERLKH LASPAGKDEY SKWVVESQRS LLEVMAEFPS AKPPLGVFFA       480
GVAPRLQPRF YSISSSPKIA ETRIHVTCAL VYEKMPTGRI HKGVCSTWMK NAVPYEKSEK       540
LFLGRPIFVR QSNFKLPSDS KVPIIMIGPG TGLAPFRGFL QERLALVESG VELGPSVLFF       600
GCRNRRMDFI YEEELQRFVE SGALAELSVA FSREGPTKEY VQHKMMDKAS DIWNMISQGA       660
YLYVCGDAKG MARDVHRSLH TIAQEQGSMD STKAEGFVKN LQTSGRYLRD VW               712
```

SEQ ID NO:93

S. rebaudiana

```
atggaagcct cttacctata catttctatt ttgcttttac tggcatcata cctgttcacc        60
actcaactta gaaggaagag cgctaatcta ccaccaaccg tgtttccatc aataccaatc       120
attggacact tatacttact caaaaagcct ctttatagaa ctttagcaaa aattgccgct       180
aagtacggac caatactgca attacaactc ggctacagac gtgttctggt gatttcctca       240
ccatcagcag cagaagagtg ctttaccaat aacgatgtaa tcttcgcaaa tagacctaag       300
acattgtttg gcaaaatagt gggtggaaca tcccttggca gtttatccta cggcgatcaa       360
tggcgtaatc taaggagagt agcttctatc gaaatcctat cagttcatag gttgaacgaa       420
tttcatgata tcagagtgga tgagaacaga ttgttaatta gaaaacttag aagttcatct       480
tctcctgtta ctcttataac agtcttttat gctctaacat tgaacgtcat tatgagaatg       540
atctctggca aaagatattt cgacagtggg gatagagaat tggaggagga aggtaagaga       600
tttcgagaaa tcttagacga aacgttgctt ctagccggtg cttctaatgt tggcgactac       660
ttaccaatat tgaactggtt gggagttaag tctcttgaaa agaaattgat cgctttgcag       720
aaaaagagag atgacttttt ccagggtttg attgaacagg ttagaaaatc tcgtggtgct       780
aaagtaggca aaggtagaaa aacgatgatc gaactcttat tatctttgca agagtcagaa       840
cctgagtact atacagatgc tatgataaga tcttttgtcc taggtctgct ggctgcaggt       900
agtgatactt cagcgggcac tatggaatgg gccatgagct tactggtcaa tcacccacat       960
gtattgaaga aagctcaagc tgaaatcgat agagttatcg gtaataacag attgattgac      1020
gagtcagaca ttggaaatat cccttacatc gggtgtatta tcaatgaaac tctaagactc      1080
tatccagcag ggccattgtt gttcccacat gaaagttctg ccgactgcgt tatttccggt      1140
tacaatatac ctagaggtac aatgttaatc gtaaaccaat gggcgattca tcacgatcct      1200
aaagtctggg atgatcctga aacctttaaa cctgaaagat ttcaaggatt agaaggaact      1260
agagatggtt tcaaacttat gccattcggt tctgggagaa gaggatgtcc aggtgaaggt      1320
ttggcaataa ggctgttagg gatgacacta ggctcagtga tccaatgttt tgattgggag      1380
agagtaggag atgagatggt tgacatgaca gaaggtttgg gtgtcacact tcctaaggcc      1440
gttccattag ttgccaaatg taagccacgt tccgaaatga ctaatctcct atccgaactt      1500
taa                                                                   1503
```

SEQ ID NO:94

S. rebaudiana

```
MEASYLYISI LLLLASYLFT TQLRRKSANL PPTVFPSIPI IGHLYLLKKP LYRTLAKIAA        60
KYGPILQLQL GYRRVLVISS PSAAEECFTN NDVIFANRPK TLFGKIVGGT SLGSLSYGDQ       120
WRNLRRVASI EILSVHRLNE FHDIRVDENR LLIRKLRSSS SPVTLITVFY ALTLNVIMRM       180
```

```
ISGKRYFDSG DRELEEEGKR FREILDETLL LAGASNVGDY LPILNWLGVK SLEKKLIALQ       240
KKRDDFFQGL IEQVRKSRGA KVGKGRKTMI ELLLSLQESE PEYYTDAMIR SFVLGLLAAG       300
SDTSAGTMEW AMSLLVNHPH VLKKAQAEID RVIGNNRLID ESDIGNIPYI GCIINETLRL       360
YPAGPLLFPH ESSADCVISG YNIPRGTMLI VNQWAIHHDP KVWDDPETFK PERFQGLEGT       420
RDGFKLMPFG SGRRGCPGEG LAIRLLGMTL GSVIQCFDWE RVGDEMVDMT EGLGVTLPKA       480
VPLVAKCKPR SEMTNLLSEL                                                   500
```

SEQ ID NO:95

```
atggaagtaa cagtagctag tagtgtagcc ctgagcctgg tctttattag catagtagta        60
agatgggcat ggagtgtggt gaattgggtg tggtttaagc cgaagaagct ggaaagattt       120
ttgagggagc aaggccttaa aggcaattcc tacaggtttt tatatggaga catgaaggag       180
aactctatcc tgctcaaaca agcaagatcc aaacccatga acctctccac ctcccatgac       240
atagcacctc aagtcacccc ttttgtcgac caaaccgtga aagcttacgg taagaactct       300
tttaattggg ttggccccat accaagggtg aacataatga atccagaaga tttgaaggac       360
gtcttaacaa aaaatgttga ctttgttaag ccaatatcaa acccacttat caagttgcta       420
gctacaggta ttgcaatcta tgaaggtgag aaatggacta aacacagaag gattatcaac       480
ccaacattcc attcggagag gctaaagcgt atgttacctt catttcacca aagttgtaat       540
gagatggtca aggaatggga gagcttggtg tcaaaagagg gttcatcatg tgagttggat       600
gtctggcctt ttcttgaaaa tatgtcggca gatgtgatct cgagaacagc atttggaact       660
agctacaaaa aaggacagaa aatctttgaa ctcttgagag agcaagtaat atatgtaacg       720
aaaggctttc aaagttttta cattccagga tggaggtttc tcccaactaa gatgaacaag       780
aggatgaatg agattaacga agaaataaaa ggattaatca ggggtattat aattgacaga       840
gagcaaatca ttaaggcagg tgaagaaacc aacgatgact tattaggtgc acttatggag       900
tcaaacttga aggacattcg ggaacatggg aaaaacaaca aaaatgttgg gatgagtatt       960
gaagatgtaa ttcaggagtg taagctgttt tactttgctg ggcaagaaac cacttcagtg      1020
ttgctggctt ggacaatggt tttacttggt caaaatcaga actggcaaga tcgagcaaga      1080
caagaggttt tgcaagtctt tggaagcagc aagccagatt ttgatggtct agctcacctt      1140
aaagtcgtaa ccatgatttt gcttgaagtt cttcgattat acccaccagt cattgaactt      1200
attcgaacca ttcacaagaa aacacaactt gggaagctct cactaccaga aggagttgaa      1260
gtccgcttac caacactgct cattcaccat gacaaggaac tgtggggtga tgatgcaaac      1320
cagttcaatc cagagaggtt ttcggaagga gtttccaaag caacaaagaa ccgactctca      1380
ttcttccccc tcggagccgg tccacgcatt tgcattggac agaacttttc tatgatggaa      1440
gcaaagttgg ccttagcatt gatcttgcaa cacttcacct ttgagctttc tccatctcat      1500
gcacatgctc cttcccatcg tataaccctt caaccacagt atggtgttcg tatcatttta      1560
catcgacgtt ag                                                         1572
```

SEQ ID NO:96

R. suavissimus

```
atggaagtca ctgtcgcctc ttctgtcgct ttatccttag tcttcatttc cattgtcgtc        60
agatgggctt ggtccgttgt caactgggtt tggttcaaac caaagaagtt ggaaagattc       120
ttgagagagc aaggtttgaa gggtaattct tatagattct tgtacggtga catgaaggaa       180
aattctattt tgttgaagca agccagatcc aaaccaatga acttgtctac ctctcatgat       240
attgctccac aagttactcc attcgtcgat caaactgtta aagcctacgg taagaactct       300
ttcaattggg ttggtccaat tcctagagtt aacatcatga acccagaaga tttgaaggat       360
gtcttgacca agaacgttga cttcgttaag ccaatttcca acccattgat taaattgttg       420
gctactggta ttgccattta cgaaggtgaa aagtggacta agcatagaag aatcatcaac       480
cctaccttcc actctgaaag attgaagaga atgttaccat ctttccatca atcctgtaat       540
gaaatggtta aggaatggga atccttggtt tctaaagaag gttcttcttg cgaattggat       600
gtttggccat tcttggaaaa tatgtctgct gatgtcattt ccagaaccgc tttcggtacc       660
tcctacaaga agggtcaaaa gattttcgaa ttgttgagag agcaagttat ttacgttacc       720
aagggtttcc aatccttcta catcccaggt tggagattct tgccaactaa aatgaacaag       780
cgtatgaacg agatcaacga agaaattaaa ggtttgatca gaggtattat tatcgacaga       840
gaacaaatta ttaaagctgg tgaagaaacc aacgatgatt tgttgggtgc tttgatggag       900
tccaacttga aggatattag agaacatggt aagaacaaca agaatgttgg tatgtctatt       960
```

```
gaagatgtta ttcaagaatg taagttattc tacttcgctg gtcaagagac cacttctgtt      1020
ttgttagcct ggactatggt cttgttaggt caaaaccaaa attggcaaga tagagctaga      1080
caagaagttt tgcaagtctt cggttcttcc aagccagact ttgatggttt ggcccacttg      1140
aaggttgtta ctatgatttt gttagaagtt ttgagattgt acccaccagt cattgagtta      1200
atcagaacca ttcataaaaa gactcaattg ggtaaattat ctttgccaga aggtgttgaa      1260
gtcagattac caaccttgtt gattcaccac gataaggaat tatggggtga cgacgctaat      1320
caatttaatc cagaaagatt ttccgaaggt gtttccaagg ctaccaaaaa ccgtttgtcc      1380
ttcttcccat ttggtgctgg tccacgtatt tgtatcggtc aaaacttttc catgatggaa      1440
gccaagttgg ctttggcttt aatcttgcaa cacttcactt tcgaattgtc tccatcccat      1500
gcccacgctc cttctcatag aatcacttta caaccacaat acggtgtcag aatcatctta      1560
cacagaagat aa                                                          1572
```

SEQ ID NO:97

R. suavissimus

```
MEVTVASSVA LSLVFISIVV RWAWSVVNWV WFKPKKLERF LREQGLKGNS YRFLYGDMKE        60
NSILLKQARS KPMNLSTSHD IAPQVTPFVD QTVKAYGKNS FNWVGPIPRV NIMNPEDLKD       120
VLTKNVDFVK PISNPLIKLL ATGIAIYEGE KWTKHRRIIN PTFHSERLKR MLPSFHQSCN       180
EMVKEWESLV SKEGSSCELD VWPFLENMSA DVISRTAFGT SYKKGQKIFE LLREQVIYVT       240
KGFQSFYIPG WRFLPTKMNK RMNEINEEIK GLIRGIIIDR EQIIKAGEET NDDLLGALME       300
SNLKDIREHG KNNKNVGMSI EDVIQECKLF YFAGQETTSV LLAWTMVLLG QNQNWQDRAR       360
QEVLQVFGSS KPDFDGLAHL KVVTMILLEV LRLYPPVIEL IRTIHKKTQL GKLSLPEGVE       420
VRLPTLLIHH DKELWGDDAN QFNPERFSEG VSKATKNRLS FFPFGAGPRI CIGQNFSMME       480
AKLALALILQ HFTFELSPSH AHAPSHRITL QPQYGVRIIL HRR                         523
```

SEQ ID NO:98

```
atggaagcat caagggctag ttgtgttgcg ctatgtgttg tttgggtgag catagtaatt        60
acattggcat ggagggtgct gaattgggtg tggttgaggc caaagaaact agaaagatgc       120
ttgagggagc aaggccttac aggcaattct tacaggcttt tgtttggaga caccaaggat       180
ctctcgaaga tgctggaaca aacacaatcc aaacccatca aactctccac ctcccatgat       240
atagcgccac gagtcacccc atttttccat cgaactgtga actctaatgg caagaattct       300
tttgtttgga tgggccctat accaagagtg cacatcatga atccagaaga tttgaaagat       360
gccttcaaca gacatgatga ttttcataag acagtaaaaa atcctatcat gaagtctcca       420
ccaccgggca ttgtaggcat tgaaggtgag caatgggcta aacacagaaa gattatcaac       480
ccagcattcc atttagagaa gctaaagggt atggtaccaa tattttacca aagttgtagc       540
gagatgatta acaaatggga gagcttggtg tccaaagaga gttcatgtga gttggatgtg       600
tggccttatc ttgaaaattt taccagcgat gtgatttccc gagctgcatt tggaagtagc       660
tatgaagagg gaaggaaaat atttcaacta ctaagagagg aagcaaaagt ttattcggta       720
gctctacgaa gtgtttacat tccaggatgg aggtttctac caaccaagca gaacaagaag       780
acgaaggaaa ttcacaatga aattaaaggc ttacttaagg gcattataaa taaaagggaa       840
gaggcgatga aggcagggga agccactaaa gatgacttac taggaatact tatggagtcc       900
aacttcaggg aaattcagga acatgggaac aacaaaaatg ctggaatgag tattgaagat       960
gtaattggag agtgtaagtt gttttacttt gctgggcaag agaccacttc ggtgttgctt      1020
gtttggacaa tgattttact aagccaaaat caggattggc aagctcgtgc aagagaagag      1080
gtcttgaaag tctttggaag caacatccca acctatgaag agctaagtca cctaaaagtt      1140
gtgaccatga ttttacttga agttcttcga ttatacccat cagtcgttgc gcttcctcga      1200
accactcaca agaaaacaca gcttggaaaa ttatcattac cagctggagt ggaagtctcc      1260
ttgcccatac tgcttgttca ccatgacaaa gagttgtggg gtgaggatgc aaatgagttc      1320
aagccagaga ggttttcaga gggagtttca aaggcaacaa agaacaaatt tacatactta      1380
cctttcggag ggggtccaag gatttgcatt ggacaaaact ttgccatggt ggaagctaaa      1440
ttggccttgg ccctgatttt acaacacttt gcctttgagc tttctccatc ctatgctcat      1500
gctccttctg cagttataac ccttcaacct caatttggtg ctcatatcat tttgcataaa      1560
cgttga                                                                 1566
```

SEQ ID NO:99

```
atggaagctt ctagagcatc ttgtgttgct ttgtgtgttg tttgggtttc catcgttatt        60
actttggctt ggagagtttt gaattgggtc tggttaagac caaaaaagtt ggaaagatgc       120
ttgagagaac aaggtttgac tggtaactct tacagattgt tgttcggtga taccaaggac       180
ttgtctaaga tgttggaaca aactcaatcc aagcctatca agttgtctac ctctcatgat       240
attgctccaa gagttactcc attcttccat agaactgtta actccaacgg taagaactct       300
tttgtttgga tgggtccaat tccaagagtc catattatga accctgaaga tttgaaggac       360
gctttcaaca gacatgatga tttccataag accgtcaaga acccaattat gaagtctcca       420
ccaccaggta tagttggtat tgaaggtgaa caatgggcca aacatagaaa gattattaac       480
ccagccttcc acttggaaaa gttgaaaggt atggttccaa tcttctacca atcctgctct       540
gaaatgatta acaagtggga atccttggtt tccaaagaat cttcctgtga attggatgtc       600
tggccatatt tggaaaactt cacctccgat gttatttcca gagctgcttt tggttcttct       660
tacgaagaag gtagaaagat cttccaatta ttgagagaag aagccaaggt ttactccgtt       720
gctttgagat ctgtttacat tccaggttgg agattcttgc caactaagca aaacaaaaag       780
accaaagaaa tccacaacga aatcaagggt ttgttgaagg gtatcatcaa caagagagaa       840
gaagctatga aggctggtga agctacaaaa gatgatttgt tgggtatctt gatggaatcc       900
aacttcagag aaatccaaga acacggtaac aacaagaatg ccggtatgtc tattgaagat       960
gttatcggtg aatgcaagtt gttctacttt gctggtcaag aaactacctc cgttttgttg      1020
gtttggacca tgattttgtt gtcccaaaat caagattggc aagctagagc tagagaagaa      1080
gtcttgaaag ttttcggttc taacatccca acctacgaag aattgtctca cttgaaggtt      1140
gtcactatga tcttgttgga agtattgaga ttatacccat ccgttgttgc attgccaaga      1200
actactcata agaaaactca attgggtaaa ttgtccttgc cagctggtgt tgaagtttct      1260
ttgccaattt tgttagtcca ccacgacaaa gaattgtggg gtgaagatgc taatgaattc      1320
aagccagaaa gattctccga aggtgtttct aaagctacca agaacaagtt cacttacttg      1380
ccatttggtg gtggtccaag aatatgtatt ggtcaaaatt tcgctatggt cgaagctaaa      1440
ttggctttgg ctttgatctt gcaacatttc gctttcgaat tgtcaccatc ttatgctcat      1500
gctccatctg ctgttattac attgcaacca caatttggtg cccatatcat cttgcataag      1560
agataac                                                                1567
```

SEQ ID NO:100

```
MEASRASCVA LCVVWVSIVI TLAWRVLNWV WLRPKKLERC LREQGLTGNS YRLLFGDTKD        60
LSKMLEQTQS KPIKLSTSHD IAPRVTPFFH RTVNSNGKNS FVWMGPIPRV HIMNPEDLKD       120
AFNRHDDFHK TVKNPIMKSP PPGIVGIEGE QWAKHRKIIN PAFHLEKLKG MVPIFYQSCS       180
EMINKWESLV SKESSCELDV WPYLENFTSD VISRAAFGSS YEEGRKIFQL LREEAKVYSV       240
ALRSVYIPGW RFLPTKQNKK TKEIHNEIKG LLKGIINKRE EAMKAGEATK DDLLGILMES       300
NFREIQEHGN NKNAGMSIED VIGECKLFYF AGQETTSVLL VWTMILLSQN QDWQARAREE       360
VLKVFGSNIP TYEELSHLKV VTMILLEVLR LYPSVVALPR TTHKKTQLGK LSLPAGVEVS       420
LPILLVHHDK ELWGEDANEF KPERFSEGVS KATKNKFTYL PFGGGPRICI GQNFAMVEAK       480
LALALILQHF AFELSPSYAH APSAVITLQP QFGAHIILHK R                          521
```

SEQ ID NO:101

```
ASWVAVLSVV WVSMVIAWAW RVLNWVWLRP KKLEKCLREQ GLAGNSYRLL FGDTKDLSKM        60
LEQTQSKPIK LSTSHDIAPH VTPFFHQTVN SYGKNSFVWM GPIPRVHIMN PEDLKDTFNR       120
HDDFHKVVKN PIMKSLPQGI VGIEGEQWAK HRKIINPAFH LEKLKGMVPI FYRSCSEMIN       180
KWESLVSKES SCELDVWPYL ENFTSDVISR AAFGSSYEEG RKIFQLLREE AKIYTVAMRS       240
VYIPGWRFLP TKQNKKAKEI HNEIKGLLKG IINKREEAMK AGEATKDDLL GILMESNFRE       300
IQEHGNNKNA GMSIEDVIGE CKLFYFAGQE TTSVLLVWTM VLLSQNQDWQ ARAREEVLQV       360
FGSNIPTYEE LSQLKVVTMI LLEVLRLYPS VVALPRTTHK KTQLGKLSLP AGVEVSLPIL       420
LVHHDKELWG EDANEFKPER FSEGVSKATK NQFTYFPFGG GPRICIGQNF AMMEAKLALS       480
LILRHFALEL SPLYAHAPSV TITLQPQYGA HIILHKR                                517
```

SEQ ID NO:102

```
MEASRPSCVA LSVVLVSIVI AWAWRVLNWV WLRPNKLERC LREQGLTGNS YRLLFGDTKE        60
```

```
ISMMVEQAQS KPIKLSTTHD IAPRVIPFSH QIVYTYGRNS FVWMGPTPRV TIMNPEDLKD       120
AFNKSDEFQR AISNPIVKSI SQGLSSLEGE KWAKHRKIIN PAFHLEKLKG MLPTFYQSCS       180
EMINKWESLV FKEGSREMDV WPYLENLTSD VISRAAFGSS YEEGRKIFQL LREEAKFYTI       240
AARSVYIPGW RFLPTKQNKR MKEIHKEVRG LLKGIINKRE DAIKAGEAAK GNLLGILMES       300
NFREIQEHGN NKNAGMSIED VIGECKLFYF AGQETTSVLL VWTLVLLSQN QDWQARAREE       360
VLQVFGTNIP TYDQLSHLKV VTMILLEVLR LYPAVVELPR TTYKKTQLGK FLLPAGVEVS       420
LHIMLAHHDK ELWGEDAKEF KPERFSEGVS KATKNQFTYF PFGAGPRICI GQNFAMLEAK       480
LALSLILQHF TFELSPSYAH APSVTITLHP QFGAHFILHK R                           521
```

SEQ ID NO:103

```
CVALSVVLVS IVIAWAWRVL NWVWLRPNKL ERCLREQGLT GNSYRLLFGD TKEISMMVEQ        60
AQSKPIKLST THDIAPRVIP FSHQIVYTYG RNSFVWMGPT PRVTIMNPED LKDAFNKSDE       120
FQRAISNPIV KSISQGLSSL EGEKWAKHRK IINPAFHLEK LKGMLPTFYQ SCSEMINKWE       180
SLVFKEGSRE MDVWPYLENL TSDVISRAAF GSSYEEGRKI FQLLREEAKF YTIAARSVYI       240
PGWRFLPTKQ NKRMKEIHKE VRGLLKGIIN KREDAIKAGE AAKGNLLGIL MESNFREIQE       300
HGNNKNAGMS IEDVIGECKL FYFAGQETTS VLLVWTLVLL SQNQDWQARA REEVLQVFGT       360
NIPTYDQLSH LKVVTMILLE VLRLYPAVVE LPRTTYKKTQ LGKFLLPAGV EVSLHIMLAH       420
HDKELWGEDA KEFKPERFSE GVSKATKNQF TYFPFGAGPR ICIGQNFAML EAKLALSLIL       480
QHFTFELSPS YAHAPSVTIT LHPQFGAHFI LHKR                                   514
```

SEQ ID NO:104

```
MGPIPRVHIM NPEDLKDTFN RHDDFHKVVK NPIMKSLPQG IVGIEGDQWA KHRKIINPAF        60
HLEKLKGMVP IFYQSCSEMI NIWKSLVSKE SSCELDVWPY LENFTSDVIS RAAFGSSYEE       120
GRKIFQLLRE EAKVYTVAVR SVYIPGWRFL PTKQNKKTKE IHNEIKGLLK GIINKREEAM       180
KAGEATKDDL LGILMESNFR EIQEHGNNKN AGMSIEDVIG ECKLFYFAGQ ETTSVLLVWT       240
MVLLSQNQDW QARAREEVLQ VFGSNIPTYE ELSHLKVVTM ILLEVLRLYP SVVALPRTTH       300
KKTQLGKLSL PAGVEVSLPI LLVHHDKELW GEDANEFKPE RFSEGVSKAT KNQFTYFPFG       360
GGPRICIGQN FAMMEAKLAL SLILQHFTFE LSPQYSHAPS VTITLQPQYG AHLILHKR         418
```

SEQ ID NO:105

```
atgggtttgt tcccattaga ggattcctac gcgctggtct ttgaaggact agcaataaca        60
ctggctttgt actatctact gtctttcatc tacaaaacat ctaaaaagac atgtacacct       120
cctaaagcat ctggtgaaat cattccaatt acaggaatca tattgaatct gctatctggc       180
tcaagtggtc tacctattat cttagcactt gcctctttag cagacagatg tggtcctatt       240
ttcaccatta ggctgggtat taggagagtg ctagtagtat caaattggga aatcgctaag       300
gagattttca ctacccacga tttgatagtt tctaatagac caaaatactt agccgctaag       360
attcttggtt tcaattatgt ttcattctct ttcgctccat acggcccata ttgggtcgga       420
atcagaaaga ttattgctac aaaactaatg tcttcttcca gacttcagaa gttgcaattt       480
gtaagagttt ttgaactaga aaactctatg aaatctatca gagaatcatg gaaggagaaa       540
aaggatgaag agggaaaggt attagttgag atgaaaaagt ggttctggga actgaatatg       600
aacatagtgt taaggacagt tgctggtaaa caatacactg gtacagttga tgatgccgat       660
gcaaagcgta tctccgagtt attcagagaa tggtttcact acactggcag atttgtcgtt       720
ggagacgctt ttccttttct aggttggttg gacctgggcg gatacaaaaa gacaatggaa       780
ttagttgcta gtagattgga ctcaatggtc agtaaatggt tagatgagca tcgtaaaaag       840
caagctaacg atgacaaaaa ggaggatatg gatttcatgg atatcatgat ctccatgaca       900
gaagcaaatt caccacttga aggatacggc actgatacta ttatcaagac cacatgtatg       960
actttgattg tttcaggagt tgatacaacc tcaatcgtac ttacttgggc cttatcactt      1020
ttgttaaaca acagagatac tttgaaaaag gcacaagagg aattagatat gtgcgtaggt      1080
aaaggaagac aagtcaacga gtctgatctt gttaacttga tatacttgga agcagtgctt      1140
aaagaggctt taagacttta cccagcagcg ttcttaggcg gaccaagagc attcttggaa      1200
gattgtactg ttgctggtta tagaattcca aagggcacct gcttgttgat taacatgtgg      1260
aaactgcata gagatccaaa catttggagt gatccttgcg aattcaagcc agaaagattt      1320
ttgacaccta atcaaaagga tgttgatgtg atcggtatgg atttcgaatt gataccattt      1380
```

```
ggtgccggca gaagatattg tccaggtact agattggctt tacagatgtt gcatatcgta      1440
ttagcgacat tgctgcaaaa cttcgaaatg tcaacaccaa acgatgcgcc agtcgatatg      1500
actgcttctg ttggcatgac aaatgccaaa gcatcacctt tagaagtctt gctatcacct      1560
cgtgttaaat ggtcctaa                                                    1578
```

SEQ ID NO:106

```
MGLFPLEDSY ALVFEGLAIT LALYYLLSFI YKTSKKTCTP PKASGEHPIT GHLNLLSGSS        60
GLPHLALASL ADRCGPIFTI RLGIRRVLVV SNWEIAKEIF TTHDLIVSNR PKYLAAKILG       120
FNYVSFSFAP YGPYWVGIRK IIATKLMSSS RLQKLQFVRV FELENSMKSI RESWKEKKDE       180
EGKVLVEMKK WFWELNMNIV LRTVAGKQYT GTVDDADAKR ISELFREWFH YTGRFVVGDA       240
FPFLGWLDLG GYKKTMELVA SRLDSMVSKW LDEHRKKQAN DDKKEDMDFM DIMISMTEAN       300
SPLEGYGTDT IIKTTCMTLI VSGVDTTSIV LTWALSLLLN NRDTLKKAQE ELDMCVGKGR       360
QVNESDLVNL IYLEAVLKEA LRLYPAAFLG GPRAFLEDCT VAGYRIPKGT CLLINMWKLH       420
RDPNIWSDPC EFKPERFLTP NQKDVDVIGM DFELIPFGAG RRYCPGTRLA LQMLHIVLAT       480
LLQNFEMSTP NDAPVDMTAS VGMTNAKASP LEVLLSPRVK WS                          522
```

SEQ ID NO:107

```
atgatacaag ttttaactcc aattctactc ttcctcatct tcttcgtttt ctggaaagtc        60
tacaaacatc aaaagactaa aatcaatcta ccaccaggtt ccttcggctg gccatttttg       120
ggtgaaacct tagccttact tagagcaggc tgggattctg agccagaaag attcgtaaga       180
gagcgtatca aaaagcatgg atctccactt gttttcaaga catcactatt tggagacaga       240
ttcgctgttc tttgcggtcc agctggtaat aagtttttgt tctgcaacga aaacaaatta       300
gtggcatctt ggtggccagt ccctgtaagg aagttgttcg gtaaaagttt actcacaata       360
agaggagatg aagcaaaatg gatgagaaaa atgctattgt cttacttggg tccagatgca       420
tttgccacac attatgccgt tactatggat gttgtaacac gtagacatat tgatgtccat       480
tggaggggca aggaggaagt taatgtattt caaacagtta agttgtacgc attcgaatta       540
gcttgtagat tattcatgaa cctagatgac ccaaaccaca tcgcgaaact cggtagtctt       600
ttcaacattt tcctcaaagg gatcatcgag cttcctatag acgttcctgg aactagattt       660
tactccagta aaaaggccgc agctgccatt agaattgaat tgaaaaagct cattaaagct       720
agaaaactcg aattgaagga gggtaaggcg tcttcttcac aggacttgct ttctcatcta       780
ttaacatcac ctgatgagaa tgggatgttc ttgacagaag aggaaatagt cgataacatt       840
ctacttttgt tattcgctgg tcacgatacc tctgcactat caataacact tttgatgaaa       900
accttaggtg aacacagtga tgtgtacgac aaggttttga aggaacaatt agaaatttcc       960
aaaacaaagg aggcttggga atcactaaag tgggaagata tccagaagat gaagtactca      1020
tggtcagtaa tctgtgaagt catgagattg aatcctcctg tcatagggac atacagagag      1080
gcgttggttg atatcgacta tgctggttac actatcccaa aaggatggaa gttgcattgg      1140
tcagctgttt ctactcaaag agacgaagcc aatttcgaag atgtaactag attcgatcca      1200
tccagatttg aaggggcagg ccctactcca ttcacatttg tgcctttcgg tggaggtcct      1260
agaatgtgtt taggcaaaga gtttgccagg ttagaagtgt tagcatttct ccacaacatt      1320
gttaccaact ttaagtggga tcttctaatc cctgatgaga agatcgaata tgatccaatg      1380
gctactccag ctaagggctt gccaattaga cttcatccac accaagtcta a               1431
```

SEQ ID NO:108

```
MIQVLTPILL FLIFFVFWKV YKHQKTKINL PPGSFGWPFL GETLALLRAG WDSEPERFVR        60
ERIKKHGSPL VFKTSLFGDR FAVLCGPAGN KFLFCNENKL VASWWPVPVR KLFGKSLLTI       120
RGDEAKWMRK MLLSYLGPDA FATHYAVTMD VVTRRHIDVH WRGKEEVNVF QTVKLYAFEL       180
ACRLFMNLDD PNHIAKLGSL FNIFLKGIIE LPIDVPGTRF YSSKKAAAAI RIELKKLIKA       240
RKLELKEGKA SSSQDLLSHL LTSPDENGMF LTEEEIVDNI LLLLFAGHDT SALSITLLMK       300
TLGEHSDVYD KVLKEQLEIS KTKEAWESLK WEDIQKMKYS WSVICEVMRL NPPVIGTYRE       360
ALVDIDYAGY TIPKGWKLHW SAVSTQRDEA NFEDVTRFDP SRFEGAGPTP FTFVPFGGGP       420
RMCLGKEFAR LEVLAFLHNI VTNFKWDLLI PDEKIEYDPM ATPAKGLPIR LHPHQV           476
```

SEQ ID NO:109

```
atggagtctt tagtggttca tacagtaaat gctatctggt gtattgtaat cgtcgggatt        60
ttctcagttg gttatcacgt ttacggtaga gctgtggtcg aacaatggag aatgagaaga       120
tcactgaagc tacaaggtgt taaaggccca ccaccatcca tcttcaatgg taacgtctca       180
gaaatgcaac gtatccaatc cgaagctaaa cactgctctg gcgataacat tatctcacat       240
gattattctt cttcattatt cccacacttc gatcactgga gaaaacagta cggcagaatc       300
tacacatact ctactggatt aaagcaacac ttgtacatca atcatccaga aatggtgaag       360
gagctatctc agactaacac attgaacttg ggtagaatca cccatataac caaaagattg       420
aatcctatct taggtaacgg aatcataacc tctaatggtc ctcattgggc ccatcagcgt       480
agaattatcg cctacgagtt tactcatgat aagatcaagg gtatggttgg tttgatggtt       540
gagtctgcta tgcctatgtt gaataagtgg gaggagatgg taaagagagg cggagaaatg       600
ggatgcgaca taagagttga tgaggacttg aaagatgttt cagcagatgt gattgcaaaa       660
gcctgtttcg gatcctcatt ttctaaaggt aaggctattt tctctatgat aagagatttg       720
cttacagcta tcacaaagag aagtgttcta ttcagattca acggattcac tgatatggtc       780
tttgggagta aaaagcatgg tgacgttgat atagacgctt tagaaatgga attggaatca       840
tccatttggg aaactgtcaa ggaacgtgaa atagaatgta aagatactca caaaaaggat       900
ctgatgcaat tgattttgga aggggcaatg cgttcatgtg acggtaacct ttgggataaa       960
tcagcatata gaagatttgt tgtagataat tgtaaatcta tctacttcgc agggcatgat      1020
agtacagctg tctcagtgtc atggtgtttg atgttactgg ccctaaaccc atcatggcaa      1080
gttaagatcc gtgatgaaat tctgtcttct tgcaaaaatg gtattccaga tgccgaaagt      1140
atcccaaacc ttaaaacagt gactatggtt attcaagaga caatgagatt ataccctcca      1200
gcaccaatcg tcgggagaga agcctctaaa gatatcagat tgggcgatct agttgttcct      1260
aaaggcgtct gtatatggac actaatacca gctttacaca gagatcctga gatttgggga      1320
ccagatgcaa acgatttcaa accagaaaga ttttctgaag gaatttcaaa ggcttgtaag      1380
tatcctcaaa gttacattcc atttggtctg ggtcctagaa catgcgttgg taaaaacttt      1440
ggcatgatgg aagtaaaggt tcttgtttcc ctgattgtct ccaagttctc tttcactcta      1500
tctcctacct accaacatag tcctagtcac aaacttttag tagaaccaca acatggggtg      1560
gtaattagag tggtttaa                                                    1578
```

SEQ ID NO:110

```
MESLVVHTVN AIWCIVIVGI FSVGYHVYGR AVVEQWRMRR SLKLQGVKGP PPSIFNGNVS        60
EMQRIQSEAK HCSGDNIISH DYSSSLFPHF DHWRKQYGRI YTYSTGLKQH LYINHPEMVK       120
ELSQTNTLNL GRITHITKRL NPILGNGIIT SNGPHWAHQR RIIAYEFTHD KIKGMVGLMV       180
ESAMPMLNKW EEMVKRGGEM GCDIRVDEDL KDVSADVIAK ACFGSSFSKG KAIFSMIRDL       240
LTAITKRSVL FRFNGFTDMV FGSKKHGDVD IDALEMELES SIWETVKERE IECKDTHKKD       300
LMQLILEGAM RSCDGNLWDK SAYRRFVVDN CKSIYFAGHD STAVSVSWCL MLLALNPSWQ       360
VKIRDEILSS CKNGIPDAES IPNLKTVTMV IQETMRLYPP APIVGREASK DIRLGDLVVP       420
KGVCIWTLIP ALHRDPEIWG PDANDFKPER FSEGISKACK YPQSYIPFGL GPRTCVGKNF       480
GMMEVKVLVS LIVSKFSFTL SPTYQHSPSH KLLVEPQHGV VIRVV                       525
```

SEQ ID NO:111

```
atgtacttcc tactacaata cctcaacatc acaaccgttg gtgtctttgc cacattgttt        60
ctctcttatt gtttacttct ctggagaagt agagcgggta acaaaaagat tgccccagaa       120
gctgccgctg catggcctat tatcggccac ctccacttac ttgcaggtgg atcccatcaa       180
ctaccacata ttacattggg taacatggca gataagtacg gtcctgtatt cacaatcaga       240
ataggcttgc atagagctgt agttgtctca tcttgggaaa tggcaaagga atgttcaaca       300
gctaatgatc aagtgtcttc ttcaagacct gaactattag cttctaagtt gttgggttat       360
aactacgcca tgtttggttt ttcaccatac ggttcatact ggagagaaat gagaaagatc       420
atctctctcg aattactatc taattccaga ttggaactat tgaaagatgt tagagcctca       480
gaagttgtca catctattaa ggaactatac aaattgtggg cggaaaagaa gaatgagtca       540
ggattggttt ctgtcgagat gaaacaatgg ttcggagatt tgactttaaa cgtgatcttg       600
agaatggtgg ctggtaaaag atacttctcc gcgagtgacg cttcagaaaa caaacaggcc       660
cagcgttgta gaagagtctt cagagaattc ttccatctct ccggcttgtt tgtggttgct       720
gatgctatac cttttcttgg atggctcgat tggggaagac acgagaagac cttgaaaaag       780
accgccatag aaatggattc catcgcccag gagtggcttg aggaacatag acgtagaaaa       840
```

gattctggag atgataattc tacccaagat ttcatggacg ttatgcaatc tgtgctagat 900
ggcaaaaatc taggcggata cgatgctgat acgattaaca aggctacatg cttaactctt 960
atatcaggtg gcagtgatac tactgtagtt tctttgacat gggctcttag tcttgtgtta 1020
aacaatagag atactttgaa aaaggcacag gaagagttag acatccaagt cggtaaggaa 1080
agattggtta acgagcaaga catcagtaag ttagtttact tgcaagcaat agtaaaagag 1140
acactcagac tttatccacc aggtcctttg ggtggtttga gacaattcac tgaagattgt 1200
acactaggtg gctatcacgt ttcaaaagga actagattaa tcatgaactt atccaagatt 1260
caaaaagatc cacgtatttg gtctgatcct actgaattcc aaccagagag attccttacg 1320
actcataaag atgtcgatcc acgtggtaaa cactttgaat tcattccatt cggtgcagga 1380
agacgtgcat gtcctggtat cacattcgga ttacaagtac tacatctaac attggcatct 1440
ttcttgcatg cgtttgaatt ttcaacacca tcaaatgagc aggttaacat gagagaatca 1500
ttaggtctta cgaatatgaa atctacccca ttagaagttt tgatttctcc aagactatcc 1560
cttaattgct tcaaccttat gaaaatttga 1590

SEQ ID NO:112

MYFLLQYLNI TTVGVFATLF LSYCLLLWRS RAGNKKIAPE AAAAWPIIGH LHLLAGGSHQ 60
LPHITLGNMA DKYGPVFTIR IGLHRAVVVS SWEMAKECST ANDQVSSSRP ELLASKLLGY 120
NYAMFGFSPY GSYWREMRKI ISLELLSNSR LELLKDVRAS EVVTSIKELY KLWAEKKNES 180
GLVSVEMKQW FGDLTLNVIL RMVAGKRYFS ASDASENKQA QRCRRVFREF FHLSGLFVVA 240
DAIPFLGWLD WGRHEKTLKK TAIEMDSIAQ EWLEEHRRRK DSGDDNSTQD FMDVMQSVLD 300
GKNLGGYDAD TINKATCLTL ISGGSDTTVV SLTWALSLVL NNRDTLKKAQ EELDIQVGKE 360
RLVNEQDISK LVYLQAIVKE TLRLYPPGPL GGLRQFTEDC TLGGYHVSKG TRLIMNLSKI 420
QKDPRIWSDP TEFQPERFLT THKDVDPRGK HFEFIPFGAG RRACPGITFG LQVLHLTLAS 480
FLHAFEFSTP SNEQVNMRES LGLTNMKSTP LEVLISPRLS SCSLYN 526

SEQ ID NO:113

atggaaccta acttttactt gtcattacta ttgttgttcg tgaccttcat ttctttaagt 60
ctgtttttca tcttttacaa acaaaagtcc ccattgaatt tgccaccagg gaaaatgggt 120
taccctatca taggtgaaag tttagaattc ctatccacag gctggaaggg acatcctgaa 180
aagttcatat ttgatagaat gcgtaagtac agtagtgagt tattcaagac ttctattgta 240
ggcgaatcca cagttgtttg ctgtggggca gctagtaaca aattcctatt ctctaacgaa 300
aacaaactgg taactgcctg gtggccagat tctgttaaca aaatcttccc aacaacttca 360
ctggattcta atttgaagga ggaatctata aagatgagaa agttgctgcc acagttcttc 420
aaaccagaag cacttcaaag atacgtcggc gttatggatg taatcgcaca aagacatttt 480
gtcactcact gggacaacaa aaatgagatc acagtttatc cacttgctaa aagatacact 540
ttcttgcttg cgtgtagact gttcatgtct gttgaggatg aaaatcatgt ggcgaaattc 600
tcagacccat tccaactaat cgctgcaggc atcatttcac ttcctatcga tcttcctggt 660
actccattca acaaggccat aaaggcttca aatttcatta gaaaagagct gataaagatt 720
atcaaacaaa gacgtgttga tctggcagag ggtacagcat ctccaaccca ggatatcttg 780
tcacatatgc tattaacatc tgatgaaaac ggtaaatcta tgaacgagtt gaacattgcc 840
gacaagattc ttggactatt gataggaggc cacgatacag cttcagtagc ttgcacattt 900
ctagtgaagt acttaggaga attaccacat atctacgata aagtctacca agagcaaatg 960
gaaattgcca agtccaaacc tgctggggaa ttgttgaatt gggatgactt gaaaaagatg 1020
aagtattcat ggaatgtggc atgtgaggta atgagattgt caccaccttt acaaggtggt 1080
tttagagagg ctataactga ctttatgttt aacggtttct ctattccaaa agggtggaag 1140
ttatactggt ccgccaactc tacacacaaa aatgcagaat gtttcccaat gcctgagaaa 1200
ttcgatccta ccagatttga aggtaatggt ccagcgcctt atacatttgt accattcggt 1260
ggaggcccta gaatgtgtcc tggaaaggaa tacgctagat tagaaatctt ggttttcatg 1320
cataatctgg tcaaacgttt taagtgggaa aaggttattc cagacgaaaa gattattgtc 1380
gatccattcc caatcccagc taaagatctt ccaatccgtt tgtatcctca caaagcttaa 1440

SEQ ID NO:114

MEPNFYLSLL LLFVTFISLS LFFIFYKQKS PLNLPPGKMG YPIIGESLEF LSTGWKGHPE 60

```
KFIFDRMRKY SSELFKTSIV GESTVVCCGA ASNKFLFSNE NKLVTAWWPD SVNKIFPTTS       120
LDSNLKEESI KMRKLLPQFF KPEALQRYVG VMDVIAQRHF VTHWDNKNEI TVYPLAKRYT       180
FLLACRLFMS VEDENHVAKF SDPFQLIAAG IISLPIDLPG TPFNKAIKAS NFIRKELIKI       240
IKQRRVDLAE GTASPTQDIL SHMLLTSDEN GKSMNELNIA DKILGLLIGG HDTASVACTF       300
LVKYLGELPH IYDKVYQEQM EIAKSKPAGE LLNWDDLKKM KYSWNVACEV MRLSPPLQGG       360
FREAITDFMF NGFSIPKGWK LYWSANSTHK NAECFPMPEK FDPTRFEGNG PAPYTFVPFG       420
GGPRMCPGKE YARLEILVFM HNLVKRFKWE KVIPDEKIIV DPFPIPAKDL PIRLYPHKA        479
```

SEQ ID NO:115

```
atggcctctg ttactttggg ttcctggatc gtcgtccacc accataacca tcaccatcca        60
tcatctatcc taactaaatc tcgttcaaga tcctgtccta ttacactaac caaaccaatc       120
tcttttcgtt caaagagaac agtttcctct agtagttcta tcgtgtcctc tagtgtcgtc       180
actaaggaag acaatctgag acagtctgaa ccttcttcct ttgatttcat gtcatatatc       240
attactaagg cagaactagt gaataaggct cttgattcag cagttccatt aagagagcca       300
ttgaaaatcc atgaagcaat gagatactct cttctagctg gcgggaagag agtcagacct       360
gtactctgca tagcagcgtg cgaattagtt ggtggcgagg aatcaaccgc tatgcctgcc       420
gcttgtgctg tagaaatgat tcatacaatg tcactgatac acgatgattt gccatgtatg       480
gataacgatg atctgagaag gggtaagcca actaaccata aggttttcgg cgaagatgtt       540
gccgtcttag ctggtgatgc tttgttatct ttcgcgttcg aacatttggc atccgcaaca       600
tcaagtgatg ttgtgtcacc agtaagagta gttagagcag ttggagaact ggctaaagct       660
attggaactg agggtttagt tgcaggtcaa gtcgtcgata tctcttccga aggtcttgat       720
ttgaatgatg taggtcttga acatctcgaa ttcatccatc ttcacaagac agctgcactt       780
ttagaagcca gtgcggttct cggcgcaatt gttggcggag ggagtgatga cgaaattgag       840
agattgagga agtttgctag atgtatagga ttactgttcc aagtagtaga cgatatacta       900
gatgtgacaa agtcttccaa agagttggga aaaacagctg gtaaagattt gattgccgac       960
aaattgacct accctaagat tatggggcta gaaaaatcaa gagaatttgc cgagaaactc      1020
aatagagagg cgcgtgatca actgttgggt ttcgattctg ataaagttgc accactctta      1080
gccttagcca actacatcgc ttacagacaa aactaa                                 1116
```

SEQ ID NO:116

```
MASVTLGSWI VVHHHNHHHP SSILTKSRSR SCPITLTKPI SFRSKRTVSS SSSIVSSSVV        60
TKEDNLRQSE PSSFDFMSYI ITKAELVNKA LDSAVPLREP LKIHEAMRYS LLAGGKRVRP       120
VLCIAACELV GGEESTAMPA ACAVEMIHTM SLIHDDLPCM DNDDLRRGKP TNHKVFGEDV       180
AVLAGDALLS FAFEHLASAT SSDVVSPVRV VRAVGELAKA IGTEGLVAGQ VVDISSEGLD       240
LNDVGLEHLE FIHLHKTAAL LEASAVLGAI VGGGSDDEIE RLRKFARCIG LLFQVVDDIL       300
DVTKSSKELG KTAGKDLIAD KLTYPKIMGL EKSREFAEKL NREARDQLLG FDSDKVAPLL       360
ALANYIAYRQ N                                                            371
```

SEQ ID NO:117

R. suavissimus

```
MATLLEHFQA MPFAIPIALA ALSWLFLFYI KVSFFSNKSA QAKLPPVPVV PGLPVIGNLL        60
QLKEKKPYQT FTRWAEEYGP IYSIRTGAST MVVLNTTQVA KEAMVTRYLS ISTRKLSNAL       120
KILTADKCMV AISDYNDFHK MIKRYILSNV LGPSAQKRHR SNRDTLRANV CSRLHSQVKN       180
SPREAVNFRR VFEWELFGIA LKQAFGKDIE KPIYVEELGT TLSRDEIFKV LVLDIMEGAI       240
EVDWRDFFPY LRWIPNTRME TKIQRLYFRR KAVMTALINE QKKRIASGEE INCYIDFLLK       300
EGKTLTMDQI SMLLWETVIE TADTTMVTTE WAMYEVAKDS KRQDRLYQEI QKVCGSEMVT       360
EEYLSQLPYL NAVFHETLRK HSPAALVPLR YAHEDTQLGG YYIPAGTEIA INIYGCNMDK       420
HQWESPEEWK PERFLDPKFD PMDLYKTMAF GAGKRVCAGS LQAMLIACPT IGRLVQEFEW       480
KLRDGEEENV DTVGLTTHKR YPMHAILKPR S                                      511
```

SEQ ID NO:118

S. cerevisiae

```
atgtcatttc aaattgaaac ggttcccacc aaaccatatg aagaccaaaa gcctggtacc        60
```

```
tctggtttgc gtaagaagac aaaggtgttt aaagacgaac ctaactacac agaaaatttc       120
attcaatcga tcatggaagc tattccagag ggttctaaag gtgccactct tgttgtcggt       180
ggtgatgggc gttactacaa tgatgtcatt cttcataaga ttgccgctat cggtgctgcc       240
aacggtatta aaaagttagt tattggccag catggtcttc tgtctacgcc agccgcttct       300
cacatcatga gaacctacga ggaaaaatgt actggtggta ttatcttaac cgcctcacat       360
aatccaggtg gtccagaaaa tgacatgggt attaagtata acttatccaa tgggggtcct       420
gctcctgaat ccgtcacaaa tgctatttgg gagatttcca aaaagcttac cagctataag       480
attatcaaag acttcccaga actagacttg ggtacgatag gcaagaacaa gaaatacggt       540
ccattactcg ttgacattat cgatattaca aaagattatg tcaacttctt gaaggaaatc       600
ttcgatttcg acttaatcaa gaaattcatc gataatcaac gttctactaa gaattggaag       660
ttactgtttg acagtatgaa cggtgtaact ggaccatacg gtaaggctat tttcgttgat       720
gaatttggtt taccggcgga tgaggtttta caaaactggc atccttctcc ggattttggt       780
ggtatgcatc cagatccaaa cttaacttat gccagttcgt tagtgaaaag agtagatcgt       840
gaaaagattg agtttggtgc tgcatccgat ggtgatggtg atagaaatat gatttacggt       900
tacggcccat ctttcgtttc tccaggtgac tccgtcgcaa ttattgccga atatgcagct       960
gaaatcccat atttcgccaa gcaaggtata tatggtctgg cccgttcatt ccctacctca      1020
ggagccatag accgtgttgc caaggcccat ggtctaaact gttatgaggt cccaactggc      1080
tggaaatttt tctgtgcttt gttcgacgct aaaaaattat ctatttgtgg tgaagaatcg      1140
tttggtactg gttccaacca cgtaagggaa aaggacggtg tttgggccat tatggcgtgg      1200
ttgaacatct tggccattta caacaagcat catccggaga acgaagcttc tattaagacg      1260
atacagaatg aattctgggc aaagtacggc cgtactttct tcactcgtta tgattttgaa      1320
aaagttgaaa cagaaaaagc taacaagatt gtcgatcaat tgagagcata tgttaccaaa      1380
tcgggtgttg ttaattccgc cttcccagcc gatgagtctc ttaaggtcac cgattgtggt      1440
gatttttcat acacagattt ggacggttct gtttctgacc atcaaggttt atatgtcaag      1500
ctttccaatg gtgcaagatt cgttctaaga ttgtcaggta caggttcttc aggtgctacc      1560
attagattgt acattgaaaa atactgcgat gataaatcac aataccaaaa gacagctgaa      1620
gaatacttga agccaattat taactcggtc atcaagttct tgaactttaa acaagtttta      1680
ggaactgaag aaccaacggt tcgtacttaa                                       1710
```

SEQ ID NO:119

S. cerevisiae

```
MSFQIETVPT KPYEDQKPGT SGLRKKTKVF KDEPNYTENF IQSIMEAIPE GSKGATLVVG        60
GDGRYYNDVI LHKIAAIGAA NGIKKLVIGQ HGLLSTPAAS HIMRTYEEKC TGGIILTASH       120
NPGGPENDMG IKYNLSNGGP APESVTNAIW EISKKLTSYK IIKDFPELDL GTIGKNKKYG       180
PLLVDIIDIT KDYVNFLKEI FDFDLIKKFI DNQRSTKNWK LLFDSMNGVT GPYGKAIFVD       240
EFGLPADEVL QNWHPSPDFG GMHPDPNLTY ASSLVKRVDR EKIEFGAASD GDGDRNMIYG       300
YGPSFVSPGD SVAIIAEYAA EIPYFAKQGI YGLARSFPTS GAIDRVAKAH GLNCYEVPTG       360
WKFFCALFDA KKLSICGEES FGTGSNHVRE KDGVWAIMAW LNILAIYNKH HPENEASIKT       420
IQNEFWAKYG RTFFTRYDFE KVETEKANKI VDQLRAYVTK SGVVNSAFPA DESLKVTDCG       480
DFSYTDLDGS VSDHQGLYVK LSNGARFVLR LSGTGSSGAT IRLYIEKYCD DKSQYQKTAE       540
EYLKPIINSV IKFLNFKQVL GTEEPTVRT                                         569
```

SEQ ID NO:120

S. cerevisiae

```
atgtccacta agaagcacac caaaacacat tccacttatg cattcgagag caacacaaac        60
agcgttgctg cctcacaaat gagaaacgcc ttaaacaagt tggcggactc tagtaaactt       120
gacgatgctg ctcgcgctaa gtttgagaac gaactggatt cgtttttcac gcttttcagg       180
agatatttgg tagagaagtc ttctagaacc accttggaat gggacaagat caagtctccc       240
aacccggatg aagtggttaa gtatgaaatt atttctcagc agcccgagaa tgtctcaaac       300
ctttccaaat tggctgtttt gaagttgaac ggtgggctgg gtacctccat gggctgcgtt       360
ggccctaaat ctgttattga agtgagagag ggaaacacct ttttggattt gtctgttcgt       420
caaattgaat acttgaacag acagtacgat agcgacgtgc cattgttatt gatgaattct       480
ttcaacactg acaaggatac ggaacacttg attaagaagt attccgctaa cagaatcaga       540
atcagatctt tcaatcaatc caggttccca agagtctaca aggattcttt attgcctgtc       600
```

```
cccaccgaat acgattctcc actggatgct tggtatccac caggtcacgg tgatttgttt       660
gaatctttac acgtatctgg tgaactggat gccttaattg cccaaggaag agaaatatta       720
tttgtttcta acggtgacaa cttgggtgct accgtcgact taaaaatttt aaaccacatg       780
atcgagactg gtgccgaata tataatggaa ttgactgata agaccagagc cgatgttaaa       840
ggtggtactt tgatttctta cgatggtcaa gtccgtttat tggaagtcgc ccaagttcca       900
aaagaacaca ttgacgaatt caaaaatatc agaaagttta ccaacttcaa cacgaataac       960
ttatggatca atctgaaagc agtaaagagg ttgatcgaat cgagcaattt ggagatggaa      1020
atcattccaa accaaaaaac tataacaaga gacggtcatg aaattaatgt cttacaatta      1080
gaaaccgctt gtggtgctgc tatcaggcat tttgatggtg ctcacggtgt tgtcgttcca      1140
agatcaagat tcttgcctgt caagacctgt tccgatttgt tgctggttaa atcagatcta      1200
ttccgtctgg aacacggttc tttgaagtta gacccatccc gttttggtcc aaacccatta      1260
atcaagttgg gctcgcattt caaaaaggtt tctggtttta acgcaagaat ccctcacatc      1320
ccaaaaatcg tcgagctaga tcatttgacc atcactggta acgtcttttt aggtaaagat      1380
gtcactttga ggggtactgt catcatcgtt tgctccgacg gtcataaaat cgatattcca      1440
aacggctcca tattggaaaa tgttgtcgtt actggtaatt tgcaaatctt ggaacattga      1500
```

SEQ ID NO:121
S. cerevisiae

```
MSTKKHTKTH STYAFESNTN SVAASQMRNA LNKLADSSKL DDAARAKFEN ELDSFFTLFR        60
RYLVEKSSRT TLEWDKIKSP NPDEVVKYEI ISQQPENVSN LSKLAVLKLN GGLGTSMGCV       120
GPKSVIEVRE GNTFLDLSVR QIEYLNRQYD SDVPLLLMNS FNTDKDTEHL IKKYSANRIR       180
IRSFNQSRFP RVYKDSLLPV PTEYDSPLDA WYPPGHGDLF ESLHVSGELD ALIAQGREIL       240
FVSNGDNLGA TVDLKILNHM IETGAEYIME LTDKTRADVK GGTLISYDGQ VRLLEVAQVP       300
KEHIDEFKNI RKFTNFNTNN LWINLKAVKR LIESSNLEME IIPNQKTITR DGHEINVLQL       360
ETACGAAIRH FDGAHGVVVP RSRFLPVKTC SDLLLVKSDL FRLEHGSLKL DPSRFGPNPL       420
IKLGSHFKKV SGFNARIPHI PKIVELDHLT ITGNVFLGKD VTLRGTVIIV CSDGHKIDIP       480
NGSILENVVV TGNLQILEH                                                    499
```

SEQ ID NO:122
S. cerevisiae

```
atgtctagtc aaacagaaag aacttttatt gcggtaaaac cagatggtgt ccagaggggc        60
ttagtatctc aaattctatc tcgttttgaa aaaaaaggtt acaaactagt tgctattaaa       120
ttagttaaag cggatgataa attactagag caacattacg cagagcatgt tggtaaacca       180
tttttcccaa agatggtatc ctttatgaag tctggtccca ttttggccac ggtctgggag       240
ggaaaagatg tggttagaca aggaagaact attcttggtg ctactaatcc tttgggcagt       300
gcaccaggta ccattagagg tgatttcggt attgacctag gcagaaacgt ctgtcacggc       360
agtgattctg ttgatagcgc tgaacgtgaa atcaatttgt ggtttaagaa ggaagagtta       420
gttgattggg aatctaatca agctaagtgg atttatgaat ga                          462
```

SEQ ID NO:123
S. cerevisiae

```
MSSQTERTFI AVKPDGVQRG LVSQILSRFE KKGYKLVAIK LVKADDKLLE QHYAEHVGKP        60
FFPKMVSFMK SGPILATVWE GKDVVRQGRT ILGATNPLGS APGTIRGDFG IDLGRNVCHG       120
SDSVDSAERE INLWFKKEEL VDWESNQAKW IYE                                    153
```

SEQ ID NO:124
S. rebaudiana

```
atggctgctg ctgatactga aaagttgaac aatttgagat ccgccgtttc tggtttgacc        60
caaatttctg ataacgaaaa gtccggtttc atcaacttgg tcagtagata tttgtctggt       120
gaagctcaac acgttgaatg gtctaaaatt caaactccaa ccgataagat cgttgttcca       180
tacgatactt tgtctgctgt tccagaagat gctgctcaaa caaaatcttt gttggataag       240
ttggtcgtct tgaagttgaa cggtggtttg ggtactacta tgggttgtac tggtccaaag       300
tctgttatcg aagttagaaa cggtttgacc ttcttggatt tgatcgtcat ccaaatcgaa       360
```

```
tccttgaaca agaagtacgg ttgttctgtt cctttgttgt tgatgaactc tttcaacacc       420
catgaagata cccaaaagat cgtcgaaaag tactccggtt ctaacattga agttcacacc       480
ttcaatcaat cccaataccc aagattggtt gtcgatgaat ttttgccatt gccatctaaa       540
ggtgaaactg gtaaagatgg ttggtatcca ccaggtcatg gtgatgtttt tccatccttg       600
atgaattccg gtaagttgga tgctttgttg tcccaaggta aagaatacgt tttcgttgcc       660
aactctgata acttgggtgc agttgttgat ttgaagatct tgaaccactt gatccaaaac       720
aagaacgaat actgcatgga agttactcca aagactttgg ctgatgttaa gggtggtact       780
ttgatttctt acgatggtaa ggttcaatta ttggaaatcg cccaagttcc agatgaacac       840
gttaatgaat tcaagtccat cgaaaagttt aagatcttta acactaacaa cttgtgggtc       900
aacttgaacg ccattaagag attggttcaa gctgatgctt tgaagatgga aattattcca       960
aatccaaaag aagtcaacgg tgtcaaggta ttgcaattgg aaactgctgc tggtgctgct      1020
attaagtttt tcgataatgc catcggtatc aacgtcccaa gatctagatt tttgcctgtt      1080
aaggcttcct ctgacttgtt gttagttcaa tcagacttgt acaccgaaaa ggatggttac      1140
gttattagaa acccagctag aaaggatcca gctaacccat ctattgaatt gggtccagaa      1200
ttcaaaaagg tcggtgattt cttgaagaga ttcaagtcta tcccatccat catcgaattg      1260
gactcattga aagtttctgg tgatgtctgg tttggttcca acgttgtttt gaaaggtaag      1320
gttgttgttg ctgccaaatc cggtgaaaaa ttggaaattc cagatggtgc cttgattgaa      1380
aacaaagaag ttcatggtgc ctccgacatt tga                                   1413
```

SEQ ID NO:125

S. rebaudiana

```
MAAADTEKLN NLRSAVSGLT QISDNEKSGF INLVSRYLSG EAQHVEWSKI QTPTDKIVVP        60
YDTLSAVPED AAQTKSLLDK LVVLKLNGGL GTTMGCTGPK SVIEVRNGLT FLDLIVIQIE       120
SLNKKYGCSV PLLLMNSFNT HEDTQKIVEK YSGSNIEVHT FNQSQYPRLV VDEFLPLPSK       180
GETGKDGWYP PGHGDVFPSL MNSGKLDALL SQGKEYVFVA NSDNLGAVVD LKILNHLIQN       240
KNEYCMEVTP KTLADVKGGT LISYDGKVQL LEIAQVPDEH VNEFKSIEKF KIFNTNNLWV       300
NLNAIKRLVQ ADALKMEIIP NPKEVNGVKV LQLETAAGAA IKFFDNAIGI NVPRSRFLPV       360
KASSDLLLVQ SDLYTEKDGY VIRNPARKDP ANPSIELGPE FKKVGDFLKR FKSIPSIIEL       420
DSLKVSGDVW FGSNVVLKGK VVVAAKSGEK LEIPDGALIE NKEVHGASDI                  470
```

SEQ ID NO:126

A. pullulans

```
atgtcctctg aaatggctac tcatttgaaa cctaatggtg gtgccgaatt cgaaaaaaga        60
catcatggta agacccaatc ccatgttgct tttgaaaaca cttctacatc tgttgctgcc       120
tcccaaatga gaaatgcttt gaatactttg tgcgattccg ttactgatcc agctgaaaag       180
caaagattcg aaaccgaaat ggataacttc ttcgccttgt ttagaagata cttgaacgat       240
aaggctaagg gtaacgaaat cgaatggtct agaattgctc caccaaaacc agaacaagtt       300
gttgcttatc aagacttgcc tgaacaagaa tccgttgaat tcttgaacaa attggccgtc       360
ttgaagttga atggtggttt gggtacttct atgggttgtg ttggtccaaa gtctgttatc       420
gaagttagag atggtatgtc cttcttggat ttgtccgtta gacaaatcga atacttgaat       480
agaacctacg gtgttaacgt tccattcgtc ttgatgaatt ctttcaacac tgatgctgat       540
accgccaaca ttatcaaaaa gtacgaaggt cacaacatcg acatcatgac cttcaatcaa       600
tctagatacc caagaatctt gaaggattct ttgttgccag ctccaaaatc tgccaactct       660
caaatttctg attggtatcc accaggtcat ggtgacgttt ttgaatcctt gtacaactct       720
ggtatcttgg ataagttgtt ggaaagaggt gtcgaaatcg ttttcttgtc caatgctgat       780
aatttgggtg ccgttgttga tttgaagatc ttgcaacata tggttgatac caaggccgaa       840
tatatcatgg aattgactga taagactaag gccgatgtta agggtggtac tattattgac       900
tatgaaggtc aagccagatt attggaaatt gcccaagttc caaaagaaca cgtcaacgaa       960
ttcaagtcca tcaagaagtt taagtacttc aacaccaaca acatctggat gaacttgaga      1020
gctgttaaga gaatcgtcga aaacaacgaa ttggccatgg aaattatccc aaacggtaaa      1080
tctattccag ccgacaaaaa aggtgaagcc gatgtttcta tagttcaatt ggaaactgct      1140
gttggtgctg ccattagaca ttttaacaat gctcatggtg tcaacgtccc aagaagaaga      1200
tttttgccag ttaagacctg ctccgatttg atgttggtta agtctgactt gtacactttg      1260
aagcacggtc aattgattat ggacccaaat agatttggtc cagccccatt gattaagttg      1320
```

```
ggtggtgatt ttaagaaggt ttcctcattc caatccagaa tcccatccat tcctaaaatc      1380
ttggaattgg atcatttgac cattaccggt ccagttaact tgggtagagg tgttactttt      1440
aagggtactg ttattatcgt tgcctccgaa ggtcaaacca ttgatattcc acctggttcc      1500
attttggaaa acgttgttgt tcaaggttcc ttgagattat tagaacatta a               1551
```

SEQ ID NO:127

A. pullulans

```
MSSEMATHLK PNGGAEFEKR HHGKTQSHVA FENTSTSVAA SQMRNALNTL CDSVTDPAEK        60
QRFETEMDNF FALFRRYLND KAKGNEIEWS RIAPPKPEQV VAYQDLPEQE SVEFLNKLAV       120
LKLNGGLGTS MGCVGPKSVI EVRDGMSFLD LSVRQIEYLN RTYGVNVPFV LMNSFNTDAD       180
TANIIKKYEG HNIDIMTFNQ SRYPRILKDS LLPAPKSANS QISDWYPPGH GDVFESLYNS       240
GILDKLLERG VEIVFLSNAD NLGAVVDLKI LQHMVDTKAE YIMELTDKTK ADVKGGTIID       300
YEGQARLLEI AQVPKEHVNE FKSIKKFKYF NTNNIWMNLR AVKRIVENNE LAMEIIPNGK       360
SIPADKKGEA DVSIVQLETA VGAAIRHFNN AHGVNVPRRR FLPVKTCSDL MLVKSDLYTL       420
KHGQLIMDPN RFGPAPLIKL GGDFKKVSSF QSRIPSIPKI LELDHLTITG PVNLGRGVTF       480
KGTVIIVASE GQTIDIPPGS ILENVVVQGS LRLLEH                                 516
```

SEQ ID NO:128

A. thaliana

```
atggctgcta ctactgaaaa cttgccacaa ttgaaatctg ccgttgatgg tttgactgaa        60
atgtccgaat ctgaaaagtc cggtttcatc tctttggtca gtagatattt gtctggtgaa       120
gcccaacata tcgaatggtc taaaattcaa actccaaccg acgaaatcgt tgtcccatac       180
gaaaaaatga ctccagtttc tcaagatgtc gccgaaacta agaatttgtt ggataagttg       240
gtcgtcttga agttgaatgg tggtttgggt actactatgg gttgtactgg tccaaagtct       300
gttatcgaag ttagagatgg tttaaccttc ttggacttga tcgtcatcca aatcgaaaac       360
ttgaacaaca agtacggttg caaggttcca ttggtcttga tgaattcttt caacacccat       420
gatgataccc acaagatcgt tgaaaagtac accaactcca acgttgatat ccacaccttc       480
aatcaatcta agtacccaag agttgttgcc gatgaatttg ttccatggcc atctaaaggt       540
aagactgaca aagaaggttg gtatccacca ggtcatggtg atgtttttcc agctttaatg       600
aactccggta agttggatac tttcttgtcc caaggtaaag aatacgtttt cgttgccaac       660
tctgataact tgggtgctat agttgatttg accatcttga agcacttgat ccaaaacaag       720
aacgaatact gcatggaagt tactccaaag actttggctg atgttaaggg tggtactttg       780
atttcttacg aaggtaaggt tcaattattg gaaatcgccc aagttccaga tgaacacgtt       840
aatgaattca agtccatcga aaagttcaag atcttcaaca ccaacaactt gtgggttaac       900
ttgaaggcca tcaagaaatt ggttgaagct gatgctttga agatggaaat tatcccaaac       960
ccaaaagaag ttgacggtgt taaggtattg caattggaaa ctgctgctgg tgctgctatt      1020
agatttttcg ataatgccat cggtgttaac gtcccaagat ctagattttt gccagttaag      1080
gcttcctccg atttgttgtt ggttcaatct gacttgtaca ccttggttga cggttttgtt      1140
acaagaaaca aggctagaac taacccatcc aacccatcta ttgaattggg tccagaattc      1200
aaaaaggttg ccacattctt gtccagattc aagtctattc catccatcgt cgaattggac      1260
tcattgaaag tttctggtga tgtctggttt ggttcctcta tagttttgaa gggtaaggtt      1320
actgttgctg ctaaatctgg tgttaagttg gaaattccag atagagccgt tgtcgaaaac      1380
aaaaacatta acggtcctga agatttgtga                                       1410
```

SEQ ID NO:129

A. thaliana

```
MAATTENLPQ LKSAVDGLTE MSESEKSGFI SLVSRYLSGE AQHIEWSKIQ TPTDEIVVPY        60
EKMTPVSQDV AETKNLLDKL VVLKLNGGLG TTMGCTGPKS VIEVRDGLTF LDLIVIQIEN       120
LNNKYGCKVP LVLMNSFNTH DDTHKIVEKY TNSNVDIHTF NQSKYPRVVA DEFVPWPSKG       180
KTDKEGWYPP GHGDVFPALM NSGKLDTFLS QGKEYVFVAN SDNLGAIVDL TILKHLIQNK       240
NEYCMEVTPK TLADVKGGTL ISYEGKVQLL EIAQVPDEHV NEFKSIEKFK IFNTNNLWVN       300
LKAIKKLVEA DALKMEIIPN PKEVDGVKVL QLETAAGAAI RFFDNAIGVN VPRSRFLPVK       360
ASSDLLLVQS DLYTLVDGFV TRNKARTNPS NPSIELGPEF KKVATFLSRF KSIPSIVELD       420
```

SLKVSGDVWF GSSIVLKGKV TVAAKSGVKL EIPDRAVVEN KNINGPEDL 469

SEQ ID NO:130
E. coli

atggctgcta ttaacaccaa ggttaagaag gctgttattc cagttgctgg tttgggtact 60
agaatgttgc cagctacaaa agccattcca aaagaaatgt taccattggt cgataagcca 120
ttgatccaat acgttgtcaa cgaatgtatt gctgctggta ttaccgaaat cgttttggtt 180
actcactcct ccaagaactc cattgaaaat catttcgaca cctcattcga attggaagcc 240
atgttggaaa agagagtcaa gagacaatta ttggacgaag tccaatctat ttgcccacca 300
catgttacta tcatgcaagt tagacaaggt ttggctaaag gtttgggtca tgctgttttg 360
tgtgctcatc cagttgttgg tgatgaacca gttgcagtta ttttgccaga tgttatcttg 420
gacgaatacg aatccgattt gtctcaagat aacttggctg aaatgatcag aagattcgac 480
gaaactggtc actcccaaat tatggttgaa cctgttgctg atgttactgc ttatggtgtt 540
gttgattgca agggtgttga attggctcca ggtgaatctg ttccaatggt tggtgttgta 600
gaaaagccaa aagctgatgt tgctccatct aatttggcta tcgttggtag atatgttttg 660
tccgctgata tttggccttt gttggctaaa actccaccag gtgctggtga cgaaattcaa 720
ttgactgatg ctatcgacat gttgatcgaa aaagaaaccg ttgaagccta ccacatgaag 780
ggtaaatctc atgattgtgg taacaagttg ggttacatgc aagcttttgt tgaatacggt 840
atcagacata acaccttagg tactgaattc aaggcttggt tggaagaaga aatgggtatc 900
aagaagtaa 909

SEQ ID NO:131
E. coli

MAAINTKVKK AVIPVAGLGT RMLPATKAIP KEMLPLVDKP LIQYVVNECI AAGITEIVLV 60
THSSKNSIEN HFDTSFELEA MLEKRVKRQL LDEVQSICPP HVTIMQVRQG LAKGLGHAVL 120
CAHPVVGDEP VAVILPDVIL DEYESDLSQD NLAEMIRRFD ETGHSQIMVE PVADVTAYGV 180
VDCKGVELAP GESVPMVGVV EKPKADVAPS NLAIVGRYVL SADIWPLLAK TPPGAGDEIQ 240
LTDAIDMLIE KETVEAYHMK GKSHDCGNKL GYMQAFVEYG IRHNTLGTEF KAWLEEEMGI 300
KK 302

SEQ ID NO:132
R. suavissimus

atggctgctg ttgctactga taagatctct aagttgaagt ctgaagttgc tgccttgtcc 60
caaatttctg aaaacgaaaa gtccggtttc atcaacttgg tcagtagata tttgtctggt 120
actgaagcta ctcacgttga atggtctaaa attcaaactc caaccgatga agttgttgtt 180
ccatatgata ctttggctcc aactccagaa gatccagctg aaactaagaa gttgttagat 240
aagttggtcg tcttgaagtt gaacggtggt ttgggtacta ctatgggttg tactggtcca 300
aagtctgtta tcgaagttag aaacggtttg accttcttgg atttgatcgt cattcaaatc 360
gaaaccttga acaacaagta cggttgtaac gttcctttgt tgttgatgaa ctctttcaac 420
acccatgatg acaccttcaa gatcgttgaa agatacacca agtccaacgt tcaaatccat 480
accttcaatc aatcccaata cccaagattg gttgtcgaag ataattctcc attgccatct 540
aagggtcaaa ctggtaaaga tggttggtat ccaccaggtc atggtgatgt ttttccatct 600
ttgagaaact ccggtaagtt ggatttgttg ttatcccaag gtaaagaata cgttttcatc 660
tccaactctg ataacttggg tgcagttgtt gatttgaaga tcttgtccca tttggtccaa 720
aaaaagaacg aatactgcat ggaagttacc ccaaaaactt tggctgatgt taagggtggt 780
actttgattt cttacgaagg tagaacccaa ttattggaaa ttgcccaagt tccagatcaa 840
cacgttaacg aattcaagtc catcgaaaag ttcaagatct ttaacaccaa caatttgtgg 900
gtcaacttga acgccattaa gagattagtt gaagctgatg ccttgaaaat ggaaatcatc 960
ccaaatccaa aagaagtcga cggtattaag gtcttgcaat tggaaactgc tgctggtgct 1020
gctattagat ttttcaatca tgccatcggt atcaacgtcc caagatctag atttttgcca 1080
gttaaggcta cctccgattt gttattggtt caatctgact tgtacaccgt cgaagatggt 1140
ttcgttatta gaaacactgc tagaaagaat ccagccaacc catctgttga attgggtcca 1200
gaattcaaaa aggttgccaa cttcttgtcc agattcaagt ctattccatc catcatcgaa 1260

```
ttggactcat tgaaggttgt tggtgatgta tggtttggtg ctggtgttgt tttgaaaggt      1320
aaggttacta ttactgctaa gccaggtgtt aagttggaaa ttccagataa ggctgtcttg      1380
gaaaacaagg atattaacgg tcctgaagat ttgtga                                1416
```

SEQ ID NO:133

R. suavissimus

```
MAAVATDKIS KLKSEVAALS QISENEKSGF INLVSRYLSG TEATHVEWSK IQTPTDEVVV        60
PYDTLAPTPE DPAETKKLLD KLVVLKLNGG LGTTMGCTGP KSVIEVRNGL TFLDLIVIQI       120
ETLNNKYGCN VPLLLMNSFN THDDTFKIVE RYTKSNVQIH TFNQSQYPRL VVEDNSPLPS       180
KGQTGKDGWY PPGHGDVFPS LRNSGKLDLL LSQGKEYVFI SNSDNLGAVV DLKILSHLVQ       240
KKNEYCMEVT PKTLADVKGG TLISYEGRTQ LLEIAQVPDQ HVNEFKSIEK FKIFNTNNLW       300
VNLNAIKRLV EADALKMEII PNPKEVDGIK VLQLETAAGA AIRFFNHAIG INVPRSRFLP       360
VKATSDLLLV QSDLYTVEDG FVIRNTARKN PANPSVELGP EFKKVANFLS RFKSIPSIIE       420
LDSLKVVGDV WFGAGVVLKG KVTITAKPGV KLEIPDKAVL ENKDINGPED L                471
```

SEQ ID NO:134

H. vulgare

```
atggctgctg ctgcagttgc tgctgattct aaaattgatg gtttgagaga tgctgttgcc        60
aagttgggtg aaatttctga aaacgaaaag gccggtttca tctccttggt ttctagatat       120
ttgtctggtg aagccgaaca aatcgaatgg tctaaaattc aaactccaac cgatgaagtt       180
gttgttccat atgatacttt ggctccacca cctgaagatt tggatgctat gaaggctttg       240
ttggataagt tggttgtctt gaagttgaat ggtggtttgg gtactactat gggttgtact       300
ggtccaaagt ctgttatcga agttagaaac ggtttcacct tcttggattt gatcgttatc       360
caaattgaat ccttgaacaa gaagtacggt tgctctgttc ctttgttgtt gatgaactct       420
ttcaacaccc atgatgacac ccaaaagatc gttgaaaagt actccaactc caacatcgaa       480
atccacacct tcaatcaatc tcaataccca agaatcgtca ccgaagattt tttgccattg       540
ccatctaaag gtcaaactgg taaagatggt tggtatccac caggtcatgg tgatgttttt       600
ccatctttga acaactccgg taagttggat accttgttgt ctcaaggtaa agaatacgtt       660
ttcgttgcca actctgataa cttgggtgct atcgttgata ttaagatctt gaaccacttg       720
atccacaatc aaaacgaata ctgcatggaa gttactccaa agactttggc tgatgttaag       780
ggtggtactt tgatttctta cgaaggtaga gttcaattat tggaaatcgc ccaagttcca       840
gatgaacacg ttgatgaatt caagtccatc gaaaagttca aaatcttcaa caccaacaac       900
ttgtgggtta acttgaaggc cattaagaga ttggttgatg ctgaagcttt gaaaatggaa       960
atcatcccaa accctaaaga agttgacggt gttaaggtat tgcaattgga aactgctgct      1020
ggtgctgcta ttagattctt tgaaaaagcc atcggtatca acgtcccaag atctagattt      1080
ttgccagtta aggctacctc tgacttgttg ttggttcaat cagacttgta caccttggtt      1140
gacggttacg ttattagaaa tccagctaga gttaagccat ccaacccatc tattgaattg      1200
ggtccagaat tcaagaaggt cgctaatttc ttggctagat tcaagtctat cccatccatc      1260
gttgaattgg actcattgaa agtttctggt gatgtctctt ttggttccgg tgttgttttg      1320
aagggtaatg ttactattgc tgctaaggct ggtgttaagt tggaaattcc agatggtgct      1380
gttttggaaa acaaggatat taacggtcca gaagatattt ga                         1422
```

SEQ ID NO:135

H. vulgare

```
MAAAAVAADS KIDGLRDAVA KLGEISENEK AGFISLVSRY LSGEAEQIEW SKIQTPTDEV        60
VVPYDTLAPP PEDLDAMKAL LDKLVVLKLN GGLGTTMGCT GPKSVIEVRN GFTFLDLIVI       120
QIESLNKKYG CSVPLLLMNS FNTHDDTQKI VEKYSNSNIE IHTFNQSQYP RIVTEDFLPL       180
PSKGQTGKDG WYPPGHGDVF PSLNNSGKLD TLLSQGKEYV FVANSDNLGA IVDIKILNHL       240
IHNQNEYCME VTPKTLADVK GGTLISYEGR VQLLEIAQVP DEHVDEFKSI EKFKIFNTNN       300
LWVNLKAIKR LVDAEALKME IIPNPKEVDG VKVLQLETAA GAAIRFFEKA IGINVPRSRF       360
LPVKATSDLL LVQSDLYTLV DGYVIRNPAR VKPSNPSIEL GPEFKKVANF LARFKSIPSI       420
VELDSLKVSG DVSFGSGVVL KGNVTIAAKA GVKLEIPDGA VLENKDINGP EDI              473
```

SEQ ID NO:136
O. sativa

```
atggctgacg aaaaattggc caaattgaga gaagctgttg ctggtttgtc tcaaatctct        60
gataacgaaa agtccggttt catttccttg gttgctagat atttgtccgg tgaagaagaa       120
catgttgaat gggctaaaat tcatacccca accgatgaag ttgttgttcc atatgatact       180
ttggaagctc caccagaaga tttggaagaa acaaaaaagt tgttgaacaa gttggccgtc       240
ttgaagttga atggtggttt gggtactact atgggttgta ctggtccaaa gtctgttatc       300
gaagttagaa acggtttcac cttcttggat ttgatcgtca tccaaatcga atccttgaac       360
aaaaagtacg gttccaacgt tcctttgttg ttgatgaact ctttcaacac ccatgaagat       420
accttgaaga tcgttgaaaa gtacaccaac tccaacatcg aagttcacac cttcaatcaa       480
tctcaatacc caagagttgt tgccgatgaa tttttgccat ggccatctaa aggtaagact       540
tgtaaagatg gttggtatcc accaggtcat ggtgatattt ttccatcctt gatgaacagt       600
ggtaagttgg acttgttgtt gtcccaaggt aaagaatacg ttttcattgc caactccgat       660
aacttgggtg ctatagttga tatgaagatt ttgaaccact tgatccacaa gcaaaacgaa       720
tactgtatgg aagttactcc aaagactttg gctgatgtta agggtggtac tttgatctct       780
tacgaagata aggttcaatt attggaaatc gcccaagttc cagatgctca tgttaatgaa       840
ttcaagtcca tcgaaaagtt caagatcttt aacaccaaca acttgtgggt taacttgaag       900
gccattaaga gattagttga agctgacgct ttgaagatgg aaattatccc aaacccaaaa       960
gaagttgacg gtgttaaggt attgcaattg gaaactgctg ctggtgctgc tattagattt      1020
ttcgatcatg ctatcggtat caacgtccca agatctagat ttttaccagt taaggctacc      1080
tccgacttgc aattagttca atctgacttg tacaccttgg ttgatggttt cgttactaga      1140
aatccagcta gaactaatcc atccaaccca tctattgaat tgggtccaga attcaagaag      1200
gttggttgtt ttttgggtag attcaagtct atcccatcca tcgttgaatt ggacactttg      1260
aaagtttctg gtgatgtttg gttcggttcc tccattacat tgaaaggtaa ggttactatt      1320
accgctcaac caggtgttaa gttggaaatt ccagatggtg ctgtcatcga aaacaaggat      1380
attaacggtc ctgaagattt gtga                                            1404
```

SEQ ID NO:137
O. sativa

```
MADEKLAKLR EAVAGLSQIS DNEKSGFISL VARYLSGEEE HVEWAKIHTP TDEVVVPYDT        60
LEAPPEDLEE TKKLLNKLAV LKLNGGLGTT MGCTGPKSVI EVRNGFTFLD LIVIQIESLN       120
KKYGSNVPLL LMNSFNTHED TLKIVEKYTN SNIEVHTFNQ SQYPRVVADE FLPWPSKGKT       180
CKDGWYPPGH GDIFPSLMNS GKLDLLLSQG KEYVFIANSD NLGAIVDMKI LNHLIHKQNE       240
YCMEVTPKTL ADVKGGTLIS YEDKVQLLEI AQVPDAHVNE FKSIEKFKIF NTNNLWVNLK       300
AIKRLVEADA LKMEIIPNPK EVDGVKVLQL ETAAGAAIRF FDHAIGINVP RSRFLPVKAT       360
SDLQLVQSDL YTLVDGFVTR NPARTNPSNP SIELGPEFKK VGCFLGRFKS IPSIVELDTL       420
KVSGDVWFGS SITLKGKVTI TAQPGVKLEI PDGAVIENKD INGPEDL                     467
```

SEQ ID NO:138
S. tuberosum

```
atggctactg ctactacttt gtctccagct gatgctgaaa agttgaacaa tttgaaatct        60
gctgtcgccg gtttgaatca aatctctgaa aacgaaaagt ccggtttcat caacttggtt       120
ggtagatatt tgtctggtga agcccaacat attgactggt ctaaaattca aactccaacc       180
gatgaagttg ttgtcccata tgataagttg gctccattgt ctgaagatcc agctgaaaca       240
aaaaagttgt tggacaagtt ggtcgtcttg aagttgaatg gtggtttggg tactactatg       300
ggttgtactg gtccaaagtc tgttatcgaa gttagaaacg gtttgacctt cttggatttg       360
atcgtcaagc aaattgaagc tttgaacgct aagttcggtt gttctgttcc tttgttgttg       420
atgaactctt tcaacaccca tgatgacacc ttgaagatcg ttgaaaagta cgccaactcc       480
aacattgata tccacacctt caatcaatcc caatacccaa gattggttac cgaagatttt       540
gctccattgc catgtaaagg taactctggt aaagatggtt ggtatccacc aggtcatggt       600
gatgtttttc catccttgat gaattccggt aagttggatg ctttgttggc taagggtaaa       660
gaatacgttt tcgttgccaa ctctgataac ttgggtgcta tcgttgattt gaaaatcttg       720
aaccacttga tcttgaacaa gaacgaatac tgcatggaag ttactccaaa gactttggct       780
gatgttaagg gtggtacttt gatttcttac gaaggtaagg ttcaattatt ggaaatcgcc       840
```

```
caagttccag atgaacacgt taatgaattc aagtccatcg aaaagtttaa gatcttcaac       900
actaacaact tgtgggtcaa cttgtctgcc attaagagat tggttgaagc tgatgccttg       960
aaaatggaaa ttattccaaa cccaaaagaa gtcgatggtg tcaaagtatt gcaattggaa      1020
actgctgctg gtgctgctat taagtttttc gatagagcta ttggtgccaa cgttccaaga      1080
tctagatttt tgccagttaa ggctacctct gacttgttgt tggttcaatc agacttgtac      1140
actttgactg atgaaggtta cgttattaga aacccagcta gatccaatcc atccaaccca      1200
tctattgaat tgggtccaga attcaagaag gtagccaatt ttttgggtag attcaagtct      1260
atcccatcca tcatcgattt ggattctttg aaagttactg gtgatgtctg gtttggttct      1320
ggtgttactt tgaaaggtaa agttaccgtt gctgctaagt caggtgttaa gttggaaatt      1380
ccagatggtg ctgttattgc caacaaggat attaacggtc cagaagatat ctaa            1434
```

SEQ ID NO:139

S. tuberosum

```
MATATTLSPA DAEKLNNLKS AVAGLNQISE NEKSGFINLV GRYLSGEAQH IDWSKIQTPT        60
DEVVVPYDKL APLSEDPAET KKLLDKLVVL KLNGGLGTTM GCTGPKSVIE VRNGLTFLDL       120
IVKQIEALNA KFGCSVPLLL MNSFNTHDDT LKIVEKYANS NIDIHTFNQS QYPRLVTEDF       180
APLPCKGNSG KDGWYPPGHG DVFPSLMNSG KLDALLAKGK EYVFVANSDN LGAIVDLKIL       240
NHLILNKNEY CMEVTPKTLA DVKGGTLISY EGKVQLLEIA QVPDEHVNEF KSIEKFKIFN       300
TNNLWVNLSA IKRLVEADAL KMEIIPNPKE VDGVKVLQLE TAAGAAIKFF DRAIGANVPR       360
SRFLPVKATS DLLLVQSDLY TLTDEGYVIR NPARSNPSNP SIELGPEFKK VANFLGRFKS       420
IPSIIDLDSL KVTGDVWFGS GVTLKGKVTV AAKSGVKLEI PDGAVIANKD INGPEDI          477
```

SEQ ID NO:140

```
atgttcttgt tggttacctc ttgcttcttg ccagattctg gttcttctgt taaggtcagt        60
ttgttcatct tcggtgtctc attggtttct acctctccaa ttgatggtca aaaaccaggt       120
acttctggtt tgagaaagaa ggtcaaggtt ttcaagcaac ctaactactt ggaaaacttc       180
gttcaagcta ctttcaacgc tttgactacc gaaaaagtta agggtgctac tttggttgtt       240
tctggtgatg gtagatatta ctccgaacaa gccattcaaa tcatcgttaa gatggctgct       300
gctaacggtg ttagaagagt ttgggttggt caaaactctt tgttgtctac tccagctgtt       360
tccgccatta ttagagaaag agttggtgct gatggttcta aagctactgg tgctttcatt       420
ttgactgctt ctcataatcc aggtggtcca actgaagatt tcggtattaa gtacaacatg       480
gaaaatggtg gtccagcccc agaatctatt actgataaga tatacgaaaa caccaagacc       540
atcaaagaat acccaattgc agaagatttg ccaagagttg atatctctac tatcggtatc       600
acttctttcg aaggtcctga aggtaaattc gacgttgaag tttttgattc cgctgatgat       660
tacgtcaagt tgatgaagtc catcttcgac ttcgaatcca tcaagaagtt gttgtcttac       720
ccaaagttca ccttttgtta cgatgcattg catggtgttg ctggtgctta tgctcataga       780
attttcgttg aagaattggg tgctccagaa tcctctttat tgaactgtgt tccaaaagaa       840
gattttggtg gtggtcatcc agatccaaat ttgacttatg ccaaagaatt ggttgccaga       900
atgggtttgt ctaagactga tgatgctggt ggtgaaccac ctgaatttgg tgctgctgca       960
gatggtgatg ctgatagaaa tatgatcttg ggtaaaagat tcttcgtcac cccatctgat      1020
tccgttgcta ttattgctgc taatgctgtt ggtgctattc catacttttc atccggtttg      1080
aaaggtgttg ctagatctat gccaacttct gctgctttgg atgttgttgc taagaatttg      1140
ggtttgaagt tcttcgaagt tccaactggt tggaaattct tcggtaattt gatggatgca      1200
ggtatgtgtt ctgtttgcgg tgaagaatca tttggtactg gttccgatca tatcagagaa      1260
aaggatggta tttgggctgt tttggcttgg ttgtctattt tggctcacaa gaacaaagaa      1320
accttggatg gtaatgccaa gttggttact gttgaagata tcgttagaca acattgggct      1380
acttacggta gacattacta cactagatac gactacgaaa acgttgatgc tacagctgct      1440
aaagaattga tgggtttatt ggtcaagttg caatcctcat tgccagaagt taacaagatc      1500
atcaagggta tccatcctga agttgctaat gttgcttctg ctgatgaatt cgaatacaag      1560
gatccagttg atggttccgt ttctaaacat caaggtatca gatacttgtt tgaagatggt      1620
tccagattgg ttttcagatt gtctggtaca ggttctgaag gtgctactat tagattgtac      1680
atcgaacaat acgaaaagga cgcctctaag attggtagag attctcaaga tgctttgggt      1740
ccattggttg atgttgcttt gaagttgtcc aagatgcaag aattcactgg tagatcttct      1800
ccaaccgtta ttacctga                                                    1818
```

SEQ ID NO:141

```
MFLLVTSCFL PDSGSSVKVS LFIFGVSLVS TSPIDGQKPG TSGLRKKVKV FKQPNYLENF        60
VQATFNALTT EKVKGATLVV SGDGRYYSEQ AIQIIVKMAA ANGVRRVWVG QNSLLSTPAV       120
SAIIRERVGA DGSKATGAFI LTASHNPGGP TEDFGIKYNM ENGGPAPESI TDKIYENTKT       180
IKEYPIAEDL PRVDISTIGI TSFEGPEGKF DVEVFDSADD YVKLMKSIFD FESIKKLLSY       240
PKFTFCYDAL HGVAGAYAHR IFVEELGAPE SSLLNCVPKE DFGGGHPDPN LTYAKELVAR       300
MGLSKTDDAG GEPPEFGAAA DGDADRNMIL GKRFFVTPSD SVAIIAANAV GAIPYFSSGL       360
KGVARSMPTS AALDVVAKNL GLKFFEVPTG WKFFGNLMDA GMCSVCGEES FGTGSDHIRE       420
KDGIWAVLAW LSILAHKNKE TLDGNAKLVT VEDIVRQHWA TYGRHYYTRY DYENVDATAA       480
KELMGLLVKL QSSLPEVNKI IKGIHPEVAN VASADEFEYK DPVDGSVSKH QGIRYLFEDG       540
SRLVFRLSGT GSEGATIRLY IEQYEKDASK IGRDSQDALG PLVDVALKLS KMQEFTGRSS       600
PTVIT                                                                   605
```

SEQ ID NO:142

```
atggccattc ataatagagc tggtcaacca gcacaacaat ccgatttgat taacgttgct        60
caattgaccg cccaatatta cgttttgaaa cctgaagctg gtaacgctga acatgctgtt       120
aagtttggta cttctggtca tagaggttct gctgctagac attcttttaa cgaaccacat       180
attttggcta tcgctcaagc tattgctgaa gaaagagcta agaacggtat tactggtcca       240
tgttacgttg gtaaagatac ccatgctttg tctgaaccag ctttcatttc tgttttggaa       300
gttttggctg ctaacggtgt tgatgttatc gttcaagaaa acaacggttt cactccaact       360
ccagctgttt ctaatgctat tttggttcac aacaaaaagg gtggtccatt ggctgatggt       420
atagttatta ctccatctca taacccacct gaagatggtg gtattaagta caatccacca       480
aatggtggtc cagctgatac aaatgttact aaggttgttg aagatagagc caacgctttg       540
ttagctgatg gtttgaaagg tgtcaagaga atctctttgg atgaagctat ggcttcaggt       600
catgtcaaag aacaagattt ggttcaacca ttcgttgaag gtttggctga tatagttgat       660
atggctgcta ttcaaaaggc tggtttgact ttgggtgttg atccattggg tggttctggt       720
attgaatact ggaaaagaat cggtgaatat tacaacttga acttgaccat cgtcaacgat       780
caagttgacc aaactttcag attcatgcac ttggataagg atggtgctat tagaatggac       840
tgttcttctg aatgtgctat ggctggttta ttggctttga gagataagtt cgatttggct       900
tttgctaacg atccagatta cgatagacat ggtatcgtta ctccagcagg tttgatgaat       960
ccaaatcatt acttggctgt tgccatcaac tacttgtttc aacatagacc acaatggggt      1020
aaggatgttg ctgttggtaa aactttggtt tcctccgcta tgatcgatag agttgttaac      1080
gatttgggta gaaagttggt tgaagttcca gttggtttca agtggtttgt tgacggtttg      1140
tttgatggtt cttttggttt tggtggtgaa gaatctgctg gtgcttcatt tttgagattt      1200
gatggtactc catggtccac tgacaaagat ggtattatca tgtgtttgtt ggctgctgaa      1260
attactgctg ttactggtaa gaatccacaa gaacactaca acgaattggc taagagattt      1320
ggtgctccat cttacaatag attgcaagct gctgctactt ctgctcaaaa agctgcttta      1380
tctaagttgt ccccagaaat ggtttctgct tctactttag ctggtgatcc aattacagct      1440
agattgactg ctgctccagg taatggtgct tctattggtg gtttaaaggt tatgactgat      1500
aacggttggt ttgctgcaag accatctggt actgaagatg cttacaaaat ctactgcgaa      1560
tccttcttgg gtgaagaaca tagaaagcaa attgaaaaag aagccgtcga aatcgtcagt      1620
gaagttttga agaatgccta a                                                1641
```

SEQ ID NO:143

```
MAIHNRAGQP AQQSDLINVA QLTAQYYVLK PEAGNAEHAV KFGTSGHRGS AARHSFNEPH        60
ILAIAQAIAE ERAKNGITGP CYVGKDTHAL SEPAFISVLE VLAANGVDVI VQENNGFTPT       120
PAVSNAILVH NKKGGPLADG IVITPSHNPP EDGGIKYNPP NGGPADTNVT KVVEDRANAL       180
LADGLKGVKR ISLDEAMASG HVKEQDLVQP FVEGLADIVD MAAIQKAGLT LGVDPLGGSG       240
IEYWKRIGEY YNLNLTIVND QVDQTFRFMH LDKDGAIRMD CSSECAMAGL LALRDKFDLA       300
FANDPDYDRH GIVTPAGLMN PNHYLAVAIN YLFQHRPQWG KDVAVGKTLV SSAMIDRVVN       360
DLGRKLVEVP VGFKWFVDGL FDGSFGFGGE ESAGASFLRF DGTPWSTDKD GIIMCLLAAE       420
ITAVTGKNPQ EHYNELAKRF GAPSYNRLQA AATSAQKAAL SKLSPEMVSA STLAGDPITA       480
```

```
RLTAAPGNGA SIGGLKVMTD NGWFAARPSG TEDAYKIYCE SFLGEEHRKQ IEKEAVEIVS       540
EVLKNA                                                                  546
```

SEQ ID NO:144
R. suavissimus

```
atgtcctccg gtaagattaa gagagttcaa actactccat tcgacggtca aaaaccaggt        60
acttctggtt tgagaaagaa ggttaaggtt ttcacccaac ctaactactt gcaaaacttc       120
gttcaatcta ccttcaacgc tttgccatct gataaggtaa aaggtgctag attggttgtt       180
tctggtgatg gtagatactt ctccaaagaa gccattcaaa tcatcattaa gatggctgct       240
ggtaacggtg ttaagtctgt ttgggttggt caaaatggtt tgttgtctac tccagctgtt       300
tctgctgttg ttagagaaag agttggtgct gatggttgta aagcttctgg tgctttcatt       360
ttgactgctt ctcataatcc aggtggtcca aatgaagatt tcggtatcaa gtacaacatg       420
gaaaatggtg gtccagctcc agaatctatt accaacaaaa tctacgaaaa caccacccaa       480
atcaaagaat acttgaccgt tgatttgcca gaagttgata ttactaagcc aggtgttact       540
accttcgaag ttgaaggtgg tactttcact gttgatgttt tcgattctgc ttccgattac       600
gtcaagttga tgaagtccat tttcgacttc gaatccatca gaaagttgtt gtcctctcca       660
aagttcacct tttgttttga tgcattgcat ggtgttggtg gtgcttacgc taaaagaatt       720
ttcgttgaag aattgggtgc caaagaatcc tctttgttga actgtgttcc taaagaagat       780
tttggtggtg gtcatccaga tccaaatttg acatatgcta aagaattggt cgccagaatg       840
ggtttgtcta agtctaatac tcaaaacgaa ccaccagaat ttggtgctgc tgcagatggt       900
gatgctgata gaaatatggt tttgggtaag agattcttcg ttaccccatc tgattccgtt       960
gctattattg ctgctaatgc tgttgaagct atcccatact tttctactgg tttgaaaggt      1020
gttgctagat ctatgccaac ttctgctgct ttggatgttg ttgctaaaca cttgaacttg      1080
aagttcttcg aagtaccaac tggttggaag tttttcggta atttgatgga tgctggtttg      1140
tgttctgttt gcggtgaaga atcttttggt actggttccg atcatatcag agaaaaggat      1200
ggtatttggg ctgttttggc ttggttgtca attattgcca tcaagaacaa ggataacatc      1260
ggtggtgata agttggttac cgttgaagat atcgttagaa aacattgggc tacttacggt      1320
agacattact acactagata cgattacgaa aacgttgatg ctggtaaggc taaagatttg      1380
atggcatcat tggtcaactt gcaatcatct ttgcctgaag ttaacaagat cgttaagggt      1440
atctgttccg atgttgcaaa tgttgttggt gccgatgaat tcgaatacaa ggattctgtt      1500
gatggttcca tctccaaaca tcaaggtatc agatacttgt tcgaagatgg ttcaagattg      1560
gttttcagat tgtctggtac aggttctgaa ggtgctacta ttagattgta catcgaacaa      1620
tacgaaaatg acccatccaa gatctccaga gaatcttctg aagctttggc tccattggtt      1680
gaagttgctt tgaaattgtc caagatgcaa gaattcactg gtagatcagc tccaactgtt      1740
attacctga                                                              1749
```

SEQ ID NO:145
R. suavissimus

```
MSSGKIKRVQ TTPFDGQKPG TSGLRKKVKV FTQPNYLQNF VQSTFNALPS DKVKGARLVV        60
SGDGRYFSKE AIQIIIKMAA GNGVKSVWVG QNGLLSTPAV SAVVRERVGA DGCKASGAFI       120
LTASHNPGGP NEDFGIKYNM ENGGPAPESI TNKIYENTTQ IKEYLTVDLP EVDITKPGVT       180
TFEVEGGTFT VDVFDSASDY VKLMKSIFDF ESIRKLLSSP KFTFCFDALH GVGGAYAKRI       240
FVEELGAKES SLLNCVPKED FGGGHPDPNL TYAKELVARM GLSKSNTQNE PPEFGAAADG       300
DADRNMVLGK RFFVTPSDSV AIIAANAVEA IPYFSTGLKG VARSMPTSAA LDVVAKHLNL       360
KFFEVPTGWK FFGNLMDAGL CSVCGEESFG TGSDHIREKD GIWAVLAWLS IIAIKNKDNI       420
GGDKLVTVED IVRKHWATYG RHYYTRYDYE NVDAGKAKDL MASLVNLQSS LPEVNKIVKG       480
ICSDVANVVG ADEFEYKDSV DGSISKHQGI RYLFEDGSRL VFRLSGTGSE GATIRLYIEQ       540
YENDPSKISR ESSEALAPLV EVALKLSKMQ EFTGRSAPTV IT                          582
```

SEQ ID NO:146

```
atggcctctt tcaaggttaa cagagttgaa tcctctccaa tcgaaggtca aaaaccaggt        60
acttctggtt tgagaaagaa ggttaaggtt ttcacccaac cacattactt gcacaacttc       120
gttcaatcta ctttcaacgc tttgtctgcc gaaaaagtta agggttctac tttggttgtt       180
```

```
tccggtgatg gtagatatta ctccaaggat gccattcaaa tcatcattaa gatggctgct       240
gctaacggtg ttagaagagt ttgggttggt caaaatggtt tgttgtctac tccagctgtt       300
tctgctgttg ttagagaaag agttggtgct gatggttcta aatctaacgg tgctttcatt       360
ttgactgcct ctcataatcc aggtggtcca aatgaagatt tcggtatcaa gtacaacatg       420
gaaaatggtg gtccagctcc agaaggtatt actgataaga tttttgaaaa caccaagacc       480
atcaaagaat acttcattgc tgaaggtttg ccagacgttg atatttccgc tattggtatc       540
tcttcattct ctggtccaga tggtcaattc gatgttgatg ttttcgattc ctcttccgac       600
tacgtcaaat tgatgaagtc catcttcgac ttccaatcca tcaagaagtt gattacctcc       660
ccacaatttt ctttctgtta cgatgcttta catggtgttg gtggtgctta tgctaagcca       720
atttttgttg atgaattggg tgccaaagaa tcctctttgt tgaactgtgt tcctaaagaa       780
gattttggtg gtggtcatcc agatccaaat ttgacttacg ctaaagaatt ggtttccaga       840
atgggtttgg gtaagaatcc agattctaat ccaccagaat ttggtgctgc tgcagatggt       900
gatgctgata gaaatatgat cttgggtaaa agattcttcg tcacccccatc tgattccgtt      960
gctattattg ctgctaatgc cgttcaatca atcccatact tttcatccgg tttgaaaggt      1020
gttgctagat ctatgccaac ttctgctgct ttggatgttg ttgctaagtc tttgaacttg      1080
aagttcttcg aagttccaac tggttggaag tttttcggta atttgatgga tgctggtttg      1140
tgttctgttt gcggtgaaga atcatttggt actggttccg atcatatcag agaaaaggat      1200
ggtatttggg ctgttttggc ttggttgtct attttggctc ataagaacaa ggacaacttg      1260
aacggtggta acttggttac tgttgaagat atcgttaagc aacattgggc tacttacggt      1320
agacattact acactagata cgactacgaa aacgttgatg ctggtgctgc aaaagaattg      1380
atggctcatt tggttaagtt gcaatcctcc atctctgatg ttaacacctt cattaagggt      1440
atcagatccg atgttgctaa tgttgcatct gctgatgaat tcgaatacaa ggatccagtt      1500
gacggttcta tttccaaaca tcaaggtatt agatacttgt ttgaagatgg ttccagattg      1560
gttttcagat tgtctggtac aggttctgaa ggtgctacta ttagattgta catcgaacaa      1620
tacgaaaagg attcctctaa gaccggtaga gattctcaag aagctttggc tccattagtt      1680
gaagttgcct tgaaattgtc caagatgcaa gaattcactg gtagatctgc tccaactgtt      1740
attacctga                                                             1749
```

SEQ ID NO:147

```
MASFKVNRVE SSPIEGQKPG TSGLRKKVKV FTQPHYLHNF VQSTFNALSA EKVKGSTLVV        60
SGDGRYYSKD AIQIIIKMAA ANGVRRVWVG QNGLLSTPAV SAVVRERVGA DGSKSNGAFI       120
LTASHNPGGP NEDFGIKYNM ENGGPAPEGI TDKIFENTKT IKEYFIAEGL PDVDISAIGI       180
SSFSGPDGQF DVDVFDSSSD YVKLMKSIFD FQSIKKLITS PQFSFCYDAL HGVGGAYAKP       240
IFVDELGAKE SSLLNCVPKE DFGGGHPDPN LTYAKELVSR MGLGKNPDSN PPEFGAAADG       300
DADRNMILGK RFFVTPSDSV AIIAANAVQS IPYFSSGLKG VARSMPTSAA LDVVAKSLNL       360
KFFEVPTGWK FFGNLMDAGL CSVCGEESFG TGSDHIREKD GIWAVLAWLS ILAHKNKDNL       420
NGGNLVTVED IVKQHWATYG RHYYTRYDYE NVDAGAAKEL MAHLVKLQSS ISDVNTFIKG       480
IRSDVANVAS ADEFEYKDPV DGSISKHQGI RYLFEDGSRL VFRLSGTGSE GATIRLYIEQ       540
YEKDSSKTGR DSQEALAPLV EVALKLSKMQ EFTGRSAPTV IT                          582
```

SEQ ID NO:148

```
gcacacacca tagcttcaaa atgtttctac tccttttta ctcttccaga ttttctcgga         60
ctccgcgcat cgccgtacca cttcaaaaca cccaagcaca gcatactaaa tttcccctct        120
ttcttcctct agggtgtcgt taattacccg tactaaaggt ttggaaaaga aaaaagagac        180
cgcctcgttt ctttttcttc gtcgaaaaag gcaataaaaa tttttatcac gtttcttttt        240
cttgaaaatt ttttttttg attttttct ctttcgatga cctcccattg atatttaagt          300
taataaacgg tcttcaattt ctcaagtttc agtttcattt ttcttgttct attacaactt        360
tttttacttc ttgctcatta gaaagaaagc atagcaatct aatctaagtt ttaattacaa        420
ggatcc                                                                  426
```

SEQ ID NO:149

```
ggaagtacct tcaaagaatg gggtcttatc ttgttttgca agtaccactg agcaggataa         60
taatagaaat gataatatac tatagtagag ataacgtcga tgacttccca tactgtaatt        120
```

gcttttagtt gtgtattttt agtgtgcaag tttctgtaaa tcgattaatt ttttttttctt 180
tcctcttttt attaacctta atttttattt tagattcctg acttcaactc aagacgcaca 240
gatattataa catctgcata ataggcattt gcaagaatta ctcgtgagta aggaaagagt 300
gaggaactat cgcatacctg catttaaaga tgccgatttg ggcgcgaatc ctttattttg 360
gcttcaccct catactatta tcagggccag aaaaaggaag tgtttccctc cttcttgaat 420
tgatgttacc ctcataaagc acgtggcctc ttatcgagaa agaaattacc gtcgctcgtg 480
atttgtttgc aaaaagaaca aaactgaaaa aacccagaca cgctcgactt cctgtcttcc 540
tattgattgc agcttccaat ttcgtcacac aacaaggtcc tagcgacggc tcacaggttt 600
tgtaacaagc aatcgaaggt tctggaatgg cgggaaaggg tttagtacca catgctatga 660
tgcccactgt gatctccaga gcaaagttcg ttcgatcgta ctgttactct ctctctttca 720
aacagaattg tccgaatcgt gtgacaacaa cagcctgttc tcacacactc ttttcttcta 780
accaaggggg tggtttagtt tagtagaacc tcgtgaaact tacatttaca tatatataaa 840
cttgcataaa ttggtcaatg caagaaatac atatttggtc ttttctaatt cgtagttttt 900
caagttctta gatgctttct ttttctcttt tttacagatc atcaaggaag taattatcta 960
ctttttacaa caaatataaa acaa 984

SEQ ID NO:150

cattatcaat actgccattt caaagaatac gtaaataatt aatagtagtg attttcctaa 60
ctttatttag tcaaaaaatt agccttttaa ttctgctgta acccgtacat gcccaaaata 120
gggggcgggt tacacagaat atataacatc gtaggtgtct gggtgaacag tttattcctg 180
gcatccacta aatataatgg agcccgcttt ttaagctggc atccagaaaa aaaaagaatc 240
ccagcaccaa aatattgttt tcttcaccaa ccatcagttc ataggtccat tctcttagcg 300
caactacaga gaacaggggc acaaacaggc aaaaaacggg cacaacctca atggagtgat 360
gcaacctgcc tggagtaaat gatgacacaa ggcaattgac ccacgcatgt atctatctca 420
ttttcttaca ccttctatta ccttctgctc tctctgattt ggaaaaagct gaaaaaaaag 480
gttgaaacca gttccctgaa attattcccc tacttgacta ataagtatat aaagacggta 540
ggtattgatt gtaattctgt aaatctattt cttaaacttc ttaaattcta cttttatagt 600
tagtcttttt tttagtttta aaacaccaag aacttagttt cgaataaaca cacataaaca 660
aacaaa 666

SEQ ID NO:151

gatctgggcc gtatacttac atatagtaga tgtcaagcgt aggcgcttcc cctgccggct 60
gtgagggcgc cataaccaag gtatctatag accgccaatc agcaaactac ctccgtacat 120
tcatgttgca cccacacatt tatacaccca gaccgcgaca aattacccat aaggttgttt 180
gtgacggcgt cgtacaagag aacgtgggaa ctttttaggc tcaccaaaaa agaaagaaaa 240
aatacgagtt gctgacagaa gcctcaagaa aaaaaaaatt cttcttcgac tatgctggag 300
gcagagatga tcgagccggt agttaactat atatagctaa attggttcca tcaccttctt 360
ttctggtgtc gctccttcta gtgctatttc tggcttttcc tatttttttt tttccatttt 420
tctttctctc tttctaatat ataaattctc ttgcattttc tatttttctc tctatctatt 480
ctacttgttt attcccttca aggttttttt ttaaggagta cttgttttta gaatatacgg 540
tcaacgaact ataattaact aaaca 565

SEQ ID NO:152

agttataata atcctacgtt agtgtgagcg ggatttaaac tgtgaggacc ttaatacatt 60
cagacacttc tgcggtatca ccctacttat tcccttcgag attatatcta ggaacccatc 120
aggttggtgg aagattaccc gttctaagac ttttcagctt cctctattga tgttacacct 180
ggacacccct tttctggcat ccagttttta atcttcagtg gcatgtgaga ttctccgaaa 240
ttaattaaag caatcacaca attctctcgg ataccacctc ggttgaaact gacaggtggt 300
ttgttacgca tgctaatgca aaggagccta tatacctttg gctcggctgc tgtaacaggg 360
aatataaagg gcagcataat ttaggagttt agtgaacttg caacatttac tattttccct 420
tcttacgtaa atatttttct ttttaattct aaatcaatct ttttcaattt tttgtttgta 480
ttcttttctt gcttaaatct ataactacaa aaaacacata cataaactaa aa 532

SEQ ID NO:153

```
gatctatgcg actgggtgag catatgttcc gctgatgtga tgtgcaagat aaacaagcaa        60
ggcagaaact aacttcttct tcatgtaata aacacacccc gcgtttattt acctatctct       120
aaacttcaac accttatatc ataactaata tttcttgaga taagcacact gcacccatac       180
cttccttaaa aacgtagctt ccagtttttg gtggttccgg cttccttccc gattccgccc       240
gctaaacgca tatttttgtt gcctggtggc atttgcaaaa tgcataacct atgcatttaa       300
aagattatgt atgctcttct gacttttcgt gtgatgaggc tcgtggaaaa aatgaataat       360
ttatgaattt gagaacaatt ttgtgttgtt acggtatttt actatggaat aatcaatcaa       420
ttgaggattt tatgcaaata tcgtttgaat atttttccga ccctttgagt acttttcttc       480
ataattgcat aatattgtcc gctgcccctt tttctgttag acggtgtctt gatctacttg       540
ctatcgttca acaccacctt attttctaac tatttttttt ttagctcatt tgaatcagct       600
tatggtgatg gcacattttt gcataaacct agctgtcctc gttgaacata ggaaaaaaaa       660
atatataaac aaggctcttt cactctcctt gcaatcagat ttgggtttgt tccctttatt       720
ttcatatttc ttgtcatatt cctttctcaa ttattatttt ctactcataa cctcacgcaa       780
aataacacag tcaaatctat caaaa                                             805
```

SEQ ID NO:154

```
atccgctcta accgaaaagg aaggagttag acaacctgaa gtctaggtcc ctatttattt        60
tttttaatag ttatgttagt attaagaacg ttatttatat ttcaaatttt tctttttttt       120
ctgtacaaac gcgtgtacgc atgtaacatt atactgaaaa ccttgcttga gaaggttttg       180
ggacgctcga ag                                                          192
```

SEQ ID NO:155

```
gtagatacgt tgttgacact tctaaataag cgaatttctt atgatttatg atttttatta        60
ttaaataagt tataaaaaaa ataagtgtat acaaatttta aagtgactct taggttttaa       120
aacgaaaatt cttattcttg agtaactctt tcctgtaggt caggttgctt tctcaggtat       180
agcatgaggt cgctc                                                       195
```

WHAT IS CLAIMED IS:

1. A recombinant host cell capable of producing one or more steviol glycosides or a steviol glycoside composition in a cell culture, comprising:
    (a) a recombinant gene encoding a polypeptide capable of synthesizing uridine 5'-triphosphate (UTP) from uridine diphosphate (UDP);
    (b) a recombinant gene encoding a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate; and/or
    (c) a recombinant gene encoding a polypeptide capable of synthesizing uridine diphosphate glucose (UDP-glucose) from UTP and glucose-1-phosphate.

2. The recombinant host cell of claim 1, wherein:
    (a) the polypeptide capable of synthesizing UTP from UDP comprises a polypeptide having at least 60% sequence identity to the amino acid sequence set forth in SEQ ID NO:123;
    (b) the polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate comprises a polypeptide having at least 60% sequence identity to the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:119, SEQ ID NO:143 or a polypeptide having at least 55% sequence identity to the amino acid sequence set forth in SEQ ID NO:141, SEQ ID NO:145, or SEQ ID NO:147; and/or
    (c) the polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate comprises a polypeptide having at least 60% sequence identity to the amino acid sequence set forth in SEQ ID NO:121, SEQ ID NO:127, a polypeptide having at least 55% sequence identity to the amino acid sequence set forth in SEQ ID NO:125, SEQ ID NO:129, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, or SEQ ID NO:139 or a polypeptide having at least 70% sequence identity to the amino acid sequence set forth in SEQ ID NO:131.

3. The recombinant host cell of claim 1 or 2, further comprising:

(a) a gene encoding a polypeptide capable of glycosylating steviol or a steviol glycoside at its C-13 hydroxyl group thereof;

(b) a gene encoding a polypeptide capable of beta 1,3 glycosylation of the C3' of the 13-*O*-glucose, 19-*O*-glucose, or both 13-*O*-glucose and 19-*O*-glucose of a steviol glycoside;

(c) a gene encoding a polypeptide capable of glycosylating steviol or a steviol glycoside at its C-19 carboxyl group thereof; and/or

(d) a gene encoding a polypeptide capable of beta 1,2 glycosylation of the C2' of the 13-*O*-glucose, 19-*O*-glucose, or both 13-*O*-glucose and 19-*O*-glucose of a steviol glycoside.

4. The recombinant host cell of claim 3, further comprising:

(e) a gene encoding a polypeptide capable of synthesizing geranylgeranyl pyrophosphate (GGPP) from farnesyl diphosphate (FPP) and isopentenyl diphosphate (IPP);

(f) a gene encoding a polypeptide capable of synthesizing *ent*-copalyl diphosphate from GGPP;

(g) a gene encoding an a polypeptide capable of synthesizing *ent*-kaurene from *ent*-copalyl diphosphate;

(h) a gene encoding a polypeptide capable of synthesizing *ent*-kaurenoic acid from *ent*-kaurene;

(i) a gene encoding a polypeptide capable of reducing cytochrome P450 complex; and/or

(j) a gene encoding a polypeptide capable of synthesizing steviol from *ent*-kaurenoic acid.

5. The recombinant host cell of claim 3 or 4, wherein:

(a) the polypeptide capable of glycosylating steviol or a steviol glycoside at its C-13 hydroxyl group thereof comprises a polypeptide having at least 55% sequence identity to the amino acid sequence set forth in SEQ ID NO:7;

(b) the polypeptide capable of beta 1,3 glycosylation of the C3' of the 13-O-glucose, 19-O-glucose, or both 13-O-glucose and 19-O-glucose of a

steviol glycoside comprises a polypeptide having at least 50% sequence identity to the amino acid sequence set forth in SEQ ID NO:9;

(c) the polypeptide capable of glycosylating steviol or a steviol glycoside at its C-19 carboxyl group thereof comprises a polypeptide having at least 55% sequence identity to the amino acid sequence set forth in SEQ ID NO:4;

(d) the polypeptide capable of beta 1,2 glycosylation of the C2' of the 13-O-glucose, 19-O-glucose, or both 13-O-glucose and 19-O-glucose of a steviol glycoside comprises a polypeptide having 80% or greater identity to the amino acid sequence set forth in SEQ ID NO:11; a polypeptide having 80% or greater identity to the amino acid sequence set forth in SEQ ID NO:13; or a polypeptide having at least 65% sequence identity to the amino acid sequence set forth in SEQ ID NO:16;

(e) the polypeptide capable of synthesizing GGPP comprises a polypeptide having at least 70% sequence identity to the amino acid sequence set forth in SEQ ID NO:20, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, or SEQ ID NO:116;

(f) the polypeptide capable of synthesizing ent-copalyl diphosphate comprises a polypeptide having at least 70% sequence identity to the amino acid sequence set forth in SEQ ID NO:34, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:40, SEQ ID NO:42, or SEQ ID NO:120;

(g) the polypeptide capable of synthesizing ent-kaurene comprises a polypeptide having at least 70% sequence identity to the amino acid sequence set forth in SEQ ID NO:44, SEQ ID NO:46, SEQ ID NO:48, SEQ ID NO:50, or SEQ ID NO:52;

(h) the polypeptide capable of synthesizing ent-kaurenoic acid comprises a polypeptide having at least 70% sequence identity to the amino acid sequence set forth in SEQ ID NO:60, SEQ ID NO:62, SEQ ID NO:117, SEQ ID NO:66, SEQ ID NO:68, SEQ ID NO:70, SEQ ID NO:72, SEQ ID NO:74, or SEQ ID NO:76;

(i) the polypeptide capable of reducing cytochrome P450 complex comprises a polypeptide having at least 70% sequence identity to the amino acid sequence set forth in SEQ ID NO:78, SEQ ID NO:80, SEQ ID NO:82,

SEQ ID NO:84, SEQ ID NO:86, SEQ ID NO:88, SEQ ID NO:90, SEQ ID NO:92; and/or

(k) the polypeptide capable of synthesizing steviol comprises a polypeptide having at least 70% sequence identity to the amino acid sequence set forth in SEQ ID NO:94, SEQ ID NO:97, SEQ ID NO:100, SEQ ID NO:101, SEQ ID NO:102, SEQ ID NO:103, SEQ ID NO:104, SEQ ID NO:106, SEQ ID NO:108, SEQ ID NO:110, SEQ ID NO:112, or SEQ ID NO:114.

6. The recombinant host cell of any one of claims 1-5, wherein the recombinant host cell comprises:

(a) a gene encoding a polypeptide capable of synthesizing uridine 5'-triphosphate (UTP) from uridine diphosphate (UDP) having at least 60% sequence identity to the amino acid sequence set forth in SEQ ID NO:123;

(b) one or more genes encoding a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate, each having at least 60% sequence identity to the amino acid sequence set forth in SEQ ID NO:2 and/or SEQ ID NO:119; and

(c) a gene encoding a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate having at least 60% sequence identity to the amino acid sequence set forth in SEQ ID NO:121;

7. The recombinant host cell of any one of claims 1-6, wherein the recombinant host cell comprises:

(a) a gene encoding a polypeptide capable of synthesizing uridine 5'-triphosphate (UTP) from uridine diphosphate (UDP);

(b) a gene encoding a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate;

(c) a gene encoding a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate having at least 60% sequence identity to the amino acid sequence set forth in SEQ ID NO:121;

(d) a gene encoding a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate having at least 55% sequence

identity to the amino acid sequence set forth in SEQ ID NO:125, SEQ ID NO:129; SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, or SEQ ID NO:139; or at least 60% sequence identity to the amino acid sequence set forth in SEQ ID NO:127; or at least 70% sequence identity to the amino acid sequence set forth in SEQ ID NO:131; and

one or more of:

(e) a gene encoding a polypeptide capable of glycosylating steviol or a steviol glycoside at its C-13 hydroxyl group thereof having at least 55% sequence identity to the amino acid sequence set forth in SEQ ID NO:7;

(b) a gene encoding a polypeptide capable of beta 1,3 glycosylation of the C3' of the 13-O-glucose, 19-O-glucose, or both 13-O-glucose and 19-O-glucose of a steviol glycoside having at least 50% sequence identity to the amino acid sequence set forth in SEQ ID NO:9;

(c) a gene encoding a polypeptide capable of glycosylating steviol or a steviol glycoside at its C-19 carboxyl group thereof having at least 55% sequence identity to the amino acid sequence set forth in SEQ ID NO:4;

(d) a gene encoding a polypeptide capable of beta 1,2 glycosylation of the C2' of the 13-O-glucose, 19-O-glucose, or both 13-O-glucose and 19-O-glucose of a steviol glycoside comprises a polypeptide having 80% or greater identity to the amino acid sequence set forth in SEQ ID NO:11; a polypeptide having 80% or greater identity to the amino acid sequence set forth in SEQ ID NO:13; or a polypeptide having at least 65% sequence identity to the amino acid sequence set forth in SEQ ID NO:16.

8. The recombinant host cell of any one of claims 1-7, wherein the recombinant host cell comprises:

(a) a recombinant gene encoding a polypeptide capable of synthesizing uridine 5'-triphosphate (UTP) from uridine diphosphate (UDP) having at least 60% sequence identity to the amino acid sequence set forth in SEQ ID NO:123;

(b) one or more recombinant genes encoding a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate, each having at

least 60% sequence identity to the amino acid sequence set forth in SEQ ID NO:2 and/or SEQ ID NO:119; and/or

(c) a recombinant gene encoding a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate having at least 60% sequence identity to the amino acid sequence set forth in SEQ ID NO:121;

wherein the gene encoding a polypeptide capable of synthesizing uridine 5'-triphosphate (UTP) from uridine diphosphate (UDP), the one or more genes encoding a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate, and/or the gene encoding a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate are overexpressed relative to a corresponding host cell lacking the one or more recombinant genes.

9. The recombinant host cell of claim 8, wherein the gene encoding a polypeptide capable of synthesizing uridine 5'-triphosphate (UTP) from uridine diphosphate (UDP), the one or more genes encoding a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate, and/or the gene encoding a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate are overexpressed by at least 10%, or at least 15%, or at least 20%, or at least 30%, or at least 40%, or at least 50%, or at least 60%, or at least 70%, or at least 80%, or at least 90%, or at least 100%, or at least 125%, or at least 150%, or at least 175%, or at least 200% relative to a corresponding host cell lacking the one or more recombinant genes.

10. The recombinant host cell of any one of claims 1-9, wherein expression of the one or more recombinant genes increase the amount of UDP-glucose accumulated by the cell relative to a corresponding host lacking the one or more recombinant genes.

11. The recombinant host cell of claim 10, wherein expression of the one or more recombinant genes increases the amount of UDP-glucose accumulated by the cell by at least about 10%, at least about 25%, or at least about 50%, at least about 100%, at least about 150%, at least about 200%, or at least about 250% relative to a corresponding host lacking the one or more recombinant genes.

12. The recombinant host cell of any one of claims 1-11, wherein expression of the one or more recombinant genes increases an amount of the one or more steviol glycosides or the steviol glycoside composition produced by the cell relative to a corresponding host lacking the one or more recombinant genes.

13. The recombinant host cell of claim 12, wherein expression of the one or more recombinant genes increases the amount of the one or more steviol glycosides produced by the cell by at least about 5%, at least about 10%, at least about 25%, at least about 50%, at least about 75%, or at least about 100% relative to a corresponding host lacking the one or more recombinant genes.

14. The recombinant host cell of claim 12 or 13, wherein expression of the one or more recombinant genes increases the amount of RebA, RebB, Reb D, and/or RebM produced by the cell relative to a corresponding host lacking the one or more recombinant genes.

15. The recombinant host cell of any one of claims 1-14, wherein expression of the one or more recombinant genes decreases the one of one or more steviol glycosides or the steviol glycoside composition accumulated by the cell relative to a corresponding host lacking the one or more recombinant genes.

16. The recombinant host cell of claim 15, wherein expression of the one or more recombinant genes decreases the amount of the one or more steviol glycosides accumulated by the cell by at least about 5%, at least about 10%, at least about 25%, or at least about 50% relative to a corresponding host lacking the one or more recombinant genes.

17. The recombinant host cell of claim 15 or 16, wherein expression of the one or more recombinant genes decreases the amount of RebB, RebD, and/or 13-SMG accumulated by the cell relative to a corresponding host lacking the one or more recombinant genes.

18. The recombinant host cell of any one of claims 1-17, wherein expression of the one or more recombinant genes increases or decreases the amount of total steviol glycosides

produced by the cell by less than 5%, less than 2.5%, or less than 1% relative to a corresponding host lacking the one or more recombinant genes.

19. The recombinant host cell of any one of claims 1-17, wherein expression of the one or more recombinant genes increases the amount of total steviol glycosides produced by the cell by at least about 5%, at least about 10%, or at least about 25% relative to a corresponding host lacking the one or more recombinant genes.

20. The recombinant host cell of any one of claims 1-18, wherein the one or more steviol glycosides is, or the steviol glycoside composition comprises, steviol-13-*O*-glucoside (13-SMG), steviol-1,2-Bioside, steviol-1,3-Bioside, steviol-19-*O*-glucoside (19-SMG), 1,2-Stevioside, 1,3-stevioside (RebG), rubusoside, rebaudioside A (RebA), rebaudioside B (RebB), rebaudioside C (RebC), rebaudioside D (RebD), rebaudioside E (RebE), rebaudioside F (RebF), rebaudioside M (RebM), rebaudioside Q (RebQ), rebaudioside I (RebI), dulcoside A, and/or an isomer thereof.

21. The recombinant host cell of any one of claims 1-20, wherein the recombinant host cell is a plant cell, a mammalian cell, an insect cell, a fungal cell, an algal cell or a bacterial cell.

22. A method of producing one or more steviol glycosides or a steviol glycoside composition in a cell culture, comprising culturing the recombinant host cell of any one of claims 1-21 in the cell culture, under conditions in which the genes are expressed, and wherein the one or more steviol glycosides or the steviol glycoside composition is produced by the recombinant host cell.

23. The method of claim 22, wherein the genes are constitutively expressed and/or expression of the genes is induced.

24. The method of claim 22 or 23, wherein the amount of UDP-glucose accumulated by the cell is increased by at least by at least about 10% relative to a corresponding host lacking the one or more recombinant genes.

25. The method of any one of claims 22-24, wherein the amount of RebA, RebB, RebD, and/or RebM produced by the cell is increased by at least about 5% relative to a corresponding host lacking the one or more recombinant genes.

26. The method of any one of claims 22-25, wherein the amount of RebB, RebD, and/or 13-SMG accumulated by the cell is decreased by at least about 5% relative to a corresponding host lacking the one or more recombinant genes.

27. The method of any one of claims 22-26, wherein the amount of total steviol glycosides produced by the cell is increased or decreased by less than about 5% relative to a corresponding host lacking the one or more recombinant genes.

28. The method of any one of claims 22-26, wherein the amount of total steviol glycosides produced by the cell is increased by at least about 5% relative to a corresponding host lacking the one or more recombinant genes.

29. The method of any one of claims 22-28, wherein the recombinant host cell is grown in a fermentor at a temperature for a period of time, wherein the temperature and period of time facilitate the production of the one or more steviol glycosides or the steviol glycoside composition.

30. The method of claim 29, wherein the amount of UDP-glucose present in the cell culture is increased by at least about 10%, at least about 25%, or at least about 50%, at least about 100%, at least about 150%, at least about 200%, or at least about 250% at any point throughout the period of time.

31. The method of any one of claims 22-30, further comprising isolating the produced one or more steviol glycosides or the steviol glycoside composition from the cell culture.

32. The method of claim 31, wherein the isolating step comprises:

(a) providing the cell culture comprising the one or more steviol glycosides or the steviol glycoside composition;

(b) separating a liquid phase of the cell culture from a solid phase of the cell culture to obtain a supernatant comprising the produced one or more steviol glycosides or the steviol glycoside composition;

(c) providing one or more adsorbent resins, comprising providing the adsorbent resins in a packed column; and

(d) contacting the supernatant of step (b) with the one or more adsorbent resins in order to obtain at least a portion of the produced one or more steviol glycosides or the steviol glycoside composition, thereby isolating the produced one or more steviol glycosides or the steviol glycoside composition;

or

(a) providing the cell culture comprising the one or more steviol glycosides or the steviol glycoside composition;

(b) separating a liquid phase of the cell culture from a solid phase of the cell culture to obtain a supernatant comprising the produced one or more steviol glycosides or the steviol glycoside composition;

(c) providing one or more ion exchange or ion exchange or reversed-phase chromatography columns; and

(d) contacting the supernatant of step (b) with the one or more ion exchange or ion exchange or reversed-phase chromatography columns in order to obtain at least a portion of the produced one or more steviol glycosides or the steviol glycoside composition, thereby isolating the produced one or more steviol glycosides or the steviol glycoside composition;

or

(a) providing the cell culture comprising the one or more steviol glycosides or the steviol glycoside composition;

(b) separating a liquid phase of the cell culture from a solid phase of the cell culture to obtain a supernatant comprising the produced one or more steviol glycosides or the steviol glycoside composition;

(c) crystallizing or extracting the produced one or more steviol glycosides or the steviol glycoside composition, thereby isolating the produced one or more steviol glycosides or the steviol glycoside composition.

33. The method of any one of claims 22-30, further comprising recovering the one or more steviol glycosides or the steviol glycoside composition from the cell culture.

34. The method of claim 33, wherein the recovered one or more steviol glycosides or the steviol glycoside composition has a reduced level of Stevia plant-derived components relative to a plant-derived Stevia extract.

35. A method for producing one or more steviol glycosides or a steviol glycoside composition, comprising whole-cell bioconversion of plant-derived or synthetic steviol and/or steviol glycosides in a cell culture medium of a recombinant host cell using:

(a) a polypeptide capable of synthesizing UTP from UDP having at least 60% sequence identity to the amino acid sequence set forth in SEQ ID NO:123;

(b) a polypeptide capable of converting glucose-6-phosphate to glucose-1-phosphate having at least 60% sequence identity to the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:119, or SEQ ID NO:143; or at least 55% sequence identity to the amino acid sequence set forth in SEQ ID NO:141, SEQ ID NO:145, or SEQ ID NO:147; and/or

(c) a polypeptide capable of synthesizing UDP-glucose from UTP and glucose-1-phosphate having at least 60% sequence identity to the amino acid sequence set forth in SEQ ID NO:121, SEQ ID NO:127; at least 55% sequence identity to the amino acid sequence set forth in SEQ ID NO:125, SEQ ID NO:129, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, or SEQ ID NO:139; or at least 70% sequence identity to the amino acid sequence set forth in SEQ ID NO:131, and

one or more of:

(d) a polypeptide capable of glycosylating steviol or a steviol glycoside at its C-13 hydroxyl group thereof;

(e) a polypeptide capable of beta 1,3 glycosylation of the C3' of the 13-O-glucose, 19-O-glucose, or both 13-O-glucose and 19-O-glucose of a steviol glycoside;

(f) a polypeptide capable of glycosylating steviol or a steviol glycoside at its C-19 carboxyl group thereof; and/or

(g) a polypeptide capable of beta 1,2 glycosylation of the C2' of the 13-O-glucose, 19-O-glucose, or both 13-O-glucose and 19-O-glucose of a steviol glycoside;

wherein at least one of the polypeptides is a recombinant polypeptide expressed in the recombinant host cell; and producing the one or more steviol glycosides or the steviol glycoside composition thereby.

36. The method of claim 35, wherein:

(d) the polypeptide capable of glycosylating steviol or a steviol glycoside at its C-13 hydroxyl group thereof comprises a polypeptide having at least 55% sequence identity to the amino acid sequence set forth in SEQ ID NO:7;

(e) the polypeptide capable of beta 1,3 glycosylation of the C3' of the 13-O-glucose, 19-O-glucose, or both 13-O-glucose and 19-O-glucose of a steviol glycoside comprises a polypeptide having at least 50% sequence identity to the amino acid sequence set forth in SEQ ID NO:9;

(f) the polypeptide capable of glycosylating steviol or a steviol glycoside at its C-19 carboxyl group thereof comprises a polypeptide having at least 55% sequence identity to the amino acid sequence set forth in SEQ ID NO:4;

(g) the polypeptide capable of beta 1,2 glycosylation of the C2' of the 13-O-glucose, 19-O-glucose, or both 13-O-glucose and 19-O-glucose of a steviol glycoside comprises a polypeptide having 80% or greater identity to the amino acid sequence set forth in SEQ ID NO:11; a polypeptide having 80% or greater identity to the amino acid sequence set forth in SEQ ID NO:13; or a polypeptide having at least 65% sequence identity to the amino acid sequence set forth in SEQ ID NO:16.

37. The method of any one of claims 22-36, wherein the recombinant host cell is a plant cell, a mammalian cell, an insect cell, a fungal cell, an algal cell or a bacterial cell.

38. The method of any one of claims 22-37, wherein the one or more steviol glycosides is, or the steviol glycoside composition comprises, steviol-13-*O*-glucoside (13-SMG), steviol-

1,2-Bioside, steviol-1,3-Bioside, steviol-19-*O*-glucoside (19-SMG), 1,2-stevioside, 1,3-stevioside (RebG), rubusoside, rebaudioside A (RebA), rebaudioside B (RebB), rebaudioside C (RebC), rebaudioside D (RebD), rebaudioside E (RebE), rebaudioside F (RebF), rebaudioside M (RebM), rebaudioside Q (RebQ), rebaudioside I (RebI), dulcoside A, and/or an isomer thereof.

39. A cell culture, comprising the recombinant host cell of any one of claims 1-21, the cell culture further comprising:

(a) the one or more steviol glycosides or the steviol glycoside composition produced by the recombinant host cell;

(b) glucose, fructose, sucrose, xylose, rhamnose, UDP-glucose, UDP-rhamnose, UDP-xylose, and/or N-acetyl-glucosamine; and

(c) supplemental nutrients comprising trace metals, vitamins, salts, YNB, and/or amino acids;

wherein the one or more steviol glycosides or the steviol glycoside composition is present at a concentration of at least 1 mg/liter of the cell culture;

wherein the cell culture is enriched for the one or more steviol glycosides or the steviol glycoside composition relative to a steviol glycoside composition from a Stevia plant and has a reduced level of Stevia plant-derived components relative to a plant-derived Stevia extract.

40. A cell culture, comprising the recombinant host cell of any one of claims 1-21, the cell culture further comprising:

(a) the one or more steviol glycosides or the steviol glycoside composition produced by the recombinant host cell;

(b) glucose, fructose, sucrose, xylose, rhamnose, UDP-glucose, UDP-rhamnose, UDP-xylose, and/or N-acetyl-glucosamine; and

(c) supplemental nutrients comprising trace metals, vitamins, salts, YNB, and/or amino acids;

wherein UDP-glucose is present in the cell culture at a concentration of at least 100 μM;

wherein the cell culture is enriched for UGP-glucose relative to a steviol glycoside composition from a Stevia plant and has a reduced level of Stevia plant-derived components relative to a plant-derived Stevia extract.

41. A cell lysate from the recombinant host cell of any one of claims 1-21 grown in the cell culture, comprising:

(a) the one or more steviol glycosides or the steviol glycoside composition produced by the recombinant host cell;

(b) glucose, fructose, sucrose, xylose, rhamnose, UDP-glucose, UDP-rhamnose, UDP-xylose, and/or N-acetyl-glucosamine; and/or

(c) supplemental nutrients comprising trace metals, vitamins, salts, yeast nitrogen base, YNB, and/or amino acids;

wherein the one or more steviol glycosides or the steviol glycoside composition produced by the recombinant host cell is present at a concentration of at least 1 mg/liter of the cell culture.

42. One or more steviol glycosides produced by the recombinant host cell of any one of claims 1-21;

wherein the one or more steviol glycosides produced by the recombinant host cell are present in relative amounts that are different from a steviol glycoside composition from a Stevia plant and have a reduced level of Stevia plant-derived components relative to a plant-derived Stevia extract.

43. One or more steviol glycosides produced by the method of any one of claims 22-38;

wherein the one or more steviol glycosides produced by the recombinant host cell are present in relative amounts that are different from a steviol glycoside composition from a Stevia plant and have a reduced level of Stevia plant-derived components relative to a plant-derived Stevia extract.

44. A sweetener composition, comprising the one or more steviol glycosides of claim 42 or 43.

45. A food product comprising, the sweetener composition of claim 44.

46. A beverage or a beverage concentrate, comprising the sweetener composition of claim 44.

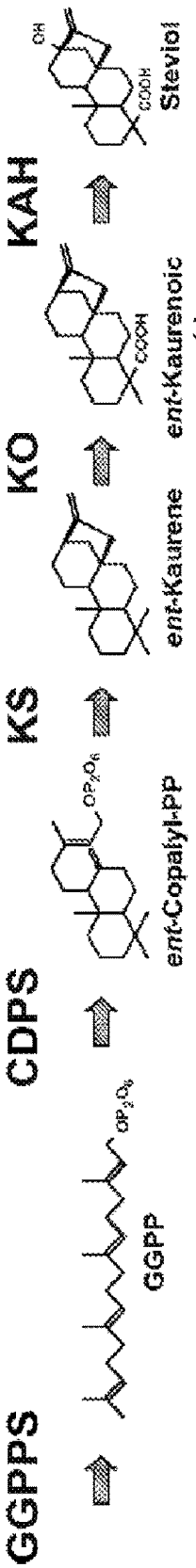
GGPPS
GGPP
CDPS
ent-Copalyl-PP
KS
ent-Kaurene
KO
ent-Kaurenoic acid
KAH
Steviol
COOH
OH

Figure 1

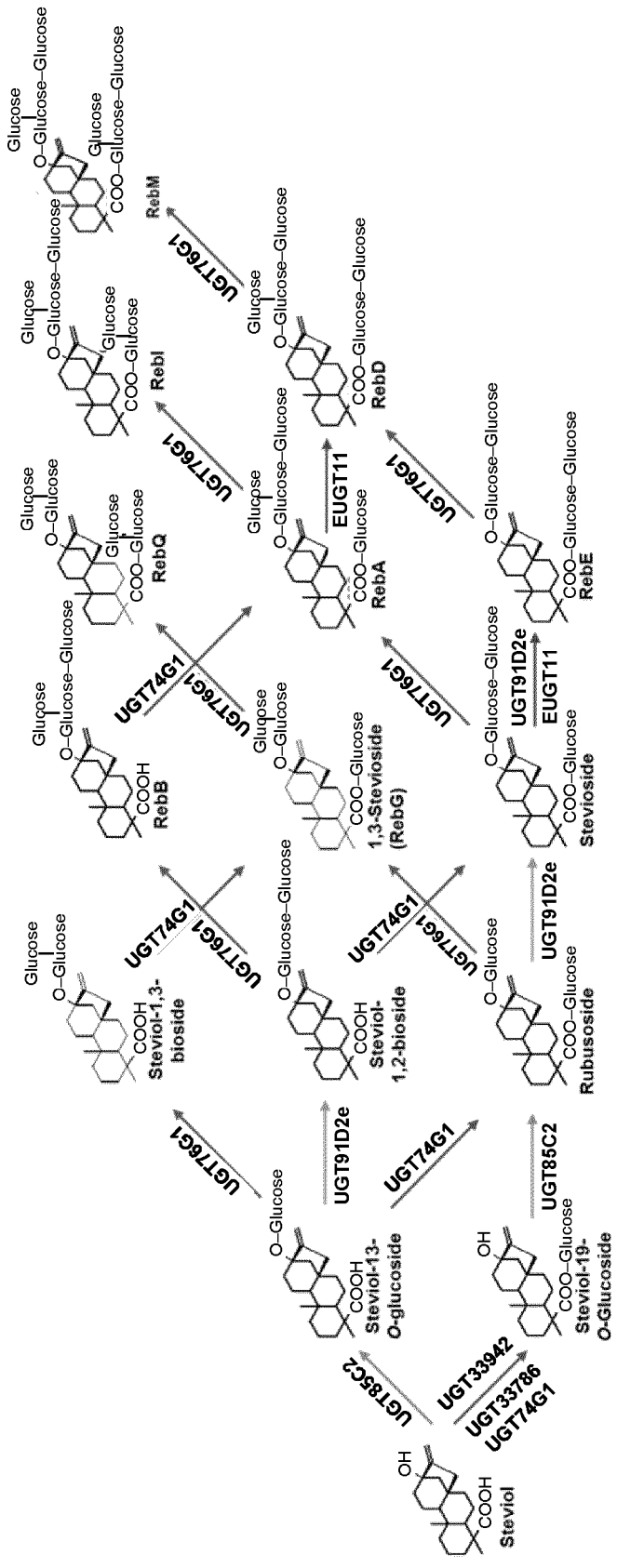
Steviol
COOH
OH
UGT85C2
UGT33942
UGT33786
UGT74G1
Steviol-13-O-glucoside
O-Glucose
COOH
COO–Glucose
Steviol-19-O-Glucoside
OH
UGT91D2e
UGT74G1
UGT85C2
Steviol-1,2-bioside
Rubusoside
UGT76G1
UGT74G1
UGT91D2e
Steviol-1,3-bioside
1,3-Stevioside (RebG)
Stevioside
UGT76G1
UGT74G1
UGT91D2e
EUGT11
RebB
RebA
RebE
UGT76G1
UGT74G1
EUGT11
RebQ
RebD
UGT76G1
RebI
RebM
UGT76G1
Glucose
O–Glucose–Glucose
COO–Glucose–Glucose

Figure 2

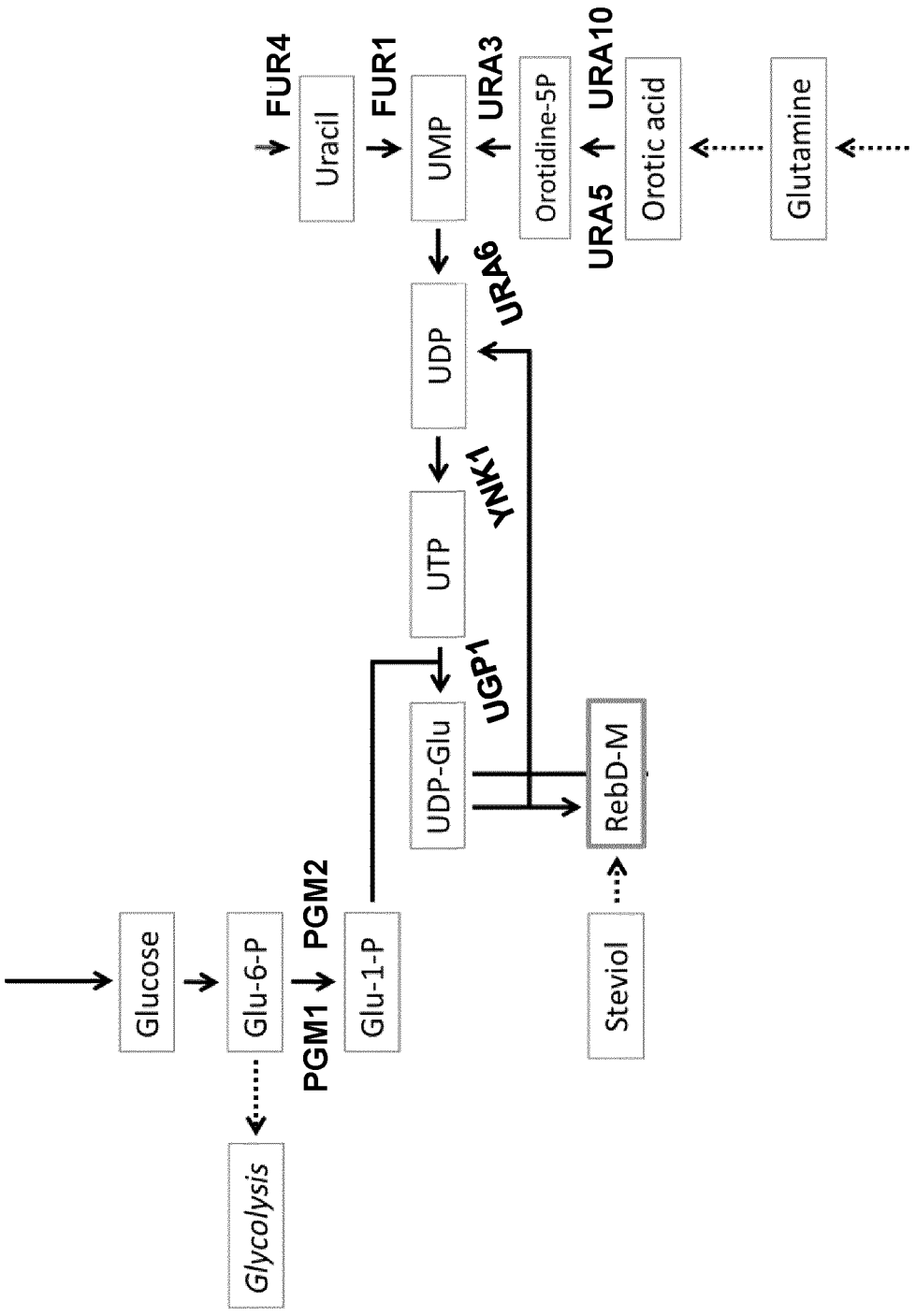
Glucose
Glu-6-P
Glycolysis
PGM1
PGM2
Glu-1-P
UDP-Glu
UGP1
UTP
YNK1
UDP
URA6
UMP
FUR1
Uracil
FUR4
Orotidine-5P
URA3
Orotic acid
URA5
URA10
Glutamine
RebD-M
Steviol

Figure 3

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2017/059028

A. CLASSIFICATION OF SUBJECT MATTER
INV. A23L27/30 A23L2/60 C12N15/63
ADD.

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
A23L C12N

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

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

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	KR 2015 0000258 A (KOREA RES INST OF BIOSCIENCE [KR]) 2 January 2015 (2015-01-02)	42-46
Y	the whole document, in particular Figure 1 and claims	1-46

X	WO 2015/014969 A1 (DSM IP ASSETS BV [NL]) 5 February 2015 (2015-02-05)	42-46
Y	the whole document, in particular the claims	1-46

X	WO 2016/038095 A2 (EVOLVA SA [CH]; ROBERTSEN HELENE LUNDE [DK]; ANDERSEN IBEN NORDMARK [D) 17 March 2016 (2016-03-17)	42-46
Y	the whole document, in particular the claims	1-46

	-/--	

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

[X] See patent family annex.

* Special categories of cited documents :

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

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

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

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

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

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
2 June 2017	27/06/2017

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

2

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

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2014/122328 A1 (EVOLVA SA [CH]) 14 August 2014 (2014-08-14)	42-46
Y	the whole document, in particular the claims	1-46

X	WO 2014/122227 A2 (EVOLVA SA [CH]) 14 August 2014 (2014-08-14)	42-46
Y	the whole document in particular the claims	1-46

A	BRANDLE ET AL: "Steviol glycoside biosynthesis", PHYTOCHEMISTRY, PERGAMON PRESS, GB, vol. 68, no. 14, 1 July 2007 (2007-07-01), pages 1855-1863, XP022145443, ISSN: 0031-9422, DOI: 10.1016/J.PHYTOCHEM.2007.02.010	1-46

A	SUDESH KUMAR YADAV ET AL: "Steviol Glycosides from Stevia: Biosynthesis Pathway Review and their Application in Foods and Medicine", CRITICAL REVIEWS IN FOOD SCIENCE AND NUTRITION, vol. 52, no. 11, 1 November 2012 (2012-11-01), pages 988-998, XP055351385, USA ISSN: 1040-8398, DOI: 10.1080/10408398.2010.519447	1-46

2

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/EP2017/059028

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
KR 20150000258 A	02-01-2015	NONE	
WO 2015014969 A1	05-02-2015	AU 2014298430 A1	11-02-2016
		CA 2917615 A1	05-02-2015
		CN 105658081 A	08-06-2016
		EP 3027048 A1	08-06-2016
		US 2016213039 A1	28-07-2016
		WO 2015014969 A1	05-02-2015
WO 2016038095 A2	17-03-2016	AU 2015314251 A1	16-03-2017
		CA 2960693 A1	17-03-2016
		KR 20170052647 A	12-05-2017
		SG 11201701677U A	27-04-2017
		WO 2016038095 A2	17-03-2016
WO 2014122328 A1	14-08-2014	AU 2014213918 A1	20-08-2015
		CA 2900882 A1	14-08-2014
		CN 105189771 A	23-12-2015
		EP 2954061 A1	16-12-2015
		JP 2016506743 A	07-03-2016
		KR 20150128705 A	18-11-2015
		SG 11201506127W A	29-09-2015
		US 2015361476 A1	17-12-2015
		WO 2014122328 A1	14-08-2014
WO 2014122227 A2	14-08-2014	AU 2014214004 A1	30-07-2015
		CA 2899276 A1	14-08-2014
		CN 105051195 A	11-11-2015
		EP 2954058 A2	16-12-2015
		JP 2016506739 A	07-03-2016
		KR 20150115002 A	13-10-2015
		SG 11201505597S A	28-08-2015
		US 2016186225 A1	30-06-2016
		WO 2014122227 A2	14-08-2014