compounds that contribute to off-flavors. These off-flavors can be more or less problematic depending on the food system or application of choice. Potential contaminants 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 triterpene, centauredin, quercitin, epi-alpha-cadinol, carophyllenes and derivatives, beta-pinene, beta-sitosterol, and gibberellin.

SUMMARY

Provided herein is a recombinant host, such as a microorganism, comprising one or more biosynthesis genes whose expression results in production of steviol. Such genes include a gene encoding a copalyl diphosphate synthase, a gene encoding a kaurene synthase, a gene encoding a kaurene oxidase; and a gene encoding a steviol synthetase. The recombinant host can include a gene encoding a bifunctional copalyl diphosphate synthase and kaurene synthase, in place of the genes encoding copalyl diphosphate synthase and kaurene synthase. At least one of the genes is a recombinant gene. In some examples the recombinant host further comprises a gene encoding a geranylgeranyl diphosphate synthase. The recombinant host can further comprise a gene encoding a truncated HMG-CoA reductase and/or a gene encoding a CPR. The expression of one or more of the genes can be inducible.

In one aspect, the present disclosure refers to a recombinant host cell, comprising a recombinant gene encoding a recombinant 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 and having at least 90% sequence identity to the amino acid sequence set forth in any one of SEQ ID NOs:5, 10, 76, 78, or 150;

wherein the polypeptide is capable of transferring a sugar moiety to the C2' of a glucose in the steviol glycoside;

wherein the steviol glycoside is steviol-13-*O*-glucoside, rubusoside, stevioside or rebaudioside A, and

wherein the recombinant host cell is capable of producing stevioside,

rebaudioside E, rebaudioside D, steviol-1,2-bioside, steviol-1,2-xylobioside, steviol-1,2-rhamnobioside, 1,2-stevioxyloside, an isomer thereof, and/or a composition thereof upon transfer of the sugar moiety by the polypeptide.

In another aspect, the present disclosure refers to a method of producing steviol, one or more steviol glycosides or a steviol glycoside composition in a cell culture, comprising growing the recombinant host cell as disclosed herein under conditions in which one or more of the genes are expressed;

wherein steviol, the steviol glycoside, or the steviol glycoside composition is produced by the recombinant host cell;

wherein the one or more steviol glycosides is, or the steviol glycoside composition comprises, stevioside, rebaudioside E, rebaudioside D, steviol-1,2-bioside, steviol-1,2-xylobioside, steviol-1,2-rhamnobioside, 1,2-stevioxyloside, and/or an isomer thereof.

In another aspect, the present disclosure refers to an *in vitro* method for producing one or more steviol glycosides or a steviol glycoside composition, comprising adding 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 and having at least 90% sequence identity to the amino acid sequence set forth in any one of SEQ ID NOs:5, 10, 76, 78, or 150;

(b) a polypeptide capable of glycosylating steviol or a steviol glycoside at its C-13 hydroxyl group,

wherein the polypeptide comprises:

(i) a polypeptide having at least 90% sequence identity to the amino acid sequence set forth in SEQ ID NO:3; or

(II) a polypeptide having one or more amino acid substitutions of residues 9, 10, 13, 15, 21, 27, 60, 65, 71, 87, 91, 220, 243, 270, 289, 298, 334, 336, 350, 368, 389, 394, 397, 418, 420, 440, 441, 444, and 471 of SEQ ID NO:3;

(c) a polypeptide capable of beta 1,3 glycosylation of the C3' of the 13-*O*-glucose of a steviol glycoside,

wherein the polypeptide comprises:

(i) a polypeptide having at least 90% sequence identity to the amino acid sequence set forth in SEQ ID NO:7; or

(ii) a polypeptide having one or more amino acid substitutions of residues 29, 74, 87, 91, 116, 123, 125, 126, 130, 145, 192, 193, 194, 196, 198, 199, 200, 203, 204, 205, 206, 207, 208, 266, 273, 274, 284, 285, 291, 330, 331, and 346 of SEQ ID NO:7; and/or

(d) a polypeptide capable of glycosylating steviol or a steviol glycoside at its C-19 carboxyl group,

wherein the polypeptide comprises a polypeptide having at least 90% sequence identity to the amino acid sequence set forth in SEQ ID NO:1;

and steviol, steviol-13-*O*-glucoside, rubusoside, stevioside or rebaudioside A to a reaction mixture; and

producing the one or more steviol glycosides or the steviol glycoside composition;

wherein the steviol glycoside comprises steviol-13-*O*-glucoside, rubusoside, stevioside or rebaudioside A; and

wherein the steviol glycoside composition comprises stevioside, rebaudioside E, rebaudioside D, steviol-1,2-bioside, steviol-1,2-xylobioside, steviol-1,2-rhamnobioside, and/or 1,2-stevioxyloside.

In yet another aspect, the present disclosure refers to one or more steviol glycosides or a steviol glycoside composition produced by the method as disclosed herein.

In yet another aspect, the present disclosure refers to a sweetener composition, comprising the one or more steviol glycosides or the steviol glycoside composition as disclosed herein.

In yet another aspect, the present disclosure refers to a food product, comprising the sweetener composition as disclosed herein.

In yet another aspect, the present disclosure refers to a beverage or a beverage concentrate, comprising the sweetener composition as disclosed herein.

In yet another aspect, the present disclosure refers to a cell culture, comprising the recombinant host cell as disclosed herein, the cell culture further comprising:

(a) stevioside, rebaudioside E, rebaudioside D, steviol-1,2-bioside, steviol-1,2-xylobioside, steviol-1,2-rhamnobioside, 1,2-stevioxyloside, an isomer thereof, and/or the composition thereof produced by the recombinant host cell;

(b) glucose, fructose, sucrose, xylose, rhamnose, uridine diphosphate (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 one or more of stevioside, rebaudioside E, rebaudioside D, steviol-1,2-bioside, steviol-1,2-xylobioside, steviol-1,2-rhamnobioside, 1,2-stevioxyloside, an isomer thereof alone, and/or in the composition thereof is present at a concentration of at least 1 mg/liter of the cell culture;

wherein the cell culture is enriched for stevioside, rebaudioside E, rebaudioside D, steviol-1,2-bioside, steviol-1,2-xylobioside, steviol-1,2-rhamnobioside, 1,2-stevioxyloside, an isomer thereof, and/or the composition thereof relative to a steviol glycoside composition from a Stevia plant and has a reduced level of Stevia plant-derived components relative to a steviol glycoside composition obtained from a plant-derived Stevia extract.

In yet another aspect, the present disclosure refers to a cell lysate from the recombinant host as disclosed herein grown in the cell culture, comprising:

(a) stevioside, rebaudioside E, rebaudioside D, steviol-1,2-bioside, steviol-1,2-xylobioside, steviol-1,2-rhamnobioside, 1,2-stevioxyloside, an isomer thereof, and/or the composition thereof produced by the recombinant host cell;

(b) glucose, fructose, sucrose, xylose, rhamnose, uridine diphosphate (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 one or more of stevioside, rebaudioside E, rebaudioside D, steviol-1,2-bioside, steviol-1,2-xylobioside, steviol-1,2-rhamnobioside, 1,2-stevioxyloside, an isomer thereof alone, and/or in the composition thereof produced by the recombinant host cell is present at a concentration of at least 1 mg/liter of the cell culture.

In yet another aspect, the present disclosure refers to a reaction mixture, comprising:

(a) one or more of stevioside, rebaudioside E, rebaudioside D, steviol-1,2-bioside, steviol-1,2-xylobioside, steviol-1,2-rhamnobioside, 1,2-stevioxyloside, an isomer thereof, and/or the steviol glycoside composition thereof produced in the reaction mixture,

(b) 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 and having at least 90% sequence identity to the amino acid sequence set forth in any one of SEQ ID NOs:5, 10, 76, 78, or 150;

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

(d) reaction buffer and/or salts.

In one example, this document features a recombinant host that includes a recombinant gene encoding a UGT91D2 polypeptide (e.g., a UGT91D2e or UGT91D2m polypeptide). The UGT91D2 polypeptide can have at least 90% identity (e.g., at least 95% or 99% identity) to the amino acid sequence set forth in SEQ ID NO:5. The UGT91D2 polypeptide can include at least one amino acid substitution at residues 1-19, 27-38, 44-87, 96-120, 125-141, 159-184, 199-202, 215-380, or 387-473 of SEQ ID NO:5. For example, the UGT91D2 polypeptide can include an amino acid substitution at one or more residues selected from the group consisting of residues 30, 93, 99, 122, 140, 142, 148, 153, 156, 195, 196, 199, 206, 207, 211, 221, 286, 343, 427, and 438 of SEQ ID NO:5. In one example, the UGT91D2 polypeptide includes an arginine at residue 206, a cysteine at residue 207, and an arginine at

residue 343 relative to SEQ ID NO:5. In one example, the UGT91D2 polypeptide includes a phenylalanine at residue 30, a glutamine at residue 93, a valine at residue 99, a phenylalanine at residue 122, a tyrosine at residue 140, a cysteine at residue 142, a threonine at residue 148, an alanine at residue 153, a serine at residue 156, a methionine at residue 195, a glutamic acid at residue 196, a glutamic acid at residue 199, a methionine at residue 211, a phenylalanine at residue 221, an alanine at residue 286, an asparagine at residue 427, or an alanine at residue 438 relative to SEQ ID NO:5. The polypeptide can have the amino acid sequence of SEQ ID NO:5 or SEQ ID NO:95.

A host described herein further can include a recombinant gene encoding a UGT85C polypeptide having at least 90% identity to the amino acid sequence set forth in SEQ ID NO:3. For example, the UGT85C polypeptide can include one or more amino acid substitutions at residues 9, 10, 13, 15, 21, 27, 60, 65, 71, 87, 91, 220, 243, 270, 289, 298, 334, 336, 350, 368, 389, 394, 397, 418, 420, 440, 441, 444, and 471 of SEQ ID NO:3.

A host described herein further can include a recombinant gene encoding a UGT76G polypeptide having at least 90% identity to the amino acid sequence set forth in SEQ ID NO:7. For example, the UGT76G polypeptide can have one or more amino acid substitutions at residues 29, 74, 87, 91, 116, 123, 125, 126, 130, 145, 192, 193, 194, 196, 198, 199, 200, 203, 204, 205, 206, 207, 208, 266, 273, 274, 284, 285, 291, 330, 331, and 346 of SEQ ID NO:7.

This document also features a recombinant host that includes a recombinant gene encoding a UGT85C polypeptide having at least 90% identity to the amino acid sequence set forth in SEQ ID NO:3, and having one or more amino acid substitutions at residues 9, 10, 13, 15, 21, 27, 60, 65, 71, 87, 91, 220, 243, 270, 289, 298, 334, 336, 350, 368, 389, 394, 397, 418, 420, 440, 441, 444, and 471 of SEQ ID NO:3. For example, the UGT85C polypeptide can include substitutions at residues 13, 15, 60, 270, 289, and 418 of SEQ ID NO:3. For example, the UGT85C polypeptide can include a) substitutions at residues 13, 60, and 270 of SEQ ID NO:3; b) substitutions at residues 60 and 87 of SEQ ID NO:3; c) substitutions at residues 65, 71, 220, 243, and 270 of SEQ ID NO:3; d) substitutions at residues 65, 71, 220, 243, 270, and 441 of SEQ ID NO:3; e) substitutions at residues 65, 71, 220, 389, and 394 of SEQ ID

NO:3; f) substitutions at residues 65, 71, 270, and 289 of SEQ ID NO:3; g) substitutions at residues 15 and 65 of SEQ ID NO:3; h) substitutions at residues 65 and 270 of SEQ ID NO:3; i) substitutions at residues 65 and 440 of SEQ ID NO:3; j) substitutions at residues 65 and 441 of SEQ ID NO:3; k) substitutions at residues 65 and 418 of SEQ ID NO:3; l) substitutions at residues 220, 243, 270, and 334 of SEQ ID NO:3; or m) substitutions at residues 270 and 289 of SEQ ID NO:3.

In another example, this document features a recombinant host that includes a recombinant gene encoding a UGT76G polypeptide having at least 90% identity to the amino acid sequence set forth in SEQ ID NO:7, and having one or more amino acid substitutions at residues 29, 74, 87, 91, 116, 123, 125, 126, 130, 145, 192, 193, 194, 196, 198, 199, 200, 203, 204, 205, 206, 207, 208, 266, 273, 274, 284, 285, 291, 330, 331, and 346. For example, the UGT76G polypeptide can have a) substitutions at amino acid residues 74 , 87, 91, 116, 123, 125, 126, 130, 145, 192, 193, 194, 196, 198, 199, 200, 203, 204, 205, 206, 207, 208, and 291; b) substitutions at residues 74 , 87, 91, 116, 123, 125, 126, 130, 145, 192, 193, 194, 196, 198, 199, 200, 203, 204, 205, 206, 207, 208, 266, 273, 274, 284, 285, and 291; or c) substitutions at residues 74 , 87, 91, 116, 123, 125, 126, 130, 145, 192, 193, 194, 196, 198, 199, 200, 203, 204, 205, 206, 207, 208, 266, 273, 274, 284, 285, 291, 330, 331, and 346.

Any of the hosts described herein further can include a gene encoding a UGT74G1 polypeptide (e.g., a recombinant gene encoding a UGT74G1 polypeptide).

Any of the hosts described herein further can include one or more of: (i) a gene encoding a geranylgeranyl diphosphate synthase; (ii) a gene encoding a bifunctional copalyl diphosphate synthase and kaurene synthase, or a gene encoding a copalyl diphosphate synthase and a gene encoding a kaurene synthase; (iii) a gene encoding a kaurene oxidase; (iv) a gene encoding a steviol synthetase; (v) a gene encoding a truncated HMG-CoA; (vi) a gene encoding a CPR; (vii) a gene encoding a rhamnose synthetase; (viii) a gene encoding a UDP-glucose dehydrogenase; and (ix) a gene encoding a UDP-glucuronic acid decarboxylase. At least one of the genes of (i), (ii), (iii), (iv), (v), (vi), (vii), (viii), or (ix) can be a recombinant gene. In some examples, each of the genes of (i), (ii), (iii), and (iv) is a recombinant gene.

This document also features an isolated nucleic acid encoding a polypeptide having at least 90% sequence identity (e.g., at least 95% or 99% sequence identity) to

the amino acid sequence set forth in SEQ ID NO:5. The polypeptide can include at least one amino acid substitution at residues 1-19, 27-38, 44-87, 96-120, 125-141, 159-184, 199-202, 215-380, or 387-473 of SEQ ID NO:5. The polypeptide can include an amino acid substitution at one or more residues selected from the group consisting of residues 30, 93, 99, 122, 140, 142, 148, 153, 156, 195, 196, 199, 206, 207, 211, 221, 286, 343, 427, and 438 of SEQ ID NO:5. The polypeptide can include an arginine at residue 206, a cysteine at residue 207, and an arginine at residue 343 of SEQ ID NO:5. In some examples, the polypeptide includes a phenylalanine at residue 30, a glutamine at residue 93, a valine at residue 99, a phenylalanine at residue 122, a tyrosine at residue 140, a cysteine at residue 142, a threonine at residue 148, an alanine at residue 153, a serine at residue 156, a methionine at residue 195, a glutamic acid at residue 196, a glutamic acid at residue 199, a methionine at residue 211, a phenylalanine at residue 221, an alanine at residue 286, an asparagine at residue 427, or an alanine at residue 438 of SEQ ID NO:5.

In another example, this document features an isolated polypeptide having an amino acid sequence with at least 90% identity to the amino acid sequence of SEQ ID NO:5.

This document also features a recombinant host that includes (i) a gene encoding a geranylgeranyl diphosphate synthase; (ii) a gene encoding a bifunctional copalyl diphosphate synthase and kaurene synthase, or a gene encoding a copalyl diphosphate synthase and a gene encoding a kaurene synthase; (iii) a gene encoding a kaurene oxidase; and (iv) a gene encoding a steviol synthetase; wherein at least one of said genes. The host can produce steviol when cultured under conditions in which each of the genes is expressed, and can accumulate to at least 1 mg/L in the culture medium. The geranylgeranyl diphosphate synthase can have greater than 90 % sequence identity to one of the amino acid sequences set forth in SEQ ID NOs: 121-128. The copalyl diphosphate synthase can have greater than 90 % sequence identity to one of the amino acid sequences set forth in SEQ ID NOs:129-131. The kaurene synthase can have greater than 90 % sequence identity to one of the amino acid sequences set forth in 132-135. The kaurene oxidase can have greater than 90 % sequence identity to one of the amino acid sequences set forth in 138-141. The steviol synthetase can have greater than 90 % sequence identity to one of the amino acid

In another example, this document features a method of producing steviol or a steviol glycoside. The method includes growing a host described herein in a culture medium, under conditions in which the genes are expressed; and recovering the steviol or steviol glycoside produced by the host. The growing step can include inducing expression of one or more of the genes. The steviol or steviol glycoside is selected from the group consisting of steviol-13-O-glucoside, steviol-19-O-glucoside, rubusoside, rebaudioside A, rebaudioside B, rebaudioside C, rebaudioside D, rebaudioside E, rebaudioside F, and dulcoside A.

Also provided herein is a method of producing steviol or a steviol glycoside. The method includes growing a microorganism in a culture medium, under conditions in which a geranylgeranyl diphosphate synthase, copalyl diphosphate synthase, kaurene synthase, kaurene oxidase, kaurenoic acid 13-hydroxylase gene and optionally a UGT74G1 and/or a UGT85C2 gene are expressed, and recovering the steviol or steviol glycoside produced by the microorganism. The microorganism can be a *Saccharomyces* spp. In some examples, the growing step comprises inducing expression of one or more of the geranylgeranyl diphosphate synthase, copalyl diphosphate synthase, kaurene synthase, kaurene oxidase, kaurenoic acid 13-hydroxylase, UGT74G1 and UGT85C2 genes. In some examples, the recovering step comprises purifying the steviol or steviol glycoside from the culture medium by HPLC. The steviol or steviol glycoside can be steviol, rubusoside, rebaudioside C, rebaudioside F, or dulcoside A.

Also provided herein is a recombinant *Saccharomyces* strain, comprising one or more biosynthesis genes whose expression results in production of ent-kaurene. The biosynthesis genes include a gene encoding a bifunctional copalyl diphosphate synthase and kaurene synthase, or a gene encoding a copalyl diphosphate synthase and a gene encoding a kaurene synthase. The strain produces ent-kaurene upon expression of the copalyl diphosphate synthase and the kaurene synthase.

In another example, this document features an isolated nucleic acid having greater than 90 % sequence identity (e.g., greater than 95% or 99% sequence identity) to one of the nucleotide sequences set forth in SEQ ID NOs: 18-25, 34-36, 4-43, 48, 49, 52-55, 60-64, 70-72, 77, or 79.

This document also features a recombinant host that includes (i) a gene encoding a UGT74G1; (ii) a gene encoding a UGT85C2; (iii) a gene encoding a UGT76G1; and (iv) a gene encoding a UGT91D2, wherein at least one of said genes is a recombinant gene. In some examples, each of the genes is a recombinant gene. The host can produce at least one steviol glycoside when cultured under conditions in which each of the genes is expressed. The host further can include (a) a gene encoding a bifunctional copalyl diphosphate synthase and kaurene synthase, or a gene encoding a copalyl diphosphate synthase and a gene encoding a kaurene synthase; (b) a gene encoding a kaurene oxidase; (c) a gene encoding a steviol synthetase; and (d) a gene encoding a geranylgeranyl diphosphate synthase. The steviol glycoside can be rebaudioside A, rebaudioside D or rebaudioside E. This document also features a steviol glycoside composition produced by such a host. The composition can have greater than 4% rebaudioside D by weight of total steviol glycosides and a reduced level of stevia plant-derived contaminants relative to a stevia extract. The composition can have greater than 4% rebaudioside E by weight of total steviol glycosides and a reduced level of stevia plant-derived contaminants relative to a stevia extract.

Also featured herein is an isolated nucleic acid encoding a polypeptide having greater than 90 % sequence identity to the amino acid sequences of UGT91D2e and UGT91D2m, excluding the amino acid sequence of UGT91D2m, as well as the isolated polypeptides having greater than 90 % sequence identity to the amino acid sequence of UGT91D2e or UGT91D2m, excluding the amino acid sequence of UGT91D2m.

This document also features steviol glycoside composition produced by the host described herein. The composition having reduced levels of stevia plant-derived contaminants relative to a stevia extract.

In another example, this document features a recombinant host. The host includes (i) a recombinant gene encoding a UGT91D2; (ii) a recombinant gene encoding a UGT74G1; (iii) a recombinant gene encoding a UGT85C2; (iv) a recombinant gene encoding a UGT76G1; and (v) a gene encoding a rhamnose synthetase, wherein the host produces at least one steviol glycoside when cultured under conditions in which each of the genes is expressed. The host further can

include (a) a gene encoding a bifunctional copalyl diphosphate synthase and kaurene synthase, or a gene encoding a copalyl diphosphate synthase and a gene encoding a kaurene synthase; (b) a gene encoding a kaurene oxidase; (c) a gene encoding a steviol synthetase; and (d) a gene encoding a geranylgeranyl diphosphate synthase. The steviol glycoside can be rebaudioside C or dulcoside A. This document also features a steviol glycoside composition produced by such a host. The composition has greater than 15% rebaudioside C by weight of total steviol glycosides and a reduced level of stevia plant-derived contaminants relative to a stevia extract. A steviol glycoside composition produced by such a host also is featured. The composition can have greater than 15% dulcoside A by weight of total steviol glycosides and a reduced level of stevia plant-derived contaminants relative to a stevia extract.

This document also features a recombinant host. The host includes (i) a recombinant gene encoding a UGT91D2; (ii) a recombinant gene encoding a UGT74G1; (iii) a recombinant gene encoding a UGT85C2; (iv) a recombinant gene encoding a UGT76G1; (v) a gene encoding a UDP-glucose dehydrogenase; and (vi) a gene encoding a UDP-glucuronic acid decarboxylase, wherein the host produces at least one steviol glycoside when cultured under conditions in which each of the genes is expressed. The host further can include (a) a gene encoding a bifunctional copalyl diphosphate synthase and kaurene synthase, or a gene encoding a copalyl diphosphate synthase and a gene encoding a kaurene synthase; (b) a gene encoding a kaurene oxidase; (c) a gene encoding a steviol synthetase; and (d) a gene encoding a geranylgeranyl diphosphate synthase. The steviol glycoside can be rebaudioside F. This document also features a steviol glycoside composition produced by such hosts. The composition can have greater than 4% rebaudioside F by weight of total steviol glycosides and a reduced level of stevia plant-derived contaminants relative to a stevia extract.

In another example, this document features a method of producing a steviol glycoside composition. The method includes growing a host described herein in a culture medium, under conditions in which each of the genes is expressed; and recovering the steviol glycoside composition produced by the host, wherein the recovered composition is enriched for rebaudioside A, rebaudioside C, rebaudioside

D, rebaudioside E, rebaudioside F or dulcoside A relative to the steviol glycoside composition of a wild-type *Stevia* plant. The steviol glycoside composition produced by the host (e.g., microorganism) can have a reduced level of stevia plant-derived contaminants relative to a stevia extract.

This document also features a food product that includes a steviol glycoside composition enriched for rebaudioside A, rebaudioside C, rebaudioside D, rebaudioside E, rebaudioside F or dulcoside A relative to the steviol glycoside composition of a wild-type *Stevia* plant.

In another example, this document features a method of identifying whether a polymorphism is associated with variation in a trait. The method includes determining whether one or more genetic polymorphisms in a population of plants is associated with the locus for a polypeptide set forth in SEQ ID NO:5 and functional homologs thereof; and measuring the correlation between variation in the trait in plants of the population and the presence of the one or more genetic polymorphisms in plants of the population, thereby identifying whether or not the one or more genetic polymorphisms are associated with variation in the trait.

In yet another example, this document features a mehod of making a plant line. The method includes determining whether one or more genetic polymorphisms in a population of plants is associated with the locus for a polypeptide set forth in SEQ ID NO:5 and functional homologs thereof; identifying one or more plants in the population in which the presence of at least one of the genetic polymorphisms is associated with variation in a trait; crossing one or more of the identified plants with itself or a different plant to produce seed; crossing at least one progeny plant grown from the seed with itself or a different plant; and repeating the crossing steps for an additional 0-5 generations to make said plant line, wherein at least one of the genetic polymorphisms is present in the plant line.

This document also features a method for transferring a second sugar moiety to the C-2' of a glucose in a steviol glycoside. The method includes contacting the steviol glycoside with a UGT91D2 polypeptide and a UDP-sugar under suitable reaction conditions for the transfer of the second sugar moiety to the steviol glycoside. The UGT91D2 polypeptide can have at least 90% sequence identity (e.g., at least 95%

or 99%) to the amino acid sequence set forth in SEQ ID NO:5. The UGT91D2 polypeptide can inlcude at least one amino acid substitution at residues 1-19, 27-38, 44-87, 96-120, 125-141, 159-184, 199-202, 215-380, or 387-473 of SEQ ID NO:5. The UGT91D2 polypeptide can include an amino acid substitution at one or more residues selected from the group consisting of residues 30, 93, 99, 122, 140, 142, 148, 153, 156, 195, 196, 199, 206, 207, 211, 221, 286, 343, 427, and 438 of SEQ ID NO:5. The steviol glycoside can be selected from the group consisting of steviol-13-*O*-glucoside, rubusoside, stevioside, and Rebaudioside A. The steviol glycoside can be rubusoside and the second sugar moiety is glucose, and stevioside is produced upon transfer of the second glucose moiety. The steviol glycoside can be stevioside and the second sugar moiety can be glucose, and Rebaudioside E is produced upon transfer of the second glucose moiety. The steviol glycoside can be stevioside, wherein stevioside is contacted with the UGT91D2 polypeptide and a UGT76G1 polypeptide under suitable reaction conditions to produce Rebaudioside D. The steviol glycoside can be steviol-13-*O*-glucoside and steviol-1,2 bioside is produced upon transfer of said second glucose moiety. The steviol glycoside can be steviol-13-*O*-glucoside and steviol-1,2-xylobioside is produced upon transfer of the second sugar moiety. The steviol glycoside can be steviol-13-*O*-glucoside and steviol-1,2-rhamnobioside can be produced upon transfer of the second sugar moiety. The steviol glycoside can be Rebaudioside A, and Rebaudioside D is produced upon transfer of a second glucose moiety.

In another example, this document features a method of determining the presence of a polynucleotide in a *Stevia* plant. The method includes contacting at least one probe or primer pair with nucleic acid from the *Stevia* plant, wherein the probe or primer pair is specific for a polynucleotide that encodes a UGT polypeptide, wherein the UGT polypeptide has at least 90% sequence identity to SEQ ID NO: 5, SEQ ID NO: 1, SEQ ID NO: 3 or SEQ ID NO:7 and determining whether or not the polynucleotide is present in said *Stevia* plant.

This document also features a kit for genotyping a *Stevia* biological sample. The kit includes a primer pair that specifically amplifies, or a probe that specifically hybridizes to, a polynucleotide that encodes a UGT polypeptide having at least 90% sequence identity to SEQ ID NO: 5, SEQ ID NO: 1, SEQ ID NO: 3 or SEQ ID NO:7.

diphosphocytidyl-2-C-methyl-D-erythritol synthase (CMS), and/or a gene encoding a 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase (CMK), and/or a gene encoding a 4-diphosphocytidyl-2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase (MCS), and/or a gene encoding a 1-hydroxy-2-methyl-2(E)-butenyl 4-diphosphate synthase (HDS), and/or a gene encoding a 1-hydroxy-2-methyl-2(E)-butenyl 4-diphosphate reductase (HDR).

The recombinant microorganisms described above can further comprise a gene encoding a acetoacetyl-CoA thiolase, and/or a gene encoding a truncated HMG-CoA reductase, and/or a gene encoding a mevalonate kinase, and/or a gene encoding a phosphomevalonate kinase, and/or a gene encoding a mevalonate pyrophosphate decarboxylase.

Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which the disclosure belongs. Although methods and materials similar or equivalent to those described herein can be used to practice the disclosure, suitable methods and materials are described below. All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. In case of conflict, the present specification, including definitions, will control. In addition, the materials, methods, and examples are illustrative only and are not intended to be limiting. Other features and advantages of the disclosure will be apparent from the following detailed description. Applicants reserve the right to alternatively claim any disclosed disclosure using the transitional phrase "comprising," "consisting essentially of," or "consisting of," according to standard practice in patent law.

DESCRIPTION OF DRAWINGS

FIG. 1 is a scheme illustrating the biosynthesis of steviol from geranylgeranyl diphosphate.

FIGS. 2A- D show representative pathways for the biosynthesis of steviol glycosides from steviol.

FIG. 3 shows chemical structures for various steviol glycosides.

additional amino acids at the N-terminus, relative to the six sequences identified by pyrosequencing. To re-test UGT91D1 family members for activity, UGT91D1 sequences were re-isolated by PCR amplification of *Stevia* leaf cDNA. The resulting PCR products were cloned into a plasmid vector and enzymatic activity for each product was measured as described above by: GST-tagged expression in *E. coli*, coupled *in vitro* transcription-translation, and/or *in vivo* expression in yeast. Steviol 1,2-glucosylation activity was detected from one clone by all three methods. This clone was designated UGT91D2e. The amino acid sequence of UGT91D2e is set forth in SEQ ID NO:5. In contrast, no 1,2-glucosylation activity was detected from a clone having the same sequence as described by Accession No. AY345980 (Protein Accession number AAR06918), but lacking the 12 amino acids of the amino terminus.

Example 11 – Analysis of UGT91D2e Sequences

Sequence variants of UGT91D2e

As evidenced in FIG. 8, a small number of amino acid modifications exist between the active (91D2e) variants and the closest inactive homologs (91D1). The 91D1 genes cloned by Ma et al., Shi Yan Sheng Wu Xue Bao. 2003 36(2):123-9 (Protein Accession number AAM53963, GI:21435782) and Brandle et al., *supra* (Protein Accession number AAR06918, GI:37993665) did not exhibit the 1,2-glycosylating activity required for RebA biosynthesis. To ascertain which amino acids are required for activity, 21 single site-directed mutants were created such that the amino acid in UGT91D2e (SEQ ID NO:5) was changed to the corresponding amino acid in an inactive homolog. See Table 12. In addition, a site-directed mutation was made such that position 364 (S→P) also was changed. The mutants were made using the QuikChange® II Site-Directed Mutagenesis kit according to manufacturer's protocols (Agilent Technologies, Santa Clara, CA), and the pGEX-4TI vectors were transformed into a XJb Autolysis *E. coli* strain (ZymoResearch, Orange, CA). A mutant was not made to change residue 162 from a glycine to an aspartic acid.

In order to assess the activity of the mutant enzymes, a substrate-feeding experiment was performed *in vitro* using protein produced in *E. coli*. Initially, *E. coli* cells were grown overnight at 30˚C, followed by induction with 3 mM arabinose and

WHAT IS CLAIMED IS

1. A recombinant host cell, comprising a recombinant gene encoding a recombinant 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 and having at least 90% sequence identity to the amino acid sequence set forth in any one of SEQ ID NOs:5, 10, 76, 78, or 150;

wherein the polypeptide is capable of transferring a sugar moiety to the C2' of a glucose in the steviol glycoside;

wherein the steviol glycoside is steviol-13-*O*-glucoside, rubusoside, stevioside or rebaudioside A, and

wherein the recombinant host cell is capable of producing stevioside, rebaudioside E, rebaudioside D, steviol-1,2-bioside, steviol-1,2-xylobioside, steviol-1,2-rhamnobioside, 1,2-stevioxyloside, an isomer thereof, and/or a composition thereof upon transfer of the sugar moiety by the polypeptide.

2. The recombinant host cell of claim 1, wherein 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 the steviol glycoside comprises:

(a) a polypeptide having at least 90% sequence identity to the amino acid sequence set forth in SEQ ID NO:5 and having one or more amino acid substitutions of residues 1-19, 27-38, 44-87, 96-120, 125-141, 159-184, 199-202, 215-380, or 387-473 of SEQ ID NO:5;

(b) a polypeptide having at least 90% sequence identity to the amino acid sequence set forth in SEQ ID NO:5 and having one or more amino acid substitutions of residues 30, 93, 99, 122, 140, 142, 144, 148, 152, 153, 156, 195, 196, 199, 206, 207, 211, 213, 221, 286, 343, 364, 384, 427, and 438 of SEQ ID NO:5;

(c) a polypeptide having at least 90% sequence identity to the amino acid sequence set forth in SEQ ID NO:5 and having an arginine at residue 206, a cysteine at residue 207, and an arginine at residue 343 relative to SEQ ID NO:5; or

(d) a polypeptide having at least 90% sequence identity to the amino acid sequence set forth in SEQ ID NO:5 and having a tyrosine or phenylalanine at residue 30, a proline or glutamine at residue 93, a serine or valine at residue 99, a tyrosine or phenylalanine at residue 122, a histidine or tyrosine at residue 140, a serine or cysteine at residue 142, an alanine or threonine at residue 148, a methionine at residue 152, an alanine at residue 153, an alanine or serine at residue 156, a glycine at residue 162, a leucine or methionine at residue 195, a glutamic acid at residue 196, a lysine or glutamic acid at residue 199, a leucine or methionine at residue 211, a leucine at residue 213, a serine or phenylalanine at residue 221, a valine or isoleucine at residue 253, a valine or alanine at residue 286, an asparagine or lysine at residue 427, or an alanine at residue 438 and an alanine or threonine at residue 462 relative to SEQ ID NO:5.

3. The recombinant host cell of claim 1 or 2, wherein the sugar moiety comprises glucose, rhamnose, and/or xylose.

4. The recombinant host cell of any one of claims 1-3, wherein the recombinant host cell further comprises:

(a) a gene encoding a polypeptide capable of glycosylating steviol or the steviol glycoside at its C-13 hydroxyl group, wherein the polypeptide comprises:

(i) a polypeptide having at least 90% sequence identity to the amino acid sequence set forth in SEQ ID NO:3; or

(II) a polypeptide having one or more amino acid substitutions of residues 9, 10, 13, 15, 21, 27, 60, 65, 71, 87, 91, 220, 243, 270, 289, 298, 334, 336, 350, 368, 389, 394, 397, 418, 420, 440, 441, 444, and 471 of SEQ ID NO:3;

(b) a gene encoding a polypeptide capable of beta 1,3 glycosylation of the C3' of the 13-*O*-glucose of the steviol glycoside, wherein the polypeptide comprises:

(i) a polypeptide having at least 90% sequence identity to the amino acid sequence set forth in SEQ ID NO:7; or

(ii) a polypeptide having one or more amino acid substitutions of residues 29, 74, 87, 91, 116, 123, 125, 126, 130, 145, 192, 193, 194, 196, 198, 199, 200, 203, 204, 205, 206, 207, 208, 266, 273, 274, 284, 285, 291, 330, 331, and 346 of SEQ ID NO:7; and/or

(c) a gene encoding a polypeptide capable of glycosylating steviol or the steviol glycoside at its C-19 carboxyl group; wherein the polypeptide comprises a polypeptide having at least 90% sequence identity to the amino acid sequence set forth in SEQ ID NO:1.

5. The recombinant host cell of any one of claims 1-4, wherein the host further comprises:

(a) at least one genetic modification in exo-1,3-□-glucanase (EXG1) and/or exo-1,3-□-glucanase (EXG2),

wherein the at least one genetic modification reduces EXG1 and/or EXG2 activity in the host; or

(b) at least one genetic modification in squalene synthase (ERG9),

wherein the at least one genetic modification reduces ERG9 activity in the host.

6. The recombinant host cell of any one of claims 1-5, wherein one or more of the genes do not comprise:

(a) a leader sequence;

(b) a signal peptide; and/or

(c) a membrane anchoring sequence.

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

(a) further comprise a nucleotide sequence coding for a fusion tag; or

(b) are expressed as fusion proteins.

8. The recombinant host cell of any one of claims 1-7, wherein the recombinant host cell comprises a plant cell, a fungal cell from *Aspergillus* genus or a yeast 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, an algal cell or a bacterial cell from *Escherichia coli* species or *Bacillus* genus.

9. A method of producing steviol, one or more steviol glycosides or a steviol glycoside composition in a cell culture, comprising growing the recombinant host cell of any one of claims 1-8 under conditions in which one or more of the genes are expressed;

wherein steviol, the steviol glycoside, or the steviol glycoside composition is produced by the recombinant host cell;

wherein the one or more steviol glycosides is, or the steviol glycoside composition comprises, stevioside, rebaudioside E, rebaudioside D, steviol-1,2-bioside, steviol-1,2-xylobioside, steviol-1,2-rhamnobioside, 1,2-stevioxyloside, and/or an isomer thereof.

10. The method of claim 9, wherein growing can include inducing expression of one or more of the genes.

11. The method of claim 10, wherein the steviol glycoside is rebaudioside D, wherein rebaudioside D is produced upon transfer of a glucose moiety to rebaudioside A.

12. The method of claim 10, further comprising isolating the Rebaudioside D, alone or together with at least one other steviol glycoside from the cell culture;

wherein the isolating step comprises separating a liquid phase of the cell culture from a solid phase of the cell culture to obtain a supernatant comprising Rebaudioside D, alone or together with at least one other steviol glycoside, and:

(a) contacting the supernatant with one or more adsorbent resins in order to obtain at least a portion of Rebaudioside D, alone or together with at least one other steviol glycoside; or

(b) contacting the supernatant with one or more ion exchange or reversed-phase chromatography columns in order to obtain at least a portion of Rebaudioside D, alone or together with at least one other steviol glycoside; or

(c) crystallizing or extracting Rebaudioside D, alone or together with at least one other steviol glycoside;

thereby isolating Rebaudioside D, alone or together with at least one other steviol glycoside.

13. The method of any one of claims 9-11, further comprising recovering Rebaudioside D, alone or together with at least one other steviol glycoside or the steviol glycoside composition from the cell culture;

wherein the recovered steviol glycoside composition is enriched for Rebaudioside D relative to a steviol glycoside composition from a Stevia plant and has a reduced level of Stevia plant-derived components relative to a steviol glycoside composition obtained from a plant-derived Stevia extract.

14. The method of claims 9-13, wherein the cell culture comprises:

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

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

15. An *in vitro* method for producing one or more steviol glycosides or a steviol glycoside composition, comprising adding 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 and having at least 90% sequence identity to the amino acid sequence set forth in any one of SEQ ID NOs:5, 10, 76, 78, or 150;

(b) a polypeptide capable of glycosylating steviol or a steviol glycoside at its C-13 hydroxyl group,

wherein the polypeptide comprises:

(i) a polypeptide having at least 90% sequence identity to the amino acid sequence set forth in SEQ ID NO:3; or

(II) a polypeptide having one or more amino acid substitutions of residues 9, 10, 13, 15, 21, 27, 60, 65, 71, 87, 91, 220, 243, 270, 289, 298, 334, 336, 350, 368, 389, 394, 397, 418, 420, 440, 441, 444, and 471 of SEQ ID NO:3;

(c) a polypeptide capable of beta 1,3 glycosylation of the C3' of the 13-*O*-glucose of a steviol glycoside,

wherein the polypeptide comprises:

(i) a polypeptide having at least 90% sequence identity to the amino acid sequence set forth in SEQ ID NO:7; or

(ii) a polypeptide having one or more amino acid substitutions of residues 29, 74, 87, 91, 116, 123, 125, 126, 130, 145, 192, 193, 194, 196, 198, 199, 200, 203, 204, 205, 206, 207, 208, 266, 273, 274, 284, 285, 291, 330, 331, and 346 of SEQ ID NO:7; and/or

(d) a polypeptide capable of glycosylating steviol or a steviol glycoside at its C-19 carboxyl group,

wherein the polypeptide comprises a polypeptide having at least 90% sequence identity to the amino acid sequence set forth in SEQ ID NO:1;

and steviol, steviol-13-*O*-glucoside, rubusoside, stevioside or rebaudioside A to a reaction mixture; and

producing the one or more steviol glycosides or the steviol glycoside composition;

wherein the steviol glycoside comprises steviol-13-*O*-glucoside, rubusoside, stevioside or rebaudioside A; and

wherein the steviol glycoside composition comprises stevioside, rebaudioside E, rebaudioside D, steviol-1,2-bioside, steviol-1,2-xylobioside, steviol-1,2-rhamnobioside, and/or 1,2-stevioxyloside.

16. The method of claim 23, wherein 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:

(a) at least 90% sequence identity to the amino acid sequence set forth in SEQ ID NO:5 and having at least one amino acid substitution at residues 1-19, 27-38, 44-87, 96-120, 125-141, 159-184, 199-202, 215-380, or 387-473 of SEQ ID NO:5;

(b) at least 90% sequence identity to the amino acid sequence set forth in SEQ ID NO:5 and having an amino acid substitution at one or more residues selected from the group consisting of residues 30, 93, 99, 122, 140, 142, 144, 148, 152, 153, 156, 195, 196, 199, 206, 207, 211, 213, 221, 286, 343, 364, 384, 427, and 438 of SEQ ID NO:5;

(c) at least 90% sequence identity to the amino acid sequence set forth in SEQ ID NO:5 and having an arginine at residue 206, a cysteine at residue 207, and an arginine at residue 343 relative to SEQ ID NO:5; or

(d) at least 90% sequence identity to the amino acid sequence set forth in SEQ ID NO:5 and having a tyrosine or phenylalanine at residue 30, a proline or glutamine at residue 93, a serine or valine at residue 99, a tyrosine or phenylalanine at residue 122, a histidine or tyrosine at residue 140, a serine or cysteine at residue 142, an alanine or threonine at residue 148, a methionine at residue 152, an alanine at residue 153, an alanine or serine at residue 156, a

glycine at residue 162, a leucine or methionine at residue 195, a glutamic acid at residue 196, a lysine or glutamic acid at residue 199, a leucine or methionine at residue 211, a leucine at residue 213, a serine or phenylalanine at residue 221, a valine or isoleucine at residue 253, a valine or alanine at residue 286, an asparagine or lysine at residue 427, or an alanine at residue 438 and an alanine or threonine at residue 462 relative to SEQ ID NO:5.

17. One or more steviol glycosides or a steviol glycoside composition produced by the method of any one of claims 9-16.

18. A sweetener composition, comprising the one or more steviol glycosides or the steviol glycoside composition of claim 17.

19. A food product, comprising the sweetener composition of claim 18.

20. A beverage or a beverage concentrate, comprising the sweetener composition of claim 18.

21. A cell culture, comprising the recombinant host cell of any one of claims 1-8, the cell culture further comprising:

(a) stevioside, rebaudioside E, rebaudioside D, steviol-1,2-bioside, steviol-1,2-xylobioside, steviol-1,2-rhamnobioside, 1,2-stevioxyloside, an isomer thereof, and/or the composition thereof produced by the recombinant host cell;

(b) glucose, fructose, sucrose, xylose, rhamnose, uridine diphosphate (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 one or more of stevioside, rebaudioside E, rebaudioside D, steviol-1,2-bioside, steviol-1,2-xylobioside, steviol-1,2-rhamnobioside, 1,2-stevioxyloside, an isomer thereof alone, and/or in the composition thereof is present at a concentration of at least 1 mg/liter of the cell culture;

wherein the cell culture is enriched for stevioside, rebaudioside E, rebaudioside D, steviol-1,2-bioside, steviol-1,2-xylobioside, steviol-1,2-rhamnobioside, 1,2-stevioxyloside, an isomer thereof, and/or the composition thereof relative to a steviol glycoside composition from a Stevia plant and has a reduced level of Stevia plant-derived components relative to a steviol glycoside composition obtained from a plant-derived Stevia extract.

22. A cell lysate from the recombinant host of any one of claims 1-8 grown in the cell culture, comprising:

(a) stevioside, rebaudioside E, rebaudioside D, steviol-1,2-bioside, steviol-1,2-xylobioside, steviol-1,2-rhamnobioside, 1,2-stevioxyloside, an isomer thereof, and/or the composition thereof produced by the recombinant host cell;

(b) glucose, fructose, sucrose, xylose, rhamnose, uridine diphosphate (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 one or more of stevioside, rebaudioside E, rebaudioside D, steviol-1,2-bioside, steviol-1,2-xylobioside, steviol-1,2-rhamnobioside, 1,2-stevioxyloside, an isomer thereof alone, and/or in the composition thereof produced by the recombinant host cell is present at a concentration of at least 1 mg/liter of the cell culture.

23. A reaction mixture, comprising:

(a) one or more of stevioside, rebaudioside E, rebaudioside D, steviol-1,2-bioside, steviol-1,2-xylobioside, steviol-1,2-rhamnobioside, 1,2-stevioxyloside, an isomer thereof, and/or the steviol glycoside composition thereof produced in the reaction mixture,

(b) 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 and having at least 90% sequence identity to the amino acid sequence set forth in any one of SEQ ID NOs:5, 10, 76, 78, or 150;

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

(d) reaction buffer and/or salts.