respectively. All three of these enzymes, UGT91D_like3, Os_UGT_91C1, and SL_UGT_101249881, are uridine diphosphate-dependent glycosyltransferases (UGT) which transfer a glucose moiety to the C-2′ position of the 19-O-glucose residue via formation of a beta(1->2) linkage (**FIG. 1A**).

[0005] To produce RebM efficiently and at high purity, improved enzymes capable of converting RebA to RebD at high efficiency are needed. The compositions and methods provided herein address this need and provide related advantages as well.

4. SUMMARY

[0006] Provided herein are compositions and methods for the improved conversion of RebA to RebD and for the improved production of RebD and/or RebM. These compositions and methods are based in part on the surprising discovery of certain uridine diphosphate-dependent glycosyltransferases (UGTs) are capable of converting RebA to RebD with remarkably high efficiency. Even a modest improvement in strain performance (*e.g.*, ten percent) with new UGTs can potentially save over ten million dollars in production cost in the future, assuming that the market demand for RebM is 5000 million tons per year.

[0007] Certain UGTs described herein are also capable of producing RebM with little or no non-natural glycosides such as RebM2 side product (*i.e.*, an isomer of RebM). *See* Ceunen S. *et al.* Steviol Glycosides: Chemical Diversity, Metabolism, and Function. *J. Nat. Prod.*, 76, 1201-1228 (2013) for a list of currently-known Stevia glycosides. As such, in certain examples, the compositions and methods described herein can reduce the costs of downstream processing to obtain a composition with high purity RebM.

[0008] Also provided herein are compositions and methods for alternative enzymes which are capable of producing steviol glycosides with different substrate specificities than previously known UGTs. The new alternative enzymes can potentially produce different mixtures or proportions of steviol glycosides compared to those produced by other known enzymes. Compositions with different mixtures or proportions of steviol glycosides can potentially impart alternative, sweet taste profiles, which may be useful in formulating various consumable or food products.

[0009] Thus, provided herein are genetically modified host cells and methods of their use for the production of industrially useful compounds. In one aspect, provided herein is a genetically modified host cell comprising: a heterologous nucleic acid encoding a UDP-glycosyltransferase (UGT40087, also refer to as Si_UGT_40087). In some examples, the

genetically modified host cell further comprises one or more enzymatic pathways capable of producing steviol and/or steviol glycosides.

[0010] In certain examples, provided herein are genetically modified host cells comprising a heterologous nucleic acid encoding a UDP-glycosyltransferase comprising an amino acid sequence having at least 80%, 85%, 90%, or 95% sequence identity to the sequence of UGT40087 (*e.g.*, SEQ ID NO:1 or SEQ ID NO:11). In certain examples, the genetically modified host cell is capable of converting RebA to RebD at an efficiency greater than 90%, 95%, 96%, or 97%. In certain examples, the genetically modified host cell comprises a UDP-glycosyltransferase comprising a sugar acceptor domain, wherein the amino acid sequence of the sugar acceptor domain has at least 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of the sugar acceptor domain of SEQ ID NO:1 or SEQ ID NO:11. In certain examples, the genetically modified host cell comprises a UDP-glycosyltransferase which comprises a loop1 amino acid sequence, a variant loop1 amino acid sequence, a loop2 amino acid sequence, a variant loop2 amino acid sequence, a loop3_1 amino acid sequence, a variant loop3_1 amino acid sequence, a loop3_2 amino acid sequence, a variant loop3_2 amino acid sequence, a loop4_1 amino acid sequence, a variant loop4_1 amino acid sequence, a loop4_2 amino acid sequence, or any combination thereof. In certain examples, the genetically modified host cell comprises a UDP-glycosyltransferase comprising an amino acid sequence having at least 61%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% sequence identity to the sugar acceptor domain of SEQ ID NO:1 or SEQ ID NO:11, and further comprises the loop4_1 amino acid sequence of SEQ ID NO:1 or SEQ ID NO:11.

[0011] In certain examples, provided herein are genetically modified host cells comprising a heterologous nucleic acid encoding a UDP-glycosyltransferase comprising an amino acid sequence having at least 80%, 85%, 90%, or 95% sequence identity to the sequence of SEQ ID NO:2, 5, or 6. In certain examples, the genetically modified host cell is capable of converting RebA to RebD at an efficiency greater than 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, or 97%. In certain examples, the genetically modified host cell comprises a UDP-glycosyltransferase comprising a sugar acceptor domain, wherein the amino acid sequence of the sugar acceptor domain has at least 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence of the sugar acceptor domain of SEQ ID NO:2, 5, or 6. In certain examples, the genetically modified host cell comprises a UDP-glycosyltransferase which comprises a loop1 amino acid sequence, a variant loop1 amino acid sequence, a loop2 amino

acid sequence, a variant loop2 amino acid sequence, a loop3_1 amino acid sequence, a variant loop3_1 amino acid sequence, a loop3_2 amino acid sequence, a variant loop3_2 amino acid sequence, a loop4_1 amino acid sequence, a variant loop4_1 amino acid sequence, a loop4_2 amino acid sequence, or any combination thereof. In certain examples, the genetically modified host cell comprises a UDP-glycosyltransferase comprising an amino acid sequence having at least 61%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% sequence identity to the sugar acceptor domain of SEQ ID NO:2, 5, or 6, and further comprises the loop4_1 amino acid sequence of SEQ ID NO:2, 5, or 6.

[0012] In another example, provided herein are methods for producing a heterologous steviol glycoside, the method comprising: culturing a population of genetically modified host cells provided herein, capable of producing the steviol glycoside as described herein, in a medium with a carbon source under conditions suitable for making said steviol glycoside compound; and recovering said steviol glycoside from the medium. In some examples, the heterologous steviol glycoside is selected from the group consisting of RebD and RebM.

[0013] In another example, provided herein is a method for increasing the production of a steviol glycoside compound in a host cell, the method comprising: expressing in the host cell a heterologous nucleic acid encoding a UGT40087; and culturing the host cell under conditions suitable for producing the steviol glycoside. In some examples, the host cell does not comprise a UGT91D_like3 enzyme, a Os_UGT_91C1 enzyme, or a Sl_UGT_101249881 enzyme.

[0014] In another example, provided herein are methods for producing RebD, the method comprising: culturing a population of genetically modified host cells provided herein, capable of producing RebD as described herein, in a medium with a carbon source under conditions suitable for making said RebD; and recovering said RebD from the medium.

[0015] In another aspect, provided herein are methods for producing RebM, the method comprising: culturing a population of genetically modified host cells provided herein, capable of producing RebM as described herein, in a medium with a carbon source under conditions suitable for making said RebM; and recovering said RebM from the medium.

[0016] In another example, provided herein are methods for producing RebD, the method comprising: contacting RebA with glucose and a UDP-glycosyltransferase described herein, capable of converting RebA to RebD, under conditions suitable for forming RebD.

[0017] In another example, provided herein are methods for producing RebM, the method comprising: contacting RebA with glucose and a UDP-glycosyltransferase described

herein, capable of converting RebA to RebD, under conditions suitable for forming RebD, and with a UDP-glycosyltransferase described herein, capable of converting RebD to RebM.

[0018] In some examples, the host cell is a yeast cell. In some examples, the yeast is *Saccharomyces cerevisiae*. In some examples, the host cell produces RebD or RebM at high efficiency. In some examples, the host cell produces an increased amount of RebD or RebM compared to a yeast cell not comprising the UGT40087 enzyme. In some examples, the host cell produces an increased amount of RebM relative to RebM2 compared to a yeast cell not comprising the UGT40087 enzyme.

[0019] In another example, other UDG-glycosyltransferases provided herein can be used in addition or in alternative to UGT40087. These include, for example, sr.UGT_9252778, Bd_UGT10840, Hv_UGT_V1, Bd_UGT10850, and Ob_UGT91B1_like.

[0019a] In another aspect, the present invention refers to a fermentation composition produced by the genetically modified host cells as described herein comprising steviol glycosides, wherein the steviol glycosides comprise RebA, RebD and RebM at a RebA:RebD:RebM ratio of at least 1:7:50.

5. BRIEF DESCRIPTION OF THE FIGURES

[0020] **FIG. 1A** provides a schematic representation of the conversion of RebA to RebD to RebM.

[0021] **FIG. 1B** provides the structure of RebM2.

[0022] **FIG. 1C** provides a schematic diagram of the mevalonate pathway.

[0023] **FIG. 2A** provides an exemplary pathway of farnesyl pyrophosphate (FPP) to steviol.

[0024] **FIG. 2B** provides an exemplary pathway of steviol to RebM.

[0025] **FIG. 2C** provides an exemplary pathway for the enzymatic production of RebM.

[0026] **FIG. 3** provides the ratio of RebA to RebD conversion *in vivo*, as measured by micromoles of Reb (D+M) / micromoles of Reb (A+D+M). The parent control strain is labelled 91D_like3 (from *Stevia rebaudiana*; this strain contains UGT's: 85C2, 74G1, 91D_like3, and 76G1 only, in addition to an empty landing pad. It is noted that 91D_like3 has very low RebA to RebD conversion (~3% *see* Table 5). A single copy of each UGT enzyme was inserted into the parent control strain and screened for improved RebA to RebD conversion. Six UGT enzymes are shown to have RebA to RebD conversion that are at least equivalent to, or better than, the previous known enzymes Os_UGT_91C1 and Sl_UGT_101249881. Three UGT enzymes (Si_UGT_40087, Ob_UGT91B1_like, and Hv_UGT_V1 are better than both previously identified UGT enzymes at converting RebA to RebD. Error bars are standard error.

WHAT IS CLAIMED:

1. A genetically modified host cell comprising a heterologous nucleic acid encoding a UDP-glycosyltransferase comprising an amino acid sequence having at least 80%, 85%, 90%, or 95% sequence identity to SEQ ID NO: 1 or SEQ ID NO: 11.

2. The genetically modified host cell of claim 1 that is capable of converting RebA to RebD at an efficiency of greater than 90%, 95%, 96%, or 97%.

3. The genetically modified host cell of any of the preceding claims, wherein the UDP-glycosyltransferase comprises an amino acid sequence having the sequence of SEQ ID NO:1 or SEQ ID NO:11.

4. The genetically modified host cell of any of the preceding claims, wherein the UDP-glycosyltransferase is capable of beta 1,2 glycosylation of the C2' position of the 19-O glucose of a steviol glycoside.

5. The genetically modified host cell of any of the preceding claims, wherein the UDP-glycosyltransferase is encoded by a heterologous nucleic acid, wherein the heterologous nucleic acid comprises a nucleotide sequence having at least 80%, 85%, 90%, or 95% sequence identity to SEQ ID NO:13.

6. The genetically modified host cell of any of the preceding claims, wherein the UDP-glycosyltransferase is encoded by a heterologous nucleic acid having the sequence of SEQ ID NO:13.

7. The genetically modified host cell of any of the preceding claims, that is capable of converting RebA to RebD at an efficiency of greater than 90%, 95%, 96%, or 97%, and wherein the UDP-glycosyltransferase comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO:1 or SEQ ID NO:11.

8. The genetically modified host cell of any of the preceding claims, wherein the UDP-glycosyltransferase comprises one or more of the following amino acid residues:

(a) valine at an amino acid position of the UDP-glycosyltransferase that corresponds to amino acid position 11 of SEQ ID NO:11;

(b) isoleucine at an amino acid position of the UDP-glycosyltransferase that corresponds to amino acid position 12 of SEQ ID NO:11;

(c) proline at an amino acid position of the UDP-glycosyltransferase that corresponds to amino acid position 55 of SEQ ID NO:11;

(d) glutamic acid at an amino acid position of the UDP-glycosyltransferase that corresponds to amino acid position 90 of SEQ ID NO:11;

(e) serine at an amino acid position of the UDP-glycosyltransferase that corresponds to amino acid position 203 of SEQ ID NO:11;

(f) glutamic acid at an amino acid position of the UDP-glycosyltransferase that corresponds to amino acid position 223 of SEQ ID NO:11; or

(g) valine at an amino acid position of the UDP-glycosyltransferase that corresponds to amino acid position 413 of SEQ ID NO:11,

wherein the amino acid positions of the UDP-glycosyltransferase that correspond to the amino acid positions of SEQ ID NO:11 are determined by sequence alignment.

9. The genetically modified host cell of any of the preceding claims, wherein the UDP-glycosyltransferase comprises an amino acid sequence having SEQ ID NO:24.

10. The genetically modified host cell of any of the preceding claims, wherein the UDP-glycosyltransferase comprises an amino acid sequence having at least 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to SEQ ID NO:1 or SEQ ID NO:11.

11. The genetically modified host cell of any of the preceding claims that is capable of producing RebD.

12. The genetically modified host cell of any of the preceding claims that is capable of producing RebM.

13. The genetically modified host cell of any of the preceding claims, wherein the genetically modified host cell further comprises one or more heterologous nucleic acids encoding one or more enzymes of a pathway for making steviol.

14. The genetically modified host cell of any of the preceding claims, wherein the genetically modified host cell further comprises one or more heterologous nucleic acids encoding one or more enzymes of a pathway for making a steviol glycoside.

15. The genetically modified host cell of any of the preceding claims, wherein the one or more enzymes of the pathway comprise a geranylgeranyl diphosphate synthase, a copalyl diphosphate synthase, a ent-kaurene synthase, a kaurene oxidase, a kaurenoic acid hydroxylase, a cytochrome P450 reductase, UGT74G1, UGT76G1, UGT85C2, and UGT91D.

16. The genetically modified host cell of any one of any of the preceding claims, wherein the cell is selected from the group consisting of a bacterial cell, a fungal cell, an algal cell, an insect cell, and a plant cell.

17. The genetically modified host cell of any of the preceding claims, wherein the cell is a yeast cell.

18. The genetically modified host cell of any of the preceding claims, wherein the yeast is *Saccharomyces cerevisiae*.

19. A method for producing RebM:

(a) culturing a population of the genetically modified host cells of any of the preceding claims in a medium with a carbon source under conditions suitable for making RebM; and

(b) recovering said RebM compound from the medium.

20. A fermentation composition produced by the genetically modified host cells of any of the preceding claims comprising steviol glycosides, wherein the steviol glycosides comprise RebA, RebD and RebM at a RebA:RebD:RebM ratio of at least 1:7:50.