

Recombinant Production of Steviol Glycosides

SEQUENCE LISTING

5 The instant application contains a Sequence Listing which has been submitted
in ASCII format via EFS-Web and is hereby incorporated by reference in its entirety.
Said ASCII copy, created on June 2, 2011, is named 25933WO1.txt and is 483,406
bytes in size.

TECHNICAL FIELD

10 This disclosure relates to the recombinant production of steviol and steviol
glycosides. In particular, this disclosure relates to the production of steviol and
steviol glycosides such as rubusoside and/or rebaudioside A by recombinant hosts
such as recombinant microorganisms, plants, or plant cells. This disclosure also
provides compositions containing steviol glycosides.

BACKGROUND

15 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,
20 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
25 sweetener in foods and in blends or alone as a tabletop sweetener.

Extracts of the Stevia plant contain rebaudiosides and other steviol glycosides
that contribute to the sweet flavor, although the amount of each glycoside often varies
among different production batches. Existing commercial products are predominantly
rebaudioside A with lesser amounts of other glycosides such as rebaudioside C, D,
30 and F. Stevia extracts may also contain contaminants such as plant-derived

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-
5 eicosatrienoic acid, 2-methyloctadecane, pentacosane, octacosane, tetracosane, octadecanol, stigmasterol, β -sitosterol, α - and β -amyrin, lupeol, β -amryin acetate, pentacyclic triterpene, centauredin, quercetin, cpi-alpha-cadinol, carophyllenes and derivatives, beta-pinene, beta-sitosterol, and gibberellin.

SUMMARY

10 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
15 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 embodiments 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
20 encoding a CPR. The expression of one or more of the genes can be inducible.

In one aspect, 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
25 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,
30 427, and 438 of SEQ ID NO:5. In one embodiment, 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 embodiment, 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
5 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.

10 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
15 471 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,
20 193, 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
25 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
30 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 aspect, 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 embodiments, 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
5 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 embodiments, the polypeptide includes a phenylalanine at residue 30, a glutamine at residue 93, a valine at residue 99, a phenylalanine at residue
10 122, 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.

15 In another aspect, 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
20 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
25 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
30 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

sequences set forth in SEQ ID NOs:142-146. The host further can include a gene encoding a truncated HMG-CoA and/or a gene encoding a CPR.

Any of the recombinant hosts further can include one or more of a gene encoding a UGT74G1 polypeptide, a UGT85C2 polypeptide, a UGT76G1
5 polypeptide, or a UGT91D2 polypeptide.

Any of the recombinant hosts can produce at least one steviol glycoside when cultured under conditions in which each of the genes is expressed. The steviol glycoside can be selected from the group consisting of steviol-13-*O*-glucoside, steviol-19-*O*-glucoside, rubusoside, rebaudioside A, rebaudioside B, rebaudioside C,
10 rebaudioside D, rebaudioside E, rebaudioside F, and dulcoside A. The steviol glycoside can accumulate to at least 1 mg/liter (e.g., at least 10 mg/liter or 20 mg/liter) of culture medium when cultured under said conditions.

Any of the recombinant hosts further can include one or more of i) a gene encoding a deoxyxylulose 5-phosphate synthase (DXS); ii) a gene encoding a D-1-
15 deoxyxylulose 5-phosphate reductoisomerase (DXR); iii) a gene encoding a 4-diphosphocytidyl-2-C-methyl-D-erythritol synthase (CMS); iv) a gene encoding a 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase (CMK); v) a gene encoding a 4-diphosphocytidyl-2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase (MCS); vi) a gene encoding a 1-hydroxy-2-methyl-2(E)-butenyl 4-diphosphate synthase (HDS);
20 or vii) a gene encoding a 1-hydroxy-2-methyl-2(E)-butenyl 4-diphosphate reductase (HDR).

Any of the recombinant hosts further can include one or more of ix) a gene encoding an acetoacetyl-CoA thiolase; x) a gene encoding a truncated HMG-CoA reductase; xi) a gene encoding a mevalonate kinase; xii) a gene encoding a
25 phosphomevalonate kinase; or xiii) a gene encoding a mevalonate pyrophosphate decarboxylase.

In any of the hosts described herein, expression of one or more of the genes can be inducible.

Any of the hosts described herein can be a microorganism (e.g., a
30 Saccharomycete such as *Saccharomyces cerevisiae*, or *Escherichia coli*), or a plant or plant cell (e.g., a *Stevia* such as a *Stevia rebaudiana*, *Physcomitrella*, or tobacco plant or plant cell).

In another aspect, 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
5 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.
10 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
15 be a *Saccharomyces* spp. In some embodiments, 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 embodiments, the recovering step comprises purifying the steviol or steviol glycoside from the culture medium by
20 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
25 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 aspect, this document features an isolated nucleic acid having greater than 90 % sequence identity (e.g., greater than 95% or 99% sequence identity)
30 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 embodiments, 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 UGT91D2c 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 UGT91D2c 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 aspect, 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.

5 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
10 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
15 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
20 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.
25 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 aspect, this document features a method of producing a steviol glycoside composition. The method includes growing a host described herein in a
30 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.

5 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 aspect, this document features a method of identifying whether a
10 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
15 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 aspect, this document features a method 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
20 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
25 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
30 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 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 UGT91D2 polypeptide can include an amino acid substitution at one or more
5 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
10 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
15 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
20 Rebaudioside A, and Rebaudioside D is produced upon transfer of a second glucose moiety.

In another aspect, 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
25 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.
30 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.

Also provided herein is a recombinant microorganism, comprising one or more biosynthesis genes whose expression results in production of one or more steviol glycosides. The biosynthesis genes include a gene encoding a geranylgeranyl diphosphate synthase, a gene encoding a copalyl diphosphate synthase and a gene encoding a kaurene synthase, a gene encoding a kaurene oxidase, a gene encoding a steviol synthetase, and a gene encoding a UGT74G1 and/or a UGT85C2. At least one of the genes is a recombinant gene. The microorganism can comprise a gene encoding a bifunctional copalyl diphosphate synthase and kaurene synthase in place of the genes encoding copalyl diphosphate synthase and kaurene synthase.

The recombinant microorganism produces at least one steviol glycoside when cultured under conditions in which each of the genes is expressed. The steviol glycoside can be rubusoside, rebaudioside C, rebaudioside F, dulcoside B, or dulcoside A.

The recombinant microorganism can be a *Saccharomycete*, e.g., *Saccharomyces cerevisiae*, and can have one or more genetic modifications that reduce *EXG1* and *EXG2* glycoside hydrolase activity relative to a control microorganism that lacks such genetic modifications, and can have one or more genetic modifications that reduce ergosterol biosynthesis relative to a control microorganism that lacks such genetic modifications. The *Saccharomycete* produces rubusoside when cultured under conditions in which each of the genes is expressed. The rubusoside can accumulate to at least 10 mg/liter of culture medium. The *Saccharomycete* can be a *Saccharomyces cerevisiae* strain designated CEY171, CEY191, or CEY213.

The recombinant microorganism can further comprise a gene encoding an SM12UGT and a gene encoding a UGT76G1, and produce a steviol glycoside when cultured under conditions in which each of the genes is expressed. The steviol glycoside can be rebaudioside A.

Also provided herein is a recombinant microorganism, comprising one or more biosynthesis genes whose expression results in production of at least one steviol glycoside. The biosynthesis genes include a gene encoding an SM12UGT, a gene encoding a UGT74G1, a gene encoding a UGT76G1 and a gene encoding a UGT85C2. The recombinant microorganism produces rebaudioside A or rebaudioside

B when cultured under conditions in which each of the genes is expressed. The rebaudioside A or rebaudioside B can accumulate to at least 1 mg/L in the culture medium.

Also featured herein is a recombinant microorganism, comprising a gene
5 encoding a UGT91D2 polypeptide, e.g., a recombinant UGT91D2 gene.

Also featured herein is a recombinant microorganism, comprising a gene
encoding a geranylgeranyl diphosphate synthase, 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), a gene encoding a
10 kaurene oxidase, a gene encoding a steviol synthetase, a gene encoding a UGT74G1,
a gene encoding a UGT85C2, a gene encoding a UGT76G1, and a gene encoding a
UGT91D2. At least one of the genes is a recombinant gene. The recombinant
microorganism can produce at least one steviol glycoside, e.g., rebaudioside A,
rebaudioside B, and/or rebaudioside F, when cultured under conditions in which each
15 of the genes is expressed. The recombinant microorganism can accumulate at least 20
mg of steviol glycoside per liter of culture medium when cultured under such
conditions. The recombinant microorganism can be a *Saccharomycete*, e.g.,
Saccharomyces cerevisiae, and can have one or more genetic modifications that
reduce *EXG1* and *EXG2* glycoside hydrolase activity relative to a control
20 microorganism that lacks such genetic modifications, and can have one or more
genetic modifications that reduce ergosterol biosynthesis relative to a control
microorganism that lacks such genetic modifications.

Also featured herein is a recombinant microorganism, comprising a gene
encoding a UGT74G1, a gene encoding a UGT85C2, a gene encoding a UGT76G1,
25 and a gene encoding a UGT91D2. At least one of the genes is a recombinant gene.
The recombinant microorganism can produce a steviol glycoside, e.g., rebaudioside A
or rebaudioside B, when cultured under conditions in which each of the genes is
expressed. The rebaudioside A or rebaudioside B can accumulate to at least 15 mg/L
in the culture medium.

30 The recombinant microorganisms described above can further comprise a gene
encoding a deoxyxylulose 5-phosphate synthase (DXS), and/or a gene encoding a D-
1-deoxyxylulose 5-phosphate reductoisomerase (DXR), and/or a gene encoding a 4-

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 invention belongs. Although methods and materials similar or equivalent to those described herein can be used to practice the invention, 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 invention will be apparent from the following detailed description. Applicants reserve the right to alternatively claim any disclosed invention 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.

FIG. 4 is a schematic representation of rebA production in *Saccharomyces cerevisiae*.

FIG. 5 is a schematic representation of the concatenation of genes to form eYACs.

5 FIG. 6 shows rubusoside production by yeast strain CEY13 under various culture conditions.

FIG. 7 shows data obtained from ^1H and ^{13}C NMR analysis of the compound produced by yeast strain CEY213, compared to literature values for rubusoside.

FIG. 8 is an alignment of UGT91D1 and UGT91D2 amino acid sequences
10 (SEQ ID NOs:14, 16, 12, 5, and 10, respectively).

FIG. 9 shows Rebaudioside A, stevioside, and rubusoside production by yeast CEY213 containing plasmid pMUS47 after 24 and 99 hours of culture.

FIG. 10A is a graph illustrating the concentrations of RebA, rubusoside and 19-SMG in supernatants. FIG. 10B is a graph of the concentrations of RebA,
15 rubusoside and 19-SMG measured in cell pellets, for experiments where yeast cells were fed with 100 μM steviol. In both graphs, the first set of bars represents the untagged control strains; the second set of bars represents the strain containing the UGT74G1, UGT76G1, and UGT91D2c fusion proteins in which the N-terminal 158 amino acids of the MDM2 protein are fused to each UGT, and a UGT85C2 fusion
20 protein in which four repeats of the synthetic PMI peptide is fused in-frame to the N-terminus of 85C2. The y-axis is concentration in micromolar units.

Like reference symbols in the various drawings indicate like elements.

DETAILED DESCRIPTION

Two glycosides, stevioside and rebaudioside A, are the primary compounds in
25 commercially-produced stevia extracts. Stevioside is reported to have a more bitter and less sweet taste than rebaudioside A and, therefore, a higher proportion of rebaudioside A in an extract preparation is preferred. However, the composition of stevia extract can vary from lot to lot depending on the soil and climate in which the plants are grown. Depending upon the sourced plant, the climate conditions, and the
30 extraction process, the amount of rebaudioside A in commercial preparations is reported to vary from 20 to 97% of the total steviol glycoside content, typically >50-

80% and sometimes as high as >95-97% of the total steviol glycosides. Moreover, other steviol glycosides are present in varying amounts in stevia extracts, which further complicates the ability to produce a sweetener with a consistent taste profile by extraction and purification from Stevia plants. For example, Rebaudioside B is typically present at less than 1-2%, whereas Rebaudioside C can be present at levels as high as 7-15%. Rebaudioside D is typically present in levels of 2% or less, and Rebaudioside F is typically present in compositions at 3.5% or less of the total steviol glycosides. Even trace amounts of the minor steviol glycosides are reported to affect the flavor profile of a Stevia extract. Additionally, it is thought that some of the contaminants from the Stevia plant, even at very low concentrations, may also provide off-flavors to some of the commercially available plant extracts.

This document is based on the discovery that recombinant hosts such as plant cells, plants, or microorganisms can be developed that express polypeptides useful for the biosynthesis of steviol. Further, such hosts can express Uridine 5'-diphospho (UDP) glycosyl transferases suitable for producing steviol glycosides such as rubusoside and rebaudioside A. Recombinant microorganisms are particularly useful hosts. Expression of these biosynthetic polypeptides in various microbial chassis allows steviol and its glycosides to be produced in a consistent, reproducible manner from energy and carbon sources such as sugars, glycerol, CO₂, H₂, and sunlight. The proportion of each steviol glycoside produced by a recombinant host can be tailored by incorporating preselected biosynthetic enzymes into the hosts and expressing them at appropriate levels, to produce a sweetener composition with a consistent taste profile. Furthermore, the concentrations of steviol glycosides produced by recombinant hosts are expected to be higher than the levels of steviol glycosides produced in the Stevia plant, which improves the efficiency of the downstream purification. Such sweetener compositions contain little or no plant based contaminants, relative to the amount of contaminants present in Stevia extracts.

At least one of the genes is a recombinant gene, the particular recombinant gene(s) depending on the species or strain selected for use. Additional genes or biosynthetic modules can be included in order to increase steviol and glycoside yield, improve efficiency with which energy and carbon sources are converted to steviol and its glycosides, and/or to enhance productivity from the cell culture or plant. Such

additional biosynthetic modules include genes involved in the synthesis of the terpenoid precursors, isopentenyl diphosphate and dimethylallyl diphosphate. Additional biosynthetic modules include terpene synthase and terpene cyclase genes, such as genes encoding geranylgeranyl diphosphate synthase and copalyl diphosphate synthase; these genes may be endogenous genes or recombinant genes.

I. Steviol and Steviol Glycoside Biosynthesis Polypeptides

A. Steviol Biosynthesis Polypeptides

Chemical structures for several of the compounds found in Stevia extracts are shown in Fig. 3, including the diterpene steviol and various steviol glycosides. CAS numbers are shown in Table A below. See also, Steviol Glycosides Chemical and Technical Assessment 69th JECFA, prepared by Harriet Wallin, Food Agric. Org. (2007).

Table A.

COMPOUND	CAS #
Steviol	471-80-7
Rebaudioside A	58543-16-1
Steviolbioside	41093-60-1
Stevioside	57817-89-7
Rebaudioside B	58543-17-2
Rebaudioside C	63550-99-2
Rebaudioside D	63279-13-0
Rebaudioside E	63279-14-1
Rebaudioside F	438045-89-7
Rubusoside	63849-39-4
Dulcoside A	64432-06-0

It has been discovered that expression of certain genes in a host such as a microorganism confers the ability to synthesize steviol upon that host. As discussed in more detail below, one or more of such genes may be present naturally in a host. Typically, however, one or more of such genes are recombinant genes that have been transformed into a host that does not naturally possess them.

The biochemical pathway to produce steviol involves formation of geranylgeranyl diphosphate, cyclization to (-) copalyl diphosphate, followed by oxidation and hydroxylation to form steviol. See FIG. 1. Thus, conversion of geranylgeranyl diphosphate to steviol in a recombinant microorganism involves the

expression of a gene encoding a kaurene synthase (KS), a gene encoding a kaurene oxidase (KO), and a gene encoding a steviol synthetase (KAH). Steviol synthetase also is known as kaurenoic acid 13-hydroxylase.

Suitable KS polypeptides are known. For example, suitable KS enzymes
5 include those made by *Stevia rebaudiana*, *Zea mays* and *Populus trichocarpa*. See, SEQ ID NOs: 132-135. Nucleotide sequences encoding these polypeptides are described in more detail below. See, for example, Table 3 and SEQ ID NOs: 40-47.

Suitable KO polypeptides are known. For example, suitable KO enzymes
10 include those made by *Stevia rebaudiana*, *Arabidopsis thaliana*, *Gibberella fujikuroi* and *Trametes versicolor*. See, SEQ ID NOs: 138-141. Nucleotide sequences encoding these polypeptides are described in more detail below. See, for example, Table 5 and SEQ ID NOs: 52-59.

Suitable KAH polypeptides are known. For example, suitable KAH enzymes
15 include those made by *Stevia rebaudiana*, *Arabidopsis thaliana*, *Vitis vinifera* and *Medicago trunculata*. See, e.g., SEQ ID NOs: 142-146; U.S. Patent Publication No. 2008-0271205; U.S. Patent Publication No. 2008-0064063 and Genbank Accession No. gi 189098312. The steviol synthetase from *Arabidopsis thaliana* is classified as a CYP714A2. Nucleotide sequences encoding these polypeptides are described in more detail below. See, for example, Table 6 and SEQ ID NOs: 60-69.

20 In some embodiments, a recombinant microorganism contains a recombinant gene encoding a KO and/or a KAH polypeptide. Such microorganisms also typically contain a recombinant gene encoding a cytochrome P450 reductase (CPR) polypeptide, since certain combinations of KO and/or KAH polypeptides require expression of an exogenous CPR polypeptide. In particular, the activity of a KO
25 and/or a KAH polypeptide of plant origin can be significantly increased by the inclusion of a recombinant gene encoding an exogenous CPR polypeptide. Suitable CPR polypeptides are known. For example, suitable CPR enzymes include those made by *Stevia rebaudiana*, *Arabidopsis thaliana*, and *Gibberella fujikuroi*. See, e.g., SEQ ID NOs: 147-149. Nucleotide sequences encoding these polypeptides are
30 described in more detail below. See, for example, Table 7 and SEQ ID NOs: 70-75.

Expression in a recombinant microorganism of these genes results in the conversion of geranylgeranyl diphosphate to steviol.

B. Steviol Glycoside Biosynthesis Polypeptides

In some embodiments, a recombinant host described herein can convert steviol to a steviol glycoside. Such a host (e.g., microorganism) contains genes encoding one or more UDP Glycosyl Transferases, also known as UGTs. UGTs transfer a monosaccharide unit from an activated nucleotide sugar to an acceptor moiety, in this case, an -OH or -COOH moiety on steviol or steviol derivative. UGTs have been classified into families and subfamilies based on sequence homology. Li et al. J. Biol. Chem. 276:4338-4343 (2001).

B. 1 Rubusoside Biosynthesis Polypeptides

The biosynthesis of rubusoside involves glycosylation of the 13-OH and the 19-COOH of steviol. See FIG. 2A. It has been discovered that conversion of steviol to rubusoside in a recombinant host such as a microorganism can be accomplished by the expression of gene(s) encoding UGTs 85C2 and 74G1, which transfer a glucose unit to the 13-OH or the 19-COOH, respectively, of steviol.

Thus, a suitable UGT85C2 functions as a uridine 5'-diphospho glucosyl: steviol 13-OH transferase, and a uridine 5'-diphospho glucosyl: steviol-19-*O*-glucoside 13-OH transferase. Functional UGT85C2 polypeptides also may catalyze glucosyl transferase reactions that utilize steviol glycoside substrates other than steviol and steviol-19-*O*-glucoside.

A suitable UGT74G1 polypeptide functions as a uridine 5'-diphospho glucosyl: steviol 19-COOH transferase and a uridine 5'-diphospho glucosyl: steviol-13-*O*-glucoside 19-COOH transferase. Functional UGT74G1 polypeptides also may catalyze glucosyl transferase reactions that utilize steviol glycoside substrates other than steviol and steviol-13-*O*-glucoside, or that transfer sugar moieties from donors other than uridine diphosphate glucose.

A recombinant microorganism expressing a functional UGT74G1 and a functional UGT85C2 can make rubusoside and both steviol monosides (i.e., Steviol 13-*O*-monoglucoside and Steviol 19-*O*-monoglucoside) when fed steviol in the medium. One or more of such genes may be present naturally in the host. Typically,

however, such genes are recombinant genes that have been transformed into a host (e.g., microorganism) that does not naturally possess them.

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 incorporated DNA sequence.

5 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 the non-recombinant host. It will be appreciated that typically the genome of a recombinant host described herein is augmented through the stable
10 introduction of one or more recombinant genes. Generally, the introduced DNA is not originally resident in the host that is the recipient of the DNA, but it is within the scope of the invention 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
15 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, plant cells, and plants.

The term "recombinant gene" refers to a gene or DNA sequence that is
20 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 may be a DNA sequence from another species, or may be a DNA sequence that originated from or is present in the same species, but
25 has been incorporated into a host by genetic engineering 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
30 that DNA.

Suitable UGT74G1 and UGT85C2 polypeptides include those made by *Stevia rebaudiana*. Genes encoding functional UGT74G1 and UGT85C2 polypeptides from

Stevia are reported in Richman, et al. Plant J. 41: 56-67 (2005). Amino acid sequences of *S. rebaudiana* UGT74G1 and UGT85C2 polypeptides are set forth in SEQ ID NOs: 1 and 3, respectively. Nucleotide sequences encoding UGT74G1 and UGT85C2 that have been optimized for expression in yeast are set forth in SEQ ID NOs: 2 and 4, respectively. See also the UGT85C2 and UGT74G1 variants described in Examples 17 and 18, respectively.

In some embodiments, the recombinant host is a microorganism. The recombinant microorganism can be grown on media containing steviol in order to produce rubusoside. In other embodiments, however, the recombinant microorganism expresses one or more recombinant genes involved in steviol biosynthesis, e.g., a CDPS gene, a KS gene, a KO gene and/or a KAH gene. Thus, a microorganism containing a CDPS gene, a KS gene, a KO gene and a KAH gene in addition to a UGT74G1 and a UGT85C2 gene is capable of producing both steviol monosides and rubusoside without the necessity for including steviol in the culture media.

In some embodiments, the recombinant microorganism further expresses a recombinant gene encoding a geranylgeranyl diphosphate synthase (GGPPS). Suitable GGPPS polypeptides are known. For example, suitable GGPPS enzymes include those made by *Stevia rebaudiana*, *Gibberella fujikuroi*, *Mus musculus*, *Thalassiosira pseudonana*, *Streptomyces clavuligerus*, *Sulfolobus acidocaldarius*, *Synechococcus* sp. and *Arabidopsis thaliana*. See, SEQ ID NOs: 121-128. Nucleotide sequences encoding these polypeptides are described in more detail below. See Table 1 and SEQ ID NOs: 18-33. In some embodiments, the recombinant microorganism further expresses recombinant genes involved in diterpene biosynthesis or production of terpenoid precursors, e.g., genes in the methylerythritol 4-phosphate (MEP) pathway or genes in the mevalonate (MEV) pathway discussed below.

B. 2 Rebaudioside A Biosynthesis Polypeptides

The biosynthesis of rebaudioside A involves glucosylation of the aglycone steviol. Specifically, rebaudioside A can be formed by glucosylation of the 13-OH of steviol which forms the 13-*O*-steviolmonoside, glucosylation of the C-2' of the 13-*O*-glucose of steviolmonoside which forms steviol-1,2-bioside, glucosylation of the C-

19 carboxyl of steviol-1,2-bioside which forms stevioside, and glucosylation of the C-3' of the C-13-O-glucose of stevioside. The order in which each glucosylation reaction occurs can vary. See FIG. 2A.

It has been discovered that conversion of steviol to rebaudioside A in a recombinant host can be accomplished by the expression of gene(s) encoding the following functional UGTs: 74G1, 85C2, 76G1 and 91D2. Thus, a recombinant microorganism expressing these four UGTs can make rebaudioside A when fed steviol in the medium. Typically, one or more of these genes are recombinant genes that have been transformed into a microorganism that does not naturally possess them. It has also been discovered that UGTs designated herein as SM12UGT can be substituted for UGT91D2.

Suitable UGT74G1 and UGT85C2 polypeptides include those discussed above. A suitable UGT76G1 adds a glucose moiety to the C-3' of the C-13-O-glucose of the acceptor molecule, a steviol 1,2 glycoside. Thus, UGT76G1 functions, for example, as a uridine 5'-diphospho glucosyl: steviol 13-O-1,2 glucoside C-3' glucosyl transferase and a uridine 5'-diphospho glucosyl: steviol-19-O-glucose, 13-O-1,2 bioside C-3' glucosyl transferase. Functional UGT76G1 polypeptides may also catalyze glucosyl transferase reactions that utilize steviol glycoside substrates that contain sugars other than glucose, e.g., steviol rhamnosides and steviol xylosides. Sec, FIGS 2A, 2B, 2C and 2D. Suitable UGT76G1 polypeptides include those made by *S. rebaudiana* and reported in Richman, et al. Plant J. 41: 56-67 (2005). The amino acid sequence of a *S. rebaudiana* UGT76G1 polypeptide is set forth in SEQ ID NO:7. The nucleotide sequence encoding the UGT76G1 polypeptide of SEQ ID NO:7 has been optimized for expression in yeast and is set forth in SEQ ID NO:8. See also the UGT76G1 variants set forth in Example 18.

A suitable UGT91D2 polypeptide functions as a uridine 5'-diphospho glucosyl: steviol-13-O-glucoside transferase (also referred to as a steviol-13-monoglucoside 1,2-glucosylase), transferring a glucose moiety to the C-2' of the 13-O-glucose of the acceptor molecule, steviol-13-O-glucoside. Typically, a suitable UGT91D2 polypeptide also functions as a uridine 5'-diphospho glucosyl: rubusoside transferase transferring a glucose moiety to the C-2' of the 13-O-glucose of the acceptor molecule, rubusoside.

Functional UGT91D2 polypeptides may also catalyze reactions that utilize steviol glycoside substrates other than steviol-13-*O*-glucoside and rubusoside, e.g., functional UGT91D2 polypeptides may utilize stevioside as a substrate, transferring a glucose moiety to the C-2' of the 19-*O*-glucose residue to produce Rebaudioside E.

5 Functional UGT91D2 polypeptides may also utilize Rebaudioside A as a substrate, transferring a glucose moiety to the C-2' of the 19-*O*-glucose residue to produce Rebaudioside D. However, a functional UGT91D2 polypeptide typically does not transfer a glucose moiety to steviol compounds having a 1,3-bound glucose at the C-13 position, i.e., transfer of a glucose moiety to steviol 1,3-bioside and 1,3-stevioside

10 does not occur.

Functional UGT91D2 polypeptides can transfer sugar moieties from donors other than uridine diphosphate glucose. For example, a functional UGT91D2 polypeptide can act as a uridine 5'-diphospho D-xylosyl: steviol-13-*O*-glucoside transferase, transferring a xylose moiety to the C-2' of the 13-*O*-glucose of the

15 acceptor molecule, steviol-13-*O*-glucoside. As another example, a functional UGT91D2 polypeptide can act as a uridine 5'-diphospho L-rhamnosyl: steviol-13-*O*-glucoside transferase, transferring a rhamnose moiety to the C-2' of the 13-*O*-glucose of the acceptor molecule, steviol-13-*O*-glucoside

Suitable functional UGT91D2 polypeptides include those disclosed herein,

20 e.g., the polypeptides designated UGT91D2e and UGT91D2m. The amino acid sequence of an exemplary UGT91D2e polypeptide from *Stevia rebaudiana* is set forth in SEQ ID NO: 5. SEQ ID NO:6 is a nucleotide sequence encoding the polypeptide of SEQ ID NO:5 that has been codon optimized for expression in yeast. The *S. rebaudiana* nucleotide sequence encoding the polypeptide of SEQ ID NO:5 is set

25 forth in SEQ ID NO:9. The amino acid sequences of exemplary UGT91D2m polypeptides from *S. rebaudiana* are set forth in SEQ ID NOs: 10 and 12, and are encoded by the nucleic acid sequences set forth in SEQ ID NOs: 11 and 13, respectively. See also the UGT91D2 variants of Example 16, e.g., a variant containing a substitution at amino acid residues 206, 207, and 343.

30 As indicated above, UGTs designated herein as SM12UGT can be substituted for UGT91D2. Suitable functional SM12UGT polypeptides include those made by *Ipomoea purpurea* (Japanese morning glory) and described in Morita et al. Plant J. 42,

353-363 (2005). The amino acid sequence encoding the *I. purpurea* IP3GGT polypeptide is set forth in SEQ ID NO:76. SEQ ID NO:77 is a nucleotide sequence encoding the polypeptide of SEQ ID NO:76 that has been codon optimized for expression in yeast. Another suitable SM12UGT polypeptide is a Bp94B1 polypeptide having an R25S mutation. See Osmani et al. Plant Phys. 148: 1295-1308 (2008) and Sawada et al. J. Biol. Chem. 280:899-906 (2005). The amino acid sequence encoding the *Bellis perennis* (red daisy) UGT94B1 polypeptide is set forth in SEQ ID NO:78. SEQ ID NO:79 is the nucleotide sequence encoding the polypeptide of SEQ ID NO:78 that has been codon optimized for expression in yeast.

10 In some embodiments, the recombinant microorganism is grown on media containing steviol-13-*O*-glucoside or steviol-19-*O*-glucoside in order to produce rebaudioside A. In such embodiments, the microorganism contains and expresses genes encoding a functional UGT91D2, a functional UGT74G1 and a functional UGT76G1, and is capable of producing rebaudioside A when it is fed steviol, one or
15 both of the steviolmonosides, or rubusoside in the culture media.

In other embodiments, the recombinant microorganism is grown on media containing rubusoside in order to produce rebaudioside A. In such embodiments, the microorganism contains and expresses genes encoding a functional UGT91D2 and a functional UGT76G1, and is capable of producing rebaudioside A when it is fed
20 rubusoside in the culture media.

In other embodiments the recombinant microorganism expresses one or more genes involved in steviol biosynthesis, e.g., a CDPS gene, a KS gene, a KO gene and/or a KAH gene. Thus, for example, a microorganism containing a CDPS gene, a KS gene, a KO gene and a KAH gene, in addition to a UGT74G1, a UGT85C2, a
25 UGT91D2 gene and a UGT76G1 gene, is capable of producing rebaudioside A without the necessity for including steviol in the culture media.

In some embodiments, the recombinant microorganism further contains and expresses a recombinant GGPPS gene in order to provide increased levels of the diterpene precursor geranylgeranyl diphosphate, for increased flux through the
30 rebaudioside A biosynthetic pathway. In some embodiments, the recombinant microorganism further contains and expresses recombinant genes involved in

diterpene biosynthesis or production of terpenoid precursors, e.g., genes in the MEP or MEV pathway discussed below.

B. 3 Dulcoside A and Rebaudioside C Biosynthesis Polypeptides

5 The biosynthesis of rebaudioside C and/or dulcoside A involves glucosylation and rhamnosylation of the aglycone steviol. Specifically, dulcoside A can be formed by glucosylation of the 13-OH of steviol which forms steviol-13-*O*-glucoside, rhamnosylation of the C-2' of the 13-*O*-glucose of steviol-13-*O*-glucoside which forms the 1,2 rhamnobioside, and glucosylation of the C-19 carboxyl of the 1,2
10 rhamnobioside. Rebaudioside C can be formed by glucosylation of the C-3' of the C-13-*O*-glucose of dulcoside A. The order in which each glycosylation reaction occurs can vary. See FIG. 2B.

 It has been discovered that conversion of steviol to dulcoside A in a recombinant host can be accomplished by the expression of gene(s) encoding the
15 following functional UGTs: 85C2, 91D2, and 74G1. Thus, a recombinant microorganism expressing these three UGTs and a rhamnose synthetase can make dulcoside A when fed steviol in the medium. Alternatively, a recombinant microorganism expressing two UGTs, 91D2 and 74G1, and rhamnose synthetase can make dulcoside A when fed the monoside, steviol-13-*O*-glucoside or steviol-19-*O*-
20 glucoside, in the medium. Similarly, conversion of steviol to rebaudioside C in a recombinant microorganism can be accomplished by the expression of gene(s) encoding UGTs 85C2, 91D2, 74G1, and 76G1 and rhamnose synthetase when fed steviol, by the expression of genes encoding UGTs 91D2, 74G1 and 76G1, and rhamnose synthetase when fed steviol-13-*O*-glucoside, by the expression of genes
25 encoding UGTs 85C2, 91D2 and 76G1, and rhamnose synthetase when fed steviol-19-*O*-glucoside, or by the expression of genes encoding UGTs 91D2 and 76G1 and rhamnose synthetase when fed rubusoside. Typically, one or more of these genes are recombinant genes that have been transformed into a microorganism that does not naturally possess them.

30 Suitable UGT91D2, UGT74G1, UGT76G1 and UGT85C2 polypeptides include the functional UGT polypeptides discussed herein. Rhamnose synthetase provides increased amounts of the UDP-rhamnose donor for rhamnosylation of the

steviol compound acceptor. Suitable rhamnose synthetases include those made by *Arabidopsis thaliana*, such as the product of the *A. thaliana* RHM2 gene.

In some embodiments, a UGT79B3 polypeptide is substituted for a UGT91D2 polypeptide. Suitable UGT79B3 polypeptides include those made by *Arabidopsis thaliana*, which are capable of rhamnosylation of steviol 13-O-monoside *in vitro*. *A. thaliana* UGT79B3 can rhamnosylate glucosylated compounds to form 1,2-rhamnosides. The amino acid sequence of an *Arabidopsis thaliana* UGT79B3 is set forth in SEQ ID NO:150. The nucleotide sequence encoding the amino acid sequence of SEQ ID NO:150 is set forth in SEQ ID NO:151.

In some embodiments rebaudioside C can be produced using *in vitro* methods while supplying the appropriate UDP-sugar or a cell-free system for regeneration of UDP-sugars. See, for example, “An integrated cell-free metabolic platform for protein production and synthetic biology” by Jewett MC, Calhoun KA, Voloshin A, Wu JJ and Swartz JR in **Molecular Systems Biology**, 4, article 220 (2008).

Reactions may be carried out together, or stepwise. For instance, rebaudioside C may be produced from rubusoside with the addition of stoichiometric amounts of UDP-rhamnose and UGT91d2e, followed by addition of UGT76G1 and an excess or stoichiometric supply of UDP-glucose. In some embodiments phosphatases are used to remove secondary products and improve the reaction yields.

In other embodiments, the recombinant host expresses one or more genes involved in steviol biosynthesis, e.g., a CDPS gene, a KS gene, a KO gene and/or a KAH gene. Thus, for example, a microorganism containing a CDPS gene, a KS gene, a KO gene and a KAH gene, in addition to a UGT85C2, a UGT74G1, a UGT91D2 gene and a UGT76G1 gene, is capable of producing rebaudioside C without the necessity for including steviol in the culture media. In addition, the recombinant host typically expresses an endogenous or a recombinant gene encoding a rhamnose synthetase. Such a gene is useful in order to provide increased amounts of the UDP-rhamnose donor for rhamnosylation of the steviol compound acceptor. Suitable rhamnose synthetases include those made by *Arabidopsis thaliana*, such as the product of the *A. thaliana* RHM2 gene.

One with skill in the art will recognize that by modulating relative expression levels of different UGT genes as well as modulating the availability of UDP-

5 rhamnose, a recombinant host can be tailored to specifically produce steviol and steviol glycoside products in a desired proportion. Transcriptional regulation of steviol biosynthesis genes, and steviol glycoside biosynthesis genes can be achieved by a combination of transcriptional activation and repression using techniques known to those in the art. For *in vitro* reactions, one with skill in the art will recognize that addition of different levels of UGT enzymes in combination or under conditions which impact the relative activities of the different UGTs in combination will direct synthesis towards a desired proportion of each steviol glycoside.

10 In some embodiments, the recombinant host further contains and expresses a recombinant GGPPS gene in order to provide increased levels of the diterpene precursor geranylgeranyl diphosphate, for increased flux through the rebaudioside A biosynthetic pathway. In some embodiments, the recombinant host further contains a genetic construct to silence or reduce the expression of non-steviol pathways consuming geranylgeranyl diphosphate, ent-Kaurenoic acid or farnesyl
15 pyrophosphate, thereby providing increased flux through the steviol and steviol glycosides biosynthetic pathways. For example, flux to sterol production pathways such as ergosterol may be reduced by downregulation of the ERG9 gene. In cells that produce gibberellins, gibberellin synthesis may be downregulated to increase flux of ent-kaurenoic acid to steviol. In carotenoid-producing organisms, flux to steviol may be increased by downregulation of one or more carotenoid biosynthetic genes.
20

In some embodiments, the recombinant host further contains and expresses recombinant genes involved in diterpene biosynthesis or production of terpenoid precursors, e.g., genes in the MEP or MEV pathway discussed below.

25 In some embodiments, a recombinant host such as a microorganism produces steviol glycoside compositions that have greater than at least 15% rebaudioside C of the total steviol glycosides, e.g., at least 20% rebaudioside C, 30-40% rebaudioside C, 40-50% rebaudioside C, 50-60% rebaudioside C, 60-70% rebaudioside C, 70-80% rebaudioside C, 80-90% rebaudioside C. In some embodiments, a recombinant host such as a microorganism produces steviol glycoside compositions that have at least
30 90% rebaudioside C, e.g., 90-99% rebaudioside C. Other steviol glycosides present may include those depicted in Figures 2 A and B such as steviol monosides, steviol glucobiosides, steviol rhamnobiosides, rebaudioside A, and Dulcoside A. In some

embodiments, the rebaudioside C-enriched composition produced by the host can be further purified and the rebaudioside C or Dulcoside A so purified may then be mixed with other steviol glycosides, flavors, or sweeteners to obtain a desired flavor system or sweetening composition. For instance, a rebaudioside C-enriched composition
5 produced by a recombinant microorganism can be combined with a rebaudioside A, F, or D-enriched composition produced by a different recombinant microorganism, with rebaudioside A, F, or D purified from a Stevia extract, or with rebaudioside A, F, or D produced *in vitro*.

10 B. 4 Rebaudioside E and Rebaudioside D Biosynthesis Polypeptides

The biosynthesis of rebaudioside E and/or rebaudioside D involves glucosylation of the aglycone steviol. Specifically, rebaudioside E can be formed by glucosylation of the 13-OH of steviol which forms steviol-13-*O*-glucoside, glucosylation of the C-2' of the 13-*O*-glucose of steviol-13-*O*-glucoside which forms
15 the steviol-1,2-bioside, glucosylation of the C-19 carboxyl of the 1,2-bioside to form 1,2-stevioside, and glucosylation of the C-2' of the 19-*O*-glucose of the 1,2-stevioside to form rebaudioside E. Rebaudioside D can be formed by glucosylation of the C-3' of the C-13-*O*-glucose of rebaudioside E. The order in which each glycosylation reaction occurs can vary. For example, the glucosylation of the C-2' of the 19-*O*-
20 glucose may be the last step in the pathway, wherein Rebaudioside A is an intermediate in the pathway. See FIG. 2C.

It has been discovered that conversion of steviol to rebaudioside D in a recombinant host can be accomplished by the expression of gene(s) encoding the following functional UGTs: 85C2, 91D2, 74G1 and 76G1. Thus, a recombinant
25 microorganism expressing these four UGTs can make rebaudioside D when fed steviol in the medium. Alternatively, a recombinant microorganism expressing two functional UGTs, 91D2 and 76G1, can make rebaudioside D when fed rubusoside or 1,2-stevioside in the medium. As another alternative, a recombinant microorganism expressing three functional UGTs, 74G1, 91D2 and 76G1, can make rebaudioside D
30 when fed the monoside, steviol-13-*O*-glucoside, in the medium. Similarly, conversion of steviol-19-*O*-glucoside to rebaudioside D in a recombinant microorganism can be accomplished by the expression of genes encoding UGTs 85C2, 91D2 and 76G1

when fed steviol-19-*O*-glucoside. Typically, one or more of these genes are recombinant genes that have been transformed into a host that does not naturally possess them.

Suitable UGT91D2, UGT74G1, UGT76G1 and UGT85C2 polypeptides
5 include the functional UGT polypeptides discussed herein. In some embodiments, a UGT79B3 polypeptide is substituted for a UGT91, as discussed above.

In some embodiments, rebaudioside D or rebaudioside E can be produced using *in vitro* methods while supplying the appropriate UDP-sugar or a cell-free system for regeneration of UDP-sugars. See, for example, Jewett MC, et al.
10 **Molecular Systems Biology**, Vol. 4, article 220 (2008). Conversions requiring multiple reactions may be carried out together, or stepwise. Rebaudioside D may be produced from Rebaudioside A that is commercially available enriched extract or produced via biosynthesis, with the addition of stoichiometric or excess amounts of UDP-glucose and UGT91D2e. In some embodiments phosphatases are used to
15 remove secondary products and improve the reaction yields.

One with skill in the art will recognize that by modulating relative expression levels of different UGT genes, a recombinant host can be tailored to specifically produce steviol and steviol glycoside products in a desired proportion. Transcriptional regulation of steviol biosynthesis genes and steviol glycoside
20 biosynthesis genes can be achieved by a combination of transcriptional activation and repression using techniques known to those in the art. For *in vitro* reactions, one with skill in the art will recognize that addition of different levels of UGT enzymes in combination or under conditions which impact the relative activities of the different UGTs in combination will direct synthesis towards a desired proportion of each
25 steviol glycoside. One with skill in the art will recognize that a higher proportion of rebaudioside D or E or more efficient conversion to rebaudioside D or E can be obtained with a diglycosylation enzyme that has a higher activity for the 19-*O*-glucoside reaction as compared to the 13-*O*-glucoside reaction (substrates rebaudioside A and stevioside).

30 In other embodiments, the recombinant host expresses one or more genes involved in steviol biosynthesis, e.g., a CDPS gene, a KS gene, a KO gene and/or a KAH gene. Thus, for example, a microorganism containing a CDPS gene, a KS gene,

a KO gene and a KAH gene, in addition to a UGT85C2, a UGT74G1, a UGT91D2 gene and a UGT76G1 gene, is capable of producing rebaudiosides E and D without the necessity for including steviol in the culture media.

5 In some embodiments, the recombinant host further contains and expresses a recombinant GGPPS gene in order to provide increased levels of the diterpene precursor geranylgeranyl diphosphate, for increased flux through the steviol biosynthetic pathway. In some embodiments, the recombinant host further contains a genetic construct to silence the expression of non-steviol pathways consuming geranylgeranyl diphosphate, ent-Kaurenoic acid or farnesyl pyrophosphate, thereby
10 providing increased flux through the steviol and steviol glycosides biosynthetic pathways. For example, flux to sterol production pathways such as ergosterol may be reduced by downregulation of the ERG9 gene. In cells that produce gibberellins, gibberellin synthesis may be downregulated to increase flux of ent-kaurenoic acid to steviol. In carotenoid-producing organisms, flux to steviol may be increased by
15 downregulation of one or more carotenoid biosynthetic genes. In some embodiments, the recombinant host further contains and expresses recombinant genes involved in diterpene biosynthesis or production of terpenoid precursors, e.g., genes in the MEP or MEV pathways discussed below.

In some embodiments, a recombinant host such as a microorganism produces
20 rebaudioside D-enriched steviol glycoside compositions that have greater than at least 3% rebaudioside D by weight total steviol glycosides, e.g., at least 4% rebaudioside D at least 5% rebaudioside D, 10-20% rebaudioside D, 20-30% rebaudioside D, 30-40% rebaudioside D, 40-50% rebaudioside D, 50-60% rebaudioside D, 60-70% rebaudioside D, 70-80% rebaudioside D. In some embodiments, a recombinant host
25 such as a microorganism produces steviol glycoside compositions that have at least 90% rebaudioside D, e.g., 90-99% rebaudioside D. Other steviol glycosides present may include those depicted in Figure 2 C such as steviol monosides, steviol glucobiosides, rebaudioside A, rebaudioside E, and stevioside. In some embodiments, the rebaudioside D-enriched composition produced by the host (e.g., microorganism)
30 can be further purified and the rebaudioside D or rebaudioside E so purified can then be mixed with other steviol glycosides, flavors, or sweeteners to obtain a desired flavor system or sweetening composition. For instance, a rebaudioside D-enriched

composition produced by a recombinant host can be combined with a rebaudioside A, C, or F-enriched composition produced by a different recombinant host, with rebaudioside A, F, or C purified from a Stevia extract, or with rebaudioside A, F, or C produced *in vitro*.

5

B. 5 Rebaudioside F Biosynthesis Polypeptides

The biosynthesis of rebaudioside F involves glucosylation and xylosylation of the aglycone steviol. Specifically, rebaudioside F can be formed by glucosylation of the 13-OH of steviol which forms steviol-13-*O*-glucoside, xylosylation of the C-2' of the 13-*O*-glucose of steviol-13-*O*-glucoside which forms steviol-1,2-xylobioside, 10 the 13-*O*-glucose of steviol-13-*O*-glucoside which forms steviol-1,2-xylobioside, glucosylation of the C-19 carboxyl of the 1,2-xylobioside to form 1,2-stevioxyloside, and glucosylation of the C-3' of the C-13-*O*-glucose of 1,2-stevioxyloside to form rebaudioside F. The order in which each glycosylation reaction occurs can vary. See FIG. 2D.

15 It has been discovered that conversion of steviol to rebaudioside F in a recombinant host can be accomplished by the expression of genes encoding the following functional UGTs: 85C2, 91D2, 74G1 and 76G1, along with endogenous or recombinantly expressed UDP-glucose dehydrogenase and UDP-glucuronic acid decarboxylase. Thus, a recombinant microorganism expressing these four UGTs 20 along with endogenous or recombinant UDP-glucose dehydrogenase and UDP-glucuronic acid decarboxylase can make rebaudioside F when fed steviol in the medium. Alternatively, a recombinant microorganism expressing two functional UGTs, 91D2 and 76G1, can make rebaudioside F when fed rubusoside in the medium. As another alternative, a recombinant microorganism expressing a functional UGT 25 76G1 can make rebaudioside F when fed 1,2 steviorhamnoside. As another alternative, a recombinant microorganism expressing three functional UGTs, 74G1, 91D2 and 76G1, can make rebaudioside F when fed the monoside, steviol-13-*O*-glucoside, in the medium. Similarly, conversion of steviol-19-*O*-glucoside to rebaudioside F in a recombinant microorganism can be accomplished by the 30 expression of genes encoding UGTs 85C2, 91D2 and 76G1 when fed steviol-19-*O*-glucoside. Typically, one or more of these genes are recombinant genes that have been transformed into a host that does not naturally possess them.

Suitable UGT91D2, UGT74G1, UGT76G1 and UGT85C2 polypeptides include the functional UGT polypeptides discussed herein. In some embodiments, a UGT79B3 polypeptide is substituted for a UGT91, as discussed above. UDP-glucose dehydrogenase and UDP-glucuronic acid decarboxylase provide increased amounts of the UDP-xylose donor for xylosylation of the steviol compound acceptor. Suitable UDP-glucose dehydrogenases and UDP-glucuronic acid decarboxylases include those made by *Arabidopsis thaliana* or *Cryptococcus neoformans*. For example, suitable UDP-glucose dehydrogenase and UDP-glucuronic acid decarboxylases polypeptides can be encoded by the *A. thaliana* UGD1 gene and UXS3 gene, respectively. See, Oka and Jigami, FEBS J. 273:2645-2657 (2006).

In some embodiments rebaudioside F can be produced using *in vitro* methods while supplying the appropriate UDP-sugar or a cell-free system for regeneration of UDP-sugars. See, for example, Jewett MC, et al. **Molecular Systems Biology**, Vol. 4, article 220 (2008). Reactions may be carried out together, or stepwise. For instance, rebaudioside F may be produced from rubusoside with the addition of stoichiometric amounts of UDP-xylose and UGT91D2e, followed by addition of UGT76G1 and an excess or stoichiometric supply of UDP-glucose. In some embodiments phosphatases are used to remove secondary products and improve the reaction yields.

In other embodiments, the recombinant host expresses one or more genes involved in steviol biosynthesis, e.g., a CDPS gene, a KS gene, a KO gene and/or a KAH gene. Thus, for example, a microorganism containing a CDPS gene, a KS gene, a KO gene and a KAH gene, in addition to a UGT85C2, a UGT74G1, a UGT91D2 gene and a UGT76G1 gene, is capable of producing rebaudioside F without the necessity for including steviol in the culture media. In addition, the recombinant host typically expresses an endogenous or a recombinant gene encoding a UDP-glucose dehydrogenase and a UDP-glucuronic acid decarboxylase. Such genes are useful in order to provide increased amounts of the UDP-xylose donor for xylosylation of the steviol compound acceptor. Suitable UDP-glucose dehydrogenases and UDP-glucuronic acid decarboxylases include those made by *Arabidopsis thaliana* or *Cryptococcus neoformans*. For example, suitable UDP-glucose dehydrogenase and UDP-glucuronic acid decarboxylases polypeptides can be encoded by the *A. thaliana*

UGD1 gene and UXS3 gene, respectively. See, Oka and Jigami, FEBS J. 273:2645-2657 (2006).

One with skill in the art will recognize that by modulating relative expression levels of different UGT genes as well as modulating the availability of UDP-xylose, a recombinant microorganism can be tailored to specifically produce steviol and steviol glycoside products in a desired proportion. Transcriptional regulation of steviol biosynthesis genes can be achieved by a combination of transcriptional activation and repression using techniques known to those in the art. For *in vitro* reactions, one with skill in the art will recognize that addition of different levels of UGT enzymes in combination or under conditions which impact the relative activities of the different UGTs in combination will direct synthesis towards a desired proportion of each steviol glycosides.

In some embodiments, the recombinant host further contains and expresses a recombinant GGPPS gene in order to provide increased levels of the diterpene precursor geranylgeranyl diphosphate, for increased flux through the steviol biosynthetic pathway. In some embodiments, the recombinant host further contains a genetic construct to silence the expression of non-steviol pathways consuming geranylgeranyl diphosphate, ent-Kaurenoic acid or farnesyl pyrophosphate, thereby providing increased flux through the steviol and steviol glycosides biosynthetic pathways. For example, flux to sterol production pathways such as ergosterol may be reduced by downregulation of the ERG9 gene. In cells that produce gibberellins, gibberellin synthesis may be downregulated to increase flux of ent-kaurenoic acid to steviol. In carotenoid-producing organisms, flux to steviol may be increased by downregulation of one or more carotenoid biosynthetic genes. In some embodiments, the recombinant host further contains and expresses recombinant genes involved in diterpene biosynthesis, e.g., genes in the MEP pathway discussed below.

In some embodiments, a recombinant host such as a microorganism produces rebaudioside F-enriched steviol glycoside compositions that have greater than at least 4% rebaudioside F by weight total steviol glycosides, e.g., at least 5% rebaudioside F, at least 6% of rebaudioside F, 10-20% rebaudioside F, 20-30% rebaudioside F, 30-40% rebaudioside F, 40-50% rebaudioside F, 50-60% rebaudioside F, 60-70% rebaudioside F, 70-80% rebaudioside F. In some embodiments, a recombinant host

such as a microorganism produces steviol glycoside compositions that have at least 90% rebaudioside F, e.g., 90-99% rebaudioside F. Other steviol glycosides present may include those depicted in Figure 2 A and D such as steviol monosides, steviol glucobiosides, steviol xylobiosides, rebaudioside A, stevioxyloside, rubusoside and stevioside. In some embodiments, the rebaudioside F-enriched composition produced by the host can be mixed with other steviol glycosides, flavors, or sweeteners to obtain a desired flavor system or sweetening composition. For instance, a rebaudioside F-enriched composition produced by a recombinant microorganism can be combined with a rebaudioside A, C, or D-enriched composition produced by a different recombinant microorganism, with rebaudioside A, C, or D purified from a Stevia extract, or with rebaudioside A, C, or D produced *in vitro*.

C. Other Polypeptides

Genes for additional polypeptides whose expression facilitates more efficient or larger scale production of steviol or a steviol glycoside can also be introduced into a recombinant host. For example, a recombinant microorganism, plant, or plant cell can also contain one or more genes encoding a geranylgeranyl diphosphate synthase (GGPPS, also referred to as GGDPS). As another example, the recombinant host can contain one or more genes encoding a rhamnose synthetase, or one or more genes encoding a UDP-glucose dehydrogenase and/or a UDP-glucuronic acid decarboxylase. As another example, a recombinant host can also contain one or more genes encoding a cytochrome P450 reductase (CPR). Expression of a recombinant CPR facilitates the cycling of NADP⁺ to regenerate NADPH, which is utilized as a cofactor for terpenoid biosynthesis. Other methods can be used to regenerate NADHP levels as well. In circumstances where NADPH becomes limiting; strains can be further modified to include exogenous transhydrogenase genes. See, e.g., Sauer et al., J. Biol. Chem. 279: 6613–6619 (2004). Other methods are known to those with skill in the art to reduce or otherwise modify the ratio of NADH/NADPH such that the desired cofactor level is increased.

As another example, the recombinant host can contain one or more genes encoding one or more enzymes in the MEP pathway or the mevalonate pathway. Such genes are useful because they can increase the flux of carbon into the diterpene

biosynthesis pathway, producing geranylgeranyl diphosphate from isopentenyl diphosphate and dimethylallyl diphosphate generated by the pathway. The geranylgeranyl diphosphate so produced can be directed towards steviol and steviol glycoside biosynthesis due to expression of steviol biosynthesis polypeptides and steviol glycoside biosynthesis polypeptides.

C. 1 MEP Biosynthesis Polypeptides

In some embodiments, a recombinant host contains one or more genes encoding enzymes involved in the methylerythritol 4-phosphate (MEP) pathway for isoprenoid biosynthesis. Enzymes in the MEP pathway include deoxyxylulose 5-phosphate synthase (DXS), D-1-deoxyxylulose 5-phosphate reductoisomerase (DXR), 4-diphosphocytidyl-2-C-methyl-D-erythritol synthase (CMS), 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase (CMK), 4-diphosphocytidyl-2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase (MCS), 1-hydroxy-2-methyl-2(E)-butenyl 4-diphosphate synthase (HDS) and 1-hydroxy-2-methyl-2(E)-butenyl 4-diphosphate reductase (HDR). One or more DXS genes, DXR genes, CMS genes, CMK genes, MCS genes, HDS genes and/or HDR genes can be incorporated into a recombinant microorganism. See, Rodríguez-Concepción and Boronat, Plant Phys. 130: 1079-1089 (2002).

Suitable genes encoding DXS, DXR, CMS, CMK, MCS, HDS and/or HDR polypeptides include those made by *E. coli*, *Arabidopsis thaliana* and *Synechococcus leopoliensis*. Nucleotide sequences encoding DXR polypeptides are described, for example, in U.S. Patent No. 7,335,815.

25 C. 2 Mevalonate Biosynthesis Polypeptides

In some embodiments, a recombinant host contains one or more genes encoding enzymes involved in the mevalonate pathway for isoprenoid biosynthesis. Genes suitable for transformation into a host encode enzymes in the mevalonate pathway such as a truncated 3-hydroxy-3-methyl-glutaryl (HMG)-CoA reductase (tHMG), and/or a gene encoding a mevalonate kinase (MK), and/or a gene encoding a phosphomevalonate kinase (PMK), and/or a gene encoding a mevalonate pyrophosphate decarboxylase (MPPD). Thus, one or more HMG-CoA reductase

genes, MK genes, PMK genes, and/or MPPD genes can be incorporated into a recombinant host such as a microorganism.

Suitable genes encoding mevalonate pathway polypeptides are known. For example, suitable polypeptides include those made by *E. coli*, *Paracoccus denitrificans*, *Saccharomyces cerevisiae*, *Arabidopsis thaliana*, *Kitasatospora griseola*, *Homo sapiens*, *Drosophila melanogaster*, *Gallus gallus*, *Streptomyces* sp. KO-3988, *Nicotiana attenuata*, *Kitasatospora griseola*, *Hevea brasiliensis*, *Enterococcus faecium* and *Haematococcus pluvialis*. See, e.g., U.S. Patent Nos. 7,183,089, 5,460,949, and 5,306,862.

D. Functional Homologs

Functional homologs of the polypeptides described above are also suitable for use in producing steviol or 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 may be natural occurring polypeptides, and the sequence similarity may 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, may 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 UGT 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.

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 or steviol glycoside biosynthesis polypeptides. Sequence analysis can involve BLAST, Reciprocal BLAST, or PSI-BLAST analysis of nonredundant databases using a GGPPS, a CDPS, a KS, a KO or a KAH 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 or 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 biosynthesis polypeptides, *e.g.*, conserved functional domains.

Conserved regions can be identified by locating a region within the primary amino acid sequence of a steviol or 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.

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.

For example, polypeptides suitable for producing steviol glycosides in a recombinant host include functional homologs of UGT91D2e, UGT91D2m, UGT85C, and UGT76G. Such homologs have greater than 90% (e.g., at least 95% or 99%) sequence identity to the amino acid sequence of UGT91D2e (SEQ ID NO:5), UGT91D2m (SEQ ID NO:10), UGT85C (SEQ ID NO:3), or UGT76G (SEQ ID NO:7). Variants of UGT91D2, UGT85C, and UGT76G polypeptides typically have 10 or fewer amino acid substitutions within the primary amino acid sequence, e.g., 7 or fewer amino acid substitutions, 5 or conservative amino acid substitutions, or between 1 and 5 substitutions. However, in some embodiments, variants of UGT91D2, UGT85C, and UGT76G polypeptides can have 10 or more amino acid substitutions (e.g., 10, 15, 20, 25, 30, 35, 10-20, 10-35, 20-30, or 25-35 amino acid substitutions). The substitutions may be conservative, or in some embodiments, non-conservative. Non-limiting examples of non-conservative changes in UGT91D2e polypeptides include glycine to arginine and tryptophan to arginine. Non-limiting examples of non-conservative substitutions in UGT76G polypeptides include valine to glutamic acid, glycine to glutamic acid, glutamine to alanine, and serine to proline. Non-limiting examples of changes to UGT85C polypeptides include histidine to aspartic acid, proline to serine, lysine to threonine, and threonine to arginine.

In some embodiments, a useful UGT91D2 homolog can have amino acid substitutions (e.g., conservative amino acid substitutions) in regions of the polypeptide that are outside of predicted loops, e.g., residues 20-26, 39-43, 88-95, 121-124, 142-158, 185-198, and 203-214 are predicted loops in the N-terminal domain and residues 381-386 are predicted loops in the C-terminal domain of SEQ ID NO:5. For example, a useful UGT91D2 homolog 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. In some embodiments, a UGT91D2 homolog can have 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. For example, a UGT91D2

functional homolog can have an amino acid substitution at one or more of residues 206, 207, and 343, such as an arginine at residue 206, a cysteine at residue 207, and an arginine at residue 343 of SEQ ID NO:5. See, SEQ ID NO:95. Other functional homologs of UGT91D2 can have one or more of the following: a tyrosine or
5 phenylalanine at residue 30, a proline or glutamine at residue 93, a serine or valine at residue 99, a tyrosine or a 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
10 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, a lysine or asparagine at residue 427, an alanine at residue 438, and either an alanine or threonine at residue 462 of SEQ ID NO:5. See, Examples 11 and 16, and Tables 12 and 14. A useful
15 variant UGT91D2 polypeptide also can be constructed based on the alignment set forth in Figure 8.

In some embodiments, a useful UGT85C homolog can have 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
20 SEQ ID NO:3. Non-limiting examples of useful UGT85C homologs include polypeptides having substitutions (with respect to SEQ ID NO:3) at residue 65; at residue 65 in combination with residue 15, 270, 418, 440, or 441; residues 13, 15, 60, 270, 289, and 418; substitutions at residues 13, 60, and 270; substitutions at residues 60 and 87; substitutions at residues 65, 71, 220, 243, and 270; substitutions at residues
25 65, 71, 220, 243, 270, and 441; substitutions at residues 65, 71, 220, 389, and 394; substitutions at residues 65, 71, 270, and 289; substitutions at residues 220, 243, 270, and 334; or substitutions at residues 270 and 289. See, Example 17 and Table 15.

In some embodiments, a useful UGT76G homolog can have one or more amino acid substitutions at residues 29, 74, 87, 91, 116, 123, 125, 126, 130, 145, 192,
30 193, 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. Non-limiting examples of useful UGT76G homologs include polypeptides having substitutions (with respect to SEQ ID NO:7) at

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; 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 residues 74 , 87, 91, 116, 123, 125, 126, 130, 145, 192, 193,
5 194, 196, 198, 199, 200, 203, 204, 205, 206, 207, 208, 266, 273, 274, 284, 285, 291, 330, 331, and 346. See, Example 18 and Table 16.

Methods to modify the substrate specificity of, for example UGT91D2e, are known to those skilled in the art, and include without limitation site-directed/rational mutagenesis approaches, random directed evolution approaches and combinations in
10 which random mutagenesis/saturation techniques are performed near the active site of the enzyme. For example see Sarah A. Osmani, et al. *Phytochemistry* 70 (2009) 325–347.

A candidate sequence typically has a length that is from 80 percent to 200 percent of the length of the reference sequence, *e.g.*, 82, 85, 87, 89, 90, 93, 95, 97, 99,
15 100, 105, 110, 115, 120, 130, 140, 150, 160, 170, 180, 190, or 200 percent of the length of the reference sequence. A percent 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) is aligned to one or more candidate sequences using the computer program
20 ClustalW (version 1.83, default parameters), which allows alignments of nucleic acid or polypeptide sequences to be carried out across their entire length (global alignment). Chenna *et al.*, *Nucleic Acids Res.*, 31(13):3497-500 (2003).

ClustalW calculates the best match between a reference and one or more candidate sequences, and aligns them so that identities, similarities and differences
25 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: percentage; number of top diagonals: 4; and gap penalty: 5. For multiple alignment of nucleic acid sequences,
30 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:

percentage; 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 (cbi.ac.uk/clustalw).

To determine percent identity of a candidate nucleic acid or amino acid sequence to a reference sequence, the sequences are aligned using ClustalW, 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 percent 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.

It will be appreciated that a functional UGT91D2 polypeptide can include additional amino acids that are not involved in glucosylation or other enzymatic activities carried out by UGT91D2, and thus such a polypeptide can be longer than would otherwise be the case. For example, a UGT91D2 polypeptide can include a purification tag, a chloroplast transit peptide, a mitochondrial transit peptide, an amyloplast peptide, signal peptide, or a secretion tag added to the amino or carboxy terminus. In some embodiments, a UGT91D2 polypeptide includes an amino acid sequence that functions as a reporter, *e.g.*, a green fluorescent protein or yellow fluorescent protein.

II. Steviol and Steviol Glycoside Biosynthesis Nucleic Acids

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.

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.

5 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,
10 *e.g.*, introns, enhancers, upstream activation regions, transcription terminators, and inducible elements.

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,
15 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
20 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.

25 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
30 bias tables for that host (*e.g.*, microorganism). SEQ ID NOs:18-25, 34-36, 40-43, 48-49, 52-55, 60-64, and 70-72 set forth nucleotide sequences encoding certain enzymes for steviol and steviol glycoside biosynthesis, modified for increased expression in

yeast. 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.

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. As another example, expression of membrane transporters involved in transport of steviol glycosides can be inhibited, such that secretion of glycosylated steviol glycosides is inhibited. Such regulation can be beneficial in that secretion of steviol glycosides can be inhibited for a desired period of time during culture of the microorganism, thereby increasing the yield of glycoside product(s) at harvest. In such cases, a nucleic acid that inhibits expression of 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 inhibit function.

III. Hosts

A. Microorganisms

A number of prokaryotes and eukaryotes are 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 or 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 assembled in one or more recombinant constructs, which are then transformed into the strain in order to supply the missing function(s).

Exemplary prokaryotic and eukaryotic species are described in more detail below. However, it will be appreciated that other species may be suitable. For

example, suitable species may be in a genus selected from the group consisting of *Agaricus*, *Aspergillus*, *Bacillus*, *Candida*, *Corynebacterium*, *Escherichia*, *Fusarium*/*Gibberella*, *Kluyveromyces*, *Laetiporus*, *Lentinus*, *Phaffia*, *Phanerochaete*, *Pichia*, *Physcomitrella*, *Rhodoturula*, *Saccharomyces*, *Schizosaccharomyces*, *Sphaceloma*, *Xanthophyllomyces* and *Yarrowia*. Exemplary species from such genera include *Lentinus tigrinus*, *Laetiporus sulphureus*, *Phanerochaete chrysosporium*, *Pichia pastoris*, *Physcomitrella patens*, *Rhodoturula glutinis* 32, *Rhodoturula mucilaginoso*, *Phaffia rhodozyma* UBV-AX, *Xanthophyllomyces dendrorhous*, *Fusarium fujikuroi*/*Gibberella fujikuroi*, *Candida utilis* and *Yarrowia lipolytica*. In some embodiments, a microorganism can be an Ascomycete such as *Gibberella fujikuroi*, *Kluyveromyces lactis*, *Schizosaccharomyces pombe*, *Aspergillus niger*, or *Saccharomyces cerevisiae*. In some embodiments, a microorganism can be a prokaryote such as *Escherichia coli*, *Rhodobacter sphaeroides*, or *Rhodobacter capsulatus*. It will be appreciated that certain microorganisms can be used to screen and test genes of interest in a high throughput manner, while other microorganisms with desired productivity or growth characteristics can be used for large-scale production of steviol glycosides.

Saccharomyces cerevisiae

Saccharomyces cerevisiae is a widely used chassis organism in synthetic biology, and can be used as the recombinant microorganism platform. 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.

A steviol biosynthesis gene cluster can be expressed in yeast using any of a number of known promoters. Strains that overproduce terpenes are known and can be used to increase the amount of geranylgeranyl diphosphate available for steviol and steviol glycoside production.

Aspergillus spp.

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 the production of food ingredients such as steviol and steviol glycosides. Example 23 describes cloning methodology for production of steviol glycosides in *Aspergillus nidulans*.

Escherichia coli

Escherichia 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.

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

Rhodobacter spp.

Rhodobacter can be used as the recombinant microorganism platform. Similar to *E. coli*, there are libraries of mutants available as well as suitable plasmid vectors, allowing for rational design of various modules to enhance product yield. Isoprenoid pathways have been engineered in membranous bacterial species of *Rhodobacter* for increased production of carotenoid and CoQ10. See, U.S. Patent Publication Nos. 20050003474 and 20040078846. Methods similar to those described above for *E. coli* can be used to make recombinant *Rhodobacter* microorganisms.

Physcomitrella spp.

Physcomitrella mosses, when grown in suspension culture, have characteristics similar to yeast or other fungal cultures. This genera is becoming an important type of cell for production of plant secondary metabolites, which can be difficult to produce in other types of cells. Example 22 describes production of active
5 UGT enzymes in the steviol glycoside pathway in *P. patens*.

B. Plant Cells or Plants

In some embodiments, the nucleic acids and polypeptides described herein are introduced into plants or plant cells to increase overall steviol glycoside production or
10 enrich for the production of specific steviol glycosides in proportion to others. Thus, a host can be a plant or a plant cell that includes at least one recombinant gene described herein. A plant or plant cell can be transformed by having a recombinant gene integrated into its genome, *i.e.*, can be stably transformed. Stably transformed cells typically retain the introduced nucleic acid with each cell division. A plant or
15 plant cell can also be transiently transformed such that the recombinant gene is not integrated into its genome. Transiently transformed cells typically lose all or some portion of the introduced nucleic acid with each cell division such that the introduced nucleic acid cannot be detected in daughter cells after a sufficient number of cell divisions. Both transiently transformed and stably transformed transgenic plants and
20 plant cells can be useful in the methods described herein.

Transgenic plant cells used in methods described herein can constitute part or all of a whole plant. Such plants can be grown in a manner suitable for the species under consideration, either in a growth chamber, a greenhouse, or in a field. Transgenic plants can be bred as desired for a particular purpose, *e.g.*, to introduce a
25 recombinant nucleic acid into other lines, to transfer a recombinant nucleic acid to other species, or for further selection of other desirable traits. Alternatively, transgenic plants can be propagated vegetatively for those species amenable to such techniques. As used herein, a transgenic plant also refers to progeny of an initial transgenic plant provided the progeny inherits the transgene. Seeds produced by a
30 transgenic plant can be grown and then selfed (or outcrossed and selfed) to obtain seeds homozygous for the nucleic acid construct.

Transgenic plants can be grown in suspension culture, or tissue or organ culture. For the purposes of this invention, solid and/or liquid tissue culture techniques can be used. When using solid medium, transgenic plant cells can be placed directly onto the medium or can be placed onto a filter that is then placed in contact with the medium. When using liquid medium, transgenic plant cells can be placed onto a flotation device, *e.g.*, a porous membrane that contacts the liquid medium.

When transiently transformed plant cells are used, a reporter sequence encoding a reporter polypeptide having a reporter activity can be included in the transformation procedure and an assay for reporter activity or expression can be performed at a suitable time after transformation. A suitable time for conducting the assay typically is about 1-21 days after transformation, *e.g.*, about 1-14 days, about 1-7 days, or about 1-3 days. The use of transient assays is particularly convenient for rapid analysis in different species, or to confirm expression of a heterologous polypeptide whose expression has not previously been confirmed in particular recipient cells.

Techniques for introducing nucleic acids into monocotyledonous and dicotyledonous plants are known in the art, and include, without limitation, *Agrobacterium*-mediated transformation, viral vector-mediated transformation, electroporation and particle gun transformation, U.S. Patent Nos 5,538,880; 5,204,253; 6,329,571; and 6,013,863. If a cell or cultured tissue is used as the recipient tissue for transformation, plants can be regenerated from transformed cultures if desired, by techniques known to those skilled in the art.

A population of transgenic plants can be screened and/or selected for those members of the population that have a trait or phenotype conferred by expression of the transgene. For example, a population of progeny of a single transformation event can be screened for those plants having a desired level of expression of a steviol or steviol glycoside biosynthesis polypeptide or nucleic acid. Physical and biochemical methods can be used to identify expression levels. These include Southern analysis or PCR amplification for detection of a polynucleotide; Northern blots, S1 RNase protection, primer-extension, or RT-PCR amplification for detecting RNA transcripts; enzymatic assays for detecting enzyme or ribozyme activity of polypeptides and

polynucleotides; and protein gel electrophoresis, Western blots, immunoprecipitation, and enzyme-linked immunoassays to detect polypeptides. Other techniques such as *in situ* hybridization, enzyme staining, and immunostaining also can be used to detect the presence or expression of polypeptides and/or nucleic acids. Methods for performing all of the referenced techniques are known. As an alternative, a population of plants comprising independent transformation events can be screened for those plants having a desired trait, such as production of a steviol glycoside or modulated biosynthesis of a steviol glycoside. Selection and/or screening can be carried out over one or more generations, and/or in more than one geographic location. In some cases, transgenic plants can be grown and selected under conditions which induce a desired phenotype or are otherwise necessary to produce a desired phenotype in a transgenic plant. In addition, selection and/or screening can be applied during a particular developmental stage in which the phenotype is expected to be exhibited by the plant. Selection and/or screening can be carried out to choose those transgenic plants having a statistically significant difference in a steviol glycoside level relative to a control plant that lacks the transgene.

The nucleic acids, recombinant genes, and constructs described herein can be used to transform a number of monocotyledonous and dicotyledonous plants and plant cell systems. Non-limiting examples of suitable monocots include, for example, cereal crops such as rice, rye, sorghum, millet, wheat, maize, and barley. The plant may be a non-cereal monocot such as asparagus, banana, or onion. The plant also may be a dicot such as stevia (*Stevia rebaudiana*), soybean, cotton, sunflower, pea, geranium, spinach, or tobacco. In some cases, the plant may contain the precursor pathways for phenyl phosphate production such as the mevalonate pathway, typically found in the cytoplasm and mitochondria. The non-mevalonate pathway is more often found in plant plastids [Dubey, et al., 2003 *J. Biosci.* **28** 637-646]. One with skill in the art may target expression of steviol glycoside biosynthesis polypeptides to the appropriate organelle through the use of leader sequences, such that steviol glycoside biosynthesis occurs in the desired location of the plant cell. One with skill in the art will use appropriate promoters to direct synthesis, e.g., to the leaf of a plant, if so desired. Expression may also occur in tissue cultures such as callus culture or hairy root culture, if so desired.

In one embodiment, one or more nucleic acid or polypeptides described herein are introduced into *Stevia* (e.g., *Stevia rebaudiana*) such that overall steviol glycoside biosynthesis is increased or that the the overall steviol glycoside composition is selectively enriched for one or more specific steviol glycosides. For example, one or more recombinant genes can be introduced into *Stevia* such that one or more of the following are expressed: a UGT91D enzyme such as UGT91D2e (e.g., SEQ ID NO:5 or a functional homolog thereof), UGT91D2m (e.g., SEQ ID NO:10); a UGT85C enzyme such as a variant set forth in Table 15, or a UGT76G1 enzyme such as a variant set forth in Example 18. Nucleic acid constructs typically include a suitable promoter (e.g., 35S, e35S, or ssRUBISCO promoters) operably linked to a nucleic acid encoding the UGT polypeptide. Nucleic acids can be introduced into *Stevia* by *Agrobacterium*-mediated transformation; electroporation-mediated gene transfer to protoplasts; or by particle bombardment. See, e.g., Singh, et al., Compendium of Transgenic Crop Plants: Transgenic Sugar, Tuber and Fiber, Edited by Chittaranjan Kole and Timothy C. Hall, Blackwell Publishing Ltd. (2008), pp. 97-115. For particle bombardment of stevia leaf derived callus, the parameters can be as follows: 6 cm distance, 1100 psi He pressure, gold particles, and one bombardment.

Stevia plants can be regenerated by somatic embryogenesis as described by Singh et al., 2008, *supra*. In particular, leaf segments (approximately 1-2cm long) can be removed from 5 to 6-week-old *in vitro* raised plants and incubated (adaxial side down) on MS medium supplemented with B5 vitamins, 30 g sucrose and 3 g Gelrite. 2,4-dichlorophenoxyacetic acid (2,4-D) can be used in combination with 6-benzyl adenine (BA), kinetin (KN), or zeatin. Proembryogenic masses appear after 8 weeks of subculture. Within 2-3 weeks of subcultures, somatic embryos will appear on the surface of cultures. Embryos can be matured in medium containing BA in combination with 2,4-D, α -naphthaleneacetic acid (NAA), or indolbutyric acid (IBA). Mature somatic embryos that germinate and form plantlets can be excised from calli. After plantlets reach 3-4 weeks, the plantlets can be transferred to pots with vermiculite and grown for 6-8 weeks in growth chambers for acclimatization and transferred to greenhouses.

In one embodiment, steviol glycosides are produced in rice. Rice and maize are readily transformable using techniques such as *Agrobacterium*-mediated

transformation. Binary vector systems are commonly utilized for *Agrobacterium* exogenous gene introduction to monocots. See, for example, U.S. Patent Nos. 6,215,051 and 6,329,571. In a binary vector system, one vector contains the T-DNA region, which includes a gene of interest (e.g., a UGT described herein) and the other
5 vector is a disarmed Ti plasmid containing the *vir* region. Co-integrated vectors and mobilizable vectors also can be used. The types and pretreatment of tissues to be transformed, the strain of *Agrobacterium* used, the duration of the inoculation, the prevention of overgrowth and necrosis by the *Agrobacterium*, can be readily adjusted by one of skill in the art. Immature embryo cells of rice can be prepared for
10 transformation with *Agrobacterium* using binary vectors. The culture medium used is supplemented with phenolic compounds. Alternatively, the transformation can be done *in planta* using vacuum infiltration. See, for example, WO 2000037663, WO 2000063400, and WO 2001012828.

IV. Methods of Producing Steviol and Steviol Glycosides

15 Recombinant hosts described herein can be used in methods to produce steviol or steviol glycosides. For example, if the recombinant host is a microorganism, the method can include growing the recombinant microorganism in a culture medium under conditions in which steviol and/or steviol glycoside biosynthesis genes are expressed. The recombinant microorganism may be grown in a fed batch or
20 continuous process. Typically, the recombinant microorganism is grown in a fermentor at a defined temperature(s) for a desired period of time. 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
25 diphosphate, dimethylallyl diphosphate, geranylgeranyl diphosphate, kaurene and kaurenoic acid, can be determined by extracting samples from culture media for analysis according to published methods.

After the recombinant microorganism has been grown in culture for the desired period of time, steviol and/or one or more steviol glycosides can then be
30 recovered from the culture using various techniques known in the art. If the recombinant host is a plant or plant cells, steviol or steviol glycosides can be extracted

from the plant tissue using various techniques known in the art. For example, a crude lysate of the cultured microorganism or plant tissue can be centrifuged to obtain a supernatant. The resulting supernatant can then be applied to a chromatography column, e.g., a C-18 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.

The amount of steviol or steviol glycoside produced can be from about 1 mg/l to about 1,500 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. 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.

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 steviol while a second microorganism comprises steviol glycoside biosynthesis genes. 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 rebaudioside A. 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 fermentor.

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 and have a consistent taste profile. Thus, the recombinant microorganisms, plants, and

plant cells 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.

Microorganisms described herein do not produce the undesired plant byproducts found in Stevia extracts. Thus, steviol glycoside compositions produced by the recombinant microorganisms described herein are distinguishable from compositions derived from Stevia plants.

V. Food Products

The steviol and steviol glycosides obtained by the methods disclosed herein can be used to make food products, dietary supplements and sweetener compositions. For example, substantially pure steviol or steviol glycoside such as rebaudioside A can be included in food products such as ice cream, carbonated beverages, 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 or growing different plants/plant cells, each producing a specific steviol or steviol glycoside, recovering the steviol or steviol glycoside in substantially pure form from each microorganism or plant/plant cells and then combining the compounds to obtain a mixture containing each compound in the desired proportion. The recombinant microorganisms, plants, and plant cells described herein permit more precise and consistent mixtures to be obtained compared to current Stevia products. 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. Patent Publication No. 2007/0128311. In some embodiments, the steviol or steviol glycoside may be

provided with a flavor (e.g., citrus) as a flavor modulator. For example, Rebaudioside C can be used as a sweetness enhancer or sweetness modulator, in particular for carbohydrate based sweeteners, such that the amount of sugar can be reduced in the food product.

5 Compositions produced by a recombinant microorganism, plant, or plant cell described herein can be incorporated into food products. For example, a steviol glycoside composition produced by a recombinant microorganism, plant, or plant cell 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
10 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, plant, or plant cell 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
15 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 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, plant, or plant cell can be incorporated into a sauce
20 (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, plant, or plant cell can be incorporated into a bread such that the food product has a maximum of 160 mg steviol glycoside/kg food on a dry weight basis. A steviol glycoside
25 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
30 jellies) such that the food product has a maximum of 1000 mg steviol glycoside/kg food on a dry weight basis.

For example, such a steviol glycoside composition can have from 90-99% rebaudioside A and an undetectable amount of stevia plant-derived contaminants, and be incorporated into a food product at from 25-1600 mg/kg, c.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.

5 Such a steviol glycoside composition can be a rebaudioside B-enriched composition having greater than 3% rebaudioside B and be incorporated into the food product such that the amount of rebaudioside B in the product is from 25-1600 mg/kg, c.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 rebaudioside B-enriched composition has
10 an undetectable amount of stevia plant-derived contaminants.

 Such a steviol glycoside composition can be a rebaudioside C-enriched composition having greater than 15% rebaudioside C and be incorporated into the food product such that the amount of rebaudioside C in the product is from 20-600 mg/kg, c.g., 100-600 mg/kg, 20-100 mg/kg, 20-95 mg/kg, 20-250 mg/kg, 50-75
15 mg/kg or 50-95 mg/kg on a dry weight basis. Typically, the rebaudioside C-enriched composition has an undetectable amount of stevia plant-derived contaminants.

 Such a steviol glycoside composition can be a rebaudioside D-enriched composition having greater than 3% rebaudioside D and be incorporated into the food product such that the amount of rebaudioside D in the product is from 25-1600 mg/kg,
20 c.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 rebaudioside D-enriched composition has an undetectable amount of stevia plant-derived contaminants.

 Such a steviol glycoside composition can be a rebaudioside E-enriched composition having greater than 3% rebaudioside E and be incorporated into the food
25 product such that the amount of rebaudioside E in the product is from 25-1600 mg/kg, c.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 rebaudioside E-enriched composition has an undetectable amount of stevia plant-derived contaminants.

 Such a steviol glycoside composition can be a rebaudioside F-enriched
30 composition having greater than 4% rebaudioside F and be incorporated into the food product such that the amount of rebaudioside F in the product is from 25-1000 mg/kg, c.g., 100-600 mg/kg, 25-100 mg/kg, 25-95 mg/kg, 50-75 mg/kg or 50-95 mg/kg on a

dry weight basis. Typically, the rebaudioside F-enriched composition has an undetectable amount of stevia plant-derived contaminants.

Such a steviol glycoside composition can be a dulcoside A-enriched composition having greater than 4% dulcoside A and be incorporated into the food product such that the amount of dulcoside A in the product is from 25-1000 mg/kg, e.g., 100-600 mg/kg, 25-100 mg/kg, 25-95 mg/kg, 50-75 mg/kg or 50-95 mg/kg on a dry weight basis. Typically, the dulcoside A-enriched composition has an undetectable amount of stevia plant-derived contaminants.

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 rebaudioside A, rebaudioside C, rebaudioside D, rebaudioside E, rebaudioside F, or dulcoside A can be packaged 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.

VI. Plant Breeding

A. Polymorphisms

Polymorphisms among the nucleic acids described herein (e.g., UGT91D2 nucleic acids) can be used as markers in plant genetic mapping and plant breeding programs in *Stevia*. See, e.g., Yao et al., Genome, 1999, 42:657-661. Thus, the polymorphisms described herein can be used in a method of identifying whether that polymorphism is associated with variation in a trait. The method involves measuring the correlation between variation in the trait in plants of a *Stevia* line or population and the presence of one or more genetic polymorphisms in those plants, thereby identifying whether or not the genetic polymorphisms are associated with variation in the trait. Typically, the trait is the total amount of steviol glycosides present in leaves of the plant, although the trait also can be the amount of a particular steviol glycoside, e.g., rebaudioside A, rebaudioside B, rebaudioside C, rebaudioside D, rebaudioside E, rebaudioside F, or dulcoside A. In some embodiments, the trait is the amount of steviol, or the amount of an isoprenoid precursor. A statistically significant correlation

between the trait and the presence of the polymorphic marker is determined using an appropriate parametric or non-parametric statistic, *e.g.*, Chi-square test, Student's t-test, Mann-Whitney test, or F-test. A statistically significant correlation between, for example, the amount of rebaudioside A in a plant and presence of a polymorphic
5 marker indicates that the marker may be useful in a marker-assisted breeding program for selection of altered rebaudioside A levels.

Polymorphisms may be detected by means known in the art, including without limitation, restriction fragment length polymorphism (RFLP), random amplified polymorphic DNA detection (RAPD), amplified fragment length polymorphism
10 (AFLP), simple sequence repeat (SSR) or microsatellites. Discovery, detection, and genotyping of polymorphisms have been described in the literature. See, *e.g.*, Henry, ed. (2001) Plant Genotyping. The DNA Fingerprinting of Plants Wallingford: CABI Publishing; and Phillips and Vasil, eds. (2001) DNA-based Markers in Plants Dordrecht: Kluwer Academic Publishers. For example, a primer or probe derived
15 from the nucleic acid sequences set forth in SEQ ID NO:6, SEQ ID NO:9, or SEQ ID NO:96, or the complements thereof, can be used to identify one or more individual plants that possess the polymorphic allele that is correlated with a desired steviol glycoside composition. Those plants then can be used in a breeding program to combine the polymorphic allele with a plurality of other alleles at other loci that are
20 correlated with the desired steviol glycoside composition. As will be evident to one of skill, the number and type of markers required can differ, depending on the trait(s) to be selected for and the degree of correlation for each marker. The methods, therefore, involve detecting a plurality of polymorphisms in the genome of the plant in certain embodiments. It will be appreciated that the method may further comprise
25 storing the results of the step of detecting the plurality of polymorphisms on a computer readable medium.

Thus, in some embodiments, a method for identifying *Stevia* plant lines or populations comprises supplying a nucleic acid sample for a *Stevia* plant, providing amplification primers for amplifying a region of a *Stevia* plant corresponding to a
30 UGT gene having 90% or greater sequence identity to a nucleic acid encoding the polypeptides set forth in SEQ ID NOs: 1, 3, 5, or 7, present in the sample, applying the amplification primers to the nucleic acid sample such that amplification of the region

occurs, and identifying plants having a desired trait based on the presence of one or more polymorphisms in the amplified nucleic acid sample that correlate with the trait.

In some embodiments, a method of determining the presence of a polynucleotide in a *Stevia* plant involves contacting at least one probe or primer pair with nucleic acid from the plant. The probe or primer pair is specific for a polynucleotide that encodes a UGT polypeptide having at least 90% sequence identity to SEQ ID NOs: 1, 3, 5, or 7. The presence or absence of the polynucleotide is then determined.

In addition to methods for detecting polymorphisms and determining the genotype of a *Stevia* plant, kits suitable for carrying out the methods are also described, as well as a computer readable medium produced by such methods that contains data generated by the methods. A kit for genotyping a *Stevia* biological sample 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 NOs: 1, 3, 5, or 7. Such kits typically have the primer or probe contained within suitable packaging material.

In some embodiments of the methods and kits described herein, one or more sets of oligonucleotides, each capable of recognizing the presence or absence of a specific and defined genomic position, is used. For polyploid *Stevia* lines or populations, more oligonucleotides are desirable. The lower limit is one oligonucleotide pair and the upper limit is set by the desired resolution capacity of the method and the test kit. Hybridization of the oligonucleotides to DNA from the *Stevia* plant is preferably recorded *in situ* by any conventional labelling system, applying for instance terminal transferase and conventional recordable labels. As an alternative to *in situ* labelling the hybridized sample DNA may be released from the solid support and subsequently hybridized with labelled polynucleotide sequences corresponding to each of the original oligonucleotide sequences attached to the solid support. Hybridization is optionally reversible and the solid support can be returned to its original state for reuse. A labelled dideoxynucleotide can be incorporated at the end of the oligonucleotide provided that the oligonucleotide is hybridized to genomic DNA as template. The nucleotide sequence at the genomic position adjacent to the region matching the oligonucleotide is known and therefore the particular nucleotide which

will be incorporated (A, C, G, T or U) is known. Co-dominant scoring is achieved using paired, *i.e.* two or parallel, *i.e.* three, flanking oligonucleotide sequences. The results obtained are recorded as full, empty, failure or null alleles and can be used to distinguish between heterozygous and/or homozygous genotypes. Optional post-hybridization treatments, including washing and digestion, are provided in order to remove sample DNA not fully hybridized to the solid support-attached oligonucleotide sequences, for example before and after labelling. The presence or absence of hybridization is recorded using a method allowing the recording of the hybridization state, typically on a computer readable medium.

B. Breeding Programs

Stevia is typically an outcrossing species, although self-pollination is occasionally observed. Thus, a *Stevia* plant breeding program typically involves the use of one or more of: recurrent selection mass selection, bulk selection, and intercrossing. These techniques can be used alone or in combination with one or more other techniques in a breeding program. See, Yadav et al., Can. J. Plant Sci. 91: 1-27 (2011). Each identified plant can be crossed to a different plant to produce seed, which is then germinated to form progeny plants. Seed from one or more progeny plants possessing the desired phenotype(s) and desired polymorphism(s) is composited and then randomly mated to form a subsequent progeny generation. The breeding program can repeat these steps for an additional 0 to 5 generations as appropriate in order to achieve the desired stability in the resulting plant population, which retains the polymorphic allele(s). In most breeding programs, analysis for the particular polymorphic allele will be carried out in each generation, although analysis can be carried out in alternate generations if desired. Selfing of progeny plants may be carried out for those stevia lines and populations in which selfing is feasible.

Recurrent selection is a method used in a plant breeding program to improve a population of plants. The method entails individual plants cross pollinating with each other to form progeny. The progeny are grown and the superior progeny selected by any number of selection methods, which include individual plant, half-sib progeny, full-sib progeny and selfed progeny. The selected progeny are self pollinated or cross pollinated with each other to form progeny for another population. This population is

planted and again superior plants are selected to self pollinate or cross pollinate with each other. Recurrent selection is a cyclical process and therefore can be repeated as many times as desired. The objective of recurrent selection is to improve the traits of a population. The improved population can then be used as a source of breeding material to obtain new varieties for commercial or breeding use, including the production of a synthetic cultivar. A synthetic cultivar is the resultant progeny formed by the intercrossing of several selected varieties. The number of parental plant varieties, populations, wild accessions, ecotypes, etc., that are used to generate a synthetic can vary from as little as 10 to as much as 500. Typically, about 100 to 300 varieties, populations, etc., are used as parents for the synthetic variety. Seed from the parental seed production plot of a synthetic variety can be sold to the farmer. Alternatively, seed from the parental seed production plot can subsequently undergo one or two generations of multiplication, depending on the amount of seed produced in the parental plot and the demand for seed.

Mass selection is a useful technique when used in conjunction with molecular marker-assisted selection. In mass selection, seeds from individuals are selected based on phenotype or genotype. These selected seeds are then bulked and used to grow the next generation. Bulk selection requires growing a population of plants in a bulk plot, allowing the plants to self-pollinate, harvesting the seed in bulk and then using a sample of the seed harvested in bulk to plant the next generation. Also, instead of self pollination, directed pollination could be used as part of the breeding program.

Thus, in some embodiments, a method of making a *Stevia* plant line or population involves identifying one or more plants in the line or population in which the presence of a polymorphism at a locus having nucleotide sequence encoding a polypeptide that is at least 90% identical to SEQ ID NOs: 1, 3, 5, or 7 is associated with variation in a trait of interest. The identified plant(s) is then crossed with itself or a different stevia plant to produce seed, and at least one progeny plant grown from the seed is again crossed with itself or a different stevia plant for an additional 0-5 generations to make a line or population that possesses the polymorphism.

In some cases, selection for other useful traits is also carried out, e.g., selection for disease resistance. Selection for such other traits can be carried out before, during

or after identification of individual plants that possess the desired polymorphic allele.

Marker-assisted breeding techniques may be used in addition to, or as an alternative to, other sorts of identification techniques.

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

VI. Examples

Example 1 -- Construction of Kaurene Biosynthesis Pathway Genes

A nucleotide sequence encoding a truncated baker's yeast HMG CoA
10 reductase was cloned into a yeast high copy episomal plasmid vector such that the coding sequence was operably linked to and under the transcriptional control of a promoter which can be repressed by the amino acid methionine. See, U.S. Patent Nos. 5,460,949 and 5,306,862.

Nucleotide sequences encoding the GGPPS enzymes shown in Table 1 were
15 modified for expression in yeast (see SEQ ID NOs:18-25) and cloned into an *E. coli* vector such that the coding sequence was operably linked to and under the transcriptional control of a yeast promoter which can be repressed by the amino acid methionine. The name for each expression cassette-containing plasmid ("entry vector") is also shown in Table 1. The nucleotide sequences from the source
20 organisms from which the polypeptides were originally identified are set forth in SEQ ID NOs: 26-33. Other entry vectors were constructed using GGPPS enzymes expressed by an unmodified nucleotide sequence from *Catharanthus roseus* designated EV270, an unmodified nucleotide sequence from *Aspergillus nidulans* designated C301 and an unmodified nucleotide sequence from *Xanthophyllomyces*
25 *dendrorhous* designated C413.

Table 1. GGPPS Clones

Enzyme Source Organism	gi Number	Accession Number	Plasmid Name	Construct Name	Length (nts)	SEQ ID (DNA)	SEQ ID (protein)
<i>Stevia rebaudiana</i>	90289577	ABD92926	pMUS14	MM-1	1086	18	121
<i>Gibberella fujikuroi</i>	3549881	CAA75568	pMUS15	MM-2	1029	19	122

<i>Mus musculus</i>	47124116	AAH69913	pMUS16	MM-3	903	20	123
<i>Thalassiosira pseudonana</i>	223997332	XP_002288339	pMUS17	MM-4	1020	21	124
<i>Streptomyces clavuligerus</i>	254389342	ZP_05004570	pMUS18	MM-5	1068	22	125
<i>Sulfolobus acidocaldarius</i>	506371	BAA43200	pMUS19	MM-6	993	23	126
<i>Synechococcus</i> sp.	86553638	ABC98596	pMUS20	MM-7	894	24	127
<i>Arabidopsis thaliana</i>	15234534	NP_195399	pMUS21	MM-8	1113	25	128

Nucleotide sequences encoding the CDPS enzymes shown in Table 2 were modified for expression in yeast (see SEQ ID NOs: 34-36) and cloned into yeast entry vectors. The nucleotide sequences from the source organisms from which the polypeptides were originally identified are set forth in SEQ ID NOs: 37-39. Other entry vectors were constructed using CDPS enzymes expressed by an unmodified nucleotide sequence from *Arabidopsis thaliana* designated EV64, an unmodified nucleotide sequence from *Zea mays* designated EV65 and an unmodified nucleotide sequence from *Lycopersicon esculentum* designated EV66.

10

Table 2. CDPS Clones

Enzyme Source Organism	gi Number	Accession Number	Plasmid Name	Construct Name	Length (nts)	SEQ ID: (DNA)	SEQ ID (protein)
<i>Stevia rebaudiana</i>	2642661	AAB87091	pMUS22	MM-9	2364	34	129
<i>Streptomyces clavuligerus</i>	197705855	EDY51667	pMUS23	MM-10	1584	35	130
<i>Bradyrhizobium japonicum</i>	529968	AAC28895.1	pMUS24	MM-11	1551	36	131

Nucleotide sequences encoding the KS enzymes shown in Table 3 were modified for expression in yeast (see SEQ ID NOs: 40 -43) and cloned into yeast entry vectors. The nucleotide sequences from the source organisms from which the polypeptides were originally identified are set forth in SEQ ID NOs:44-47. Other entry vectors were constructed using KS enzymes expressed by an unmodified nucleotide sequence from *Arabidopsis thaliana* designated EV70, an unmodified

nucleotide sequence from *Cucurbita maxima* designated EV71 and an unmodified nucleotide sequence from *Cucumis sativus* designated EV72.

Table 3. KS Clones

Enzyme Source Organism	gi Number	Accession Number	Plasmid Name	Construct Name	Length (nts)	SEQ ID (DNA)	SEQ ID (protein)
<i>Stevia rebaudiana</i>	4959241	AAD34295	pMUS25	MM-12	2355	40	132
<i>Stevia rebaudiana</i>	4959239	AAD34294	pMUS26	MM-13	2355	41	133
<i>Zea mays</i>	162458963	NP_001105097	pMUS27	MM-14	1773	42	134
<i>Populus trichocarpa</i>	224098838	XP_002311286	pMUS28	MM-15	2232	43	135

5

Nucleotide sequences encoding the CDPS-KS fusion enzymes shown in Table 4 were modified for expression in yeast (see SEQ ID NOs: 48 and 49) and cloned into yeast entry vectors. The nucleotide sequences from the source organisms from which the polypeptides were originally identified are set forth in SEQ ID NOs: 50 and 51.

10

Table 4. CDPS-KS Clones

Enzyme Source Organism	gi Number	Accession Number	Plasmid Name	Construct Name	Length (nts)	SEQ ID (DNA)	SEQ ID (protein)
<i>Phomopsis amygdali</i>	186704306	BAG30962	pMUS29	MM-16	2952	48	136
<i>Physcomitrella patens</i>	146325986	BAF61135	pMUS30	MM-17	2646	49	137

Nucleotide sequences encoding the KO enzymes shown in Table 5 were modified for expression in yeast (see SEQ ID NOs: 52-55) and cloned into yeast entry vectors. The nucleotide sequences from the source organisms from which the polypeptides were originally identified are set forth in SEQ ID NOs: 56-59.

15

Table 5. KO Clones

Enzyme Source Organism	gi Number	Accession Number	Plasmid Name	Construct Name	Length (nts)	SEQ ID (DNA)	SEQ ID (protein)
<i>Stevia</i>	76446107	ABA42921	pMUS31	MM-18	1542	52	138

<i>rebaudiana</i>							
<i>Arabidopsis thaliana</i>	3342249	AAC39505	pMUS32	MM-19	1530	53	139
<i>Gibberella fujikuroi</i>	4127832	CAA76703	pMUS33	MM-20	1578	54	140
<i>Trametes versicolor</i>	14278967	BAB59027	pMUS34	MM-21	1500	55	141

- Nucleotide sequences encoding the KAH enzymes shown in Table 6 were modified for expression in yeast (see SEQ ID NOs: 60-64) and cloned into yeast entry vectors. The nucleotide sequences from the source organisms from which the polypeptides were originally identified are set forth in SEQ ID NOs: 65-69.

Table 6. KAH Clones

Enzyme Source Organism	gi Number	Accession Number	Plasmid Name	Construct Name	Length (nts)	SEQ ID (DNA)	SEQ ID (protein)
<i>Stevia rebaudiana</i>		--*	pMUS35	MM-22	1578	60	142
<i>Stevia rebaudiana</i>	189418962	ACD93722	pMUS36	MM-23	1431	61	143
<i>Arabidopsis thaliana</i>	15238644	NP_197872	pMUS37	MM-24	1578	62	144
<i>Vitis vinifera</i>	225458454	XP_002282091	pMUS38	MM-25	1590	63	145
<i>Medicago truncatula</i>	84514135	ABC59076	pMUS39	MM-26	1440	64	146

* = Sequence is shown in U.S. Patent Publication No. 2008-0064063.

- Nucleotide sequences encoding the CPR enzymes shown in Table 7 were modified for expression in yeast (see SEQ ID NOs: 70-72) and cloned into yeast entry vectors. The nucleotide sequences from the source organisms from which the polypeptides were originally identified are set forth in SEQ ID NOs:73-75.

Table 7. CPR Clones

Enzyme Source Organism	gi Number	Accession Number	Plasmid Name	Construct Name	Length (nts)	SEQ ID (DNA)	SEQ ID (protein)
<i>Stevia rebaudiana</i>	93211213	ABB88839	pMUS40	MM-27	2133	70	147
<i>Arabidopsis thaliana</i>	15233853	NP_194183	pMUS41	MM-28	2079	71	148
<i>Giberella fujikuroi</i>	32562989	CAE09055	pMUS42	MM-29	2142	72	149

Example 2 – Construction of Steviol Glycoside Pathway Genes

Integration vectors containing nucleotide sequences encoding the UGT85C2 and UGT74G1 enzymes listed in Table 8 were transformed into yeast. Transformants were obtained that contained UGT85C2, or UGT85C2 and UGT74G1, integrated into the genome.

Table 8. UGT Clones

Source Organism	UGT No.	gi Number	Accession Number	Type	Plasmid Name	Length (nucleotides)	SEQ ID
<i>Stevia rebaudiana</i>	UGT85C2	37993660	AY345978.1	Integration vector	pMUS11	1446	4
<i>Stevia rebaudiana</i>	UGT74G1	37993668	AY345982	Integration vector	pMUS12	1383	2
<i>Stevia rebaudiana</i>	UGT76G1	37993652	AY345974	Integration vector	pMUS13	1377	8
<i>Ipomoea purpurea</i>	IP3GGT	62857205	AB192315.1	High copy vector	pMUS10	1380	77
<i>Bellis perennis</i>	UGT94B1 R25S mutant	56550538 (wild type)	AB190262.1 (wild type)	High copy vector	pEF1156	1317 (wild type)	79
<i>Arabidopsis thaliana</i>	UGT79B3	28951020	BT005370.1	High copy vector	pEF1153	1362	151

Nucleotide sequences encoding the IP3GGT and UGT94B1 R25S enzymes were modified for expression in yeast (see SEQ ID NOs: 77 and 79) and cloned into yeast entry vectors. Amino acid sequences for IP3GGT and UGT94B1 R25S are set forth in SEQ ID NOs: 76 and 78, respectively. The high copy episomal vector containing a modified IP3GGT nucleotide sequence was designated pEF1155. The high copy episomal vector containing a modified UGT94B1 R25S nucleotide sequence was designated pEF1156.

Example 3 – Construction of Yeast Strains

A yeast strain designated EFSC301 was modified by replacing the endogenous *ERG9* promoter with the copper inducible *CUP1* promoter. Strain EFSC301 is a derivative of EUROSCARF collection yeast strain BY4742. See, the world wide web at uni-frankfurt.de/fb15/mikro/euroscarf/data/by.html. In standard yeast growth medium, the *ERG9* gene is transcribed at very low levels, since the concentration of copper in such medium is low. The decrease in ergosterol production in this strain results in increased amounts of isoprene units available for steviol biosynthesis. The

yeast strain was also modified by genomically integrating the *Stevia* UGT85C2 and UGT74G1 genes, each under the transcriptional control of the strong constitutive *GPD1* promoter. See Table 8. The strain has one copy of each of the *Stevia* UGT85C2 and UGT74G1 genes integrated in the MUS1241 strain genome.

5

Example 4 -- Analysis of Steviol Glycoside Pathway Gene Expression in Yeast

To examine steviol glycoside biosynthesis in yeast, the expression cassettes of the 36 entry vectors of Tables 1-7 and Example 1 were randomly concatenated in ligation reactions to create artificial yeast chromosomes ("cYACs"). The process is shown schematically in FIG 5.

Two different sets of ligations were carried out. Ligation set A included all genes listed in Tables 1-7, except that no bi-functional CDPS-KS genes (Table 4) were included. Ligation set B included all genes listed in Tables 1-7 except that no mono-functional CDPS and KS genes (Tables 2-3) were included.

From 30 to 200 µg of DNA was prepared from each of the cassette-containing entry vectors. The gene expression cassettes were released from each vector by digestion with the restriction enzyme *AscI*. The cassettes were then randomly concatenated into eYACs by ligation with T4 ligase in a 3 hour reaction. The success of the concatenation reaction was assessed by the viscosity of the reaction mixture, since concatenated DNA is highly viscous. DNA fragments ("arms") containing a centromere, two telomeres and the *LEU2* and *TRP1 selection* markers were added to the end of the concatenated expression cassettes, thereby creating functional eYACs.

The eYACs were transformed into spheroplasts of the competent yeast strain MUS1243 by zymolyase digestion of the yeast cell wall, followed by treatment with a CaCl_2 /PEG buffer, making the spheroplasts permeable to large molecules such as eYACs.

After transformation, the yeast spheroplasts were embedded in a noble agar based solid growth medium, in which regeneration of the cell wall can take place. Colonies appeared from 4-8 days after inoculation. The regeneration medium lacked the amino acids leucine and tryptophan, thus selecting for the presence of double-armed eYACs in the yeast cells.

About 3,000 transformants were obtained for each set. Each transformant was re-streaked and tested for yeast strain markers and the genetic presence of both arms of the cYAC, i.e., the *LEU2* and *TRP1* markers. More than 97% of the transformants had the correct genotype. Each transformant was given a CEY designation number.

5 Initially, 24 CEYs from each set were grown for 24 hours in 2 ml of Synthetic Complete medium (SC), without methionine, so as to induce gene expression from the eYACs. After 24 hours, the supernatant from each culture was collected and subjected to LC-MS (Liquid Chromatography-coupled Mass Spectrometry (Triple Quadropole)) analysis for the presence of rubusoside. Since the *Stevia* UGT74G1 and
10 UGT85C2 genes are co-expressed in each CEY transformant, the expected end product when steviol is produced is rubusoside (steviol-(13- β -D-glucopyranosyloxy)- β -D-glucopyranosyl ester).

None of the CEYs from set B produced detectable levels of rubusoside, whereas 7 of the CEYs from set A did. Strain CEY19 was the top producer. CEY19
15 produced a compound with a mass of 665.2, which could correspond to a sodium adduct of rubusoside. A compound with a mass of 643.2 also was seen, and probably corresponds to protonated rubusoside. MS-MS-based molecular fractionation of the 665.2 mass compound resulted in a break down mass of 503.2, which corresponds to steviol monoside as a sodium adduct. Since the mass, the fractionation pattern, the
20 HPLC spectrum, and the retention standard of this compound corresponded exactly to that of a rubusoside standard produced *in vitro* by the glucosylation of steviol using *Stevia* enzymes 85C2 and 74G1, the compound produced by CEY was determined to be rubusoside.

Additional screening for rubusoside production

25 An additional 95 clones from set A and 95 clones from set B were grown in 96 deep-well trays in 1 ml SC medium without methionine. Supernatants from each of these cultures were combined in pools of two clones, analyzed by LC-MS, and the MS signal/noise ratio determined. The MS s/n ratio is an approximate measure of the relative rubusoside content. When a pool of 2 CEYs was found to produce
30 rubusoside, each clone in that pool was analyzed separately. The results showed that no set B CEYs produced rubusoside, while at least 28 CEYs from set A produced detectable levels of rubusoside.

Identification of genes present in rubusoside producing CEY clones

To correlate the gene content of cYACs to rubusoside production, a PCR protocol was developed in which similar sized fragments (0.5 kb) of all the possible cYAC-borne genes could be amplified. Internal primers of 20-25 nt were placed so that a similar annealing temperature could be used to amplify all genes. Genomic DNA, which includes cYAC DNA, was prepared from 4 CEYs with no rubusoside production, 4 with low rubusoside production and 6 with high to very high rubusoside production. Using equimolar amounts of these 14 DNA preparations, analytical PCR was performed for all 37 genes for these 14 CEYs, as well as positive and negative controls. All genes were amplified except one, apparently due to primer failure.

The genes present in the six high rubusoside-producing CEY strains are shown in Table 9. The genes present in the eight low or no rubusoside-producing CEY strains are shown in Table 10.

Table 9. Genes Present in High Rubusoside-Producing CEY Strains

Gene	HIGH production			VERY high production		
	CEY50	CEY176	CEY19	CEY173	CEY191	CEY213
tHMG1	+	+	+	+	-	+
MM-1	-	+	+	+	+	-
MM-2	-	+	+	+	+	-
MM-3	+	+	+	+	+	+
MM-4	+	+	+	-	+	+
MM-5	+	+	+	+	+	+
MM-6	+	+	+	+	+	+
MM-7	-	+	-	+	+	-
MM-8	+	+	+	+	-	+
EV270	+	+	-	+	+	+
C301	+	+	+	+	+	+
C413	+	+	-	+	+	+
MM-9	+	+	+	+	+	+
MM-10	+	-	-	+	+	+
MM-11	+	+	-	+	+	+
EV64	+	+	+	+	+	+
EV65	-	-	+	+	+	+
EV66	+	+	+	+	+	+
MM-12	+	-	-	+	+	+
MM-13	+	+	+	+	+	+
MM-14	+	+	+	+	+	+
MM-15	-	-	-	-	+	-
EV70	-	+	+	+	-	-
EV71	Primers failed					
EV72	+	+	+	+	+	+
MM-18	+	+	+	+	+	-
MM-19	+	-	+	-	+	+

MM-20	+	+	+	+	+	+
MM-21	-	-	+	+	-	+
MM-22	+	+	+	+	+	+
MM-23	+	-	+	+	-	+
MM-24	+	+	+	+	+	+
MM-25	+	+	+	+	+	+
MM-26	+	+	+	+	+	+
MM-27	+	+	+	+	+	+
MM-28	-	-	-	-	-	-
MM-29	+	+	+	+	+	+

Table 10. Genes Present in CEY Strains Producing Low or No Rubusoside

Gene	NO rubusoside production				LOW production			
	CEY162	CEY169	CEY171	CEY188	CEY75	CEY147	CEY214	CEY87
tHMG1	-	-	-	-	-	-	+	+
MM-1	+	+	+	+	-	+	-	-
MM-2	+	-	+	+	+	+	+	+
MM-3	+	+	+	+	+	+	+	+
MM-4	-	-	+	-	-	+	-	+
MM-5	+	+	+	+	+	+	+	+
MM-6	+	+	+	-	+	+	+	+
MM-7	+	-	+	+	+	+	+	+
MM-8	+	+	+	+	+	+	+	+
EV270	+	+	+	+	+	+	+	+
C301	+	+	+	+	+	+	+	+
C413	+	+	+	+	+	+	+	+
MM-9	+	+	+	+	-	+	+	+
MM-10	+	+	+	+	-	+	+	+
MM-11	+	+	+	+	+	+	+	-
EV64	+	+	+	+	-	+	+	+
EV65	+	-	-	-	+	-	+	-
EV66	+	+	+	+	+	+	+	+
MM-12	+	+	+	+	+	+	+	+
MM-13	+	+	+	+	+	+	+	+
MM-14	+	+	+	+	+	+	+	+
MM-15	+	-	+	-	+	+	-	+
EV70	+	+	+	+	+	+	+	+
EV71	Primers failed							
EV72	+	+	+	+	+	+	+	+
MM-18	+	+	+	+	+	+	+	+
MM-19	+	+	+	+	+	+	+	+
MM-20	+	+	+	+	+	+	+	+
MM-21	-	+	-	-	-	+	-	+
MM-22	+	+	+	+	+	-	+	+
MM-23	+	-	+	-	+	+	-	+
MM-24	+	-	+	+	+	+	+	+
MM-25	+	-	+	+	+	+	+	+
MM-26	+	+	+	+	+	-	+	+
MM-27	+	+	+	+	+	+	+	+
MM-28	-	-	+	-	-	-	-	+

MM-29	+	+	+	+	-	+	+	+
-------	---	---	---	---	---	---	---	---

Example 5 – Modification of Yeast Culture Conditions

Experiments were carried out with strain CEY213 in order to determine culture conditions conducive to maximum rubusoside production. The starting material was a glycerol freezer stock (-80°C) of CEY213. Frozen cells originally came from an agar plate containing SC yeast medium without tryptophan, leucine and histidine (SC-TLH), and containing 2 mM methionine. Five ml of liquid SC-TLH medium containing 2 mM methionine was inoculated with a loop-full of freeze stock CEY213 yeast cells. eYAC expression in CEY213 is repressed under these conditions. The cells were grown overnight at 30° C with slow shaking (170 rpm) and were designated as “pre-cultures.”

The CEY 213 pre-cultures were used to inoculate 25-50 ml of SC media without methionine, in which the parameters indicated below were varied. Rubusoside production under each of the growth conditions was measured by centrifuging 500 µl of each culture medium, transferring 250 µl of the supernatant to a new tube, adding 250 µl methanol, shaking thoroughly and centrifuging for 10 minutes at maximum speed. An aliquot of the supernatant was analyzed for rubusoside production by LC-MS.

Copper Levels

CEY213 precultures were grown in SC medium to which 50 µM bathocuproinedisulfonic acid was added. Bathocuproinedisulfonic acid chelates copper in the growth medium. The *ERG9* gene in CEY213 has been modified so that expression is controlled by the *CUP1* promoter. A decrease in copper levels in the medium will further decrease *ERG9* activity and thereby increase the amount of isoprene units available for steviol biosynthesis.

Chelation of copper ions in the growth medium had a detrimental effect on growth of the yeast culture and rubusoside production was decreased proportionally. These results suggested that even without copper chelation, strain CEY213 is at its minimum rate of ergosterol biosynthesis, and no more isoprene units can be diverted from ergosterol biosynthesis towards steviol glycoside production.

Glucose

Doubling the available glucose from 2 to 4% had a marginal effect on rubusoside production, about a 5-10% increase in rubusoside production.

Limiting available nitrogen

5 CEY213 pre-cultures were grown under conditions of limited available nitrogen. Limiting nitrogen during growth of yeast in culture is known to increase production of ergosterol. When the concentration of NH_4SO_4 was decreased from 4g/l to 2, 1 or 0.4 g/l, the growth rate of CEY213 decreased in proportion to the amount of nitrogen. Rubusoside production decreased proportionally with the
10 decrease in growth.

Aeration of cultures

CEY213 was grown in Ehrlenmeyer flasks with or without baffles. The results indicated that there was at best a marginal effect of increased aeration via the use of baffles. If anything, the lack of aeration via the lack of baffles increased
15 production.

Optical Density at Initiation, Fermentation Time and Growth temperature

Cultures were initiated at two different optical densities, $\text{OD}_{600}=0.1$ or $\text{OD}_{600}=1.0$ of pre-cultured CEY213. Fermentation was then carried out for 24, 48, 72 or 144 hours at a temperature of 20, 25 or 30° C.

20 As shown in FIG 6, the density of the batch culture at fermentation start, the culture temperature and the length of time in fermentation, in combination, had a significant effect on the amount of rubusoside produced by CEY213. Thus, 144 hours growth of a culture with a starting density of $\text{OD}_{600}=1.0$, at 30° C, resulted in the production of no less than 8.5 mgs/liter of rubusoside.

25

Example 6 -- Large Scale Production of Rubusoside

A series of fermentation experiments with CEY213 were performed using 3 kinds of yeast medium (rich medium and two types of synthetic medium), varying inoculation density, and changing timing of eYAC gene cassette expression.

Batch Fermentation Conditions

30 Batch fermentation was carried out by centrifuging a CEY213 pre-culture, discarding the supernatant and re-suspending the cells in 6 liters of SC-TLH medium

containing 100µM methionine and 4% glucose. The OD₆₀₀ was adjusted to 1.0 in a 100 ml Ehrlenmeyer flask without baffles and the cells were allowed to grow for 144 hours at 30° C with slow shaking.

Recovery of Rubusoside

5 After fermentation, the culture was centrifuged and the supernatant was mixed with an equal volume of methanol, shaken thoroughly, and centrifuged to remove precipitated material. The resulting supernatant was purified by flash C18-silica column chromatography with methanol as the eluent, followed by preparative HPLC to obtain one major compound, with one additional minor compound detected.

10 The purified compound was analyzed by ¹H and ¹³C NMR, and the data are shown in FIG. 7. The compound was confirmed to be rubusoside based on comparison to ¹H and ¹³C NMR literature values for rubusoside. Quantitative analysis indicated that CEY213 fermentation produced 12.8 mgs/liter of rubusoside.

Example 7 -- IP3GGT Activity

1. Enzymatic Activity of *Ipomoea purpurea* 3GGT glycosyltransferase *in vitro*

 The enzymatic activity of *Ipomoea purpurea* 3GGT glycosyltransferase (IP3GGT) using steviol as a substrate was determined *in vitro*. Genes for *Stevia rebaudiana* UGT85C2 and IP3GGT glycosyltransferase were each expressed in *E.*
20 *coli* and each enzyme was purified.

 The enzymatic reaction was performed in two steps. First, 0.5 mM steviol (9.55 mgs total) was incubated with ca. 0.5 µg UGT85C2 enzyme for 16 hours at 30 °C in a reaction buffer (containing 1 mM UDP-glucose, 100 mM Tris-HCl (pH 8.0), 5 mM MgCl₂, 1 mM KCl, 0.1 U/ul calf intestine phosphatase). Then ca. 0.5 µg
25 IP3GGT enzyme was added and the reaction mixture incubated for an additional 20 hours at 30°C.

 Analysis of the reaction products indicated about 100% conversion of steviol to steviol-13-O-monoside, 25% of which was further glycosylated into stevio-13-O-1,2-bioside. The theoretical steviol-13-O-1,2-bioside yield was about 4.8 mg. The
30 reaction mixture was then subjected to preparative HPLC, which yielded 2.5 mg steviol-13-O-1,2-bioside (52% purification yield). Using LC-MS, the mass of the purified compound had a different retention time than rubusoside and steviol-13-O-

1,3-bioside. The purified compound was subjected to ^1H NMR, heteronuclear single quantum coherence (HSQC)-NMR and heteronuclear multiple bond correlation (HMBC)-NMR analysis, which confirmed that the compound was steviol-13-O-1,2-bioside. .

5

2. *In vivo* expression of IP3GGT in steviol- or steviol monoside-fed yeast

To determine whether the IP3GGT was active in yeast, the 2 μ high copy (episomal) plasmid, pMUS10, containing an unmodified IP3GGT coding sequence operably linked to a strong *GPD1* promoter was transformed into the yeast strain
10 MUS1245. MUS1245 contains a genomically integrated UGT85C2 expression cassette. The resulting yeast strain was grown in SC medium without histidine to select for the continued presence of the IP3GGT expression plasmid, at a starting density of $\text{OD}_{600}=0.2$. Steviol or steviol monoside was added to the medium at 3 mM. After growth for 72 hours at 30°C, culture supernatants were assayed for the presence
15 of steviol and steviol glucosides by HPLC.

LC-MS analysis indicated that no 1,2-glucosylated steviol-13-O-glucoside was detected after feeding with steviol, although steviol-13-O-monoside could be detected. In contrast, low but detectable amounts of the steviol 1,2-bioside were produced by MUS1245 carrying pMUS10 after feeding with steviol-13-O-monoside.
20 These results show that the native *Ipomoea purpurea* 3GGT coding sequence is expressed in yeast at levels sufficient to obtain detectable *in vivo* conversion of steviol monoside to steviol 1,2-bioside.

Example 8 -- Modification of Yeast Strains

25

EXG1 and EXG2

S. cerevisiae may contain enzymes that degrade the 1,2 or 1,3 sugar bonds in steviol 1,2- and steviol 1,3-biosides. To test this possibility, yeast strain CEY213 was grown for 3 days at 30° C on media containing 0.1 mM of each of the two biosides. LC-MS analysis of the culture showed the level of 1,2-bioside to be stable, whereas
30 the 1,3-bond in the 1,3-bioside appeared to completely hydrolyse within the limits of detection of the assay.

Twenty-five *S. cerevisiae* mutants, each disrupted in one known or putative glycoside hydrolase gene, were examined for their ability to degrade steviol biosides. A culture of each yeast mutant was grown as described above on media containing steviol 1,3-bioside and analyzed by LC-MS. The yeast strain carrying a mutation in the *EXG1* (exo-1,3- β -glucanase) gene was found to have lost most of the 1,3-bioside hydrolysing activity. The nucleotide sequence of the yeast EXG1 gene is reported in Vazquez de Aldana et al. Gene 97:173-182 (1991). The yeast strain carrying a mutation in the EXG2 gene (another exo-1,3- β -glucanase) showed a small decrease in hydrolysing activity. Correa, et al., Current Genetics 22:283-288 (1992).

A double mutant yeast strain (*exg1 exg2*) was made. When the double mutant strain was grown on media containing steviol 1,3-bioside, no hydrolysis of the bioside was detected.

Example 9 – Increased Titer of Steviol Biosynthesis

Individual clones of enzymes from each of the different enzyme classes tested in Example 4 (and Table 11) were examined using cYAC technology to identify particular clones that exhibited the greatest production of steviol from isopentenyl pyrophosphate and farnesyl pyrophosphate. The GGPPS, KO and KAH enzymes have been tested on cYACs, individually or in the case of GGPPS enzymes individually or in pools of two (e.g., *Synechococcus* sp. + *S. acidocaldarius* GGPPS or *Aspergillus nidulans* GGPPS alone), in a *S. cerevisiae* strain expressing all remaining enzymatic steps in the steviol pathway. The results indicated that the *Synechococcus* spp. GGPPS clone MM-7 (encoded by SEQ ID NO:24) was the most efficient. GGPPS clones from *Aspergillus nidulans* and *Sulfolobus acidocaldarius* also were quite active. The results also indicated that among the KO and KAH clones, the *Stevia* KO clone MM-18 (encoded by SEQ ID NO:52) and the *A. thaliana* KAH clone MM-24 (encoded by SEQ ID NO:62) resulted in the greatest steviol production.

Table 11.

Source Organism	Enzyme	gi Number	Accession Number	Coding Sequence	Coding Sequence Length (nucleotides)
<i>Stevia rebaudiana</i>	GGPPS-1	158104429	ABD92926	MM-1	1086
<i>Gibberella fujikuroi</i>	GGPPS-2	3549881	CAA75568	MM-2	1029

<i>Mus musculus</i>	GGPPS-3		BC069913.1	MM-3	903
<i>Thalassiosira pseudonana</i>	GGPPS-4	223997332	XP_002288339	MM-4	1020
<i>Sulfolobus acidocaldarius</i>	GGPPS-6	506371	BAA43200	MM-6	993
<i>Synechococcus</i> sp.	GGPPS-7	86553638	ABC98596	MM-7	894
<i>Cantharanthus roseus</i>	GGPPS-9	1063275	X92893	EV270	1074
<i>Aspergillus nidulans</i>	GGPPS-10	29468175	AF479566	C301	1191
<i>Xanthophyllomyces dendrorhous</i>	GGPPS11	63145970	DQ016502	C413	1131
<i>Stevia rebaudiana</i>	CDPS-1	2642661	AAB87091	MM-9	2364
<i>Streptomyces clavuligerus</i>	CDPS-2	197705855	EDY51667	MM-10	1584
<i>Bradyrhizobium japonicum</i>	CDPS-3	529968	AAC28895.1	MM-11	1551
<i>Arabidopsis thaliana</i>	CDPS-4	18412041	NM_116512	EV-64	2409
<i>Zea mays</i>	CDPS-5	50082774	AY562490	EV-65	2484
<i>Lycopersicon esculentum</i>	CDPS-6	6009477	AB015675	EV-66	2403
<i>Stevia rebaudiana</i>	KS-1	4959241	AAD34295	MM-12	2355
<i>Stevia rebaudiana</i>	KS-2	4959239	AAD34294	MM-13	2355
<i>Zea mays</i>	KS-3	162458963	NP_001105097	MM-14	1773
<i>Populus trichocarpa</i>	KS-4	224098838	XP_002311286	MM-15	2232
<i>Arabidopsis thaliana</i>	KS-5	3056724	AF034774	EV-70	2358
<i>Cucurbita maxima</i>	KS-6	1431869	U43904	EV-71	2370
<i>Cucumis sativus</i>	KS-7	21326756	AB045310	EV-72	2358
<i>Stevia rebaudiana</i>	KO-1	76446107	ABA42921	MM-18	1542
<i>Arabidopsis thaliana</i>	KO-2	3342249	AAC39505	MM-19	1530
<i>Gibberella fujikuroi</i>	KO-3	74676162	O94142	MM-20	1578
<i>Trametes versicolor</i>	KO-4	14278966	AB057426	MM-21	1500
<i>Stevia rebaudiana</i>	KAH-1	*		MM-22	1578
<i>Stevia rebaudiana</i>	KAH-2	189418962	ACD93722	MM-23	1431
<i>Arabidopsis thaliana</i>	KAH-3	15238644	NM_122399	MM-24	1578

<i>Vitis vinifera</i>	KAH4	225458453	XM_002282055	MM-25	1590
<i>Medicago trunculata</i>	KAH5	84514134	DQ335781	MM-26	1440
<i>Stevia rebaudiana</i>	CPR-1	189098311	DQ269454.4	MM-27	2133
<i>Arabidopsis thaliana</i>	CPR-2	145343899	NM_118585	MM-28	2079
<i>Gibberella fujikuroi</i>	CPR-3	32562988	AJ576025.1	MM-29	2142

* U.S. Patent Publication No. 20080064063

S. cerevisiae strain CEY213, described in Example 4, was transformed with high copy plasmids carrying one of the CDPS or KS genes shown in Table 11, operably linked to the strong GPD1 promoter. Preliminary experiments indicated that overexpression of the *Stevia rebaudiana* CDPS (CDPS-1, encoded by SEQ ID NO:34) in CEY213 gave an increase in rubusoside production relative to CEY213 that lacked the high copy CDPS-1 overexpressing plasmid. The experiments also indicated that the *Stevia rebaudiana* KO (KO-1, encoded by SEQ ID NO:52) was the most active KO of the two tested.

To construct a yeast strain with consistently high levels of steviol glycoside production, expression cassettes containing the GGPPS-10 clone, the KO-1 clone (SEQ ID NO:52) and the KAH-3 clone (SEQ ID NO:62) were stably integrated into the genome of the *S. cerevisiae* strain CEN.PK 111-61A. Expression of these cassettes was driven by the constitutive *GPD1* and *TPH1* promoters. In addition, expression cassettes containing KS-1 (SEQ ID NO:40), CDPS-1 (SEQ ID NO:34) and UGT74G1 (SEQ ID NO:2) were stably integrated into the genome. The resulting yeast strain, EFSC1751, however, did not produce any steviol-19-O-monoside when grown at laboratory scale under the conditions described in Example 6.

To determine the basis for the lack of steviol glycoside production in EFSC1751, CDPS-3, CDPS -4, CDPS -5 and CPR-1 genes, alone or in combination, were expressed in strain EFSC1751. CPR-1 is from *Stevia rebaudiana* and its sequence can be found at Genbank Accession DQ269454.4. The results showed that CPR-1, when expressed with either CDPS-3, CDPS-4 or CDPS-5, resulted in production of steviol-19-O-monoside in EFSC1751. None of these genes alone in the same strain resulted in any production. These results indicate that the genomically integrated copy of CDPS-1, *Stevia* enzyme, is non-functional in this yeast construct,

whereas the *Bradyrhizobium*, *Arabidopsis* or *Zea* CDPS clones were functional in this construct. In addition, the plant-derived KAH and/or KO genes integrated into the chromosome for this construct appear to require an exogenous CPR for activity. The CPR from *Gibberella fujikuroi* (MM-29) also appears to be able to work with plant-
5 derived KAH and/or KO polypeptides.

The two leading GGPPS candidates, GGPPS-6 (encoded by SEQ ID NO:23) and GGPPS-7 (encoded by SEQ ID NO:24), were further expressed individually in a *S. cerevisiae* strain that has a functional steviol glycoside pathway (including UGT74G1) but no GGPPS genes. Transformants then were analyzed for the
10 production of 19-SMG by LC-MS analysis of culture samples that had been boiled in 50% DMSO for 5 minutes and centrifuged at 16000 relative centrifugal force (RCF) for 5 minutes. It was found that many transformants containing the GGPPS-6-expressing plasmid did not produce 19-SMG.

Very few transformants were obtained containing GGPPS-7, indicating that
15 GGPPS-7 (*Synechococcus sp.*) may be the more active of the two enzymes, and that the activity could be high enough to confer toxicity. For example, a dramatic increase in GGPP production could result in a drain on a downstream pathway such as ergosterol production. To test this hypothesis, a *UPC2-1* gene was co-expressed with GGPPS-7, and ergosterol feeding of the cells was attempted to see if this would
20 rescue growth of cells. However, cell growth was not rescued.

Cell toxicity also may be due to an accumulation of GGPP or a metabolite of GGPP. To test this hypothesis, CDPS-5 was further overexpressed in the GGPPS-7-expressing yeast strain to see if the toxicity could be alleviated by increased GGPP usage. CDPS5 over-expression did appear to rescue growth to some extent since
25 transformants with a plasmid overexpressing this enzyme along with the GGPPS-7 gave rise to a few colonies. The number of transformants was still low. Over-expression of CDPS-5 in a similar strain but with GGPPS-10 instead of GGPPS-7 resulted in a doubling of steviol glycoside production, and these results together could suggest that CDPS is a limiting bottleneck in the introduced steviol glycoside
30 biosynthesis pathway.

In summary, based upon production of 19-SMG or rubusoside in test tube cell cultures at 30°C with yeast medium + 2% glucose, for 24-72 hours, the following

conclusions were made with the eYAC constructs: KS-1 (*Stevia rebaudiana*, encoded by SEQ ID NO:40), KO-1 (*S. rebaudiana*, encoded by SEQ ID NO:52) and KAH-1 (*S. rebaudiana*) or KAH-3 (*Arabidopsis thaliana*, encoded by SEQ ID NO:62) appear to be the best combinations for the steviol pathway. GGPPS-7 (*Synechococcus* sp.) appears to show the highest amount of activity for this step, but if downstream bottlenecks occur overexpression also could lead to toxicity and overall lower levels of steviol glycosides. All combinations of CDPS and CPR gene analogs were tested and it was found that all 3 CPRs in Table 11 were active, and that combinations of CPR-1 (*S. rebaudiana*, encoded by SEQ ID NO:70) or CPR-3 (*Gibberella fujikuroi*, encoded by SEQ ID NO:72) with either CDPS-5 (*Zea mays*) or CDPS-4 (*A. thaliana*) were particularly useful. CDPS-5 appears to be the optimal CDPS in the pathway. Combinations can be further tested in a reporter strain with reduced flux to sterol pathways.

To investigate the potential for even higher activity of the CDPS from *Zea mays* (CDPS-5), this gene was expressed from a 2micron multicopy plasmid using the GPD promoter, with and without a plastid signal peptide, to determine if activity is higher in the cytoplasm when targeting sequences are removed. The nucleotide sequence and amino acid sequence of the CDPS-5 from *Zea mays* and containing the chloroplast signal peptide are set forth in SEQ ID NOs:80 and 81, respectively. The chloroplast signal peptide is encoded by nucleotides 1-150 of SEQ ID NO:80, and corresponds to amino acids 1 to 50 of SEQ ID NO:81. The plasmid was transformed into the stable rubusoside producer strain (EFSC1859) that has GGPPS-10, CDPS-5, KS-1, KO-1, KAH-3, CPR-1 and UGT74G1 (SEQ ID NO:2) integrated into the genome and expressed from the strong constitutive GPD and TPI promoters. Furthermore, in strain EFSC1859, expression of squalene synthase, which is encoded by ERG9, was downregulated by displacement of the endogenous promoter with the CUP1 inducible promoter. In addition to these genes, strain EFSC1859 also expresses UGT85C2 (SEQ ID NO:3) from a 2micron multicopy vector using a GPD1 promoter. Rubusoside and 19-SMG production were measured by LC-MS to estimate the production level. The removal of the plastid leader sequence did not appear to increase steviol glycoside production as compared to the wild-type sequence.

However, this work demonstrates that the leader sequences can be removed without causing a loss of steviol pathway function.

Similarly, plasmids were constructed for CPR-3, KAH-3 and KO-1 without membrane anchoring sequences (i.e., nucleotides 4-63 of SEQ ID NO:72; nucleotides 4-87 of SEQ ID NO:62; and nucleotides 1-117 of SEQ ID NO:52) and were transformed into strain EFSC1859 with the UGT85C2 integrated on the chromosome rather than on a plasmid. It is expected that these enzymes will be functional without the anchoring sequence.

Example 10 – Identification of Steviol-1,3-O-Monoglucoside 1,2-Glucosyltransferase Sequences

Stevia EST Analysis

A tBLASTN search of a *Stevia* (*Stevia rebaudiana*) leaf EST (Expressed Sequence Tags) database (Brandle et al., Plant Mol. Biol. 50:613-622, 2002) was carried out using complete *Ipomoea* (*Ipomoea purpurea*) UGT79 type UGT (IP3GGT), *Bellis* (*Bellis perennis*) UGT94B1, *Stevia* UGT79A2, *Stevia* UGT76G1 and *Stevia* UGT91D1 amino acid sequences as queries, thus representing UGTs from all Family 1 glycosyltransferase sub-families known to primarily contain diglycosyltransferases. Partial sequences for 9 previously undescribed UGT genes were identified. One of the partial sequences was from the UGT 79 sub-family (“79-EV1”), one from the UGT 76 sub-family (“76-EV1”) and two from the UGT 91 sub-family (“91-EV-1” and “91-EV2”), as well as members of the UGT 71, 72, 78, 84 and 88 sub-families. Seven of the partial sequences were isolated using *Stevia* cDNA or cDNA libraries as the PCR template for isolation. In addition, two *Stevia* members of the UGT 76 sub-family were isolated, GenBank accession ACT33422.1 which is a member of the 76G1 sub-family (Mohankumar), and GenBank accession ACM47734.1 which is a member of the 76G2 (Yang) sub-family.

Pyrosequencing

Additional UGT clones were identified and isolated by performing pyrosequencing with *Stevia* cDNA as follows. *Stevia* mRNA was prepared from *Stevia* leaves, using the Ambion® Micro Poly Purist™ mRNA preparation kit. As a quality control, reverse transcribed mRNA was tested for the presence of the *Stevia*

Rebaudioside A pathway UGT genes 85C2, 74G1 and 76G1, by employing analytical PCR with oligonucleotide primers identical to 21 nucleotides at the 5'- and 3'-termini of each sequence. The amplified full length mRNA was then used for pyrosequencing and contig assembly (MOgene, St. Louis, MO USA). About 3.4 million reads of an average length of 393 nucleotides were performed, and the resulting raw sequences used to obtain 25907 sequence contigs. A database was constructed, containing publicly available amino acid sequences of a total of ca. 1,500 UGTs. About 150 of the sequenced UGTs were fully annotated UGTs from a wide variety of sub-families. The remaining sequenced UGTs were partially annotated homologs of these. A BLASTX search was performed (CLC Genomics, Muelheim, Germany), using the 25907 Stevia EST contigs as query, to the fabricated UGT database (Genetic code = 1, Low complexity = Yes, Expect value = 10.0, Word size = 3, No of processors = 2, Matrix = BLOSUM62, Gap cost (open) = 11, Gap cost (extension)= 1). The results suggested that sequences for more than 90 previously unknown UGTs from Stevia were present in the pyrosequencing database.

No additional members of the UGT 79 sub-family or the UGT 94 sub-family were identified in the pyrosequencing database. However, the analysis showed new members of the UGT 76 and 91 sub-families. For a few of the genes, full length sequence data was immediately available from the pyrosequencing EST data. A previously constructed *Stevia* plasmid cDNA library was used to obtain full-length sequences for those members for which partial sequence data was obtained. An oligonucleotide primer identical to each specific, partial UGT sequence was combined with an oligonucleotide primer identical to the library plasmid vector sequence. These primers were employed in PCR to obtain the full length product, which was subsequently sequenced. Based on the full length sequence, a second PCR was performed using a proof-reading PCR polymerase enzyme for amplification of the full length UGT gene from a *Stevia* cDNA library as the template for the reaction. Using this strategy, five members of the UGT 76 sub-family, six members of the UGT 91 sub-family, as well as ten members of other UGT sub-families were isolated.

Each of the 7 UGTs identified from the *Stevia* EST database, the 2 publicly available *Stevia* UGT 76 sequences, and the 21 UGTs identified from pyrosequencing was cloned into the *E. coli* expression vectors pET30A+ or pETDuet (making use of

the HIS-tag for purification purposes) and expressed in the autolysis-prone *E. coli* strains XjA and XjB. For a large number of these UGTs, expression of the UGT protein resulted in the formation of inclusion bodies. In order to overcome formation of those inclusion bodies, some of these UGTs were expressed in the low temperature expression strain "Arctic Express" (Agilent Technologies). For those which failed to express in this system, coupled *in vitro* transcription-translation of PCR products (TNT[®]T7 Quick for PCR DNA kit, Promega) was attempted, allowing successful expression of the remaining UGTs. Efficiency of the reaction was ensured by labeling with ³⁵S-methionine, separation on SDS-PAGE and phosphorimaging detection of a protein band of the expected size for the UGT protein in question.

UGT polypeptides from each clone, expressed as described above, were tested for 1,2-glycosylation activity, using steviol-13-O-monoglucoside as substrate. *In vitro* transcribed/translated protein, corresponding to approximately one fifth of the total protein formed in a 25 µL reaction, was used in an *in vitro* reaction, using 0.5 mM steviol-13-O-monoglucoside (SMG) as substrate, in a reaction buffer (containing 1 mM UDP-glucose, 100 mM Tris-HCl (pH 8.0), 5 mM MgCl₂, 1 mM KCl, 0.1 U/µl calf intestine phosphatase). The reaction mixture was incubated at 30 °C for 20 hours. The reaction mixture was then analyzed by LC-MS analysis for the presence of Steviol-1,2-bioside. LC-MS analyses were performed using an Agilent 1100 Series HPLC system (Agilent Technologies) fitted with a Phenomenex[®] Synergy Hydro-RP column (250 x 3 mm, 3 µm particles, 80 Å pore size) and hyphenated to a TSQ Quantum (ThermoFisher Scientific) triple quadrupole mass spectrometer with electrospray ionization. Elution was carried out using a mobile phase (30 °C) containing MeCN (0.01% Formic acid) and H₂O (0.01% Formic acid) by applying a gradient composed of 0.6→0.4 ml/min, 5% MeCN for 4 min; 0.4 ml/min, 5→40% MeCN for 2 min; 0.4 ml/min, 40→55% MeCN for 11 min; 0.4→1.0 ml/min, 55→100% MeCN for 3 min. Steviol biosides were detected using SIM (Single Ion Monitoring) on Mw 665.2 [M+Na⁺]. None of the 30 UGT enzymes tested exhibited detectable steviol-13-O-monoglucoside glycosylation activity.

The nucleotide sequences of the six UGT91 members identified by pyrosequencing were compared to the sequence of *Stevia* UGT91D1 in Genbank Accession No. AY345980. It appeared that the GenBank sequence encoded 12

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 UGT91D2c. The amino acid sequence of UGT91D2c 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 UGT91D2c Sequences

Sequence variants of UGT91D2c

As evidenced in FIG. 19B, a small number of amino acid modifications exist between the active (91D2c) 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 UGT91D2c (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

0.1 mM IPTG, and further incubation at 20°C. For the *in vitro* assay, cells were induced overnight at 20°C, lysed by a freeze/thaw cycle, and the crude cell extract used for an enzymatic reaction in which the substrates were 0.5 mM steviol-13-*O*-glucoside and 0.5 mM rubusoside.

5 The results are shown in Table 12 for the steviol monoglucoside (SMG) and Rubusoside (Rub) substrates. A “+” indicates that diglycosylation activity was detected, a “-” indicates activity was not detected, and “NA” indicates the assay was not performed. The noted mutations are based on the numbering of the 91D2c sequence (SEQ ID NO:5).

10 As some of the genes have a tendency to express in inclusion bodies in *E. coli*, the coding sequences that did not show activity in the *E. coli* experiments also were produced by coupled *in vitro* transcription-translation of PCR products (TNT[®]T7 Quick for PCR DNA kit, Promega) as above in Example 10. Briefly, 2 µL of DNA from the PCR amplification of the five single mutants and the wild type enzyme were
15 incubated for 90 minutes at 30°C with the kit master mix and 1 µL L-[³⁵S]-Methionine, in a total of 25 µL reaction. For each sample, a volume of 2 µL final reaction was run on a SDS-PAGE gel. All six proteins showed similar levels of soluble recombinant protein as judged by visual observation of the SDS-PAGE gel. The results for the *in vitro*-translated proteins are shown on the right side of Table 12.
20 The percentages in this table indicate the approximate amount of conversion of substrate to product based on relative peak areas of substrate and product.

Table 12

Mutation	<i>E. coli</i> protein SMG	<i>E. coli</i> protein Rub	<i>in vitro</i> protein SMG	<i>in vitro</i> protein Rub
Y30→F	+	+	NA	NA
P93→Q	+	+	NA	NA
S99→V	+	+	NA	NA
Y122→F	+	+	NA	NA
H→I40Y	+	+	NA	NA

S142→C	+	+	NA	NA
T144→I	-	-	5.9%	0.05%
A148→T	+	+	NA	NA
M152→L	-	-	25.1%	0.85%
G153→A	+	+	NA	NA
A156→S	+	+	NA	NA
L195→M	+	+	NA	NA
V196→E	+	+	NA	NA
K199→E	+	+	NA	NA
L211→M	+	+	NA	NA
L213→F	-	-	29.4%	1.59%
S221→F	+	+	NA	NA
V286→A	+	+	NA	NA
S364→P	-	-	4.1%	0.4%
G384→C	-	-	14.1%	1.28%
K427→N	+	+	NA	NA
E438→A	+	+	NA	NA

The approximate amount of diglycosylation activity as compared to UGT91D2e (SEQ ID NO:5) was found to be: 6.1% for T144S, 26.2% for M152L, 30.7% for L213F, 4.3% for S364P, and 14.7% for G384C using 13-SMG as substrate.

5 For rubusoside, the approximate amount of diglycosylation activity as compared to UGT91D2e was 1.4%, 23.4%, 43.7%, 10.9% and 35.2% for T144S, M152L, L213F, S364P, and G384C, respectively.

These results indicate that 5 of the 22 amino acid mutations were noticeably deleterious for activity when done in isolation. It is also possible that combinations of
10 the other 17 mutations also could result in inactivity or loss of activity.

By aligning the 91D2e sequences and the variants described above with proteins termed At72B1, Mt85H2, VvGT1 and Mt71G1 (Osmani et al (2009) Phytochemistry 70, 325-347), and analyzing predicted tertiary structures (alpha helices, beta-sheets, and coil regions), regions can be identified where mutations are
15 likely to result in loss of diglycosylation activity. The first three mutations that are

deleterious are found in the N-terminal domain, in regions that are thought to be loops. The N-terminal domain (amino acid residues 1-240), in particular the predicted loop regions of the N-terminal domain (amino acids 20-26, 39-43, 88-95, 121-124, 142-158, 185-198, and 203-214), are thought to be primarily responsible for binding of the glucose acceptor molecule substrate. The fourth mutation that appears to be deleterious for activity is found in the C-terminal domain, in a region that is believed to be the C5 loop (corresponding to amino acids 381-386). This loop is also thought to be important for glucose acceptor substrate specificity. Nineteen of the twenty-two mutations that separate the inactive versus the active rubusoside diglycosylase enzymes are located within five amino acids of the predicted acceptor substrate binding regions of 91D2c. Therefore it is likely that the published 91D1 enzymes catalyze a glycosyl transferase reaction between UDP-glucose and an alternative acceptor substrate.

Example 12 -- Production of Rebaudioside A in Yeast

Production of Rebaudioside A in Steviol-Fed Yeast

The yeast strain EFSC1580, which contains a genomically integrated UGT74G1 expression cassette, was transformed with three different 2 μ high copy (episomal) plasmids for co-expression of *Stevia* UGTs 91D2c (SEQ ID NO:5), 85C2 (SEQ ID NO:3), and 76G1 (SEQ ID NO:7). The three plasmids, designated pMUS44, pMUS7 and pMUS9, contain coding sequences for UGT91D2c, UGT85C2 and UGT76G1, respectively, operably linked to the strong *GPD1* promoter. The resulting yeast strain was grown in SC medium without uracil, histidine, and leucine to select for the continued presence of the pMUS44, pMUS7 and pMUS9 expression plasmids. Steviol was added to the medium to a final concentration of 250 μ M, and the strain was cultured at 30°C. At 18 hours and 72 hours of culture, aliquots of the supernatants and cell pellets were analyzed for the presence of Rebaudioside A by LC-MS. LC-MS analyses were performed using an Agilent 1100 Series HPLC system (Agilent Technologies, Wilmington, DE, USA) fitted with a Phenomenex® Synergy Hydro-RP column (250 x 3 mm, 3 μ m particles, 80 Å pore size) and hyphenated to a TSQ Quantum (ThermoFisher Scientific) triple quadrupole mass spectrometer with electrospray ionization. Elution was carried out using a mobile

phase (30 °C) containing MeCN (0.01% Formic acid) and H₂O (0.01% Formic acid) by applying a gradient composed of 0.6→0.4 ml/min, 5% MeCN for 4 min; 0.4 ml/min, 5→40% MeCN for 2 min; 0.4 ml/min, 40→55% MeCN for 11 min; 0.4→1.0 ml/min, 55→100% MeCN for 3 min. Steviol biosides were detected using SIM
5 (Single Ion Monitoring).

LC-MS results showed that detectable amounts of Rebaudioside A were found in the supernatant at 18 and 72 hours of culture when strain EFSC1580 containing pMUS44, pMUS7 and pMUS9 was grown in the presence of steviol. The product co-eluted with a Rebaudioside A standard and the expected mass was confirmed as the
10 $[M+Na]^+ = 989$. By comparing the absorbance of the product to the absorbance of a 10 μM Rebaudioside A standard, the accumulation in the supernatant of the cell culture was estimated to be more than 6mg/L at 18 hours, and more than 15 mg/L at 72 hours.

15 Production of Rebaudioside A and Rebaudioside D in Glucose-Fed Yeast

Yeast strain CEY213, described in Example 4, contains steviol biosynthetic pathway genes expressed from cYACs as well as genomically integrated UGT74G1 and UGT85C2 expression cassettes. Strain CEY213 produces rubusoside, as described in Example 6.

20 Strain CEY213 was transformed with a 2 μ high copy (episomal) dual expression plasmid, pMUS47, for simultaneous expression of UGT91D2e (SEQ ID NO:5) and UGT76G1 (SEQ ID NO:7). The pMUS47 plasmid contains two expression cassettes, one having the coding sequence of UGT91D2e and the other having the coding sequence of UGT76G1. Both coding sequences are operably
25 linked to the strong constitutive *GPD1* promoter. The resulting yeast strain was pre-cultured overnight at 30°C in SC medium without histidine, leucine and tryptophan in order to maintain selection for the presence of cYACs, without uracil in order to maintain selection for the presence pMUS47, and finally with methionine (2mM) in order to suppress promoters present on the cYACs. The next day, the cells were
30 washed and transferred to an identical medium, but without methionine, for induction of the cYAC promoters. Samples were collected after 24 hours and 99 hours of

incubation, and supernatants and cell pellets analyzed for the presence of Rebaudioside A and Rebaudioside D, using LC-MS as described above.

The results showed that detectable amounts of Rebaudioside A were found in the supernatants at both 24 and 99 hours. The product co-eluted with a Rebaudioside A standard and the expected mass was confirmed as the $[M+Na]^+ = 989$. By comparing the absorbance of the product to a 10 μ M Rebaudioside A standard, the accumulation of Rebaudioside A in the supernatant was estimated to be more than 3mg/L at 24 hours and more than 6 mg/L at 99 hours. See FIG. 9. The results also indicated that small amounts of stevioside and rubusoside were present in the yeast cell pellet and that detectable amounts of stevioside and rubusoside were present in the culture supernatant. See FIG. 9.

The results also showed that small but detectable amounts of Rebaudioside D were produced, suggesting that UGT91D2e is capable of conjugating an additional glucose to the 19-O glucose of either stevioside producing Rebaudioside E or directly to the 19-O glucose of Rebaudioside A. These results also suggest that UGT76G1 may be capable of accepting Rebaudioside E as a substrate to produce Rebaudioside D. See FIG. 2C.

Example 13 -- Production of Rebaudioside A with codon optimized sequences for

UGT sequences

Optimal coding sequences for UGT 91d2c, 74G1, 76G1, and 85C2 were designed and synthesized for yeast expression using two methodologies, supplied by GeneArt (Regensburg, Germany) (SEQ ID NOs: 6, 2, 8, and 4, respectively) or DNA 2.0 (Menlo Park, CA) (SEQ ID NOs: 84, 83, 85, and 82, respectively). The amino acid sequences of UGT 91d2c, 74G1, 76G1, and 85C2 (SEQ ID NOs: 5, 1, 7, and 3, respectively) were not changed.

High copy number plasmids containing expression cassettes with all four optimized UGTs were constructed and expressed, and their activity compared to expression products of similar constructs containing wild-type sequences. The plasmids were transformed into the universal Watchmaker strain, EFSC301 (described in Example 3). UGTs were inserted in high copy (2 μ) vectors and expressed from a strong constitutive promoter (*GPD1*) (vectors P423-GPD, P424-

GPD, P425-GPD, and P426-GPD). After overnight growth and re-inoculation in fresh media at an OD₆₀₀ of 0.25, the culture medium (SC -leu-trp-ura-his) was supplemented with 25 μ M steviol (final concentration), and production of Rubusoside (Rub), 19-SMG (19SMG) and RebA (RebA) was measured in the media after 24h.

5 The experiment was repeated, in part due to the fact that 19-SMG was undetectable in one of the first samples.

The results from the two separate studies, shown in Table 13 below, indicate that all eight of the codon-optimized UGTs were active. However, enzyme expression for at least one of the codon-optimized UGTs in each strategy was reduced
10 by the new codon optimization algorithm used to make the constructs. It appears that in the GeneArt modified constructs (SEQ ID NOs: 6, 2, 8, and 4), a bottleneck was potentially created between rubusoside and RebA. It is expected that individual enzyme activity assays and expression analyses of these coding sequences expressed in the yeast strains will allow for the optimal combination of UGT genes in the
15 pathway.

Table 13

	RebA (μ M)	19SMG (μ M)	Rub (μ M)
Wild-type	3.2	17.2	4.9
	1.7	14.0	3.2
DNA2.0	4.4	12.4	4.6
	1.7	10.8	3.1
GeneArt	1.2	nd	4.6
	0.8	11.1	4.5

nd = below detection limit

Example 14 - Production of Rebaudioside A using UGTs with sequence tags

20 Fusions of small peptides or protein binding domains with the UGT proteins 85C2, 91D2c, 74G1, and 76G1 can promote interactions between the UGTs (channeling) or aid in targeting/anchoring the UGTs to specific components of the yeast cells.

To assess if scaffolding of the UGTs in the RebA pathway could result in
25 active pathway enzymes, the DNA 2.0 codon-optimized UGTs 85C2 and 74G1 were fused in-frame to a string of 4 high-affinity, short (also known as PMI) peptides that

resemble the p53 protein motif. The p53 protein motif interacts with the MDM2 protein in humans (see Li et al., *J Mol Biol.* 2010, 398(2):200-13). DNA 2.0 codon-optimized UGTs 85C2, 91D2e, 74G1 and 76G1 (SEQ ID NOs: 82, 84, 83, and 85, respectively) were fused in-frame to the first 158 amino acids of the human protein MDM2 (gene accession number ABT17086). A small GS-rich linker region also was fused just prior to the N-terminal methionine of the UGTs. Unfused, the affinity of PMI/MDM2 binding is in the low nM range representing a high-affinity binding. Yeast cells transformed with the above constructs are expected to produce a UGT scaffold around the 4 X PMI (P53-like) peptide repeat fused N-terminally to the 85C2 protein (designated 85C2_P53) scaffold.

The laboratory yeast strain BY4741, deleted for *TRP1*, was transformed with expression plasmids p423-426 GPD (Mumberg et al, *Gene*, 156 (1995), 119-122) expressing *Stevia rebaudiana* UGTs 74G1, 76G1 and 91D2e with N-terminal, in-frame fusions of the first 158 amino acids of human MDM2 protein, and expressing *Stevia rebaudiana* UGT85C2 with an N-terminal in-frame fusion of 4 repeats of the synthetic PMI peptide (4 X TSFAEYWNLLSP, SEQ ID NO:86). See SEQ ID NOs: 88, 90, 92, and 94 for the amino acid sequences of the 85C2, 74G1, 91D2e, and 76G1 fusion proteins, respectively; see SEQ ID NOs: 89, 92, 93, and 95 for the nucleotide sequences encoding the fusion proteins. This yeast strain and a control strain (expressing the four UGT's without any fusions) were grown overnight in synthetic yeast medium selecting for the presence of plasmids and then transferred the next day to a 96 deep-well tray containing synthetic yeast medium to a cell density giving an OD₆₀₀ of 1. A final concentration of 100µM steviol was added. After 72 hours, samples were taken and analysed by LC-MS, as described in Example 12. As indicated in FIGs. 10A and 10B, the UGTs are active in yeast when expressed with the various fusion tags.

Example 15 –UGT91D2e Activity

Additional sub-family 91 UGTs were cloned using cDNA/library preparations made from 3 *Stevia* sources of different genetic backgrounds. Oligonucleotide primers identical to UGT91D1/91D2e were used for PCR amplification of the cDNA preparations, and the resulting PCR products of correct size were cloned into

appropriate plasmid vectors. Numerous clones from each experiment were sequenced, and the sequencing results showed that UGT91D nucleic acids with slight variations in sequence could be amplified. The twenty UGT91D variants with the greatest differences in sequence relative to UGT91D2e were expressed by *in vitro* transcription-translation followed by enzymatic testing for steviol-13-O-monoglucoside-1,2-glucosylating activity. One of the variants showed weak 1,2-bioside glucosylation activity, while the reminder showed no detectable glucosylation activity. It therefore appears that UGT91D2 polypeptides are the primary steviol-13-O-monoglucoside-1,2-glucosylating enzymes in *Stevia*.

Enzymatic activity of UGT91D2e

UGT91D2e (SEQ ID NO:5), made by coupled *in vitro* transcription-translation, was tested for the ability to xylosylate and rhamnosylate steviol-13-O-monoglucoside in an *in vitro* enzyme assay, using UDP-xylose or UDP-rhamnose as the sugar donors rather than UDP-glucose.

The xylosylation assay was performed as follows: 3 mM UDP-glucuronic acid was mixed with ca. 1 µg *Arabidopsis thaliana*-encoded UDP-glucuronic acid decarboxylase UXS3 (produced in *E. coli* and then purified), 100 mM Tris-HCl (pH 8.0), 1 mM DTT, 6 µg BSA, 1 mM MgCl₂, and 1% calf intestine phosphatase. The reaction mixture was incubated for 30 minutes at 30 °C, in order for UDP-glucuronic acid to be turned into UDP-xylose. Then 1.5 mM steviol-13-O-monoglucoside substrate and ca. 0.5 µg UGT91D2e enzyme made as described in Example 9 was added to the mixture, which was allowed to incubate at 30 °C for an additional 20 hours.

The rhamnosylation assay was performed in the following way: 3 mM UDP-glucose was mixed with 0.6 µg of each of the N-terminal and C-terminal parts of *Arabidopsis thaliana*-encoded RHM2 rhamnose synthetase (produced in *E. coli* and then purified), 100 mM Tris-HCl (pH 8.0), 1 mM DTT, 1.5 mM NADPH, 1.5 mM NAD⁺, 6 µg BSA, 1 mM MgCl₂, and 1% calf intestine phosphatase. The reaction mixture was incubated for 30 minutes at 30 °C, in order for UDP-glucose to be turned into UDP-rhamnose. Then 1.5 mM steviol-13-O-monoglucoside substrate and ca. 0.5 µg UGT91D2e enzyme was added to the mixture, which was allowed to incubate at 30 °C for an additional 20 hours.

The results indicated that UGT91D2e was capable of carrying out xylosylation of the steviol-13-*O*-monoglucoside substrate at about one half to one third the rate observed with UDP-glucose, forming 1,2-xylosylated steviol-13-*O*-monoside, which is a precursor to Rebaudioside F. UGT91D2e was capable of carrying out
 5 rhamnosylation of the steviol-13-*O*-monoglucoside substrate at about the same rate as the rate observed with UDP-glucose, forming 1,2-rhamnosylated steviol-13-*O*-monoside, which is a precursor for Rebaudioside C (Dulcoside B). These results indicate that synthesis of appropriate precursor molecules and expression of appropriate UGTs *in vivo* should result in the production of Rebaudioside F and C *in*
 10 *vivo*. See FIGs. 2B and 2D.

UGT91D2c also was tested for its ability to 1,2-glucosylate substrates other than steviol-13-*O*-monoglucoside *in vitro*, i.e., rubusoside, steviol-1,3-bioside and 1,3-stevioside. The results indicated that UGT 91D2c was not active when a 1,3-bound glucose was present (e.g., steviol 1,3-bioside and 1,3-stevioside), while UGT
 15 91D2c was active regardless of primary glucosylation at the 19-*O* position. These results suggest that steviol 1,3-bioside and 1,3-stevioside are likely not present in the *in vivo* *Stevia* pathway for rebA formation. See FIG. 2A and FIG. 3.

Example 16 –UGT91D homologs

20 Different ecotypes of *S. rebaudiana* are genetically diverse. Investigation of 96 clones of 91Ds from different *Stevia* RNA accessions revealed many amino acid changes between six investigated ecotypes (e.g., at nucleotide 74 (resulting in an amino acid change of G to D), 89 (Y to F), 131 (V to A), 137 (F to S), 278 (P to Q), 295 (S to V or P), 331 (E to Q), 365 (Y to F), 395 (A to V), 418 (H to Y), 425 (S to
 25 G), 431 (T to I), 442 (A to T), 454 (M to L), 458 (G to A), 466 (A to S), 485 (G to D), 583 (L to M), 587 (V to E), 595 (K to E), 614 (D to G), 616 (G to R), 631 (L to M), 637 (L to F), 662 (S to F), 664 (K to E), 671 (Y to C), 857 (V to A), 867 (S to R), 919 (F to L), 989 (V to A), 1000 (R to C), 1090 (S to P), 1150 (G to C), 1232 (L to S), 1281 (K to N), 1313 (E to A), 1354 (Q to R), and 1369 (V to I)), as numbered with
 30 respect to the nucleotide sequence of 91D2e set forth in SEQ ID NO:9. Some additional variation from these polymorphisms was noted, which is likely due to sequencing or PCR errors, particularly if the polymorphisms were found only once.

Twenty coding regions were chosen for further analysis. See Table 14 for descriptions of clones that were isolated. The numbering of the amino acids in Table 14 is based on the amino acid sequence of UGT91D2e set forth in SEQ ID NO:5.

TABLE 14

Clone	Mutations as compared to UGT91D2e (SEQ ID NO:5)
1	+1 frameshift between residues 119-145 in the nucleotide sequence, G165V, I367V, L388P
2	27 bp deletion starting at nucleotide 728, K214R
3	D205G, V286A, Y443C
4	L28P, Y30F, P93Q, S99V, E111Q, I118V, Y122F, H140Y, S142C, T144I, A148T, M152L, G153A, A156S, G162D, L195M, V196E, K199E, L211M, L213F, S221F, L411S, V425A
5	G206R, Y207C, W343R
6	Q13R, F46S, S99P, D395G
7	Y30F, S364P, G384C, K427N, E438A
8	Y94C, A132V, Y224C, G384C, K427N, E438A, Q455R
9	K222E, T341M, G384C
10	Y94C, A132V, Y224C, K313N, R334C, G384C
11	Y30F, K222E, V286A, G384C, K427N, E438A
12	Y30F, P93Q, S99V, Y122F, H140Y, S142C, T144I, T145N, A148T, M152L, G153A, A156S, G162D, L195M, V196E, K199E, L211M, L213F, S221F, V286A S289R, R334C, G384C, K427N, E438A
13	V44A, I136V, G374D, V457I, N463S
14	I60S, K97R, Q103R, F181S, L411S
15	V244A, F307L
16	H140Y, S142C, T144I, A148T, M152L, G153A, A156S
17	L195M, V196E, K199E, L211M, L213F, S221F, V286A, R334C, G384C, K427N, E438A
18	V169A, R334C, G384C, K427N, E438A
19	G25D, Y30F, P93Q, S99V, Y122F, H140Y, S142C, T144I, A148T, M152L, G153A, A156S, G162D, L195M, V196E, K199E, L211M, L213F,

	S221F, V286A, G384C
20	I64T, V323A, V330A, G384C, K427N, E438A

All of the clones in Table 14 were tested for activity using 13-SMG as a substrate. Clone 5 had weak 1,2-glycosylating activity whereas the remaining nineteen did not appear to have activity under the conditions tested. The sequence of clone 5 is set forth in SEQ ID NO:95 and has the following mutations with respect to wild-type UGT92D2e (SEQ ID NO:5): G206R, Y207C, and W343R.

Example 17 –UGT85C homologs

The genetic diversity of UGT85Cs from six different *S. rebaudiana* ecotypes was examined to identify homologs that have the same or enhanced activity in pathways for steviol glycoside production. PCR primers were designed that were specific for UGT85C genes, and PCR reactions were carried out on cDNA (some were done on cDNA libraries, some were done on cDNA preparations). The resulting PCR products were cloned and 96 clones were sequenced. Amino acid polymorphisms were mapped and 16 UGT 85C clones were chosen with varying common polymorphism representation. See Table 15. Additional modifications were also noted for some clones, but could be due to PCR errors or were not common polymorphisms. Polymorphisms are described with respect to the nucleotide and amino acid numbering of the wild-type *S. rebaudiana* UGT85C nucleotide sequence set forth in Accession No. AY345978.1 (see Table 8).

TABLE 15

Clone	Allele	38	44	179	194	212	260	659	728	809	866	1001	1007	1116	1181	1190	1253	1319	1322
	Amino Acid #	13	15	60	65	71	87	220	243	270	289	334	336	389	394	397	418	440	441
1	V-F	F-L	H-D							T-M	Q-H						E-V		
2	V-F		H-D							T-M									
3			H-D														E-V		
4			H-D				I-F												
6			H-D													P-S			
7			H-D																
8			H-D							T-M									
13					A-S					T-M				A-V	I-V				
16					A-S	E-Q		K-T	R- W	T-R									
17					A-S	E-Q		K-T	R- W	T-R									H-N
19					A-S	E-Q		K-T						A-V	I-V				
20					A-S	E-Q		K-T		T-M	Q-H								
21			F-L		A-S														
22					A-S		I-F										G-D		
23					A-S					T-A									
24					A-S					T-M	Q-H								
26					A-S					T-M									
27					A-S								I-T						
28					A-S														

The clones were expressed through coupled *in vitro* transcription-translation of PCR products (TNT[®]T7 Quick for PCR DNA kit, Promega) and assayed for glycosylation activity on the substrates steviol and steviol-19-*O*-glucoside (0.5 mM), as described in previous examples. The UGT85Cs produced from clones 1, 4, 16, 17, 19, 20, 21, 26, 29, 30, 31, 37, and 39 were soluble and were able to convert 19-SMG to rubusoside in a 90 min assay. The UGT85C produced from clone 27 was considered insoluble. Although UGT85Cs produced from clones 2 and 33 were considered insoluble, trace amounts of rubusoside were produced despite the protein band not being visible. These experiments were independently performed three times. The experiments showed that the following amino acid mutations did not result in a loss of activity: V13F, F15L, H60D, A65S, E71Q, I87F, K220T, R243W, T270M, T270R, Q289H, L334S, A389V, I394V, P397S, E418V, G440D, and H441N. Additional mutations that were seen in active clones include K9E in clone 37, K10R in clone 26, Q21H in clone 2, M27V in clone 30, L91P in clone 4, Y298C in clone 31, K350T in clone 37, H368R in clone 1, G420R in clone 19, L431P in clone 4, R444G in clone 16, and M471T in clone 30.

The only common polymorphisms that were not tested were T270A and I336T, which are both fairly conservative substitutions. Clone 17 had the most changes incorporated as compared to UGT85C, 6/480 amino acids. The 17-20 amino acids that appear to be changeable represent approximately a 4% difference at the amino acid level.

Generally, there is low genetic diversity among the 85Cs and it is likely that all of the 85C homologs with the common polymorphisms set forth in Table 15 will be active.

Example 18 –UGT76G homologs

The genetic diversity of UGT76Gs from six different *S. rebaudiana* ecotypes was examined to identify homologs that have the same or enhanced activity in pathways for steviol glycoside production. PCR primers were designed that were specific for UGT76G, and PCR reactions were carried out on preparations of cDNA (cDNA libraries or cDNA preparations). The resulting PCR fragments were cloned and 96 clones were sequenced. Common amino acid polymorphisms were mapped and sixteen UGT76G

clones chosen, with varying polymorphism representation, including (amino acid numbering): R10S, I16L, F22V, M29I, K52S, V74K/E, P80S, L85A, V87S/G, L91P, I92F, I93F, H96Y, G97R, L108V, E113D, G116E, A123T, Q125A, I126L, Y128H, T130A, L142I, V145M, S147N, N151T, F152I, H153L, H155Y, V156D, Q160L,
5 E163D, L167F, P169L, K188N, K191Q, C192S/F, S193G/A, F194Y, M196N, K198Q, K199(I, V, Q), Y200(L, A, G), Y203I, F204L, E205G, N206K, I207M, T208I, V217I/F, E226Q, S228P, L230V, V233I, I234T, E236D, I237F, S253P, P266Q, S273P, R274S, G284T/A, T285S, 287-3 bp deletion, R298H, P326A, L330V, G331A, P341L, L346I, S376L, D377A, G379A, L380F, S438P, and K441N. Generally, there was very high
10 diversity among the 76Gs.

The clones were expressed through *in vitro* translation and assayed for glycosylation activity using 0.5 mM steviol-13-*O*-glucoside and 0.5 mM stevioside as substrates, as described in previous examples. Reactions were carried out for 90 min at 30 °C. The native 76G1 activity was found in three new 76Gs designated 76G_C4,
15 76G_G7 and 76G_H12, by formation of 1,3-bioside when steviol-13-*O*-glucoside was used as substrate. Activity in this case was determined comparatively to the positive control, the functional 76G1. Clones 76G_G7 and 76G_H12 produced slightly higher levels of Reb A than the control but 76G_C4 had slightly less Reb A than the control. The number of changes in these clones represents a difference of about 7% at the amino
20 acid level, from the control enzyme. SEQ ID NOs: 98, 100, and 102 set forth the amino acid sequence of 76G_C4, 76G_G7, and 76G_H12, respectively. SEQ ID NOs: 97, 99, and 101 set forth the nucleotide sequences encoding 76G_C4, 76G_G7, and 76G_H12, respectively. SEQ ID NOs: 98, 100, and 102 set forth the amino acid sequence of
25 76G_C4, 76G_G7, and 76G_H12, respectively. SEQ ID NOs: 97, 99, and 101 set forth the nucleotide sequences encoding 76G_C4, 76G_G7, and 76G_H12, respectively.

Table 16 summarizes the amino acid changes of the 76G clones that had activity, as compared to the wildtype enzyme. There are a large number of overlapping polymorphisms in the active clones, thus it is expected that these polymorphisms do not cause a loss of activity for the enzyme. It appears that certain mutations are frequent in

inactive clones, such as the P→S mutation at position 80 or the F→V mutation at position 22.

TABLE 16

Clone	Mutations
76G_G7	M29I, V74E, V87G, L91P, G116E, A123T, Q125A, I126L, T130A, V145M, C192S, S193A, F194Y, M196N, K198Q, K199I, Y200L, Y203I, F204L, E205G, N206K, I207M, T208I, P266Q, S273P, R274S, G284T, T285S, 287-3 bp deletion, L330V, G331A, L346I
76G_H12	M29I, V74E, V87G, L91P, G116E, A123T, Q125A, I126L, T130A, V145M, C192S, S193A, F194Y, M196N, K198Q, K199I, Y200L, Y203I, F204L, E205G, N206K, I207M, T208I, P266Q, S273P, R274S, G284T, T285S, 287-3 bp deletion
76G_C4	M29I, V74E, V87G, L91P, G116E, A123T, Q125A, I126L, T130A, V145M, C192S, S193A, F194Y, M196N, K198Q, K199I, Y200L, Y203I, F204L, E205G, N206K, I207M, T208I

5

Example 19 – Expression of truncated yeast HMG-CoA reductase and other HMG-CoA reductases

In *S. cerevisiae*, the mevalonate pathway is heavily regulated, for example, at the level of the enzyme 3-Hydroxy-3-methylglutaryl-coenzyme A (HMG-CoA) reductase.

10 Expressing a truncated HMG-CoA reductase (tHMG1, encoding an enzyme stabilized from degradation) is one method in which flux towards PPP production can be increased in yeast. For example, expression of tHMG1 in yeast has led to dramatic overproduction of β-carotene. See, Verwaal et al., 2007, Appl. Environ. Microbiol. 73:4342.

15 Interestingly, such yeast did not show a darker orange coloration on solid growth medium as was expected, but rather a stronger yellow color, likely due to even higher over-production of the intermediate phytoene.

To determine if expression of HMG-CoA reductase could be used to improve flux to the steviol and steviol glycoside pathways, a yeast reporter strain for testing isoprenoid

flux was prepared by substituting the inherent promoter of the *ERG9* gene with a *CUP1* promoter. See, U.S. Patent Application No. 61/346853, filed May 20, 2010.

The genes used to produce the yeast strain are shown in Table 17. The genes from the source organisms were codon optimized according to DNA 2.0 IncTM. For the purpose of monitoring the cellular prenyl phosphate availability, a construct was produced which had a high copy number plasmid containing gene expression cassettes (methionine-repressible promoters) with the genes for the three enzymes needed to turn prenyl phosphates into β -carotene (GGPP synthase from *Xanthophyllomyces dendrorhous*, phytoene synthase and beta carotene synthase from *X. dendrorhous*, and zeta carotene synthase and delta carotene synthase from *Neurospora crassa*). See, Verwaal et al., 2007 supra; and U.S. Patent Application No. 61/346853.

Table 17 Sources of HMG CoA Reductases and other Mevalonate Genes

Accession#	Organism	Enzyme	Size (nt)	Gene name	SEQ ID (codon optimized)	SEQ ID (protein)
XM_001467423	<i>Leishmania infantum</i>	Acetyl-CoA C-acetyltransferase	1323	MEV-4	103	104
YML075C	<i>Saccharomyces cerevisiae</i>	Truncated HMG (tHMG1)	1584	tHMG1	105	106
EU263989	<i>Ganoderma lucidum</i>	3-HMG-CoA reductase	3681	MEV-11	107	108
BC153262	<i>Bos taurus</i>	3-HMG-CoA reductase	2667	MEV-12	109	110
AAD47596	<i>Artemisia annua</i>	3-HMG-CoA reductase	1704	MEV-13	111	112
AAB62280	<i>Trypanosoma cruzi</i>	3-HMG-CoA reductase	1308	MEV-14	113	114
CAG41604	<i>Staph aureus</i>	3-HMG-CoA reductase	1281	MEV-15	115	116
DNA2.0 sequence	<i>Archaeoglobus fulgidus</i>	3-HMG-CoA reductase	1311	HMG reductase	117	118
DNA2.0 sequence	<i>Pseudomonas mevalonii</i>	3-HMG-CoA reductase	1287	HMG reductase	119	120

The yeast tHMG1 was expressed in the CEN.PK-based yeast strain that produces β -carotene, resulting in a color change from orange to light yellow. Interestingly, expression of the full length HMGs from *Artemisia annua*, *Trypanosoma cruzi* and *Staphylococcus aureus*, as well as the NADH-dependent HMG's from *Pseudomonas mevalonii* and *Archeoglobus fulgidus* produced a similar result, indicating these genes also improve the flux through the mevalonate pathway in yeast (similar overexpression of *Bos taurus* HMG had no such effect). Finally, the same color change was seen after overexpression of *Leishmania infantum* acetyl-CoA C-acetyltransferase (first enzyme of mevalonate pathway, described in Table 17) or native *S. cerevisiae* (*CAB1*, YDR531W) or *B. subtilis*, (acc. No. YP004204141) pantothenate kinases (known to result in increased acetyl-CoA production).

To test if the color change in these experiments were indeed due to higher GGPP availability, the yeast tHMG1, *P. mevalonii* or *S. aureus* HMGs, or *B. subtilis* pantothenate kinase were expressed in a stable 19-SMG producer strain. None of these constructs appeared to produce an increase in 19-SMG or rubusoside production (UGT85C2 co-expressed) under the conditions tested. Mevalonate feeding to the yeast reporter strain also did not result in increased rubusoside production. The rubusoside reporter strain, however, has not been genetically modified to reduce the *ERG9*-encoded flux towards ergosterol biosynthesis. It is expected that control of flux to ergosterol production would result in increased steviol glycoside production using the HMG reductase genes and other mevalonate pathway genes found to be beneficial to beta-carotene production.

Example 20 – Production of RebC *in vivo*

The synthesis of a precursor molecule to Rebaudioside C, steviol-13-O-glucopyranosyl-1,2-rhamnoside, was shown *in vitro* in Example 15. In that example steviol-13-O-monoglucoside was used as a substrate, along with UDP-glucose and the *Arabidopsis thaliana* RHM2 enzyme (locus tag AT1G53500) and UGT91D2c. To further demonstrate the pathway shown in Fig. 2B, production of Rebaudioside C from steviol was accomplished *in vivo*.

A yeast strain capable of producing Rebaudioside C was constructed, and production of rebaudioside C and rebaudioside A was assayed by LC-MS. A modified *Saccharomyces cerevisiae* strain BY4742 was constructed and designated EYS583-7A. The use of BY4742 has been described by Naesby et al., Microb Cell Fact. 8:45 (2009)

5 All four UGTs (91D2d, 76G1, 74G1, and 85C2) were constitutively expressed in plasmids with GPD promoters. This type of strain has been described by Naesby et. al, Microb Cell Fact. 8:45 (2009). UGT85C2 was inserted in plasmid P423 GPD (ATCC#87355), UGT74G1 was cloned into P424 GPD (ATCC#87357) and both UGT91D2c and UGT76G1 were cloned into P425-GPD (ATCC#87359) with 91D2c in

10 the original multiple cloning site (MCS), and 76G1 inserted with an additional GPD promoter and a CYC terminator. The resulting strain was transformed with plasmid P426 GPD (ATCC#87361) containing the *RHM2* gene expressed from the GPD promoter. The strain was grown on SC medium lacking histidine, leucine, tryptophan and uracil for 24 hours. The culture was then re-inoculated to an OD₆₀₀ of 0.4 in fresh media containing

15 25 μ M steviol, and the yeast was allowed to grow for 72 more hours before detecting if Rebaudioside C was present in the supernatant and the cell pellets. Rebaudioside C was quantified using an authentic Rebaudioside C standard (Chromadex, Irvine CA). A total of $1.27 \mu\text{M} \pm 0.36 \mu\text{M}$ of RebC was detected in the supernatant. Similarly, $3.17 \mu\text{M} \pm 1.09 \mu\text{M}$ RebA was detected in the cell pellet One of skill in the art will recognize that

20 different ratios of RebC to RebA can be obtained by modulation of the activity of the *RHM2* enzyme and/or by usage of UGT91D2e or UGT76G1-like enzymes with higher activity for the UDP-rhamnose reactions. The alternative UGTs can be mutagenized versions of the wildtype enzymes or unique enzymes that are obtained through discovery initiatives.

25 One of skill in the art will recognize that a yeast strain capable of production of Rebaudioside A from glucose, such as strain CEY213 transformed with a plasmid containing UGT91D2c and UGT76G1 in Example 12 would produce Rebaudioside C with the addition of the *RHM2* gene either via a vector or integrated into the chromosome.

30

**Example 21 -- Production of Steviol Glycosides using UGTs expressed
in *Escherichia coli***

Activity of UGT enzymes in gram negative bacteria

5 The wildtype genes for UGTs 91D2e, 74G1, 76G1, and 85C2 were cloned individually into *E. coli* XJB-autolysis BL21(DE3) cells using the pET30 vector system from Novagen (EMD4 Biosciences, Madison, WI), except for UGT91D2e, which was cloned into a pGEX 4T-1 (GE Healthcare, Uppsala, Sweden) vector. Similar cloning was described in Examples 7 and 10. All vectors use an IPTG-inducible promoter. Plasmid DNA was transformed into chemically competent cells as described by the vendor.

10 Transformants displaying the desired antibiotic resistance were grown overnight at 30°C in 2 mL cultures using NZCYM- media and antibiotic. For *in vivo* feedings, 5 cultures were grown: UGT 91d2e, 74G1, 76G1, and 85C2 individually, and a mix of all 4 clones. The following day, the cultures were induced to a final concentration of 0.3 mM IPTG and 3 mM arabinose, and grown 2 days at 20°C in the presence of 50 µM steviol (UGT74G1, UGT85C2 and the quadruple mix) or 50µM rubusoside (UGT91D2e and 15 UGT76G1). The temperature was raised to 30°C and the cells were grown for one more day. The cells were then harvested by centrifugation at 4000 rpm for 5 min., and the supernatants were removed for LC-MS analysis. The cells were resuspended in 50% DMSO, lysed at 80°C for 5 min and the lysates were analyzed by LC-MS.

20 For *in vitro* assays, transformants displaying the desired antibiotic resistance were grown overnight at 30°C in 2 mL cultures using NZCYM- media and antibiotic. The following day, the cultures were induced to a final concentration of 0.3 mM IPTG and 3 mM arabinose, and grown for 24h at 20°C. The cells were then harvested by centrifugation at 4000 rpm for 5 min and resuspended in 200 µL GT-buffer (RBC 25 Bioscience) and 3 tablets/100ml of Complete mini, protease inhibitor (Roche), transferred to Eppendorf tubes, vortexed and frozen at -80°C for 1.5 hour. Cells were thawed on ice, and left at room temperature for 3 minutes. When approximately half-way thawed, 15 µl of 0.14 mg/ml H₂O DNase solution + 30 µl 0.05M MgCl₂ was added to each tube and the samples were incubated for approximately 5 minutes at room temperature. The cells were

centrifuged at maximum speed for 5 minutes. One-hundred μL of supernatant (lysate) was transferred to fresh microfuge tubes, and 100 μL of glycerol was added.

Enzyme assays were performed by adding 15.15 μL H_2O , 7.5 μL 4X Buffer (400mM Tris, 20mM MgCl_2 , 4mM KCl), 0.3 μL FastAPTM (1u/ μL) from Fermentas, 0.45 μL of a 100 mM stock of UDP-glucose, 0.6 μL of substrate (steviol or rubusoside) and 6 μL of the crude enzyme preparations described above. UGT74G1, UGT85C2, as well as all four UGTs mixed were incubated with steviol. UGT 76G1 and 85C2 were incubated with rubusoside. The enzyme assays were incubated overnight at 37°C.

Following centrifugation at 4000 rpm for 5 minutes, 30 μL samples were transferred to a fresh 96 well plate and 30 μL of DMSO was added. The samples were then subjected to LC-MS analysis. Similar *in vitro* experiments were also done using steviol 1,2-bioside (for UGT76G1 and UGT74G1) or Rebaudioside B (for UGT74G1) as substrates.

No activity was detected in the *in vivo* feedings. Table 18 illustrates the results for the *in vitro* assays.

Table 18

Tube	UGT Clone(s)	Substrate fed	Product detected
1	74G1	Steviol	19-SMG, low levels of rubusoside
2	85G1	Steviol	13-SMG, low levels of rubusoside
3	76G1	Rubusoside	1,3-stevioside, an unknown tetra-glycoside
4	91D2c	Rubusoside	stevioside
5	Mix of 4 crude UGT preparations	Steviol	Rubusoside, 1,3-stevioside, trace RebA (no monosides)
6	76G1	Steviol 1,2-bioside	Rebaudioside B
7	74G1	Steviol 1,2-bioside	Stevioside
8	74G1	Rebaudioside B	Rebaudioside A

These results indicate that the UGT enzymes are all active in *E. coli* cells. However, the substrates may not be readily imported into the cytoplasm. It is expected

that if the steviol were produced in *E. coli* from precursor pathways, the production of the various steviol glycoside products would be feasible from glucose. It is unexpected that the 74G1 and 85G1 UGTs, which have slightly overlapping substrate specificities, can produce rubusoside from steviol singly. The mix of the four crude enzyme preparations gave very low levels of the monosides, which indicates that the conversion to di- and tri-glycosides was efficient. With respect to UGT91D2e, the preparation that was used had lost some of its original activity after long-term storage. It is expected that a fresh preparation of the enzyme would have yielded higher levels of Rebaudioside A.

Example 22 -- Production of Steviol Glycosides in *Physcomitrella patens*

Feeding experiments in moss cells

The genes for UGT 91d2e, 74G1, 76G1, and 85C2 were cloned into *Physcomitrella patens* using the pTHUbi:Gateway vector system described in U.S. Patent Publication No. 20100297722. This vector uses a strong maize Ubiquitin promoter. PCR primers were designed to amplify the coding regions in previous examples (native sequences) with the addition of "CACC" upstream of the start codon. Plasmid DNA was digested with SmaI and used for transformation into protoplasts (generally around 0.5×10^6 protoplasts). Transformants displaying the desired resistance were grown 1 day in 10 mL cultures and then fed either steviol, rubusoside, or buffer + DMSO as indicated by Table 19. One-half mL of buffer containing substrate was added per 10 mL of culture, and final concentrations of 0.1% DMSO, 50 μ M steviol or rubusoside, and 0.125 mM phosphate buffer were added to the cultures. A positive control was done where the YFP (yellow fluorescent protein) was expressed in the presence of steviol or just buffer and DMSO. Cultures were grown 2 more days prior to separation of cells and freezing in liquid nitrogen until further analysis. In some cases multiple UGT-containing plasmids were transformed into the same protoplast cells, to illustrate conversion of multiple steps within the moss cells.

Table 19

Tube	UGT Gene(s)	Substrate fed
-------------	--------------------	----------------------

1	YFP (control)	none
2	YFP	Steviol (50 μ M)
3	74G1	none
4	76G1	none
5	85C2	none
6	91D2E	none
7	74G1	Steviol (50 μ M)
8	76G1	Steviol (50 μ M)
9	85C2	Steviol (50 μ M)
10	91D2E	Steviol (50 μ M)
11	74G1/85C2	none
12	74G1/85C2	Steviol (50 μ M)
13	74G1/85C2/91D2E	none
14	74G1/85C2/91D2E	Steviol (50 μ M)
15	76G1	Rubusoside (50 μ M)
16	91D2E	Rubusoside (50 μ M)
17	76G1/91D2E	none
18	76G1/91D2E	Rubusoside (50 μ M)

Expression was positive in the controls (tubes 1 and 2) as measured by fluorescent signal observation. The supernatants from the experiments were analyzed by LC-MS; 200 μ L of each supernatant sample was mixed with an equal volume of 50 percent DMSO. The samples were spun (15,700 relative centrifugal force, 10 minutes) and 100 microliters of the resulting supernatant was analyzed by LC-MS.

Protoplast pellets were thawed on ice and 10 mM Tris-HCl pH 8 containing 3 tablets/100 ml of Complete Mini Protease Inhibitor (Roche) was added to reach a final volume of 150 μ L. The solutions were divided in two: 75 μ L was transferred to a new tube and protoplasts were pelleted (15,700 relative centrifugal force, 1 minute). Pellets were washed with 75 μ L Milli-Q water before resuspension in 150 μ L DMSO (50

percent). Samples were then heated (80 degrees Celsius, 10 minutes), vortexed and centrifuged (15,700 relative centrifugal force, 10 minutes). Fifty μ L of the resulting supernatant was analyzed by LC-MS.

No steviol glycoside production was detectable in supernatants or pellets. It is unknown if the steviol and rubusoside can be transported into moss cells.

In vitro feeding of pellet extracts

In vitro feeding experiments were conducted with samples 1, 3, 4, 5, 6, 11, 13 and 17). Glass beads (425-600 microns) were added to the remaining 75 μ L of the original resuspensions and protoplasts were mechanically lysed by vortexing 3 times, 2 minutes each time, at 4 degrees Celsius and storage on ice in between vortexing. The samples were spun (15,700 relative centrifugal force, 10 minutes, 4 degrees Celsius) and 6 μ L of resulting supernatants was used in *in vitro* enzyme reactions. For the enzyme reactions FastAPTM phosphatase (Fermentas) was used (0.3 U/reaction) and the UDP-glucose:substrate ratio was 5. The samples were fed either steviol or rubusoside according to Table 20.

Table 20

Cell extract from tube	UGT Gene(s)	Substrate fed
1	YFP	None
1	YFP	0.5 mM steviol
1	YFP	0.5 mM rubusoside
3	74G1	0.5 mM steviol
4	76G1	0.5 mM rubusoside
5	85C2	0.5 mM steviol
6	91D2E	0.5 mM rubusoside
11	74G1/85C2	0.5 mM steviol
13	74G1/85C2/91D2E	0.5 mM steviol
17	76G1/91D2E	0.5 mM rubusoside

Reactions were incubated at 30 °C overnight. After incubation, an equal amount of DMSO (100 percent) was added to the samples and mixed, then the sample was spun (15,700 relative centrifugal force, 10 minutes) and 30 µL of the resulting supernatant was analyzed by LC-MS.

LC-MS analysis showed conversion of rubusoside to 1,3-stevioside by UGT76G1. None of the other steviol glycosides were detectable. It is unknown if soluble expression of the UGTs occurred in *Physcomitrella*. It is expected if one UGT is active in the moss cells, the others would also be active if expression occurred. In addition, the cloning was done in a transient manner. Stable integration of the genes is expected to produce additional clones that are active for UGT activity when tested.

Methods are known to those with skill in the art for increasing soluble expression of recombinant proteins. Alternative promoters, ribosome binding sites, codon usage, co-expression with chaperones, and change in temperature are non-limiting examples of methods for increasing soluble expression of recombinant proteins.

Example 23 -- Production of Steviol Glycosides in *Aspergillus nidulans*

Activity of UGT enzymes in fungal cells

The native genes for UGT 91D2c, 74G1, 76G1, and 85C2 were cloned into *Aspergillus nidulans* using a PCR-fabricated expression cassette and the USER vector system. Cloning methods are described in Hansen et al., Appl. Environ. Microbiol. 77: 3044-3051 (2011). Briefly, a nucleotide sequence encoding each UGT was inserted between the constitutive P_{gpdA} promoter and the T_{trpC} terminator, in a vector containing additionally two targeting sequences for genomic integration and *argB* as selection marker. Plasmid DNA was transformed into *A. nidulans* protoplasts according to Nielsen et al., Fungal Genet. Biol. 43:54–64 (2006) and Johnstone et al., EMBO J. 4:1307-1311 (1985). Transformants displaying the desired resistance were grown for 48 hours in 150 mL cultures using minimal media (1% Glucose; 10 mM NaNO₃; mineral mix).

Cell lysates prepared by disruption of the mycelia with glass beads were used to

determine the activities of the individual UGTs in *in vitro*. The cell lysates of strains expressing 74G1 and 85C2 were incubated with 0.5 mM steviol and the strains expressing 76G1 and 91D2e were incubated with 0.5 mM steviol-13-*O*-glucoside for 24 hours, and the supernatants further analyzed using LC/MS. No steviol glycosides were detected.

It is unknown whether soluble expression of the UGT enzymes was achieved as these products are not typically visible on SDS-PAGE. Since *Aspergillus* and *Saccharomyces* are both fungi, it is expected that additional experimentation would result in active clones. Methods are known to those with skill in the art for increasing soluble expression of recombinant proteins. Alternative promoters, inducer levels, ribosome binding sites, codon usage, co-expression with chaperones, and change in temperature are non-limiting examples of methods for increasing soluble expression of recombinant proteins.

OTHER EMBODIMENTS

It is to be understood that while the invention has been described in conjunction with the detailed description thereof, the foregoing description is intended to illustrate and not limit the scope of the invention, which is defined by the scope of the appended claims. Other aspects, advantages, and modifications are within the scope of the following claims.

WHAT IS CLAIMED IS:

- 1 1. A recombinant host comprising a recombinant gene encoding a UGT91D2
2 polypeptide.
3
- 4 2. The recombinant host of claim 1, wherein said UGT91D2 polypeptide is
5 UGT91D2e or UGT91D2m.
6
- 7 3. The recombinant host of claims 1 or 2, wherein said UGT91D2
8 polypeptide has at least 90% identity to the amino acid sequence set forth in SEQ ID
9 NO:5.
10
- 11 4. The recombinant host of claim 3, wherein said UGT91D2 polypeptide has
12 at least 95% identity to the amino acid sequence set forth in SEQ ID NO:5.
13
- 14 5. The recombinant host of claim 4, wherein said UGT91D2 polypeptide has
15 at least 99% identity to the amino acid sequence set forth in SEQ ID NO:5.
16
- 17 6. The host of any one of claims 1-5, wherein said host is a microorganism.
18
- 19 7. The host of claim 6, wherein said microorganism is a *Saccharomycete*.
20
- 21 8. The host of claim 6, wherein said *Saccharomycete* is *Saccharomyces*
22 *cerevisiae*.
23
- 24 9. The host of claim 6, wherein said microorganism is *Escherichia coli*.
25
- 26 10. The host of any of claims 1-5, wherein said host is a plant or plant cell.
27

28 11. The host of claim 10, wherein said plant or plant cell is a *Stevia*,
29 *Physcomitrella*, or tobacco plant or plant cell.

30
31 12. The host of claim 11, wherein said *Stevia* plant or plant cell is a *Stevia*
32 *rebaudiana* plant or plant cell.

33
34 13. The host of any one of claims 1-12, wherein said polypeptide comprises at
35 least one amino acid substitution at residues 1-19, 27-38, 44-87, 96-120, 125-141, 159-
36 184, 199-202, 215-380, or 387-473 of SEQ ID NO:5.

37
38 14. The host of any one of claims 1-13, wherein said polypeptide comprises
39 an amino acid substitution at one or more residues selected from the group consisting of
40 residues 30, 93, 99, 122, 140, 142, 148, 153, 156, 195, 196, 199, 206, 207, 211, 221, 286,
41 343, 427, and 438 of SEQ ID NO:5.

42
43 15. The host of claim 14, wherein said polypeptide comprises an arginine at
44 residue 206, a cysteine at residue 207, and an arginine at residue 343 relative to SEQ ID
45 NO:5.

46
47 16. The host of claim 14, wherein said polypeptide comprises a phenylalanine
48 at residue 30, a glutamine at residue 93, a valine at residue 99, a phenylalanine at residue
49 122, a tyrosine at residue 140, a cysteine at residue 142, a threonine at residue 148, an
50 alanine at residue 153, a serine at residue 156, a methionine at residue 195, a glutamic
51 acid at residue 196, a glutamic acid at residue 199, a methionine at residue 211, a
52 phenylalanine at residue 221, an alanine at residue 286, an asparagine at residue 427, or
53 an alanine at residue 438 relative to SEQ ID NO:5.

54
55 17. The host of any one of claims 1-15, wherein said polypeptide comprises
56 the amino acid sequence of SEQ ID NO:95.

57

58 18. The host of any one of claims 1-12, wherein said polypeptide comprises
59 the amino acid sequence of SEQ ID NO:5.

60
61 19. The host of any one of claims 1-18, said host further comprising a
62 recombinant gene encoding a UGT85C polypeptide having at least 90% identity to the
63 amino acid sequence set forth in SEQ ID NO:3.

64
65 20. The host of claim 19, wherein said UGT85C polypeptide comprises one or
66 more amino acid substitutions at residues 9, 10, 13, 15, 21, 27, 60, 65, 71, 87, 91, 220,
67 243, 270, 289, 298, 334, 336, 350, 368, 389, 394, 397, 418, 420, 440, 441, 444, and 471
68 of SEQ ID NO:3.

69
70 21. The host of any one of claims 1-20, said host further comprising a
71 recombinant gene encoding a UGT76G polypeptide having at least 90% identity to the
72 amino acid sequence set forth in SEQ ID NO:7.

73
74 22. The host of claim 21, said UGT76G polypeptide having one or more
75 amino acid substitutions at residues 29, 74, 87, 91, 116, 123, 125, 126, 130, 145, 192,
76 193, 194, 196, 198, 199, 200, 203, 204, 205, 206, 207, 208, 266, 273, 274, 284, 285, 291,
77 330, 331, and 346 of SEQ ID NO:7.

78
79 23. A recombinant host, said host comprising a recombinant gene encoding a
80 UGT85C polypeptide having at least 90% identity to the amino acid sequence set forth in
81 SEQ ID NO:3, and having one or more amino acid substitutions at residues 9, 10, 13, 15,
82 21, 27, 60, 65, 71, 87, 91, 220, 243, 270, 289, 298, 334, 336, 350, 368, 389, 394, 397,
83 418, 420, 440, 441, 444, and 471 of SEQ ID NO:3.

84
85 24. The host of claim 23, wherein said polypeptide comprises substitutions at
86 residues 13, 15, 60, 270, 289, and 418 of SEQ ID NO:3.

87

88 25. The host of claim 23, wherein said polypeptide comprises a) substitutions at
89 residues 13, 60, and 270 of SEQ ID NO:3; b) substitutions at residues 60 and 87 of SEQ
90 ID NO:3; c) substitutions at residues 65, 71, 220, 243, and 270 of SEQ ID NO:3; d)
91 substitutions at residues 65, 71, 220, 243, 270, and 441 of SEQ ID NO:3; e) substitutions
92 at residues 65, 71, 220, 389, and 394 of SEQ ID NO:3; f) substitutions at residues 65, 71,
93 270, and 289 of SEQ ID NO:3; g) substitutions at residues 15 and 65 of SEQ ID NO:3; h)
94 substitutions at residues 65 and 270 of SEQ ID NO:3; i) substitutions at residues 65 and
95 440 of SEQ ID NO:3; j) substitutions at residues 65 and 441 of SEQ ID NO:3; k)
96 substitutions at residues 65 and 418 of SEQ ID NO:3; l) substitutions at residues 220,
97 243, 270, and 334 of SEQ ID NO:3; or m) substitutions at residues 270 and 289 of SEQ
98 ID NO:3.

99

100 26. A recombinant host comprising a recombinant gene encoding a UGT76G
101 polypeptide having at least 90% identity to the amino acid sequence set forth in SEQ ID
102 NO:7, and having one or more amino acid substitutions at residues 29, 74, 87, 91, 116,
103 123, 125, 126, 130, 145, 192, 193, 194, 196, 198, 199, 200, 203, 204, 205, 206, 207, 208,
104 266, 273, 274, 284, 285, 291, 330, 331, and 346.

105

106 27. The host of claim 26, wherein said polypeptide has a) substitutions at amino
107 acid residues 74 , 87, 91, 116, 123, 125, 126, 130, 145, 192, 193, 194, 196, 198, 199, 200,
108 203, 204, 205, 206, 207, 208, and 291; b) substitutions at residues 74 , 87, 91, 116, 123,
109 125, 126, 130, 145, 192, 193, 194, 196, 198, 199, 200, 203, 204, 205, 206, 207, 208, 266,
110 273, 274, 284, 285, and 291; or c) substitutions at residues 74 , 87, 91, 116, 123, 125,
111 126, 130, 145, 192, 193, 194, 196, 198, 199, 200, 203, 204, 205, 206, 207, 208, 266, 273,
112 274, 284, 285, 291, 330, 331, and 346.

113

114 28. The host of any one of claims 23-27, wherein said host is a microorganism.

115

116 29. The host of claim 28, wherein said microorganism is a Saccharomycete.

117

118 30. The host of claim 29, whercin said Saccharomycete is *Saccharomyces*
119 *cerevisiae*.
120
121 31. The host of claim 28, whercin said microorganism is *Escherichia coli*.
122
123 32. The host of any one of claims 23-27, wherein said host is a plant or plant
124 cell.
125
126 33. The host of claim 32, whercin said plant or plant cell is a *Stevia*,
127 *Physcomitrella*, or tobacco plant or plant cell.
128
129 34. The host of claim 33, whercin said *Stevia* plant or plant cell is a *Stevia*
130 *rebaudiana* plant or plant cell.
131
132 35. The host of any one of claims 1-34, said host further comprising a gene
133 encoding a UGT74G1 polypeptide.
134
135 36. The host of claim 35, whercin said gene encoding said UGT74G1
136 polypeptide is a recombinant gene.
137
138 37. The host of any one of claims 1-36, said host further comprising one or
139 more of: (i) a gene encoding a geranylgeranyl diphosphate synthase;
140 (ii) a gene encoding a bifunctional copalyl diphosphate synthase and kaurene
141 synthase, or a gene encoding a copalyl diphosphate synthase and a gene encoding a
142 kaurene synthase;
143 (iii) a gene encoding a kaurene oxidase; and
144 (iv) a gene encoding a steviol synthetase.
145
146 38. The host of any one of claims 1-37, said host further comprising one or
147 more of (v) a gene encoding a truncated HMG-CoA;

148 (vi) a gene encoding a CPR;
149 (vii) a gene encoding a rhamnose synthetase;
150 (viii) a gene encoding a UDP-glucose dehydrogenase; and
151 (ix) a gene encoding a UDP-glucuronic acid decarboxylase.

152

153 39. The host of claim 38, wherein at least one of the genes of (i), (ii), (iii),
154 (iv), (v), (vi), (vii), (viii), or (ix) is a recombinant gene.

155

156 40. The host of claim 37, wherein each of the genes of (i), (ii), (iii), and (iv) is
157 a recombinant gene.

158

159 41. An isolated nucleic acid encoding a polypeptide having at least 90%
160 identity to the amino acid sequence set forth in SEQ ID NO:5.

161

162 42. The nucleic acid of claim 41, wherein said polypeptide has at least 95%
163 identity to the amino acid sequence set forth in SEQ ID NO:5.

164

165 43. The nucleic acid of claim 41 or claim 42, wherein said polypeptide has at
166 least 99% identity to the amino acid sequence set forth in SEQ ID NO:5.

167

168 44. The nucleic acid of any one of claims 41-43, wherein said polypeptide
169 comprises at least one amino acid substitution at residues 1-19, 27-38, 44-87, 96-120,
170 125-141, 159-184, 199-202, 215-380, or 387-473 of SEQ ID NO:5.

171

172 45. The nucleic acid of any one of claims 41-44, wherein said polypeptide
173 comprises an amino acid substitution at one or more residues selected from the group
174 consisting of residues 30, 93, 99, 122, 140, 142, 148, 153, 156, 195, 196, 199, 206, 207,
175 211, 221, 286, 343, 427, and 438 of SEQ ID NO:5.

176

177 46. The nucleic acid of any one of claims 41-45, wherein said polypeptide
178 comprises an arginine at residue 206, a cysteine at residue 207, and an arginine at residue
179 343 of SEQ ID NO:5.

180
181 47. The nucleic acid of any one of claims 41-45, wherein said polypeptide
182 comprises a phenylalanine at residue 30, a glutamine at residue 93, a valine at residue 99,
183 a phenylalanine at residue 122, a tyrosine at residue 140, a cysteine at residue 142, a
184 threonine at residue 148, an alanine at residue 153, a serine at residue 156, a methionine
185 at residue 195, a glutamic acid at residue 196, a glutamic acid at residue 199, a
186 methionine at residue 211, a phenylalanine at residue 221, an alanine at residue 286, an
187 asparagine at residue 427, or an alanine at residue 438 of SEQ ID NO:5.

188
189 48. An isolated polypeptide having an amino acid sequence with at least 90%
190 identity to the amino acid sequence of SEQ ID NO:5.

191
192 49. A recombinant host, said host comprising:
193 (i) a gene encoding a geranylgeranyl diphosphate synthase;
194 (ii) a gene encoding a bifunctional copalyl diphosphate synthase and kaurene
195 synthase, or a gene encoding a copalyl diphosphate synthase and a gene encoding a
196 kaurene synthase;
197 (iii) a gene encoding a kaurene oxidase; and
198 (iv) a gene encoding a steviol synthetase;
199 wherein at least one of said genes is a recombinant gene.

200
201 50. The host of claim 49, wherein said host produces steviol when cultured
202 under conditions in which each of said genes is expressed.

203
204 51. The host of claim 50, wherein said steviol accumulates to at least 1 mg/L
205 of culture medium when cultured under said conditions.

206

207 52. The host of any one of claims 49-51, wherein said geranylgeranyl
208 diphosphate synthase has greater than 90 % sequence identity to one of the amino acid
209 sequences set forth in SEQ ID NOs: 121-128.
210
211 53. The host of any one of claims 49-52, wherein said copalyl diphosphate
212 synthase has greater than 90 % sequence identity to the amino acid sequence set forth in
213 SEQ ID NOs: 129-131.
214
215 54. The host of any one of claims 49-53, wherein said kaurene synthase has
216 greater than 90 % sequence identity to one of the amino acid sequences set forth in SEQ
217 ID NOs: 132-135.
218
219 55. The host of any one of claims 49-54, wherein said kaurene oxidase has
220 greater than 90 % sequence identity to one of the amino acid sequences set forth in SEQ
221 ID NOs: 138-141.
222
223 56. The host of any one of claims 49-55, wherein said steviol synthetase has
224 greater than 90 % sequence identity to one of the amino acid sequences set forth in SEQ
225 ID NOs: 142-146.
226
227 57. The host of any one of claims 49-56, further comprising a gene encoding a
228 truncated HMG-CoA.
229
230 58. The host of any one of claims 49-57, further comprising a gene encoding a
231 CPR.
232
233 59. The host of any one of claims 49-58, said host further comprising a gene
234 encoding a UGT74G1 polypeptide.
235

236 60. The host of any one of claims 49-59, said host further comprising a gene
237 encoding a UGT85C2 polypeptide.

238

239 61. The host of any one of claims 49-60, said host further comprising a gene
240 encoding a UGT76G1 polypeptide.

241

242 62. The host of any one of claims 49-61, said host further comprising a gene
243 encoding a UGT91D2 polypeptide.

244

245 63. The host of any one of claims 59-62, wherein said host produces at least
246 one steviol glycoside when cultured under conditions in which each of said genes is
247 expressed.

248

249 64. The host of claim 63, wherein said steviol glycoside is selected from the
250 group consisting of steviol-13-*O*-glucoside, steviol-19-*O*-glucoside, rubusoside,
251 rebaudioside A, rebaudioside B, rebaudioside C, rebaudioside D, rebaudioside E,
252 rebaudioside F, and dulcoside A.

253

254 65. The host of claim 64, wherein said steviol glycoside accumulates to at
255 least 1 mg/liter of culture medium when cultured under said conditions.

256

257 66. The host of claim 65, wherein said steviol glycoside accumulates to at
258 least 10 mg/liter of culture medium when cultured under said conditions.

259

260 67. The host of claim 66, wherein steviol glycoside accumulates to at least 20
261 mg/liter of culture medium when cultured under said conditions.

262

263 68. The host of any one of claims 49-67, said host further comprising one or
264 more of

265 i) a gene encoding a deoxyxylulose 5-phosphate synthase (DXS);

266 ii) a gene encoding a D-1-deoxyxylulose 5-phosphate reductoisomerase (DXR);
267 iii) a gene encoding a 4-diphosphocytidyl-2-C-methyl-D-erythritol synthase
268 (CMS);
269 iv) a gene encoding a 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase (CMK);
270 v) a gene encoding a 4-diphosphocytidyl-2-C-methyl-D-erythritol 2,4-
271 cyclodiphosphate synthase (MCS);
272 vi) a gene encoding a 1-hydroxy-2-methyl-2(E)-butenyl 4-diphosphate synthase
273 (HDS);
274 vii) a gene encoding a 1-hydroxy-2-methyl-2(E)-butenyl 4-diphosphate reductase
275 (HDR).

276
277 69. The host of any one of claims 49-68, said host further comprising one or
278 more of:

279 ix) a gene encoding a acetoacetyl-CoA thiolase;
280 x) a gene encoding a truncated HMG-CoA reductase;
281 xi) a gene encoding a mevalonate kinase;
282 xii) a gene encoding a phosphomevalonate kinase;
283 xiii) a gene encoding a mevalonate pyrophosphate decarboxylase.

284
285 70. The host of any one of claims 49-69, wherein expression of one or more of
286 said genes is inducible.

287
288 71. The host of any one of claims 49-70, wherein said host is a
289 microorganism, plant, or plant cell.

290
291 72. The host of claim 71, wherein said host is *Saccharomyces cerevisiae*.

292
293 73. A method of producing steviol or a steviol glycoside, comprising:
294 a) growing the host of claim 71 or claim 72 in a culture medium, under
295 conditions in which said genes are expressed; and

296 b) recovering the steviol or steviol glycoside produced by said host.

297

298 74. The method of claim 73, wherein said growing step comprises inducing
299 expression of one or more of said genes.

300

301 75. The method of claim 73 or 74, wherein said steviol or steviol glycoside is
302 selected from the group consisting of steviol-13-O-glucoside, steviol-19-O-glucoside,
303 rubusoside, rebaudioside A, rebaudioside B, rebaudioside C, rebaudioside D,
304 rebaudioside E, rebaudioside F, and dulcoside A.

305

306 76. A recombinant *Saccharomyces* strain comprising:

307 (i) a gene encoding a bifunctional copalyl diphosphate synthase and kaurene
308 synthase, or

309 (ii) a gene encoding a copalyl diphosphate synthase and a gene encoding a
310 kaurene synthase;

311 wherein said strain produces ent-kaurene upon expression of said copalyl
312 diphosphate synthase and said kaurene synthase.

313

314 77. An isolated nucleic acid having greater than 90 % sequence identity to one
315 of the nucleotide sequences set forth in SEQ ID NOs: 18-25, 34-36, 4-43, 48, 49, 52-55,
316 60-64, 70-72, 77, or 79.

317

318 78. A recombinant host, said host comprising:

319 (i) a gene encoding a UGT74G1;

320 (ii) a gene encoding a UGT85C2;

321 (iii) a gene encoding a UGT76G1; and

322 (iv) a gene encoding a UGT91D2,

323 wherein at least one of said genes is a recombinant gene.

324

325 79. The host of claim 78, wherein each of said genes is a recombinant gene.

326

327 80. The host of claim 78 or 79, wherein said host produces at least one steviol
328 glycoside when cultured under conditions in which each of said genes is expressed.

329

330 81. The host of any one of claims 78-80, further comprising

331 (a) a gene encoding a bifunctional copalyl diphosphate synthase and kaurene
332 synthase, or a gene encoding a copalyl diphosphate synthase and a gene encoding a
333 kaurene synthase;

334 (b) a gene encoding a kaurene oxidase;

335 (c) a gene encoding a steviol synthetase;

336 (d) a gene encoding a geranylgeranyl diphosphate synthase.

337

338 82. The host of any one of claims 78-81, wherein said steviol glycoside is
339 rebaudioside A.

340

341 83. An isolated nucleic acid encoding a polypeptide having greater than 90 %
342 sequence identity to the amino acid sequence of UGT91D2c or UGT91D2m, excluding
343 the amino acid sequence of UGT91D2m.

344

345 84. An isolated polypeptide having greater than 90 % sequence identity to the
346 amino acid sequence of UGT91D2c or UGT91D2m, excluding the amino acid sequence
347 of UGT91D2m.

348

349 85. A steviol glycoside composition produced by the host of any one of claims
350 78-81, said composition having reduced levels of stevia plant-derived contaminants
351 relative to a stevia extract.

352

353 86. A recombinant host, said host comprising:

354 (i) a recombinant gene encoding a UGT91D2;

355 (ii) a recombinant gene encoding a UGT74G1;

356 (iii) a recombinant gene encoding a UGT85C2; and
357 (iv) a recombinant gene encoding a UGT76G1; and
358 (v) a gene encoding a rhamnose synthetase,
359 wherein said host produces at least one steviol glycoside when cultured under
360 conditions in which each of said genes is expressed.

361

362 87. The host of claim 86, further comprising
363 (a) a gene encoding a bifunctional copalyl diphosphate synthase and kaurene
364 synthase, or a gene encoding a copalyl diphosphate synthase and a gene encoding a
365 kaurene synthase;
366 (b) a gene encoding a kaurene oxidase;
367 (c) a gene encoding a steviol synthetase; and
368 (d) a gene encoding a geranylgeranyl diphosphate synthase.

369

370 88. The host of claim 86 or 87, wherein said steviol glycoside is rebaudioside
371 C or dulcoside A.

372

373 89. A steviol glycoside composition produced by the host of claim 86 or 87,
374 said composition having greater than 15% rebaudioside C by weight total steviol
375 glycosides and a reduced level of stevia plant-derived contaminants relative to a stevia
376 extract.

377

378 90. A steviol glycoside composition produced by the host of claim 86 or 87,
379 said composition having greater than 15% dulcoside A by weight total steviol glycosides
380 and a reduced level of stevia plant-derived contaminants relative to a stevia extract.

381

382 91. A recombinant host, said host comprising:
383 (i) a recombinant gene encoding a UGT91D2;
384 (ii) a recombinant gene encoding a UGT74G1;
385 (iii) a recombinant gene encoding a UGT85C2; and

386 (iv) a recombinant gene encoding a UGT76G1;
387 wherein said host produces at least one steviol glycoside when cultured under
388 conditions in which each of said genes is expressed.

389

390 92. The host of claim 91, further comprising

391 (a) a gene encoding a bifunctional copalyl diphosphate synthase and kaurene
392 synthase, or a gene encoding a copalyl diphosphate synthase and a gene encoding a
393 kaurene synthase;

394 (b) a gene encoding a kaurene oxidase;

395 (c) a gene encoding a steviol synthetase; and

396 (d) a gene encoding a geranylgeranyl diphosphate synthase.

397

398 93. The host of claim 91 or 92, wherein said steviol glycoside is rebaudioside
399 D or rebaudioside E.

400

401 94. A steviol glycoside composition produced by the host of claim 91 or 92,
402 said composition having greater than 4% rebaudioside D by weight total steviol
403 glycosides and a reduced level of stevia plant-derived contaminants relative to a stevia
404 extract.

405

406 95. A steviol glycoside composition produced by the host of claim 91 or 92,
407 said composition having greater than 4% rebaudioside E by weight total steviol
408 glycosides and a reduced level of stevia plant-derived contaminants relative to a stevia
409 extract.

410

411 96. A recombinant host, said host comprising:

412 (i) a recombinant gene encoding a UGT91D2;

413 (ii) a recombinant gene encoding a UGT74G1;

414 (iii) a recombinant gene encoding a UGT85C2;

415 (iv) a recombinant gene encoding a UGT76G1;

416 (v) a gene encoding a UDP-glucose dehydrogenase; and
417 (vi) a gene encoding a UDP-glucuronic acid decarboxylase,
418 wherein said host produces at least one steviol glycoside when cultured under
419 conditions in which each of said genes is expressed.

420
421 97. The host of claim 96, further comprising

422 (a) a gene encoding a bifunctional copalyl diphosphate synthase and kaurene
423 synthase, or a gene encoding a copalyl diphosphate synthase and a gene encoding a
424 kaurene synthase;

425 (b) a gene encoding a kaurene oxidase;

426 (c) a gene encoding a steviol synthetase; and

427 (d) a gene encoding a geranylgeranyl diphosphate synthase.

428

429 98. The microorganism of claim 96 or 97, wherein said steviol glycoside is
430 rebaudioside F.

431

432 99. A steviol glycoside composition produced by the host of claim 96 or 97,
433 said composition having greater than 4% rebaudioside F by weight total steviol
434 glycosides and a reduced level of stevia plant-derived contaminants relative to a stevia
435 extract.

436

437 100. A method of producing a steviol glycoside composition, comprising:

438 a) growing the host of any one of claims 86-88, 91-93, or 96-98 in a culture
439 medium, under conditions in which each of said genes is expressed; and

440 b) recovering the steviol glycoside composition produced by said host, wherein
441 said composition is enriched for rebaudioside A, rebaudioside C, rebaudioside D,
442 rebaudioside E, rebaudioside F or dulcoside A relative to the steviol glycoside
443 composition of a wild-type Stevia plant.

444

445 101. The method of claim 100, wherein said steviol glycoside composition
446 produced by said microorganism has a reduced level of stevia plant-derived contaminants
447 relative to a stevia extract.

448
449 102. A food product comprising a steviol glycoside composition enriched for
450 rebaudioside A, rebaudioside C, rebaudioside D, rebaudioside E, rebaudioside F or
451 dulcoside A relative to the steviol glycoside composition of a wild-type Stevia plant.

452
453 103. A method of identifying whether a polymorphism is associated with
454 variation in a trait, said method comprising:

455 a) determining whether one or more genetic polymorphisms in a population of
456 plants is associated with the locus for a polypeptide set forth in SEQ ID NO:5 and
457 functional homologs thereof; and

458 b) measuring the correlation between variation in said trait in plants of said
459 population and the presence of said one or more genetic polymorphisms in plants of said
460 population, thereby identifying whether or not said one or more genetic polymorphisms
461 are associated with variation in said trait.

462
463 104. A method of making a plant line, said method comprising:

464 a) determining whether one or more genetic polymorphisms in a population of
465 plants is associated with the locus for a polypeptide set forth in SEQ ID NO:5 and
466 functional homologs thereof;

467 b) identifying one or more plants in said population in which the presence of at
468 least one of said genetic polymorphisms is associated with variation in a trait;

469 c) crossing one or more of said identified plants with itself or a different plant to
470 produce seed;

471 d) crossing at least one progeny plant grown from said seed with itself or a
472 different plant; and

473 e) repeating steps c) and d) for an additional 0-5 generations to make said plant
474 line, wherein at least one of said genetic polymorphisms is present in said plant line.

475

476 105. A method for transferring a second sugar moiety to the C-2' of a glucose
477 in a steviol glycoside, said method comprising contacting said steviol glycoside with a
478 UGT91D2 polypeptide and a UDP-sugar under suitable reaction conditions for the
479 transfer of said second sugar moiety to said steviol glycoside.

480

481 106. The method of claim 105, wherein said UGT91D2 polypeptide has at least
482 90% identity to the amino acid sequence set forth in SEQ ID NO:5.

483

484 107. The method of claim 106, wherein said UGT91D2 polypeptide has at least
485 95% identity to the amino acid sequence set forth in SEQ ID NO:5.

486

487 108. The method of claim 107, wherein said UGT91D2 polypeptide has at least
488 99% identity to the amino acid sequence set forth in SEQ ID NO:5.

489

490 109. The method of any one of claims 105-108, wherein said polypeptide
491 comprises at least one amino acid substitution at residues 1-19, 27-38, 44-87, 96-120,
492 125-141, 159-184, 199-202, 215-380, or 387-473 of SEQ ID NO:5.

493

494 110. The method of any one of claims 105-109, wherein said polypeptide
495 comprises an amino acid substitution at one or more residues selected from the group
496 consisting of residues 30, 93, 99, 122, 140, 142, 148, 153, 156, 195, 196, 199, 206, 207,
497 211, 221, 286, 343, 427, and 438 of SEQ ID NO:5.

498

499 111. The method of any one of claims 105-110, wherein said steviol glycoside
500 is selected from the group consisting of steviol-13-*O*-glucoside, rubusoside, stevioside,
501 and Rebaudioside A.

502

503 112. The method of claim 111, wherein said steviol glycoside is rubusoside,
504 wherein said second sugar moiety is glucose, and stevioside is produced upon transfer of
505 said second glucose moiety.

506

507 113. The method of claim 111, wherein said steviol glycoside is stevioside,
508 wherein said second sugar moiety is glucose, and Rebaudioside E is produced upon
509 transfer of said second glucose moiety.

510

511 114. The method of claim 111, wherein said steviol glycoside is stevioside,
512 wherein stevioside is contacted with said UGT91D2 polypeptide and a UGT76G1
513 polypeptide under said suitable reaction conditions to produce Rebaudioside D.

514

515 115. The method of claim 111, wherein said steviol glycoside is steviol-13-*O*-
516 glucoside and steviol-1,2 bioside is produced upon transfer of said second glucose
517 moiety.

518

519 116. The method of claim 111, wherein said steviol glycoside is steviol-13-*O*-
520 glucoside and steviol-1,2-xylobioside is produced upon transfer of said second sugar
521 moiety.

522

523 117. The method of claim 111, wherein said steviol glycoside is steviol-13-*O*-
524 glucoside and steviol-1,2-rhamnobioside is produced upon transfer of said second sugar
525 moiety.

526

527 118. The method of claim 111, wherein said steviol glycoside is Rebaudioside
528 A, and Rebaudioside D is produced upon transfer of said second glucose moiety.

529

530 119. A method of determining the presence of a polynucleotide in a *Stevia*
531 plant, comprising:

532 a) contacting at least one probe or primer pair with nucleic acid from said stevia
533 plant, wherein said probe or primer pair is specific for a polynucleotide that encodes a
534 UGT polypeptide, wherein said UGT polypeptide has at least 90% sequence identity to
535 SEQ ID NO: 5, SEQ ID NO: 1, SEQ ID NO: 3 or SEQ ID NO:7 and,
536 b) determining whether or not said polynucleotide is present in said *Stevia* plant.
537

538 120. A kit for genotyping a *Stevia* biological sample, said kit comprising a
539 primer pair that specifically amplifies, or a probe that specifically hybridizes to, a
540 polynucleotide that encodes a UGT polypeptide having at least 90% sequence identity to
541 SEQ ID NO: 5, SEQ ID NO: 1, SEQ ID NO: 3 or SEQ ID NO:7.

Page 1
SEQUENCE LISTING

<110> EVOLVA SA
ABUNDA NUTRITION, INC.

<120> RECOMBINANT PRODUCTION OF STEVIOL GLYCOSIDES

<130> 25903-0003w01

<140>

<141>

<150> 61/471,622

<151> 2011-04-04

<150> 61/434,582

<151> 2011-01-20

<150> 61/350,553

<151> 2010-06-02

<160> 151

<170> PatentIn version 3.5

<210> 1

<211> 460

<212> PRT

<213> Stevia rebaudiana

<400> 1

Met Ala Glu Gln Gln Lys Ile Lys Lys Ser Pro His Val Leu Leu Ile
1 5 10 15

Pro Phe Pro Leu Gln Gly His Ile Asn Pro Phe Ile Gln Phe Gly Lys
20 25 30

Arg Leu Ile Ser Lys Gly Val Lys Thr Thr Leu Val Thr Thr Ile His
35 40 45

Thr Leu Asn Ser Thr Leu Asn His Ser Asn Thr Thr Thr Thr Ser Ile
50 55 60

Glu Ile Gln Ala Ile Ser Asp Gly Cys Asp Glu Gly Gly Phe Met Ser
65 70 75 80

Ala Gly Glu Ser Tyr Leu Glu Thr Phe Lys Gln Val Gly Ser Lys Ser
85 90 95

Leu Ala Asp Leu Ile Lys Lys Leu Gln Ser Glu Gly Thr Thr Ile Asp
100 105 110

Ala Ile Ile Tyr Asp Ser Met Thr Glu Trp Val Leu Asp Val Ala Ile
115 120 125

Glu Phe Gly Ile Asp Gly Gly Ser Phe Phe Thr Gln Ala Cys Val Val
130 135 140

Asn Ser Leu Tyr Tyr His Val His Lys Gly Leu Ile Ser Leu Pro Leu
145 150 155 160

Gly Glu Thr Val Ser Val Pro Gly Phe Pro Val Leu Gln Arg Trp Glu
 165 170 175
 Thr Pro Leu Ile Leu Gln Asn His Glu Gln Ile Gln Ser Pro Trp Ser
 180 185 190
 Gln Met Leu Phe Gly Gln Phe Ala Asn Ile Asp Gln Ala Arg Trp Val
 195 200 205
 Phe Thr Asn Ser Phe Tyr Lys Leu Glu Glu Glu Val Ile Glu Trp Thr
 210 215 220
 Arg Lys Ile Trp Asn Leu Lys Val Ile Gly Pro Thr Leu Pro Ser Met
 225 230 235 240
 Tyr Leu Asp Lys Arg Leu Asp Asp Asp Lys Asp Asn Gly Phe Asn Leu
 245 250 255
 Tyr Lys Ala Asn His His Glu Cys Met Asn Trp Leu Asp Asp Lys Pro
 260 265 270
 Lys Glu Ser Val Val Tyr Val Ala Phe Gly Ser Leu Val Lys His Gly
 275 280 285
 Pro Glu Gln Val Glu Glu Ile Thr Arg Ala Leu Ile Asp Ser Asp Val
 290 295 300
 Asn Phe Leu Trp Val Ile Lys His Lys Glu Glu Gly Lys Leu Pro Glu
 305 310 315 320
 Asn Leu Ser Glu Val Ile Lys Thr Gly Lys Gly Leu Ile Val Ala Trp
 325 330 335
 Cys Lys Gln Leu Asp Val Leu Ala His Glu Ser Val Gly Cys Phe Val
 340 345 350
 Thr His Cys Gly Phe Asn Ser Thr Leu Glu Ala Ile Ser Leu Gly Val
 355 360 365
 Pro Val Val Ala Met Pro Gln Phe Ser Asp Gln Thr Thr Asn Ala Lys
 370 375 380
 Leu Leu Asp Glu Ile Leu Gly Val Gly Val Arg Val Lys Ala Asp Glu
 385 390 395 400
 Asn Gly Ile Val Arg Arg Gly Asn Leu Ala Ser Cys Ile Lys Met Ile
 405 410 415
 Met Glu Glu Glu Arg Gly Val Ile Ile Arg Lys Asn Ala Val Lys Trp
 420 425 430

Lys Asp Leu Ala Lys Val Ala Val His Glu Gly Gly Ser Ser Asp Asn
 435 440 445

Asp Ile Val Glu Phe Val Ser Glu Leu Ile Lys Ala
 450 455 460

<210> 2
 <211> 1383
 <212> DNA
 <213> Stevia rebaudiana

<400> 2
 atggcagagc aacaaaagat caaaaagtca cctcacgtct tacttattcc atttcctctg 60
 caaggacata tcaaccatt catacaattt gggaaaagat tgattagtaa ggggtgtaaag 120
 acaacactgg taaccactat ccacactttg aattctactc tgaaccactc aaatactact 180
 actacaagta tagaaattca agctatatca gacggatgag atgaggggtg ctttatgtct 240
 gccggtgaat cttacttggg 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 ccattgatc 540
 ttacaaaatc atgaacaaat acaatcacct tgggtcccaga tgttgttttg tcaattcgct 600
 aacatcgatc aagcaagatg ggtctttact aattcattct ataagttaga ggaagaggta 660
 attgaatgga ctagggaagat ctggaatttg aaagtcattg gtccaacatt gccatcaatg 720
 tatttgagca aaagacttga tgatgataaa gataatggtt tcaatttgta caaggcta 780
 catcacgaat gtatgaattg gctggatgac aaaccaaagg aatcagttgt atatgttgct 840
 ttcggctctc ttgttaaaca tgggtccagaa caagttgagg agattacaag agcacttata 900
 gactctgacg taaacttttt gtgggtcatt aagcacaaag aggaggggaa actgccagaa 960
 aacctttctg aagtgataaa gaccggaaaa ggtctaactg ttgcttggtg taaacaattg 1020
 gatgttttag ctcatgaatc tgtaggctgt tttgtaacac attgcggatt caactctaca 1080
 ctagaagcca tttccttagg cgtacctgtc gttgcaatgc ctgagttctc cgatcagaca 1140
 accaacgcta aacttttggg 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

<210> 3
 <211> 481
 <212> PRT
 <213> Stevia rebaudiana

<400> 3

Met Asp Ala Met Ala Thr Thr Glu Lys Lys Pro His Val Ile Phe Ile
1 5 10 15

Pro Phe Pro Ala Gln Ser His Ile Lys Ala Met Leu Lys Leu Ala Gln
20 25 30

Leu Leu His His Lys Gly Leu Gln Ile Thr Phe Val Asn Thr Asp Phe
35 40 45

Ile His Asn Gln Phe Leu Glu Ser Ser Gly Pro His Cys Leu Asp Gly
50 55 60

Ala Pro Gly Phe Arg Phe Glu Thr Ile Pro Asp Gly Val Ser His Ser
65 70 75 80

Pro Glu Ala Ser Ile Pro Ile Arg Glu Ser Leu Leu Arg Ser Ile Glu
85 90 95

Thr Asn Phe Leu Asp Arg Phe Ile Asp Leu Val Thr Lys Leu Pro Asp
100 105 110

Pro Pro Thr Cys Ile Ile Ser Asp Gly Phe Leu Ser Val Phe Thr Ile
115 120 125

Asp Ala Ala Lys Lys Leu Gly Ile Pro Val Met Met Tyr Trp Thr Leu
130 135 140

Ala Ala Cys Gly Phe Met Gly Phe Tyr His Ile His Ser Leu Ile Glu
145 150 155 160

Lys Gly Phe Ala Pro Leu Lys Asp Ala Ser Tyr Leu Thr Asn Gly Tyr
165 170 175

Leu Asp Thr Val Ile Asp Trp Val Pro Gly Met Glu Gly Ile Arg Leu
180 185 190

Lys Asp Phe Pro Leu Asp Trp Ser Thr Asp Leu Asn Asp Lys Val Leu
195 200 205

Met Phe Thr Thr Glu Ala Pro Gln Arg Ser His Lys Val Ser His His
210 215 220

Ile Phe His Thr Phe Asp Glu Leu Glu Pro Ser Ile Ile Lys Thr Leu
225 230 235 240

Ser Leu Arg Tyr Asn His Ile Tyr Thr Ile Gly Pro Leu Gln Leu Leu
245 250 255

Leu Asp Gln Ile Pro Glu Glu Lys Lys Gln Thr Gly Ile Thr Ser Leu
260 265 270

His Gly Tyr Ser Leu Val Lys Glu Glu Pro Glu Cys Phe Gln Trp Leu
275 280 285

Gln Ser Lys Glu Pro Asn Ser Val Val Tyr Val Asn Phe Gly Ser Thr
290 295 300

Thr Val Met Ser Leu Glu Asp Met Thr Glu Phe Gly Trp Gly Leu Ala
305 310 315 320

Asn Ser Asn His Tyr Phe Leu Trp Ile Ile Arg Ser Asn Leu Val Ile
325 330 335

Gly Glu Asn Ala Val Leu Pro Pro Glu Leu Glu Glu His Ile Lys Lys
340 345 350

Arg Gly Phe Ile Ala Ser Trp Cys Ser Gln Glu Lys Val Leu Lys His
355 360 365

Pro Ser Val Gly Gly Phe Leu Thr His Cys Gly Trp Gly Ser Thr Ile
370 375 380

Glu Ser Leu Ser Ala Gly Val Pro Met Ile Cys Trp Pro Tyr Ser Trp
385 390 395 400

Asp Gln Leu Thr Asn Cys Arg Tyr Ile Cys Lys Glu Trp Glu Val Gly
405 410 415

Leu Glu Met Gly Thr Lys Val Lys Arg Asp Glu Val Lys Arg Leu Val
420 425 430

Gln Glu Leu Met Gly Glu Gly Gly His Lys Met Arg Asn Lys Ala Lys
435 440 445

Asp Trp Lys Glu Lys Ala Arg Ile Ala Ile Ala Pro Asn Gly Ser Ser
450 455 460

Ser Leu Asn Ile Asp Lys Met Val Lys Glu Ile Thr Val Leu Ala Arg
465 470 475 480

Asn

<210> 4

<211> 1446

<212> DNA

<213> Stevia rebaudiana

<400> 4

atggatgcaa tggcaactac tgagaaaaag cctcatgtga tcttcattcc atttcctgca 60

caatctcaca taaaggcaat gctaaagtta gcacaactat tacaccataa gggattacag 120

ataactttcg tgaataccga cttcatccat aatcaatttc tggaatctag tggccctcat 180

tgtttgacg gagccccagg gtttagattc gaaacaattc ctgacggtgt ttcacattcc 240


```

ccagaggcct ccatcccaat aagagagagt ttactgaggt caatagaaac caactttttg      300
gatcgtttca ttgacttggt cacaaaactt ccagaccac caacttgc atctctgat      360
ggctttctgt cagtgtttac tatcgacgct gccaaaaagt tgggtatccc agttatgat      420
tactggactc ttgctgcatg cggtttcatg ggtttctatc acatccattc tcttatcgaa      480
aagggttttg ctccactgaa agatgcatca tacttaacca acggctacct ggatactggt      540
attgactggg taccaggtat ggaaggata 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 tggattaca tccttacacg gctactcttt agtgaaagag      840
gaaccagaat gttttcaatg gctacaaagt aaagagccta attctgtggt ctacgtcaac      900
ttcggaahta caacagtc atgcttggaa gatatgactg aatttggttg gggccttgct      960
aattcaaata attactttct atggattatc aggtccaatt tggtaataagg ggaaaacgcc     1020
gtattacctc cagaattgga ggaacacatc aaaaagagag gtttcattgc ttcctggtgt     1080
tctcaggaaa aggtattgaa acatccttct gttggtggtt tccttactca ttgcggttg     1140
ggctctacaa tcgaatcact aagtgcagga gttccaatga tttgttggcc atattcatgg     1200
gaccaactta caaattgtag gtatatctgt aaagagtggg aagtggatt agaaatggga     1260
acaaaggtta aacgtgatga agtgaaaaga ttggttcagg agttgatggg ggaagggtggc     1320
cacaagatga gaaacaaggc caaagattgg aaggaaaaag ccagaattgc tattgctcct     1380
aacgggtcat cctctctaaa cattgataag atggtcaaag agattacagt cttagccaga     1440
aactaa

```

<210> 5

<211> 473

<212> PRT

<213> Stevia rebaudiana

<400> 5

```

Met Ala Thr Ser Asp Ser Ile Val Asp Asp Arg Lys Gln Leu His Val
1          5          10          15

```

```

Ala Thr Phe Pro Trp Leu Ala Phe Gly His Ile Leu Pro Tyr Leu Gln
          20          25          30

```

```

Leu Ser Lys Leu Ile Ala Glu Lys Gly His Lys Val Ser Phe Leu Ser
          35          40          45

```

```

Thr Thr Arg Asn Ile Gln Arg Leu Ser Ser His Ile Ser Pro Leu Ile
          50          55          60

```

```

Asn Val Val Gln Leu Thr Leu Pro Arg Val Gln Glu Leu Pro Glu Asp
65          70          75          80

```

Ala Glu Ala Thr Thr Asp Val His Pro Glu Asp Ile Pro Tyr Leu Lys
85 90 95

Lys Ala Ser Asp Gly Leu Gln Pro Glu Val Thr Arg Phe Leu Glu Gln
100 105 110

His Ser Pro Asp Trp Ile Ile Tyr Asp Tyr Thr His Tyr Trp Leu Pro
115 120 125

Ser Ile Ala Ala Ser Leu Gly Ile Ser Arg Ala His Phe Ser Val Thr
130 135 140

Thr Pro Trp Ala Ile Ala Tyr Met Gly Pro Ser Ala Asp Ala Met Ile
145 150 155 160

Asn Gly Ser Asp Gly Arg Thr Thr Val Glu Asp Leu Thr Thr Pro Pro
165 170 175

Lys Trp Phe Pro Phe Pro Thr Lys Val Cys Trp Arg Lys His Asp Leu
180 185 190

Ala Arg Leu Val Pro Tyr Lys Ala Pro Gly Ile Ser Asp Gly Tyr Arg
195 200 205

Met Gly Leu Val Leu Lys Gly Ser Asp Cys Leu Leu Ser Lys Cys Tyr
210 215 220

His Glu Phe Gly Thr Gln Trp Leu Pro Leu Leu Glu Thr Leu His Gln
225 230 235 240

Val Pro Val Val Pro Val Gly Leu Leu Pro Pro Glu Ile Pro Gly Asp
245 250 255

Glu Lys Asp Glu Thr Trp Val Ser Ile Lys Lys Trp Leu Asp Gly Lys
260 265 270

Gln Lys Gly Ser Val Val Tyr Val Ala Leu Gly Ser Glu Val Leu Val
275 280 285

Ser Gln Thr Glu Val Val Glu Leu Ala Leu Gly Leu Glu Leu Ser Gly
290 295 300

Leu Pro Phe Val Trp Ala Tyr Arg Lys Pro Lys Gly Pro Ala Lys Ser
305 310 315 320

Asp Ser Val Glu Leu Pro Asp Gly Phe Val Glu Arg Thr Arg Asp Arg
325 330 335

Gly Leu Val Trp Thr Ser Trp Ala Pro Gln Leu Arg Ile Leu Ser His
340 345 350

Glu Ser Val Cys Gly Phe Leu Thr His Cys Gly Ser Gly Ser Ile Val
355 360 365

Glu Gly Leu Met Phe Gly His Pro Leu Ile Met Leu Pro Ile Phe Gly
370 375 380

Asp Gln Pro Leu Asn Ala Arg Leu Leu Glu Asp Lys Gln Val Gly Ile
385 390 395 400

Glu Ile Pro Arg Asn Glu Glu Asp Gly Cys Leu Thr Lys Glu Ser Val
405 410 415

Ala Arg Ser Leu Arg Ser Val Val Val Glu Lys Glu Gly Glu Ile Tyr
420 425 430

Lys Ala Asn Ala Arg Glu Leu Ser Lys Ile Tyr Asn Asp Thr Lys Val
435 440 445

Glu Lys Glu Tyr Val Ser Gln Phe Val Asp Tyr Leu Glu Lys Asn Ala
450 455 460

Arg Ala Val Ala Ile Asp His Glu Ser
465 470

<210> 6
<211> 1422
<212> DNA
<213> Stevia rebaudiana

<400> 6
atggctacat ctgattctat tgttgatgac aggaagcagt tgcattgtggc tactttccct 60
tggcttgctt tcggtcatat actgccttac ctacaactat caaaactgat agctgaaaaa 120
ggacataaag tgtcattcct ttcaacaact agaaacattc aaagattatc ttccacata 180
tcaccattga ttaacgtcgt tcaattgaca cttccaagag tacaggaatt accagaagat 240
gctgaagcta caacagatgt gcatcctgaa gatattccctt acttgaaaaa ggcattccgat 300
ggattacagc ctgaggtcac tagattcctt gagcaacaca gtccagattg gatcatatac 360
gactacactc actattgggt gccttcaatt gcagcatcac taggcatttc tagggcacat 420
ttcagtgtaa ccacacctg ggccattgct tacatgggtc catccgctga tgctatgatt 480
aacggcagtg atggtagaac taccgttgaa gatttgacaa cccacacaaa gtgggtttcca 540
tttccaacta aagtctgttg gagaaaacac gacttagcaa gactgggttc atacaaggca 600
ccaggaatct cagacggcta tagaatgggt ttagtcctta aagggcttga ctgcctattg 660
tctaagtgtt accatgagtt tgggacacaa tggctaccac ttttgaaaac attacaccaa 720
gttcctgtcg taccagtgg tctattacct ccagaaatcc ctgggtgatga gaaggacgag 780
acttgggttt caatcaaaaa gtgggttagac ggggaagcaaa 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 gtcattccact tatcatgttg 1140
 ccaatctttg gtgaccagcc tttgaatgca cgtctgttag aagataaaca agttggaatt 1200
 gaaatccac 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

<210> 7
 <211> 458
 <212> PRT
 <213> Stevia rebaudiana

<400> 7
 Met Glu Asn Lys Thr Glu Thr Thr Val Arg Arg Arg Arg Arg Ile Ile
 1 5 10 15

Leu Phe Pro Val Pro Phe Gln Gly His Ile Asn Pro Ile Leu Gln Leu
 20 25 30

Ala Asn Val Leu Tyr Ser Lys Gly Phe Ser Ile Thr Ile Phe His Thr
 35 40 45

Asn Phe Asn Lys Pro Lys Thr Ser Asn Tyr Pro His Phe Thr Phe Arg
 50 55 60

Phe Ile Leu Asp Asn Asp Pro Gln Asp Glu Arg Ile Ser Asn Leu Pro
 65 70 75 80

Thr His Gly Pro Leu Ala Gly Met Arg Ile Pro Ile Ile Asn Glu His
 85 90 95

Gly Ala Asp Glu Leu Arg Arg Glu Leu Glu Leu Leu Met Leu Ala Ser
 100 105 110

Glu Glu Asp Glu Glu Val Ser Cys Leu Ile Thr Asp Ala Leu Trp Tyr
 115 120 125

Phe Ala Gln Ser Val Ala Asp Ser Leu Asn Leu Arg Arg Leu Val Leu
 130 135 140

Met Thr Ser Ser Leu Phe Asn Phe His Ala His Val Ser Leu Pro Gln
 145 150 155 160

Phe Asp Glu Leu Gly Tyr Leu Asp Pro Asp Asp Lys Thr Arg Leu Glu
 165 170 175

Glu Gln Ala Ser Gly Phe Pro Met Leu Lys Val Lys Asp Ile Lys Ser

180		185		190
Ala Tyr Ser	Asn Trp Gln Ile	Leu Lys Glu Ile	Leu Gly Lys Met Ile	
195		200	205	
Lys Gln Thr	Lys Ala Ser	Ser Gly Val Ile Trp	Asn Ser Phe Lys Glu	
210		215	220	
Leu Glu Glu	Ser Glu Leu	Glu Thr Val Ile	Arg Glu Ile Pro Ala Pro	
225	230		235	240
Ser Phe Leu	Ile Pro Leu	Pro Lys His	Leu Thr Ala Ser	Ser Ser Ser
	245		250	255
Leu Leu Asp	His Asp Arg Thr	Val Phe Gln Trp	Leu Asp Gln Gln Pro	
	260	265	270	
Pro Ser Ser	Val Leu Tyr Val	Ser Phe Gly Ser Thr	Ser Glu Val Asp	
275		280	285	
Glu Lys Asp	Phe Leu Glu	Ile Ala Arg Gly	Leu Val Asp Ser Lys Gln	
290		295	300	
Ser Phe Leu	Trp Val Val Arg Pro	Gly Phe Val Lys Gly	Ser Thr Trp	
305	310	315	320	
Val Glu Pro	Leu Pro Asp Gly	Phe Leu Gly Glu Arg	Gly Arg Ile Val	
	325	330	335	
Lys Trp Val	Pro Gln Gln Glu	Val Leu Ala His Gly	Ala Ile Gly Ala	
	340	345	350	
Phe Trp Thr	His Ser Gly Trp	Asn Ser Thr Leu	Glu Ser Val Cys Glu	
	355	360	365	
Gly Val Pro	Met Ile Phe Ser	Asp Phe Gly Leu	Asp Gln Pro Leu Asn	
370		375	380	
Ala Arg Tyr	Met Ser Asp Val	Leu Lys Val Gly	Val Tyr Leu Glu Asn	
385	390	395	400	
Gly Trp Glu	Arg Gly Glu Ile	Ala Asn Ala Ile	Arg Arg Val Met Val	
	405	410	415	
Asp Glu Glu	Gly Glu Tyr Ile	Arg Gln Asn Ala	Arg Val Leu Lys Gln	
	420	425	430	
Lys Ala Asp	Val Ser Leu Met	Lys Gly Gly Ser	Ser Tyr Glu Ser Leu	
435		440	445	
Glu Ser Leu	Val Ser Tyr Ile	Ser Ser Leu		
450	455			

<210> 8
 <211> 1377
 <212> DNA
 <213> Stevia rebaudiana

<400> 8
 atggaaaaca agaccgaaac aacagttaga cgtaggcgta gaatcattct gtttccagta 60
 ccttttcaag ggcacatcaa tccaatacta caactagcca acgttttgta ctctaaaggt 120
 ttttctatta caatctttca caccaatttc aacaaaccaa aaacatccaa ttaccacat 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 caatctgttg ctgatagttt gaatttgagg 420
 agattggtac taatgacatc cagtctgttt aactttcacg ctcatgttag tttaccacaa 480
 tttgacgaat tgggatactt ggaccctgat gacaagacta ggtagagga acaggcctct 540
 ggttttccta tgttgaaagt caaagatata aagtctgcct attctaattg gcaaactctg 600
 aaagagatct taggaaagat gatcaaacag acaaaggctt catctggagt gatttggaac 660
 agtttcaag 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 cgtccagggt tcgtgaaagg ctcaacatgg 960
 gtcgaaccac ttccagatgg ttttctaggc gaaagaggta gaatagtcaa atgggttcct 1020
 caacaggaag ttttagctca tggcgctatt ggggcattct ggactcattc cggatggaat 1080
 tcaactttag aatcagtatg cgaaggggta cctatgatct tttcagattt tgggtcttgat 1140
 caaccactga acgcaagata catgtctgat gttttgaaag tgggtgtata tctagaaaat 1200
 ggctgggaaa ggggtgaaat agctaataca ataagacgtg ttatggttga tgaagagggg 1260
 gagtatatca gacaaaacgc aagagtgttg aagcaaaagg ccgacgtttc tctaataaag 1320
 ggaggctctt catacgaatc cttagaatct cttgtttcct acatttcac actgtaa 1377

<210> 9
 <211> 1422
 <212> DNA
 <213> Stevia rebaudiana

<400> 9
 atggctacca gtgactccat agttgacgac cgtaagcagc ttcattgttg gacgttccca 60
 tggcttgctt tcggtcacat cctcccttac cttcagcttt cgaaattgat agctgaaaag 120
 ggtcaciaag tctcgtttct ttctaccacc agaaacattc aacgtctctc ttctcatatc 180
 tcgccactca taaatgttgt tcaactcaca cttccacgtg tccaagagct gccggaggat 240

gcagaggcga ccactgacgt ccaccctgaa gatattccat atctcaagaa ggcttctgat 300
 ggtcttcaac cggagggtcac ccggtttcta gaacaacact ctccggactg gattatttat 360
 gattatactc actactggtt gccatccatc gcggctagcc tcggtatctc acgagcccac 420
 ttctccgtca ccactccatg ggccattgct tatatgggac cctcagctga cgccatgata 480
 aatggttcag atggtcgaac cacggttgag gatctcacga caccgcccac gtggtttccc 540
 tttccgacca aagtatgctg gcggaagcat gatcttggcc gactggtgcc ttacaaagct 600
 ccggggatat ctgatggata ccgtatgggg ctggttctta agggatctga ttgtttgctt 660
 tccaaatggt accatgagtt tggaactcaa tggctacctc ttttggagac actacaccaa 720
 gtaccggtgg ttccggtggg attactgcca ccggaatac ccggagacga gaaagatgaa 780
 acatgggtgt caatcaagaa atggctcgat ggtaaacaac aaggcagtgt ggtgtacgtt 840
 gcattaggaa gcgaggtttt ggtgagccaa accgaggttg ttgagttagc attgggtctc 900
 gagctttctg ggttgccatt tggttgggct tatagaaaac caaaagggtc cgcgaggtca 960
 gactcggtgg agttgccaga cgggttcgtg gaacgaactc gtgaccgtgg gttggtctgg 1020
 acgagttggg cacctcagtt acgaatactg agccatgagt cggtttgggg tttcttgact 1080
 cattgtggtt ctggatcaat tgtggaaggg ctaatgtttg gtcaccctct aatcatgcta 1140
 ccgatttttg gggaccaacc tctgaatgct cgattactgg aggacaaaca ggtgggaatc 1200
 gagataccaa gaaatgagga agatggttgc ttgaccaagg agtcggttgc tagatcactg 1260
 aggtccgttg ttgtggaaaa agaaggggag atctacaagg cgaacgcgag ggagctgagt 1320
 aaaatctata acgacactaa ggttgaaaaa gaatatgtaa gccaattcgt agactatttg 1380
 gaaaagaatg cgcgtgcggt tgccatcgat catgagagtt aa 1422

<210> 10
 <211> 473
 <212> PRT
 <213> Stevia rebaudiana

<400> 10
 Met Ala Thr Ser Asp Ser Ile Val Asp Asp Arg Lys Gln Leu His Val
 1 5 10 15

Ala Thr Phe Pro Trp Leu Ala Phe Gly His Ile Leu Pro Tyr Leu Gln
 20 25 30

Leu Ser Lys Leu Ile Ala Glu Lys Gly His Lys Val Ser Phe Leu Ser
 35 40 45

Thr Thr Arg Asn Ile Gln Arg Leu Ser Ser His Ile Ser Pro Leu Ile
 50 55 60

Asn Val Val Gln Leu Thr Leu Pro Arg Val Gln Glu Leu Pro Glu Asp
 65 70 75 80

Ala Glu Ala Thr Thr Asp Val His Pro Glu Asp Ile Pro Tyr Leu Lys
 85 90 95

Lys Ala Ser Asp Gly Leu Gln Pro Glu Val Thr Arg Phe Leu Glu Gln
 100 105 110
 His Ser Pro Asp Trp Ile Ile Tyr Asp Tyr Thr His Tyr Trp Leu Pro
 115 120 125
 Ser Ile Ala Ala Ser Leu Gly Ile Ser Arg Ala His Phe Ser Val Thr
 130 135 140
 Thr Pro Trp Ala Ile Ala Tyr Met Gly Pro Ser Ala Asp Ala Met Ile
 145 150 155 160
 Asn Gly Ser Asp Gly Arg Thr Thr Val Glu Asp Leu Thr Thr Pro Pro
 165 170 175
 Lys Trp Phe Pro Phe Pro Thr Lys Val Cys Trp Arg Lys His Asp Leu
 180 185 190
 Ala Arg Leu Val Pro Tyr Lys Ala Pro Gly Ile Ser Asp Gly Tyr Arg
 195 200 205
 Met Gly Leu Val Leu Lys Gly Ser Asp Cys Leu Leu Ser Lys Cys Tyr
 210 215 220
 His Glu Phe Gly Thr Gln Trp Leu Pro Leu Leu Glu Thr Leu His Gln
 225 230 235 240
 Val Pro Val Val Pro Val Gly Leu Leu Pro Pro Glu Val Pro Gly Asp
 245 250 255
 Glu Lys Asp Glu Thr Trp Val Ser Ile Lys Lys Trp Leu Asp Gly Lys
 260 265 270
 Gln Lys Gly Ser Val Val Tyr Val Ala Leu Gly Ser Glu Val Leu Val
 275 280 285
 Ser Gln Thr Glu Val Val Glu Leu Ala Leu Gly Leu Glu Leu Ser Gly
 290 295 300
 Leu Pro Phe Val Trp Ala Tyr Arg Lys Pro Lys Gly Pro Ala Lys Ser
 305 310 315 320
 Asp Ser Val Glu Leu Pro Asp Gly Phe Val Glu Arg Thr Arg Asp Arg
 325 330 335
 Gly Leu Val Trp Thr Ser Trp Ala Pro Gln Leu Arg Ile Leu Ser His
 340 345 350
 Glu Ser Val Cys Gly Phe Leu Thr His Cys Gly Ser Gly Ser Ile Val
 355 360 365

Glu Gly Leu Met Phe Gly His Pro Leu Ile Met Leu Pro Ile Phe Gly
 370 375 380

Asp Gln Pro Leu Asn Ala Arg Leu Leu Glu Asp Lys Gln Val Gly Ile
 385 390 395 400

Glu Ile Pro Arg Asn Glu Glu Asp Gly Cys Leu Thr Lys Glu Ser Val
 405 410 415

Ala Arg Ser Leu Arg Ser Val Val Val Glu Lys Glu Gly Glu Ile Tyr
 420 425 430

Lys Ala Asn Ala Arg Glu Leu Ser Lys Ile Tyr Asn Asp Thr Lys Val
 435 440 445

Glu Lys Glu Tyr Val Ser Gln Phe Val Asp Tyr Leu Glu Lys Asn Thr
 450 455 460

Arg Ala Val Ala Ile Asp His Glu Ser
 465 470

<210> 11
 <211> 1422
 <212> DNA
 <213> Stevia rebaudiana

<400> 11
 atggctacca gtgactccat agttgacgac cgtaagcagc ttcattgttg gacgttccca 60
 tggcttgctt tcggtcacat cctcccttac cttcagcttt cgaaattgat agctgaaaag 120
 ggtcacaag tctcgtttct ttctaccacc agaaacattc aacgtctctc ttctcatatc 180
 tcgccactca taaatgttgt tcaactcaca cttccacgtg tccaagagct gccggaggat 240
 gcagaggcga ccaactgacgt ccaccctgaa gatattccat atctcaagaa ggcttctgat 300
 ggtcttcaac cggaggtcac ccggtttcta gaacaacact ctccggactg gattatttat 360
 gattatactc actactgggt gccatccatc gcggctagcc tcggtatctc acgagcccac 420
 ttctccgtca ccaactccatg ggccattgct tatatgggac cctcagctga cgccatgata 480
 aatgggttcag atgggtcgaac cacggttgag gatctcacga caccgcccac gtgggtttccc 540
 tttccgacca aagtatgctg gcggaagcat gatcttgccc gactgggtgcc ttacaaagct 600
 ccggggatat ctgatggata ccgtatgggg ctggttctta agggatctga ttgtttgctt 660
 tccaaatgtt accatgagtt tggaactcaa tggctacctc ttttgagac actacaccaa 720
 gtaccggtgg ttccggtggg attactgcca ccggaagtac ccggagacga gaaagatgaa 780
 acatgggtgt caatcaagaa atggctcgat ggtaaacaaa aaggcagtg ggtgtacgtt 840
 gcattaggaa gcgaggtttt ggtgagccaa accgaggttg ttgagttagc attgggtctc 900
 gagctttctg ggttgccatt tgtttgggct tatagaaaac caaaagggtcc cgcgaagtca 960
 gactcggtag agttgccaga cgggttcgtg gaacgaactc gtgaccgtgg gttggtctgg 1020

acgagttggg cacctcagtt acgaatactg agccatgagt cggtttgtgg tttcttgact 1080
cattgtgggt ctggatcaat tgtggaagg ctaatgtttg gtcaccctct aatcatgcta 1140
ccgatttttg gggaccaacc tctgaatgct cgattactgg aggacaaaca ggtgggaatc 1200
gagataccaa gaaatgagga agatggttgc ttgaccaagg agtcggttgc tagatcactg 1260
aggtccgttg ttgtggaaaa agaaggggag atctacaagg cgaacgcgag ggagctgagt 1320
aaaatctata acgacactaa ggttgaaaaa gaatatgtaa gccaatcgt agactatttg 1380
gaaaagaata cgcgtgcggt tgccatcgat catgagagtt aa 1422

<210> 12
<211> 473
<212> PRT
<213> Stevia rebaudiana

<400> 12
Met Ala Thr Ser Asp Ser Ile Val Asp Asp Arg Lys Gln Leu His Val
1 5 10 15

Ala Thr Phe Pro Trp Leu Ala Phe Gly His Ile Leu Pro Phe Leu Gln
20 25 30

Leu Ser Lys Leu Ile Ala Glu Lys Gly His Lys Val Ser Phe Leu Ser
35 40 45

Thr Thr Arg Asn Ile Gln Arg Leu Ser Ser His Ile Ser Pro Leu Ile
50 55 60

Asn Val Val Gln Leu Thr Leu Pro Arg Val Gln Glu Leu Pro Glu Asp
65 70 75 80

Ala Glu Ala Thr Thr Asp Val His Pro Glu Asp Ile Gln Tyr Leu Lys
85 90 95

Lys Ala Val Asp Gly Leu Gln Pro Glu Val Thr Arg Phe Leu Glu Gln
100 105 110

His Ser Pro Asp Trp Ile Ile Tyr Asp Phe Thr His Tyr Trp Leu Pro
115 120 125

Ser Ile Ala Ala Ser Leu Gly Ile Ser Arg Ala Tyr Phe Cys Val Ile
130 135 140

Thr Pro Trp Thr Ile Ala Tyr Leu Ala Pro Ser Ser Asp Ala Met Ile
145 150 155 160

Asn Asp Ser Asp Gly Arg Thr Thr Val Glu Asp Leu Thr Thr Pro Pro
165 170 175

Lys Trp Phe Pro Phe Pro Thr Lys Val Cys Trp Arg Lys His Asp Leu
180 185 190

Ala Arg Met Glu Pro Tyr Glu Ala Pro Gly Ile Ser Asp Gly Tyr Arg
 195 200 205
 Met Gly Met Val Phe Lys Gly Ser Asp Cys Leu Leu Phe Lys Cys Tyr
 210 215 220
 His Glu Phe Gly Thr Gln Trp Leu Pro Leu Leu Glu Thr Leu His Gln
 225 230 235 240
 Val Pro Val Val Pro Val Gly Leu Leu Pro Pro Glu Ile Pro Gly Asp
 245 250 255
 Glu Lys Asp Glu Thr Trp Val Ser Ile Lys Lys Trp Leu Asp Gly Lys
 260 265 270
 Gln Lys Gly Ser Val Val Tyr Val Ala Leu Gly Ser Glu Ala Leu Val
 275 280 285
 Ser Gln Thr Glu Val Val Glu Leu Ala Leu Gly Leu Glu Leu Ser Gly
 290 295 300
 Leu Pro Phe Val Trp Ala Tyr Arg Lys Pro Lys Gly Pro Ala Lys Ser
 305 310 315 320
 Asp Ser Val Glu Leu Pro Asp Gly Phe Val Glu Arg Thr Arg Asp Arg
 325 330 335
 Gly Leu Val Trp Thr Ser Trp Ala Pro Gln Leu Arg Ile Leu Ser His
 340 345 350
 Glu Ser Val Cys Gly Phe Leu Thr His Cys Gly Ser Gly Ser Ile Val
 355 360 365
 Glu Gly Leu Met Phe Gly His Pro Leu Ile Met Leu Pro Leu Phe Gly
 370 375 380
 Asp Gln Pro Leu Asn Ala Arg Leu Leu Glu Asp Lys Gln Val Gly Ile
 385 390 395 400
 Glu Ile Pro Arg Asn Glu Glu Asp Gly Cys Leu Thr Lys Glu Ser Val
 405 410 415
 Ala Arg Ser Leu Arg Ser Val Val Val Glu Asn Glu Gly Glu Ile Tyr
 420 425 430
 Lys Ala Asn Ala Arg Glu Leu Ser Lys Ile Tyr Asn Asp Thr Lys Val
 435 440 445
 Glu Lys Glu Tyr Val Ser Gln Phe Val Asp Tyr Leu Glu Lys Asn Ala
 450 455 460
 Arg Ala Val Ala Ile Asp His Glu Ser

465

470

<210> 13
 <211> 1422
 <212> DNA
 <213> Stevia rebaudiana

<400> 13
 atggctacca gtgactccat agttgacgac cgtaagcagc ttcattgttg gacgttccca 60
 tggcttgctt tcggtcacat cctccctttc cttcagcttt cgaaattgat agctgaaaag 120
 ggtcaciaaag tctcgtttct ttctaccacc agaaacattc aacgtctctc ttctcatatc 180
 tcgccactca taaatgttgt tcaactcaca cttccacgtg tccaagagct gccggaggat 240
 gcagaggcga ccactgacgt ccaccctgaa gatattcaat atctcaagaa ggctgttgat 300
 ggtcttcaac cggagggtcac ccggtttcta gaacaacact ctccggactg gattatttat 360
 gattttactc actactgggt gccatccatc gcggctagcc tcggtatctc acgagcctac 420
 ttctgcgtca tcaactccatg gaccattgct tttttggcac cctcatctga cgccatgata 480
 aatgattcag atggtcgaac cacggttgag gatctcacga caccgccaa gtggtttccc 540
 tttccgacca aagtatgctg gcggaagcat gatcttgccc gaatggagcc ttacgaagct 600
 ccagggatat ctgatggata ccgtatgggg atggttttta agggatctga ttgtttgctt 660
 ttcaaagtgt accatgagtt tggaactcaa tggctacctc ttttgagac actacaccaa 720
 gtaccggtgg ttccggtggg attactgccg ccggaatac ccggagacga gaaagatgaa 780
 acatgggtgt caatcaagaa atggctcgat ggtaaacaaa aaggcagtgt ggtgtacgtt 840
 gcattaggaa gcgaggcttt ggtgagccaa accgaggttg ttgagttagc attgggtctc 900
 gagctttctg ggttgccatt tgtttgggct tatagaaaac caaaagggtc gcggaagtca 960
 gactcggtag agttgccaga cgggttcgtg gaacgaactc gtgaccgtgg gttggtctgg 1020
 acgagttggg cacctcagtt acgaatactg agccatgagt cggtttgtag tttcttgact 1080
 cattgtggtt ctggatcaat tgtggaaggg ctaatgtttg gtcaccctct aatcatgcta 1140
 ccgctttttg gggaccaacc tctgaatgct cgattactgg aggacaaaca ggtgggaatc 1200
 gagataccea gaaatgagga agatggttgc ttgaccaagg agtcggttgc tagatcactg 1260
 aggtccgttg ttgtggaaaa cgaaggggag atctacaagg cgaacgcgag ggagctgagt 1320
 aaaatctata acgacactaa ggtggaaaaa gaatatgtaa gccaattcgt agactatttg 1380
 gaaaagaatg cgcgtgagggt tgccatcgat catgagagtt aa 1422

<210> 14
 <211> 485
 <212> PRT
 <213> Stevia rebaudiana

<400> 14
 Met Tyr Asn Val Thr Tyr His Gln Asn Ser Lys Ala Met Ala Thr Ser
 1 5 10 15

Asp Ser Ile Val Asp Asp Arg Lys Gln Leu His Val Ala Thr Phe Pro

Trp	Leu	Ala	Phe	Gly	His	Ile	Leu	Pro	Phe	Leu	Gln	Leu	Ser	Lys	Leu
	35						40					45			
Ile	Ala	Glu	Lys	Gly	His	Lys	Val	Ser	Phe	Leu	Ser	Thr	Thr	Arg	Asn
	50					55					60				
Ile	Gln	Arg	Leu	Ser	Ser	His	Ile	Ser	Pro	Leu	Ile	Asn	Val	Val	Gln
	65				70					75					80
Leu	Thr	Leu	Pro	Arg	Val	Gln	Glu	Leu	Pro	Glu	Asp	Ala	Glu	Ala	Thr
				85					90					95	
Thr	Asp	Val	His	Pro	Glu	Asp	Ile	Gln	Tyr	Leu	Lys	Lys	Ala	Val	Asp
			100					105					110		
Gly	Leu	Gln	Pro	Glu	Val	Thr	Arg	Phe	Leu	Glu	Gln	His	Ser	Pro	Asp
		115					120					125			
Trp	Ile	Ile	Tyr	Asp	Phe	Thr	His	Tyr	Trp	Leu	Pro	Ser	Ile	Ala	Ala
	130					135					140				
Ser	Leu	Gly	Ile	Ser	Arg	Ala	Tyr	Phe	Cys	Val	Ile	Thr	Pro	Trp	Thr
	145				150					155					160
Ile	Ala	Tyr	Leu	Ala	Pro	Ser	Ser	Asp	Ala	Met	Ile	Asn	Asp	Ser	Asp
				165					170					175	
Gly	Arg	Thr	Thr	Val	Glu	Asp	Leu	Thr	Thr	Pro	Pro	Lys	Trp	Phe	Pro
			180					185					190		
Phe	Pro	Thr	Lys	Val	Cys	Trp	Arg	Lys	His	Asp	Leu	Ala	Arg	Met	Glu
		195					200					205			
Pro	Tyr	Glu	Ala	Pro	Gly	Ile	Ser	Asp	Gly	Tyr	Arg	Met	Gly	Met	Val
	210					215					220				
Phe	Lys	Gly	Ser	Asp	Cys	Leu	Leu	Phe	Lys	Cys	Tyr	His	Glu	Phe	Gly
	225				230					235					240
Thr	Gln	Trp	Leu	Pro	Leu	Leu	Glu	Thr	Leu	His	Gln	Val	Pro	Val	Val
				245					250					255	
Pro	Val	Gly	Leu	Leu	Pro	Pro	Glu	Ile	Pro	Gly	Asp	Glu	Lys	Asp	Glu
			260					265					270		
Thr	Trp	Val	Ser	Ile	Lys	Lys	Trp	Leu	Asp	Gly	Lys	Gln	Lys	Gly	Ser
		275					280					285			
Val	Val	Tyr	Val	Ala	Leu	Gly	Ser	Glu	Ala	Leu	Val	Ser	Gln	Thr	Glu
	290					295					300				

Val Val Glu Leu Ala Leu Gly Leu Glu Leu Ser Gly Leu Pro Phe Val
305 310 315 320

Trp Ala Tyr Arg Lys Pro Lys Gly Pro Ala Lys Ser Asp Ser Val Glu
325 330 335

Leu Pro Asp Gly Phe Val Glu Arg Thr Arg Asp Arg Gly Leu Val Trp
340 345 350

Thr Ser Trp Ala Pro Gln Leu Arg Ile Leu Ser His Glu Ser Val Cys
355 360 365

Gly Phe Leu Thr His Cys Gly Ser Gly Ser Ile Val Glu Gly Leu Met
370 375 380

Phe Gly His Pro Leu Ile Met Leu Pro Ile Phe Cys Asp Gln Pro Leu
385 390 395 400

Asn Ala Arg Leu Leu Glu Asp Lys Gln Val Gly Ile Glu Ile Pro Arg
405 410 415

Asn Glu Glu Asp Gly Cys Leu Thr Lys Glu Ser Val Ala Arg Ser Leu
420 425 430

Arg Ser Val Val Val Glu Asn Glu Gly Glu Ile Tyr Lys Ala Asn Ala
435 440 445

Arg Ala Leu Ser Lys Ile Tyr Asn Asp Thr Lys Val Glu Lys Glu Tyr
450 455 460

Val Ser Gln Phe Val Asp Tyr Leu Glu Lys Asn Ala Arg Ala Val Ala
465 470 475 480

Ile Asp His Glu Ser
485

<210> 15

<211> 1458

<212> DNA

<213> Stevia rebaudiana

<400> 15

atgtacaacg ttacttatca tcaaaattca aaagcaatgg ctaccagtga ctccatagtt	60
gacgaccgta agcagcttca tgttgcgacg ttcccatggc ttgctttcgg tcacatcctc	120
cctttccttc agctttcgaa attgatagct gaaaagggtc acaaagtctc gtttctttct	180
accaccagaa acattcaacg tctctcttct catatctcgc cactcataaa tggtgttcaa	240
ctcacacttc cacgtgtcca agagctgccg gaggatgcag aggcgaccac tgacgtccac	300
cctgaagata ttcaatatct caagaaggct gttgatggc ttcaaccgga ggtcaccggg	360
tttctagaac aacactctcc ggactggatt atttatgatt ttactcacta ctggttgcca	420

```

tccatcgcg ctagcctcgg tatctcacga gcctacttct gcgtcatcac tccatggacc 480
attgcttatt tggcaccctc atctgacgcc atgataaatg attcagatgg tcgaaccacg 540
gttgaggatc tcacgacacc gcccaagtgg tttccctttc cgaccaaagt atgctggcgg 600
aagcatgatc ttgcccgaat ggagccttac gaagctccgg ggatatctga tggataccgt 660
atggggatgg tttttaaggg atctgattgt ttgcttttca aatgttacca tgagtttgga 720
actcaatggc tacctctttt ggagacacta caccaagtac cggtggttcc ggtgggatta 780
ctgccgccgg aaatacccg agacgagaaa gatgaaacat ggggtgtcaat caagaaatgg 840
ctcgatggta aacaaaaagg cagtgtggtg tacgttgcac taggaagcga ggctttggtg 900
agccaaaccg aggttggtga gttagcattg ggtctcgagc tttctggggt gccatttggt 960
tgggcttata gaaaaccaa aggtcccgcg aagtcagact cggtgagggt gccagacggg 1020
ttcgtggaac gaactcgtga ccgtggggtg gtctggacga gttgggcacc tcagttacga 1080
atactgagcc acgagtcagt ttgtggtttc ttgactcatt gtggttctgg atcaattgtg 1140
gaagggctaa tgtttggtca cctctaatac atgctaccga ttttttgtga ccaacctctg 1200
aatgctcgat tactggagga caaacagggt ggaatcgaga taccaagaaa tgaggaagat 1260
ggttgcttga ccaaggagtc ggttgctaga tcactgaggt ccgttggtgt ggaaaacgaa 1320
ggggagatct acaaggcgaa cgcgagggcg ctgagtaaaa tctataacga cactaagggt 1380
gaaaaagaat atgtaagcca attcgtagac tttttgaaa agaatgcgcg tgcggttgcc 1440
atcgatcatg agagttaa 1458

```

<210> 16
 <211> 473
 <212> PRT
 <213> Stevia rebaudiana

<400> 16
 Met Ala Thr Ser Asp Ser Ile Val Asp Asp Arg Lys Gln Leu His Val
 1 5 10 15

Ala Thr Phe Pro Trp Leu Ala Phe Gly His Ile Leu Pro Phe Leu Gln
 20 25 30

Leu Ser Lys Leu Ile Ala Glu Lys Gly His Lys Val Ser Phe Leu Ser
 35 40 45

Thr Thr Arg Asn Ile Gln Arg Leu Ser Ser His Ile Ser Pro Leu Ile
 50 55 60

Asn Val Val Gln Leu Thr Leu Pro Arg Val Gln Glu Leu Pro Glu Asp
 65 70 75 80

Ala Glu Ala Thr Thr Asp Val His Pro Glu Asp Ile Gln Tyr Leu Lys
 85 90 95

Lys Ala Val Asp Gly Leu Gln Pro Glu Val Thr Arg Phe Leu Glu Gln

100								105						110
His	Ser	Pro	Asp	Trp	Ile	Ile	Tyr	Asp	Phe	Thr	His	Tyr	Trp	Leu
		115					120					125		Pro
Ser	Ile	Ala	Ala	Ser	Leu	Gly	Ile	Ser	Arg	Ala	Tyr	Phe	Cys	Val
	130					135					140			Ile
Thr	Pro	Trp	Thr	Ile	Ala	Tyr	Leu	Ala	Pro	Ser	Ser	Asp	Ala	Met
145					150					155				160
Asn	Asp	Ser	Asp	Gly	Arg	Thr	Thr	Val	Glu	Asp	Leu	Thr	Thr	Pro
				165					170					175
Lys	Trp	Phe	Pro	Phe	Pro	Thr	Lys	Val	Cys	Trp	Arg	Lys	His	Asp
			180					185					190	Leu
Ala	Arg	Met	Glu	Pro	Tyr	Glu	Ala	Pro	Gly	Ile	Ser	Asp	Gly	Tyr
		195					200					205		Arg
Met	Gly	Met	Val	Phe	Lys	Gly	Ser	Asp	Cys	Leu	Leu	Phe	Lys	Cys
	210					215					220			Tyr
His	Glu	Phe	Gly	Thr	Gln	Trp	Leu	Pro	Leu	Leu	Glu	Thr	Leu	His
225					230					235				240
Val	Pro	Val	Val	Pro	Val	Gly	Leu	Leu	Pro	Pro	Glu	Ile	Pro	Gly
				245					250					255
Glu	Lys	Asp	Glu	Thr	Trp	Val	Ser	Ile	Lys	Lys	Trp	Leu	Asp	Gly
			260					265					270	Lys
Gln	Lys	Gly	Ser	Val	Val	Tyr	Val	Ala	Leu	Gly	Ser	Glu	Ala	Leu
		275					280					285		Val
Ser	Gln	Thr	Glu	Val	Val	Glu	Leu	Ala	Leu	Gly	Leu	Glu	Leu	Ser
	290					295					300			Gly
Leu	Pro	Phe	Val	Trp	Ala	Tyr	Arg	Lys	Pro	Lys	Gly	Pro	Ala	Lys
305					310					315				320
Asp	Ser	Val	Glu	Leu	Pro	Asp	Gly	Phe	Val	Glu	Arg	Thr	Arg	Asp
				325					330					335
Gly	Leu	Val	Trp	Thr	Ser	Trp	Ala	Pro	Gln	Leu	Arg	Ile	Leu	Ser
			340					345					350	His
Glu	Ser	Val	Cys	Gly	Phe	Leu	Thr	His	Cys	Gly	Ser	Gly	Ser	Ile
		355					360					365		Val
Glu	Gly	Leu	Met	Phe	Gly	His	Pro	Leu	Ile	Met	Leu	Pro	Ile	Phe
	370					375					380			Cys

Asp Gln Pro Leu Asn Ala Arg Leu Leu Glu Asp Lys Gln Val Gly Ile
385 390 395 400

Glu Ile Pro Arg Asn Glu Glu Asp Gly Cys Leu Thr Lys Glu Ser Val
405 410 415

Ala Arg Ser Leu Arg Ser Val Val Val Glu Asn Glu Gly Glu Ile Tyr
420 425 430

Lys Ala Asn Ala Arg Ala Leu Ser Lys Ile Tyr Asn Asp Thr Lys Val
435 440 445

Glu Lys Glu Tyr Val Ser Gln Phe Val Asp Tyr Leu Glu Lys Asn Ala
450 455 460

Arg Ala Val Ala Ile Asp His Glu Ser
465 470

<210> 17
<211> 1422
<212> DNA
<213> Stevia rebaudiana

<400> 17
atggctacca gtgactccat agttgacgac cgtaagcagc ttcatgttgc gacgttccca 60
tggtttgctt tcggtcacat cctccctttc cttcagcttt cgaaattgat agctgaaaag 120
ggtcacaaag ttctgtttct ttctaccacc agaaacattc aacgtctctc ttctcatatc 180
tcgccactca taaatgttgt tcaactcaca cttccacgtg tccaagagct gccggaggat 240
gcagaggcga ccaactgacgt ccaccctgaa gatattcaat atctcaagaa ggctgttgat 300
ggctttcaac cggagggtcac ccggtttcta gaacaacact ctccggactg gattatttat 360
gattttactc actactgggt gccatccatc gcggctagcc tcggtatctc acgagcctac 420
ttctgctga tcaactcatg gaccattgct tttttggcac cctcatctga cgccatgata 480
aatgattcag atgggtcgaa cacggttgag gatctcacga caccgccaa gtggtttccc 540
tttccgacca aagtatgctg gcggaagcat gatcttgccc gaatggagcc ttacgaagct 600
ccggggatat ctgatggata ccgatgggg atggttttta agggatctga ttgtttgctt 660
ttcaaagtgt accatgagtt tggaactcaa tggctacctc ttttggagac actacaccaa 720
gtaccggtgg ttccggtggg attactgccg ccggaaatac ccggagacga gaaagatgaa 780
acatgggtgt caatcaagaa atggctcgat ggtaaacaaa aaggcagtgt ggtgtacgtt 840
gcattaggaa gcgaggcttt ggtgagccaa accgagggtg ttgagttagc attgggtctc 900
gagctttctg ggttgccatt tgtttgggct tatagaaaac caaaagggtc cgcaagtc 960
gactcgttg agttgccaga cgggttcgtg gaacgaactc gtgaccgtgg gttggtctgg 1020
acgagttggg cacctcagtt acgaatactg agccacgagt cagtttgtgg tttcttgact 1080
cattgtggtt ctggatcaat tgtggaaggg ctaatgtttg gtcaccctct aatcatgcta 1140

ccgatttttt	gtgaccaacc	tctgaatgct	cgattactgg	aggacaaaca	ggtgggaatc	1200
gagataccaa	gaaatgagga	agatgggtgc	ttgaccaagg	agtcgggtgc	tagatcactg	1260
agggccgttg	ttgtggaaaa	cgaaggggag	atctacaagg	cgaacgcgag	ggcgtgagt	1320
aaaatctata	acgacactaa	ggtggaaaaa	gaatatgtaa	gccaatcgt	agactatttg	1380
gaaaagaatg	cgcgtgcggt	tgccatcgat	catgagagtt	aa		1422

<210> 18
 <211> 1086
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 18	
atggcttttg	taaaccac
cgctcttttc	tatggtacct
ctatcagaac	aagacctaca
	60
aacttactaa	atccaactca
aaagctaaga	ccagtttcat
catcttcctt	accttctttc
	120
tcatcagtta	gtgcgattct
tactgaaaaa	catcaatcta
atccttctga	gaacaacaat
	180
ttgcaaaactc	atctagaaac
tcctttcaac	tttgatagtt
atatgttgga	aaaagtcaac
	240
atgggttaacg	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	aggttttagtg
	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	gatgttacia
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

<210> 19
 <211> 1029
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 19
 atggctgagc aacaaatatt 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 cttttatcga ttccctgaac 360
 gtatggttgg aggttccaga ggatgaaaca tcagtcatca aggaagttat tggtagctc 420
 cacaactctt cattaatcat tgatgacttc caagataatt ctccacttag aagaggaaag 480
 ccatctaccc atacagtctt cggccctgcc caggctatca atactgctac ttacgttata 540
 gttaaagcaa tcgaaaagat acaagacata gtgggacacg atgcattggc agatgttacg 600
 ggtactatta caactatctt ccaaggctcag gccatggact tgtgggtggac agcaaattga 660
 atcgttccat caatacagga atacttactt atggtaaacg ataaaaccgg tgctctcttt 720
 agactgagtt tggagttggt 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
 tggaatga 1029

<210> 20
 <211> 903
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 20
 atggaaaaga ctaaggagaa agcagaacgt atcttgctgg agccatacag atacttatta 60
 caactaccag gaaagcaagt ccgttctaaa ctatcacaag cgttcaatca ctgggttaaaa 120
 gttcctgaag ataagttaca aatcattatt gaagtcacag aaatgctaca caatgcttct 180
 ttactgatcg atgatataga ggattcttcc aaactgagaa gaggttttcc tgctgctcat 240
 tccatatacg gggtagcaag tgtaatcaac tcagctaatt acgtctactt cttgggattg 300
 gaaaaagtat tgacattaga tcatccagac gctgtaaagc tattcaccag acaacttctt 360
 gaattgcatc aagggtcaagg tttggatatc tattggagag acacttatac ttgccaaca 420
 gaagaggagt acaaaagcaat ggttctacaa aagactggcg gtttggttcgg acttgccgtt 480
 ggtctgatgc aacttttctc tgattacaag gaggacttaa agcctctgtt ggataccttg 540
 ggcttggttt tccagattag agatgactac gctaacttac attcaaagga atattcagaa 600

aacaaatcat tctgtgaaga tttgactgaa gggaagttta gttttccaac aatccacgcc	660
atttgggtcaa gaccagaatc tactcaagtg caaaacattc tgcgtcagag aacagagaat	720
attgacatca aaaagtattg tggtcagtag ttggaagatg ttgggttcttt tgcttacaca	780
agacatacac ttagagaatt agaggcaaaa gcatacaagc aaatagaagc ctgtggaggc	840
aatccttctc tagtggcatt ggtaaacaat ttgtccaaaa tggtcaccga ggaaaacaag	900
taa	903

<210> 21
 <211> 1020
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 21	
atggcaagat tctatcttct taacgcacta ttgatgggta tctcattaca atcaactaca	60
gccttcactc cagctaaact tgcttatcca acaacaacaa cagctctaaa tgcgcctcc	120
gccgaaactt ctttcagtct agatgaatac ttggcctcta agataggacc tatagagtct	180
gccttggaag catcagtcaa atccagaatt ccacagaccg ataagatctg cgaatctatg	240
gcctactctt tgatggcagg aggcaagaga attagaccag tggtgtgtat cgctgcatgt	300
gagatgttcg gtggatccca agatgtcgct atgcctactg ctgtggcatt agaaatgata	360
cacacaatgt ctttgattca tgatgatttg ccatccatgg ataacgatga cttgagaaga	420
ggtaaaccacaa caaacatgt cgttttcggc gaagatgtag ctattcttgc aggtgactct	480
ttattgtcaa ctcccttcga 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
gggtgtgcaa ctccctgaaga 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

<210> 22
 <211> 1068
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 22	
atgcacttag caccacgtag agtccttaga ggtagaagat caccacctga cagagttcct	60

gaaagacaag gtgccttggg tagaagacgt ggagctggct ctactggctg tgcccgtgct	120
gctgctggtg ttcaccgtag aagaggagga ggcgaggctg atccatcagc tgctgtgcat	180
agaggctggc aagccggtg tggcaccggt ttgcctgatg aggtggtgtc taccgcagcc	240
gccttagaaa tgtttcatgc ttttgcttta atccatgatg atatcatgga tgatagtga	300
actagaagag gctccccaac tgttcacaga gccctagctg atcgtttagg cgctgctctg	360
gaccagatc aggccggtca actaggagtt tctactgcta tcttggttgg agatctggct	420
ttgacatggt ccgatgaatt gttatacgtc 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
cttagagggtg gaaagcatac tgtcttagtc gccttgcaa gagaacatgc cactccagaa	840
cagagacaca cattggatac attattgggt acaccaggctc ttgatagaca aggcgcttca	900
agactaagat gcgtattggt agcaactggt gcaagagccg aagccgaaag acttattaca	960
gagagaagag atcaagcatt aactgcattg aacgcattaa cactgccacc tccttttagct	1020
gaggcattag caagattgac attagggctc acagctcatc ctgcctaa	1068

<210> 23

<211> 993

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 23

atgtcatatt tcgataacta cttcaatgag atagttaatt ccgtgaacga catcattaag	60
tcttacatct ctggcgacgt accaaaacta tacgaagcct cctaccattt gtttacatca	120
ggaggaaaga gactaagacc attgatcctt acaatttctt ctgatctttt cgggtggacag	180
agagaaagag catactatgc tggcgcagca atcgaagttt tgcacacatt cactttggtt	240
cacgatgata tcatggatca agataacatt cgtagaggctc ttcctactgt acatgtcaag	300
tatggcctac ctttggccat tttagctggt gacttattgc atgcaaaaagc ctttcaattg	360
ttgactcagg cattgagagg tctaccatct gaaactatca tcaaggcggt tgatatcttt	420
acaagatcta tcattatcat atcagaaggt caagctgtcg atatggaatt cgaagataga	480
attgatatca aggaacaaga gtatttggt atgatatctc gtaaaaccgc tgccttattc	540
tcagcttctt cttccattgg ggcgttgata gctggagcta atgataacga tgtgagatta	600
atgtccgatt tcggtacaaa tcttgggac gcatttcaaa ttgtagatga tatacttggt	660
ttaacagctg atgaaaaaga gctaggaaaa cctgttttca gtgatatcag agaaggtaaa	720

aagaccatat tagtcattaa gacttttagaa ttgtgtaagg aagacgagaa aaagattgtg	780
ttaaaagcgc taggcaacaa gtcagcatca aaggaagagt tgatgagttc tgctgacata	840
atcaaaaagt actcattgga ttacgcctac aacttagctg agaaatacta caaaaacgcc	900
atcgattctc taaatcaagt ttcaagtaaa agtgatattc caggggaaggc attgaaatat	960
cttgctgaat tcaccatcag aagacgtaag taa	993

<210> 24

<211> 894

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 24

atggtcgcac aaactttcaa cctggatacc tacttatccc aaagacaaca acaagttgaa	60
gaggccctaa gtgctgctct tgtgccagct taccctgaga gaatatacga agctatgaga	120
tactccctcc tggcaggtgg caaaagatta agacctatct tatgttttagc 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 gaggaccacc tcaattggtg	420
ctacaagtta ttgctagaat cggacacgcc gttgctgcaa caggcctcgt tggaggccaa	480
gtcgtagacc ttgaatctga aggtaaagct atttccttag aaacattgga gtatattcac	540
tcacataaga ctggagcctt gctggaagca tcagttgtct caggcgggtat 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

<210> 25

<211> 1116

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 25

atggcctctg ttactttggg ttcttgatc gtcgtccacc accataacca tcaccatcca	60
tcatctatcc taactaaatc tcgttcaaga tcctgtccta ttactaac 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	gcggaagag	agtcagacct	360
gtactctgca	tagcagcgtg	cgaattagtt	ggtggcgagg	aatcaaccgc	tatgcctgcc	420
gcttgctgctg	tagaaatgat	tcatacaatg	tcactgatac	acgatgattt	gccatgtatg	480
gataacgatg	atctgagaag	gggtaagcca	actaaccata	aggttttcgg	cgaagatggt	540
gccgtcttag	ctggtgatgc	tttgttatct	ttcgcgttcg	aacatttggc	atccgcaaca	600
tcaagtgatg	ttgtgtcacc	agtaagagta	gtagagcag	ttggagaact	ggctaaagct	660
attggaactg	agggtttagt	tgcaggtaa	gtcgtcgata	tctcttccga	aggctttgat	720
ttgaatgatg	taggtcttga	acatctcgaa	ttcatccatc	ttcacaagac	agctgcactt	780
ttagaagcca	gtgcggttct	cggcgcaatt	ggtggcgagg	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	cgctgatca	actgttgggt	ttcgattctg	ataaagttgc	accactctta	1080
gccttagcca	actacatcgc	ttacagacaa	aactaa			1116

<210> 26

<211> 1086

<212> DNA

<213> Stevia rebaudiana

<400> 26

atggctcttg	taaatcccac	agctttgttc	tatggaacct	ccataagaac	cagaccacaca	60
aacttgctca	acccgaccca	aaaacttcga	cccgtttcct	cgtcttcttt	gccttccttc	120
tcttcagttt	ctgcaatctt	gacggaaaaa	caccaatcaa	acccatcaga	aaacaataac	180
ttgcaaacc	atctcgaaac	accattcaat	ttcgactctt	acatgctgga	gaaagtaaac	240
atggtgaatg	aagctctgga	cgctcgggt	ccactcaaag	acccgataaa	gatccatgaa	300
tccatgcggt	actcccttct	agctggcggg	aaacgcattc	gaccgatgat	gtgcatcgcc	360
gcttgcgaaa	tagtcggagg	caacatatta	aacgccatgc	cagctgcatg	cgcggtcgag	420
atgattcaca	ccatgtcact	agttcatgac	gaccttccat	gcatggataa	cgacgacttc	480
cgacgtggaa	aaccaataag	ccacaagggt	tacgggtgaag	aaatggcggg	tctaaccggg	540
gacgcgttac	tctcattatc	cttcgaacat	atcgcgaccg	cgacaaaagg	cgatatccaaa	600
gacaggatcg	tccgagccat	tggtgaactc	gcaagggtccg	ttggctcgga	gggtttgggtc	660
gccggtcagg	tggttgatat	tttatccgaa	ggggctgatg	ttgggttaga	ccacttgag	720
tatattcata	tacacaagac	tgcaatgttg	cttgagagct	cggtcgtgat	cggcgcgac	780
atgggcgggtg	ggtctgacca	acagatcgaa	aagttgcgaa	agtttgcgag	atcgattggt	840
ttgttgtttc	aggtggtaga	tgatattctt	gatgtcacaa	agtcgactga	ggaattgggg	900

aaaacggcgg gaaaagattt gctgacggac aagacaacgt atccgaagtt gttggggatc	960
gaaaaatcga gagaatttgc ggagaaatta aacaaggaag cgcaagaaca attgtcgggg	1020
tttgatcgcc gcaaggcggc tccgttaatt gcccttgcta attacaatgc ttataggcaa	1080
aactga	1086

<210> 27
 <211> 1029
 <212> DNA
 <213> *Gibberella fujikuroi*

<400> 27	
atggctgaac aacagatctc caaccttctt tcaatgtttg atgcttctca cgcaagccag	60
aagttggaga ttacggttca gatgatggat acctaccatt acagagaaac tcctccagac	120
tcttctctctt cagaaggcgg ttccttatct cgctatgatg agcgacgggt ctcccttccg	180
ctctctcaca atgcagcctc cccagacata gtctcccagt tatgcttctc aacagctatg	240
agctcggagc tcaatcacag gtggaagtca cagcgccctca aggttgctga ctctccctac	300
aactacatcc tgactcttcc atctaaagggt attcgtgggg ctttcattga ctactgaat	360
gtctggctcg aggtccccga agacgagacc tcggtgatca aagaggatg tgccatgctc	420
cacaactcgt ctctcataat cgatgacttc caagacaact cccacttcg gcggggcaag	480
ccatctacac atactgtctt cgggtccagca caagcaatca acacagcaac atatgtcatc	540
gtcaaggcca tcgagaaaaat acaggatatc gtcggtcacg atgcattggc agatgtaact	600
ggcactataa ccacaatctt ccagggtcag gcaatggatc tgtgggtggac tgctaatacc	660
attgttccgt ctatccaaga atatctcctg atgggtcaatg acaagactgg tgccctgttc	720
aggttatcgc ttgaactact ggcgctgaac tctgaagcat ccatcagtga cagcgcgctt	780
gaatctctca gcagcgctgt ctactgctc gggcagtatt tccagataag agatgattac	840
atgaatctca ttgacaacaa gtatactgat cagaaaggat tttgcgagga tctggacgag	900
gggaaatact cgttgactct aatccatgct ctgcagaccg actccagcga ctttctcacc	960
aacatcttat cgatgagaag agtccaagga aaacttacgg cgcagaaaaag atgctggttt	1020
tggaagtga	1029

<210> 28
 <211> 903
 <212> DNA
 <213> *Mus musculus*

<400> 28	
atggagaaaa ctaaagagaa agctgagagg attcttctag agccctatag gtacttactt	60
cagttaccag gtaaacagggt gagaagcaaa ctttcacagg catttaatca ctggctgaaa	120
gttccagaag acaagctaca gattatcatt gaagtgactg aaatgttgca taatgccagt	180
ttactcattg atgatattga agacagttca aagctccgac gtggtttccc agtggctcac	240
agcatctatg gtgtcccatc tgtcattaat tctgccatt acgtctactt ccttgactg	300
gaaaaagtct taacccttga tcaccgggat gcggtgaagc ttttcacacg ccagcttctg	360

gaacttcatc agggacaagg cctcgatatt tactggaggg acacctacac ttgtccaact	420
gaagaagaat ataaagccat ggtgttgag aagacaggtg gtttgtttgg attagcagta	480
ggtcttatgc agctgttctc tgattacaaa gaagatctaa agccactgct tgacacactt	540
gggctctttt tccagattag agatgattat gccaatctac actccaaaga atacagtga	600
aacaaaagtt tctgtgaaga ctgacagaa ggggaagttct cattccccac tatccatgcc	660
atgttggtcaa ggccagaaag caccaggtg cagaacatcc tgcgccagag aacagagaat	720
atagatatta aaaagtattg tgtgcagtac ctggaggatg taggttcttt tgcatacact	780
cgacacactc ttagagagct tgaagctaaa gcctacaaac aaattgaggc ctgtggtggg	840
aacccttcac tagtggtctt agtcaagcac ttaagtaaga tgttcacaga agaaaataaa	900
taa	903

<210> 29

<211> 1020

<212> DNA

<213> *Thalassiosira pseudonana*

<400> 29

atggctcgtt tctacttcct gaacgctctc ctcatggtga tttctttaca aagcaccacg	60
gcattcaccc cggcaaaaact cgcctaccca acaaccacca ctgcattaaa cgttgcctct	120
gccgaaacat catttagcct cgatgaatac ctagcctcca aaatcggacc cattgaatca	180
gctctcgagg catctgtcaa atctcgatt cctcaaactg acaagatatg cgagtctatg	240
gcatactcac tcatggctgg aggaaagcgt atccgtcccg ttttgtgcat tgctgcttgt	300
gaaatgtttg ggggaagtca agatgtggct atgccgacgg ctgtggcttt ggagatgatt	360
catactatga gtcttattca tgacgatttg cttcaatgg acaacgatga tctccgacga	420
ggaaagccaa ctaatcatgt tgtctttgga gaggatgttg ctattcttgc tggggattct	480
cttctcagta cgtcttttga acatgttgcc cgtgaaacca aaggagtgtc agctgaaaag	540
attgtagatg ttatcgctcg cctcgggaag tctgtgggtg cagaggggtct tgctggtgga	600
caggttatgg atcttgagtg tgaggcgaag ccaggaacta ccctcgacga tctcaagtgg	660
attcacattc acaaaaactgc cactcttctt caagtggcag tggcatcagg tgctgttctt	720
ggaggggcca caccagagga ggttgctgct tgtgaactgt tcgcaatgaa tattggactt	780
gccttccagg tcgctgatga tatcttgac gtgacggcat cgagtgagga tcttgcaaaa	840
actgctggaa aggatgaagc cacagataag acaacttatc ctaagctttt gggattggag	900
gagagtaagg catacgctcg acaactcata gacgaagcaa aggaatcttt ggctcctttc	960
ggtgatcgtg ctgctccatt gttggcaatt gccgacttta tcattgatcg aaagaactag	1020

<210> 30

<211> 1068

<212> DNA

<213> *Streptomyces clavuligerus*

<400> 30

atgcacctgg	ctccccgccg	agtaccgcgc	ggccgtcgaa	gcccacctga	ccgcgttcct	60
gaacgccaag	gagcgctcgg	ccgccgccgg	ggggccggtt	ccacaggatg	tgccccgcgt	120
gctgcgggag	ttcatcggcg	ccgggggggg	ggggaagcgg	atccgtccgc	tgctgtgcat	180
cgcgctggc	aggccggcgg	cggaacagga	ctgccggacg	agggtggtgc	cacagcggcg	240
gcgctggaga	tgttccacgc	gttcgcgctg	atccacgacg	acatcatgga	tgactccgcg	300
accaggcgcg	gcagcccgcg	ggtgcaccgg	gcactcgccg	accggctcgg	cgccgctctc	360
gaccccgacc	aagccggaca	actgggggtg	agcacggcga	tcctcgtcgg	ggacctcgcc	420
ctgacctggg	cggacgaact	gctgtacgct	cccctgaccc	cccaccggct	ggccgcggta	480
ctgcccctgg	tcacggccat	gcgcgcggaa	acggtccacg	gccagtacct	ggacatcacc	540
tccgcccgcc	ggccccggc	ggacacctca	ctggcgctgc	gaatcgcgcg	ctacaaaacc	600
gctgcttaca	ccatggaacg	ccccctgcac	atcggagcag	cgctcgccgg	cgcacgaccg	660
gaactcctgg	cagggctcag	cgcttacgcg	ctgccggcgg	gcgaggcatt	ccagctcgcc	720
gacgacctcc	tgggagtgtt	cggcgatcca	cggagaaccg	gcaaaccgga	cctcgacgac	780
ctccgcggcg	gcaagcacac	cgctctcgctg	gccctcgccc	gggaacacgc	cacacctgaa	840
cagcggcaca	ccctggacac	cctgctcggc	acaccaggcc	tcgaccggca	gggcgcgtcc	900
cggctgcgct	gcgtcctcgt	cgccaccggg	gcccgggcgg	aagccgaacg	cctgatcacc	960
gaacggcgcg	accaggccct	caccgcgctc	aacgccctga	cactgcccc	accgctcgcc	1020
gaggcactcg	ccgcctcac	cctcgggagt	accgcacacc	cggcctga		1068

<210> 31

<211> 993

<212> DNA

<213> Sulfolobus acidicaldarius

<400> 31

atgagttact	ttgacaacta	ttttaatgag	attgttaatt	ctgtaaacga	cattattaag	60
agctatatat	ctggagatgt	tcctaaacta	tatgaagcct	catatcattt	gtttacatct	120
ggaggtaaga	ggttaagacc	attaatctta	actatatcat	cagattttatt	cggaggacag	180
agagaaagag	cttattatgc	aggtgcagct	attgaagttc	ttcatacttt	tacgcttggtg	240
catgatgata	ttatggatca	agataatatc	agaagagggg	taccacacagt	ccacgtgaaa	300
tacggcttac	ccttagcaat	attagctggg	gatttactac	atgcaaaggc	ttttcagctc	360
ttaaccagg	ctcttagagg	tttgccaagt	gaaaccataa	ttaaggcttt	cgatattttc	420
actcgttcaa	taataattat	atccgaagga	caggcagtag	atatggaatt	tgaggacaga	480
attgatataa	aggagcagga	ataccttgac	atgatctcac	gtaagacagc	tgcatatttc	540
tcggcatcct	caagtatagg	cgcacttatt	gctggtgcta	atgataatga	tgtaagactg	600
atgtctgatt	tcggtacgaa	tctaggtatt	gcatttcaga	ttgttgacga	tatcttaggt	660
ctaacagcag	acgaaaagga	acttggaag	cctgttttta	gtgatattag	ggagggtaaa	720
aagactatac	ttgtaataaa	aacactggag	ctttgtaaag	aggacgagaa	gaagattgtc	780

ctaaaggcgt taggtaataa gtcagcctca aaagaagaat taatgagctc agcagatata	840
attaagaaat actcttttaga ttatgcatac aatttagcag agaaatatta taaaaatgct	900
atagactctt taaatcaagt ctcttctaag agtgcatac ctggaaaggc tttaaaatat	960
ctagctgaat ttacgataag aaggagaaaa taa	993

<210> 32
 <211> 894
 <212> DNA
 <213> *Synechococcus* sp.

<400> 32	
ttggttgccc aaaccttcaa cctggacacc tacttgagcc aacgccagca acaggtggaa	60
gaggcgcttt ctgcggcatt ggttcccgcc tatccggagc gcatttacga ggcgatgcgc	120
tacagcctgc tggcgggggg gaaacgcctg aggccgatcc tctgtctggc ggcctgtgag	180
ttggccggcg gctctgtgga gcaggccatg cccaccgcct gcgccctgga gatgatccac	240
accatgtcgc tgatccacga cgatctgccg gcgatggaca acgacgattt tcgccgcggc	300
aagcccacca atcacaaggt attcggcgag gatatcgcca ttttggcagg agatgccctg	360
ttggcctatg cttttgagca tatcgccagc caaacgcggg gggcgccgcc gcagttgggtg	420
ctgcaagtca ttgcccgcct tggccatgct gtggcggcaa ccggcttggg agggggccag	480
gtggtggatc tggagtccga aggcaaagcc atttccctag aaactttgga gtacatccac	540
agtcacaaga cgggtgctct gctggaggcc tcggtggttt cgggagggat cctggcaggg	600
gccgatgagg agctgctggc gcggctgagc cactacgctc gggacatcgg cctggctttt	660
cagatcgtgg acgacatttt ggatgttact gccaccagcg agcaactggg caaaacggca	720
ggcaaggatc aagctgccgc caaagccacc taccagct tgttgggcct agaggcttcc	780
cggcagaaag ctgaggaact gatccaatcg gccaaaggagg cgttgcgccc ctacggatcc	840
caggccgagc ccctgttggc tctggccgat ttcacaccc gccgccagca ttga	894

<210> 33
 <211> 1116
 <212> DNA
 <213> *Arabidopsis thaliana*

<400> 33	
atggcttcag tgactctagg ttcattgatt gttgttcacc accacaatca tcattcatcca	60
tcttcaatcc ttaccaaata cagatccaga tcttgccta taactcttac taaaccatc	120
tcctttcgat caaaacgcac cgtttcatca tcttcttcaa tcgtttcttc ttccgttgtt	180
acaaaagaag acaatctacg ccaatctgaa ccattctctt tcgatttcat gtcgtacatc	240
atcaccaaag ccgaattagt caacaaagct ttagattcag ctgttcctct ccgtgagcca	300
ctcaagatcc acgaagcgat gcgttactct cttctcgccg gtggcaaaag agtttagacca	360
gttctctgca tcgctgcttg tgaactcgct ggagggtgaag aatcaaccgc tatgccagca	420
gcttgcgccc tcgagatgat tcacaccatg tcgttgatcc acgacgatct cccttgatg	480
gataacgacg atctccgccg tggaaaaccg accaaccaca aagtgtttgg tgaagacgtc	540

gctgttttag	ccggagacgc	gcttctctct	ttcgctttcg	agcatttagc	ttcggcgacg	600
agttctgatg	ttgtttctcc	ggtgagagtg	gttcgagccg	ttggagaatt	ggctaaagcg	660
ataggaacag	aagggtagt	ggcgggtcaa	gtcgtggata	ttagtagtga	agggtagat	720
ttaaacgacg	tcggtttaga	gcatttggag	tttatccatt	tgcataaaac	ggcggcgttg	780
cttgaagcct	ctgctgtttt	gggagctatt	gttgggtggag	gaagtgatga	tgagattgag	840
aggttaagaa	agtttgcgag	atgtattggt	ttgttgtttc	aggtggttga	tgatatcttg	900
gatgtgacga	aatcgtcgaa	agagttaggg	aaaactgctg	ggaaagattt	gattgctgat	960
aagttgacgt	atcctaagat	tatgggtttg	gagaaatcga	gagagtgtgc	tgagaaattg	1020
aatagagagg	ctcgtgatca	gcttttaggg	tttgattctg	ataagggtgc	tcctttgttg	1080
gctttggcta	attacattgc	ctatagacag	aactga			1116

<210> 34

<211> 2364

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 34

atgaaaaccg	ggtttatctc	accagcaaca	gtatttcatac	acagaatctc	accagcgacc	60
actttcagac	atcacttatc	acctgtact	acaaactcta	caggcattgt	cgcttaaga	120
gacatcaact	tcagatgtaa	agcagtttct	aaagagtact	ctgatctgtt	gcagaaagat	180
gaggcttctt	tcacaaaatg	ggacgatgac	aagggtgaaag	atcatcttga	taccaacaaa	240
aacttatacc	caaatgatga	gattaaggaa	tttgttgaat	cagtaaaggc	tatgttcggt	300
agtatgaatg	acggggagat	aaacgtctct	gcatacgata	ctgcatgggt	tgctttgggt	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	atttgcaa	atagaagatga	aaacgcagaa	600
catatgccaa	ttggttttga	agtaacattc	ccatcactaa	ttgatatcgc	gaaaaagttg	660
aacattgaag	tacctgagga	tactccagca	cttaaagaga	tctacgcacg	tagagatatc	720
aagttaacta	agatcccaat	ggaagttctt	cacaaggtag	ctactacttt	gttacattct	780
ttggaaggaa	tgcttgattt	ggagtgggaa	aaactgttaa	agctacaatg	taaagatggt	840
agtttcttgt	tttccccatc	tagtaccgca	ttcgccctaa	tgcaaacaaa	agatgagaaa	900
tgcttacagt	atctaacaaa	tatcgtcact	aagttcaacg	gtggcgtgcc	taatgtgtac	960
ccagtcgatt	tgtttgaaaca	tatttgggtt	gttgatagac	tgagagatt	ggggattgcc	1020
agatacttca	aatcagagat	aaaagattgt	gtagagtata	tcaataagta	ctggacccaaa	1080
aatggaattt	gttgggctag	aaatactcac	gttcaagata	tcgatgatac	agccatggga	1140

ttcagagtgt	tgagagcgca	cggttatgac	gtcactccag	atgttttttag	acaatttgaa	1200
aaagatggta	aattcgtttg	ctttgcaggg	caatcaacac	aagccgtgac	aggaatgttt	1260
aacgtttaca	gagcctctca	aatgttgttc	ccaggggaga	gaattttgga	agatgccaaa	1320
aagtctctct	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
gtttcttact	acttggctgc	ggcttcaata	ttcgaacctg	agagatctaa	ggagagaatc	1740
gcttgggcaa	agacaacaat	cttagtcgat	aagatcacat	caattttcga	ttcctctcag	1800
tcaagtaagg	aagatattac	tgcctttatt	gacaagtttc	gtaacaagtc	ctcctctaaa	1860
aagcactcta	tcaacgggta	accatggcat	gaagttatgg	tagctttgaa	aaagacctta	1920
cacggctttg	ctctggatgc	tcttatgact	cattctcaag	atatacatcc	acagttacat	1980
caagcctggg	aatgtgggtt	gactaaacta	caagacggcg	tagatgttac	tgctgagcta	2040
atgggtccaaa	tgatcaacat	gactgctggc	agatgggtat	caaaggaatt	acttactcat	2100
ccacaatatc	aaagattgtc	tactgtgaca	aattctgtgt	gtcacgatat	taccaaactt	2160
cacaatttca	aggagaattc	caccacagtg	gattcaaagg	ttcaggaact	agtccagttg	2220
gttttttagtg	acacaccaga	tgatttggat	caagatatga	aacaaacatt	cctgacagtg	2280
atgaagacat	tctactacaa	ggcgtgggtg	gatccaaaca	ctataaacga	tcatatatct	2340
aaagttttcg	aatcgtaat	ttga				2364

<210> 35

<211> 1584

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 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	gtcttgccctc	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
gcttttagat cacacgccgc agcaggtaga 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
attacccta 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 gggcttatgg cattcaactg ttgaagagac tgcttatgcc	1380
ttacagatct tggccccacc ttctgggtgg ggcaatatcc cagtccaaca agcacttact	1440
agaggcagag caagattgtg tggagccttg ccactgactc ctttatggca tgataaggat	1500
ttgtatactc cagtaagagt agtcagagct gccagagctg ctgctctgta cactaccaga	1560
gatctattgt taccaccatt gtaa	1584

<210> 36

<211> 1551

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 36

atgaacgccc tatccgaaca cattttgtct gaattgagaa gattattgtc tgaaatgagt	60
gatggcggat ctgttggtcc atctgtgtat gatacgcccc 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 ccctattgggt	360
gctgaactga tcttgccctca gttttgtgga gaggctgctt ggttgttggg aggtgtggcc	420
ttccctagac acccagccct attaccatta agacaggctt gtttagtcaa actgggtgca	480
gtcgccatgt tgccttcagg acaccattg ctccactcct gggaggcatg gggtaactct	540
ccaacaacag cctgtccaga cgatgatgggt tctataggta tctcaccagc agctacagcc	600

gcctggagag	cccaggctgt	gaccagagggc	tcaactcctc	aagtgggcag	agctgacgca	660
tacttacaaa	tggcttcaag	agcaacgaga	tcaggcatag	aaggagtcctt	ccctaattgtt	720
tggcctataa	acgtattcga	accatgctgg	tactgtaca	ctctccatct	tgccggtctg	780
ttcgcccatc	cagcactggc	tgaggctgta	agagttatcg	ttgctcaact	tgaagcaaga	840
ttgggagtg	atggcctcgg	accagcttta	cattttgctg	ccgacgctga	tgatactgca	900
gttgcccttat	gcgttctgca	tttggctggc	agagatcctg	cagttgacgc	attgagacat	960
tttgaaattg	gtgagctctt	tgttacattc	ccaggagaga	gaaatgctag	tgtctctacg	1020
aacattcacg	ctcttcatgc	tttgagattg	ttaggtaaac	cagctgccgg	agcaagtgca	1080
tacgtcgaag	caaatagaaa	tccacatggg	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	ggtgctgggtg	ctgcacctta	a	1551

<210> 37

<211> 2364

<212> DNA

<213> Stevia rebaudiana

<400> 37

atgaagaccg	gcttcatctc	tcccgccacc	gtcttccacc	accgtatttc	tccggcaacc	60
accttccgcc	accacctttc	tccggcgacc	accaactcca	ctggaattgt	agctcttaga	120
gacatcaact	tccggtgtaa	agcggtatcc	aaagagtact	ctgatttact	acaaaaagat	180
gaggcttcat	ttaccaagtg	ggacgatgac	aaagtgaagg	accatttgga	cacaaataag	240
aatttgtatc	caaacgatga	gatcaaggag	tttgttgaga	gcgtgaaagc	aatgtttggt	300
tctatgaatg	acggagaaat	aaatgtgtca	gcgtatgata	cggcttgggt	tgactcgtg	360
caagatgttg	atggaagtgg	ttccctcaa	tttccatcaa	gtttggagtg	gatcggaac	420
aatcaactct	cagatgggtc	ttggggcgat	catttggttat	tttcgggtca	tgataggatc	480
attaacacgt	tggcatgtgt	tatagcgctt	acttcttggg	acgtccatcc	aagtaaattgt	540
gaaaaaggac	tgaattttct	tagagaaaac	atatgtaaac	tcgaagacga	gaacgcggaa	600
catatgcaa	ttggttttga	agtcacgttc	ccgtcgctaa	tagatatcgc	aaagaagcta	660
aatattgaag	ttcctgagga	tactctgcc	ttaaaagaaa	tttatgcaag	aagagacata	720
aaactcacia	agataccaat	ggaagtattg	cacaaagtgc	ccacaacttt	acttcatagt	780
ttggaaggaa	tgccagattt	ggaatgggaa	aaacttctga	aattgcaatg	caaagatgga	840
tcatttctgt	tttctccatc	atctactgct	tttgactca	tgcaaacaaa	agatgaaaag	900

tgtcttcagt	atttgacaaa	tattgttacc	aaattcaatg	gtggagttcc	gaatgtgtac	960
ccggtggatc	tattcgaaca	tatttgggta	gttgatcgac	ttcaacgact	tgggattgct	1020
cgttatttca	aatcagagat	caaagattgc	gttgaatata	ttaacaagta	ttggacaaag	1080
aatgggattt	gttgggcaag	aaacacgcac	gtacaagata	ttgatgatac	cgcaatggga	1140
tttagggttt	taagagcaca	tggttatgat	gttactccag	atgtatttcg	acaatttgag	1200
aaggatggta	aattcgtatg	tttcgctgga	cagtcaacac	aagccgtcac	cggaatgttc	1260
aatgtgtata	gagcgtcaca	aatgctcttt	cccggagaaa	gaattcttga	agatgcaaag	1320
aaattttcat	ataattattt	gaaagaaaaa	caatcgacaa	atgagcttct	tgataaatgg	1380
atcatcgcca	aagacttacc	tggagaggtt	ggatatgctc	tagacatacc	atggtatgca	1440
agcttaccgc	gactcgagac	aagatattac	ttagagcaat	acgggggcga	ggatgatgtt	1500
tggattggaa	aaactctata	caggatggga	tatgtgagca	ataatacgta	ccttgaaatg	1560
gccaaattgg	actacaataa	ctatgtggcc	gtgcttcaac	tcgaatggta	cactatccag	1620
caatggtatg	ttgatatcgg	tatcgaaaag	tttgaaagtg	acaatatcaa	aagcgtatta	1680
gtgtcgtatt	acttggctgc	agccagcata	ttcgagccgg	aaagggtcaa	ggaacgaatc	1740
gcgtgggcta	aaaccaccat	attagttgac	aagatcacct	caatttttga	ttcatcacia	1800
tcctcaaaag	aggacataac	agcctttata	gacaaattta	ggaacaaatc	gtctttctaag	1860
aagcattcaa	taaatggaga	accatggcac	gaggtgatgg	ttgcactgaa	aaagacccta	1920
cacggcttcg	ctttggatgc	actcatgact	catagtcaag	acatccaccc	gcaactccat	1980
caagcttggg	agatgtgggt	gacgaaattg	caagatggag	tagatgtgac	agcggaatta	2040
atggtacaaa	tgataaatat	gacagctggt	cgttgggtat	ccaaagaact	tttaactcat	2100
cctcaatacc	aacgcctctc	aaccgtcaca	aatagtgtgt	gtcacgatat	aactaagctc	2160
cataacttca	aggagaattc	cacgacggta	gactcgaaag	ttcaagaact	agtgcaactt	2220
gtgttttagc	acacgcccga	tgatcttgat	caggatatga	aacagacggt	tctaaccgtc	2280
atgaaaacct	tctactacaa	ggcgtggtgt	gatccgaaca	cgataaatga	ccatatctcc	2340
aagggtgttcg	agattgtaat	atga				2364

<210> 38

<211> 1584

<212> DNA

<213> Streptomyces clavuligerus

<400> 38

ttgcccagcg	cgcatgatgc	ccctccgcct	cagatacgac	agcggaccct	tgatgatgag	60
gcgacgcaac	tcctcacgga	gtcggccgag	gacgcctggg	gtgaggtgtc	cggtgtccgag	120
tacgaaacgg	cgcggtggtt	ggcccacgcc	acctggctcg	gcggtcacgc	cacacgggtg	180
gccttctctg	tggagcggca	gcatgaggac	ggctcgtggg	gcccgcccg	cggtgtaccgt	240
ctcgtaccca	cgctgagtgc	cgtacacgcc	ctgctcacct	gtctggcgtc	tcccgcgcag	300
gaccacggag	tgctcatga	ccggctcctg	cgcgcagttg	acgcgggcct	gacggcactg	360

cgtcgtcttg ggacgagcga cagcccgcgc gacaccattg cggtcgaact ggtcataccc 420
 tcgtctcttg agggcatcca gcacctctg gaccggcgcc acccgcatte ccgacccgct 480
 ttttcgcaac accgcggcag cctcgtctgc cccggggggc tcgacggccg cacgctgggg 540
 gccttgcgct cccacgccgc agccggcaca cctgtcccgc gcaagggtgtg gcacgcctcg 600
 gaaaccttgg ggctatcgac cgaggcagcc tcccaccttc aaccgcacca gggcatcatc 660
 ggtggctccg ccgccgcgac agcaacatgg ctcaccaggg tcgccccgtc gcaacagagc 720
 gacagcgcac ggcgtacct ggaagaactc cagcaccgat acagcggccc ggtgccctcc 780
 atcaccccgga tcacctattt cgaacggggc tggctgctca acaacttcgc tgccgcgggg 840
 gttccatgcg aggtctccgc agcccttctc gacagcctgg aggcagcgct cacaccacag 900
 ggcgctccag cgggtgcggg actgccgccg gacgcggatg acaccgccgc cgttctgctg 960
 gcgcttgcca cgcacggccg cgggcgccgt cccgaggtcc tcatggacta ccgcacggac 1020
 ggctacttcc agtgcttcat cggcgaacgc accccttcca tcagcaccaa tgcccatgtc 1080
 ctggagacgc tcggtcacca cgtcgcccaa caccctcagg acagggcccc atacggctca 1140
 gccatggaca ccgcatcagc gtggctcctc gcggctcaga agcaggatgg cagctggctc 1200
 gacaagtggc acgcctcccc ctactacgcc accgtctgct gacccaggc actggcagcc 1260
 cacgcttccc ctgccaccgc cccgcacgg cagcgtgctg tgcggtgggt gctggcaaca 1320
 caacgctcgg acggcggtg gggcctgtgg cactccacgg tcgaggagac cgcctacgcc 1380
 ctgcagatcc tcgccccacc ttccggcggc gggaacatcc ccgtgcaaca ggcgctcacc 1440
 agggggcgcg cccgcctctg cggcgctttg ccgctgactc ccctatggca tgacaaggac 1500
 ctgtacacgc cgttacgtgt cgtccgcgcc gcccggtgcc cgcacctgta caccaccgct 1560
 gacctgttcc tgccgcccct gtga 1584

<210> 39

<211> 1551

<212> DNA

<213> *Bradyrhizobium japonicum*

<400> 39

gtgaacgcgc tgtccgaaca tatectttcc gaattgcgcc gcctgctgag cgaaatgagc 60
 gatggcggca gcgtcgttcc gtccgtctac gacacggcgcc aggcgctgcg cttccacggc 120
 aacgtcaccg gtcggcagga cgcatacgcg tggctcatcg cgcagcaaca ggccgacggc 180
 ggatggggaa gcgcggactt cccgctgttc cgccatgcgc ccacgtgggc ggcgttactg 240
 gcattgcagc gtgccgatcc tcttcccgga gctgcggacg cagtccagac tgcaacgagg 300
 ttctccagc gccagcccga tccctacgca catgcggtgc cagaagacgc gccgatcggc 360
 gcggagctga tcctgccgca gttttgcggt gagggccgat ggttgctggg tggcgtagcg 420
 tttccgcgcc atcctgcgct gttgccattg cggcaagcgt gcctgggtcaa gctgggggag 480
 gtggcgatgt tgccgagcgg ccatccgttg ctacactcct gggaagcctg ggggacgtcg 540
 ccgaccaccg catgcccgga tgacgacggc agcatcggca tcagtccggc ggccaccgcc 600

gcgtggcgtg	cccagggcgt	gacacggggg	agcacgccgc	aggtcgggcg	cgccgatgcg	660
tatctgcaga	tggcatcgcg	ggcgacgcgc	agcggcatcg	aaggtgtctt	tcccaacgtc	720
tggccgatca	atgtgttcga	gccatgctgg	tcgctgtaca	ccctgcatct	ggccgggctt	780
ttcgcgcatc	ccgcgctcgc	ggaggcgggt	cgcgatgatcg	tcgcgagct	cgaggcccgt	840
ctgggcgtgc	acgggtctggg	cccggccttg	cacttcgcgg	ctgatgcgga	cgacaccgcc	900
gttgcgttgt	gcgtcctgca	ccttgaggc	cgtgacccgg	cggtcgatgc	gttgcgccat	960
ttcgaaatcg	gcgagctggt	cgtcaccttc	cccggcgaac	gcaatgcctc	ggtgtcgacc	1020
aacattcatg	ccctgcatgc	gttgcgactg	ttgggaaagc	ccgccgcggg	cgccagcgcg	1080
tacgtcgagg	ccaatcgcaa	cccgcacggt	ctatgggaca	acgaaaaatg	gcacgtttcg	1140
tggctgtatc	ccaccgcgca	tgcggtcgct	gcgctggcgc	aaggcaagcc	ccagtggcga	1200
gatgagcgcg	cgctggcggc	gctgctgcag	gcgcagcgcg	acgacggtgg	ctggggcgcg	1260
ggtcgcgggt	ccacgttcga	ggaaaccgcc	tatgcgctgt	ttgcgttgca	cgtgatggat	1320
gggagcgaag	aggcgacagg	gcgccggcgc	atcgcgcagg	tggcggcgcg	tcgctggag	1380
tggatgctcg	cccgccatgc	ggcgcatgga	ttgccgcaga	cgccgctgtg	gatcggcaag	1440
gaactgtatt	gccccactcg	ggtcgtgcgc	gtggccgaac	tcgccggggt	gtggctggcg	1500
cttcgttggg	ggcggcgcg	cctggccgag	ggggcaggag	cggcgccatg	a	1551

<210> 40

<211> 2355

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 40

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	atcttgggggt	300
ttagtcaatc	acacgcacaa	tcacaaccat	ccacttttga	aagattcttt	atcctcaact	360
ttggcttgca	tcgtggccct	aaagagatgg	aacgtagggtg	aggatcagat	taacaagggg	420
cttagtttca	ttgaatctaa	cttggtctcc	gcgactgaaa	aatctcaacc	atctccaata	480
ggattcgata	tcatctttcc	aggtctgtta	gagtacgcca	aaaatctaga	tatcaactta	540
ctgtctaagc	aaactgattt	ctcactaatg	ttacacaaga	gagaattaga	acaaaagaga	600
tgtcattcaa	acgaaatgga	tggttaccta	gcttatatct	ctgaagggtct	tggtaatctt	660
tacgattgga	atatggtgaa	aaagtaccag	atgaaaaatg	gctcagtttt	caattcccct	720
tctgcaactg	cggcagcatt	cattaaccat	caaaatccag	gatgcctgaa	ctatttgaat	780

tcactactag	acaaattcgg	caacgcagtt	ccaactgtat	accctcacga	tttgtttatc	840
agattgagta	tgggtggatac	aattgaaaga	cttgggtatat	cccaccactt	tagagtcgag	900
atcaaaaaatg	ttttggatga	gacataccgt	tgttgggttg	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	ctaattggcg	aagtggtaaa	2100
gtcgaagagg	aagtagttga	ggaaatgatg	atgatgatca	aaaacaagag	aaaggagttg	2160
atgaaactaa	tcttcgaaga	gaacggttca	attgttccta	gagcatgtaa	ggatgcattt	2220
tggaacatgt	gtcatgtgct	aaactttttc	tacgcaaacg	acgatgggtt	tactgggaac	2280
acaatactag	atacagtaaa	agacatcata	tacaaccctt	tgggtcttagt	aaacgaaaac	2340
gaggagcaaa	gataa					2355

<210> 41

<211> 2355

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 41

atgaatctgt	ccctttgtat	agctagtcca	ctgttgacaa	aatcttctag	accaactgct	60
ctttctgcaa	ttcatactgc	cagtactagt	catggaggtc	aaacaaaccc	aacaaatttg	120

ataatcgata ctactaagga gagaatccaa aagctattca aaaatgttga aatctcagta	180
tcattcttatg acaccgcatg gggtgcaatg gtgccatcac ctaattcccc aaaaagtcca	240
tgttttccag agtgcttgaa ttggttaatc aataatcagt taaacgatgg ttcttggggt	300
ttagtcaacc acactcataa ccacaatcat ccattattga aggactcttt atcatcaaca	360
ttagcctgta ttgttgcatg gaaaagatgg aatgtagggtg aagatcaaat caacaagggg	420
ttatcattca tagaatccaa tctagcttct gctaccgaca aatcacaacc atctccaatc	480
gggttcgaca taatcttccc tggtttgctg gagtatgcca aaaaccttga tatcaactta	540
ctgtctaaac aaacagatgt ctctttgatg ctacacaaaa gagagttaga gcagaaaaga	600
tgccattcta acgaaattga cgggtactta gcatatatct cagaagggtt gggtaatgtg	660
tatgactgga acatgggtcaa aaagtatcag atgaaaaatg gatccgtatt caattctcct	720
tctgcaactg ccgcagcatt cattaatcat caaaacctg ggtgtcttaa ctacttgaac	780
tcactattag ataagtttgg aaatgcagtt ccaacagtct atcctttgga cttgtacatc	840
agattatcta tggttgacac tatagagaga ttaggtatgt ctcattcatt cagagttag	900
atcaaaaatg ttttgacga gacatacaga tgttgggtcg aaagagatga gcaaattctt	960
atggatgtcg tgacctgctg tctggctttt agattgctaa ggatacacgg atacaaagta	1020
tctctgatc aactggctga gattacaaac gaactggctt tcaaagacga atacgccgca	1080
ttagaaacat accatgcac ccaaatactt taccaggaag acctaagttc aggaaaacaa	1140
atcttgaagt ctgcagatgt cctgaaaggc attctgtcta cagatagtaa taggttgtct	1200
aaattgatac acaaggaagt agaaaacgca ctaaagttt ctattaacac tggtttagag	1260
agaatcaata ctaggagaaa cattcagctg tacaacgtag ataatacaag gattcttaag	1320
accacctacc atagttcaaa ctttccaac acctattact taagattagc tgtcgaagac	1380
ttttacactt gtcaatcaat ctacagagag gagttaaagg gcctagaaag atgggtagtt	1440
caaaacaagt tggatcaact gaagtttgct agacagaaga cagcatactg ttattttctt	1500
gttgctgcta ccctttcatc cccagaattg tctgatgcca gaataagttg ggccaaaaat	1560
ggtattctta caactgtagt cgatgatgtt tttgatattg gaggtactat tgatgaactg	1620
acaaatctta ttcaatgtgt tgaagatgg 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 tacgctaattg atgatggctt cactggaaat	2280
acaatatggg atacagttaa agatatcatc tacaaccac ttgttttggg caatgagaac	2340
gaggaacaaa gataa	2355

<210> 42

<211> 1773

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 42

atggctatgc cagtgaagct aacacctgcg tcattatcct taaaagctgt gtgctgcaga	60
ttctcatccg gtggccatgc tttagattc gggagtagtc tgccatgttg gagaaggacc	120
cctacccaaa gatctacttc ttctctact actagaccag ctgccgaagt gtcacaggt	180
aagagtaaac aacatgatca ggaagctagt gaagcgacta tcagacaaca attacaactt	240
gtggatgtcc tggagaatat gggaaatacc agacattttg ctgcagagat aaagtgcata	300
ctagacagaa cttacagatc ttggttaca 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
ggcgactga gaaagccttc ttattcaaa gaggttgaa atgcactgga tggacctttt	660
tacaccacac ttgatagact tcatcatagg tggaaatttg aaaacttcaa cattattgag	720
caacacatgt tggagactcc atacttatct aaccagcata catcaaggga taccctagca	780
ttgtcaatta gagatttttc ctctcaca ttacttatc aacaagagct acagcatctg	840
gagagtggg ttaaggaatg tagattagat caactacagt tcgcaagaca gaaattagcg	900
tacttttacc tatcagccgc aggcaccatg ttttctctg agctttctga tgcgagaaca	960
ttatgggcca aaaacggggg gttgacaact attgttgatg atttctttga tgttgccggg	1020
tctaaagagg aattggaaaa cttagtcatg ctggtcgaaa tgtgggatga acatcacaaa	1080
gttgaaattct attctgagca ggtcgaaatc atcttctctt ccatctacga ttctgtcaac	1140
caattgggtg agaaggcctc ttggttcaa gacagatcaa ttacaaaaca cttgttgaa	1200
atatggttag acttggttaa gtccatgatg acggaagttg aatggagact gtcaaaatac	1260
gtgcctacag aaaaggaata catgattaat gcctctctta tcttcggcct aggtccaatc	1320
gttttaccag ctttgtattt cgttggtcca aagatttcag aaagtatagt aaaggacca	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	atacactggg	atctgatggt	taa			1773

<210> 43

<211> 2232

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 43

atgcagaact	tccatggtac	aaaggaaagg	atcaaaaaga	tgtttgacaa	gattgaattg	60
tccgtttctt	cttatgatac	agcctgggtt	gcaatgggtcc	catcccctga	ttgcccagaa	120
acaccttggt	ttccagaatg	tactaaatgg	atcctagaaa	atcagttggg	tgatggtagt	180
tggtcacttc	ctcatggcaa	tccacttcta	gttaaagatg	cattatcttc	cactcttgct	240
tgtattctgg	ctcttaaaag	atggggaatc	ggtgagggaac	agattaacaa	aggactgaga	300
ttcatagaac	tcaactctgc	tagtgaacc	gataacgaac	aacacaaacc	aattggattt	360
gacattatct	ttccaggtat	gattgaatac	gctatagact	tagacctgaa	tctaccacta	420
aaaccaactg	acattaactc	catgttgcac	cgtagagccc	ttgaattgac	atcaggtgga	480
ggcaaaaatc	tagaaggtag	aagagcttac	ttggcctacg	tctctgaagg	aatcggttaag	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	cgtcttggtg	ttgatagaca	tttcagaaag	780
gagagaaagt	tcgttctgga	tgaacatac	agattttggt	tgcaaggaga	agaggagatt	840
ttctccgata	acgcaacctg	tgctttggcc	ttcagaatat	tgagacttaa	tggttacgat	900
gtctctcttg	aagatcactt	ctctaactct	ctgggcgggt	acttaaagga	ctcaggagca	960
gctttagaac	tgtacagagc	cctccaattg	tcttaccag	acgagtcctt	cctggaaaag	1020
caaaattcta	gaacttctta	cttcttaaaa	caagggttat	ccaatgtctc	cctctgtggt	1080
gacagattgc	gtaaaaacat	aattggagag	gtgcatgatg	ctttaaactt	ttccgaccac	1140
gctaacttac	aaagattagc	tattcgtaga	aggattaagc	attacgctac	tgacgataca	1200
aggattctaa	aaacttccta	cagatgctca	acaatcggtg	accaagattt	tctaaaactt	1260
gcagtgggaag	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 atcgagcggt gggatgtgaa tggcagtgc	1560
gatttttcta gtgaggaagt tgagattatc tattctgcta tccactcaac tatctctgaa	1620
ataggtgata agtcatttgg ctggcaagggt agagatgtaa agtctcaagt tatcaagatc	1680
tggctggact tattgaaatc aatgttaact gaagctcaat ggtcttcaaa caagtctgtt	1740
cctaccctag atgagtatat gacaaccgcc catgtttcat tcgcacttgg tccaattgta	1800
cttcagcct tatacttcgt tggcccaaag ttgtcagaag aggttgcagg tcatcctgaa	1860
ctactaaacc tctacaaagt cacatctact tgtggcagac tactgaatga ttggagaagt	1920
tttaagagag aatccgagga aggtaagctc aacgctatta gttatacat 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

<210> 44
 <211> 2355
 <212> DNA
 <213> Stevia rebaudiana

<400> 44	
atgaatcttt cactatgcat cgcgtccctt ttgttaacca aatcaaatcg acccgcggt	60
ctgtcagcta ttcatacagc atcaacttca catggtggac aaactaatcc cactaatctg	120
atcattgata caaccaaaga acggatccaa aaacagttta aaaatgtaga aatttctgtt	180
tcttcatatg acacagcatg ggtagccatg gtcccttctc caaactcacc caaatcgcct	240
tgtttccctg agtgtctcaa ttggttaatt aataatcagc ttaatgatgg ttcatggggt	300
cttggttaatc acactcataa tcataatcac ccgttgctta aagattctct atcttcaaca	360
ttagcatgta ttgttgcat aaaaagatgg aatgttgggg aagatcaaat aaataaaggt	420
ctaagtttta ttgagtcaaa tcttgcttca gctactgaaa aaagtcaacc atctcccatt	480
ggttttgaca tcatatttcc tggtttgctt gagtatgca aaaacttgga cataaacctc	540
ctttcaaaac aaacagattt tagtttgatg ctacataaga gggaattgga gcaaaaaaga	600
tgccattcaa atgagatgga tggatacttg gcgtatatct ctgaaggact cggtaattta	660
tatgattgga atatggtgaa gaaatatcag atgaaaaatg gttctgtttt caactcacca	720
tcagcaacag ctgctgcttt cattaatcat caaaatcctg gttgtcttaa ttatttaa	780
tcacttttgg acaagtttgg taatgcagtc ccaacagttt atcctcatga tttatttatc	840
cgacttttcta tggttgacac aattgaaaga ttaggaattt cacaccattt cagagtggaa	900
attaaaaatg ttttagatga aacatacaga tgttgggttg aacgagatga gcaaatattc	960
atggatgttg taacatgtgc tttagccttt cggttattaa ggatcaatgg gtatgaagtt	1020

tccccagatc cattggctga aattactaat gaattagctt tgaaagacga atatgcagct	1080
cttgaaacat atcatgctgc acatatatta taccaagagg atttatcttc tggaaaacaa	1140
atcttgaagt cagctgattt cctcaaagag ataatatcca ctgattcaaa caggctttct	1200
aaattaattc acaaagaggt ggaaaatgct ctttaagttcc ctatcaatac cggtttagaa	1260
cgcataaaca ctagacgaaa tatacagctt tacaatgtag acaatacaag aattctgaaa	1320
actacatatc actcatcaaa tattagtaac actgattacc taaggttggc tgttgaagat	1380
ttctacacct gccaatctat ttatcgtgaa gaattaaaag gtcttgaaag gtgggtggta	1440
gagaataagt tggaccagct caagtttgct aggcaaaaga ccgcctactg ttatttctct	1500
gttgctgcaa cactttcgtc tcccgaatta tcagatgctc gtatttcatg ggccaaaaat	1560
ggcatattaa ctacagtagt tgatgacttt ttgatatcg gtggtacaat cgatgaattg	1620
accaacctga ttcaatgtgt tgaaaaatgg aatgtagatg tcgacaagga ttgttgttca	1680
gagcatgttc ggattttatt tttagcatta aaagatgcaa tctgttggat tggagatgaa	1740
gcttttaaat ggcaagcgcg cgatgtaact agccatgtta ttcaaacttg gttggaacta	1800
atgaatagta tgttgagaga agctatatgg acaagagatg cttatgtgcc aacattaaat	1860
gaatatatgg aaaacgctta cgtgtcattt gcattaggcc cgattgtcaa gccggctatt	1920
tactttgtgg ggcccaaatt atcagaggag attgttgaaa gctctgaata tcataatcta	1980
tttaagctaa tgagcacgca gggtcgactt ctaaacgata tccatagctt caagagggaa	2040
tttaaggaag gcaaattaaa cgcggtagca ttgcatttga gtaacggaga aagtgggaaa	2100
gtggaagaag aggttgtgga ggagatgatg atgatgatta aaaacaagag gaaagaatta	2160
atgaaattaa tttttgaaga aaatggtagc attgttccta gagcttgtaa agatgcattt	2220
tggaacatgt gtcacgtgtt gaattttttt tacgcaaacg atgacgggtt tactggaaac	2280
acgattcttg atactgtgaa ggacatcatt tacaaccggt tgggtgcttgt gaatgaaaat	2340
gaagaacaaa ggtaa	2355

<210> 45

<211> 2355

<212> DNA

<213> Stevia rebaudiana

<400> 45

atgaatcttt cactatgcat tgcgtccctt ttgttaacca aatcaagtcg acccacggct	60
ctgtcagcta ttcatacagc atcaacttca catggtggac aaactaatcc cactaatctg	120
atcattgata caaccaaaga acggatccaa aaactgttta aaaatgtaga aatttctgtt	180
tcttcatatg acacagcatg ggtagccatg gtcccttctc caaactcacc caaatcgcct	240
tgtttccctg agtgtctcaa ttggttaatt aataatcagc ttaatgatgg ttcattgggt	300
cttggttaatc acactcataa tcataatcac ccgttgctta aagattctct atcttcaaca	360
ttagcatgta ttgttgcat aaaaagatgg aatgttgggg aagatcaaat aaataaaggt	420
ctaagtttta ttgagtcaaa tcttgcttca gcaactgaca aaagtcaacc atctcccatt	480

ggttttgata tcatatattcc tggtttgctt gagtatgCGa	aaaacttgga cataaacctc	540	
ctttcaaac aaacagattt tagtttgatg ctacataaga	gggaattgga gcaaaaaaga	600	
tgccattcaa atgagattga tggatacttg gcgtatatct	ctgaaggact cggtaattta	660	
tatgattgga atatggtgaa gaaatatcag atgaaaaatg	gttctgtttt caactcacca	720	
tcagcaacag cagctgcttt cattaatcat caaaatcccg	gttgtcttaa ttatttaa	780	
tcacttttgg acaagtttgg taatgcagtc ccaacagttt	atcctcttga tttatatatc	840	
cggcttttcta tggttgacac aattgaaaga ttaggaattt	cacaccattt cagagtggaa	900	
attaaaaatg ttttagatga aacatacaga tgttgggtgg	aacgagatga gcaaatattc	960	
atggatgttg taacatgtgc tttagccttt cggttattaa	ggatccacgg gtataaagtc	1020	
tccccagatc aattggctga aattactaat gaattagctt	tcaaagacga atacgcagct	1080	
cttgaaacat atcatgcac acagatatta taccaagagg	atttatcttc tggaaaacaa	1140	
atcttgaagt cagctgattt cctcaaaggg atattatcca	ctgattcaaa caggctttct	1200	
aaattaattc acaaagaggt ggaaaatgct ctttaagttcc	ctatcaatac cggtttagaa	1260	
cgcataaaca ctagacgaaa tatacagctt tacaatgtag	acaatacaag aattctgaaa	1320	
actacatatac actcatcaaa tattagtaac acttattacc	taaggttggc tgttgaagat	1380	
ttctacacct gccaatctat ttatcgtgaa gaattaaaag	gtcttgaaag gtgggtggta	1440	
cagaataagt tggaccagct caagtttgct aggcaaaaga	ccgcctactg ttatttctct	1500	
gttgctgcaa cactttcgtc tcccgaatta tcagatgcgc	gtatttcatg ggccaaaaat	1560	
ggcatattaa ctacagtagt tgatgacttt tttgatatcg	gtggtacaat cgatgaattg	1620	
accaacctga ttcaatgtgt tgaaaaatgg aatgtagatg	tcgacaagga ttgttgttca	1680	
gagcatgttc ggattttatt tttagcatta aaagatgcaa	tctgttggat tggagatgaa	1740	
gctttttaa	at ggcaagcgcg cgatgtaact agccatgtta	ttcaaacttg gttggaacta	1800
atgaatagta	tgttgagaga agctatatgg acaagagatg	cttatgtgcc aacattaaat	1860
gaatatatgg	aaaacgctta cgtgtcattt gcattaggcc	cgattgtcaa gccggctatt	1920
tactttgtgg	ggcccaaatt atcagaggag attgttgaaa	gctctgaata tcataatcta	1980
tttaagctaa	tgagcacgca gggtcgactt ctaaacgata	tccatagctt caagagggaa	2040
tttaagggaag	gcaaattaaa cgcggtagca ttgcatttga	gtaacggaga aagtgggaaa	2100
gtggaagaag	aggttgtgga ggagatgatg atgatgatta	aaaacaagag gaaagaatta	2160
atgaaattaa	tttttgaaga aaatggtagc attgttccta	gagcttgtaa agatgcattt	2220
tggaacatgt	gtcacgtgtt gaattttttt tacgcaaacg	atgacggggt tactggaaac	2280
acgattcttg	atactgtgaa ggacatcatt tacaaccggt	tgggtgcttgt gaatgaaaat	2340
gaagaacaaa	ggtaa	2355	

<210> 46
 <211> 1773
 <212> DNA
 <213> Zea mays

<400> 46
 atggccatgc cagtgaagct gactcctgcc tccctctcgc tgaaggcggc ctgctgccgc 60
 tttagctccg gagggcatgc gctgcgcttc ggctcgtcgc taccgtgctg gaggaggacg 120
 ccgacgcaac ggagcacgtc gtcgtctacg acgcgccctg cggctgaggt tagctctggc 180
 aaaagcaagc agcacgatca agaagcatcg gagggctacga taagacagca gctccagcta 240
 gtcgatgtgc ttgagaacat ggggatttct cggcattttg ctgctgaaat caaatgcatc 300
 cttgacagga catacagaag ttggttacag agacatgagg aaattatgct ggacacaatg 360
 acctgtgcga tggcatttct tattctaagg ttgaatggat acaatgtctc ttctgatgag 420
 ttgtatcatg ttgttgaagc ttccggactc cataattcac ttggaggata tctcaatgat 480
 acaagaacct tgtagaatt acacaaggcc tcgacagtta gtatctctga agatgagtct 540
 atcctggata gcataggctc aaggtcacgt accttactga gggaacaact agagtctggc 600
 ggtgctctac gaaaaccttc actctttaaa gaggtggaac atgctctgga cggtccttc 660
 tacaccacat tggaccgtct acaccatagg tggaacatcg aaaatttcaa tattatagag 720
 cagcacatgc tagagacacc atacttgtca aatcaacata ccagtagaga tattctagcg 780
 ttgagtatta gagacttcag ttcctctcag ttacttacc agcaagaact tcaacatctt 840
 gaaagctggg tgaaagagtg caggtagac cagctacaat ttgcgcgaca gaagttggca 900
 tacttctact tgtctgctgc tggcaccatg ttctctcctg agctgtctga tgctcgaact 960
 ttgtgggcca aaaatggtgt gctcacaact attgttgacg acttcttga tgttgcggga 1020
 tcaaaagaag aacttgaaaa ccttgtcatg ttggttgaga tgtgggacga gcatcacaaa 1080
 gttgagttct actcagaaca agtagagatt atattttctt caatttatga ctcagttaac 1140
 caacttgggtg aaaaggcttc tttggtacaa gaccgcagta ttaccaaaaca cctagtagaa 1200
 atatggttgg atttgctaaa gtctatgatg acagaggtag agtggcgttt gagcaaatat 1260
 gtgccaacag agaaggaata catgataaat gcatctttaa tatttggact aggccccatt 1320
 gtattgccag cattatatat ttgtgggcca aagatctcag agtctattgt taaagatcca 1380
 gaatatgatg aattgttcaa actgatgagc acatgtggtc gcctcttgaa tgatgttcag 1440
 acttttgaga gggagtacaa cgagggcaag ttgaatagtg tttctctcct cgttcttcat 1500
 ggtggcccca tgtccatatc agacgccaaa aggaaattac agaagcccat agacacatgc 1560
 agaagagacc tcctaagttt agttcttcgt gaagaaagtg ttgttcctag gccctgcaag 1620
 gaattatatt ggaaaatgtg caaggtgtgc tacttcttct actcgacgac ggatgggttt 1680
 agtcacaag tggagagggc taaagaagtg gatgcggtga tcaatgagcc actaaagcta 1740
 caaggaagtc atacgctggg gtctgatgtg tga 1773

<210> 47

<211> 2232

<212> DNA

<213> *Populus trichocarpa*

<400> 47

atgcagaact	ttcatggaac	taaggaaagg	atcaagaaga	tgtttgataa	gattgaattg	60
tcagtgtctt	catatgacac	tgcttggtg	gcaatggtcc	catctccaga	ttgtccggaa	120
actccttggt	ttccagagtg	cacaaaatgg	atcttgga	atcaacttgg	tgatggctcc	180
tggagtcttc	ctcatggcaa	tccattatta	gttaaggatg	ctctttcatc	tacattagcg	240
tgcatccttg	cattgaagcg	atggggatc	ggtgaagaac	aaataaataa	aggccttcga	300
tttattgagt	tgaattccgc	ttcagttacg	gataacgagc	aacataaacc	aattggattt	360
gatataatat	ttcctggcat	gattgaatat	gccatagatt	tggatttgaa	cctccctttg	420
aagccgacag	atataaattc	catgctccac	aggagggctt	tggagcttac	aagtggcggg	480
ggcaagaact	tggagggag	aagagcctac	ttagcatatg	tttcggaagg	aattggaaaa	540
ttacaggatt	gggaaatggc	catgaaatat	caaagaaaga	atggatcact	gttcaattca	600
ccatccacca	cagcagctgc	ctttattcat	attcaagatg	ctgagtgtct	ccattatatt	660
cgttcactct	tacagaagtt	tgggaatgca	gttccaacca	tttatccttt	ggatatatat	720
gctcgtcttt	ctatggttga	tgctcttgaa	aggttgggaa	tcgatcggca	ttttaggaag	780
gaaagaaaat	ttgttttggg	cgaacatac	cgattttggg	tgcaggggga	ggaagagata	840
ttttctgata	atgccacttg	tgctttggca	tttaggatat	tacgtttgaa	cggatatgat	900
gtctctctag	aagatcattt	ctctaattca	ctgggaggat	atttgaagga	ttcgggagct	960
gccttagagt	tgtacagagc	tctgcagcta	agttatccag	atgaatcact	tctggaaaaa	1020
caaaattctc	ggacaagcta	tttcctgaaa	cagggattat	ccaacgtttc	actttgtgga	1080
gataggcttc	gtaaaaatat	tatcggagag	gtgcatgatg	ctctcaattt	ttctgaccat	1140
gcaaatttgc	aacgcttagc	tatcagaaga	agaattaaac	attatgctac	agatgatacg	1200
aggattttga	aaacttcgta	tcgttggtcg	actattggta	accaggattt	tctcaaattg	1260
gctgtagaag	acttcaatat	ctgtcaatca	atacagcgtg	aagaatttaa	acatatcgag	1320
agggtgggtt	tagagaggag	actggacaag	ctaaagtttg	ctaggcagaa	ggaggcctac	1380
tgttacttct	ctgctgcagc	aactctcttc	gctccagaac	tatctgatgc	acgcatgtca	1440
tgggcaaaaa	atggtgtgct	tactactgtt	gttgatgact	tctttgatgt	tgggtggttct	1500
gaagaagaac	tggtaaacct	tattgaattg	attgagaggt	gggatgtcaa	tggcagtgct	1560
gatttttggt	ctgaggaagt	tgagatcata	tattcggcaa	ttcacagcac	tataagtgag	1620
ataggagaca	aatctttcgg	atggcaagga	cgcatgtga	aaagtcaggt	tatcaagatt	1680
tggttggatt	tgctcaaatc	catgttgaca	gaagcacaat	ggtcaagtaa	caaatcagtg	1740
ccgacccttg	atgaatatat	gacaactgca	catgtatcgt	tcgctctagg	gcctattggt	1800
cttcagctc	tgtattttgt	ggggcctaag	ctttcagagg	aggttgctgg	acatcctgaa	1860
ttgcttaatc	tatacaaggt	tacgagcact	tgcgggcgtc	tgctcaatga	ctggagaagc	1920
tttaagagag	aatctgaaga	agggaaattg	aatgccatct	cattgtacat	gattcacagc	1980
ggtggtgctt	caactgaaga	agagaccatc	gaacatttta	aaggattgat	cgacagccag	2040
agaagacaat	tgcttcaatt	agttttgcag	gaaaaggata	gtataattcc	tagaccctgc	2100

aaggatttgt tttggaacat gataaaatta ttgcacacgt tctacatgaa ggatgatgga	2160
ttcacttcaa acgagatgag aaatgttgct aaggcaataa taaatgaacc catctctcta	2220
gatgaattat aa	2232

<210> 48
 <211> 2952
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 48	
atggaatttg atgaaccatt ggttgacgaa gcaagatctt tagtgcagcg tactttacaa	60
gattatgatg acagatacgg cttcggctact atgtcatgtg ctgcttatga tacagcctgg	120
gtgtcttttag 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 acacgtcggg	420
tttgagataa ttgttcctgc aatgctagac ccattagaag ccgaagatcc atctctagtt	480
ttcgattttc cagctaggaa acctttgatg aagattcatg atgctaagat gagtagattc	540
aggccagaat acttgatggg caaacaacca atgaccgcct tacattcatt agaggctttc	600
ataggcaaaa tcgacttcga taaggtaaga caccaccgta cccatggggtc tatgatgggt	660
tctccttcat ctaccgcagc ctacttaatg cacgcttcac aatgggatgg tgactcagag	720
gcttacctta gacacgtgat taaacacgca gcagggcagg gaactgggtgc tgtaccatct	780
gctttcccat caacacattt tgagtcattt tggattctta ccacattggt tagagctgga	840
ttttcagctt ctcatcttgc ctgtgatgag ttgaacaagt tggctgagat acttgagggc	900
tcattcgaga aggaaggtgg ggcaatcggg tacgctccag ggtttcaagc agatgttgat	960
gatactgcta aaacaataag tacattagca gtccttgga gagatgctac accaagacaa	1020
atgatcaagg tatttgaagc taatacacat tttagaacat accctgggtga aagagatcct	1080
tctttgacag ctaattgtaa tgctctatca gccttactac accaaccaga tgcagcaatg	1140
tatggatctc aaattcaaaa gattacaaa tttgtctgtg actattgggtg gaagtctgat	1200
ggtaagatta aagataagtg gaacacttgc tacttgtagc catctgtctt attagttgag	1260
gttttggttg atcttggttg 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	tttgtttagca	1620
gcaagatggg	ctgctaagtc	tccttttaggc	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
ttccttttacg	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	ctatacaag	tgcctctgta	tgggatagaa	aactacttgc	tagagagatg	2280
aaggcttact	tacttgctca	tatccaacaa	gcagaagatt	caactccatt	gtctgaattg	2340
aaagatgtgc	ctcaaagac	tgatgtaaca	agagtttcta	catctactac	taccttcttt	2400
aactgggtta	gaacaacttc	cgagaccat	atatcctgcc	catactcctt	ccactttgta	2460
gcatgccatc	taggcgcagc	attgtcacct	aaagggctta	acggtgattg	ctatccttca	2520
gctggtgaga	agttcttggc	agctgcagtc	tgcagacatt	tggccaccat	gtgtagaatg	2580
tacaacgatc	ttggatcagc	tgaacgtgat	tctgatgaag	gtaatttgaa	ctccttggac	2640
ttccctgaat	tcgccgattc	cgaggaac	ggagggatag	aaattcagaa	ggccgctcta	2700
ttaaggttag	ctgagtttga	gagagattca	tacttagagg	ccttccgtcg	tttacaagat	2760
gaatccaata	gagttcacgg	tccagccggt	ggatgatgaag	ccagattgtc	cagaaggaga	2820
atggcaatcc	ttgaattctt	cgccagcag	gtagatttgt	acggatcaagt	atacgatcatt	2880
agggatatatt	ccgctcgat	tcctaaaaac	gaggttgaga	aaaagagaaa	attggatgat	2940
gctttcaatt	ga					2952

<210> 49

<211> 2646

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 49

atggcttcta	gtacacttat	ccaaaacaga	tcatgtggcg	tcacatcatc	tatgtcaagt	60
tttcaaatct	tcagagggtca	accactaaga	tttcctggca	ctagaacccc	agctgcagtt	120
caatgcttga	aaaagaggag	atgccttagg	ccaaccgaat	ccgtactaga	atcatctcct	180
ggctctgggt	catatagaat	agtaactggc	ccttctggaa	ttaaccctag	ttctaacggg	240
cacttgcaag	agggttcctt	gactcacagg	ttaccaatac	caatggaaaa	atctatcgat	300

aacttccaat ctactctata tgtgtcagat atttggctctg aaacactaca gagaactgaa	360
tgtttgctac aagtaactga aaacgtccag atgaatgagt ggattgagga aattagaatg	420
tacttttagaa 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
caaacgttg aaagaggaat tcagttccta caatctaaca tatacaagat ggaggaagat	720
gacgctaadc atatgccaat aggattcgaa atcgtattcc ctgctatgat ggaagatgcc	780
aaagcattag gtttggtttt gccatacgat gctactattt tgcaacagat ttcagccgaa	840
agagagaaaa agatgaaaa gatcccaatg gcaatgggtg acaaataccc aaccacttta	900
cttcactcct tagaaggctt gcatagagaa gttgattgga ataagttggtt acaattacaa	960
tctgaaaatg gtagttttct ttattcacct gcttcaaccg catgcgctt 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 gctggtgaa tcgagtataa cttgaccttc	1560
ccatggatg cctctttgcc tagattagaa cataggacat acttagatca atatggaatc	1620
gatgatatct ggataggcaa atctttatac aaaatgcctg ctggttacc aaagattttc	1680
ctaaagtgg caaaggcaga cttaacatg tgtcaagctc tacacaaaaa ggaattggaa	1740
caagtgataa agtggaacgc gtcctgtcaa ttcagagatc ttgaattcgc cagacaaaaa	1800
tcagtagaat gctattttgc tgggtgcagcc acaatgttcg aaccagaaat ggttcaagct	1860
agattagtct gggcaagatg ttgtgtattg acaactgtct tagacgatta ctttgaccac	1920
gggacacctg ttgaggaact tagagtgttt gttcaagctg tcagaacatg gaatccagag	1980
ttgatcaacg gtttgccaga gcaagctaaa atcttgttta tgggcttata caaacagtt	2040
aacacaattg cagaggaagc attcatggca cagaaaagag acgtccatca tcatttgaaa	2100
cactattggg acaagttgat aacaagtgcc cttaaaggagg ccgaatgggc agagtcaggt	2160
tacgtcccaa catttgatga atacatggaa gtagctgaaa tttctgttgc tctagaacca	2220
attgtctgta gtacctgtt ctttgcggtt catagactag atgaggatgt tctagatagt	2280
tacgattacc atctagttat gcatttggtg aacagagtcg gtagaatctt gaatgatata	2340

caaggcatga agagggaggc ttcacaaggt aagatctcat cagttcaaat ctacatggag	2400
gaacatccat ctgttccatc tgaggccatg gcgatcgctc atcttcaaga gttagttagat	2460
aattcaatgc agcaattgac atacgaagtt cttaggttca ctgcggttcc aaaaagttgt	2520
aagagaatcc acttgaatat ggctaaaatc atgcatgcct tctacaagga tactgatgga	2580
ttctcatccc ttactgcaat gacaggattc gtcaaaaagg ttcttttcga acctgtgcct	2640
gagtaa	2646

<210> 50
 <211> 2952
 <212> DNA
 <213> *Phomopsis amygdali*

<400> 50	
atggagttcg atgaaccact tgtggacgag gcgaggtcct tgggtccaaag aacctgcaa	60
gattatgacg accgctatgg ctttggcact atgagctgtg cggcctatga cacagcatgg	120
gtatcgctgg tgactaaaac agtcgatggg cgtaaacaaat ggttgttccc tgagtgttct	180
gaattttctc tagaaacgca gtccgatgct ggcggctggg aaatcggcaa cagcgcaccc	240
atcgatggga tccttaacac tgctgcttca ctgctggcat tgaagcgcca cgtccaaaca	300
gagcagatta ttcagccgca acacgaccat aaagacctgg ccgggctgtc ggaaagagcg	360
gcggcgtctt tgcgagcaca gttggcggct ctggatgtgt cgacaacgga gcatgtgggc	420
ttcgaaatca tcgtcccggc catgctcgac cctctcgagg ccgaagaccc gtctttggtg	480
ttcgactttc cagcacgcaa accactgatg aagatccacg acgctaagat gtcgcatc	540
cgaccagagt acctctacgg taaacagccg atgacggcat tgcattcgct cgaggccttt	600
atcgggaaaa tagacttcga caaagtacgg catcacagga cacacggttc gatgatgggg	660
tcgccctcgt cgacggctgc atacctgatg catgcttctc agtgggacgg cgactctgag	720
gcctatctac gccatgtcat caagcacgca gctggccagg gcaccggagc tgttccgagt	780
gcatttcctt cgacgcattt cgagtcttct tggattttga caacattgtt tcgagctggg	840
ttctcagcct ctcatctagc atgcgacgaa ttaaacaagc tggtagagat cctcgaaggc	900
tcatttgaga aagaaggggg agccatcggg tatgctcctg ggtttcaagc agatgtggat	960
gataccgcaa agaccatctc cactttggct gtgcttggga gagatgccac tccccggcaa	1020
atgatcaagg tttttgaagc caatacacac tttcggactt accctgggtga aagagatcca	1080
agcttgactg ccaattgcaa cgcgtctctg gctcttcttc accagccaga cgcagcaatg	1140
tacggcagcc agatccagaa gatcaciaag tttgtttgtg actactggtg gaaaagtgac	1200
ggcaaaatca aggacaagtg gaatactgc tacttgatc catcggtcct cctcgtcgag	1260
gtgttagtag accttgtgtc cctgttggag caaggaaagc taccgcagct gctggatcag	1320
gagctgcaat acagggtcgc cattacgtta ttccaggcct gcttgcgacc gctacttgat	1380
caagatgctg aagggtcatg gaacaaatcc attgaagcca cagcctacgg cattctaadc	1440
cttacggagg cgcggcgagt atgctttttt gaccgtctga gtgagcctct gaatgaggct	1500

attcgacgcg ggattgcggtt tgcagattcg atgagcggta ctgaagctca gctgaattat	1560
atatggatcg agaaagtgag ctacgcacct gctcttctga ccaaatacata cctcctcgca	1620
gctcgggtggg cggcaaagtc cccgcttggc gcttccggtg gatccagcct ttggacgcct	1680
ccaagagaag gcttggataa gcacgtccgt ctattccacc aggagagct cttcaggctcg	1740
ttgccggagt gggagctgcg cgcgtccatg atcgaggcag ccctgttcac tcctttgctg	1800
cgtgcgcata ggctggatgt atttccacgc caagacgtcg gcgaggacaa gtacctggac	1860
gttgtgacct tcttctggac ggccgccaat aaccgcgac gcacgtacgc atccactctg	1920
tttctgtatg acatgtgctt tatcgccatg ctttaacttcc agctggatga gttcatggag	1980
gctacagcgg gaatcctctt ccgggacct atggatgatt tgcgccaact catccacgac	2040
ctgcttgccg aaaagacgag cccaagtca tcgggcagaa gtagccaagg aaccaagac	2100
gcggactcgg gcacgaaga agacgtttct atgagcgact cagcgtcaga ctcccaggac	2160
cgcagccctg aatacgacct ggtcttctct gcgctctcta ccttcaccaa acatgtcctg	2220
cagcaccctt caatccagtc agccagtgtc tgggatagga aactactcgc tcgcgagatg	2280
aaagcatacc tcttagctca tattcaacag gctgaggaca gcacgccctt gagtgagctc	2340
aaggacgtcc ctcaaaaaac tgacgtgaca cgcgtctcaa cgtccacaac gactttcttc	2400
aactgggtac gcacaacatc cgcagaccac atatcctgcc catattcatt ccatttcgtg	2460
gcgtgtcacc tcggcgccgc gctgagcccc aagggcagca acggcgactg ttaccctgca	2520
gccggtgaaa agttcctcgc ggccgccgta tgccgccatt tggccacgat gtgccgcatg	2580
tacaatgact tgggatcggc ggagcgcgac agtgacgagg gaaatttgaa ttcactcgac	2640
tttcccgagt tcgccgactc agcggggaat ggtgggattg agatccagaa agctgccttg	2700
ctcaggctgg ccgagttcga acgcgactcg tatctcgagg ctttccggcg acttcaggat	2760
gaaagcaacc gcgttcacgg accggctggt ggggatgaag ccagactcag caggcggcgc	2820
atggccatcc ttgagttctt tgcccagcag gtggacttgt atggccaggt ctacgttatt	2880
cgcgatatct cggccaggat tccaaagaac gaggttgaga agaaaaggaa actagatgat	2940
gctttcaatt ag	2952

<210> 51

<211> 2646

<212> DNA

<213> *Physcomitrella patens*

<400> 51

atggcttcca gcaccttgat acagaatcgc tcttgtggcg ttacgtcaag catgtcttcc	60
tttcagattt ttgaggggca acctctacgt tttccaggca ctagaactcc tgctgcagtt	120
caatgcctaa agaagcgtcg atgtttgcga cctactgaat cagtcctcga gagctctcct	180
ggtagcggat cttacaggat tgtaactgga ccctccggca tcaatccttc ttcaaacggc	240
cacttgcaag aggggtccct tactcacaga cttccgatac ccatggaaaa atccattgat	300
aacttccagt ctactttgta cgtatcagac atatggtcag aaaccttgca aagaacggaa	360

tgtttgttgc	aggtgactga	gaatgtacag	atgaacgagt	ggattgagga	aatcagaatg	420
tacttccgaa	atatgacact	gggggaaata	tccatgtctc	catacgacac	agcttgggta	480
gcgcgagtgc	cagcgctgga	tggtcacat	ggccctcagt	tccatcggtc	tttgacgtgg	540
attattgata	atcagctccc	ggatggcgat	tgggggtgaac	cgctctcttt	ccttggatac	600
gatcgcgttt	gcaacactct	cgctgtgta	attgccctga	aaacttgggg	tgttggggct	660
cagaacgtag	agcgtggaat	ccagtttctg	caatctaaca	tctacaaaat	ggaggaagat	720
gacgccaatc	atatgccgat	tggatttgag	attgtcttcc	cagcgatgat	ggaagatgcc	780
aaggcactgg	gactggattt	accatacgat	gccactatct	tgcaacaaat	ctcggctgaa	840
agagagaaga	aaatgaaaaa	gattcctatg	gcgatgggtg	acaagtaccc	cactactttg	900
ctgcattctc	tggaaggcct	gcaccgggaa	gtggactgga	acaagctcct	ccagctacag	960
tccgagaatg	gctcctttct	gtattcacc	gcattccactg	catgcgcact	tatgtacaca	1020
aaagatgtga	agtgtcttga	ctacttgaac	cagctcctca	tcaagttcga	ccacgcttgt	1080
ccaaacgtgt	accccgttga	tctcttcgag	cgtttgtgga	tggtagaccg	cctacaaagg	1140
ctgggaatat	cccgtacttt	cgagcgagaa	atcagagact	gtctacaata	tgtataaccga	1200
tactggaagg	attgtggtat	tggctgggca	agcaattcgt	ccgtgcagga	cgtggacgac	1260
acggccatgg	ccttccgcct	tctccgcaca	cacggattcg	acgtcaagga	ggactgcttc	1320
agacagtttt	tcaaagatgg	tgagttcttc	tgcttcgccg	gccagtccag	ccaagccgtc	1380
acgggaatgt	tcaacctcag	cagagcatcg	caaacgctct	tcccagggga	atcactccta	1440
aaaaaggcca	gaaccttttc	cagaaacttt	ttgagaacca	agcatgaaaa	caatgaatgc	1500
ttcgacaagt	ggataatcac	gaaggatcta	gcgggcgagg	tggaatacaa	tctcacattc	1560
ccctggtatg	ctagccttcc	tcgtcttgag	catcgcacct	acttggacca	atatgggatt	1620
gatgatattc	ggattggcaa	gtcgtcttac	aaaatgccgg	ccgtcaccaa	cgaagtgttt	1680
ctcaaattgg	ccaaagccga	cttcaacatg	tgccaagctc	ttcacaagaa	ggaactcgag	1740
caggtcatca	aatggaatgc	cagctgccaa	tttagagacc	tcgagtttgc	tagacagaaa	1800
tccgtggagt	gctacttcgc	aggcgtgca	accatgtttg	agcccgaaat	ggtgcaggcg	1860
aggctcgttt	gggcacgctg	ttgcgtgctc	accaccgttc	tagacgatta	cttcgatcac	1920
ggtacacctg	tggaagagct	tcgggttttt	gtgcaggccg	taaggacttg	gaatcccag	1980
ctcatcaacg	gactacctga	gcaagccaag	attctcttta	tgggactgta	caagactgtg	2040
aacactatcg	ccgaggaggc	attcatggca	cagaaacgag	acgtacatca	tcatctcaag	2100
cattactggg	acaaattgat	cacttcagct	ttgaaagaag	ccgaatgggc	agagtccggc	2160
tacgtcccca	ccttcgacga	gtatatggaa	gtcgtgaaa	tctccgtcgc	actagagccc	2220
attgtatgta	gcactctctt	cttcgccggc	cataggctcg	atgaggatgt	gcttgacagt	2280
tatgactacc	atcttgtcat	gcattctcgtc	aaccgcgtag	gtcgcaccc	caacgacatc	2340
caaggaatga	agaggggaagc	cagccaaggg	aagatatcga	gcgtgcagat	ctacatggag	2400
gagcatccaa	gtgtgccttc	agaggccatg	gccatcgtc	atctgcagga	attggtcgac	2460

aactccatgc aacagctgac atacgaagtg ctgcgcttca ctgcagtccc gaagtcctgt	2520
aagagaatcc attttaaocat ggcgaagatc atgcacgctt tctacaagga cactgatggg	2580
ttttcgtcac tgacagccat gacagggttt gtgaagaagg tgctcttcga gccagtacct	2640
gaatag	2646

<210> 52

<211> 1542

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 52

atggatgctg tgacgggttt gttaactgtc ccagcaaccg ctataactat tgggtggaact	60
gctgtagcat tggcggtagc gctaattctt tggtagctga aatcctacac atcagctaga	120
agatcccaat caaatcatct tccaagagtg cctgaagtcc cagggtgtcc attgttagga	180
aatctgttac aattgaagga gaaaaagcca tacatgactt ttacgagatg ggcagcgaca	240
tatggaccta tctatagtat caaaactggg gctacaagta tggttgtggg atcatcta	300
gagatagcca aggaggcatt ggtgaccaga ttccaatcca tatctacaag gaacttatct	360
aaagccctga aagtacttac agcagataag acaatggctg caatgtcaga ttatgatgat	420
tatcataaaa cagttaagag acacatactg accgccgtct tgggtcctaa tgcacagaaa	480
aagcatagaa ttacacagaga tatcatgatg gataacatat ctactcaact tcatgaattc	540
gtgaaaaaca acccagaaca ggaagaggta gaccttagaa aaatctttca atctgagtta	600
ttcggccttag ctatgagaca agccttagga aaggatgttg aaagtgtgta cgttgaagac	660
ctgaaaatca ctatgaatag agacgaaatc tttcaagtcc ttgttgttga tccaatgatg	720
ggagcaatcg atgttgattg gagagacttc tttccatacc taaagtgggt cccaaacaaa	780
aagttcgaat atactattca acaaatgtac atcagaagag aagctgttat gaaatcttta	840
atcaaagagc acaaaaagag aatagcgtca ggcgaaaagc taaatagtta tatcgattac	900
cttttatctg aagctcaaac ttaaccgat cagcaactat tgatgtcctt gtgggaacca	960
atcattgaat cttcagatac aacaatggc acaacagaat gggcaatgta cgaattagct	1020
aaaaacccta aattgcaaga taggtgttac agagacatta agtccgtctg tggatctgaa	1080
aagataaccg aagagcatct atcacagctg cttacatta cagctatctt 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 gtggtgaagag agtttgtgct	1380
ggttccttgc aagccctttt aactgcatct attgggattg ggagaatggg tcaagagttc	1440
gaatggaaac tgaaggatat gactcaagag gaagtgaaca cgataggcct aactacacaa	1500

atgttaagac cattgagagc tattatcaaa cctaggatct aa 1542

<210> 53
 <211> 1530
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 53
 atggcatctt tctctatgat ttcaatcttg ttgggatttg ttatttcttc tttcatcttc 60
 atctttttct tcaaaaagtt acttagtttt agtaggaaaa acatgtcaga agtttctact 120
 ttgccaaagt ttccagtagt gcctggtttt ccagttattg ggaatttggt gcaactaaag 180
 gagaaaaagc ctcataaaac tttcactaga tggtcagaga tatatggacc tatctactct 240
 ataaagatgg gttcttcatc tcttattgta ttgaacagta cagaaactgc taaggaagca 300
 atggctacta gattttcatc aatatctacc agaaaattgt caaacgccct aacagttcta 360
 acctgcgata agtctatggt cgccacttct gattatgatg acttccacaa attagttaag 420
 agatgtttgc taaatggact tcttgggtgct aatgctcaaa agagaaaaag acactacaga 480
 gatgctttga ttgaaaatgt gagttccaag ctacatgcac acgctagaga tcatccacaa 540
 gagccagtta acttttagagc aattttcgaa cacgaattgt ttggtgtagc attaaagcaa 600
 gccttcggta aagacgtaga atccatatac gtcaaggagt taggcgtaac attatcaaaa 660
 gatgaaatct ttaaggtgct tgtacatgat atgatggagg gtgcaattga tgtagattgg 720
 agagatttct tcccatatct gaaatggatc cctaataagt cttttgaagc taggatacaa 780
 caaaagcaca agagaagact agctgttatg aacgcactta tacaggacag attgaagcaa 840
 aatgggtctg aatcagatga tgattgttac cttaacttct taatgtctga ggctaaaaca 900
 ttgactaagg aacagatcgc aatccttgct tgggaaacaa tcattgaaac agcagatact 960
 accttagtca caactgaatg ggccatatac gagctagcca aacatccatc tgtgcaagat 1020
 aggttggtga aggagatcca gaacgtgtgt ggtggagaga aattcaagga agagcagttg 1080
 tcacaagttc cttaccttaa cggcgttttc catgaaacct tgagaaaata ctcacctgca 1140
 ccattagtct ctattagata cgcccacgaa gatacacaaa tcggtggcta ccatgttcca 1200
 gctgggtccg aaattgctat aaacatctac ggggtgcaaca tggacaaaaa gagatgggaa 1260
 agaccagaag attggtggcc agaaagattc ttagatgatg gcaaatatga aacatctgat 1320
 ttgcataaaa caatggcttt cggagctggc aaaagagtgt gtgccggtgc tctacaagcc 1380
 tccctaattg ctggtatcgc tattggtaga ttggtccaag agttcgaatg gaaacttaga 1440
 gatggtgaag aggaaaatgt cgatacttat gggttaacat ctcaaaagtt ataccacta 1500
 atggcaatca tcaatcctag aagatcctaa 1530

<210> 54
 <211> 1578

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 54

atgagtaagt ctaatagtat gaattctaca tcacacgaaa ccctttttca acaattggtc	60
ttgggttttg accgtatgcc attgatggat gttcactggg tgatctacgt tgctttcggc	120
gcatggttat gttcttatgt gatacatgtt ttatcatctt cctctacagt aaaagtgcc	180
gttggtggat acaggtctgt attcgaacct acatgggtgc ttagacttag attcgtctgg	240
gaaggtggct ctatcatagg tcaagggtag aataagttaa aagactctat tttccaagtt	300
aggaaattgg gaactgatat tgtcattata ccacctaact atattgatga agtgagaaaa	360
ttgtcacagg acaagactag atcagttgaa cctttcatta atgattttgc aggtcaatac	420
acaagaggca tggttttctt gcaatctgac ttacaaaacc gtgttatata acaaagacta	480
actccaaaat tggtttcctt gaccaagggtc atgaaggaag agttggatta tgctttaaca	540
aaagagatgc ctgatatgaa aaatgacgaa tgggtagaag tagatatcag tagtataatg	600
gtgagattga tttccaggat ctccgccaga gtctttctag ggcctgaaca ctgtcgtaac	660
caggaatggt tgactactac agcagaatat tcagaatcac ttttcattac agggtttatc	720
ttaagagttg tacctcatat ctaagacca ttcacgccc ctctattacc ttcatacagg	780
actctactta gaaacgtttc aagtggtaga agagtcacgt gtgacatcat aagatctcag	840
caaggggatg gtaacgaaga tatactttcc tggatgagag atgctgccac aggagaggaa	900
aagcaaatcg ataacattgc tcagagaatg ttaattcttt ctttagcatc aatccacact	960
actgcatgga ccatgacaca tgccatgtac gatctatgtg cttgccctga gtacattgaa	1020
ccattaagag atgaagttaa atctgttggt ggggcttctg gctgggacaa gacagcgta	1080
aacagatttc ataagttgga ctcttccta aaagagtcac aaagattcaa cccagtattc	1140
ttattgacat tcaatagaat ctaccatcaa tctatgacct tatcagatgg cactaacatt	1200
ccatctggaa cacgtattgc tgttccatca cacgcaatgt tgcaagattc tgcacatgtc	1260
ccaggtccaa cccacctac tgaatttgat ggattcagat atagtaagat acgttctgat	1320
agtaactacg cacaaaagta cctattctcc atgaccgatt cttcaaacat ggctttcgg	1380
tacggcaagt atgcttgacc aggtagattt tacgctgcta atgagatgaa actaacatta	1440
gccattttgt tgctacaatt tgagttcaaa ctaccagatg gtaaagggtcgc tcctagaaat	1500
atcactatcg attctgatat gattccagac ccaagagcta gactttgcgt cagaaaaaga	1560
tcacttagag atgaatga	1578

<210> 55

<211> 1500

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 55

atggaagatc ctactgtctt atatgcttgt cttgccattg cagttgcaac tttcgttggt	60
agatggtaca gagatccatt gagatccatc ccaacagttg gtggttccga tttgcctatt	120
ctatcttaca tcggcgcaact aagatggaca agacgtggca gagagatact tcaagagga	180
tatgatggct acagaggatc tacattcaaa atcgcgatgt tagaccgttg gatcgtgatc	240
gcaaatggct ctaaactagc tgatgaagtc agacgtagac cagatgaaga gttaaacttt	300
atggacggat taggagcatt cgtccaaact aagtacacct taggtgaagc tattcataac	360
gatccatacc atgtcgatat cataagagaa aaactaaca gaggccttcc agccgtgctt	420
cctgatgtca ttgaagagtt gacacttgcg gttagacagt acattccaac agaagggtgat	480
gaatgggtgt ccgtaaaactg ttcaaaggcc gcaagagata ttgttgctag agcttcta	540
agagtctttg taggtttgcc tgcttgaga aaccaagggt acttagattt ggcaatagac	600
tttacattgt ctgttgtaa ggatagagcc atcatcaata tgtttccaga attgttgaag	660
ccaatagttg gcagagttgt aggtaacgcc accagaaatg ttcgtagagc tgttcctttt	720
gttgctccat tgggtggagga aagacgtaga cttatggaag agtacggtga agactggtct	780
gaaaaaccta atgatatgtt acagtggata atggatgaag ctgcatccag agatagttca	840
gtgaaggcaa tcgcagagag attgttaatg gtgaacttcg cggctattca tacctcatca	900
aacactatca ctcatgcttt gtaccacctt gccgaaatgc ctgaaacttt gcaaccactt	960
agagaagaga tcgaaccatt agtcaaagag gagggctgga ccaaggctgc tatgggaaaa	1020
atgtggtggt tagattcatt tctaagagaa tctcaaagat acaatggcat taacatcgta	1080
tctttaacta gaatggctga caaagatatt acattgagtg atggcacatt tttgcaaaa	1140
ggtactctag tggccgttcc agcgtattct actcatagag atgatgctgt ctacgctgat	1200
gccttagtat tcgatccttt cagattctca cgtatgagag cgagagaagg tgaaggatca	1260
aagcaccagt tcgttaatac ttcagtcgag tacgttccat ttggtcacgg aaagcatgct	1320
tgtccaggaa gattcttcgc cgcaaacgaa ttgaaagcaa tgttggtcta cattgttcta	1380
aactatgatg taaagttgcc tgggtgacgg aaacgtccat tgaacatgta ttggggcca	1440
acagttttgc ctgcaccagc aggccaaagta ttgttcagaa agagacaagt tagtctataa	1500

<210> 56

<211> 1542

<212> DNA

<213> *Stevia rebaudiana*

<400> 56

atggatgccg tcaccggttt gctgacagtt ccggcaaccg caataaccat cggcgggtacg	60
gccgtcgcac tcgccgtcgc tctgatattc tggtagctca aaagctacac atctgcacgc	120
aggagccaat caaaccatct ccctcgggtt cccgaggtac ctggtgtgcc attattgggg	180
aatttattgc agttgaagga gaagaaacct tacatgactt ttacaagatg ggcggcaact	240

tatggtccga	tttattcgat	taaaaccgga	gcaacttcta	tgggtggtcgt	cagttcaaat	300
gaaattgcaa	aggaggcatt	ggttaccaga	tttcaatcta	tctcaaccag	aaacctatca	360
aaggcattaa	aggtttctcac	agcagataaa	accatggtgg	cgatgagtga	ttatgatgat	420
tatcataaga	ctgtcaaacg	ccatatactg	accgctgttt	tgggaccaa	tgctcagaag	480
aaacaccgca	tccatagggg	catcatgatg	gataatatat	caacccaact	tcatgaattt	540
gttaaaaata	atcctgaaca	agaggaagtg	gatctaagga	aaatattcca	atccgaactt	600
tttggattag	ctatgagaca	agcattggga	aaggatgtgg	agagcttata	tgttgaggat	660
cttaaaatca	ccatgaaccg	agacgagata	tttcagggtat	tgggtgttga	cccgatgatg	720
ggtgcaattg	acgtcgactg	gagagatttc	ttcccgtatc	taaagtgggt	cccgaataaa	780
aagtttgaag	acacgatcca	acaaatgtat	atccggagag	aagctgtgat	gaagtctctt	840
attaaagaac	ataaaaaacg	tattgcatcc	ggagagaaat	taaacagcta	cattgattac	900
ttgctatcgg	aagcaciaac	gttaaccgat	caacaactac	ttatgtctct	atgggaacct	960
attattgaat	catcagacac	cactatgggt	acaactgaat	gggctatgta	tgaacttgca	1020
aaaaacccca	aacttcagga	tcgtttgtat	cgggatatca	aaagtgtttg	cgggtcagag	1080
aagattacag	aagaacactt	gtctcaactg	ccatacataa	ctgccatttt	tcatgaaacc	1140
ttgagaaggc	atagtccagt	tcctataatt	ccattaagac	acgtgcatga	agacacagtg	1200
ttaggagggt	accatgtgcc	agctggaacc	gagctagcgg	taaacattta	tggatgtaac	1260
atggataaga	atgtgtggga	gaatcctgaa	gaatggaatc	cagagagatt	catgaaggaa	1320
aatgaaacga	tagattttcca	gaaaacaatg	gcgtttggag	gtggaaagcg	cgtatgtgct	1380
ggttcgcttc	aagcattgtt	gactgcttcc	attggaattg	gaagaatggt	gcaagagttt	1440
gagtggaaac	tgaagayat	gacccaagaa	gaagttaata	cgattgggct	tacgaccag	1500
atgcttcgct	cactgcgggc	cataataaag	cccaggatat	ga		1542

<210> 57

<211> 1530

<212> DNA

<213> *Arabidopsis thaliana*

<400> 57

atggccttct	tctccatgat	ctccattctc	cttggctttg	ttatctcctc	cttcatcttc	60
atcttcttct	tcaagaaact	tctctccttc	tccagaaaga	acatgtctga	agtctccact	120
ctcccctctg	ttccagtggg	accaggggtt	cctgttattg	ggaacttgct	gcaactaaaa	180
gagaagaaac	ctcacaagac	tttactaga	tggtcagaga	tttatgggtc	tatttactct	240
ataaagatgg	gttcttcttc	tcttattgtc	ctcaattcta	ctgagactgc	caaagaggcc	300
atggtgacgc	ggttttcgtc	tatctcaacg	aggaagtgtg	caaatgcggt	gacagtcctt	360
acttgtgaca	aatctatggg	tgctactagt	gattatgatg	atttccacaa	gttggtgaaa	420
cgggtgtctc	tgaacgggtc	tttgggtgct	aatgcacaga	aacgaaaaag	acattacaga	480
gatgcactca	ttgaaaatgt	gtcttccaag	ttgcatgccc	atgctagggg	ccatccacaa	540

gaacctgtaa	acttcagagc	tatatttgag	catgagcttt	tcggtgtagc	attgaagcaa	600
gcttttgga	aagatgtgga	atccatttat	gttaaagaac	tcggtgtgac	tttgtcgaaa	660
gacgagatct	tcaaggtttt	agtacatgac	atgatggaag	gtgcaattga	tgttgattgg	720
agagacttct	tcccatactt	gaaatggatt	ccaaataaaa	gttttgaagc	aagaatccag	780
caaaagcata	aacgtagact	cgcggtgatg	aatgctctga	ttcaagatcg	actgaagcag	840
aatggttcag	aatcggtatg	tgattgctat	ctcaacttct	tgatgtcgga	agcgaaaaca	900
ctaaccaagg	agcaaattgc	tatcttggtt	tgggagacga	ttatcgagac	agctgacact	960
actttgggta	caactgaatg	ggccatctat	gagctcgcta	agcatccaag	tgtccaagat	1020
cgtctgtgta	aagaaatcca	aaatgtctgc	ggaggagaaa	agttcaaaga	agagcaattg	1080
tctcaagttc	cttatctcaa	tggagtattc	catgaaacgc	ttaggaaata	cagtcctgct	1140
cctctagtcc	ccattcgcta	tgcccacgaa	gatacgcaaa	tcggaggcta	tcatgtccct	1200
gcaggaagtg	agattgcaat	aaacatctat	ggatgcaaca	tggataagaa	gcgttgggag	1260
agaccagagg	actggtggcc	ggagcggttt	cttgatgatg	gcaaatacga	aacgtcggat	1320
cttcacaaga	caatggcggt	tggagcggga	aagagggttt	gtgctgggtg	tcttcaagca	1380
tctctcatgg	caggcattgc	cattgggagg	ttagtgcaag	aattcgagtg	gaagcttaga	1440
gacggtgaag	aagagaatgt	ggatacatat	ggcttgacct	ctcagaagct	ttatcctctt	1500
atggctatta	tcaatccaag	gcgttcttaa				1530

<210> 58

<211> 1578

<212> DNA

<213> *Gibberella fujikoro*

<400> 58

atgagtaagt	ccaacagcat	gaacagtacc	agccatgaaa	cgttattcca	gcagctcgtc	60
ttaggtcttg	acagaatgcc	gctaattggac	gttacttggc	tgatctacgt	ggcctttggc	120
gcttggttat	gctcttatgt	catccatgtc	ctatcgctct	cttctacagt	caaagtgtcc	180
gtcgtaggct	accgcagcgt	ctttgagcct	acatggcttc	tccgtttgcg	ctttgtttgg	240
gaaggggggat	ctatcatcgg	ccaaggctac	aacaaattta	aagactctat	cttccagggtg	300
cgaaagcttg	gtaccgatat	cgtcatcatc	ccgccaaact	acatcgatga	ggtcagaaag	360
ctgtcccaag	acaagactcg	ctcggtcgag	cccttcatca	atgactttgc	gggacagtat	420
acacggggca	tgggtctttct	gcaaagtgat	ttgcagaacc	gtgtgattca	gcagcgggtg	480
acgccaaaac	tcgtatcggt	gacaaaggta	atgaaggagg	agcttgacta	tgccttgacc	540
aaagagatgc	ctgacatgaa	gaatgatgaa	tgggttgaag	tcgacatttc	ttccatcatg	600
gtcagggtca	tatcacgcat	ctcagccaga	gtgtttctcg	gtccagagca	ctgccgcaac	660
caagaatggt	tgacgaccac	tgagagtac	agcgagagcc	tgttcataac	tggctttatt	720
ctccgcgttg	tccccatat	tctaagacca	ttcatagccc	cgctgctacc	ctcctacaga	780
acactacttc	gcaacgtctc	gtcaggctcg	agagttattg	gagacatcat	tcgctcccag	840

caaggtgatg gcaacgagga catcctgtca tggatgaggg atgctgcgac aggggaagaa	900
aagcaaattg acaacattgc ccagcggatg cttatcctga gtctcgcgtc tattcacact	960
acggcaatga cgatgacgca tgctatgtat gacttatgtg cttgccctga gtacatagag	1020
cctcttagag atgaggtcaa aagtgtcgtt ggcgctagtg gttgggacaa gacggcgttg	1080
aatcgattcc acaaactcga cagctttctc aaagagtcac aacgcttcaa ccccggttgc	1140
ctcttaacgt tcaatcgcat ttatcaccaa tccatgacac tctcagatgg caccaacatc	1200
ccatcaggca ctgcgcatgc ggttccctct cacgcgatgc ttcaggactc agcgcatgtc	1260
ccaggcccga cgccaccaac cgagtttgat ggatttagat actcaaagat tcgctcagac	1320
tcaaactatg cacagaaata tctcttctcc atgactgatt ctagtaacat ggcgtttggg	1380
tatgggaaat acgcctgccc agggcggttc tatgcatcta atgagatgaa gctgactttg	1440
gcgatactcc ttttacaatt tgagttcaag ttgccagatg ggaaaggaag accacgaaat	1500
atcactattg atagtgcacat gatacctgat ccgagagcta ggctgtgcgt taggaagcga	1560
tcactgagag atgaatga	1578

<210> 59

<211> 1500

<212> DNA

<213> *Trametes versicolor*

<400> 59

atggaggatc ccaccgtact ctacgcttgc ctgccatcg ctgtcgctac tttcgttgtc	60
agatggtaca gagacccgct tcggtccatt cctacggttg ggggctctga ctttcccatc	120
ctctcataca tcggggcgct caggtggacc cgccgcggaa gagagatact gcaagaaggt	180
tatgatgggt atcgcgatc cacgttcaag atcgcgatgc tcgaccggtg gatcgatc	240
gccaacggcc caaagctcgc cgacgaggtg aggaggcgtc ctgacgaaga gctaaacttc	300
atggacggac tgggagcgtt cgtgcagacg aagtataccc ttggggaagc aatccacaat	360
gacccgtacc acgtggacat tattcgtgag aagctgacgc gaggcctccc ggcagtcctg	420
ccggacgtca tcgaggaact cacgctagcc gttcgccagt acatcccgc ggaaggagat	480
gaatgggtca gcgtgaactg ctccaaagca gcgcgggaca tcgtcgcccg ggcaagcaac	540
cgcgctctttg tcgggttgcc cgcttgccgc aaccagggtt atctcgacct cgccattgac	600
ttcacctga gcgttgtaa agacagggcg atcatcaata tgttcccga gttgctgaaa	660
cctatcgctg gacgcgtggt tggaaatgcc actaggaacg tgcgccgcgc ggtccattc	720
gtagcgccgt tgggtggagga acgtcgccgc ctcattggagg agtacggtga ggattggtcg	780
gagaaaccga acgacatgct ccagtggatc atggacgagg cagcctcgcg ggactcctcc	840
gtcaaagcga tcgctgagcg tcttctcatg gtcaactttg ccgcaattca cacgtcgctg	900
aacaccatca cccacgtctt ttaccacctc gccgagatgc cggagaccct acagccgctg	960
cggaagaga tcgagccgct cgtcaaggaa gaaggctgga cgaaggccgc catgggcaag	1020
atgtggtggc tcgacagctt cctgcgggag tcacagcgct acaatggcat caacatcgct	1080

tccctgacgc gcatggccga caaggacata acgctcagcg acggcacggt cctcccgaag	1140
ggcacgctcg tcgcgggtccc cgcgtactcg acgcaccgcg acgacgcggt gtacgcggac	1200
gcgctgggtct tcgacccggt ccgcttctcc cgcgtcgcg cccgcgaggg cgagggcacg	1260
aagcaccagt tcgtcaacac ctccgtggag tacgtgccct tcggccacgg gaagcacgcc	1320
tgccccgggc ggttcttcgc ggccaacgag ctgaaggcga tgctcgcgta catcgtgctc	1380
aactacgacg tgaagctgcc cggcgatggc aagcgcccc tgaacatgta ctggggcccg	1440
acggtcttgc ctgctccggc tgggcaggtg ctcttccgca agaggcaggt gtcgtgtag	1500

<210> 60

<211> 1578

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 60

atgggtttgt tcccattaga ggattcctac gcgctgggtct ttgaaggact agcaataaca	60
ctggctttgt actatctact gtctttcatc tacaaaacat ctaaaaagac atgtacacct	120
cctaaagcat ctggtgaaat cattccaatt acaggaatca tattgaatct gctatctggc	180
tcaagtggtc tacctattat cttagcactt gcctctttag cagacagatg tggctctatt	240
ttcaccatta ggctgggtat taggagagt ctagtagtat caaattggga aatcgctaag	300
gagattttca ctaccacga tttgatagtt tctaatagac caaaatactt agccgctaag	360
attcttgggt tcaattatgt ttcatctctt ttcgctccat acggcccata ttgggtcggg	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 tggtttcaact aactggcag atttgtcgtt	720
ggagacgctt ttccttttct aggttggttg gacctgggcg gatacaaaaa gacaatggaa	780
ttagttgcta gtagattgga ctcaatgggt 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 agcagtgcct	1140
aaagaggctt taagacttta cccagcagcg ttcttaggcg gaccaagagc attcttggaa	1200
gattgtactg ttgctgggta tagaattcca aagggcacct gcttgttgat taacatgtgg	1260
aaactgcata gagatccaaa catttgaggt gatccttgcg aattcaagcc agaaagattt	1320

ttgacaccta atcaaaagga tgttgatgtg atcgggatgg atttcgaatt gataccattt	1380
ggtgccggca gaagatatgg tccaggtact agattggctt tacagatgtt gcatatcgta	1440
ttagcgacat tgctgcaaaa cttcgaaatg tcaacaccaa acgatgcgcc agtcgatatg	1500
actgcttctg ttggcatgac aaatgccaaa gcatcacctt tagaagtctt gctatcacct	1560
cgtgttaaat ggtcctaa	1578

<210> 61

<211> 1431

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 61

atgatacaag ttttaactcc aattctactc ttcctcatct tcttcgtttt ctggaaagtc	60
tacaaacatc aaaagactaa aatcaatcta ccaccagggt ccttcggctg gccatttttg	120
ggtgaaacct tagccttact tagagcaggc tgggattctg agccagaaag attcgtaaga	180
gagcgtatca aaaagcatgg atctccactt gttttcaaga catcactatt tggagacaga	240
ttcgtgttc 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 tcgcaaaact cggtagtctt	600
ttcaacattt tcctcaaagg gatcatcgag cttcctatag acgttcctgg aactagattt	660
tactccagta aaaaggccgc agctgccatt agaattgaat tgaaaaagct cattaagct	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 tggaggctct	1260
agaatgtgtt taggcaaaga gtttgccagg ttagaagtgt tagcatttct ccacaacatt	1320
gttaccaact ttaagtggga tcttctaata cctgatgaga agatcgaata tgatccaatg	1380
gctactccag ctaagggtt gccaataga cttcatccac accaagtcta a	1431

<210> 62
 <211> 1578
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic polynucleotide

```

<400> 62
atggagtctt tagtggttca tacagtaa at gctatctggt gtattgta at cgtcgggatt      60
ttctcagttg gttatcacgt ttacggtaga gctgtggctg aacaatggag aatgagaaga      120
tactgaagc tacaagggtg taaaggccca ccaccatcca tcttcaatgg taacgtctca      180
gaaatgcaac gtatccaatc cgaagctaaa cactgctctg gcgataacat tatctcacat      240
gattattctt cttcattatt cccacacttc gatcactgga gaaaacagta cggcagaatc      300
tacacatact ctactggatt aaagcaacac ttgtacatca atcatccaga aatgggtgaag      360
gagctatctc agactaacac attgaacttg ggtagaatca cccatataac caaaagattg      420
aatcctatct taggtaacgg aatcataacc tctaattggtc ctcatgggc ccatcagcgt      480
agaattatcg cctacgagtt tactcatgat aagatcaagg gtatgggttg tttgatggtt      540
gagtctgcta tgcctatggt gaataagtgg gaggagatgg taaagagagg cggagaaatg      600
ggatgcgaca taagagttga tgaggacttg aaagatgttt cagcagatgt gattgcaaaa      660
gcctgtttcg gatcctcatt ttctaaagg t aaggctat t tctctatgat aagagatttg      720
cttacagcta tcacaaagag aagtgttcta ttcagattca acggattcac tgatatggtc      780
tttgggagta aaaagcatgg tgacgttgat atagacgctt tagaaatgga attggaatca      840
tccatttggt aaactgtcaa ggaacgtgaa atagaatgta aagatactca caaaaaggat      900
ctgatgcaat tgattttgga aggggcaatg cgttcattgt 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 gactatgggt attcaagaga caatgagatt atacctcca     1200
gcaccaatcg tcgggagaga agcctctaaa gatatcagat tgggcgatct agttgttcct     1260
aaaggcgtct gtatatggac actaatacca gctttacaca gagatcctga gatttgggga     1320
ccagatgcaa acgatttcaa accagaaaga ttttctgaag gaatttcaa ggcttgtaag     1380
tatcctcaaa gttacattcc atttggctct ggtcctagaa catgcgttgg taaaaacttt     1440
ggcatgatgg aagtaaagg tcttgtttcc ctgattgtct ccaagttctc tttcactcta     1500
tctcctacct accaakatag tcctagtcac aaacttttag tagaaccaca acatgggggtg     1560
gtaattagag tggtttaa                                     1578

```

<210> 63
 <211> 1590
 <212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 63

atgtacttcc tactacaata cctcaacatc acaaccgttg gtgtctttgc cacattgttt	60
ctctcttatt gtttacttct ctggagaagt agagcgggta acaaaaagat tgccccagaa	120
gctgccgctg catggcctat tatcggccac ctccacttac ttgcagggtg atcccatcaa	180
ctaccacata ttacattggg taacatggca gataagtacg gtcctgtatt cacaatcaga	240
ataggcttgc atagagctgt agttgtctca tcttgggaaa tggcaaagga atgttcaaca	300
gctaataatc aagtgtcttc ttcaagacct gaactattag cttctaagtt gttgggttat	360
aactacgcca tgtttggttt ttcaccatac gggtcatact ggagagaaat gagaaagatc	420
atctctctcg aattactatc taattccaga ttggaactat tgaaagatgt tagagcctca	480
gaagttgtca catctattaa ggaactatac aaattgtggg cggaagaa gaatgagtca	540
ggattgggtt ctgtcgagat gaaacaatgg ttcggagatt tgactttaaa cgtgatcttg	600
agaatgggtg ctggtaaaag atacttctcc gcgagtgcg cttcagaaaa caaacaggcc	660
cagcgttgta gaagagtctt cagagaattc ttccatctct ccggcttggt tgtggttgct	720
gatgctatac cttttcttgg atggctcgat tggggaagac acgagaagac cttgaaaaag	780
accgccatag aaatggattc catcgcccag gagtggcttg aggaacatag acgtagaaaa	840
gattctggag atgataattc tacccaagat ttcattggacg ttatgcaatc tgtgctagat	900
ggcaaaaatc taggcggata cgatgctgat acgattaaca aggctacatg ctttaactctt	960
atatcaggtg gcagtgtatc tactgtagtt tctttgacat gggctcttag tcttgtgtta	1020
aacaatagag atactttgaa aaaggcacag gaagagttag acatccaagt cggttaaggaa	1080
agattgggta acgagcaaga catcagtaag ttagtttact tgcaagcaat agtaaaagag	1140
acactcagac tttatccacc aggtcctttg ggtgggttga gacaattcac tgaagattgt	1200
acactagggtg gctatcacgt ttcaaaagga actagattaa tcatgaactt atccaagatt	1260
caaaaagatc cacgtatttg gtctgatcct actgaattcc aaccagagag attccttacg	1320
actcataaag atgtcgatcc acgtggtaaa cactttgaat tcattccatt cgggtgcagga	1380
agacgtgcat gtcctggtat cacattcgga ttacaagtac tacatctaac attggcatct	1440
ttcttgcatg cgtttgatc ttcaacacca tcaaatgagc aggttaacat gagagaatca	1500
ttaggtctta cgaatatgaa atctaccca ttagaagttt tgatttctcc aagactatcc	1560
cttaattgct tcaaccttat gaaaatttga	1590

<210> 64

<211> 1440

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 64

atggaaccta acttttactt gtcattacta ttgttggtcg 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 taatcgaca aagacatttt	480
gtcactcact gggacaacaa aaatgagatc acagtttatc cacttgctaa aagatacact	540
ttcttgcttg cgtgtagact gttcatgtct gttgaggatg aaaatcatgt ggcgaaattc	600
tcagacccat tccaactaat cgctgcaggc atcatttcac ttcctatcga tcttctgggt	660
actccattca acaaggccat aaaggcttca aatttcatta gaaaagagct gataaagatt	720
atcaaacaaa gacgtgttga tctggcagag ggtacagcat ctccaacca ggatatcttg	780
tcacatatgc tattaacatc tgatgaaaac ggtaaactta tgaacgagtt gaacattgcc	840
gacaagattc ttggactatt gataggaggc cacgatacag cttcagtagc ttgcacattt	900
ctagtgaagt acttaggaga attaccacat atctacgata aagtctacca agagcaaatg	960
gaaattgcc agtccaaacc tgctggggaa ttgttgaatt gggatgactt gaaaaagatg	1020
aagtattcat ggaatgtggc atgtgaggta atgagattgt caccaccttt acaagggtgt	1080
tttagagagg ctataactga ctttatgttt aacggtttct ctattccaaa aggggtggaag	1140
ttatactggc cgcacaactc tacacacaaa aatgcagaat gtttcccaat gcctgagaaa	1200
ttcgatccta ccagatttga aggtaatggc 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

<210> 65

<211> 1572

<212> DNA

<213> Stevia rebaudiana

<400> 65

atgggtctct tccctttgga agatagttac gcactcgtct ttgaagggtt agcaataact	60
actctagctc tctactactt attatccttc atctataaaa cctctaataa gacttgtact	120
ccacctaaag caagcggtag gcaccctata acaggccact taaaccttct tagtggttca	180
tccggtcttc cccatctagc cttagcatct ttggctgacc gatgtgggcc catattcacc	240
atccgacttg gcatacgtag agttttggtg gttagtaatt gggaaattgc taaggagatc	300
ttcactaccc atgatttgat tgtttcaaac cgtcccaaat acctcgctgc aaagattttg	360

ggattcaact atgtgtcctt ttcgtttgct ccatatggcc cctattgggt tggaatccgt	420
aagatcatcg ccacaaaact gatgtcaagt agcaggctcc agaagcttca gtttgtccga	480
gttttcgaac tagaaaactc catgaaaagc atacgcgagt cttggaaaga gaaaaaagac	540
gaagaaggta aagtgttggg ggagatgaaa aaatggtttt gggaattgaa tatgaatata	600
gttcttagaa ctgttgctgg taaacagtac actggaactg ttgatgatgc ggatgcgaag	660
aggattagtg aattgttttag agaatggttt cattacacag gaaggtttgt tgtgggagat	720
gcttttcctt ttcttgggtg gttggatttg ggtggatata agaagaccat ggaactagt	780
gcttccagac tagattccat ggtctcaaaa tggtagacg agcatcgcaa aaagcaggct	840
aacgacgaca aaaaagagga catggatttc atggacatca tgatatcgat gactgaagcc	900
aattcccctt tggaggggta tggtagcgat acaataatta aaaccacttg catgactctt	960
attgtcagtg gtgtagatac aacctccatc gtgctaactt gggcactctc gttactactg	1020
aacaaccgtg acactcttaa gaaagctcaa gaagagctag acatgtgtgt gggaaaagg	1080
cgacaagtaa acgaatcaga tctagtaaac ctaatctacc ttgaagccgt attaaaagaa	1140
gcattgcgac tatacccagc agcattcctt ggaggtccta gagccttttt agaagactgc	1200
accgtggcag ggtaccgtat cccaaaaggc acatgtctac ttattaacat gtggaaactt	1260
catcgtgatc caaacatatg gtcagaccca tgtgagttta aaccagagag gttcttaacc	1320
ccaaacaaa aggacgtaga tgttattgga atggattttg agttaatccc atttgggtcg	1380
ggaagaagggt attgtccagg gacacgtttg gcattacaaa tgttacacat agttctggcc	1440
actctactac aaaactttga gatgtcaact ccaaatgatg caccggttga tatgaccg	1500
agtgttgga tgacaaatgc gaaggcaagt ccacttgaag ttctactttc gccacgtgtt	1560
aagtggcat ag	1572

<210> 66

<211> 1431

<212> DNA

<213> Stevia rebaudiana

<400> 66

atgattcaag ttctaacc gatccttctc ttctcatatt tcttcgtttt ctggaaggtt	60
tacaagcacc agaaaaccaa aatcaatctt ccaccgggaa gcttcggatg gccatttctg	120
ggcgaaactc tggcactcct acgtgcagggt tgggactcag agccggagag atttgttcgt	180
gaacggatca agaaacacgg aagtcctcta gtgtttaaga cgtcgttggt tggcgaccgt	240
tttgcggtgt tgtgtggacc tgccggaaac aagttcctgt tctgcaacga gaacaagctg	300
gtggcgtcgt ggtggccggt tccggtgagg aagcttttcg gcaagtctct gtcacgatt	360
cgtggtgatg aagctaagt gatgaggaag atgttggtat cgtatctcgg tcctgatgct	420
ttcgcaactc attatgccgt caccatggac gtcgtcacc gtcggcatat cgacgttcat	480
tggcgagggg aggaagagg gaacgtattc caaacgtta agttatatgc ctttgagctt	540
gcatgtcgtt tattcatgaa cctagacgac ccaaaccaca ttgcaaaact cggttccttg	600

ttcaacattt	tcttgaaagg	catcattgag	cttccaatcg	acgtcccagg	gacacgattt	660
tatagctcca	aaaaagcagc	agcagctatc	aggattgaac	taaaaaaatt	gattaaagca	720
agaaaactgg	aactgaaaga	aggggaaggca	tcatcttcac	aagacctctt	atcacatttg	780
cttacatctc	cagatgaaaa	tggtatgttt	ctaaccgaag	aagagattgt	agacaacatc	840
ttgttactac	tctttgcggg	tcatgatacc	tcggctcttt	caatcacttt	gctcatgaag	900
actcttggcg	aacattctga	tgtttatgac	aagggtgtta	aagagcaact	agagatatcg	960
aagacgaaag	aagcatggga	gtccctgaaa	tgggaggaca	tacaaaagat	gaaatactcc	1020
tggagtgtta	tatgtgaagt	catgagacta	aatccacctg	ttataggaac	ctatagagag	1080
gcccttgtgg	atattgatta	tgcggtttat	accatcccca	aaggatggaa	gctgcactgg	1140
agtgtctgtat	cgacacaaag	ggacgaggct	aactttgaag	acgtaacacg	ttttgaccca	1200
tcacggtttg	aaggcgcagg	accgactcca	ttcacctttg	ttccgtttgg	aggggggcct	1260
agaatgtgtt	tagggaaaga	atttgctcga	ttggaagtac	ttgcgtttct	tcacaatatt	1320
gtcaccaatt	tcaaatggga	cctgttgata	cctgatgaga	aaatagaata	tgatcccatg	1380
gctaccccag	caaaggggct	tccaattcgt	cttcatcccc	atcaagtttg	a	1431

<210> 67

<211> 1578

<212> DNA

<213> *Arabidopsis thaliana*

<400> 67

atggagagtt	tggttgttca	tacggtaa	aatgca	atttgg	gcatagttat	tgtcggaa	tc	60
ttcagcgtag	gttatcatgt	gtatggaaga	gcggtggtgg	agcagtggag	gatgcggagg			120
agtttaaagt	tgcaaggcgt	gaagggtcct	ccaccgtcga	tctttaacgg	caatgtgtcg			180
gagatgcaac	ggattcagtc	ggaggctaaa	cactgtttccg	gcgataacat	catttctcat			240
gactattctt	cttctctatt	tcctcatttc	gatcactggc	gaaaacaata	cggaaggatt			300
tacacatact	caacgggggt	aaagcagcac	ctttacataa	accacccgga	aatgggtgaag			360
gagcttagcc	aaaccaacac	acttaacctt	ggtagaatca	ctcacatcac	caaacgcctt			420
aaccccatc	tcggaatgg	catcatcacc	tctaattggc	ctcattgggc	ccatcaacgt			480
cgtatcattg	cctatgagtt	taccacgac	aaaatcaagg	gaatgggttg	tttaatgggtg			540
gaatctgcca	tgccaatgtt	gaacaaatgg	gaagagatgg	tgaaaagagg	aggagaaatg			600
ggttgtgaca	taagagtgg	cgaagacctt	aaggatgtct	cagctgatgt	catcgctaag			660
gcttgctttg	ggagctcttt	ttcaaaaggc	aaagcaatat	tctctatgat	tagggatctt			720
ttaaccgcc	ttactaaacg	aagcgtcctc	ttcagattca	atggcttcac	tgatatgggtg			780
tttggaagta	agaagcatgg	tgatgtggat	attgatgcgc	ttgagatgga	attagaatct			840
tctatatggg	aaacgggtta	ggagagggaa	attgaatgta	aggatactca	caagaaggat			900
ctaatagcagt	tgatactcga	gggagcgatg	cgaagctgcg	atggtaactt	gtgggacaag			960
tcagcctata	gacggtttgt	ggtggacaat	tgcaagagca	tctatttcgc	cggacatgat			1020

tcaaccgcag	tctcagtgtc	ttggtgcctt	atgctcctcg	ctctcaatcc	tagttggcag	1080
gttaaaattc	gcatgaaat	cttgagttct	tgcaagaatg	gcattcccga	cgcagaatca	1140
attcctaadc	tcaaaacggt	gacaatggta	atacaagaaa	caatgagact	ataccaccca	1200
gcaccaatcg	tgggaagaga	agcatccaaa	gacataagac	ttggagacct	tgtggtgcc	1260
aaaggagtgt	gcatttggac	actcattcct	gccttacacc	gagaccccga	gatctgggga	1320
ccagacgcaa	acgacttcaa	gccagagagg	tttagtgagg	gaatctctaa	ggcttgcaaa	1380
taccctcagt	catacatccc	atttggcctt	ggaccaagaa	catgcgtagg	caaaaacttt	1440
ggtatgatgg	aagtgaaggt	gcttgtttca	cttattgtct	caaagttag	ttttactctt	1500
tccccgactt	atcagcactc	tccaagccat	aaactccttg	tagagcctca	acatggtggt	1560
gtcattaggg	ttgtttga					1578

<210> 68

<211> 1590

<212> DNA

<213> Vitis vinifera

<400> 68

atgtatttcc	ttctccaata	cctaaacatc	accacggctcg	gagtctttgc	cacacttttc	60
ctttcctact	gtctattatt	atggaggtct	agagctggta	acaaaaaat	agcacctgaa	120
gctgctgctg	catggcccat	aatcggtcac	ctacacctgt	tagctggtgg	ttctcatcag	180
cttccccaca	taaccttggg	aaacatggcc	gacaaatatg	gaccggtctt	cacaattcgg	240
attgggttgc	atcgagctgt	ggtggttaagt	tcttgggaga	tggctaaaga	atgctcgacc	300
gccaatgacc	aggtttcatc	ctcgcgtccc	gaacttttag	cctcaaaact	tttgggctac	360
aactacgcc	tgtttggttt	ctctccatac	ggttcttact	ggcgtgaaat	gcgcaagata	420
atcagcctag	agctactctc	taacagccgc	ttagagctgc	tgaaggacgt	ccgagcttca	480
gaagtgggtg	catccataaa	agagctatac	aagctctggg	cagagaaaaa	aaatgaatcg	540
ggccttgtct	cgggtggagat	gaagcagtgg	tttggagact	tgactctgaa	cgtaattctt	600
aggatggtgg	caggaagcg	ttatttcagt	gcttcagatg	caagtgaaa	taaacaggcg	660
cagaggtgcc	ggagagtgtt	caggaattc	tttcatttgt	cagggctctt	tgtggtggcg	720
gacgctattc	catttcttgg	atggctcgac	tgggggagac	atgagaaaac	cctaaagaag	780
acagcaatag	aaatggacag	tattgtctca	gaatgggttag	aggagaccg	tcggaggaaa	840
gactccggtg	atgataatag	tacgcaagac	ttcatggatg	tgatgcagtc	agttcttgat	900
ggcaaaaacc	ttggtgggtta	cgacgctgat	accatcaata	aagccacatg	cctgactcta	960
atctccggag	gtagcgacac	aactgttgtc	tctctaacat	gggcactctc	tcttgactta	1020
aacaaccgtg	acaccttaaa	aaaagctcaa	gaagaattag	acatccaagt	tggttaaggaa	1080
agattagtga	atgaacaaga	tataagtaag	ttggtctatc	tccaagccat	tggttaagag	1140
acattacggt	tatatccacc	aggaccactt	ggaggactac	gccaatttac	cgaggattgc	1200
accttgggtg	gataccatgt	ctctaaaggc	acccgtttta	taatgaacct	ttcgaagatc	1260

caaaaggatc	caagaatttg	gtcagatccg	acagaattcc	aaccagagag	gtttctcacc	1320
acccataaag	atgttgatcc	tcggggaaaa	cattttgagt	ttataaccatt	tggagctggg	1380
cgaagagcat	gtccaggaat	aacttttggt	cttcaagtat	tacatttaac	attggctagt	1440
ttcttacatg	cgtttgaatt	ttcaactcca	tcaaataaac	aggtaaatat	gcgcgagagc	1500
cttggaacta	caaataatgaa	atctacccca	cttgaagttc	tcattttctcc	acgcttatca	1560
ttgaattggt	ttaacctaata	gaagatatata				1590

<210> 69

<211> 1440

<212> DNA

<213> Medicago trunculata

<400> 69

atggagccta	atttctatct	ctcccttctc	cttctctttg	tcactttcat	atctctctct	60
ctttttttca	tattctacaa	acagaaatct	ccattaaatt	tgccacctgg	taaaatgggt	120
taccaatca	taggtgaaag	ccttgagttc	ttatcaacag	gatggaaagg	acatcctgaa	180
aaattcattt	tcgaccgtat	gcgtaaatat	tcctcagaac	tcitttaaac	atcaatcgta	240
ggagaatcta	cggtggtttg	ttgcggagca	gcaagtaaca	agtttttggt	ttcaaacgag	300
aataaacttg	tgactgcatg	gtggccagat	agtgtaaaca	aaatcttccc	tactacttct	360
cttgactcta	acttgaagga	agaatccatc	aagatgagaa	aattgcttcc	acaattcttt	420
aaacccgaag	ctctacaacg	ttatgttggt	gtcatggatg	ttattgctca	aagacatttt	480
gttactcatt	gggataataa	aaatgaaatc	accgtctacc	ccttggccaa	gaggtacacc	540
tttttgttag	cttgctcggtt	gttcatgagc	gttgaagacg	agaatcatgt	agcaaaattt	600
agtgatccat	ttcagttaat	tgcggccgga	atcatatctc	taccaattga	tttgccagga	660
acaccattca	acaaagctat	aaaggcctca	aactttataa	gaaaggaggt	gattaagatc	720
ataaagcaaa	ggagggtaga	tttgccagaa	gggacagcat	caccaacaca	agatatattg	780
tctcacatgt	tgttgacaag	tgatgaaaat	ggaaagagta	tgaatgaact	taatattgct	840
gataagattc	ttggcctttt	gatcggagga	catgacactg	ctagcgtcgc	atgcactttc	900
cttgtcaaat	atctcggcga	gttacctcac	atttatgata	aagtctatca	agagcaaagt	960
gaaattgcaa	aatcgaaacc	agcaggagaa	ttgttgaatt	gggatgacct	gaagaaaatg	1020
aaatactctt	ggaacgtagc	ttgtgaagta	atgagacttt	ccccctcact	ccaaggaggt	1080
ttcaggggaag	ccatcactga	ctttatgttc	aatggattct	caattcctaa	gggatggaag	1140
ctttatttga	gtgcaaattc	aacacataag	aacgcagaat	gttttcccat	gccagagaaa	1200
tttgacccaa	caagatttga	aggaaatgga	ccagctcctt	atacttttgt	tccctttggg	1260
ggaggaccaa	ggatgtgtcc	tggaaaagag	tatgcaagat	tagaaatact	tgttttcatg	1320
cacaatttgg	tgaagaggtt	taagtgggaa	aagggtgattc	cagatgagaa	gattattggt	1380
gatccattcc	ccatccctgc	aaaggatctt	ccaattcgcc	tttatccaca	caaagcttaa	1440

<210> 70
 <211> 2133
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 70
 atgcaatcag attcagtcaa agtctctcca tttgatttgg tttccgctgc tatgaatggc 60
 aaggcaatgg aaaagttgaa cgctagttaa tctgaagatc caacaacatt gcctgcacta 120
 aagatgctag ttgaaaatag agaattgttg aactgttca caacttcctt cgcagttctt 180
 attgggtgtc ttgtatttct aatgtggaga cgttcatcct ctaaaaagct ggtacaagat 240
 ccagttccac aagttatcgt tgtaaagaag aaagagaagg agtcagaggt tgatgacggg 300
 aaaaagaaag tttctatttt ctacggcaca caaacaggaa ctgccgaagg ttttgctaaa 360
 gcattagtgc aggaagcaaa agtgagatat gaaaagacct ctttcaagggt tatcgatcta 420
 gatgactacg ctgcagatga tgatgaatat gagggaaaac tgaaaaagga atccttagcc 480
 ttcttcttct tggccacata cggtgatggg gaacctactg ataatgctgc taacttctac 540
 aagtggttca cagaaggcga cgataaagggt gaatggctga aaaagttaca atacggagta 600
 tttggtttag gtaacagaca atatgaacat ttcaacaaga tcgctattgt agttgatgat 660
 aaacttactg aaatgggagc caaaagatta gtaccagtag gattagggga tgatgatcag 720
 tgtatagaag atgacttcac cgcttgaag gaattgggtat ggccagaatt ggatcaactt 780
 ttaagggacg aagatgatac ttctgtgact acccataca ctgcagccgt attggagtac 840
 agagtggttt accatgataa accagcagac tcatatgctg aagatcaaac ccatacaaac 900
 ggtcatgttg ttcattgatgc acagcatcct tcaagatcta atgtggcttt caaaaaggaa 960
 ctacacacct ctcaatcaga taggtcttgt actcacttag aattcgatat ttctcacaca 1020
 ggactgtctt acgaaactgg cgatcacggt ggcgtttatt ccgagaactt gtccgaagtt 1080
 gtcgatgaag cactaaaact gttaggggta tcaccagaca catacttctc agtccatgct 1140
 gataaggagg atgggacacc tatcgggtgt gttcactac caccaccttt tcctccttgc 1200
 acattgagag acgctctaac cagatacgca gatgtcttat cctcacctaa aaaggtagct 1260
 ttgctggcat tggctgtcga tgctagtgat cctagtgaag ccgatagggt aaagttcctg 1320
 gcttcaccag ccggaaaaga tgaatatgca caatggatcg tcgccaacca acgttctttg 1380
 ctagaagtga tgcaaagttt tccatctgcc aagcctccat taggtgtgtt cttcgagca 1440
 gtagctccac gtttacaacc aagatactac tctatcagtt catctcctaa gatgtctcct 1500
 aacagaatac atgttacatg tgctttgggtg tacgagacta ctccagcagg cagaattcac 1560
 agaggattgt gttcaacctg gatgaaaaat gctgtccctt taacagagtc acctgattgc 1620
 tctcaagcat ccatthtcgt tagaacatca aatttcagac ttccagtga tccaaaagtt 1680
 ccagtcatta tgataggacc aggcactggg cttgccccat tcaggggctt tcttcaagag 1740
 agattggcct tgaaggaatc tggtagagaa ttgggttctt ctatcttttt ctttgggtgc 1800

cgtaatagaa aagttgactt tatctacgag gacgagctta acaattttgt tgagacagga	1860
gcattgtcag aattgatcgt cgcattttca agagaagggga 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

<210> 71

<211> 2079

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 71

atgacttctg cactttatgc ctccgatctt ttcaacaat tgaaaagtat catgggaacg	60
gattctttgt ccgatgatgt tgtattagtt attgctacaa cttctctggc actggttgct	120
ggtttcgttg tcttattgtg gaaaaagacc acggcagatc gttccggcga gctaaagcca	180
ctaagatgcc 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
tacaagtggg ttactgaaga gaacgaaaga gatatcaagt tgcagcaact tgcttacggc	540
gtttttgcct taggtaacag acaatacgag cactttaaca agataggtat tgtcttagat	600
gaagagttat gcaaaaaggg tgcgaagaga ttgattgaag tcggtttagg agatgatgat	660
caatctatcg aggatgactt taatgcatgg aaggaatctt tgtggctctga attagataag	720
ttacttaagg acgaagatga taaatccgtt gccactccat acacagccgt cattccagaa	780
tatagagtag ttactcatga tccaagattc acaacacaga aatcaatgga aagtaatgtg	840
gctaattgga atactaccat cgatattcat catccatgta gagtagacgt tgcagttcaa	900
aaggaattgc aactcatga atcagacaga tcttgcatat atcttgaatt tgatatatca	960
cgtactggta tcacttacga aacaggtgat cacgtgggtg tctacgtga aaaccatgtt	1020
gaaattgtag aggaagctgg aaagtgtgtg 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 ctcttttggg tgttttcttc	1380

gccgcaatag cgcctagact gcaaccaaga tactattcaa tttcatcctc acctagactg	1440
gcaccatcaa gagttcatgt cacatccgct ttagtgtagc gtccaactcc tactggtaga	1500
atccataagg gcgtttgttc aacatggatg aaaaacgcgg ttccagcaga gaagtctcac	1560
gaatgttctg gtgctccaat ctttatcaga gcctccaact tcaaactgcc ttccaatcct	1620
tctactccta ttgtcatggt cggtcctggt acaggctctg ctccattcag aggtttctta	1680
caagagagaa tggccttaaa ggaggatggt gaagagttgg gatcttcttt gttgtttttc	1740
ggctgtagaa acagacaaat ggatttcac tacgaagatg aactgaataa cttttagat	1800
caaggagtta tttcagagtt gataatggct ttttctagag aagggtgctca gaaggagtac	1860
gtccaacaca aaatgatgga aaaggccgca caagtttggg acttaataca agaggaaggc	1920
tatctatatg tctgtggtga tgcaaaggg atggcaagag atgttcacag aacacttcat	1980
actatagtcc aggaacagga aggcgttagt tcttctgaag cggaagcaat tgtgaaaaag	2040
ttacaaacag agggaagata cttgagagat gtgtggtaa	2079

<210> 72

<211> 2142

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 72

atggcagaat tagatacact tgatatagta gtattaggtg ttatcttttt ggggtactgtg	60
gcatacttta ctaagggtaa attgtggggt gttaccaagg atccatacgc taacggattc	120
gctgcagggt 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 acgggtgaagg cgaaccaaca gataacgccg tggatttcta tgagttcatt	420
actggcgaag atgcctcttt caatgagggc aacgatcctc cactaggtaa cttgaattac	480
gttgcgttcg gtctgggcaa caatacctac gaacactaca actcaatggt caggaacggt	540
aacaaggctc tagaaaagtt aggagctcat agaattggag aagcagggtga 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 ctacagtaag gatagaaatt gtctgcatat ggaaattgat	900
atttctggta gtaatctaaa gtatgaaaca ggcgaccata tcgcgatctg gcctaccaac	960
ccagggtgaag aggtcaacaa atttcttgac attctagatc tgtctggtaa gcaacattcc	1020

gtcgtaacag	tgaaagcctt	agaacctaca	gccaaagttc	cttttccaaa	tccaactacc	1080
tacgatgcta	tattgagata	ccatctggaa	atatgcgctc	cagtttctag	acagtttgtc	1140
tcaacttttag	cagcattcgc	ccctaattgat	gatatcaaag	ctgagatgaa	ccgtttggga	1200
tcagacaaaag	attactttcca	cgaaaagaca	ggaccacatt	actacaatat	cgctagattt	1260
ttggcctcag	tctctaaagg	tgaaaaatgg	acaaagatac	catttttctgc	tttcatagaa	1320
ggccttacaa	aactacaacc	aagatactat	tctatctctt	cctctagttt	agttcagcct	1380
aaaaagatta	gtattactgc	tgtgtgcgaa	tctcagcaaa	ttccaggtag	agatgaccca	1440
ttcagagggtg	tagcgactaa	ctacttgttc	gctttgaagc	agaaacaaaa	cggtgatcca	1500
aatccagctc	cttttggcca	atcatacgag	ttgacaggac	caaggaataa	gtatgatggt	1560
atacatgttc	cagtccatgt	aagacattct	aactttaagc	taccatctga	tccaggcaaa	1620
cctattatca	tgatcgggtc	aggtaccggt	gttgcccctt	ttagaggctt	cgtccaagag	1680
agggcaaaac	aagccagaga	tgggtgtagaa	gttggtaaaa	cactgctggt	ctttggatgt	1740
agaaagagta	cagaagattt	catgtatcaa	aaagagtggc	aagagtacaa	ggaagctctt	1800
ggcgacaaat	tcgaaatgat	tacagctttt	tcaagagaag	gatctaaaaa	ggtttatggt	1860
caacacagac	tgaaggaaaag	atcaaaggaa	gtttctgatc	ttctatccca	aaaagcatac	1920
ttctacgttt	gcggagacgc	cgcacatatg	gcacgtgaag	tgaacactgt	gttagcacag	1980
atcatagcag	aaggccgtgg	tgtatcagaa	gccaaggggtg	aggaaattgt	caaaaacatg	2040
agatcagcaa	atcaatacca	agtgtgttct	gatttcgtaa	ctttacactg	taaagagaca	2100
acatacgcg	attcagaatt	gcaagaggat	gtctggagtt	aa		2142

<210> 73

<211> 2133

<212> DNA

<213> Stevia rebaudiana

<400> 73

atgcaatcag	attccgtaaa	agtgtcgccg	ttcgatctcg	tatctgcagc	tatgaacgga	60
aaagcaatgg	agaaattgaa	cgcatcgga	tcggaagatc	cgacgacgct	accggcggtg	120
aagatgctgg	tggagaatcg	cgagctgctg	acactgttta	cgacgtcggt	tgctgtattg	180
atcggatgtc	tcgtgttttt	gatgtggcgg	agatcgtcct	cgaagaaact	ggttcaggat	240
ccggtgccgc	aggtaatcgt	tgtaagaag	aaagagaagg	agtctgaggt	tgatgatggc	300
aagaagaaaag	tttcgatatt	ctacggaact	caaacaggaa	ccgctgaagg	ttttgccaaag	360
gcactttag	aggaagctaa	agttcgatat	gaaaagacat	cctttaaagt	tattgatctg	420
gatgattatg	ctgctgatga	cgatgagtat	gaggagaagc	ttaagaaaga	atctttggcg	480
tttttctttt	tggcaacgta	tggagatggt	gaaccaacag	ataatgcagc	caatttttac	540
aaatggttta	cagagggaga	tgacaaaggc	gaatggctga	agaaacttca	atatggcgtg	600
tttggcctcg	gtaacagaca	atatgagcat	ttcaataaga	ttgcaatagt	ggttgatgac	660
aaactcacag	aaatgggcgc	aaaacgcctt	gttcctgtgg	gtcttgagga	tgacgatcaa	720

tgtatagaag atgactttac agcatggaaa gagttagtgt ggcccagagt ggatcaattg	780
ttgcgtgatg aggatgacac gagtggttacg actccttaca ctgctgcggt tttggaatac	840
cgagttgtat atcatgataa acctgcagac tcgtatgcag aagatcaaac tcatacaaat	900
ggatcatgtt ttcatgatgc tcaacatcca tctagatcca atgtggcatt taaaaaggaa	960
ttgcacacct ctcaatctga ccggtcttgc actcatttgg aatttgatat ctctcacacc	1020
gggctatcat acgagacggg ggatcatgtt ggtgtctaca gtgagaatct aagtgaagtt	1080
gtagatgaag ctttaaaatt actcggtttg tcacccgaca cttatttctc agtccatgct	1140
gacaaggaag acggaacacc tattggcggc gcctccttgc cgccacctt ccttccatgc	1200
actttaagag atgcattaac gcgctacgca gatgttttga gttctcctaa aaagggttgc	1260
ttgcttgctc tggctgctca tgcttctgat cctagcgaag ccgatcgatt aaaatttcta	1320
gcatctccgg ctggcaagga tgaatatgct caatggatag ttgcaaacca aagaagtctt	1380
cttgaagtta tgcagtcatt tccgtcagct aaaccgccac ttggtgtttt cttcgcagct	1440
gtcgcctcac gtttacaacc tcgatattac tcgatttctt cttctccaaa gatgtcacca	1500
aacagaattc atgtgacttg tgcattagtt tatgagacaa cacctgcagg acgtattcac	1560
agaggattgt gttcaacatg gatgaagaat gctgtgcctt tgaccgaaag tccagattgt	1620
agtcaggcgt cgatttttgt tagaacgtct aacttccgac ttccggttga cccgaaagtc	1680
ccggtcatca tgatcgggtc cgggactggg ttagccctt tcagagggtt tcttcaagaa	1740
cgggttagct tgaaggaatc tggaaaccgaa ctccgggtcat ctattttctt tttcggatgc	1800
agaaaccgca aagtggattt tatatacgaa gacgaactaa acaactttgt ggagaccggt	1860
gctttatcgg agcttattgt tgcattctcc cgtgaaggaa ccgcaaagga gtatgtgcaa	1920
cataaaatga gccagaaggc ttcagatatt tggaaagttg tttcagaggg agcatattta	1980
tatgtatgtg gtgatgctaa aggcattggc aaagatgtac acagaaccct tcacacaatt	2040
gtacaagaac agggatctct agattcttcc aaggcagaat tgtatgtaaa gaacctacaa	2100
atgtcgggaa gatattctcg tgatgtttgg taa	2133

<210> 74

<211> 2079

<212> DNA

<213> Arabidopsis thaliana

<400> 74

atgacttctg ctttgtatgc ttccgatttg ttttaagcagc tcaagtcaat tatggggaca	60
gattcgttat ccgacgatgt tgtacttggt attgcaacga cgtctttggc actagtagct	120
ggatttgtgg tggtgttatg gaagaaaacg acggcggatc ggagcgggga gctgaagcct	180
ttgatgatcc ctaagtctct tatggctaag gacgaggatg atgatttggg tttgggatcc	240
gggaagacta gagtctctat ctcttccggt acgcagactg gaacagctga gggatttgc	300
aaggcattat ccgaagaaat caaagcgaga tatgaaaaag cagcagtcaa agtcattgac	360
ttggatgact atgctgccga tgatgaccag tatgaagaga aattgaagaa ggaaactttg	420

gcatttttct	gtgttgctac	ttatggagat	ggagagccta	ctgacaatgc	tgccagattt	480
tacaaatggg	ttacggagga	aaatgaacgg	gatataaagc	ttcaacaact	agcatatggg	540
gtgtttgctc	ttggtaatcg	ccaatatgaa	cattttaata	agatcgggat	agttcttgat	600
gaagagttat	gtaagaaagg	tgcaaagcgt	cttattgaag	tcgggtctagg	agatgatgat	660
cagagcattg	aggatgattt	taatgcctgg	aaagaatcac	tatgggtctga	gctagacaag	720
ctcctcaaag	acgaggatga	taaaagtgtg	gcaactcctt	atacagctgt	tattcctgaa	780
taccgggtgg	tgactcatga	tcctcggttt	acaactcaaa	aatcaatgga	atcaaagtgt	840
gccaatggaa	atactactat	tgacattcat	catccctgca	gagttgatgt	tgctgtgcag	900
aaggagcttc	acacacatga	atctgatcgg	tcttgcatte	atctcgagtt	cgacatatcc	960
aggacgggta	ttacatatga	aacagggtgac	catgtaggtg	tatatgctga	aaatcatggt	1020
gaaatagttg	aagaagctgg	aaaattgctt	ggccactctt	tagatttagt	attttccata	1080
catgctgaca	aggaagatgg	ctccccattg	gaaagcgcag	tgccgcctcc	tttccctggt	1140
ccatgcacac	ttgggactgg	tttggaaga	tacgcagacc	ttttgaacc	tcctcgaaag	1200
tctgcgttag	ttgccttggc	ggcctatgcc	actgaaccaa	gtgaagccga	gaaacttaag	1260
cacctgacat	cacctgatgg	aaaggatgag	tactcacaat	ggattgttgc	aagtcagaga	1320
agtcttttag	aggtgatggc	tgctttttcca	tctgcaaaac	ccccactagg	tgtatttttt	1380
gctgcaatag	ctcctcgtct	acaacctcgt	tactactcca	tctcatcctc	gccaagattg	1440
gcgccaagta	gagttcatgt	tacatccgca	ctagtatatg	gtccaactcc	tactggtaga	1500
atccacaagg	gtgtgtgttc	tacgtggatg	aagaatgcag	ttcctgcgga	gaaaagtcac	1560
gaatgtagtg	gagccccaat	ctttattcga	gcattctaatt	tcaagttacc	atccaaccct	1620
tcaactccaa	tcgttatggg	gggacctggg	actgggctgg	caccttttag	aggttttctg	1680
caggaaagga	tggcactaaa	agaagatgga	gaagaactag	gttcatcttt	gctcttcttt	1740
gggtgtagaa	atcgacagat	ggactttata	tacgaggatg	agctcaataa	ttttgttgat	1800
caaggcgtaa	tatctgagct	catcatggca	ttctcccgtg	aaggagctca	gaaggagtat	1860
gttcaacata	agatgatgga	gaaggcagca	caagtttggg	atctaataaa	ggaagaagga	1920
tatctctatg	tatgcggtga	tgctaagggc	atggcgaggg	acgtccaccg	aactctacac	1980
accattgttc	aggagcagga	aggtgtgagt	tcgtcagagg	cagaggctat	agttaagaaa	2040
cttcaaaccg	aaggaagata	cctcagagat	gtctggtga			2079

<210> 75

<211> 2142

<212> DNA

<213> Giberella fujikuroi

<400> 75

atggctgaac	tcgacactct	ggacatcgtc	gtcctcggcg	ttatcttcct	cggaacggtt	60
gcatacttta	caaagggcaa	gctatggggg	gttaccaagg	atccctacgc	gaatggcttc	120
gctgccggcg	gcgcttctaa	gccgggtcgc	acgaggaaca	tcgtcgaggc	aatggaagaa	180

tccggcaaga	actgtgttgt	cttctatggt	tctcagaccg	gtactgctga	agattatgct	240
tctcgccctcg	ccaaggaggg	taagagtcga	ttcggactaa	acaccatgat	tgccgatctt	300
gaggactacg	atttcgacaa	cctggatacc	gttcccagtg	acaacattgt	catgttctgt	360
ctcgcaactt	atggtgaagg	tgagcctacc	gataacgcgg	tcgacttcta	tgaattcatt	420
accggcgagg	atgccagctt	caatgagggc	aatgatcctc	cgctgggcaa	cctcaactac	480
gttgctttcg	gtctcggaag	caacacgtac	gagcactaca	actctatggt	ccgcaatgtt	540
aacaaggctc	tcgagaagct	tggcgctcac	cgcatcggtg	aagctgggtga	gggtgatgat	600
ggtgctggta	ccatggaaga	ggacttcttg	gcctggaagg	atcccatgtg	ggaagccctc	660
gctaagaaaa	tgggactgga	agagcgtgaa	gcagtctacg	agcctatatt	tgccattaac	720
gaacgcgacg	acctgactcc	tgaagccaat	gaagtgtatc	tcggtgagcc	caacaagctg	780
catctcgaag	gcaccgccaa	gggaccattc	aactctcaca	acccctacat	tgcccctatc	840
gctgaatctt	atgagttggt	ctccgccaag	gacagaaact	gcctccacat	ggaaattgac	900
atcagcgggt	ctaacctcaa	gtacgaaact	ggagaccata	ttgctatctg	gcctaccaac	960
cctggtgagg	aggtcaacaa	attcctggat	attctcgacc	tctctggaaa	gcagcacagc	1020
gttgtcactg	tcaaggctct	cgagcctacc	gccaagggtc	ctttcccaa	ccctacaacc	1080
tacgatgcc	ttctgcgata	ccacctcgag	atctgcgctc	ctgtttcacg	tcaattcgtc	1140
tctactctcg	ccgcatttgc	tccaacgat	gatatcaagg	ctgagatgaa	ccgccttggc	1200
agcgataagg	attattttcca	cgagaagact	ggccccgcat	actacaacat	tgcccgtttc	1260
cttgccagcg	tcagcaaggg	cgagaagtgg	acaaaaatcc	cgttctctgc	cttcatcgag	1320
ggtctcacca	agctccagcc	ccgttactac	tccatttctt	cctcgtctct	ggttcagccc	1380
aagaaaatct	cgatcactgc	cgtcgttgaa	tcccagcaga	ttcctggccg	ggatgatcct	1440
ttccgtggtg	ttgctacaaa	ctatcttttt	gccctaaagc	aaaagcagaa	cggtgacccc	1500
aaccctgcac	cttttggtca	gagctacgag	cttacaggcc	cccgcataaa	gtatgatggc	1560
atccacgttc	ctgtccatgt	tcgtcactcc	aacttcaagc	tcccctcgga	ccccggtaag	1620
cccatcatca	tgatttggtc	tggtactggt	gtcgtctccct	tccgcgggtt	cgtgcaggag	1680
cgtgctaagc	aagcccgtga	tggtgttgag	gttggaagaa	cactcttggt	ctttggttgc	1740
cgaaagtcaa	ccgaggattt	catgtaccaa	aaggagtggc	aggaatacaa	ggaggctctt	1800
ggcgataagt	ttgaaatgat	caccgccttt	tctcgagagg	gctccaagaa	ggtttatggt	1860
cagcaccgac	ttaaggagcg	atctaaggag	gtcagcgatc	tgctctccca	gaaggcttat	1920
ttctatgtct	gcggtgatgc	agcccacatg	gcccgcgagg	tcaataaccgt	cttggcacia	1980
atcattgccg	agggacgtgg	ggtgtctgag	gccaagggcg	aggagatcgt	gaagaacatg	2040
agatcagcga	accaatacca	ggtatgtagt	gactttgtta	ctcttctactg	caaagaaacc	2100
acatatgcta	actcagaatt	acaggaggat	gtttggtcat	ag		2142

<211> 459

<212> PRT

<213> Ipomoea purpurea

<400> 76

Met Gly Ser Gln Ala Thr Thr Tyr His Met Ala Met Tyr Pro Trp Phe
 1 5 10 15

Gly Val Gly His Leu Thr Gly Phe Phe Arg Leu Ala Asn Lys Leu Ala
 20 25 30

Gly Lys Gly His Arg Ile Ser Phe Leu Ile Pro Lys Asn Thr Gln Ser
 35 40 45

Lys Leu Glu Ser Phe Asn Leu His Pro His Leu Ile Ser Phe Val Pro
 50 55 60

Ile Val Val Pro Ser Ile Pro Gly Leu Pro Pro Gly Ala Glu Thr Thr
 65 70 75 80

Ser Asp Val Pro Phe Pro Ser Thr His Leu Leu Met Glu Ala Met Asp
 85 90 95

Lys Thr Gln Asn Asp Ile Glu Ile Ile Leu Lys Asp Leu Lys Val Asp
 100 105 110

Val Val Phe Tyr Asp Phe Thr His Trp Leu Pro Ser Leu Ala Arg Lys
 115 120 125

Ile Gly Ile Lys Ser Val Phe Tyr Ser Thr Ile Ser Pro Leu Met His
 130 135 140

Gly Tyr Ala Leu Ser Pro Glu Arg Arg Val Val Gly Lys Gln Leu Thr
 145 150 155 160

Glu Ala Asp Met Met Lys Ala Pro Ala Ser Phe Pro Asp Pro Ser Ile
 165 170 175

Lys Leu His Ala His Glu Ala Arg Gly Phe Thr Ala Arg Thr Val Met
 180 185 190

Lys Phe Gly Gly Asp Ile Thr Phe Phe Asp Arg Ile Phe Thr Ala Val
 195 200 205

Ser Glu Ser Asp Gly Leu Ala Tyr Ser Thr Cys Arg Glu Ile Glu Gly
 210 215 220

Gln Phe Cys Asp Tyr Ile Glu Thr Gln Phe Gln Lys Pro Val Leu Leu
 225 230 235 240

Ala Gly Pro Ala Leu Pro Val Pro Ser Lys Ser Thr Met Glu Gln Lys
 245 250 255

Trp Ser Asp Trp Leu Gly Lys Phe Lys Glu Gly Ser Val Ile Tyr Cys
260 270

Ala Phe Gly Ser Glu Cys Thr Leu Arg Lys Asp Lys Phe Gln Glu Leu
275 280 285

Leu Trp Gly Leu Glu Leu Thr Gly Met Pro Phe Phe Ala Ala Leu Lys
290 295 300

Pro Pro Phe Glu Thr Glu Ser Val Glu Ala Ala Ile Pro Glu Glu Leu
305 310 315 320

Lys Glu Lys Ile Gln Gly Arg Gly Ile Val His Gly Glu Trp Val Gln
325 330 335

Gln Gln Leu Phe Leu Gln His Pro Ser Val Gly Cys Phe Val Ser His
340 345 350

Cys Gly Trp Ala Ser Leu Ser Glu Ala Leu Val Asn Asp Cys Gln Ile
355 360 365

Val Leu Leu Pro Gln Val Gly Asp Gln Ile Ile Asn Ala Arg Ile Met
370 375 380

Ser Val Ser Leu Lys Val Gly Val Glu Val Glu Lys Gly Glu Glu Asp
385 390 395 400

Gly Val Phe Ser Arg Glu Ser Val Cys Lys Ala Val Lys Ala Val Met
405 410 415

Asp Glu Lys Ser Glu Ile Gly Arg Glu Val Arg Gly Asn His Asp Lys
420 425 430

Leu Arg Gly Phe Leu Met Asn Ala Asp Leu Asp Ser Lys Tyr Met Asp
435 440 445

Ser Phe Asn Gln Lys Leu Gln Asp Leu Leu Gly
450 455

<210> 77

<211> 1380

<212> DNA

<213> Ipomoea purpurea

<400> 77

atgggttctc aagctacaac ttaccatgat gccatgtatc catgggtttgg ggttggacat	60
ttgactgggtt tcttccgttt ggcaaacaaa ttagctggca aaggacatag aatctcattt	120
ctaattccta aaaacactca atctaagtta gaatctttca accttcatcc acacttaatc	180
tcttttgtgc ctatcgttgt cccaagtata ccaggcctgc cacctggtgc agagactaca	240
tcagatgttc ctttcccaag tacacatttg ctaatggaag caatggacaa gactcaaaac	300
gatatagaga ttatcctgaa ggatcttaaa gtagatgttg ttttctatga ttttactcac	360

tggttgcctt ctctggccag aaagattggc attaagagtg tcttttactc caccatttct 420
 cctttaatgc atggatatgc tttatcacca gaaagacgtg tagttggtaa gcaattgaca 480
 gaggcagata tgatgaaggc cccagcttct ttcccagacc catccattaa gctacatgca 540
 catgaagcta ggggttttac agccagaacc gttatgaaat tcggtggtga catcaccttt 600
 ttcgatagaa tattcacagc agtttccgaa agtgatggcc tggcctactc tacttgtaga 660
 gagatcgagg gacaattctg tgattacatt gaaacacaat tccagaagcc agtcttgta 720
 gccggtccag ctttgccagt cccatccaaa tccactatgg aacaaaagtg gtcagattgg 780
 ttggggaaat tcaaggaagg ctccgtcatc tactgtgctt tcgggtctga atgtacattg 840
 agaaaggaca aatttcagga acttttatgg ggtttggaat tgacaggaat gcctttcttc 900
 gctgctctga agccaccttt tgagactgag tctgttgagg ctgctatccc tgaggaacta 960
 aaggaaaaga ttcaggaag aggtatagta catggagaat ggtacaaca acaattgttt 1020
 cttcaacacc catctgtcgg gtgcttcgtt tctactgcg gctgggcaag tttatctgaa 1080
 gcccttgta atgattgtca aatcgtgtta cttccacaag ttggcgatca gattatcaac 1140
 gccagaataa tgtcagtatc acttaaagtg ggcgtggaag ttgaaaaggg tgaggaggac 1200
 ggtgtctttt caagagaatc tgtgtgcaag gctgttaaag cagtaatgga tgaaaaatct 1260
 gaaatcggtg gagaagtcag aggtaatcat gataaactga ggggtttctt gatgaatgca 1320
 gacttagatt caaagtacat ggattcattc aatcaaaagc tacaagattt gctagggttaa 1380

<210> 78
 <211> 438
 <212> PRT
 <213> Bellis perennis

<400> 78
 Met Asp Ser Lys Ile Asp Ser Lys Thr Phe Arg Val Val Met Leu Pro
 1 5 10 15
 Trp Leu Ala Tyr Ser His Ile Ser Ser Phe Leu Val Phe Ala Lys Arg
 20 25 30
 Leu Thr Asn His Asn Phe His Ile Tyr Ile Cys Ser Ser Gln Thr Asn
 35 40 45
 Met Gln Tyr Leu Lys Asn Asn Leu Thr Ser Gln Tyr Ser Lys Ser Ile
 50 55 60
 Gln Leu Ile Glu Leu Asn Leu Pro Ser Ser Ser Glu Leu Pro Leu Gln
 65 70 75 80
 Tyr His Thr Thr His Gly Leu Pro Pro His Leu Thr Lys Thr Leu Ser
 85 90 95
 Asp Asp Tyr Gln Lys Ser Gly Pro Asp Phe Glu Thr Ile Leu Ile Lys
 100 105 110

Leu Asn Pro His Leu Val Ile Tyr Asp Phe Asn Gln Leu Trp Ala Pro
 115 120 125

Glu Val Ala Ser Thr Leu His Ile Pro Ser Ile Gln Leu Leu Ser Gly
 130 135 140

Cys Val Ala Leu Tyr Ala Leu Asp Ala His Leu Tyr Thr Lys Pro Leu
 145 150 155 160

Asp Glu Asn Leu Ala Lys Phe Pro Phe Pro Glu Ile Tyr Pro Lys Asn
 165 170 175

Arg Asp Ile Pro Lys Gly Gly Ser Lys Tyr Ile Glu Arg Phe Val Asp
 180 185 190

Cys Met Arg Arg Ser Cys Glu Ile Ile Leu Val Arg Ser Thr Met Glu
 195 200 205

Leu Glu Gly Lys Tyr Ile Asp Tyr Leu Ser Lys Thr Leu Gly Lys Lys
 210 215 220

Val Leu Pro Val Gly Pro Leu Val Gln Glu Ala Ser Leu Leu Gln Asp
 225 230 235 240

Asp His Ile Trp Ile Met Lys Trp Leu Asp Lys Lys Glu Glu Ser Ser
 245 250 255

Val Val Phe Val Cys Phe Gly Ser Glu Tyr Ile Leu Ser Asp Asn Glu
 260 265 270

Ile Glu Asp Ile Ala Tyr Gly Leu Glu Leu Ser Gln Val Ser Phe Val
 275 280 285

Trp Ala Ile Arg Ala Lys Thr Ser Ala Leu Asn Gly Phe Ile Asp Arg
 290 295 300

Val Gly Asp Lys Gly Leu Val Ile Asp Lys Trp Val Pro Gln Ala Asn
 305 310 315 320

Ile Leu Ser His Ser Ser Thr Gly Gly Phe Ile Ser His Cys Gly Trp
 325 330 335

Ser Ser Thr Met Glu Ser Ile Arg Tyr Gly Val Pro Ile Ile Ala Met
 340 345 350

Pro Met Gln Phe Asp Gln Pro Tyr Asn Ala Arg Leu Met Glu Thr Val
 355 360 365

Gly Ala Gly Ile Glu Val Gly Arg Asp Gly Glu Gly Arg Leu Lys Arg
 370 375 380

Glu Glu Ile Ala Ala Val Val Arg Lys Val Val Val Glu Asp Ser Gly
385 390 395 400

Glu Ser Ile Arg Glu Lys Ala Lys Glu Leu Gly Glu Ile Met Lys Lys
405 410 415

Asn Met Glu Ala Glu Val Asp Gly Ile Val Ile Glu Asn Leu Val Lys
420 425 430

Leu Cys Glu Met Asn Asn
435

<210> 79
<211> 1317
<212> DNA
<213> *Bellis perennis*

<400> 79
atggattcta aaatcgattc aaagacattc agagtcgtta tgttgccttg gcttgcatac 60
tcacacattt catcattcct agtgtttgcc aagagactaa caaatcataa cttccacatc 120
tacatttggt cctctcaaac aaatatgcaa tacctgaaaa acaacttgac gtctcagtat 180
tcaaaatcta tacaactgat tgagttgaat cttccatcta gttccgaatt gcctctgcag 240
tatcatacta ctacacggact accaccacac cttacgaaaa cattgtctga tgattatcaa 300
aagtccggac ctgactttga aaccattttg atcaaattga acccacatct ggtaatctac 360
gactttaatc aactttgggc tccagagggt gctagtacac ttcattattcc atccatacag 420
ttactgtctg gttgcgtcgc cttatatgcc ttagacgccc atctgtacac aaagccacta 480
gacgaaaact tggctaagtt tcctttccca gaaatctatc ctaaaaacag agatattcct 540
aagggaggta gtaaatacat cgaaagggtc gtagactgta tgagaagatc ttgtgaaatc 600
atattagtca gaagtaccat ggaacttgaa ggaaaataca ttgattactt gtctaagaca 660
ttagggaaaa aggtgttgcc agtagggcct ctggtgcaag aggcttcttt gttgcaagat 720
gatcatatat ggattatgaa gtggttagac aaaaaggagg agtcatccgt cgtgtttggt 780
tgttttggtt ctgagtacat cttatcagac aacgaaatag aagatattgc ttatggccta 840
gagttgtccc aagtaagttt cgtttgggca ataagagcta agacttctgc cttaaattggc 900
ttcattgata gagtgggtga taaaggctta gtcatcgata aatgggttcc acaggctaac 960
atcttatctc actcttctac tgggtggattc attagtcatt gcggttggtc atcaacaatg 1020
gaatctatta gatatggggg tcctattatc gccatgccaa tgcaattcga tcaaccttac 1080
aatgctaggt tgatggaaac tgttggtgca ggtatcgaag ttggcagaga tggcgaagg 1140
agattgaaaa gagaagagat tgctgccgtg gttagaaagg tcgttggtga agattctggg 1200
gaatccataa gggagaaggc aaaggaattg ggagaaatca tgaaaaaaaa catggaggcc 1260
gaagtagatg gtatagtgat tgaaaatcta gttaagctat gtgagatgaa caattaa 1317

<210> 80
<211> 2490

<212> DNA

<213> Zea mays

<400> 80

atgggtttt	gtg	cttcttcttg	tactacagta	ccacacttat	cttcattagc	tgctcgtgcaa	60
cttggtcctt	ggagcagtag	gattaaaaag	aaaaccgata	ctgttgagct	accagccgct		120
gcaggaaggt	ggagaagggc	cttggtctaga	gcacagcaca	catcagaatc	cgcagctgtc		180
gcaaagggca	gcagtttgac	ccctatagtg	agaactgacg	ctgagtcaag	gagaacaaga		240
tggccaaccg	atgacgatga	cgccgaacct	ttagtggatg	agatcagggc	aatgcttact		300
tccatgtctg	atgggtgacat	ttccgtgagc	gcatacgata	cagcctgggt	cggattgggt		360
ccaagattag	acggcgggtga	aggtcctcaa	tttccagcag	ctgtgagatg	gataagaaat		420
aaccagttgc	ctgacggaag	ttggggcgat	gccgcattat	tctctgccta	tgacaggcctt		480
atcaataccc	ttgcctgcgt	tgtaactttg	acaagggtgg	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	atgcataaccg	ttccaacatc	aatattgcac		780
agtttggagg	gtatgcctgg	cctagattgg	gctaaactac	ttaaactaca	gagcagcgac		840
ggaagttttt	tgttctcacc	agctgccact	gcatatgctt	taatgaatac	cggagatgac		900
aggtgtttta	gctacatcga	tagaacagta	aagaaattca	acggcggcgt	ccctaattgtt		960
tatccagtgg	atctatttga	acatatttgg	gccgttgata	gacttgaaag	attaggaatc		1020
tccaggtact	tccaaaagga	gatcgaacaa	tgcatggatt	atgtaaacag	gcattggact		1080
gaggacggta	tttgttgggc	aaggaaactct	gatgtcaaag	aggtggacga	cacagctatg		1140
gccttttagac	ttcttaggtt	gcacggctac	agcgtcagtc	ctgatgtgtt	taaaaacttc		1200
gaaaaggacg	gtgaattttt	cgcatttgtc	ggacagtcta	atcaagctgt	taccggtatg		1260
tacaacttaa	acagagcaag	ccagatatcc	ttcccaggcg	aggatgtgct	tcatagagct		1320
ggtgccttct	catatgagtt	cttgaggaga	aaagaagcag	aggagcttt	gagggacaag		1380
tggatcattt	ctaaagatct	acctggtgaa	gttgtgtata	ctttggattt	tccatggtac		1440
ggcaacttac	ctagagtcga	ggccagagac	tacctagagc	aatacggagg	tggatgatgac		1500
gtttggattg	gcaagacatt	gtataggatg	ccacttgtaa	acaatgatgt	atatttggaa		1560
ttggcaagaa	tggatttcaa	ccactgccag	gctttgcata	agttagagtg	gcaaggacta		1620
aaaagatggg	atactgaaaa	taggttgatg	gactttgggtg	tcgccaaga	agatgccctt		1680
agagcttatt	ttcttgcagc	cgcactgtgt	tacgagcctt	gtagagctgc	cagagaggctt		1740
gcatgggcta	gagccgcaat	actagctaac	gccgtgagca	cccacttaag	aaatagccca		1800
tcattcagag	aaagggttaga	gcattctctt	aggtgtagac	ctagtgaaga	gacagatggc		1860
tcctggttta	actcctcaag	tggctctgat	gcagtttttag	taaaggctgt	cttaagactt		1920
actgattcat	tagccagggga	agcacagcca	atccatggag	gtgaccacga	agatattata		1980

cacaagttgt taagatctgc ttgggccgag tgggttaggg aaaaggcaga cgctgccgat 2040
 agcgtgtgca atggtagttc tgcagtagaa caagagggat caagaatggg ccatgataaa 2100
 cagacctgtc tattattggc tagaatgata gaaatttctg ccggtagggc agctggtgaa 2160
 gcagccagtg aggacggcga tagaagaata attcaattaa caggctccat ctgacgacagt 2220
 cttaagcaaa aaatgctagt ttcacaggac cctgaaaaaa atgaagagat gatgtctcac 2280
 gtggatgacg aattgaagtt gaggattaga gagttcgttc aatatttgct tagactaggt 2340
 gaaaaaaaga ctggatctag cgaaaccagg caaacatttt taagtatagt gaaatcatgt 2400
 tactatgctg ctcatgccc acctcatgtc gttgatagac acattagtag agtgattttc 2460
 gagccagtaa gtgccgcaaa gtaaccgcgg 2490

<210> 81
 <211> 827
 <212> PRT
 <213> Zea mays

<400> 81
 Met Val Leu Ser Ser Ser Cys Thr Thr Val Pro His Leu Ser Ser Leu
 1 5 10 15

Ala Val Val Gln Leu Gly Pro Trp Ser Ser Arg Ile Lys Lys Lys Thr
 20 25 30

Asp Thr Val Ala Val Pro Ala Ala Ala Gly Arg Trp Arg Arg Ala Leu
 35 40 45

Ala Arg Ala Gln His Thr Ser Glu Ser Ala Ala Val Ala Lys Gly Ser
 50 55 60

Ser Leu Thr Pro Ile Val Arg Thr Asp Ala Glu Ser Arg Arg Thr Arg
 65 70 75 80

Trp Pro Thr Asp Asp Asp Asp Ala Glu Pro Leu Val Asp Glu Ile Arg
 85 90 95

Ala Met Leu Thr Ser Met Ser Asp Gly Asp Ile Ser Val Ser Ala Tyr
 100 105 110

Asp Thr Ala Trp Val Gly Leu Val Pro Arg Leu Asp Gly Gly Glu Gly
 115 120 125

Pro Gln Phe Pro Ala Ala Val Arg Trp Ile Arg Asn Asn Gln Leu Pro
 130 135 140

Asp Gly Ser Trp Gly Asp Ala Ala Leu Phe Ser Ala Tyr Asp Arg Leu
 145 150 155 160

Ile Asn Thr Leu Ala Cys Val Val Thr Leu Thr Arg Trp Ser Leu Glu
 165 170 175

Pro Glu Met Arg Gly Arg Gly Leu Ser Phe Leu Gly Arg Asn Met Trp
 180 185 190

Lys Leu Ala Thr Glu Asp Glu Glu Ser Met Pro Ile Gly Phe Glu Leu
 195 200 205

Ala Phe Pro Ser Leu Ile Glu Leu Ala Lys Ser Leu Gly Val His Asp
 210 215 220

Phe Pro Tyr Asp His Gln Ala Leu Gln Gly Ile Tyr Ser Ser Arg Glu
 225 230 235 240

Ile Lys Met Lys Arg Ile Pro Lys Glu Val Met His Thr Val Pro Thr
 245 250 255

Ser Ile Leu His Ser Leu Glu Gly Met Pro Gly Leu Asp Trp Ala Lys
 260 265 270

Leu Leu Lys Leu Gln Ser Ser Asp Gly Ser Phe Leu Phe Ser Pro Ala
 275 280 285

Ala Thr Ala Tyr Ala Leu Met Asn Thr Gly Asp Asp Arg Cys Phe Ser
 290 295 300

Tyr Ile Asp Arg Thr Val Lys Lys Phe Asn Gly Gly Val Pro Asn Val
 305 310 315 320

Tyr Pro Val Asp Leu Phe Glu His Ile Trp Ala Val Asp Arg Leu Glu
 325 330 335

Arg Leu Gly Ile Ser Arg Tyr Phe Gln Lys Glu Ile Glu Gln Cys Met
 340 345 350

Asp Tyr Val Asn Arg His Trp Thr Glu Asp Gly Ile Cys Trp Ala Arg
 355 360 365

Asn Ser Asp Val Lys Glu Val Asp Asp Thr Ala Met Ala Phe Arg Leu
 370 375 380

Leu Arg Leu His Gly Tyr Ser Val Ser Pro Asp Val Phe Lys Asn Phe
 385 390 395 400

Glu Lys Asp Gly Glu Phe Phe Ala Phe Val Gly Gln Ser Asn Gln Ala
 405 410 415

Val Thr Gly Met Tyr Asn Leu Asn Arg Ala Ser Gln Ile Ser Phe Pro
 420 425 430

Gly Glu Asp Val Leu His Arg Ala Gly Ala Phe Ser Tyr Glu Phe Leu
 435 440 445

Arg Arg Lys Glu Ala Glu Gly Ala Leu Arg Asp Lys Trp Ile Ile Ser
 450 455 460

Lys Asp Leu Pro Gly Glu Val Val Tyr Thr Leu Asp Phe Pro Trp Tyr
 465 470 475 480

Gly Asn Leu Pro Arg Val Glu Ala Arg Asp Tyr Leu Glu Gln Tyr Gly
 485 490 495

Gly Gly Asp Asp Val Trp Ile Gly Lys Thr Leu Tyr Arg Met Pro Leu
 500 505 510

Val Asn Asn Asp Val Tyr Leu Glu Leu Ala Arg Met Asp Phe Asn His
 515 520 525

Cys Gln Ala Leu His Gln Leu Glu Trp Gln Gly Leu Lys Arg Trp Tyr
 530 535 540

Thr Glu Asn Arg Leu Met Asp Phe Gly Val Ala Gln Glu Asp Ala Leu
 545 550 555 560

Arg Ala Tyr Phe Leu Ala Ala Ala Ser Val Tyr Glu Pro Cys Arg Ala
 565 570 575

Ala Glu Arg Leu Ala Trp Ala Arg Ala Ala Ile Leu Ala Asn Ala Val
 580 585 590

Ser Thr His Leu Arg Asn Ser Pro Ser Phe Arg Glu Arg Leu Glu His
 595 600 605

Ser Leu Arg Cys Arg Pro Ser Glu Glu Thr Asp Gly Ser Trp Phe Asn
 610 615 620

Ser Ser Ser Gly Ser Asp Ala Val Leu Val Lys Ala Val Leu Arg Leu
 625 630 635 640

Thr Asp Ser Leu Ala Arg Glu Ala Gln Pro Ile His Gly Gly Asp Pro
 645 650 655

Glu Asp Ile Ile His Lys Leu Leu Arg Ser Ala Trp Ala Glu Trp Val
 660 665 670

Arg Glu Lys Ala Asp Ala Ala Asp Ser Val Cys Asn Gly Ser Ser Ala
 675 680 685

Val Glu Gln Glu Gly Ser Arg Met Val His Asp Lys Gln Thr Cys Leu
 690 695 700

Leu Leu Ala Arg Met Ile Glu Ile Ser Ala Gly Arg Ala Ala Gly Glu
 705 710 715 720

Ala Ala Ser Glu Asp Gly Asp Arg Arg Ile Ile Gln Leu Thr Gly Ser

Ile Cys Asp Ser Leu Lys Gln Lys Met Leu Val Ser Gln Asp Pro Glu
740 745 750

Lys Asn Glu Glu Met Met Ser His Val Asp Asp Glu Leu Lys Leu Arg
755 760 765

Ile Arg Glu Phe Val Gln Tyr Leu Leu Arg Leu Gly Glu Lys Lys Thr
770 775 780

Gly Ser Ser Glu Thr Arg Gln Thr Phe Leu Ser Ile Val Lys Ser Cys
785 790 795 800

Tyr Tyr Ala Ala His Cys Pro Pro His Val Val Asp Arg His Ile Ser
805 810 815

Arg Val Ile Phe Glu Pro Val Ser Ala Ala Lys
820 825

<210> 82
<211> 1461
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 82
actagtaaaa tggatgcaat ggcaactact gagaaaaagc ctcatgtgat cttcattcca 60
tttctgcac aatctcacat aaaggcaatg cttaaagttag cacaactatt acaccataag 120
ggattacaga taactttcgt gaataccgac ttcattcata atcaatttct ggaatctagt 180
ggccctcatt gtttgacgg agccccaggg tttagattcg aaacaattcc tgacggtgtt 240
tcacattccc cagaggcctc catccaata agagagagtt tactgaggtc aatagaaacc 300
aactttttgg atcgtttcat tgacttggtc aaaaaactc cagaccacc aacttgcata 360
atctctgatg gctttctgtc agtgtttact atcgacgctg ccaaaaagtt gggatatcca 420
gttatgatgt actggactct tgctgcatgc ggtttcatgg gtttctatca catccattct 480
cttatcgaag agggttttgc tccactgaaa gatgcatcat acttaaccaa cggctacctg 540
gatactgtta ttgactgggt accaggtatg gaaggataa gacttaaaga ttttcctttg 600
gattggtcta cagacctta tgataaagta ttgatgttta ctacagaagc tccacaaaga 660
tctcataagg ttccacatca tatctttcac acctttgatg aattggaacc atcaatcatc 720
aaaaccttgt ctctaagata caatcatatc tacactattg gtccattaca attacttcta 780
gatcaaattc ctgaagagaa aaagcaaact ggtattacat ccttacacgg ctactcttta 840
gtgaaagagg aaccagaatg ttttcaatgg ctacaaagta aagagcctaa ttctgtggtc 900
tacgtcaact tcggaagtac aacagtcatg tccttggaag atatgactga atttggttgg 960

ggccttgcta attcaaatca ttactttcta tggattatca ggtccaattt ggtaataggg	1020
gaaaacgccg tattacctcc agaattggag gaacacatca aaaagagagg ttccattgct	1080
tcctggtggt ctcaggaaaa ggtattgaaa catccttctg ttggtgggtt ccttactcat	1140
tgcggttggg gctctacaat cgaatcacta agtgcaggag ttccaatgat ttggtggcca	1200
tattcatggg accaacttac aaattgtagg tatactctgta aagagtggga agttggatta	1260
gaaatgggaa caaagggttaa acgtgatgaa gtgaaaagat tggttcagga gttgatgggg	1320
gaagggtggc acaagatgag aaacaaggcc aaagattgga aggaaaaagc cagaattgct	1380
attgctccta acgggtcatc ctctctaaac attgataaga tggtaaaga gattacagtc	1440
ttagccagaa actaagtcga c	1461

<210> 83

<211> 1398

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 83

actagtaaaa tggcagagca acaaaagatc aaaaagtcac ctacgtctt acttattcca	60
tttctctgc aaggacatat caaccattc atacaatttg ggaaaagatt gattagtaag	120
ggtgtaaaga caaactggt aaccactatc cacactttga attctactct gaaccactca	180
aatactacta ctacaagtat agaaattcaa gctatatcag acggatgcga tgagggtggc	240
tttatgtctg ccggtgaatc ttacttggaa acattcaagc aagtgggatc caagtctctg	300
gccgatctaa tcaaaaagtt acagagtga ggcaccacaa ttgacgcat aatctacgat	360
tctatgacag agtgggtttt agacgttgct atcgaatttg gtattgatgg aggttccttt	420
ttcacacaag catgtgttgt gaattctcta tactaccatg tgcataaagg gtaatctct	480
ttaccattgg gtgaaactgt ttcagttcca ggttttccag tgttacaacg ttgggaaacc	540
ccattgatct taaaaatca tgaacaaata caatcacctt ggtcccagat gttgtttggt	600
caattcgcta acatcgatca agcaagatgg gtctttacta attcattcta taagttagag	660
gaagaggtaa ttgaatggac taggaagatc tggaatttga aagtcattgg tccaacattg	720
ccatcaatgt atttggacaa aagacttgat gatgataaag ataatgggtt caatttgtag	780
aaggctaatc atcacgaatg tatgaattgg ctggatgaca aaccaaagga atcagttgta	840
tatgttgctt tcggctctct tgttaaacad ggtccagaac aagttgagga gattacaaga	900
gcatttatag actctgacgt aaactttttg tgggtcatta agcacaaga ggaggggaaa	960
ctgccagaaa acctttctga agtgataaag accggaaaag gtctaatacgt tgcttggtgt	1020
aaacaattgg atgttttagc tcatgaatct gtaggctgtt ttgtaacaca ttgcggattc	1080
aactctacac tagaagccat ttccttaggc gtacctgtcg ttgcaatgcc tcagttctcc	1140
gatcagacaa ccaacgctaa acttttggac gaaatactag ggggtgggtgt cagagttaaa	1200

gcagacgaga atggtatcgt cagaagaggg aacctagctt catgtatcaa aatgatcatg	1260
gaagaggaaa gaggagttaa cataaggaaa aacgcagtta agtggaggga tcttgcaaag	1320
gttgccgtcc atgaaggcgg ctcttcagat aatgatattg ttgaatttgt gtccgaacta	1380
atcaaagcct aagtcgac	1398

<210> 84

<211> 1437

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 84

actagtaaaa tggctacatc tgattctatt gttgatgaca ggaagcagtt gcatgtggct	60
actttccctt ggcttgcttt cggcatata ctgccttacc tacaactatc aaaactgata	120
gctgaaaaag gacataaagt gtcattcctt tcaacaacta gaaacattca aagattatct	180
tcccacatat caccattgat taacgtcgtt caattgacac ttccaagagt acaggaatta	240
ccagaagatg ctgaagctac aacagatgtg catcctgaag atatccctta cttgaaaaag	300
gcatccgatg gattacagcc tgaggctact agattccttg agcaacacag tccagattgg	360
atcatatacg actacactca ctattgggtg ccttcaattg cagcatcact aggcatttct	420
agggcacatt tcagtgtaac cacaccttgg gccattgctt acatgggtcc atccgctgat	480
gctatgatta acggcagtga tggtagaact accgttgaag atttgacaac cccaccaaag	540
tggtttccat ttccaactaa agtctgttgg agaaaacacg acttagcaag actggttcca	600
tacaaggcac caggaatctc agacggctat agaatgggtt tagtccttaa agggctctgac	660
tgcctattgt ctaagtgtta ccatgagttt gggacacaat ggctaccact tttggaaca	720
ttacaccaag ttctgtcgt accagttggg ctattacctc cagaaatccc tggatgatgag	780
aaggacgaga cttgggtttc aatcaaaaag tggtttagacg ggaagcaaaa aggcctcagt	840
gtatatgtgg cactgggttc cgaagtttta gtatctcaaa cagaagttgt ggaacttgcc	900
ttaggtttgg aactatctgg attgccattt gtctgggcct acagaaaacc aaaaggccct	960
gcaaagtccg attcagttga attgccagac ggctttgtcg agagaactag agatagaggg	1020
ttggtatgga ctcatgggc tccacaattg agaatcctga gtcacgaatc tgtgtgctgt	1080
ttcctaacac attgtggttc tggttctata gttgaaggac tgatgtttgg tcatccactt	1140
atcatgttgc caatctttgg tgaccagcct ttgaatgcac gtctgttaga agataaacia	1200
gttggaattg aaatcccacg taatgaggaa gatggatgtt taaccaagga gtctgtggcc	1260
agatcattac gttccgttgt cgttgaaaag gaaggcgaat tctacaaggc caatgcccgt	1320
gaactttcaa agatctacaa tgacacaaaa gtagagaagg aatatgtttc tcaatttgta	1380
gattacctag agaaaaacgc tagagccgta gctattgatc atgaatccta agtcgac	1437

<210> 85

<211> 1392

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 85

```

actagtaaaa tggaaaacaa gaccgaaaca acagttagac gtaggcgtag aatcattctg      60
tttccagtac cttttcaagg gcacatcaat ccaatactac aactagccaa cgttttgtac      120
tctaaagggt tttctattac aatctttcac accaatttca acaaaccaa aacatccaat      180
taccacatt tcacattcag attcatactt gataatgatc cacaagatga acgtatttca      240
aacttaccta cccacggtcc tttagctgga atgagaattc caatcatcaa tgaacatggt      300
gccgatgagc ttagaagaga attagagtta cttatgttgg catccgaaga ggacgaggaa      360
gtctcttgtc tgattactga cgctctatgg tactttgccc aatctgtggc tgatagtttg      420
aatttgagga gattggtact aatgacatcc agtctgttta actttcacgc tcatgttagt      480
ttaccacaat ttgacgaatt gggatacttg gaccctgatg acaagactag gttagaggaa      540
caggcctctg gttttcctat gttgaaagtc aaagatatca agtctgccta ttctaattgg      600
caaatcttga aagagatctt aggaaagatg atcaaacaga caaaggcttc atctggagtg      660
at ttggaaca gtttcaaaga gttagaagag tctgaattgg agactgtaat cagagaaatt      720
ccagcacctt cattcctgat accattacca aaacatttga ctgcttcctc ttcctctttg      780
ttggatcatg acagaacagt ttttcaatgg ttggaccaac aaccacctag ttctgttttg      840
tacgtgtcat ttggtagtac ttctgaagtc gatgaaaagg acttccttga aatcgcaaga      900
ggcttagtcg atagtaagca gtcattcctt tgggtcgtgc gtccagggtt cgtgaaaggc      960
tcaacatggg tcgaaccact tccagatggg tttctaggcg aaagaggtag aatagtcaaa     1020
tggttctcctc aacaggaagt tttagctcat ggcgctattg gggcattctg gactcattcc     1080
ggatggaatt caactttaga atcagtatgc gaaggggtac ctatgatctt ttcagatttt     1140
ggctcttgatc aaccactgaa cgcaagatac atgtctgatg ttttgaaagt ggggtgtatat     1200
ctagaaaatg gctgggaaag gggtgaaata gctaattgcaa taagacgtgt tatggttgat     1260
gaagaggggg agtatatcag acaaaacgca agagtgctga agcaaaaaggc cgacgtttct     1320
ctaataaagg gaggtcttct atacgaatcc ttagaatctc ttgtttccta catttcatca     1380
ctgtaagtcg ac                                                    1392

```

<210> 86

<211> 12

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic peptide

<400> 86

Thr Ser Phe Ala Glu Tyr Trp Asn Leu Leu Ser Pro

<210> 87
 <211> 1602
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 87
 atgaccagct ttgccgagta ttggaatctg ttaagtccca cttcttttgc agaattattgg 60
 aaccttctat caccgacgag tttcgcgag tactggaatt tgttttctcc aacatcgttc 120
 gctgaatact ggaacttact cagccctgct agtaaaatgg atgcaatggc aactactgag 180
 aaaaagcctc atgtgatctt cattccattt cctgcacaat ctcacataaa ggcaatgcta 240
 aagttagcac aactattaca ccataagga ttacagataa ctttcgtgaa taccgacttc 300
 atccataatc aatttctgga atctagtggc cctcattgtt tggacggagc cccagggttt 360
 agattcgaaa caattcctga cgggtgttca cattccccag aggcctccat cccaataaga 420
 gagagtttac tgagggtcaat agaaaccaac tttttggatc gtttcattga cttggtcaca 480
 aaacttccag acccaccaac ttgcataatc tctgatggct ttctgtcagt gtttactatc 540
 gacgtgccca aaaagttggg tatcccagtt atgatgtact ggactcttgc tgcattgggt 600
 ttcattgggt tctatcacat ccattctctt atcgaaaagg gttttgctcc actgaaagat 660
 gcatcact taaccaacgg ctacctggat actgttattg actgggtacc aggtatggaa 720
 ggtataagac ttaaagattt tcctttggat tgggtctacag acctaatga taaagtattg 780
 atgtttacta cagaagctcc acaaagatct cataagggtt cacatcatat ctttcacacc 840
 tttgatgaat tggaaccatc aatcatcaaa acctgtctc taagatacaa tcatatctac 900
 actattgggc cattacaatt acttctagat caaattcctg aagagaaaaa gcaaactggt 960
 attacatcct tacacggcta ctcttttagtg aaagaggaac cagaatgttt tcaatggcta 1020
 caaagtaaag agcctaattc tgtggtctac gtcaacttcg gaagtacaac agtcatgtcc 1080
 ttggaagata tgactgaatt tgggtggggc cttgctaatt caaatcatta ctttctatgg 1140
 attatcaggt ccaatttggt aataggggaa aacgccgtat tacctccaga attggaggaa 1200
 cacatcaaaa agagagggtt cattgcttcc tgggtgttctc aggaaaagggt attgaaacat 1260
 ctttctgttg gtggtttcct tactcattgc ggttggggct ctacaatcga atcactaagt 1320
 gcaggagttc caatgatttg ttggccatat tcatgggacc aacttacaaa ttgtaggat 1380
 atctgtaaag agtgggaagt tggattagaa atgggaacaa aggttaaacy tgatgaagt 1440
 aaaagattgg ttcaggagtt gatgggggaa ggtggccaca agatgagaaa caaggccaaa 1500
 gattggaagg aaaaagccag aattgctatt gctcctaacy ggtcatcctc tctaaacatt 1560
 gataagatgg tcaaagagat tacagtctta gccagaaact aa 1602

<210> 88

<211> 533

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polypeptide

<400> 88

Met Thr Ser Phe Ala Glu Tyr Trp Asn Leu Leu Ser Pro Thr Ser Phe
 1 5 10 15

Ala Glu Tyr Trp Asn Leu Leu Ser Pro Thr Ser Phe Ala Glu Tyr Trp
 20 25 30

Asn Leu Phe Ser Pro Thr Ser Phe Ala Glu Tyr Trp Asn Leu Leu Ser
 35 40 45

Pro Ala Ser Lys Met Asp Ala Met Ala Thr Thr Glu Lys Lys Pro His
 50 55 60

Val Ile Phe Ile Pro Phe Pro Ala Gln Ser His Ile Lys Ala Met Leu
 65 70 75 80

Lys Leu Ala Gln Leu Leu His His Lys Gly Leu Gln Ile Thr Phe Val
 85 90 95

Asn Thr Asp Phe Ile His Asn Gln Phe Leu Glu Ser Ser Gly Pro His
 100 105 110

Cys Leu Asp Gly Ala Pro Gly Phe Arg Phe Glu Thr Ile Pro Asp Gly
 115 120 125

Val Ser His Ser Pro Glu Ala Ser Ile Pro Ile Arg Glu Ser Leu Leu
 130 135 140

Arg Ser Ile Glu Thr Asn Phe Leu Asp Arg Phe Ile Asp Leu Val Thr
 145 150 155 160

Lys Leu Pro Asp Pro Pro Thr Cys Ile Ile Ser Asp Gly Phe Leu Ser
 165 170 175

Val Phe Thr Ile Asp Ala Ala Lys Lys Leu Gly Ile Pro Val Met Met
 180 185 190

Tyr Trp Thr Leu Ala Ala Cys Gly Phe Met Gly Phe Tyr His Ile His
 195 200 205

Ser Leu Ile Glu Lys Gly Phe Ala Pro Leu Lys Asp Ala Ser Tyr Leu
 210 215 220

Thr Asn Gly Tyr Leu Asp Thr Val Ile Asp Trp Val Pro Gly Met Glu
 225 230 235 240

Gly Ile Arg Leu Lys Asp Phe Pro Leu Asp Trp Ser Thr Asp Leu Asn
245 250 255

Asp Lys Val Leu Met Phe Thr Thr Glu Ala Pro Gln Arg Ser His Lys
260 265 270

Val Ser His His Ile Phe His Thr Phe Asp Glu Leu Glu Pro Ser Ile
275 280 285

Ile Lys Thr Leu Ser Leu Arg Tyr Asn His Ile Tyr Thr Ile Gly Pro
290 295 300

Leu Gln Leu Leu Leu Asp Gln Ile Pro Glu Glu Lys Lys Gln Thr Gly
305 310 315 320

Ile Thr Ser Leu His Gly Tyr Ser Leu Val Lys Glu Glu Pro Glu Cys
325 330 335

Phe Gln Trp Leu Gln Ser Lys Glu Pro Asn Ser Val Val Tyr Val Asn
340 345 350

Phe Gly Ser Thr Thr Val Met Ser Leu Glu Asp Met Thr Glu Phe Gly
355 360 365

Trp Gly Leu Ala Asn Ser Asn His Tyr Phe Leu Trp Ile Ile Arg Ser
370 375 380

Asn Leu Val Ile Gly Glu Asn Ala Val Leu Pro Pro Glu Leu Glu Glu
385 390 395 400

His Ile Lys Lys Arg Gly Phe Ile Ala Ser Trp Cys Ser Gln Glu Lys
405 410 415

Val Leu Lys His Pro Ser Val Gly Gly Phe Leu Thr His Cys Gly Trp
420 425 430

Gly Ser Thr Ile Glu Ser Leu Ser Ala Gly Val Pro Met Ile Cys Trp
435 440 445

Pro Tyr Ser Trp Asp Gln Leu Thr Asn Cys Arg Tyr Ile Cys Lys Glu
450 455 460

Trp Glu Val Gly Leu Glu Met Gly Thr Lys Val Lys Arg Asp Glu Val
465 470 475 480

Lys Arg Leu Val Gln Glu Leu Met Gly Glu Gly Gly His Lys Met Arg
485 490 495

Asn Lys Ala Lys Asp Trp Lys Glu Lys Ala Arg Ile Ala Ile Ala Pro
500 505 510

Asn Gly Ser Ser Ser Leu Asn Ile Asp Lys Met Val Lys Glu Ile Thr

Val Leu Ala Arg Asn
530

<210> 89
<211> 1893
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 89
atgtgcaata ccaacatgtc tgtacctact gatggtgctg taaccacctc acagattcca 60
gcttcggaac aagagaccct ggtagacca aagccattgc ttttgaagtt attaaagtct 120
gttggtgcac aaaaagacac ttatactatg aaagagggtc ttttttatct tggccagtat 180
attatgacta aacgattata tgatgagaag caacaacata ttgtatattg ttcaaagtat 240
cttctaggag atttgtttgg cgtgccaagc ttctctgtga aagagcacag gaaaatatat 300
accatgatct acaggaactt ggtagtagtc aatcagcagg aatcatcgga ctcaggtaca 360
tctgtgagtg agaacaggtg tcaccttgaa ggtgggagtg atcaaaagga ccttgtaaaa 420
gagcttcagg aagagaaaacc ttcattcttc catttggttt ctagaccatc taccggtggt 480
agcggatcct ctggaggcag tgctagtaaa atggcagagc aaaaaagat caaaaagtca 540
cctcacgtct tacttatttc atttcctctg caaggacata tcaaccatt catacaattt 600
gggaaaagat tgattagtaa ggggtgtaaag acaacactgg taaccactat ccacactttg 660
aattctactc tgaaccactc aaatactact actacaagta tagaaattca agctatatca 720
gacggatgag atgaggggtg ctttatgtct gccgggtgaat ctacttgga aacattcaag 780
caagtgggat ccaagtctct ggccgatcta atcaaaaagt tacagagtga aggcaccaca 840
attgacgcca taatctacga ttctatgaca gagtgggttt tagacgttgc tatcgaattt 900
gggtattgatg gaggttcctt tttcacacaa gcatgtgttg tgaattctct atactacat 960
gtgcataaag gggttaatctc tttaccattg ggtgaaactg tttcagttcc aggttttcca 1020
gtgttacaac gttgggaaac cccattgatc ttacaaaatc atgaacaaat acaatcacct 1080
tggtcccaga tggtgttttg tcaattcgct aacatcgatc aagcaagatg ggtctttact 1140
aattcattct ataagttaga ggaagaggta attgaatgga cttaggaagat ctggaatttg 1200
aaagtcattg gtccaacatt gccatcaatg tatttggaac aaagacttga tgatgataaa 1260
gataatggtt tcaatttgta caaggctaata catcacgaat gtatgaattg gctggatgac 1320
aaaccaaagg aatcagttgt atatgttgct ttcggctctc ttgttaaaca tgggtccagaa 1380
caagttgagg agattacaag agcacttata gactctgacg taaacttttt gtgggtcatt 1440
aagcacaag aggaggggaa actgccagaa aacctttctg aagtgataaa gaccggaaaa 1500
ggctaatcg ttgcttggtg taaacaattg gatgttttag ctcattgaatc tgtaggctgt 1560

```

tttgaacac attgCGgatt caactctaca ctagaagcca tttccttagg cgtacctgtc      1620
gttgcaatgc ctcagttctc cgatcagaca accaacgcta aacttttgga cgaaatacta      1680
ggggtgggtg tcagagttaa agcagacgag aatggtatcg tcagaagagg gaacctagct      1740
tcatgtatca aaatgatcat ggaagaggaa agaggagtta tcataaggaa aaacgcagtt      1800
aagtggaagg atcttgcaaa ggttgccgtc catgaaggcg gctcttcaga taatgatatt      1860
gttgaatttg tgtccgaact aatcaaagcc taa                                     1893

```

<210> 90
 <211> 630
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic polypeptide

<400> 90
 Met Cys Asn Thr Asn Met Ser Val Pro Thr Asp Gly Ala Val Thr Thr
 1 5 10 15

 Ser Gln Ile Pro Ala Ser Glu Gln Glu Thr Leu Val Arg Pro Lys Pro
 20 25 30

 Leu Leu Leu Lys Leu Leu Lys Ser Val Gly Ala Gln Lys Asp Thr Tyr
 35 40 45

 Thr Met Lys Glu Val Leu Phe Tyr Leu Gly Gln Tyr Ile Met Thr Lys
 50 55 60

 Arg Leu Tyr Asp Glu Lys Gln Gln His Ile Val Tyr Cys Ser Asn Asp
 65 70 75 80

 Leu Leu Gly Asp Leu Phe Gly Val Pro Ser Phe Ser Val Lys Glu His
 85 90 95

 Arg Lys Ile Tyr Thr Met Ile Tyr Arg Asn Leu Val Val Val Asn Gln
 100 105 110

 Gln Glu Ser Ser Asp Ser Gly Thr Ser Val Ser Glu Asn Arg Cys His
 115 120 125

 Leu Glu Gly Gly Ser Asp Gln Lys Asp Leu Val Gln Glu Leu Gln Glu
 130 135 140

 Glu Lys Pro Ser Ser Ser His Leu Val Ser Arg Pro Ser Thr Gly Gly
 145 150 155 160

 Ser Gly Ser Ser Gly Gly Ser Ala Ser Lys Met Ala Glu Gln Gln Lys
 165 170 175

 Ile Lys Lys Ser Pro His Val Leu Leu Ile Pro Phe Pro Leu Gln Gly
 180 185 190

His Ile Asn Pro Phe Ile Gln Phe Gly Lys Arg Leu Ile Ser Lys Gly
 195 200 205
 Val Lys Thr Thr Leu Val Thr Thr Ile His Thr Leu Asn Ser Thr Leu
 210 215 220
 Asn His Ser Asn Thr Thr Thr Thr Ser Ile Glu Ile Gln Ala Ile Ser
 225 230 235 240
 Asp Gly Cys Asp Glu Gly Gly Phe Met Ser Ala Gly Glu Ser Tyr Leu
 245 250 255
 Glu Thr Phe Lys Gln Val Gly Ser Lys Ser Leu Ala Asp Leu Ile Lys
 260 265 270
 Lys Leu Gln Ser Glu Gly Thr Thr Ile Asp Ala Ile Ile Tyr Asp Ser
 275 280 285
 Met Thr Glu Trp Val Leu Asp Val Ala Ile Glu Phe Gly Ile Asp Gly
 290 295 300
 Gly Ser Phe Phe Thr Gln Ala Cys Val Val Asn Ser Leu Tyr Tyr His
 305 310 315 320
 Val His Lys Gly Leu Ile Ser Leu Pro Leu Gly Glu Thr Val Ser Val
 325 330 335
 Pro Gly Phe Pro Val Leu Gln Arg Trp Glu Thr Pro Leu Ile Leu Gln
 340 345 350
 Asn His Glu Gln Ile Gln Ser Pro Trp Ser Gln Met Leu Phe Gly Gln
 355 360 365
 Phe Ala Asn Ile Asp Gln Ala Arg Trp Val Phe Thr Asn Ser Phe Tyr
 370 375 380
 Lys Leu Glu Glu Glu Val Ile Glu Trp Thr Arg Lys Ile Trp Asn Leu
 385 390 395 400
 Lys Val Ile Gly Pro Thr Leu Pro Ser Met Tyr Leu Asp Lys Arg Leu
 405 410 415
 Asp Asp Asp Lys Asp Asn Gly Phe Asn Leu Tyr Lys Ala Asn His His
 420 425 430
 Glu Cys Met Asn Trp Leu Asp Asp Lys Pro Lys Glu Ser Val Val Tyr
 435 440 445
 Val Ala Phe Gly Ser Leu Val Lys His Gly Pro Glu Gln Val Glu Glu
 450 455 460

Ile Thr Arg Ala Leu Ile Asp Ser Asp Val Asn Phe Leu Trp Val Ile
465 470 475 480

Lys His Lys Glu Glu Gly Lys Leu Pro Glu Asn Leu Ser Glu Val Ile
485 490 495

Lys Thr Gly Lys Gly Leu Ile Val Ala Trp Cys Lys Gln Leu Asp Val
500 505 510

Leu Ala His Glu Ser Val Gly Cys Phe Val Thr His Cys Gly Phe Asn
515 520 525

Ser Thr Leu Glu Ala Ile Ser Leu Gly Val Pro Val Val Ala Met Pro
530 535 540

Gln Phe Ser Asp Gln Thr Thr Asn Ala Lys Leu Leu Asp Glu Ile Leu
545 550 555 560

Gly Val Gly Val Arg Val Lys Ala Asp Glu Asn Gly Ile Val Arg Arg
565 570 575

Gly Asn Leu Ala Ser Cys Ile Lys Met Ile Met Glu Glu Glu Arg Gly
580 585 590

Val Ile Ile Arg Lys Asn Ala Val Lys Trp Lys Asp Leu Ala Lys Val
595 600 605

Ala Val His Glu Gly Gly Ser Ser Asp Asn Asp Ile Val Glu Phe Val
610 615 620

Ser Glu Leu Ile Lys Ala
625 630

<210> 91
<211> 1932
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 91
atgtgcaata ccaacatgtc tgtacctact gatggtgctg taaccacctc acagattcca 60
gcttcggaac aagagaccct ggtagacca aagcattgc ttttgaagtt attaaagtct 120
gttggtgcac aaaaagacac ttatactatg aaagaggttc ttttttatct tggccagtat 180
attatgacta aacgattata tgatgagaag caacaacata ttgtatattg ttcaaatgat 240
cttctaggag atttgtttgg cgtgccaaagc ttctctgtga aagagcacag gaaaatatat 300
accatgatct acaggaactt ggtagtagtc aatcagcagg aatcatcgga ctcaggtaca 360
tctgtgagtg agaacaggtg tcaccttgaa ggtgggagtg atcaaaagga ccttgtacaa 420

```

gagcttcagg aagagaaacc ttcatcttca cttttgggtt ctagaccatc taccggtggt      480
agcggatcct ctggaggcag tgctagtaaa atggctacat ctgattctat tgttgatgac      540
aggaagcagt tgcattgtgg tactttccct tggttgcct tcggtcatat actgccttac      600
ctacaactat caaaactgat agctgaaaaa ggacataaag tgtcattcct ttcaacaact      660
agaaacattc aaagattatc ttcccacata tcaccattga ttaacgtcgt tcaattgaca      720
cttccaagag tacaggaatt accagaagat gctgaagcta caacagatgt gcattcctgaa      780
gatatccctt acttgaaaaa ggcattccgat ggattacagc ctgaggtcac tagattcctt      840
gagcaacaca gtccagattg gatcatatac gactacactc actattgggt gccttcaatt      900
gcagcatcac taggcatttc tagggcacat ttcagtgtaa ccacaccttg ggccattgct      960
tacatgggtc catccgctga tgctatgatt aacggcagtg atggtagaac taccgttgaa     1020
gatttgacaa cccacacaaa gtggtttcca tttccaacta aagtctgttg gagaaaacac     1080
gacttagcaa gactggttcc atacaaggca ccaggaatct cagacggcta tagaatgggt     1140
ttagtcctta aagggtctga ctgcctattg tctaagtgtt accatgagtt tgggacacaa     1200
tggtaccac ttttggaac attacaccaa gttcctgtcg taccagttgg tctattacct     1260
ccagaaatcc ctggtgatga gaaggacgag acttgggttt caatcaaaaa gtggttagac     1320
gggaagcaaa aaggctcagt ggtatatgtg gcactgggtt ccgaagtttt agtatctcaa     1380
acagaagttg tggaacttgc cttaggtttg gaactatctg gattgccatt tgtctgggcc     1440
tacagaaaac caaaaggccc tgcaaagtcc gattcagttg aattgccaga cggctttgtc     1500
gagagaacta gagatagagg gttggtatgg acttcatggg ctccacaatt gagaatcctg     1560
agtcacgaat ctgtgtgagg tttcctaaca cattgtggtt ctggttctat agttgaagga     1620
ctgatgtttg gtcattccat tatcatgttg ccaatctttg gtgaccagcc tttgaatgca     1680
cgtctgttag aagataaaca agttggaatt gaaatccac gtaatgagga agatggatgt     1740
ttaaccaagg agtctgtggc cagatcatta cgttccgttg tcgttgaaaa ggaaggcgaa     1800
atctacaagg ccaatgcccg tgaactttca aagatctaca atgacacaaa agtagagaag     1860
gaatatgttt ctcaatttgt agattaccta gagaaaaacg ctagagccgt agctattgat     1920
catgaatcct aa                                                    1932

```

<210> 92

<211> 643

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polypeptide

<400> 92

```

Met Cys Asn Thr Asn Met Ser Val Pro Thr Asp Gly Ala Val Thr Thr
1           5           10           15

```

```

Ser Gln Ile Pro Ala Ser Glu Gln Glu Thr Leu Val Arg Pro Lys Pro
          20           25           30

```

Leu Leu Leu Lys Leu Leu Lys Ser Val Gly Ala Gln Lys Asp Thr Tyr
 35 40 45
 Thr Met Lys Glu Val Leu Phe Tyr Leu Gly Gln Tyr Ile Met Thr Lys
 50 55 60
 Arg Leu Tyr Asp Glu Lys Gln Gln His Ile Val Tyr Cys Ser Asn Asp
 65 70 75 80
 Leu Leu Gly Asp Leu Phe Gly Val Pro Ser Phe Ser Val Lys Glu His
 85 90 95
 Arg Lys Ile Tyr Thr Met Ile Tyr Arg Asn Leu Val Val Val Asn Gln
 100 105 110
 Gln Glu Ser Ser Asp Ser Gly Thr Ser Val Ser Glu Asn Arg Cys His
 115 120 125
 Leu Glu Gly Gly Ser Asp Gln Lys Asp Leu Val Gln Glu Leu Gln Glu
 130 135 140
 Glu Lys Pro Ser Ser Ser His Leu Val Ser Arg Pro Ser Thr Gly Gly
 145 150 155 160
 Ser Gly Ser Ser Gly Gly Ser Ala Ser Lys Met Ala Thr Ser Asp Ser
 165 170 175
 Ile Val Asp Asp Arg Lys Gln Leu His Val Ala Thr Phe Pro Trp Leu
 180 185 190
 Ala Phe Gly His Ile Leu Pro Tyr Leu Gln Leu Ser Lys Leu Ile Ala
 195 200 205
 Glu Lys Gly His Lys Val Ser Phe Leu Ser Thr Thr Arg Asn Ile Gln
 210 215 220
 Arg Leu Ser Ser His Ile Ser Pro Leu Ile Asn Val Val Gln Leu Thr
 225 230 235 240
 Leu Pro Arg Val Gln Glu Leu Pro Glu Asp Ala Glu Ala Thr Thr Asp
 245 250 255
 Val His Pro Glu Asp Ile Pro Tyr Leu Lys Lys Ala Ser Asp Gly Leu
 260 265 270
 Gln Pro Glu Val Thr Arg Phe Leu Glu Gln His Ser Pro Asp Trp Ile
 275 280 285
 Ile Tyr Asp Tyr Thr His Tyr Trp Leu Pro Ser Ile Ala Ala Ser Leu
 290 295 300

Gly Ile Ser Arg Ala His Phe Ser Val Thr Thr Pro Trp Ala Ile Ala
305 310 315 320

Tyr Met Gly Pro Ser Ala Asp Ala Met Ile Asn Gly Ser Asp Gly Arg
325 330 335

Thr Thr Val Glu Asp Leu Thr Thr Pro Pro Lys Trp Phe Pro Phe Pro
340 345 350

Thr Lys Val Cys Trp Arg Lys His Asp Leu Ala Arg Leu Val Pro Tyr
355 360 365

Lys Ala Pro Gly Ile Ser Asp Gly Tyr Arg Met Gly Leu Val Leu Lys
370 375 380

Gly Ser Asp Cys Leu Leu Ser Lys Cys Tyr His Glu Phe Gly Thr Gln
385 390 395 400

Trp Leu Pro Leu Leu Glu Thr Leu His Gln Val Pro Val Val Pro Val
405 410 415

Gly Leu Leu Pro Pro Glu Ile Pro Gly Asp Glu Lys Asp Glu Thr Trp
420 425 430

Val Ser Ile Lys Lys Trp Leu Asp Gly Lys Gln Lys Gly Ser Val Val
435 440 445

Tyr Val Ala Leu Gly Ser Glu Val Leu Val Ser Gln Thr Glu Val Val
450 455 460

Glu Leu Ala Leu Gly Leu Glu Leu Ser Gly Leu Pro Phe Val Trp Ala
465 470 475 480

Tyr Arg Lys Pro Lys Gly Pro Ala Lys Ser Asp Ser Val Glu Leu Pro
485 490 495

Asp Gly Phe Val Glu Arg Thr Arg Asp Arg Gly Leu Val Trp Thr Ser
500 505 510

Trp Ala Pro Gln Leu Arg Ile Leu Ser His Glu Ser Val Cys Gly Phe
515 520 525

Leu Thr His Cys Gly Ser Gly Ser Ile Val Glu Gly Leu Met Phe Gly
530 535 540

His Pro Leu Ile Met Leu Pro Ile Phe Gly Asp Gln Pro Leu Asn Ala
545 550 555 560

Arg Leu Leu Glu Asp Lys Gln Val Gly Ile Glu Ile Pro Arg Asn Glu
565 570 575

Glu Asp Gly Cys Leu Thr Lys Glu Ser Val Ala Arg Ser Leu Arg Ser
580 585 590

Val Val Val Glu Lys Glu Gly Glu Ile Tyr Lys Ala Asn Ala Arg Glu
595 600 605

Leu Ser Lys Ile Tyr Asn Asp Thr Lys Val Glu Lys Glu Tyr Val Ser
610 615 620

Gln Phe Val Asp Tyr Leu Glu Lys Asn Ala Arg Ala Val Ala Ile Asp
625 630 635 640

His Glu Ser

<210> 93
<211> 1887
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 93
atgtgcaata ccaacatgtc tgtacctact gatggtgctg taaccacctc acagattcca 60
gcttcggaac aagagaccct ggtagacca aagccattgc ttttgaagt attaaagtct 120
gttggtgcac aaaaagacac ttatactatg aaagagggtc ttttttatct tggccagtat 180
attatgacta aacgattata tgatgagaag caacaacata ttgtatattg ttcaaatgat 240
cttctaggag atttgtttgg cgtgccaagc ttctctgtga aagagcacag gaaaatatat 300
accatgatct acaggaactt ggtagtagtc aatcagcagg aatcatcgga ctcaggtaca 360
tctgtgagtg agaacagggt tcaccttgaa ggtgggagt atcaaaaagga cttgtacaa 420
gagcttcagg aagagaaacc ttcattctca catttggttt ctagaccatc taccggtggt 480
agcggatcct ctggaggcag tgctagtaaa atggaaaaca agaccgaaac aacagttaga 540
cgtaggcgta gaatcattct gtttcagta cttttcaag ggcacatcaa tccaatacta 600
caactagcca acgttttcta ctctaaaggt ttttctatta caatctttca caccaatttc 660
aacaaccaa aaacatcaa ttaccacat ttcacattca gattcatact tgataatgat 720
ccacaagatg aacgtatttc aaacttacct acccacggtc ctttagctgg aatgagaatt 780
ccaatcatca atgaacatgg tgccgatgag cttagaagag aattagagtt acttatgttg 840
gcatccgaag aggacgagga agtctcttgt ctgattactg acgctctatg gtactttgcc 900
caatctgtgg ctgatagttt gaatttgagg agattgggtac taatgacatc cagtctgttt 960
aactttcacg ctcatgtag tttaccacaa ttgacgaat tgggatactt ggaccctgat 1020
gacaagacta ggtagagga acaggcctct ggttttccta tgttgaaagt caaagatatc 1080
aagtctgcct attctaattg gcaaatcttg aaagagatct taggaaagat gatcaaacag 1140
acaaaggctt catctggagt gatttggaac agtttcaaag agttagaaga gtctgaattg 1200

gagactgtaa tcagagaaat tccagcacct tcattcctga taccattacc aaaacatttg 1260
 actgcttcct cttcctcttt gttggatcat gacagaacag tttttcaatg gttggaccaa 1320
 caaccaccta gttctgtttt gtacgtgtca tttggtagta cttctgaagt cgatgaaaag 1380
 gacttccttg aaatcgcaag aggcttagtc gatagtaagc agtcattcct ttgggtcgtg 1440
 cgtccagggt tcgtgaaagg ctcaacatgg gtcgaaccac ttccagatgg ttttctaggc 1500
 gaaagaggta gaatagtcaa atgggttcct caacaggaag ttttagctca tggcgctatt 1560
 ggggcattct ggactcattc cggatggaat tcaactttag aatcagtatg cgaaggggta 1620
 cctatgatct tttcagattt tggcttgat caaccactga acgcaagata catgtctgat 1680
 gttttgaaag tgggtgtata tctagaaaat ggctgggaaa ggggtgaaat agctaattga 1740
 ataagacgtg ttatggttga tgaagagggg gagtatatca gacaaaacgc aagagtgtg 1800
 aagcaaaagg ccgacgtttc tctaataag ggaggctctt catacgaatc cttagaatct 1860
 cttgtttcct acatttcac actgtaa 1887

<210> 94

<211> 628

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polypeptide

<400> 94

Met Cys Asn Thr Asn Met Ser Val Pro Thr Asp Gly Ala Val Thr Thr
1 5 10 15

Ser Gln Ile Pro Ala Ser Glu Gln Glu Thr Leu Val Arg Pro Lys Pro
20 25 30

Leu Leu Leu Lys Leu Leu Lys Ser Val Gly Ala Gln Lys Asp Thr Tyr
35 40 45

Thr Met Lys Glu Val Leu Phe Tyr Leu Gly Gln Tyr Ile Met Thr Lys
50 55 60

Arg Leu Tyr Asp Glu Lys Gln Gln His Ile Val Tyr Cys Ser Asn Asp
65 70 75 80

Leu Leu Gly Asp Leu Phe Gly Val Pro Ser Phe Ser Val Lys Glu His
85 90 95

Arg Lys Ile Tyr Thr Met Ile Tyr Arg Asn Leu Val Val Val Asn Gln
100 105 110

Gln Glu Ser Ser Asp Ser Gly Thr Ser Val Ser Glu Asn Arg Cys His
115 120 125

Leu Glu Gly Gly Ser Asp Gln Lys Asp Leu Val Gln Glu Leu Gln Glu

130 135 140
 Glu Lys Pro Ser Ser Ser His Leu Val Ser Arg Pro Ser Thr Gly Gly
 145 150 155 160
 Ser Gly Ser Ser Gly Gly Ser Ala Ser Lys Met Glu Asn Lys Thr Glu
 165 170 175
 Thr Thr Val Arg Arg Arg Arg Arg Ile Ile Leu Phe Pro Val Pro Phe
 180 185 190
 Gln Gly His Ile Asn Pro Ile Leu Gln Leu Ala Asn Val Leu Tyr Ser
 195 200 205
 Lys Gly Phe Ser Ile Thr Ile Phe His Thr Asn Phe Asn Lys Pro Lys
 210 215 220
 Thr Ser Asn Tyr Pro His Phe Thr Phe Arg Phe Ile Leu Asp Asn Asp
 225 230 235 240
 Pro Gln Asp Glu Arg Ile Ser Asn Leu Pro Thr His Gly Pro Leu Ala
 245 250 255
 Gly Met Arg Ile Pro Ile Ile Asn Glu His Gly Ala Asp Glu Leu Arg
 260 265 270
 Arg Glu Leu Glu Leu Leu Met Leu Ala Ser Glu Glu Asp Glu Glu Val
 275 280 285
 Ser Cys Leu Ile Thr Asp Ala Leu Trp Tyr Phe Ala Gln Ser Val Ala
 290 295 300
 Asp Ser Leu Asn Leu Arg Arg Leu Val Leu Met Thr Ser Ser Leu Phe
 305 310 315 320
 Asn Phe His Ala His Val Ser Leu Pro Gln Phe Asp Glu Leu Gly Tyr
 325 330 335
 Leu Asp Pro Asp Asp Lys Thr Arg Leu Glu Glu Gln Ala Ser Gly Phe
 340 345 350
 Pro Met Leu Lys Val Lys Asp Ile Lys Ser Ala Tyr Ser Asn Trp Gln
 355 360 365
 Ile Leu Lys Glu Ile Leu Gly Lys Met Ile Lys Gln Thr Lys Ala Ser
 370 375 380
 Ser Gly Val Ile Trp Asn Ser Phe Lys Glu Leu Glu Glu Ser Glu Leu
 385 390 395 400
 Glu Thr Val Ile Arg Glu Ile Pro Ala Pro Ser Phe Leu Ile Pro Leu
 405 410 415

Pro Lys His Leu Thr Ala Ser Ser Ser Ser Leu Leu Asp His Asp Arg
 420 425 430

Thr Val Phe Gln Trp Leu Asp Gln Gln Pro Pro Ser Ser Val Leu Tyr
 435 440 445

Val Ser Phe Gly Ser Thr Ser Glu Val Asp Glu Lys Asp Phe Leu Glu
 450 455 460

Ile Ala Arg Gly Leu Val Asp Ser Lys Gln Ser Phe Leu Trp Val Val
 465 470 475 480

Arg Pro Gly Phe Val Lys Gly Ser Thr Trp Val Glu Pro Leu Pro Asp
 485 490 495

Gly Phe Leu Gly Glu Arg Gly Arg Ile Val Lys Trp Val Pro Gln Gln
 500 505 510

Glu Val Leu Ala His Gly Ala Ile Gly Ala Phe Trp Thr His Ser Gly
 515 520 525

Trp Asn Ser Thr Leu Glu Ser Val Cys Glu Gly Val Pro Met Ile Phe
 530 535 540

Ser Asp Phe Gly Leu Asp Gln Pro Leu Asn Ala Arg Tyr Met Ser Asp
 545 550 555 560

Val Leu Lys Val Gly Val Tyr Leu Glu Asn Gly Trp Glu Arg Gly Glu
 565 570 575

Ile Ala Asn Ala Ile Arg Arg Val Met Val Asp Glu Glu Gly Glu Tyr
 580 585 590

Ile Arg Gln Asn Ala Arg Val Leu Lys Gln Lys Ala Asp Val Ser Leu
 595 600 605

Met Lys Gly Gly Ser Ser Tyr Glu Ser Leu Glu Ser Leu Val Ser Tyr
 610 615 620

Ile Ser Ser Leu
 625

<210> 95
 <211> 473
 <212> PRT
 <213> Stevia rebaudiana

<400> 95
 Met Ala Thr Ser Asp Ser Ile Val Asp Asp Arg Lys Gln Leu His Val
 1 5 10 15

Ala Thr Phe Pro Trp Leu Ala Phe Gly His Ile Leu Pro Tyr Leu Gln

20

25

30

Leu Ser Lys₃₅ Leu Ile Ala Glu Lys₄₀ Gly His Lys Val Ser₄₅ Phe Leu Ser

Thr Thr₅₀ Arg Asn Ile Gln Arg₅₅ Leu Ser Ser His Ile₆₀ Ser Pro Leu Ile

Asn Val Val Gln Leu Thr₇₀ Leu Pro Arg Val₇₅ Gln Glu Leu Pro Glu Asp₈₀

Ala Glu Ala Thr₈₅ Thr Asp Val His Pro Glu₉₀ Asp Ile Pro Tyr Leu Lys₉₅

Lys Ala Ser Asp₁₀₀ Gly Leu Gln Pro Glu₁₀₅ Val Thr Arg Phe Leu₁₁₀ Glu Gln

His Ser Pro₁₁₅ Asp Trp Ile Ile Tyr₁₂₀ Asp Tyr Thr His Tyr₁₂₅ Trp Leu Pro

Ser Ile₁₃₀ Ala Ala Ser Leu Gly₁₃₅ Ile Ser Arg Ala His₁₄₀ Phe Ser Val Thr

Thr Pro Trp Ala Ile Ala Tyr Met Gly Pro Ser₁₅₅ Ala Asp Ala Met Ile₁₆₀

Asn Gly Ser Asp Gly₁₆₅ Arg Thr Thr Val Glu₁₇₀ Asp Leu Thr Thr Pro₁₇₅ Pro

Lys Trp Phe Pro₁₈₀ Phe Pro Thr Lys Val₁₈₅ Cys Trp Arg Lys His₁₉₀ Asp Leu

Ala Arg Leu₁₉₅ Val Pro Tyr Lys Ala₂₀₀ Pro Gly Ile Ser Asp₂₀₅ Arg Cys Arg

Met Gly₂₁₀ Leu Val Leu Lys Gly₂₁₅ Ser Asp Cys Leu₂₂₀ Leu Ser Lys Cys Tyr

His Glu Phe Gly Thr Gln₂₃₀ Trp Leu Pro Leu Leu₂₃₅ Glu Thr Leu His Gln₂₄₀

Val Pro Val Val₂₄₅ Pro Val Gly Leu Leu Pro₂₅₀ Pro Glu Ile Pro Gly₂₅₅ Asp

Glu Lys Asp Glu₂₆₀ Thr Trp Val Ser Ile₂₆₅ Lys Lys Trp Leu Asp₂₇₀ Gly Lys

Gln Lys Gly₂₇₅ Ser Val Val Tyr Val₂₈₀ Ala Leu Gly Ser Glu₂₈₅ Val Leu Val

Ser Gln Thr Glu Val Val Glu₂₉₅ Leu Ala Leu Gly Leu₃₀₀ Glu Leu Ser Gly

Leu Pro Phe Val Trp Ala Tyr Arg Lys Pro Lys Gly Pro Ala Lys Ser
305 310 315 320

Asp Ser Val Glu Leu Pro Asp Gly Phe Val Glu Arg Thr Arg Asp Arg
325 330 335

Gly Leu Val Trp Thr Ser Arg Ala Pro Gln Leu Arg Ile Leu Ser His
340 345 350

Glu Ser Val Cys Gly Phe Leu Thr His Cys Gly Ser Gly Ser Ile Val
355 360 365

Glu Gly Leu Met Phe Gly His Pro Leu Ile Met Leu Pro Ile Phe Gly
370 375 380

Asp Gln Pro Leu Asn Ala Arg Leu Leu Glu Asp Lys Gln Val Gly Ile
385 390 395 400

Glu Ile Pro Arg Asn Glu Glu Asp Gly Cys Leu Thr Lys Glu Ser Val
405 410 415

Ala Arg Ser Leu Arg Ser Val Val Val Glu Lys Glu Gly Glu Ile Tyr
420 425 430

Lys Ala Asn Ala Arg Glu Leu Ser Lys Ile Tyr Asn Asp Thr Lys Val
435 440 445

Glu Lys Glu Tyr Val Ser Gln Phe Val Asp Tyr Leu Glu Lys Asn Ala
450 455 460

Arg Ala Val Ala Ile Asp His Glu Ser
465 470

<210> 96
<211> 1422
<212> DNA
<213> Stevia rebaudiana

<400> 96
atggctacca gtgactccat agttgacgac cgtaagcagc ttcatgttgc gacgttccca 60
tggttgctt tcggtcacat cctcccttac cttcagcttt cgaaattgat agctgaaaag 120
ggtcacaaag tctcgtttct ttctaccacc agaaacattc aacgtctctc ttctcatatc 180
tcgccactca taaatgttgt tcaactcaca cttccacgtg tccaagagct gccggaggat 240
gcagaggcga ccaactgacgt ccaccctgaa gatattccat atctcaagaa ggcttctgat 300
ggcttcaac cggagggtcac ccggtttcta gaacaacact ctccggactg gattatttat 360
gattatactc actactggtt gccatccatc gcggctagcc tcggtatctc acgagcccac 420
ttctccgtca ccactccatg ggccattgct tatatgggac cctcagctga cgccatgata 480
aatggttcag atgggtcgaac cacggttgag gatctcacga caccgcccaa gtggtttccc 540

tttccgacca aagtatgctg gcggaagcat gatcttgccc gactggtgcc ttacaaagct	600
ccggggatat ctgatcgatg ccgtatgggg ctggttctta agggatctga ttgtttgctc	660
tccaaatgtt accatgagtt tggaactcaa tggctacctc ttttggagac actacaccaa	720
gtaccggtgg ttccggtggg attactgcca ccggaaatac ccggagacga gaaagatgaa	780
acatgggtgt caatcaagaa atggctcgat ggtaaacaaa aaggcagtgt ggtgtacgtt	840
gcattaggaa gcgaggtttt ggtgagccaa accgagggtt ttgagttagc attgggtctc	900
gagctttctg ggttgccatt tgtttgggct tatagaaaac caaaagggtcc cgcgaggtca	960
gactcgggtg agttgccaga cgggttcgtg gaacgaactc gtgaccgtgg gttggtctgg	1020
acgagtcggg cacctcagtt acgaatactg agccatgagt cggtttgtgg gttcttgacg	1080
catttgtggtt ctggatcaat tgtggaaggg ctaatgtttg gtcaccctct aatcatgcta	1140
ccgatttttg gggaccaacc tctgaatgct cgattactgg aggacaaaca ggtgggaatc	1200
gagataccaa gaaatgagga agatggttgc ttgaccaagg agtcggttgc tagatcactg	1260
aggtccgttg ttgtggaaaa agaaggggag atctacaagg cgaacgcgag ggagctgagt	1320
aaaatctata acgacactaa ggttgaaaaa gaatatgtaa gccaattcgt agactatttg	1380
gaaaagaatg cgcggtgcgtg tgccatcgat catgagagtt aa	1422

<210> 97

<211> 1380

<212> DNA

<213> Stevia rebaudiana

<400> 97

atggaaaata aaacgggagac caccgttcgc cggcgccgga gaataatatt attcccggta	60
ccatttcaag gccacattaa cccaattctt cagctagcca atgtgttgta ctctaaagga	120
ttcagtatca ccatctttca caccaacttc aacaaacca aaacatctaa ttaccctcac	180
ttcactttca gattcatcct cgacaacgac ccacaagacg aacgcatttc caatctaccg	240
actcatggtc cgctcgctgg tatgcggatt ccgattatca acgaacacgg agctgacgaa	300
ttacgacgcg aactggaact gttgatgtta gcttctgaag aagatgaaga ggtatcgtgt	360
ttaatcacgg atgctctttg gtacttcgag caatctgttg ctgacagtct taacctccga	420
cggcttgttt tgatgacaag cagcttgttt aattttcatg cacatgtttc acttcctcag	480
tttgatgagc ttggttacct cgatcctgat gacaaaaccc gtttggaaga acaagcgagt	540
gggtttccta tgctaaaagt gaaagacatc aagtctgcgt attcgaactg gcaaatactc	600
aaagagatat tagggaagat gataaaacaa acaaaagcat cttcaggagt catctggaac	660
tcatttaagg aactcgaaga gtctgagctc gaaactgtta tccgtgagat cccggctcca	720
agtttcttga taccactccc caagcatttg acagcctctt ccagcagctt actagaccac	780
gatcgaaccg tttttccatg gttagaccaa caaccgtcac gttcggtagt gtatgttagt	840
tttggtagtg gtactgaagt actggatgag aaagatttct tggaaatagc tcgtgggttg	900
gttgatagca agcagtcgtt tttatgggtg gttcgacctg gggttcgacg	960

tggttcgaac cgttgccaga tgggttcttg ggtgaaagag gacgtattgt gaaatggggt	1020
ccacagcaag aagtgctagc tcatggagca ataggcgcat tctggactca tagcggatgg	1080
aactctacgt tggaagcgt ttgtgaaggt gttcctatga ttttctcgga ttttgggctc	1140
gatcaaccgt tgaatgctag atacatgagt gatgttttga aggtaggggt gtatttggaa	1200
aatgggtggg aaagaggaga gatagcaaat gcaataagaa gagttatggt ggatgaagaa	1260
ggagaataca ttagacagaa tgcaagagtt ttgaaacaaa aggcagatgt ttctttgatg	1320
aagggtgggt cgtcttacga atcattagag tctctagttt cttacatttc atcgttgtaa	1380

<210> 98
 <211> 459
 <212> PRT
 <213> Stevia rebaudiana

<400> 98
 Met Glu Asn Lys Thr Glu Thr Thr Val Arg Arg Arg Arg Arg Ile Ile
 1 5 10 15

Leu Phe Pro Val Pro Phe Gln Gly His Ile Asn Pro Ile Leu Gln Leu
 20 25 30

Ala Asn Val Leu Tyr Ser Lys Gly Phe Ser Ile Thr Ile Phe His Thr
 35 40 45

Asn Phe Asn Lys Pro Lys Thr Ser Asn Tyr Pro His Phe Thr Phe Arg
 50 55 60

Phe Ile Leu Asp Asn Asp Pro Gln Asp Glu Arg Ile Ser Asn Leu Pro
 65 70 75 80

Thr His Gly Pro Leu Ala Gly Met Arg Ile Pro Ile Ile Asn Glu His
 85 90 95

Gly Ala Asp Glu Leu Arg Arg Glu Leu Glu Leu Leu Met Leu Ala Ser
 100 105 110

Glu Glu Asp Glu Glu Val Ser Cys Leu Ile Thr Asp Ala Leu Trp Tyr
 115 120 125

Phe Ala Gln Ser Val Ala Asp Ser Leu Asn Leu Arg Arg Leu Val Leu
 130 135 140

Met Thr Ser Ser Leu Phe Asn Phe His Ala His Val Ser Leu Pro Gln
 145 150 155 160

Phe Asp Glu Leu Gly Tyr Leu Asp Pro Asp Asp Lys Thr Arg Leu Glu
 165 170 175

Glu Gln Ala Ser Gly Phe Pro Met Leu Lys Val Lys Asp Ile Lys Ser
 180 185 190

Ala Tyr Ser Asn Trp Gln Ile Leu Lys Glu Ile Leu Gly Lys Met Ile
 195 200 205
 Lys Gln Thr Lys Ala Ser Ser Gly Val Ile Trp Asn Ser Phe Lys Glu
 210 215 220
 Leu Glu Glu Ser Glu Leu Glu Thr Val Ile Arg Glu Ile Pro Ala Pro
 225 230 235 240
 Ser Phe Leu Ile Pro Leu Pro Lys His Leu Thr Ala Ser Ser Ser Ser
 245 250 255
 Leu Leu Asp His Asp Arg Thr Val Phe Pro Trp Leu Asp Gln Gln Pro
 260 265 270
 Ser Arg Ser Val Leu Tyr Val Ser Phe Gly Ser Gly Thr Glu Val Leu
 275 280 285
 Asp Glu Lys Asp Phe Leu Glu Ile Ala Arg Gly Leu Val Asp Ser Lys
 290 295 300
 Gln Ser Phe Leu Trp Val Val Arg Pro Gly Phe Val Lys Gly Ser Thr
 305 310 315 320
 Trp Val Glu Pro Leu Pro Asp Gly Phe Leu Gly Glu Arg Gly Arg Ile
 325 330 335
 Val Lys Trp Val Pro Gln Gln Glu Val Leu Ala His Gly Ala Ile Gly
 340 345 350
 Ala Phe Trp Thr His Ser Gly Trp Asn Ser Thr Leu Glu Ser Val Cys
 355 360 365
 Glu Gly Val Pro Met Ile Phe Ser Asp Phe Gly Leu Asp Gln Pro Leu
 370 375 380
 Asn Ala Arg Tyr Met Ser Asp Val Leu Lys Val Gly Val Tyr Leu Glu
 385 390 395 400
 Asn Gly Trp Glu Arg Gly Glu Ile Ala Asn Ala Ile Arg Arg Val Met
 405 410 415
 Val Asp Glu Glu Gly Glu Tyr Ile Arg Gln Asn Ala Arg Val Leu Lys
 420 425 430
 Gln Lys Ala Asp Val Ser Leu Met Lys Gly Gly Ser Ser Tyr Glu Ser
 435 440 445
 Leu Glu Ser Leu Val Ser Tyr Ile Ser Ser Leu
 450 455

<210> 99
 <211> 1380
 <212> DNA
 <213> *Stevia rebaudiana*

<220>
 <221> modified_base
 <222> (861)..(863)
 <223> a, c, t, g, unknown or other

<400> 99
 atggaaaata aaacggagac caccgttcgc cggcgccgga gaataatatt attccccgta 60
 ccatttcaag gccacattaa cccaattcct cagctagcca atgtgttgta ctctaaagga 120
 ttcagtatca ccatctttca caccaacttc aacaaacca aaacatctaa ttaccctcac 180
 ttcactttca gattcatcct cgacaacgac ccacaagacg aacgcatttc caatctaccg 240
 actcatggtc cgctcgctgg tatgctgatt ccgattatca acgaacacgg agctgacgaa 300
 ttacgacgcg aactggaact gttgatgtta gcttctgaag aagatgaaga ggtatcgtgt 360
 ttaatcacgg atgctctttg gtacttcgcg caatctgttg ctgacagtct taacctccga 420
 cggcttggtt tgatgacaag cagcttggtt aattttcatg cacatgtttc acttcctcag 480
 tttgatgagc ttggttacct cgatcctgat gacaaaaccc gtttggaga acaagcgagt 540
 gggtttccta tgctaaaagt gaaagacatc aagtctgctg attcgaactg gcaaatactc 600
 aaagagatat tagggaagat gataaaacaa acaaaagcat cttcaggagt catctggaac 660
 tcatttaagg aactcgaaga gtctgagctc gaaactgtta tccgtgagat cccggctcca 720
 agtttcttga taccactccc caagcatttg acagcctctt ccagcagctt actagaccac 780
 gatcgaaccg tttttcaatg gtttagacaa caaccgcaa gttcggtagt gtatgttagt 840
 tttggtagta ctagtgaagt nnnngatgag aaagatttct tggaaaatagc tcgtgggttg 900
 gttgatagca agcagtcgtt tttatgggtg gttcgacctg ggtttgtcaa gggttcgacg 960
 tgggtcgaac cgttgccaga tgggttcgtg gccgaaagag ggcgtattgt gaaatgggtt 1020
 ccgcaacagg aagtgatagc tcatggagca atcgggtgcat tctggactca tagcggatgg 1080
 aactctacat tggaaagcgt ttgtgaaggt gttcctatga ttttctcga ttttgggctc 1140
 gatcaaccgt tgaatgctag atacatgagt gatgttttga aggtaggggt gtatttggaa 1200
 aatgggtggg aaagaggaga gatagcaaat gcaatacgaa gagttatggt ggatgaagaa 1260
 ggagaatata ttagacagaa tgcaagagtt ttgaaacaaa aggcagatgt ttctttgatg 1320
 aagggtggtt catcttacga atcattagag tctctagttt cttacatttc atcgttgtaa 1380

<210> 100
 <211> 459
 <212> PRT
 <213> *Stevia rebaudiana*

<220>
 <221> MOD_RES
 <222> (288)..(288)
 <223> Any amino acid

<400> 100

Met Glu Asn Lys Thr Glu Thr Thr Val Arg Arg Arg Arg Arg Ile Ile
1 5 10 15

Leu Phe Pro Val Pro Phe Gln Gly His Ile Asn Pro Ile Leu Gln Leu
20 25 30

Ala Asn Val Leu Tyr Ser Lys Gly Phe Ser Ile Thr Ile Phe His Thr
35 40 45

Asn Phe Asn Lys Pro Lys Thr Ser Asn Tyr Pro His Phe Thr Phe Arg
50 55 60

Phe Ile Leu Asp Asn Asp Pro Gln Asp Glu Arg Ile Ser Asn Leu Pro
65 70 75 80

Thr His Gly Pro Leu Ala Gly Met Arg Ile Pro Ile Ile Asn Glu His
85 90 95

Gly Ala Asp Glu Leu Arg Arg Glu Leu Glu Leu Leu Met Leu Ala Ser
100 105 110

Glu Glu Asp Glu Glu Val Ser Cys Leu Ile Thr Asp Ala Leu Trp Tyr
115 120 125

Phe Ala Gln Ser Val Ala Asp Ser Leu Asn Leu Arg Arg Leu Val Leu
130 135 140

Met Thr Ser Ser Leu Phe Asn Phe His Ala His Val Ser Leu Pro Gln
145 150 155 160

Phe Asp Glu Leu Gly Tyr Leu Asp Pro Asp Asp Lys Thr Arg Leu Glu
165 170 175

Glu Gln Ala Ser Gly Phe Pro Met Leu Lys Val Lys Asp Ile Lys Ser
180 185 190

Ala Tyr Ser Asn Trp Gln Ile Leu Lys Glu Ile Leu Gly Lys Met Ile
195 200 205

Lys Gln Thr Lys Ala Ser Ser Gly Val Ile Trp Asn Ser Phe Lys Glu
210 215 220

Leu Glu Glu Ser Glu Leu Glu Thr Val Ile Arg Glu Ile Pro Ala Pro
225 230 235 240

Ser Phe Leu Ile Pro Leu Pro Lys His Leu Thr Ala Ser Ser Ser Ser
245 250 255

Leu Leu Asp His Asp Arg Thr Val Phe Gln Trp Leu Asp Gln Gln Pro
260 265 270

Pro Ser Ser Val Leu Tyr Val Ser Phe Gly Ser Thr Ser Glu Val Xaa
 275 280 285

Asp Glu Lys Asp Phe Leu Glu Ile Ala Arg Gly Leu Val Asp Ser Lys
 290 295 300

Gln Ser Phe Leu Trp Val Val Arg Pro Gly Phe Val Lys Gly Ser Thr
 305 310 315 320

Trp Val Glu Pro Leu Pro Asp Gly Phe Val Ala Glu Arg Gly Arg Ile
 325 330 335

Val Lys Trp Val Pro Gln Gln Glu Val Ile Ala His Gly Ala Ile Gly
 340 345 350

Ala Phe Trp Thr His Ser Gly Trp Asn Ser Thr Leu Glu Ser Val Cys
 355 360 365

Glu Gly Val Pro Met Ile Phe Ser Asp Phe Gly Leu Asp Gln Pro Leu
 370 375 380

Asn Ala Arg Tyr Met Ser Asp Val Leu Lys Val Gly Val Tyr Leu Glu
 385 390 395 400

Asn Gly Trp Glu Arg Gly Glu Ile Ala Asn Ala Ile Arg Arg Val Met
 405 410 415

Val Asp Glu Glu Gly Glu Tyr Ile Arg Gln Asn Ala Arg Val Leu Lys
 420 425 430

Gln Lys Ala Asp Val Ser Leu Met Lys Gly Gly Ser Ser Tyr Glu Ser
 435 440 445

Leu Glu Ser Leu Val Ser Tyr Ile Ser Ser Leu
 450 455

<210> 101
 <211> 1380
 <212> DNA
 <213> Stevia rebaudiana

<220>
 <221> modified_base
 <222> (861)..(863)
 <223> a, c, t, g, unknown or other

<400> 101
 atggaata aaacggagac caccgttcgc cggcgccgga gaataatatt attcccggta 60
 ccatttcaag gccacattaa cccaattcct cagctagcca atgtgttgta ctctaaagga 120
 ttcagtatca ccatctttca caccaacttc aacaaacca aaacatctaa ttaccctcac 180
 ttcactttca gattcatcct cgacaacgac ccacaagacg aacgcatttc caatctaccg 240

```

actcatggtc cgctcgctgg tatgcggatt ccgattatca acgaacacgg agctgacgaa      300
ttacgacgcg aactggaact gttgatgtta gcttctgaag aagatgaaga ggtatcgtgt      360
ttaatcacgg atgctctttg gtacttcgcg caatctgttg ctgacagtct taacctccga      420
cggcttgttt tgatgacaag cagcttgttt aattttcatg cacatgtttc acttcctcag      480
tttgatgagc ttggttacct cgatcctgat gacaaaaccc gtttgaaga acaagcgagt      540
gggtttccta tgctaaaagt gaaagacatc aagtctgcgt attcgaactg gcaaatactc      600
aaagagatat tagggaagat gataaaacaa acaaaagcat cttcaggagt catctggaac      660
tcatttaagg aactcgaaga gtctgagctc gaaactgtta tccgtgagat cccggctcca      720
agtttcttga taccactccc caagcatttg acagcctctt ccagcagctt actagaccac      780
gatcgaaccg tttttcaatg gttagaccaa caaccgcaa gttcggtagt gtatgttagt      840
tttggtagta ctagtgaagt nnnggatgag aaagatttct tggaaatagc tcgtgggttg      900
gttgatagca agcagtcgtt tttatgggtg gttcgacctg ggtttgtcaa gggttcgacg      960
tgggtcgaac cgttgccaga tgggttcttg ggtgaaagag gacgtattgt gaaatgggtt     1020
ccacagcaag aagtgctagc tcatggagca ataggcgcat tctggactca tagcggatgg     1080
aactctacgt tggaaagcgt ttgtgaaggt gttcctatga ttttctcgga ttttgggctc     1140
gatcaaccgt tgaatgctag atacatgagt gatgttttga aggtaggggt gtatttggaa     1200
aatgggtggg aaagaggaga gatagcaaat gcaataagaa gagttatggt ggatgaagaa     1260
ggagaataca ttagacagaa tgcaagagtt ttgaaacaaa aggcagatgt ttctttgatg     1320
aagggtgggt cgtcttacga atcattagag tctctagttt cttacatttc atcgttgtaa     1380

```

```

<210> 102
<211> 459
<212> PRT
<213> Stevia rebaudiana

```

```

<220>
<221> MOD_RES
<222> (288)..(288)
<223> Any amino acid

```

```

<400> 102
Met Glu Asn Lys Thr Glu Thr Thr Val Arg Arg Arg Arg Arg Ile Ile
1          5          10          15

```

```

Leu Phe Pro Val Pro Phe Gln Gly His Ile Asn Pro Ile Leu Gln Leu
20          25          30

```

```

Ala Asn Val Leu Tyr Ser Lys Gly Phe Ser Ile Thr Ile Phe His Thr
35          40          45

```

```

Asn Phe Asn Lys Pro Lys Thr Ser Asn Tyr Pro His Phe Thr Phe Arg
50          55          60

```

```

Phe Ile Leu Asp Asn Asp Pro Gln Asp Glu Arg Ile Ser Asn Leu Pro
65          70          75          80

```

Thr His Gly Pro Leu Ala Gly Met Arg Ile Pro Ile Ile Asn Glu His
 85 90 95
 Gly Ala Asp Glu Leu Arg Arg Glu Leu 105 Glu Leu Leu Met Leu Ala Ser
 100 110
 Glu Glu Asp Glu Glu Val Ser Cys Leu Ile Thr Asp Ala Leu Trp Tyr
 115 120 125
 Phe Ala Gln Ser Val Ala Asp Ser Leu Asn Leu Arg Arg Leu Val Leu
 130 135 140
 Met Thr Ser Ser Leu Phe Asn Phe His Ala His Val Ser Leu Pro Gln
 145 150 155 160
 Phe Asp Glu Leu Gly Tyr Leu Asp Pro Asp Asp Lys Thr Arg Leu Glu
 165 170 175
 Glu Gln Ala Ser Gly Phe Pro Met Leu Lys Val Lys Asp Ile Lys Ser
 180 185 190
 Ala Tyr Ser Asn Trp Gln Ile Leu Lys Glu Ile Leu Gly Lys Met Ile
 195 200 205
 Lys Gln Thr Lys Ala Ser Ser Gly Val Ile Trp Asn Ser Phe Lys Glu
 210 215 220
 Leu Glu Glu Ser Glu Leu Glu Thr Val Ile Arg Glu Ile Pro Ala Pro
 225 230 235 240
 Ser Phe Leu Ile Pro Leu Pro Lys His Leu Thr Ala Ser Ser Ser Ser
 245 250 255
 Leu Leu Asp His Asp Arg Thr Val Phe Gln Trp Leu Asp Gln Gln Pro
 260 265 270
 Pro Ser Ser Val Leu Tyr Val Ser Phe Gly Ser Thr Ser Glu Val Xaa
 275 280 285
 Asp Glu Lys Asp Phe Leu Glu Ile Ala Arg Gly Leu Val Asp Ser Lys
 290 295 300
 Gln Ser Phe Leu Trp Val Val Arg Pro Gly Phe Val Lys Gly Ser Thr
 305 310 315 320
 Trp Val Glu Pro Leu Pro Asp Gly Phe Leu Gly Glu Arg Gly Arg Ile
 325 330 335
 Val Lys Trp Val Pro Gln Gln Glu Val Leu Ala His Gly Ala Ile Gly
 340 345 350

Ala Phe Trp Thr His Ser Gly Trp Asn Ser Thr Leu Glu Ser Val Cys
 355 360 365

Glu Gly Val Pro Met Ile Phe Ser Asp Phe Gly Leu Asp Gln Pro Leu
 370 375 380

Asn Ala Arg Tyr Met Ser Asp Val Leu Lys Val Gly Val Tyr Leu Glu
 385 390 395 400

Asn Gly Trp Glu Arg Gly Glu Ile Ala Asn Ala Ile Arg Arg Val Met
 405 410 415

Val Asp Glu Glu Gly Glu Tyr Ile Arg Gln Asn Ala Arg Val Leu Lys
 420 425 430

Gln Lys Ala Asp Val Ser Leu Met Lys Gly Gly Ser Ser Tyr Glu Ser
 435 440 445

Leu Glu Ser Leu Val Ser Tyr Ile Ser Ser Leu
 450 455

<210> 103

<211> 1323

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 103

atgcattcta ccagacatat cttaagacaa agggccgtcc tagttacagg cgctagaaca	60
ccattcgtga aatcatttgg ggctcttatg aaagcagata ccttggaatt ggcacagca	120
tcagtcgctg ggttgctgaa caagacctca ctggacccta gagatatcga tcatatcgtt	180
tggggtaatg ttgtacttca aggatcagct cataactgcg ccagagaaat agttatcgac	240
cttaacatgc ctaaaaagat catcggtaat ttgacatcta tggcctgtgc ttcaggctta	300
tcttctttgt cacaagcctg tatgctaata gagggtgggtc atgccgatgt cgtcattgct	360
ggcggttctg attcagtcctc caacactgaa gtgcctttgc caagatccgt cacttacggt	420
ctaattgatgg cccaaaggaa ggggtttatg ggcttcttta aggaagcagg atacaacca	480
ttcaaagtgt ttccaggcgg tattgcttta accgaacgta gtacaggaaa aactatgggt	540
tggcatggag acttaattgc tgagttaaac tctatatcta gagatgacca ggaagccctg	600
gctgtggcct ctcatgcaaa tgctgctaga gcagaaaaag ctgggtactt taaggaggaa	660
attgtacctg tgacaatcga caaaaagggc aaaaagactg aagtaacatg tgatgatgtt	720
atgcaaagag atacagaaaa gatgaaggcc aagatgccat cattgaagcc tgttttcaga	780
aaagaggggag gtacaataac agcagccact tccagtactc tgactgatgg tggctctgca	840
atgttggtta tgtcagagga aaaggccaaa aagttgggtt atccaactga tgtctgcgtg	900

aagtcttgggt atttcagtgg tatcgatcct tacccacaac ttttgtagc accagttcta 960
ggttgggggtc cagctttgaa aaaggccgga ttaacccta aagatatcga tttgtacgaa 1020
attcacgaag catttgctgc acaagttcta gccacaatta agtgtttgaa gtctcaggaa 1080
ttcttcgata ggtacgctaa cggtgcaaag ccagtattaa ctgaggatat tgatctttct 1140
aaactaaatg ttaatggcgg ttccttagca cttggccacc cattcgccgc tacaggaggt 1200
agaatcgtaa tctctctagc aaatgagttg agaagatccg gaaagagaca cgggctgggtc 1260
agtattttgtg cagctggagg gttaggcgga gtagctatac ttgagcatac agcaagtaag 1320
taa 1323

<210> 104

<211> 440

<212> PRT

<213> Leishmania infantum

<400> 104

Met His Ser Thr Arg His Ile Leu Arg Gln Arg Ala Val Leu Val Thr
1 5 10 15

Gly Ala Arg Thr Pro Phe Val Lys Ser Phe Gly Ala Leu Met Lys Ala
20 25 30

Asp Thr Leu Glu Leu Ala Ser Ala Ser Val Ala Gly Leu Leu Asn Lys
35 40 45

Thr Ser Leu Asp Pro Arg Asp Ile Asp His Ile Val Trp Gly Asn Val
50 55 60

Val Leu Gln Gly Ser Ala His Asn Cys Ala Arg Glu Ile Val Ile Asp
65 70 75 80

Leu Asn Met Pro Lys Lys Ile Ile Gly Asn Leu Thr Ser Met Ala Cys
85 90 95

Ala Ser Gly Leu Ser Ser Leu Ser Gln Ala Cys Met Leu Ile Glu Gly
100 105 110

Gly His Ala Asp Val Val Ile Ala Gly Gly Ser Asp Ser Val Ser Asn
115 120 125

Thr Glu Val Pro Leu Pro Arg Ser Val Thr Tyr Gly Leu Met Met Ala
130 135 140

Gln Arg Lys Gly Val Met Gly Phe Phe Lys Glu Ala Gly Tyr Asn Pro
145 150 155 160

Phe Lys Trp Phe Pro Gly Gly Ile Ala Leu Thr Glu Arg Ser Thr Gly
165 170 175

Lys Thr Met Gly Trp His Gly Asp Leu Ile Ala Glu Leu Asn Ser Ile
180 185 190

Ser Arg Asp Asp Gln Glu Ala Leu Ala Val Ala Ser His Ala Asn Ala
 195 200 205
 Ala Arg Ala Glu Lys Ala Gly Tyr Phe Lys Glu Glu Ile Val Pro Val
 210 215 220
 Thr Ile Asp Lys Lys Gly Lys Lys Thr Glu Val Thr Cys Asp Asp Val
 225 230 235 240
 Met Gln Arg Asp Thr Glu Lys Met Lys Ala Lys Met Pro Ser Leu Lys
 245 250 255
 Pro Val Phe Arg Lys Glu Gly Gly Thr Ile Thr Ala Ala Thr Ser Ser
 260 265 270
 Thr Leu Thr Asp Gly Gly Ser Ala Met Leu Val Met Ser Glu Glu Lys
 275 280 285
 Ala Lys Lys Leu Gly Tyr Pro Thr Asp Val Cys Val Lys Ser Trp Tyr
 290 295 300
 Phe Ser Gly Ile Asp Pro Tyr Pro Gln Leu Leu Leu Ala Pro Val Leu
 305 310 315 320
 Gly Trp Gly Pro Ala Leu Lys Lys Ala Gly Leu Thr Pro Lys Asp Ile
 325 330 335
 Asp Leu Tyr Glu Ile His Glu Ala Phe Ala Ala Gln Val Leu Ala Thr
 340 345 350
 Ile Lys Cys Leu Lys Ser Gln Glu Phe Phe Asp Arg Tyr Ala Asn Gly
 355 360 365
 Ala Lys Pro Val Leu Thr Glu Asp Ile Asp Leu Ser Lys Leu Asn Val
 370 375 380
 Asn Gly Gly Ser Leu Ala Leu Gly His Pro Phe Ala Ala Thr Gly Gly
 385 390 395 400
 Arg Ile Val Ile Ser Leu Ala Asn Glu Leu Arg Arg Ser Gly Lys Arg
 405 410 415
 His Gly Leu Val Ser Ile Cys Ala Ala Gly Gly Leu Gly Gly Val Ala
 420 425 430
 Ile Leu Glu His Thr Ala Ser Lys
 435 440

<210> 105
 <211> 1584
 <212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 105

atggcagctg accaattggt gaaaactgaa gtcaccaaga agtcttttac tgctcctgta	60
caaaaggctt ctacaccagt ttttaaccaat aaaacagtca tttctggatc gaaagtcaaa	120
agtttatcat ctgcgcaatc gagctcatca ggaccttcac catctagtga ggaagatgat	180
tcccgcgata ttgaaagctt ggataagaaa atacgtcctt tagaagaatt agaagcatta	240
ttaagtagtg gaaatacaaa acaattgaag aacaaagagg tcgctgcctt ggttattcac	300
ggtaagttac ctttgtacgc tttggagaaa aaattagggtg atactacgag agcgggtgcg	360
gtacgtagga aggtctcttc aattttggca gaagctcctg tattagcatc tgatcgttta	420
ccatataaaa attatgacta cgaccgcgta tttggcgctt gttgtgaaaa tgttataggt	480
tacatgcctt tgcccgttgg tgttataggc cccttgggtta tcgatggtac atcttatcat	540
ataccaatgg caactacaga gggttgtttg gtagcttctg ccatgcgtgg ctgtaaggca	600
atcaatgctg gcggtggtgc aacaactggt ttaactaagg atggtatgac aagaggccca	660
gtagtccgtt tcccaacttt gaaaagatct ggtgcctgta agatatggtt agactcagaa	720
gagggacaaa acgcaattaa aaaagctttt aactctacat caagatttgc acgtctgcaa	780
catattcaaa cttgtctagc aggagattta ctcttcatga gatttagaac aactactggt	840
gacgcaatgg gtatgaatat gatttctaaa ggtgtcgaat actcattaa gcaaatggta	900
gaagagtatg gctgggaaga tatggaggtt gtctccgttt ctggtaacta ctgtaccgac	960
aaaaaaccag ctgccatcaa ctggatcgaa ggtcgtggta agagtgtcgt cgcagaagct	1020
actattcctg gtgatgttgt cagaaaagtg ttaaaaagtg atgtttccgc attggttgag	1080
ttgaacattg ctaagaattht ggttggatct gcaatggctg ggtctgttgg tggatttaac	1140
gcacatgcag ctaatttagt gacagctggt ttcttggcat taggacaaga tcctgcacaa	1200
aatgttgaaa gttccaactg tataacattg atgaaagaag tggacggtga tttgagaatt	1260
tccgtatcca tgccatccat cgaagtaggt accatcgggt gtggtactgt tctagaacca	1320
caaggtgccg tggtggactt attaggtgta agaggcccg atgctaccgc tcctggtacc	1380
aacgcacgtc aattagcaag aatagttgcc tgtgccgtct tggcaggtga attatcctta	1440
tgtgtgcccc tagcagccgg ccatttgggt caaagtcata tgaccacaa caggaaacct	1500
gctgaaccaa caaaacctaa caatttggac gccactgata taaatcggtt gaaagatggg	1560
tccgtcacct gcattaaatc ctaa	1584

<210> 106

<211> 527

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 106

Met Ala Ala Asp Gln Leu Val Lys Thr Glu Val Thr Lys Lys Ser Phe

1

5

10

15

Thr Ala Pro Val Gln Lys Ala Ser Thr Pro Val Leu Thr Asn Lys Thr
20 25 30

Val Ile Ser Gly Ser Lys Val Lys Ser Leu Ser Ser Ala Gln Ser Ser
35 40 45

Ser Ser Gly Pro Ser Ser Ser Ser Glu Glu Asp Asp Ser Arg Asp Ile
50 55 60

Glu Ser Leu Asp Lys Lys Ile Arg Pro Leu Glu Glu Leu Glu Ala Leu
65 70 75 80

Leu Ser Ser Gly Asn Thr Lys Gln Leu Lys Asn Lys Glu Val Ala Ala
85 90 95

Leu Val Ile His Gly Lys Leu Pro Leu Tyr Ala Leu Glu Lys Lys Leu
100 105 110

Gly Asp Thr Thr Arg Ala Val Ala Val Arg Arg Lys Ala Leu Ser Ile
115 120 125

Leu Ala Glu Ala Pro Val Leu Ala Ser Asp Arg Leu Pro Tyr Lys Asn
130 135 140

Tyr Asp Tyr Asp Arg Val Phe Gly Ala Cys Cys Glu Asn Val Ile Gly
145 150 155 160

Tyr Met Pro Leu Pro Val Gly Val Ile Gly Pro Leu Val Ile Asp Gly
165 170 175

Thr Ser Tyr His Ile Pro Met Ala Thr Thr Glu Gly Cys Leu Val Ala
180 185 190

Ser Ala Met Arg Gly Cys Lys Ala Ile Asn Ala Gly Gly Gly Ala Thr
195 200 205

Thr Val Leu Thr Lys Asp Gly Met Thr Arg Gly Pro Val Val Arg Phe
210 215 220

Pro Thr Leu Lys Arg Ser Gly Ala Cys Lys Ile Trp Leu Asp Ser Glu
225 230 235 240

Glu Gly Gln Asn Ala Ile Lys Lys Ala Phe Asn Ser Thr Ser Arg Phe
245 250 255

Ala Arg Leu Gln His Ile Gln Thr Cys Leu Ala Gly Asp Leu Leu Phe
260 265 270

Met Arg Phe Arg Thr Thr Thr Gly Asp Ala Met Gly Met Asn Met Ile
275 280 285

Ser Lys Gly Val Glu Tyr Ser Leu Lys Gln Met Val Glu Glu Tyr Gly
 290 295 300

Trp Glu Asp Met Glu Val Val Ser Val Ser Gly Asn Tyr Cys Thr Asp
 305 310 315 320

Lys Lys Pro Ala Ala Ile Asn Trp Ile Glu Gly Arg Gly Lys Ser Val
 325 330 335

Val Ala Glu Ala Thr Ile Pro Gly Asp Val Val Arg Lys Val Leu Lys
 340 345 350

Ser Asp Val Ser Ala Leu Val Glu Leu Asn Ile Ala Lys Asn Leu Val
 355 360 365

Gly Ser Ala Met Ala Gly Ser Val Gly Gly Phe Asn Ala His Ala Ala
 370 375 380

Asn Leu Val Thr Ala Val Phe Leu Ala Leu Gly Gln Asp Pro Ala Gln
 385 390 395 400

Asn Val Glu Ser Ser Asn Cys Ile Thr Leu Met Lys Glu Val Asp Gly
 405 410 415

Asp Leu Arg Ile Ser Val Ser Met Pro Ser Ile Glu Val Gly Thr Ile
 420 425 430

Gly Gly Gly Thr Val Leu Glu Pro Gln Gly Ala Met Leu Asp Leu Leu
 435 440 445

Gly Val Arg Gly Pro His Ala Thr Ala Pro Gly Thr Asn Ala Arg Gln
 450 455 460

Leu Ala Arg Ile Val Ala Cys Ala Val Leu Ala Gly Glu Leu Ser Leu
 465 470 475 480

Cys Ala Ala Leu Ala Ala Gly His Leu Val Gln Ser His Met Thr His
 485 490 495

Asn Arg Lys Pro Ala Glu Pro Thr Lys Pro Asn Asn Leu Asp Ala Thr
 500 505 510

Asp Ile Asn Arg Leu Lys Asp Gly Ser Val Thr Cys Ile Lys Ser
 515 520 525

<210> 107

<211> 3681

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial sequence: Synthetic

polynucleotide

<400> 107

atgagagctg	tccttagatt	gttatcaaca	catactgttt	tctctcctat	tgaacaatt	60
gtatctgttt	tcgtgttagc	tacattagct	tacttcaca	tcttgtccg	aatcaagcac	120
tcaagtttct	ttgcatcttc	tcctcctcct	gctatcagac	ctgcttttgc	acatctgacc	180
aacggggaat	gggttgccgt	ctcccaacat	gattggactg	aagcatggaa	gcctcctggc	240
ggttcacttg	atgcattaga	acttcaacaa	gtagttttca	ctttagatga	caagactcaa	300
ccatctgctg	tgctagatgc	atccgcaatt	agtcagcact	tagtttccaa	tgttcctgca	360
ttatctggaa	aagcctactc	ttcattgtgc	caccatccaa	atgtatcagg	cacctcctgt	420
tttacatcag	tttctggtcc	aggagcttca	ccaatcttga	cactgagttt	taagcctgga	480
actagagacg	attggttagg	atcattaagg	aaggagaaaa	ctatcacact	agatgggggtt	540
aagtacgacg	ttggagccgg	aaaaagacaa	gagtcaatcg	gcgatatgga	atcatctaag	600
tgggttgctt	atgcattatc	agctttggta	cttagatttt	gggaattaac	aaaggcagat	660
tccttagata	tactagtggg	tctaactggg	tacatcctaa	tgcacgtaac	attcatgaga	720
ttgttcttgg	catccagagc	acttggcagt	aacttttggg	tatcagctgg	catattctcc	780
tccgcaacaa	tttctttcct	attcacttta	ccaatgtgta	gatctatgga	tattccactt	840
gatccaattg	ccttgacaga	agccctgcca	ttcttggtgt	gtaccgtagg	ttttgacaaa	900
ccacttagat	tggcaagagc	tgtgatggct	catcctaata	tccttaaacc	tcaagatgat	960
ggtaggatga	aagctgccgg	agatgtcatt	cttgaggcac	tggacagagt	tggtaacatg	1020
atattgagag	attacgcttt	agagatcgca	gttctattcg	ttggcgtaa	ctccagagtt	1080
ggcgggtctta	aggaattttg	tgctgtagct	gcagcattac	ttgctatgga	cagattaatg	1140
acattcacac	tttatacagc	agtgttaacc	atcatgggtg	aggtaaggcg	tatcaaaaag	1200
gtcagagata	tgactaaggc	tagatctaga	agttcttcta	ttaccgccgt	tacagccaac	1260
ggcaccgcca	taagaggcgt	tttgagtaga	aaatcttcaa	aacaatctgt	gacagaacca	1320
gagacaacta	aaaacctaag	acaaagagcc	actgattcag	ccatcggtgt	taagggttca	1380
ttgctgaaag	atggaggcag	attgcaggaa	gccgaggaga	atccaatggc	aagattaaag	1440
ctattgttaa	tcgcttcctt	cttaacacta	cacatcttga	acttttgtag	tactttgact	1500
tcagccacag	ctaacgcaag	acatcaaaga	catcctttta	gaaccgttca	agaggtagta	1560
ccaattccta	gagttgacat	tactaccca	gccatagcca	atatcttgct	tcattctagct	1620
gtggctcagg	aacctatggt	cactgttggt	ggcagtgaac	ctatcgaact	tcttggttaa	1680
gtcgtgctc	cagtctacgt	ccatgctcta	ccattggccc	ctgctttaag	agcttcaaac	1740
actaatactg	gagaagctat	tgaaaacttt	atgagttcat	ggctagtct	ggtaggtgac	1800
ccagttgtta	gtaagtggat	cgtagcattg	ctagctgtct	ctgttgcat	gaatggatac	1860
ttgttaaagg	gtatagccgc	aggttccggg	ttggctgcca	tgagagctgt	tagatctcaa	1920
gggttctggt	tcagatctag	agctagaagt	atcgtaaaga	tatctgatga	acctgagcca	1980

gagccagaac actctatcga cccagcacca gtagtgttct tcgcttccgc agcaccagct 2040
 gtagaggccc ctgctccagc tcctgcacct gaaccagaac caccagtcaa cagaccacca 2100
 ccattgacta ttttctcaag accactgaac ttagaaacag tggacaaaaa gttacaagat 2160
 gctctgccaa taagatcccc accacctgtt gaaccaatca ctccagaatc tagagaagtg 2220
 gaaccaaccc aagtagaagt aagatctcta gctgaatgtg tggatgtgtt cgagaatggg 2280
 ccaagaccag tctcagtggc tttaaagact ctgaatgatg aggaagttat cctgctttgc 2340
 caaacaggta agatagctcc atatgcattg gttaagatgt tggctgattt cgatagggcc 2400
 gtacgtgtca gaagagcact tattagtaga gcttcacgta caaaaacttt agaaaactca 2460
 ctggttccta tgaaagatta tgattacgcc agagtcattg gtgcctgttg tgaaaacgtt 2520
 atcggataca tgccattacc actagggatt gcagggtccat tgaagattga tggcttgatg 2580
 tatcctatac caatggcaac cgcagaaggc accttggttg catctacttc taggggctgt 2640
 aaggccttaa atgctggtgg aggggtcaca actgtcttga cagcagatgg catgacaaga 2700
 gggccagcta tagactttcc ttccatcgtc agagctgcag aggctaaggc cttcattgaa 2760
 tcagaagatg gatacgctac aatcaggag gctttcagat ctacttctag atttgccaa 2820
 ttgcaaaaga tcaagtgtgc actagctggt cgtactcttt ttgtcagatt tgctactaga 2880
 acaggagatg ccatgggtat gaacatgatt tctaaggcta ccgaaaaggc acttgatgtc 2940
 ctgagtcacg agttccctga aatggtcgtc cttgctttgt ctggtaacta ctgcacagac 3000
 aaaaagcctg cagctatttc atggatcgaa ggtaggggaa aatctattgt agcagaagca 3060
 gttattcctg gtaaggctgt taagtcagtc ctgaaaacaa cagtcgagtc tctttgcaat 3120
 gtcaacacta agaaaaacct gattggttca gccatggcag gttctgttgg tggtttcaac 3180
 gctcatgccg ccaacatcct aacagctgtg ttcctagcca caggtcagga tcctgctcaa 3240
 aatgtcgaat cttctaattg catgacttta atggaaccaa caaacggcgg tgaggatttg 3300
 ctaatgacaa tttcaatgcc atgtatagag gtaggaaccg ttggtggagg gacaattctg 3360
 gaaccacaag gtgcagtttt ggatttggtg ggcgttagag gggctcacc tactaatcct 3420
 ggtcaaaacg ctcaacagtt agccagaatt atcgcatcag ctgtaatggc aggcgaattg 3480
 tctttgataa gtgccttagc cgcaggtcat ttggttagag ctcatcttgc ccacaatcgt 3540
 tctcaattga atacaccaat gccatccaga ccacatactc ctggccctga ggatgtctca 3600
 catgtgcagc agctacctac accatctgca tctgatgata aagggtgttac agctcaaggc 3660
 tacgttgtcg aagcaaaata a 3681

<210> 108

<211> 1226

<212> PRT

<213> Ganoderma lucidum

<400> 108

Met Arg Ala Val Leu Arg Leu Leu Ser Thr His Thr Val Phe Ser Pro
 1 5 10 15

Ile Glu Thr Ile Val Ser Val Phe Val Leu Ala Thr Leu Ala Tyr Phe
20 25 30

His Ile Leu Ser Gly Ile Lys His Ser Ser Phe Phe Ala Ser Ser His
35 40 45

Pro Pro Ala Ile Arg Pro Ala Phe Ala His Leu Thr Asn Gly Glu Trp
50 55 60

Val Ala Val Ser Gln His Asp Trp Thr Glu Ala Trp Lys His Pro Gly
65 70 75 80

Gly Ser Leu Asp Ala Leu Glu Leu Gln Gln Val Val Phe Thr Leu Asp
85 90 95

Asp Lys Thr Gln Pro Ser Ala Val Leu Asp Ala Ser Ala Ile Ser Gln
100 105 110

His Leu Val Ser Asn Val Pro Ala Leu Ser Gly Lys Ala Tyr Ser Ser
115 120 125

Leu Cys His His Pro Asn Val Ser Gly Thr Ser Cys Phe Thr Ser Val
130 135 140

Ser Gly Pro Gly Ala Ser Pro Ile Leu Thr Leu Ser Phe Lys Pro Gly
145 150 155 160

Thr Arg Asp Asp Trp Leu Gly Ser Leu Arg Lys Glu Lys Thr Ile Thr
165 170 175

Leu Asp Gly Val Lys Tyr Asp Val Gly Ala Gly Lys Arg Gln Glu Ser
180 185 190

Ile Gly Asp Met Glu Ser Ser Lys Trp Val Ala Tyr Ala Leu Ser Ala
195 200 205

Leu Val Leu Arg Phe Trp Glu Leu Thr Lys Ala Asp Ser Leu Asp Ile
210 215 220

Leu Val Val Leu Thr Gly Tyr Ile Leu Met His Val Thr Phe Met Arg
225 230 235 240

Leu Phe Leu Ala Ser Arg Ala Leu Gly Ser Asn Phe Trp Leu Ser Ala
245 250 255

Gly Ile Phe Ser Ser Ala Thr Ile Ser Phe Leu Phe Thr Leu Pro Met
260 265 270

Cys Arg Ser Met Asp Ile Pro Leu Asp Pro Ile Ala Leu Thr Glu Ala
275 280 285

Leu Pro Phe Leu Val Cys Thr Val Gly Phe Asp Lys Pro Leu Arg Leu

290 295

Ala Arg Ala Val Met Ala His Pro Asn Ile Leu Lys Pro Gln Asp Asp
305 310 315 320

Gly Arg Met Lys Ala Ala Gly Asp Val Ile Leu Glu Ala Leu Asp Arg
325 330 335

Val Gly Asn Met Ile Leu Arg Asp Tyr Ala Leu Glu Ile Ala Val Leu
340 345 350

Phe Val Gly Val Asn Ser Arg Val Gly Gly Leu Lys Glu Phe Cys Ala
355 360 365

Val Ala Ala Ala Leu Leu Ala Met Asp Arg Leu Met Thr Phe Thr Leu
370 375 380

Tyr Thr Ala Val Leu Thr Ile Met Val Glu Val Arg Arg Ile Lys Lys
385 390 395 400

Val Arg Asp Met Thr Lys Ala Arg Ser Arg Ser Ser Ser Ile Thr Ala
405 410 415

Val Thr Ala Asn Gly Thr Ala Ile Arg Gly Val Leu Ser Arg Lys Ser
420 425 430

Ser Lys Gln Ser Val Thr Glu Pro Glu Thr Thr Lys Asn Leu Arg Gln
435 440 445

Arg Ala Thr Asp Ser Ala Ile Gly Val Lys Gly Ser Leu Leu Lys Asp
450 455 460

Gly Gly Arg Leu Gln Glu Ala Glu Glu Asn Pro Met Ala Arg Leu Lys
465 470 475 480

Leu Leu Leu Ile Ala Ser Phe Leu Thr Leu His Ile Leu Asn Phe Cys
485 490 495

Thr Thr Leu Thr Ser Ala Thr Ala Asn Ala Arg His Gln Arg His Pro
500 505 510

Phe Arg Thr Val Gln Glu Val Val Pro Ile Pro Arg Val Asp Ile Thr
515 520 525

Thr Pro Ala Ile Ala Asn Ile Leu Ser His Leu Ala Val Ala Gln Glu
530 535 540

Pro Met Phe Thr Val Val Gly Ser Glu Pro Ile Glu Leu Leu Val Lys
545 550 555 560

Val Ala Ala Pro Val Tyr Val His Ala Leu Pro Leu Ala Pro Ala Leu
565 570 575

Arg Ala Ser Asn Thr Asn Thr Gly Glu Ala Ile Glu Asn Phe Met Ser
 580 585 590
 Ser Trp Ser Ser Leu Val Gly Asp Pro Val Val Ser Lys Trp Ile Val
 595 600 605
 Ala Leu Leu Ala Val Ser Val Ala Leu Asn Gly Tyr Leu Leu Lys Gly
 610 615 620
 Ile Ala Ala Gly Ser Gly Leu Ala Ala Met Arg Ala Val Arg Ser Gln
 625 630 635 640
 Gly Val Arg Phe Arg Ser Arg Ala Arg Ser Ile Val Lys Ile Ser Asp
 645 650 655
 Glu Pro Glu Pro Glu Pro Glu His Ser Ile Asp Pro Ala Pro Val Val
 660 665 670
 Phe Phe Ala Ser Ala Ala Pro Ala Val Glu Ala Pro Ala Pro Ala Pro
 675 680 685
 Ala Pro Glu Pro Glu Pro Pro Val Asn Arg Pro Pro Pro Leu Thr Ile
 690 695 700
 Phe Ser Arg Pro Leu Asn Leu Glu Thr Val Asp Lys Lys Leu Gln Asp
 705 710 715 720
 Ala Leu Pro Ile Arg Ser Pro Pro Pro Val Glu Pro Ile Thr Pro Glu
 725 730 735
 Ser Arg Glu Val Glu Pro Thr Gln Val Glu Val Arg Ser Leu Ala Glu
 740 745 750
 Cys Val Asp Val Phe Glu Asn Gly Pro Arg Pro Val Ser Val Ala Leu
 755 760 765
 Lys Thr Leu Asn Asp Glu Glu Val Ile Leu Leu Cys Gln Thr Gly Lys
 770 775 780
 Ile Ala Pro Tyr Ala Leu Val Lys Met Leu Ala Asp Phe Asp Arg Ala
 785 790 795 800
 Val Arg Val Arg Arg Ala Leu Ile Ser Arg Ala Ser Arg Thr Lys Thr
 805 810 815
 Leu Glu Asn Ser Leu Val Pro Met Lys Asp Tyr Asp Tyr Ala Arg Val
 820 825 830
 Met Gly Ala Cys Cys Glu Asn Val Ile Gly Tyr Met Pro Leu Pro Leu
 835 840 845

Gly Ile Ala Gly Pro Leu Lys Ile Asp Gly Leu Met Tyr Pro Ile Pro
 850 855 860

Met Ala Thr Ala Glu Gly Thr Leu Val Ala Ser Thr Ser Arg Gly Cys
 865 870 875 880

Lys Ala Leu Asn Ala Gly Gly Gly Val Thr Thr Val Leu Thr Ala Asp
 885 890 895

Gly Met Thr Arg Gly Pro Ala Ile Asp Phe Pro Ser Ile Val Arg Ala
 900 905 910

Ala Glu Ala Lys Ala Phe Ile Glu Ser Glu Asp Gly Tyr Ala Thr Ile
 915 920 925

Arg Glu Ala Phe Glu Ser Thr Ser Arg Phe Ala Lys Leu Gln Lys Ile
 930 935 940

Lys Cys Ala Leu Ala Gly Arg Thr Leu Phe Val Arg Phe Ala Thr Arg
 945 950 955 960

Thr Gly Asp Ala Met Gly Met Asn Met Ile Ser Lys Ala Thr Glu Lys
 965 970 975

Ala Leu Asp Val Leu Ser His Glu Phe Pro Glu Met Val Val Leu Ala
 980 985 990

Leu Ser Gly Asn Tyr Cys Thr Asp Lys Lys Pro Ala Ala Ile Ser Trp
 995 1000 1005

Ile Glu Gly Arg Gly Lys Ser Ile Val Ala Glu Ala Val Ile Pro
 1010 1015 1020

Gly Lys Val Val Lys Ser Val Leu Lys Thr Thr Val Glu Ser Leu
 1025 1030 1035

Cys Asn Val Asn Thr Lys Lys Asn Leu Ile Gly Ser Ala Met Ala
 1040 1045 1050

Gly Ser Val Gly Gly Phe Asn Ala His Ala Ala Asn Ile Leu Thr
 1055 1060 1065

Ala Val Phe Leu Ala Thr Gly Gln Asp Pro Ala Gln Asn Val Glu
 1070 1075 1080

Ser Ser Asn Cys Met Thr Leu Met Glu Pro Thr Asn Gly Gly Glu
 1085 1090 1095

Asp Leu Leu Met Thr Ile Ser Met Pro Cys Ile Glu Val Gly Thr
 1100 1105 1110

Val Gly Gly Gly Thr Ile Leu Glu Pro Gln Gly Ala Val Leu Asp
1115 1120 1125

Leu Leu Gly Val Arg Gly Ala His Pro Thr Asn Pro Gly Gln Asn
1130 1135 1140

Ala Gln Gln Leu Ala Arg Ile Ile Ala Ser Ala Val Met Ala Gly
1145 1150 1155

Glu Leu Ser Leu Ile Ser Ala Leu Ala Ala Gly His Leu Val Arg
1160 1165 1170

Ala His Leu Ala His Asn Arg Ser Gln Leu Asn Thr Pro Met Pro
1175 1180 1185

Ser Arg Pro His Thr Pro Gly Pro Glu Asp Val Ser His Val Gln
1190 1195 1200

Gln Leu Pro Thr Pro Ser Ala Ser Asp Asp Lys Gly Val Thr Ala
1205 1210 1215

Gln Gly Tyr Val Val Glu Ala Lys
1220 1225

<210> 109

<211> 2667

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 109

atgttatcaa gattgttcag aatgcatggt ctatttgttg cttctcacc	60
ttgggaagta atagttggta ctgtaacatt aacgatctgt atgatgtcta tgaacatg	120
ttaccggaac aacaagattt gtggttggaa ttatgagtgt cctaagctgg aagaggat	180
gttgagttca gacatcatca tacttactat aacaagatgc attgcaatat tgtatatcta	240
cttccaattt caaaacctta gacaattggg tagtaaatac atcctaggca tcgccggatt	300
gttcactatt ttctctagtt ttgttttctc aaccgtcggt attcactttt tggacaaaga	360
gttaactggt ttgaacgaag ctctaccatt cttcttgctg ctggtagatt tgtccagagc	420
ttccgcttta gctaaattcg ctctgtcctc taattctcaa gatgaagtta gagagaatat	480
agcaagggga atggccatac ttggacctac ttccacactt gatgcccttg tcgaatgttt	540
ggttattggg gttggcacia tgtccggcgt tagacagtta gaaatcatgt gttgttttgg	600
ctgtatgagt gtcttggcta actactttgt ctttatgaca ttctttccag cttgcgtttc	660
tttggatttg gagctgtcaa gagaatcaag agaaggcaga ccaatatggc aactatcaca	720
tttcgccaga gtgttagaag aggaggaaaa caaacctaata cctgtcacac agagagtga	780
aatgatcatg tctttggggt tagtcctagt gcatgctcat tctagatgga tcgcagatcc	840
atcccctcag	

aattctacag	ctgataactc	taaagttagt	ttaggttttag	atgaaaatgt	aagtaagagg	900
attgaacctt	ccgtgtcttt	gtggcaattc	tacttatcaa	aatgatttc	catggatatt	960
gaacaagtga	taacgttgtc	tttggttta	ttgttagccg	ttaagtacat	tttctttgag	1020
caagccgaaa	cggaatctac	attatcactg	aaaaacccaa	ttacatcccc	agtcgttacc	1080
cagaaaaaga	taactgatga	ttgctgtaga	agagatccag	tgttggtcag	gaatgatcaa	1140
aagttccacg	ccatggagga	ggaaactagg	aaaaacagag	aaaggaaagt	tgaagttatc	1200
aagcctctat	tagcagaaaa	tgacacttca	catagggcca	ctttcgttgt	cggcaattca	1260
tctcttttag	gtacgtcatt	ggagctggaa	acacaggaac	cagaaatgga	actaccagtt	1320
gaaccaagac	caaagagga	atgtttgcaa	atactagaga	acgctgaaaa	gggagccaag	1380
ttcctatctg	atgccgagat	tatccagctg	gtcaatgcca	agcacattcc	tgcttacaag	1440
ttggaaaccc	ttatggagac	acatgagaga	ggtgtgtcta	ttaggagaca	attactatct	1500
aaaaagttac	ctgaaccaag	ttccctacaa	tacctgcctt	atagagatta	caattactcc	1560
ttggtaatgg	gagcttgttg	tgaaaatgtc	attgggtaca	tgccaattcc	agtgggtgtc	1620
gccggtccac	tatgtttgga	cggtaaggaa	tttcaagtac	ctatggcaac	gactgaaggc	1680
tgcttagttg	catctacaaa	cagaggttgt	agagccattg	gattaggtgg	cggtgcttct	1740
tcaagagtct	tggctgacgg	tatgactaga	ggtcctgttg	tgagatttcc	tagggcctgt	1800
gactctgcag	aagttaaggc	ttggttgga	actccagaag	gtttcaccgt	aatcaaagag	1860
gcctttgatt	ccacatcaag	ggtggccaga	ttacaaaaac	tacacatgtc	tgctgctggg	1920
agaaatctgt	atatcagatt	tcaatccaga	tccggcgacg	caatgggtat	gaatatgatt	1980
tcaaaagga	cagaaaaggc	tttgtcaaag	ctgcaggagt	atttcccaga	gatgcaaatc	2040
ttggccgtat	ctggcaacta	ttgcacagac	aaaaagcctg	ccgccatcaa	ctggattgaa	2100
ggaagaggca	aatctgtggg	ttgtgaagct	gtaattccag	caaagttgt	tagagaagtg	2160
ttaaagacca	caacagaagc	tatgattgaa	gtaaacataa	acaaaaactt	agtaggttct	2220
gccatggctg	gttcaattgg	aggatacaac	gctcatgctg	ccaatattgt	aaccgctatc	2280
tacatcgcat	gtggacaaga	tgctgcccaa	aatgtcgggt	cctcaaattg	catcacattg	2340
atggaagcat	ctggccctac	aaacgaggat	ttgtatatca	gttgcacaa	gccatctata	2400
gaaatagga	ctgtgggagg	aggaactaac	ttacttccac	agcaagcctg	cttacaaatg	2460
ctgggtgtac	aaggagcctg	tagagataat	ccaggggaga	acgctagaca	acttgccaga	2520
attgtttgtg	ggacagttat	ggctggtgaa	cttagtctaa	tggcagcttt	ggctgctggg	2580
cacctggtga	gatctcatat	gattcataat	agaagtaaga	ttaaccttca	agatttgcaa	2640
ggtacgtgta	cgaaaaaggc	tgccctaa				2667

<210> 110
 <211> 888
 <212> PRT
 <213> Bos taurus

<400> 110

Met Leu Ser Arg Leu Phe Arg Met His Gly Leu Phe Val Ala Ser His
 1 5 10 15

Pro Trp Glu Val Ile Val Gly Thr Val Thr Leu Thr Ile Cys Met Met
 20 25 30

Ser Met Asn Met Phe Thr Gly Asn Asn Lys Ile Cys Gly Trp Asn Tyr
 35 40 45

Glu Cys Pro Lys Leu Glu Glu Asp Val Leu Ser Ser Asp Ile Ile Ile
 50 55 60

Leu Thr Ile Thr Arg Cys Ile Ala Ile Leu Tyr Ile Tyr Phe Gln Phe
 65 70 75 80

Gln Asn Leu Arg Gln Leu Gly Ser Lys Tyr Ile Leu Gly Ile Ala Gly
 85 90 95

Leu Phe Thr Ile Phe Ser Ser Phe Val Phe Ser Thr Val Val Ile His
 100 105 110

Phe Leu Asp Lys Glu Leu Thr Gly Leu Asn Glu Ala Leu Pro Phe Phe
 115 120 125

Leu Leu Leu Val Asp Leu Ser Arg Ala Ser Ala Leu Ala Lys Phe Ala
 130 135 140

Leu Ser Ser Asn Ser Gln Asp Glu Val Arg Glu Asn Ile Ala Arg Gly
 145 150 155 160

Met Ala Ile Leu Gly Pro Thr Phe Thr Leu Asp Ala Leu Val Glu Cys
 165 170 175

Leu Val Ile Gly Val Gly Thr Met Ser Gly Val Arg Gln Leu Glu Ile
 180 185 190

Met Cys Cys Phe Gly Cys Met Ser Val Leu Ala Asn Tyr Phe Val Phe
 195 200 205

Met Thr Phe Phe Pro Ala Cys Val Ser Leu Val Leu Glu Leu Ser Arg
 210 215 220

Glu Ser Arg Glu Gly Arg Pro Ile Trp Gln Leu Ser His Phe Ala Arg
 225 230 235 240

Val Leu Glu Glu Glu Glu Asn Lys Pro Asn Pro Val Thr Gln Arg Val
 245 250 255

Lys Met Ile Met Ser Leu Gly Leu Val Leu Val His Ala His Ser Arg
 260 265 270

Trp Ile Ala Asp Pro Ser Pro Gln Asn Ser Thr Ala Asp Asn Ser Lys
 275 280 285
 Val Ser Leu Gly Leu Asp Glu Asn Val Ser Lys Arg Ile Glu Pro Ser
 290 295 300
 Val Ser Leu Trp Gln Phe Tyr Leu Ser Lys Met Ile Ser Met Asp Ile
 305 310 315 320
 Glu Gln Val Ile Thr Leu Ser Leu Ala Leu Leu Ala Val Lys Tyr
 325 330 335
 Ile Phe Phe Glu Gln Ala Glu Thr Glu Ser Thr Leu Ser Leu Lys Asn
 340 345 350
 Pro Ile Thr Ser Pro Val Val Thr Gln Lys Lys Ile Thr Asp Asp Cys
 355 360 365
 Cys Arg Arg Asp Pro Val Leu Val Arg Asn Asp Gln Lys Phe His Ala
 370 375 380
 Met Glu Glu Glu Thr Arg Lys Asn Arg Glu Arg Lys Val Glu Val Ile
 385 390 395 400
 Lys Pro Leu Leu Ala Glu Asn Asp Thr Ser His Arg Ala Thr Phe Val
 405 410 415
 Val Gly Asn Ser Ser Leu Leu Gly Thr Ser Leu Glu Leu Glu Thr Gln
 420 425 430
 Glu Pro Glu Met Glu Leu Pro Val Glu Pro Arg Pro Asn Glu Glu Cys
 435 440 445
 Leu Gln Ile Leu Glu Asn Ala Glu Lys Gly Ala Lys Phe Leu Ser Asp
 450 455 460
 Ala Glu Ile Ile Gln Leu Val Asn Ala Lys His Ile Pro Ala Tyr Lys
 465 470 475 480
 Leu Glu Thr Leu Met Glu Thr His Glu Arg Gly Val Ser Ile Arg Arg
 485 490 495
 Gln Leu Leu Ser Lys Lys Leu Pro Glu Pro Ser Ser Leu Gln Tyr Leu
 500 505 510
 Pro Tyr Arg Asp Tyr Asn Tyr Ser Leu Val Met Gly Ala Cys Cys Glu
 515 520 525
 Asn Val Ile Gly Tyr Met Pro Ile Pro Val Gly Val Ala Gly Pro Leu
 530 535 540
 Cys Leu Asp Gly Lys Glu Phe Gln Val Pro Met Ala Thr Thr Glu Gly

545 550 560

Cys Leu Val Ala Ser Thr Asn Arg Gly Cys Arg Ala Ile Gly Leu Gly
 565 570 575

Gly Gly Ala Ser Ser Arg Val Leu Ala Asp Gly Met Thr Arg Gly Pro
 580 585 590

Val Val Arg Phe Pro Arg Ala Cys Asp Ser Ala Glu Val Lys Ala Trp
 595 600 605

Leu Glu Thr Pro Glu Gly Phe Thr Val Ile Lys Glu Ala Phe Asp Ser
 610 615 620

Thr Ser Arg Val Ala Arg Leu Gln Lys Leu His Met Ser Val Ala Gly
625 630 635 640

Arg Asn Leu Tyr Ile Arg Phe Gln Ser Arg Ser Gly Asp Ala Met Gly
 645 650 655

Met Asn Met Ile Ser Lys Gly Thr Glu Lys Ala Leu Ser Lys Leu Gln
 660 665 670

Glu Tyr Phe Pro Glu Met Gln Ile Leu Ala Val Ser Gly Asn Tyr Cys
 675 680 685

Thr Asp Lys Lys Pro Ala Ala Ile Asn Trp Ile Glu Gly Arg Gly Lys
690 695 700

Ser Val Val Cys Glu Ala Val Ile Pro Ala Lys Val Val Arg Glu Val
705 710 715 720

Leu Lys Thr Thr Thr Glu Ala Met Ile Glu Val Asn Ile Asn Lys Asn
 725 730 735

Leu Val Gly Ser Ala Met Ala Gly Ser Ile Gly Gly Tyr Asn Ala His
 740 745 750

Ala Ala Asn Ile Val Thr Ala Ile Tyr Ile Ala Cys Gly Gln Asp Ala
 755 760 765

Ala Gln Asn Val Gly Ser Ser Asn Cys Ile Thr Leu Met Glu Ala Ser
770 775 780

Gly Pro Thr Asn Glu Asp Leu Tyr Ile Ser Cys Thr Met Pro Ser Ile
785 790 795 800

Glu Ile Gly Thr Val Gly Gly Gly Thr Asn Leu Leu Pro Gln Gln Ala
 805 810 815

Cys Leu Gln Met Leu Gly Val Gln Gly Ala Cys Arg Asp Asn Pro Gly
 820 825 830

Glu Asn Ala Arg Gln Leu Ala Arg Ile Val Cys Gly Thr Val Met Ala
 835 840 845

Gly Glu Leu Ser Leu Met Ala Ala Leu Ala Ala Gly His Leu Val Arg
 850 855 860

Ser His Met Ile His Asn Arg Ser Lys Ile Asn Leu Gln Asp Leu Gln
 865 870 875 880

Gly Thr Cys Thr Lys Lys Ala Ala
 885

<210> 111

<211> 1704

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 111

atggatttga gaaggaaatt accacctaag cctccatctt caacaacaac aaaacagcca	60
agtcataaggc cccattctcc tacgccaatt ccaaaggctt cagatgcatt gcctcttcca	120
ttgtacctga ccaatacggtt tttcttcact cttttctttt ccgtagcata ttacctgttg	180
cataggtgga gagacaagat tagatccgga acacctttac acgttgtgac actgactgaa	240
ctatccgcaa ttgtactgct gattgcttcc ttcattctatc ttttaggctt tttcgggtatt	300
gattttgtgc aatctttcac atcaagagaa aatgagcaac taaacaacga tgatcacaac	360
gtcgtgtcaa caaacaatgt tttatctgat agaagggttag tttacgacta tggattcgat	420
gtgacaggag acaacgataa cgataatgat gacgatgtta ttgtgaaaag tgtcgtttct	480
ggggaagtta attcttatag tttggaggct tccctaggag attgttacag agccgcaaag	540
attagaaaga gagccgtcga gagaattgtc gggagagaag tattaggctt gggtttcgag	600
ggatttgatt atgaatctat cctggggcaa tgttgtgaaa tgcctatcgg gtacgtccaa	660
gtgccagtag gtgtcgctgg acctttattg ttaaattggtg gggaattcat ggttccaatg	720
gctacaactg aaggctgtct tgtagcttcc actaatagag gttgtaaagc catatgctta	780
tcagggtggtg ccaactgcat attgctaata gatggtatga caagagcccc agtagtgaga	840
ttcgccacag ctgagagagc ttcacaacta aagttttact tgggaagatgg tgtcaatttc	900
gatacattgt ctgttgtctt taacaaaagt tcaagatttg ccagattgca aaacatccaa	960
tgctcaattg ccggtaaaaa cttgtacatt aggtttactt gctccacagg cgacgccatg	1020
ggtatgaaca tggtttcaaa aggagtacaa aatgtattag actttttaca aaatgatttt	1080
cctgatatgg acgtaattgg gatctcttgg aagttctgct ctgacaaaaa gccaacagct	1140
gtcaactgga ttgagggcag aggaaagtct gtcgttttcc aggccgtaat taccaaaaag	1200
gtgggttagaa agtctgcact gaaccctcaa acttgcacat gtagaacttt gacctgttta	1260

agaccattat tggttctgct acttctgggtt ttgctagtgg acttaatgca tatgcttcat 1320
 atcgtgtctg ccgtgttcat cgctaccggt caagatccag ctcagaatat cgaatctagt 1380
 cactgtatca ctatgatgga ggctgtcaac aatggtaagg atttgacgt taatgttacg 1440
 atgccatcta tagaagttgg cacggtggga ggtggcactc agctagcctc tcaatcagcc 1500
 tgtttgaact tgcttggtgt aaagggtgcc tgtatagaat ccccaggatc aaacgcccag 1560
 ttgttagcta gaatcgttgc tggttctggt ctggcaggcg aattaagttt gatgtcagct 1620
 ataagtgctg ggcaactagt taaatctcat atgaaataca ataggtctag tagagatatg 1680
 tcagcaatag cttctaaggt ctaa 1704

<210> 112

<211> 567

<212> PRT

<213> Artemisia annua

<400> 112

Met Asp Leu Arg Arg Lys Leu Pro Pro Lys Pro Pro Ser Ser Thr Thr
1 5 10 15

Thr Lys Gln Pro Ser His Arg Ser His Ser Pro Thr Pro Ile Pro Lys
20 25 30

Ala Ser Asp Ala Leu Pro Leu Pro Leu Tyr Leu Thr Asn Thr Phe Phe
35 40 45

Phe Thr Leu Phe Phe Ser Val Ala Tyr Tyr Leu Leu His Arg Trp Arg
50 55 60

Asp Lys Ile Arg Ser Gly Thr Pro Leu His Val Val Thr Leu Thr Glu
65 70 75 80

Leu Ser Ala Ile Val Leu Leu Ile Ala Ser Phe Ile Tyr Leu Leu Gly
85 90 95

Phe Phe Gly Ile Asp Phe Val Gln Ser Phe Thr Ser Arg Glu Asn Glu
100 105 110

Gln Leu Asn Asn Asp Asp His Asn Val Val Ser Thr Asn Asn Val Leu
115 120 125

Ser Asp Arg Arg Leu Val Tyr Asp Tyr Gly Phe Asp Val Thr Gly Asp
130 135 140

Asn Asp Asn Asp Asn Asp Asp Asp Val Ile Val Lys Ser Val Val Ser
145 150 155 160

Gly Glu Val Asn Ser Tyr Ser Leu Glu Ala Ser Leu Gly Asp Cys Tyr
165 170 175

Arg Ala Ala Lys Ile Arg Lys Arg Ala Val Glu Arg Ile Val Gly Arg

180		185		190
Glu Val Leu Gly Leu Gly Phe Glu Gly Phe Asp Tyr Glu Ser Ile Leu	195	200	205	
Gly Gln Cys Cys Glu Met Pro Ile Gly Tyr Val Gln Val Pro Val Gly	210	215	220	
Val Ala Gly Pro Leu Leu Leu Asn Gly Gly Glu Phe Met Val Pro Met	225	230	235	240
Ala Thr Thr Glu Gly Cys Leu Val Ala Ser Thr Asn Arg Gly Cys Lys	245	250	255	
Ala Ile Cys Leu Ser Gly Gly Ala Thr Ala Ile Leu Leu Lys Asp Gly	260	265	270	
Met Thr Arg Ala Pro Val Val Arg Phe Ala Thr Ala Glu Arg Ala Ser	275	280	285	
Gln Leu Lys Phe Tyr Leu Glu Asp Gly Val Asn Phe Asp Thr Leu Ser	290	295	300	
Val Val Phe Asn Lys Ser Ser Arg Phe Ala Arg Leu Gln Asn Ile Gln	305	310	315	320
Cys Ser Ile Ala Gly Lys Asn Leu Tyr Ile Arg Phe Thr Cys Ser Thr	325	330	335	
Gly Asp Ala Met Gly Met Asn Met Val Ser Lys Gly Val Gln Asn Val	340	345	350	
Leu Asp Phe Leu Gln Asn Asp Phe Pro Asp Met Asp Val Ile Gly Ile	355	360	365	
Ser Trp Lys Phe Cys Ser Asp Lys Lys Pro Thr Ala Val Asn Trp Ile	370	375	380	
Glu Gly Arg Gly Lys Ser Val Val Phe Gln Ala Val Ile Thr Lys Lys	385	390	395	400
Val Val Arg Lys Ser Ala Leu Asn Pro Gln Thr Cys Thr Cys Arg Thr	405	410	415	
Leu Thr Cys Leu Arg Pro Leu Leu Val Leu Leu Leu Val Leu Leu	420	425	430	
Val Asp Leu Met His Met Leu His Ile Val Ser Ala Val Phe Ile Ala	435	440	445	
Thr Gly Gln Asp Pro Ala Gln Asn Ile Glu Ser Ser His Cys Ile Thr	450	455	460	

Met Met Glu Ala Val Asn Asn Gly Lys Asp Leu His Val Asn Val Thr
465 470 475 480

Met Pro Ser Ile Glu Val Gly Thr Val Gly Gly Gly Thr Gln Leu Ala
485 490 495

Ser Gln Ser Ala Cys Leu Asn Leu Leu Gly Val Lys Gly Ala Cys Ile
500 505 510

Glu Ser Pro Gly Ser Asn Ala Gln Leu Leu Ala Arg Ile Val Ala Gly
515 520 525

Ser Val Leu Ala Gly Glu Leu Ser Leu Met Ser Ala Ile Ser Ala Gly
530 535 540

Gln Leu Val Lys Ser His Met Lys Tyr Asn Arg Ser Ser Arg Asp Met
545 550 555 560

Ser Ala Ile Ala Ser Lys Val
565

<210> 113

<211> 1308

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 113

atgttttagaa gagctatact gttaggatgc tctgctgccca agacaccatg gtctgagtgt	60
tctaacgctc aattagttga tgcagttaag tctagaaaga tctcattcta cggctcttgaa	120
caagccttgg aaccagatta tagaagggtc atcgaagtaa ggagagaggt tgtctctgaa	180
atcgccctcac aacagccaga agcaaaaaag aagcaatccg cattgcacac aataccattt	240
gagaattatg attggaataa ggtcgttggc caaaactgtg aaaacattat tggatacgtc	300
ccaataccac tgggcgttgc tggccctatt ttgattgatg gtaaagagta cccaatacca	360
atggctacaa cagaaggcgc tttggtcgct agtactcata gaggtgctag agctattaca	420
agatccggag gttgtaagac attgttatta ggtgaaggta tgacaagagc accagtgggt	480
gaattgcctt cattagagga agctgggcgt ttgcacaagt actgtaatga gaacttctta	540
tctttaaagg aagcatttga atcaactacc caatatggaa aacttaattc tttaaagtgc	600
gtactagctg gtagaaaagc atacccttaga ttcagagcca ctacaggcga tgctatgggc	660
atgaacatga taacaaaggg ttagacaaa gcactgtctg ttctacagca acatttcctt	720
tcaatggaaa tcctagccct aagtggtaat tactgtaccg acaaaaagcc atctgctgta	780
aattggattg atggcagagg taaatcagtg gttgcagaag ccactttatt ggctgatgtt	840
gtcgaagata ctctgaaatg tacagtcgat tctttggtat ccttgaatat cgacaaaaac	900

```

cttgttgggt cagctatggc tggttctggt ggaggtttta acgcccaggc tgcaaacgct      960
gtggcagcca ttttcattgc aaccggtcaa gatcctgctc aagtggtaga aagttcaatg      1020
tgtatcacta caatgtccaa ggtaggtaac gatctattga tctctgtgac catgccttct      1080
atcgaggctc gggtcgtggg aggagggact ggtcttgctg cccaaagagg atgcttagag      1140
ttaatagggt gcggaggccc atctaaggag tctcctggta ctaatgcca acttctaagt      1200
agagttgttg cagctggcgt tttatcagcc gaactttcct tgatgtccgg actggcagca      1260
ggtcacttat tgtcagcaca tatgagattg aacagaaaga agaaataa      1308

```

```

<210> 114
<211> 435
<212> PRT
<213> Trypanosoma cruzi

```

```

<400> 114
Met Phe Arg Arg Ala Ile Leu Leu Gly Cys Ser Ala Ala Lys Thr Pro
1           5           10           15

```

```

Trp Ser Glu Cys Ser Asn Ala Gln Leu Val Asp Ala Val Lys Ser Arg
           20           25           30

```

```

Lys Ile Ser Phe Tyr Gly Leu Glu Gln Ala Leu Glu Pro Asp Tyr Arg
           35           40           45

```

```

Arg Ala Ile Glu Val Arg Arg Glu Val Val Ser Glu Ile Ala Ser Gln
           50           55           60

```

```

Gln Pro Glu Ala Lys Lys Lys Gln Ser Ala Leu His Thr Ile Pro Phe
65           70           75           80

```

```

Glu Asn Tyr Asp Trp Asn Lys Val Val Gly Gln Asn Cys Glu Asn Ile
           85           90           95

```

```

Ile Gly Tyr Val Pro Ile Pro Leu Gly Val Ala Gly Pro Ile Leu Ile
           100           105           110

```

```

Asp Gly Lys Glu Tyr Pro Ile Pro Met Ala Thr Thr Glu Gly Ala Leu
           115           120           125

```

```

Val Ala Ser Thr His Arg Gly Ala Arg Ala Ile Thr Arg Ser Gly Gly
           130           135           140

```

```

Cys Lys Thr Leu Leu Leu Gly Glu Gly Met Thr Arg Ala Pro Val Val
145           150           155           160

```

```

Glu Leu Pro Ser Leu Glu Glu Ala Gly Arg Leu His Lys Tyr Cys Asn
           165           170           175

```

```

Glu Asn Phe Leu Ser Leu Lys Glu Ala Phe Glu Ser Thr Thr Gln Tyr
           180           185           190

```

Gly Lys Leu Asn Ser Leu Lys Cys Val Leu Ala Gly Arg Lys Ala Tyr
 195 200 205
 Leu Arg Phe Arg Ala Thr Thr Gly Asp Ala Met Gly Met Asn Met Ile
 210 215 220
 Thr Lys Gly Val Asp Lys Ala Leu Ser Val Leu Gln Gln His Phe Pro
 225 230 235 240
 Ser Met Glu Ile Leu Ala Leu Ser Gly Asn Tyr Cys Thr Asp Lys Lys
 245 250 255
 Pro Ser Ala Val Asn Trp Ile Asp Gly Arg Gly Lys Ser Val Val Ala
 260 265 270
 Glu Ala Thr Leu Leu Ala Asp Val Val Glu Asp Thr Leu Lys Cys Thr
 275 280 285
 Val Asp Ser Leu Val Ser Leu Asn Ile Asp Lys Asn Leu Val Gly Ser
 290 295 300
 Ala Met Ala Gly Ser Val Gly Gly Phe Asn Ala Gln Ala Ala Asn Ala
 305 310 315 320
 Val Ala Ala Ile Phe Ile Ala Thr Gly Gln Asp Pro Ala Gln Val Val
 325 330 335
 Glu Ser Ser Met Cys Ile Thr Thr Met Ser Lys Val Gly Asn Asp Leu
 340 345 350
 Leu Ile Ser Val Thr Met Pro Ser Ile Glu Val Gly Val Val Gly Gly
 355 360 365
 Gly Thr Gly Leu Ala Ala Gln Arg Gly Cys Leu Glu Leu Ile Gly Cys
 370 375 380
 Gly Gly Pro Ser Lys Glu Ser Pro Gly Thr Asn Ala Gln Leu Leu Ser
 385 390 395 400
 Arg Val Val Ala Ala Gly Val Leu Ser Ala Glu Leu Ser Leu Met Ser
 405 410 415
 Gly Leu Ala Ala Gly His Leu Leu Ser Ala His Met Arg Leu Asn Arg
 420 425 430
 Lys Lys Lys
 435

<210> 115

<211> 1281

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 115

```

atgcaatccc tggacaaaaa ctttagacac ttatcaagac aacagaagtt acaacagcta      60
gttgataaac aatggctatc agaggaacaa ttcaatattc tacttaacca cccacttatt      120
gatgaagagg tagcaaactc attgatagaa aatgtcatcg cacagggcgc actgcctgtt      180
ggtttactac caaatatcat cgttgatgac aaagcatacg tcgtgcctat gatggtggaa      240
gagccatctg ttgttgccgc tgcttcatac ggcgctaaat tggatgaacca aacaggtggt      300
ttcaaaaccg tgtcctcaga acgtatcatg atagggtcaaa tagtatttga tggagtcgat      360
gataccgaga aactgtctgc agatatcaag gctcttgaaa aacaaatcca tcagattgca      420
gatgaggcctt acccttctat taaggccaga ggtggaggct atcaaaggat cgccatcgat      480
acattcccag aacaacagtt gctttcattg aagggttttcg ttgatactaa ggatgctatg      540
ggcgctaata tgttaaaccac aatcctagaa gcaatcacag cttttttgaa aaacgaattc      600
ccacaatctg atatcttgat gtctatcctt tccaaccacg caacagccag tgttgtcaag      660
gtccagggtg aaatagacgt taaggatttg gcaagaggag aacgtactgg agaagaggtc      720
gctaagagaa tggaaagagc atctgtgtta gctcaagtgg acattcatag agcagcaaca      780
cacaataagg gtgttatgaa tggcattcat gctgtagtct tggctacagg taatgatact      840
agagggtcag aagcctctgc tcacgcttac gcttccaaag acggtcaata tagagggata      900
gctacatgga gatacgatca agagagacaa aggttaatag gaactataga agttccaatg      960
actctggcca ttgttggttg cggtagcaag gtactgccta ttgctaaggc ctctttagaa     1020
ctgttaaacg tagaaagtgc ccaagagttg ggacatgttg tcgctgccgt tggactagct     1080
caaaacttcg ctgcatgtag agctttggtt tccgaaggta ttcaacaagg gcatatgtct     1140
ttgcaataca agtcttttagc catcgtagtc ggggctaagg gcgatgaaat tgctcaggta     1200
gccgaagcac taaagcaaga gccaagagca aacactcaag ttgcagagag aattttgcaa     1260
gatttgagaa gtcaacaata a                                     1281

```

<210> 116

<211> 426

<212> PRT

<213> Staphylococcus aureus

<400> 116

```

Met Gln Ser Leu Asp Lys Asn Phe Arg His Leu Ser Arg Gln Gln Lys
 1           5           10          15

```

```

Leu Gln Gln Leu Val Asp Lys Gln Trp Leu Ser Glu Glu Gln Phe Asn
          20          25          30

```

```

Ile Leu Leu Asn His Pro Leu Ile Asp Glu Glu Val Ala Asn Ser Leu
 35          40          45

```

Ile Glu Asn Val Ile Ala Gln Gly Ala Leu Pro Val Gly Leu Leu Pro
 50 55 60
 Asn Ile Ile Val Asp Asp Lys Ala Tyr Val Val Pro Met Met Val Glu
 65 70 75 80
 Glu Pro Ser Val Val Ala Ala Ala Ser Tyr Gly Ala Lys Leu Val Asn
 85 90 95
 Gln Thr Gly Gly Phe Lys Thr Val Ser Ser Glu Arg Ile Met Ile Gly
 100 105 110
 Gln Ile Val Phe Asp Gly Val Asp Asp Thr Glu Lys Leu Ser Ala Asp
 115 120 125
 Ile Lys Ala Leu Glu Lys Gln Ile His Gln Ile Ala Asp Glu Ala Tyr
 130 135 140
 Pro Ser Ile Lys Ala Arg Gly Gly Gly Tyr Gln Arg Ile Ala Ile Asp
 145 150 155 160
 Thr Phe Pro Glu Gln Gln Leu Leu Ser Leu Lys Val Phe Val Asp Thr
 165 170 175
 Lys Asp Ala Met Gly Ala Asn Met Leu Asn Thr Ile Leu Glu Ala Ile
 180 185 190
 Thr Ala Phe Leu Lys Asn Glu Phe Pro Gln Ser Asp Ile Leu Met Ser
 195 200 205
 Ile Leu Ser Asn His Ala Thr Ala Ser Val Val Lys Val Gln Gly Glu
 210 215 220
 Ile Asp Val Lys Asp Leu Ala Arg Gly Glu Arg Thr Gly Glu Glu Val
 225 230 235 240
 Ala Lys Arg Met Glu Arg Ala Ser Val Leu Ala Gln Val Asp Ile His
 245 250 255
 Arg Ala Ala Thr His Asn Lys Gly Val Met Asn Gly Ile His Ala Val
 260 265 270
 Val Leu Ala Thr Gly Asn Asp Thr Arg Gly Ala Glu Ala Ser Ala His
 275 280 285
 Ala Tyr Ala Ser Lys Asp Gly Gln Tyr Arg Gly Ile Ala Thr Trp Arg
 290 295 300
 Tyr Asp Gln Glu Arg Gln Arg Leu Ile Gly Thr Ile Glu Val Pro Met
 305 310 315 320
 Thr Leu Ala Ile Val Gly Gly Gly Thr Lys Val Leu Pro Ile Ala Lys

325

335

Ala Ser Leu Glu Leu Leu Asn Val Glu Ser Ala Gln Glu Leu Gly His
340 345 350

Val Val Ala Ala Val Gly Leu Ala Gln Asn Phe Ala Ala Cys Arg Ala
355 360 365

Leu Val Ser Glu Gly Ile Gln Gln Gly His Met Ser Leu Gln Tyr Lys
370 375 380

Ser Leu Ala Ile Val Val Gly Ala Lys Gly Asp Glu Ile Ala Gln Val
385 390 395 400

Ala Glu Ala Leu Lys Gln Glu Pro Arg Ala Asn Thr Gln Val Ala Glu
405 410 415

Arg Ile Leu Gln Asp Leu Arg Ser Gln Gln
420 425

<210> 117

<211> 1311

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 117

atgcaggtct taagattgga taggagacat tacaaaagtg gcaagattag aagagcaatg	60
agttctagaa ttcctggttt ctacaaattg tcagtcgagg aaagactgaa aaagggttgc	120
gaatttgca ggttatctga tgaggaagtg aaagctgttt tgtcacaagg ttacctttg	180
gacgtagctg atagaatgat cgaaaatgtg atcggtacat ttgaattacc acttggtata	240
gcaaccaatt tccttattga tggcaaggat tatctaattcc ctatggctat agaggaacca	300
tcagtagttg cagctgcttc taacgcagct agaattggcca gagagtctgg cgggtttaca	360
actgattaca cagggtccct gatgattggt caaattcaag tcacaaaact gttgaatcca	420
aatgcagcta agttcgaagt tctacgtcaa aaagacgaaa tcatagaaag agcaaagag	480
tgtgatccaa tgttggtgaa tttgggcggt ggatgtaaag atatagaagc aagggtgatc	540
gatacaatca tgggtaagat gctaattgtt catctgatcg ttgatgttaa agacgctatg	600
ggtgcaaattg ctgtcaacac tatgtgtgaa aaagttgctc ctttcatcga acgtattact	660
gggggaaagg tctatcttag aatcatttcc aacttggtcg catatagact tgctagagca	720
aaggccgttt ttgacaaaga cgttattggc ggagaggagg ttgtagaagg gatcatgctt	780
gcatacgctt tcgctgccgc tgaccattt cgttgcgcca cccacaataa gggatatcatg	840
aatggcatat cagccttaat gatcgctaca ggaaacgact ttagagccat tgaagcagga	900
gctcattcct atgctgcaat aggtggatac aaaccactaa ctacctacga agttgataga	960

aaaggtaatc tagtaggcac aattgaaata cctatggcag taggcgtgat tgggtggtgca 1020
 accaaagtca acccactagc caagatctct ctttaagatac taggagtgaa cactgctgaa 1080
 gagttagcca gagtcgcagc cgtcttaggt ttggctcaaa actttgctgc ctttaagagcc 1140
 ttggccacag aagggtatcca aagaggtcac atggaattac atgccaggaa cttagcaatc 1200
 atggctggag ctactggaga tgaggttgac agagttgtag agattatggt gagagatggc 1260
 aaaatcagat tggactacgc taaggaagta ttggagagac tgcgttccta a 1311

<210> 118

<211> 436

<212> PRT

<213> Archaeoglobus fulgidus

<400> 118

Met Gln Val Leu Arg Leu Asp Arg Arg His Tyr Lys Ser Gly Lys Ile
1 5 10 15

Arg Arg Ala Met Ser Ser Arg Ile Pro Gly Phe Tyr Lys Leu Ser Val
20 25 30

Glu Glu Arg Leu Lys Lys Val Ala Glu Phe Ala Gly Leu Ser Asp Glu
35 40 45

Glu Val Lys Ala Val Leu Ser Gln Gly Leu Pro Leu Asp Val Ala Asp
50 55 60

Arg Met Ile Glu Asn Val Ile Gly Thr Phe Glu Leu Pro Leu Gly Ile
65 70 75 80

Ala Thr Asn Phe Leu Ile Asp Gly Lys Asp Tyr Leu Ile Pro Met Ala
85 90 95

Ile Glu Glu Pro Ser Val Val Ala Ala Ala Ser Asn Ala Ala Arg Met
100 105 110

Ala Arg Glu Ser Gly Gly Phe Thr Thr Asp Tyr Thr Gly Ser Leu Met
115 120 125

Ile Gly Gln Ile Gln Val Thr Lys Leu Leu Asn Pro Asn Ala Ala Lys
130 135 140

Phe Glu Val Leu Arg Gln Lys Asp Glu Ile Ile Glu Arg Ala Asn Glu
145 150 155 160

Cys Asp Pro Met Leu Val Asn Leu Gly Gly Gly Cys Lys Asp Ile Glu
165 170 175

Ala Arg Val Ile Asp Thr Ile Met Gly Lys Met Leu Ile Val His Leu
180 185 190

Ile Val Asp Val Lys Asp Ala Met Gly Ala Asn Ala Val Asn Thr Met
195 200 205

Cys Glu Lys Val Ala Pro Phe Ile Glu Arg Ile Thr Gly Gly Lys Val
 210 215 220

Tyr Leu Arg Ile Ile Ser Asn Leu Ala Ala Tyr Arg Leu Ala Arg Ala
 225 230 235 240

Lys Ala Val Phe Asp Lys Asp Val Ile Gly Gly Glu Glu Val Val Glu
 245 250 255

Gly Ile Met Leu Ala Tyr Ala Phe Ala Ala Ala Asp Pro Phe Arg Cys
 260 265 270

Ala Thr His Asn Lys Gly Ile Met Asn Gly Ile Ser Ala Leu Met Ile
 275 280 285

Ala Thr Gly Asn Asp Phe Arg Ala Ile Glu Ala Gly Ala His Ser Tyr
 290 295 300

Ala Ala Ile Gly Gly Tyr Lys Pro Leu Thr Thr Tyr Glu Val Asp Arg
 305 310 315 320

Lys Gly Asn Leu Val Gly Thr Ile Glu Ile Pro Met Ala Val Gly Val
 325 330 335

Ile Gly Gly Ala Thr Lys Val Asn Pro Leu Ala Lys Ile Ser Leu Lys
 340 345 350

Ile Leu Gly Val Asn Thr Ala Glu Glu Leu Ala Arg Val Ala Ala Ala
 355 360 365

Leu Gly Leu Ala Gln Asn Phe Ala Ala Leu Arg Ala Leu Ala Thr Glu
 370 375 380

Gly Ile Gln Arg Gly His Met Glu Leu His Ala Arg Asn Leu Ala Ile
 385 390 395 400

Met Ala Gly Ala Thr Gly Asp Glu Val Asp Arg Val Val Glu Ile Met
 405 410 415

Val Arg Asp Gly Lys Ile Arg Leu Asp Tyr Ala Lys Glu Val Leu Glu
 420 425 430

Arg Leu Arg Ser
 435

<210> 119

<211> 1287

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 119
 atgtccttag attcaagact gccagctttc agaaatctgt ctccagctgc aagactagat 60
 cacattggcc aacttttggg actaagtcac gacgacgttt cccttttagc aaacgccggt 120
 gctttaccaa tggatatcgc taatggtatg attgaaaatg taatcgggac ctttgaactg 180
 ccatatgcag tggccagtaa ctttcagatc aatggccgtg acgtcttagt accattagtt 240
 gtggaggaac ctagtatcgt tgctgcagcc tcttacatgg caaagttagc tagagccaat 300
 ggtgggttca ctacatcttc atctgctcca ctaatgcatg cacaagtaca aattgtcggc 360
 attcaggatc cactaaacgc aagattgtct ttactgcgta gaaaggatga gatcatagaa 420
 ttagccaata ggaaggacca acttctgaat tcattgggcg gtggttgagc agacatagag 480
 gtgcatacat ttgccgatac tccaagagga ccaatgcttg tagcacacct tattgtcgat 540
 gtgcgtgatg ccatgggagc taatactgtt aacactatgg ctgaagcagt agcacctctg 600
 atggaagcca taacagggtg ccaggtaaga ttgagaatcc tttccaattt ggctgatctt 660
 agattggcca gagccaagt gagaatcact cctcagcaat tggaaactgc cgaattctca 720
 ggtgaggcag taattgaggg tatcttgagc gcatatgctt ttgccgctgt ggacccttac 780
 agagccgcta cccacaacaa aggcataatg aacgggtatcg atcctttgat cgctcgctaca 840
 ggaaatgatt ggagagctgt tgaggcagga gctcatgcat acgctttagt atccggacat 900
 tacggttcat taacaacatg ggaaaaagat aacaatggac acttggtcgg gacattggaa 960
 atgcctatgc cagttggttt agttgggggt gctacaaaaa cccatcctct tgctcaattg 1020
 tctttgagga tacttggtgt caaaactgct caagcactag ccgaaattgc cgttgctgtt 1080
 ggtttggcac aaaacttggg tgcaatgcgt gcttttagcta cagaaggcat ccaaagagga 1140
 catatggctc tacacgctag aaacattgca gttgttgagc gagccagagg tgatgagggtt 1200
 gattgggtgg ctagacaact tgtcgaatat catgatgtca gagcagacag ggctgtggca 1260
 ttactgaaac agaagagagg tcaataa 1287

<210> 120

<211> 428

<212> PRT

<213> *Pseudomonas mevalonii*

<400> 120

Met Ser Leu Asp Ser Arg Leu Pro Ala Phe Arg Asn Leu Ser Pro Ala
 1 5 10 15

Ala Arg Leu Asp His Ile Gly Gln Leu Leu Gly Leu Ser His Asp Asp
 20 25 30

Val Ser Leu Leu Ala Asn Ala Gly Ala Leu Pro Met Asp Ile Ala Asn
 35 40 45

Gly Met Ile Glu Asn Val Ile Gly Thr Phe Glu Leu Pro Tyr Ala Val
 50 55 60

Ala Ser Asn Phe Gln Ile Asn Gly Arg Asp Val Leu Val Pro Leu Val
 65 70 75 80
 Val Glu Glu Pro Ser Ile Val Ala Ala Ala Ser Tyr Met Ala Lys Leu
 85 90 95
 Ala Arg Ala Asn Gly Gly Phe Thr Thr Ser Ser Ser Ala Pro Leu Met
 100 105 110
 His Ala Gln Val Gln Ile Val Gly Ile Gln Asp Pro Leu Asn Ala Arg
 115 120 125
 Leu Ser Leu Leu Arg Arg Lys Asp Glu Ile Ile Glu Leu Ala Asn Arg
 130 135 140
 Lys Asp Gln Leu Leu Asn Ser Leu Gly Gly Gly Cys Arg Asp Ile Glu
 145 150 155 160
 Val His Thr Phe Ala Asp Thr Pro Arg Gly Pro Met Leu Val Ala His
 165 170 175
 Leu Ile Val Asp Val Arg Asp Ala Met Gly Ala Asn Thr Val Asn Thr
 180 185 190
 Met Ala Glu Ala Val Ala Pro Leu Met Glu Ala Ile Thr Gly Gly Gln
 195 200 205
 Val Arg Leu Arg Ile Leu Ser Asn Leu Ala Asp Leu Arg Leu Ala Arg
 210 215 220
 Ala Gln Val Arg Ile Thr Pro Gln Gln Leu Glu Thr Ala Glu Phe Ser
 225 230 235 240
 Gly Glu Ala Val Ile Glu Gly Ile Leu Asp Ala Tyr Ala Phe Ala Ala
 245 250 255
 Val Asp Pro Tyr Arg Ala Ala Thr His Asn Lys Gly Ile Met Asn Gly
 260 265 270
 Ile Asp Pro Leu Ile Val Ala Thr Gly Asn Asp Trp Arg Ala Val Glu
 275 280 285
 Ala Gly Ala His Ala Tyr Ala Cys Arg Ser Gly His Tyr Gly Ser Leu
 290 295 300
 Thr Thr Trp Glu Lys Asp Asn Asn Gly His Leu Val Gly Thr Leu Glu
 305 310 315 320
 Met Pro Met Pro Val Gly Leu Val Gly Gly Ala Thr Lys Thr His Pro
 325 330 335

Leu Ala Gln Leu Ser Leu Arg Ile Leu Gly Val Lys Thr Ala Gln Ala
 340 345 350

Leu Ala Glu Ile Ala Val Ala Val Gly Leu Ala Gln Asn Leu Gly Ala
 355 360 365

Met Arg Ala Leu Ala Thr Glu Gly Ile Gln Arg Gly His Met Ala Leu
 370 375 380

His Ala Arg Asn Ile Ala Val Val Ala Gly Ala Arg Gly Asp Glu Val
 385 390 395 400

Asp Trp Val Ala Arg Gln Leu Val Glu Tyr His Asp Val Arg Ala Asp
 405 410 415

Arg Ala Val Ala Leu Leu Lys Gln Lys Arg Gly Gln
 420 425

<210> 121
 <211> 361
 <212> PRT
 <213> Stevia rebaudiana

<400> 121
 Met Ala Leu Val Asn Pro Thr Ala Leu Phe Tyr Gly Thr Ser Ile Arg
 1 5 10 15

Thr Arg Pro Thr Asn Leu Leu Asn Pro Thr Gln Lys Leu Arg Pro Val
 20 25 30

Ser Ser Ser Ser Leu Pro Ser Phe Ser Ser Val Ser Ala Ile Leu Thr
 35 40 45

Glu Lys His Gln Ser Asn Pro Ser Glu Asn Asn Asn Leu Gln Thr His
 50 55 60

Leu Glu Thr Pro Phe Asn Phe Asp Ser Tyr Met Leu Glu Lys Val Asn
 65 70 75 80

Met Val Asn Glu Ala Leu Asp Ala Ser Val Pro Leu Lys Asp Pro Ile
 85 90 95

Lys Ile His Glu Ser Met Arg Tyr Ser Leu Leu Ala Gly Gly Lys Arg
 100 105 110

Ile Arg Pro Met Met Cys Ile Ala Ala Cys Glu Ile Val Gly Gly Asn
 115 120 125

Ile Leu Asn Ala Met Pro Ala Ala Cys Ala Val Glu Met Ile His Thr
 130 135 140

Met Ser Leu Val His Asp Asp Leu Pro Cys Met Asp Asn Asp Asp Phe
 145 150 155 160

Arg Arg Gly Lys Pro Ile Ser His Lys Val Tyr Gly Glu Glu Met Ala
165 170 175

Val Leu Thr Gly Asp Ala Leu Leu Ser Leu Ser Phe Glu His Ile Ala
180 185 190

Thr Ala Thr Lys Gly Val Ser Lys Asp Arg Ile Val Arg Ala Ile Gly
195 200 205

Glu Leu Ala Arg Ser Val Gly Ser Glu Gly Leu Val Ala Gly Gln Val
210 215 220

Val Asp Ile Leu Ser Glu Gly Ala Asp Val Gly Leu Asp His Leu Glu
225 230 235 240

Tyr Ile His Ile His Lys Thr Ala Met Leu Leu Glu Ser Ser Val Val
245 250 255

Ile Gly Ala Ile Met Gly Gly Gly Ser Asp Gln Gln Ile Glu Lys Leu
260 265 270

Arg Lys Phe Ala Arg Ser Ile Gly Leu Leu Phe Gln Val Val Asp Asp
275 280 285

Ile Leu Asp Val Thr Lys Ser Thr Glu Glu Leu Gly Lys Thr Ala Gly
290 295 300

Lys Asp Leu Leu Thr Asp Lys Thr Thr Tyr Pro Lys Leu Leu Gly Ile
305 310 315 320

Glu Lys Ser Arg Glu Phe Ala Glu Lys Leu Asn Lys Glu Ala Gln Glu
325 330 335

Gln Leu Ser Gly Phe Asp Arg Arg Lys Ala Ala Pro Leu Ile Ala Leu
340 345 350

Ala Asn Tyr Asn Ala Tyr Arg Gln Asn
355 360

<210> 122

<211> 342

<212> PRT

<213> Gibberella fujikuroi

<400> 122

Met Ala Glu Gln Gln Ile Ser Asn Leu Leu Ser Met Phe Asp Ala Ser
1 5 10 15

His Ala Ser Gln Lys Leu Glu Ile Thr Val Gln Met Met Asp Thr Tyr
20 25 30

His Tyr Arg Glu Thr Pro Pro Asp Ser Ser Ser Ser Glu Gly Gly Ser
35 40 45

Leu Ser Arg Tyr Asp Glu Arg Arg Val Ser Leu Pro Leu Ser His Asn
 50 55 60
 Ala Ala Ser Pro Asp Ile Val Ser Gln Leu Cys Phe Ser Thr Ala Met
 65 70 75 80
 Ser Ser Glu Leu Asn His Arg Trp Lys Ser Gln Arg Leu Lys Val Ala
 85 90 95
 Asp Ser Pro Tyr Asn Tyr Ile Leu Thr Leu Pro Ser Lys Gly Ile Arg
 100 105 110
 Gly Ala Phe Ile Asp Ser Leu Asn Val Trp Leu Glu Val Pro Glu Asp
 115 120 125
 Glu Thr Ser Val Ile Lys Glu Val Ile Gly Met Leu His Asn Ser Ser
 130 135 140
 Leu Ile Ile Asp Asp Phe Gln Asp Asn Ser Pro Leu Arg Arg Gly Lys
 145 150 155 160
 Pro Ser Thr His Thr Val Phe Gly Pro Ala Gln Ala Ile Asn Thr Ala
 165 170 175
 Thr Tyr Val Ile Val Lys Ala Ile Glu Lys Ile Gln Asp Ile Val Gly
 180 185 190
 His Asp Ala Leu Ala Asp Val Thr Gly Thr Ile Thr Thr Ile Phe Gln
 195 200 205
 Gly Gln Ala Met Asp Leu Trp Trp Thr Ala Asn Ala Ile Val Pro Ser
 210 215 220
 Ile Gln Glu Tyr Leu Leu Met Val Asn Asp Lys Thr Gly Ala Leu Phe
 225 230 235 240
 Arg Leu Ser Leu Glu Leu Leu Ala Leu Asn Ser Glu Ala Ser Ile Ser
 245 250 255
 Asp Ser Ala Leu Glu Ser Leu Ser Ser Ala Val Ser Leu Leu Gly Gln
 260 265 270
 Tyr Phe Gln Ile Arg Asp Asp Tyr Met Asn Leu Ile Asp Asn Lys Tyr
 275 280 285
 Thr Asp Gln Lys Gly Phe Cys Glu Asp Leu Asp Glu Gly Lys Tyr Ser
 290 295 300
 Leu Thr Leu Ile His Ala Leu Gln Thr Asp Ser Ser Asp Leu Leu Thr
 305 310 315 320

Asn Ile Leu Ser Met Arg Arg Val Gln Gly Lys Leu Thr Ala Gln Lys
 325 330 335

Arg Cys Trp Phe Trp Lys
 340

<210> 123

<211> 300

<212> PRT

<213> Mus musculus

<400> 123

Met Glu Lys Thr Lys Glu Lys Ala Glu Arg Ile Leu Leu Glu Pro Tyr
 1 5 10 15

Arg Tyr Leu Leu Gln Leu Pro Gly Lys Gln Val Arg Ser Lys Leu Ser
 20 25 30

Gln Ala Phe Asn His Trp Leu Lys Val Pro Glu Asp Lys Leu Gln Ile
 35 40 45

Ile Ile Glu Val Thr Glu Met Leu His Asn Ala Ser Leu Leu Ile Asp
 50 55 60

Asp Ile Glu Asp Ser Ser Lys Leu Arg Arg Gly Phe Pro Val Ala His
 65 70 75 80

Ser Ile Tyr Gly Val Pro Ser Val Ile Asn Ser Ala Asn Tyr Val Tyr
 85 90 95

Phe Leu Gly Leu Glu Lys Val Leu Thr Leu Asp His Pro Asp Ala Val
 100 105 110

Lys Leu Phe Thr Arg Gln Leu Leu Glu Leu His Gln Gly Gln Gly Leu
 115 120 125

Asp Ile Tyr Trp Arg Asp Thr Tyr Thr Cys Pro Thr Glu Glu Glu Tyr
 130 135 140

Lys Ala Met Val Leu Gln Lys Thr Gly Gly Leu Phe Gly Leu Ala Val
 145 150 155 160

Gly Leu Met Gln Leu Phe Ser Asp Tyr Lys Glu Asp Leu Lys Pro Leu
 165 170 175

Leu Asp Thr Leu Gly Leu Phe Phe Gln Ile Arg Asp Asp Tyr Ala Asn
 180 185 190

Leu His Ser Lys Glu Tyr Ser Glu Asn Lys Ser Phe Cys Glu Asp Leu
 195 200 205

Thr Glu Gly Lys Phe Ser Phe Pro Thr Ile His Ala Ile Trp Ser Arg
 210 215 220

Pro Glu Ser Thr Gln Val Gln Asn Ile Leu Arg Gln Arg Thr Glu Asn
225 230 235 240

Ile Asp Ile Lys Lys Tyr Cys Val Gln Tyr Leu Glu Asp Val Gly Ser
245 250 255

Phe Ala Tyr Thr Arg His Thr Leu Arg Glu Leu Glu Ala Lys Ala Tyr
260 265 270

Lys Gln Ile Glu Ala Cys Gly Gly Asn Pro Ser Leu Val Ala Leu Val
275 280 285

Lys His Leu Ser Lys Met Phe Thr Glu Glu Asn Lys
290 295 300

<210> 124

<211> 339

<212> PRT

<213> Thalassiosira pseudonana

<400> 124

Met Ala Arg Phe Tyr Phe Leu Asn Ala Leu Leu Met Val Ile Ser Leu
1 5 10 15

Gln Ser Thr Thr Ala Phe Thr Pro Ala Lys Leu Ala Tyr Pro Thr Thr
20 25 30

Thr Thr Ala Leu Asn Val Ala Ser Ala Glu Thr Ser Phe Ser Leu Asp
35 40 45

Glu Tyr Leu Ala Ser Lys Ile Gly Pro Ile Glu Ser Ala Leu Glu Ala
50 55 60

Ser Val Lys Ser Arg Ile Pro Gln Thr Asp Lys Ile Cys Glu Ser Met
65 70 75 80

Ala Tyr Ser Leu Met Ala Gly Gly Lys Arg Ile Arg Pro Val Leu Cys
85 90 95

Ile Ala Ala Cys Glu Met Phe Gly Gly Ser Gln Asp Val Ala Met Pro
100 105 110

Thr Ala Val Ala Leu Glu Met Ile His Thr Met Ser Leu Ile His Asp
115 120 125

Asp Leu Pro Ser Met Asp Asn Asp Asp Leu Arg Arg Gly Lys Pro Thr
130 135 140

Asn His Val Val Phe Gly Glu Asp Val Ala Ile Leu Ala Gly Asp Ser
145 150 155 160

Leu Leu Ser Thr Ser Phe Glu His Val Ala Arg Glu Thr Lys Gly Val

Ser Ala Glu Lys Ile Val Asp Val Ile Ala Arg Leu Gly Lys Ser Val
180 185 190

Gly Ala Glu Gly Leu Ala Gly Gly Gln Val Met Asp Leu Glu Cys Glu
195 200 205

Ala Lys Pro Gly Thr Thr Leu Asp Asp Leu Lys Trp Ile His Ile His
210 215 220

Lys Thr Ala Thr Leu Leu Gln Val Ala Val Ala Ser Gly Ala Val Leu
225 230 235 240

Gly Gly Ala Thr Pro Glu Glu Val Ala Ala Cys Glu Leu Phe Ala Met
245 250 255

Asn Ile Gly Leu Ala Phe Gln Val Ala Asp Asp Ile Leu Asp Val Thr
260 265 270

Ala Ser Ser Glu Asp Leu Gly Lys Thr Ala Gly Lys Asp Glu Ala Thr
275 280 285

Asp Lys Thr Thr Tyr Pro Lys Leu Leu Gly Leu Glu Glu Ser Lys Ala
290 295 300

Tyr Ala Arg Gln Leu Ile Asp Glu Ala Lys Glu Ser Leu Ala Pro Phe
305 310 315 320

Gly Asp Arg Ala Ala Pro Leu Leu Ala Ile Ala Asp Phe Ile Ile Asp
325 330 335

Arg Lys Asn

<210> 125
<211> 355
<212> PRT
<213> Streptomyces clavuligerus

<400> 125
Met His Leu Ala Pro Arg Arg Val Pro Arg Gly Arg Arg Ser Pro Pro
1 5 10 15

Asp Arg Val Pro Glu Arg Gln Gly Ala Leu Gly Arg Arg Arg Gly Ala
20 25 30

Gly Ser Thr Gly Cys Ala Arg Ala Ala Ala Gly Val His Arg Arg Arg
35 40 45

Gly Gly Gly Glu Ala Asp Pro Ser Ala Ala Val His Arg Gly Trp Gln
50 55 60

Ala Gly Gly Gly Thr Gly Leu Pro Asp Glu Val Val Ser Thr Ala Ala
65 70 75 80

Ala Leu Glu Met Phe His Ala Phe Ala Leu Ile His Asp Asp Ile Met
85 90 95

Asp Asp Ser Ala Thr Arg Arg Gly Ser Pro Thr Val His Arg Ala Leu
100 105 110

Ala Asp Arg Leu Gly Ala Ala Leu Asp Pro Asp Gln Ala Gly Gln Leu
115 120 125

Gly Val Ser Thr Ala Ile Leu Val Gly Asp Leu Ala Leu Thr Trp Ser
130 135 140

Asp Glu Leu Leu Tyr Ala Pro Leu Thr Pro His Arg Leu Ala Ala Val
145 150 155 160

Leu Pro Leu Val Thr Ala Met Arg Ala Glu Thr Val His Gly Gln Tyr
165 170 175

Leu Asp Ile Thr Ser Ala Arg Arg Pro Gly Thr Asp Thr Ser Leu Ala
180 185 190

Leu Arg Ile Ala Arg Tyr Lys Thr Ala Ala Tyr Thr Met Glu Arg Pro
195 200 205

Leu His Ile Gly Ala Ala Leu Ala Gly Ala Arg Pro Glu Leu Leu Ala
210 215 220

Gly Leu Ser Ala Tyr Ala Leu Pro Ala Gly Glu Ala Phe Gln Leu Ala
225 230 235 240

Asp Asp Leu Leu Gly Val Phe Gly Asp Pro Arg Arg Thr Gly Lys Pro
245 250 255

Asp Leu Asp Asp Leu Arg Gly Gly Lys His Thr Val Leu Val Ala Leu
260 265 270

Ala Arg Glu His Ala Thr Pro Glu Gln Arg His Thr Leu Asp Thr Leu
275 280 285

Leu Gly Thr Pro Gly Leu Asp Arg Gln Gly Ala Ser Arg Leu Arg Cys
290 295 300

Val Leu Val Ala Thr Gly Ala Arg Ala Glu Ala Glu Arg Leu Ile Thr
305 310 315 320

Glu Arg Arg Asp Gln Ala Leu Thr Ala Leu Asn Ala Leu Thr Leu Pro
325 330 335

Pro Pro Leu Ala Glu Ala Leu Ala Arg Leu Thr Leu Gly Ser Thr Ala

340

345

350

His Pro Ala
355

<210> 126

<211> 330

<212> PRT

<213> Sulfolobus acidicaldarius

<400> 126

Met Ser Tyr Phe Asp Asn Tyr Phe Asn Glu Ile Val Asn Ser Val Asn
1 5 10 15

Asp Ile Ile Lys Ser Tyr Ile Ser Gly Asp Val Pro Lys Leu Tyr Glu
20 25 30

Ala Ser Tyr His Leu Phe Thr Ser Gly Gly Lys Arg Leu Arg Pro Leu
35 40 45

Ile Leu Thr Ile Ser Ser Asp Leu Phe Gly Gly Gln Arg Glu Arg Ala
50 55 60

Tyr Tyr Ala Gly Ala Ala Ile Glu Val Leu His Thr Phe Thr Leu Val
65 70 75 80

His Asp Asp Ile Met Asp Gln Asp Asn Ile Arg Arg Gly Leu Pro Thr
85 90 95

Val His Val Lys Tyr Gly Leu Pro Leu Ala Ile Leu Ala Gly Asp Leu
100 105 110

Leu His Ala Lys Ala Phe Gln Leu Leu Thr Gln Ala Leu Arg Gly Leu
115 120 125

Pro Ser Glu Thr Ile Ile Lys Ala Phe Asp Ile Phe Thr Arg Ser Ile
130 135 140

Ile Ile Ile Ser Glu Gly Gln Ala Val Asp Met Glu Phe Glu Asp Arg
145 150 155 160

Ile Asp Ile Lys Glu Gln Glu Tyr Leu Asp Met Ile Ser Arg Lys Thr
165 170 175

Ala Ala Leu Phe Ser Ala Ser Ser Ser Ile Gly Ala Leu Ile Ala Gly
180 185 190

Ala Asn Asp Asn Asp Val Arg Leu Met Ser Asp Phe Gly Thr Asn Leu
195 200 205

Gly Ile Ala Phe Gln Ile Val Asp Asp Ile Leu Gly Leu Thr Ala Asp
210 215 220

Glu Lys Glu Leu Gly Lys Pro Val Phe Ser Asp Ile Arg Glu Gly Lys
225 230 235 240

Lys Thr Ile Leu Val Ile Lys Thr Leu Glu Leu Cys Lys Glu Asp Glu
245 250 255

Lys Lys Ile Val Leu Lys Ala Leu Gly Asn Lys Ser Ala Ser Lys Glu
260 265 270

Glu Leu Met Ser Ser Ala Asp Ile Ile Lys Lys Tyr Ser Leu Asp Tyr
275 280 285

Ala Tyr Asn Leu Ala Glu Lys Tyr Tyr Lys Asn Ala Ile Asp Ser Leu
290 295 300

Asn Gln Val Ser Ser Lys Ser Asp Ile Pro Gly Lys Ala Leu Lys Tyr
305 310 315 320

Leu Ala Glu Phe Thr Ile Arg Arg Arg Lys
325 330

<210> 127

<211> 297

<212> PRT

<213> Synechococcus sp.

<400> 127

Met Val Ala Gln Thr Phe Asn Leu Asp Thr Tyr Leu Ser Gln Arg Gln
1 5 10 15

Gln Gln Val Glu Glu Ala Leu Ser Ala Ala Leu Val Pro Ala Tyr Pro
20 25 30

Glu Arg Ile Tyr Glu Ala Met Arg Tyr Ser Leu Leu Ala Gly Gly Lys
35 40 45

Arg Leu Arg Pro Ile Leu Cys Leu Ala Ala Cys Glu Leu Ala Gly Gly
50 55 60

Ser Val Glu Gln Ala Met Pro Thr Ala Cys Ala Leu Glu Met Ile His
65 70 75 80

Thr Met Ser Leu Ile His Asp Asp Leu Pro Ala Met Asp Asn Asp Asp
85 90 95

Phe Arg Arg Gly Lys Pro Thr Asn His Lys Val Phe Gly Glu Asp Ile
100 105 110

Ala Ile Leu Ala Gly Asp Ala Leu Leu Ala Tyr Ala Phe Glu His Ile
115 120 125

Ala Ser Gln Thr Arg Gly Val Pro Pro Gln Leu Val Leu Gln Val Ile
130 135 140

Ala Arg Ile Gly His Ala Val Ala Ala Thr Gly Leu Val Gly Gly Gln
145 150 155 160

Val Val Asp Leu Glu Ser Glu Gly Lys Ala Ile Ser Leu Glu Thr Leu
165 170 175

Glu Tyr Ile His Ser His Lys Thr Gly Ala Leu Leu Glu Ala Ser Val
180 185 190

Val Ser Gly Gly Ile Leu Ala Gly Ala Asp Glu Glu Leu Leu Ala Arg
195 200 205

Leu Ser His Tyr Ala Arg Asp Ile Gly Leu Ala Phe Gln Ile Val Asp
210 215 220

Asp Ile Leu Asp Val Thr Ala Thr Ser Glu Gln Leu Gly Lys Thr Ala
225 230 235 240

Gly Lys Asp Gln Ala Ala Ala Lys Ala Thr Tyr Pro Ser Leu Leu Gly
245 250 255

Leu Glu Ala Ser Arg Gln Lys Ala Glu Glu Leu Ile Gln Ser Ala Lys
260 265 270

Glu Ala Leu Arg Pro Tyr Gly Ser Gln Ala Glu Pro Leu Leu Ala Leu
275 280 285

Ala Asp Phe Ile Thr Arg Arg Gln His
290 295

<210> 128

<211> 371

<212> PRT

<213> Arabidopsis thaliana

<400> 128

Met Ala Ser Val Thr Leu Gly Ser Trp Ile Val Val His His Asn
1 5 10 15

His His His Pro Ser Ser Ile Leu Thr Lys Ser Arg Ser Arg Ser Cys
20 25 30

Pro Ile Thr Leu Thr Lys Pro Ile Ser Phe Arg Ser Lys Arg Thr Val
35 40 45

Ser Ser Ser Ser Ser Ile Val Ser Ser Ser Val Val Thr Lys Glu Asp
50 55 60

Asn Leu Arg Gln Ser Glu Pro Ser Ser Phe Asp Phe Met Ser Tyr Ile
65 70 75 80

Ile Thr Lys Ala Glu Leu Val Asn Lys Ala Leu Asp Ser Ala Val Pro
85 90 95

Leu Arg Glu Pro Leu Lys Ile His Glu Ala Met Arg Tyr Ser Leu Leu
 100 105 110
 Ala Gly Gly Lys Arg Val Arg Pro Val Leu Cys Ile Ala Ala Cys Glu
 115 120 125
 Leu Val Gly Gly Glu Glu Ser Thr Ala Met Pro Ala Ala Cys Ala Val
 130 135 140
 Glu Met Ile His Thr Met Ser Leu Ile His Asp Asp Leu Pro Cys Met
 145 150 155 160
 Asp Asn Asp Asp Leu Arg Arg Gly Lys Pro Thr Asn His Lys Val Phe
 165 170 175
 Gly Glu Asp Val Ala Val Leu Ala Gly Asp Ala Leu Leu Ser Phe Ala
 180 185 190
 Phe Glu His Leu Ala Ser Ala Thr Ser Ser Asp Val Val Ser Pro Val
 195 200 205
 Arg Val Val Arg Ala Val Gly Glu Leu Ala Lys Ala Ile Gly Thr Glu
 210 215 220
 Gly Leu Val Ala Gly Gln Val Val Asp Ile Ser Ser Glu Gly Leu Asp
 225 230 235 240
 Leu Asn Asp Val Gly Leu Glu His Leu Glu Phe Ile His Leu His Lys
 245 250 255
 Thr Ala Ala Leu Leu Glu Ala Ser Ala Val Leu Gly Ala Ile Val Gly
 260 265 270
 Gly Gly Ser Asp Asp Glu Ile Glu Arg Leu Arg Lys Phe Ala Arg Cys
 275 280 285
 Ile Gly Leu Leu Phe Gln Val Val Asp Asp Ile Leu Asp Val Thr Lys
 290 295 300
 Ser Ser Lys Glu Leu Gly Lys Thr Ala Gly Lys Asp Leu Ile Ala Asp
 305 310 315 320
 Lys Leu Thr Tyr Pro Lys Ile Met Gly Leu Glu Lys Ser Arg Glu Phe
 325 330 335
 Ala Glu Lys Leu Asn Arg Glu Ala Arg Asp Gln Leu Leu Gly Phe Asp
 340 345 350
 Ser Asp Lys Val Ala Pro Leu Leu Ala Leu Ala Asn Tyr Ile Ala Tyr
 355 360 365

Arg Gln Asn
370

<210> 129

<211> 787

<212> PRT

<213> Stevia rebaudiana

<400> 129

Met Lys Thr Gly Phe Ile Ser Pro Ala Thr Val Phe His His Arg Ile
1 5 10 15

Ser Pro Ala Thr Thr Phe Arg His His Leu Ser Pro Ala Thr Thr Asn
20 25 30

Ser Thr Gly Ile Val Ala Leu Arg Asp Ile Asn Phe Arg Cys Lys Ala
35 40 45

Val Ser Lys Glu Tyr Ser Asp Leu Leu Gln Lys Asp Glu Ala Ser Phe
50 55 60

Thr Lys Trp Asp Asp Asp Lys Val Lys Asp His Leu Asp Thr Asn Lys
65 70 75 80

Asn Leu Tyr Pro Asn Asp Glu Ile Lys Glu Phe Val Glu Ser Val Lys
85 90 95

Ala Met Phe Gly Ser Met Asn Asp Gly Glu Ile Asn Val Ser Ala Tyr
100 105 110

Asp Thr Ala Trp Val Ala Leu Val Gln Asp Val Asp Gly Ser Gly Ser
115 120 125

Pro Gln Phe Pro Ser Ser Leu Glu Trp Ile Ala Asn Asn Gln Leu Ser
130 135 140

Asp Gly Ser Trp Gly Asp His Leu Leu Phe Ser Ala His Asp Arg Ile
145 150 155 160

Ile Asn Thr Leu Ala Cys Val Ile Ala Leu Thr Ser Trp Asn Val His
165 170 175

Pro Ser Lys Cys Glu Lys Gly Leu Asn Phe Leu Arg Glu Asn Ile Cys
180 185 190

Lys Leu Glu Asp Glu Asn Ala Glu His Met Pro Ile Gly Phe Glu Val
195 200 205

Thr Phe Pro Ser Leu Ile Asp Ile Ala Lys Lys Leu Asn Ile Glu Val
210 215 220

Pro Glu Asp Thr Pro Ala Leu Lys Glu Ile Tyr Ala Arg Arg Asp Ile
225 230 235 240

Lys Leu Thr Lys Ile Pro Met Glu Val Leu His Lys Val Pro Thr Thr
245 250 255

Leu Leu His Ser Leu Glu Gly Met Pro Asp Leu Glu Trp Glu Lys Leu
260 265 270

Leu Lys Leu Gln Cys Lys Asp Gly Ser Phe Leu Phe Ser Pro Ser Ser
275 280 285

Thr Ala Phe Ala Leu Met Gln Thr Lys Asp Glu Lys Cys Leu Gln Tyr
290 295 300

Leu Thr Asn Ile Val Thr Lys Phe Asn Gly Gly Val Pro Asn Val Tyr
305 310 315 320

Pro Val Asp Leu Phe Glu His Ile Trp Val Val Asp Arg Leu Gln Arg
325 330 335

Leu Gly Ile Ala Arg Tyr Phe Lys Ser Glu Ile Lys Asp Cys Val Glu
340 345 350

Tyr Ile Asn Lys Tyr Trp Thr Lys Asn Gly Ile Cys Trp Ala Arg Asn
355 360 365

Thr His Val Gln Asp Ile Asp Asp Thr Ala Met Gly Phe Arg Val Leu
370 375 380

Arg Ala His Gly Tyr Asp Val Thr Pro Asp Val Phe Arg Gln Phe Glu
385 390 395 400

Lys Asp Gly Lys Phe Val Cys Phe Ala Gly Gln Ser Thr Gln Ala Val
405 410 415

Thr Gly Met Phe Asn Val Tyr Arg Ala Ser Gln Met Leu Phe Pro Gly
420 425 430

Glu Arg Ile Leu Glu Asp Ala Lys Lys Phe Ser Tyr Asn Tyr Leu Lys
435 440 445

Glu Lys Gln Ser Thr Asn Glu Leu Leu Asp Lys Trp Ile Ile Ala Lys
450 455 460

Asp Leu Pro Gly Glu Val Gly Tyr Ala Leu Asp Ile Pro Trp Tyr Ala
465 470 475 480

Ser Leu Pro Arg Leu Glu Thr Arg Tyr Tyr Leu Glu Gln Tyr Gly Gly
485 490 495

Glu Asp Asp Val Trp Ile Gly Lys Thr Leu Tyr Arg Met Gly Tyr Val
500 505 510

Ser Asn Asn Thr Tyr Leu Glu Met Ala Lys Leu Asp Tyr Asn Asn Tyr
 515 520 525
 Val Ala Val Leu Gln Leu Glu Trp Tyr Thr Ile Gln Gln Trp Tyr Val
 530 535 540
 Asp Ile Gly Ile Glu Lys Phe Glu Ser Asp Asn Ile Lys Ser Val Leu
 545 550 555 560
 Val Ser Tyr Tyr Leu Ala Ala Ala Ser Ile Phe Glu Pro Glu Arg Ser
 565 570 575
 Lys Glu Arg Ile Ala Trp Ala Lys Thr Thr Ile Leu Val Asp Lys Ile
 580 585 590
 Thr Ser Ile Phe Asp Ser Ser Gln Ser Ser Lys Glu Asp Ile Thr Ala
 595 600 605
 Phe Ile Asp Lys Phe Arg Asn Lys Ser Ser Ser Lys Lys His Ser Ile
 610 615 620
 Asn Gly Glu Pro Trp His Glu Val Met Val Ala Leu Lys Lys Thr Leu
 625 630 635 640
 His Gly Phe Ala Leu Asp Ala Leu Met Thr His Ser Gln Asp Ile His
 645 650 655
 Pro Gln Leu His Gln Ala Trp Glu Met Trp Leu Thr Lys Leu Gln Asp
 660 665 670
 Gly Val Asp Val Thr Ala Glu Leu Met Val Gln Met Ile Asn Met Thr
 675 680 685
 Ala Gly Arg Trp Val Ser Lys Glu Leu Leu Thr His Pro Gln Tyr Gln
 690 695 700
 Arg Leu Ser Thr Val Thr Asn Ser Val Cys His Asp Ile Thr Lys Leu
 705 710 715 720
 His Asn Phe Lys Glu Asn Ser Thr Thr Val Asp Ser Lys Val Gln Glu
 725 730 735
 Leu Val Gln Leu Val Phe Ser Asp Thr Pro Asp Asp Leu Asp Gln Asp
 740 745 750
 Met Lys Gln Thr Phe Leu Thr Val Met Lys Thr Phe Tyr Tyr Lys Ala
 755 760 765
 Trp Cys Asp Pro Asn Thr Ile Asn Asp His Ile Ser Lys Val Phe Glu
 770 775 780

Ile Val Ile
785

<210> 130
<211> 527
<212> PRT
<213> Streptomyces clavuligerus

<400> 130
Met Pro Asp Ala His Asp Ala Pro Pro Pro Gln Ile Arg Gln Arg Thr
1 5 10 15

Leu Val Asp Glu Ala Thr Gln Leu Leu Thr Glu Ser Ala Glu Asp Ala
20 25 30

Trp Gly Glu Val Ser Val Ser Glu Tyr Glu Thr Ala Arg Leu Val Ala
35 40 45

His Ala Thr Trp Leu Gly Gly His Ala Thr Arg Val Ala Phe Leu Leu
50 55 60

Glu Arg Gln His Glu Asp Gly Ser Trp Gly Pro Pro Gly Gly Tyr Arg
65 70 75 80

Leu Val Pro Thr Leu Ser Ala Val His Ala Leu Leu Thr Cys Leu Ala
85 90 95

Ser Pro Ala Gln Asp His Gly Val Pro His Asp Arg Leu Leu Arg Ala
100 105 110

Val Asp Ala Gly Leu Thr Ala Leu Arg Arg Leu Gly Thr Ser Asp Ser
115 120 125

Pro Pro Asp Thr Ile Ala Val Glu Leu Val Ile Pro Ser Leu Leu Glu
130 135 140

Gly Ile Gln His Leu Leu Asp Pro Ala His Pro His Ser Arg Pro Ala
145 150 155 160

Phe Ser Gln His Arg Gly Ser Leu Val Cys Pro Gly Gly Leu Asp Gly
165 170 175

Arg Thr Leu Gly Ala Leu Arg Ser His Ala Ala Ala Gly Thr Pro Val
180 185 190

Pro Gly Lys Val Trp His Ala Ser Glu Thr Leu Gly Leu Ser Thr Glu
195 200 205

Ala Ala Ser His Leu Gln Pro Ala Gln Gly Ile Ile Gly Gly Ser Ala
210 215 220

Ala Ala Thr Ala Thr Trp Leu Thr Arg Val Ala Pro Ser Gln Gln Ser
225 230 235 240

Asp Ser Ala Arg Arg Tyr Leu Glu Glu Leu Gln His Arg Tyr Ser Gly
 245 250 255
 Pro Val Pro Ser Ile Thr Pro Ile Thr Tyr Phe Glu Arg Ala Trp Leu
 260 265 270
 Leu Asn Asn Phe Ala Ala Ala Gly Val Pro Cys Glu Ala Pro Ala Ala
 275 280 285
 Leu Leu Asp Ser Leu Glu Ala Ala Leu Thr Pro Gln Gly Ala Pro Ala
 290 300
 Gly Ala Gly Leu Pro Pro Asp Ala Asp Asp Thr Ala Ala Val Leu Leu
 305 310 315 320
 Ala Leu Ala Thr His Gly Arg Gly Arg Arg Pro Glu Val Leu Met Asp
 325 330 335
 Tyr Arg Thr Asp Gly Tyr Phe Gln Cys Phe Ile Gly Glu Arg Thr Pro
 340 345 350
 Ser Ile Ser Thr Asn Ala His Val Leu Glu Thr Leu Gly His His Val
 355 360 365
 Ala Gln His Pro Gln Asp Arg Ala Arg Tyr Gly Ser Ala Met Asp Thr
 370 375 380
 Ala Ser Ala Trp Leu Leu Ala Ala Gln Lys Gln Asp Gly Ser Trp Leu
 385 390 395 400
 Asp Lys Trp His Ala Ser Pro Tyr Tyr Ala Thr Val Cys Cys Thr Gln
 405 410 415
 Ala Leu Ala Ala His Ala Ser Pro Ala Thr Ala Pro Ala Arg Gln Arg
 420 425 430
 Ala Val Arg Trp Val Leu Ala Thr Gln Arg Ser Asp Gly Gly Trp Gly
 435 440 445
 Leu Trp His Ser Thr Val Glu Glu Thr Ala Tyr Ala Leu Gln Ile Leu
 450 455 460
 Ala Pro Pro Ser Gly Gly Gly Asn Ile Pro Val Gln Gln Ala Leu Thr
 465 470 475 480
 Arg Gly Arg Ala Arg Leu Cys Gly Ala Leu Pro Leu Thr Pro Leu Trp
 485 490 495
 His Asp Lys Asp Leu Tyr Thr Pro Val Arg Val Val Arg Ala Ala Arg
 500 505 510

Ala Ala Ala Leu Tyr Thr Thr Arg Asp Leu Leu Leu Pro Pro Leu
 515 520 525

<210> 131

<211> 516

<212> PRT

<213> Bradyrhizobium japonicum

<400> 131

Met Asn Ala Leu Ser Glu His Ile Leu Ser Glu Leu Arg Arg Leu Leu
 1 5 10 15

Ser Glu Met Ser Asp Gly Gly Ser Val Gly Pro Ser Val Tyr Asp Thr
 20 25 30

Ala Gln Ala Leu Arg Phe His Gly Asn Val Thr Gly Arg Gln Asp Ala
 35 40 45

Tyr Ala Trp Leu Ile Ala Gln Gln Ala Asp Gly Gly Trp Gly Ser
 50 55 60

Ala Asp Phe Pro Leu Phe Arg His Ala Pro Thr Trp Ala Ala Leu Leu
 65 70 75 80

Ala Leu Gln Arg Ala Asp Pro Leu Pro Gly Ala Ala Asp Ala Val Gln
 85 90 95

Thr Ala Thr Arg Phe Leu Gln Arg Gln Pro Asp Pro Tyr Ala His Ala
 100 105 110

Val Pro Glu Asp Ala Pro Ile Gly Ala Glu Leu Ile Leu Pro Gln Phe
 115 120 125

Cys Gly Glu Ala Ala Trp Leu Leu Gly Gly Val Ala Phe Pro Arg His
 130 135 140

Pro Ala Leu Leu Pro Leu Arg Gln Ala Cys Leu Val Lys Leu Gly Ala
 145 150 155 160

Val Ala Met Leu Pro Ser Gly His Pro Leu Leu His Ser Trp Glu Ala
 165 170 175

Trp Gly Thr Ser Pro Thr Thr Ala Cys Pro Asp Asp Asp Gly Ser Ile
 180 185 190

Gly Ile Ser Pro Ala Ala Thr Ala Ala Trp Arg Ala Gln Ala Val Thr
 195 200 205

Arg Gly Ser Thr Pro Gln Val Gly Arg Ala Asp Ala Tyr Leu Gln Met
 210 215 220

Ala Ser Arg Ala Thr Arg Ser Gly Ile Glu Gly Val Phe Pro Asn Val
 225 230 235 240

Trp Pro Ile Asn Val Phe Glu Pro Cys Trp Ser Leu Tyr Thr Leu His
 245 250 255
 Leu Ala Gly Leu Phe Ala His Pro Ala Leu Ala Glu Ala Val Arg Val
 260 265 270
 Ile Val Ala Gln Leu Glu Ala Arg Leu Gly Val His Gly Leu Gly Pro
 275 280 285
 Ala Leu His Phe Ala Ala Asp Ala Asp Asp Thr Ala Val Ala Leu Cys
 290 295 300
 Val Leu His Leu Ala Gly Arg Asp Pro Ala Val Asp Ala Leu Arg His
 305 310 315 320
 Phe Glu Ile Gly Glu Leu Phe Val Thr Phe Pro Gly Glu Arg Asn Ala
 325 330 335
 Ser Val Ser Thr Asn Ile His Ala Leu His Ala Leu Arg Leu Leu Gly
 340 345 350
 Lys Pro Ala Ala Gly Ala Ser Ala Tyr Val Glu Ala Asn Arg Asn Pro
 355 360 365
 His Gly Leu Trp Asp Asn Glu Lys Trp His Val Ser Trp Leu Tyr Pro
 370 375 380
 Thr Ala His Ala Val Ala Ala Leu Ala Gln Gly Lys Pro Gln Trp Arg
 385 390 395 400
 Asp Glu Arg Ala Leu Ala Ala Leu Leu Gln Ala Gln Arg Asp Asp Gly
 405 410 415
 Gly Trp Gly Ala Gly Arg Gly Ser Thr Phe Glu Glu Thr Ala Tyr Ala
 420 425 430
 Leu Phe Ala Leu His Val Met Asp Gly Ser Glu Glu Ala Thr Gly Arg
 435 440 445
 Arg Arg Ile Ala Gln Val Val Ala Arg Ala Leu Glu Trp Met Leu Ala
 450 455 460
 Arg His Ala Ala His Gly Leu Pro Gln Thr Pro Leu Trp Ile Gly Lys
 465 470 475 480
 Glu Leu Tyr Cys Pro Thr Arg Val Val Arg Val Ala Glu Leu Ala Gly
 485 490 495
 Leu Trp Leu Ala Leu Arg Trp Gly Arg Arg Val Leu Ala Glu Gly Ala
 500 505 510

Gly Ala Ala Pro
515

<210> 132
<211> 784
<212> PRT
<213> Stevia rebaudiana

<400> 132
Met Asn Leu Ser Leu Cys Ile Ala Ser Pro Leu Leu Thr Lys Ser Asn
1 5 10 15

Arg Pro Ala Ala Leu Ser Ala Ile His Thr Ala Ser Thr Ser His Gly
20 25 30

Gly Gln Thr Asn Pro Thr Asn Leu Ile Ile Asp Thr Thr Lys Glu Arg
35 40 45

Ile Gln Lys Gln Phe Lys Asn Val Glu Ile Ser Val Ser Ser Tyr Asp
50 55 60

Thr Ala Trp Val Ala Met Val Pro Ser Pro Asn Ser Pro Lys Ser Pro
65 70 75 80

Cys Phe Pro Glu Cys Leu Asn Trp Leu Ile Asn Asn Gln Leu Asn Asp
85 90 95

Gly Ser Trp Gly Leu Val Asn His Thr His Asn His Asn His Pro Leu
100 105 110

Leu Lys Asp Ser Leu Ser Ser Thr Leu Ala Cys Ile Val Ala Leu Lys
115 120 125

Arg Trp Asn Val Gly Glu Asp Gln Ile Asn Lys Gly Leu Ser Phe Ile
130 135 140

Glu Ser Asn Leu Ala Ser Ala Thr Glu Lys Ser Gln Pro Ser Pro Ile
145 150 155 160

Gly Phe Asp Ile Ile Phe Pro Gly Leu Leu Glu Tyr Ala Lys Asn Leu
165 170 175

Asp Ile Asn Leu Leu Ser Lys Gln Thr Asp Phe Ser Leu Met Leu His
180 185 190

Lys Arg Glu Leu Glu Gln Lys Arg Cys His Ser Asn Glu Met Asp Gly
195 200 205

Tyr Leu Ala Tyr Ile Ser Glu Gly Leu Gly Asn Leu Tyr Asp Trp Asn
210 215 220

Met Val Lys Lys Tyr Gln Met Lys Asn Gly Ser Val Phe Asn Ser Pro
225 230 235 240

Ser Ala Thr Ala Ala Ala Phe Ile Asn His Gln Asn Pro Gly Cys Leu
 245 250 255
 Asn Tyr Leu Asn Ser Leu Leu Asp Lys Phe Gly Asn Ala Val Pro Thr
 260 265 270
 Val Tyr Pro His Asp Leu Phe Ile Arg Leu Ser Met Val Asp Thr Ile
 275 280 285
 Glu Arg Leu Gly Ile Ser His His Phe Arg Val Glu Ile Lys Asn Val
 290 295 300
 Leu Asp Glu Thr Tyr Arg Cys Trp Val Glu Arg Asp Glu Gln Ile Phe
 305 310 315 320
 Met Asp Val Val Thr Cys Ala Leu Ala Phe Arg Leu Leu Arg Ile Asn
 325 330 335
 Gly Tyr Glu Val Ser Pro Asp Pro Leu Ala Glu Ile Thr Asn Glu Leu
 340 345 350
 Ala Leu Lys Asp Glu Tyr Ala Ala Leu Glu Thr Tyr His Ala Ser His
 355 360 365
 Ile Leu Tyr Gln Glu Asp Leu Ser Ser Gly Lys Gln Ile Leu Lys Ser
 370 375 380
 Ala Asp Phe Leu Lys Glu Ile Ile Ser Thr Asp Ser Asn Arg Leu Ser
 385 390 395 400
 Lys Leu Ile His Lys Glu Val Glu Asn Ala Leu Lys Phe Pro Ile Asn
 405 410 415
 Thr Gly Leu Glu Arg Ile Asn Thr Arg Arg Asn Ile Gln Leu Tyr Asn
 420 425 430
 Val Asp Asn Thr Arg Ile Leu Lys Thr Thr Tyr His Ser Ser Asn Ile
 435 440 445
 Ser Asn Thr Asp Tyr Leu Arg Leu Ala Val Glu Asp Phe Tyr Thr Cys
 450 455 460
 Gln Ser Ile Tyr Arg Glu Glu Leu Lys Gly Leu Glu Arg Trp Val Val
 465 470 475 480
 Glu Asn Lys Leu Asp Gln Leu Lys Phe Ala Arg Gln Lys Thr Ala Tyr
 485 490 495
 Cys Tyr Phe Ser Val Ala Ala Thr Leu Ser Ser Pro Glu Leu Ser Asp
 500 505 510

Ala Arg Ile Ser Trp Ala Lys Asn Gly Ile Leu Thr Thr Val Val Asp
515 520 525

Asp Phe Phe Asp Ile Gly Gly Thr Ile Asp Glu Leu Thr Asn Leu Ile
530 535 540

Gln Cys Val Glu Lys Trp Asn Val Asp Val Asp Lys Asp Cys Cys Ser
545 550 555 560

Glu His Val Arg Ile Leu Phe Leu Ala Leu Lys Asp Ala Ile Cys Trp
565 570 575

Ile Gly Asp Glu Ala Phe Lys Trp Gln Ala Arg Asp Val Thr Ser His
580 585 590

Val Ile Gln Thr Trp Leu Glu Leu Met Asn Ser Met Leu Arg Glu Ala
595 600 605

Ile Trp Thr Arg Asp Ala Tyr Val Pro Thr Leu Asn Glu Tyr Met Glu
610 615 620

Asn Ala Tyr Val Ser Phe Ala Leu Gly Pro Ile Val Lys Pro Ala Ile
625 630 635 640

Tyr Phe Val Gly Pro Lys Leu Ser Glu Glu Ile Val Glu Ser Ser Glu
645 650 655

Tyr His Asn Leu Phe Lys Leu Met Ser Thr Gln Gly Arg Leu Leu Asn
660 665 670

Asp Ile His Ser Phe Lys Arg Glu Phe Lys Glu Gly Lys Leu Asn Ala
675 680 685

Val Ala Leu His Leu Ser Asn Gly Glu Ser Gly Lys Val Glu Glu Glu
690 695 700

Val Val Glu Glu Met Met Met Met Ile Lys Asn Lys Arg Lys Glu Leu
705 710 715 720

Met Lys Leu Ile Phe Glu Glu Asn Gly Ser Ile Val Pro Arg Ala Cys
725 730 735

Lys Asp Ala Phe Trp Asn Met Cys His Val Leu Asn Phe Phe Tyr Ala
740 745 750

Asn Asp Asp Gly Phe Thr Gly Asn Thr Ile Leu Asp Thr Val Lys Asp
755 760 765

Ile Ile Tyr Asn Pro Leu Val Leu Val Asn Glu Asn Glu Glu Gln Arg
770 775 780

<211> 784

<212> PRT

<213> Stevia rebaudiana

<400> 133

Met Asn Leu Ser Leu Cys Ile Ala Ser Pro Leu Leu Thr Lys Ser Ser
 1 5 10 15

Arg Pro Thr Ala Leu Ser Ala Ile His Thr Ala Ser Thr Ser His Gly
 20 25 30

Gly Gln Thr Asn Pro Thr Asn Leu Ile Ile Asp Thr Thr Lys Glu Arg
 35 40 45

Ile Gln Lys Leu Phe Lys Asn Val Glu Ile Ser Val Ser Ser Tyr Asp
 50 55 60

Thr Ala Trp Val Ala Met Val Pro Ser Pro Asn Ser Pro Lys Ser Pro
 65 70 75 80

Cys Phe Pro Glu Cys Leu Asn Trp Leu Ile Asn Asn Gln Leu Asn Asp
 85 90 95

Gly Ser Trp Gly Leu Val Asn His Thr His Asn His Asn His Pro Leu
 100 105 110

Leu Lys Asp Ser Leu Ser Ser Thr Leu Ala Cys Ile Val Ala Leu Lys
 115 120 125

Arg Trp Asn Val Gly Glu Asp Gln Ile Asn Lys Gly Leu Ser Phe Ile
 130 135 140

Glu Ser Asn Leu Ala Ser Ala Thr Asp Lys Ser Gln Pro Ser Pro Ile
 145 150 155 160

Gly Phe Asp Ile Ile Phe Pro Gly Leu Leu Glu Tyr Ala Lys Asn Leu
 165 170 175

Asp Ile Asn Leu Leu Ser Lys Gln Thr Asp Phe Ser Leu Met Leu His
 180 185 190

Lys Arg Glu Leu Glu Gln Lys Arg Cys His Ser Asn Glu Ile Asp Gly
 195 200 205

Tyr Leu Ala Tyr Ile Ser Glu Gly Leu Gly Asn Leu Tyr Asp Trp Asn
 210 215 220

Met Val Lys Lys Tyr Gln Met Lys Asn Gly Ser Val Phe Asn Ser Pro
 225 230 235 240

Ser Ala Thr Ala Ala Ala Phe Ile Asn His Gln Asn Pro Gly Cys Leu
 245 250 255

Asn Tyr Leu Asn Ser Leu Leu Asp Lys Phe Gly Asn Ala Val Pro Thr
 260 265 270
 Val Tyr Pro Leu Asp Leu Tyr Ile Arg Leu Ser Met Val Asp Thr Ile
 275 280 285
 Glu Arg Leu Gly Ile Ser His His Phe Arg Val Glu Ile Lys Asn Val
 290 295 300
 Leu Asp Glu Thr Tyr Arg Cys Trp Val Glu Arg Asp Glu Gln Ile Phe
 305 310 315 320
 Met Asp Val Val Thr Cys Ala Leu Ala Phe Arg Leu Leu Arg Ile His
 325 330 335
 Gly Tyr Lys Val Ser Pro Asp Gln Leu Ala Glu Ile Thr Asn Glu Leu
 340 345 350
 Ala Phe Lys Asp Glu Tyr Ala Ala Leu Glu Thr Tyr His Ala Ser Gln
 355 360 365
 Ile Leu Tyr Gln Glu Asp Leu Ser Ser Gly Lys Gln Ile Leu Lys Ser
 370 375 380
 Ala Asp Phe Leu Lys Gly Ile Leu Ser Thr Asp Ser Asn Arg Leu Ser
 385 390 395 400
 Lys Leu Ile His Lys Glu Val Glu Asn Ala Leu Lys Phe Pro Ile Asn
 405 410 415
 Thr Gly Leu Glu Arg Ile Asn Thr Arg Arg Asn Ile Gln Leu Tyr Asn
 420 425 430
 Val Asp Asn Thr Arg Ile Leu Lys Thr Thr Tyr His Ser Ser Asn Ile
 435 440 445
 Ser Asn Thr Tyr Tyr Leu Arg Leu Ala Val Glu Asp Phe Tyr Thr Cys
 450 455 460
 Gln Ser Ile Tyr Arg Glu Glu Leu Lys Gly Leu Glu Arg Trp Val Val
 465 470 475 480
 Gln Asn Lys Leu Asp Gln Leu Lys Phe Ala Arg Gln Lys Thr Ala Tyr
 485 490 495
 Cys Tyr Phe Ser Val Ala Ala Thr Leu Ser Ser Pro Glu Leu Ser Asp
 500 505 510
 Ala Arg Ile Ser Trp Ala Lys Asn Gly Ile Leu Thr Thr Val Val Asp
 515 520 525
 Asp Phe Phe Asp Ile Gly Gly Thr Ile Asp Glu Leu Thr Asn Leu Ile

530 535 540

Gln Cys Val Glu Lys Trp Asn Val Asp Val Asp Lys Asp Cys Cys Ser
545 550 555 560

Glu His Val Arg Ile Leu Phe Leu Ala Leu Lys Asp Ala Ile Cys Trp
565 570 575

Ile Gly Asp Glu Ala Phe Lys Trp Gln Ala Arg Asp Val Thr Ser His
580 585 590

Val Ile Gln Thr Trp Leu Glu Leu Met Asn Ser Met Leu Arg Glu Ala
595 600 605

Ile Trp Thr Arg Asp Ala Tyr Val Pro Thr Leu Asn Glu Tyr Met Glu
610 615 620

Asn Ala Tyr Val Ser Phe Ala Leu Gly Pro Ile Val Lys Pro Ala Ile
625 630 635 640

Tyr Phe Val Gly Pro Lys Leu Ser Glu Glu Ile Val Glu Ser Ser Glu
645 650 655

Tyr His Asn Leu Phe Lys Leu Met Ser Thr Gln Gly Arg Leu Leu Asn
660 665 670

Asp Ile His Ser Phe Lys Arg Glu Phe Lys Glu Gly Lys Leu Asn Ala
675 680 685

Val Ala Leu His Leu Ser Asn Gly Glu Ser Gly Lys Val Glu Glu Glu
690 695 700

Val Val Glu Glu Met Met Met Met Ile Lys Asn Lys Arg Lys Glu Leu
705 710 715 720

Met Lys Leu Ile Phe Glu Glu Asn Gly Ser Ile Val Pro Arg Ala Cys
725 730 735

Lys Asp Ala Phe Trp Asn Met Cys His Val Leu Asn Phe Phe Tyr Ala
740 745 750

Asn Asp Asp Gly Phe Thr Gly Asn Thr Ile Leu Asp Thr Val Lys Asp
755 760 765

Ile Ile Tyr Asn Pro Leu Val Leu Val Asn Glu Asn Glu Glu Gln Arg
770 775 780

<210> 134
<211> 590
<212> PRT
<213> Zea mays
<400> 134

Met Ala Met Pro Val Lys Leu Thr Pro Ala Ser Leu Ser Leu Lys Ala
 1 5 10 15
 Val Cys Cys Arg Phe Ser Ser Gly Gly His Ala Leu Arg Phe Gly Ser
 20 25 30
 Ser Leu Pro Cys Trp Arg Arg Thr Pro Thr Gln Arg Ser Thr Ser Ser
 35 40 45
 Ser Thr Thr Arg Pro Ala Ala Glu Val Ser Ser Gly Lys Ser Lys Gln
 50 55 60
 His Asp Gln Glu Ala Ser Glu Ala Thr Ile Arg Gln Gln Leu Gln Leu
 65 70 75 80
 Val Asp Val Leu Glu Asn Met Gly Ile Ser Arg His Phe Ala Ala Glu
 85 90 95
 Ile Lys Cys Ile Leu Asp Arg Thr Tyr Arg Ser Trp Leu Gln Arg His
 100 105 110
 Glu Glu Ile Met Leu Asp Thr Met Thr Cys Ala Met Ala Phe Arg Ile
 115 120 125
 Leu Arg Leu Asn Gly Tyr Asn Val Ser Ser Asp Glu Leu Tyr His Val
 130 135 140
 Val Glu Ala Ser Gly Leu His Asn Ser Leu Gly Gly Tyr Leu Asn Asp
 145 150 155 160
 Thr Arg Thr Leu Leu Glu Leu His Lys Ala Ser Thr Val Ser Ile Ser
 165 170 175
 Glu Asp Glu Ser Ile Leu Asp Ser Ile Gly Ser Arg Ser Arg Thr Leu
 180 185 190
 Leu Arg Glu Gln Leu Glu Ser Gly Gly Ala Leu Arg Lys Pro Ser Leu
 195 200 205
 Phe Lys Glu Val Glu His Ala Leu Asp Gly Pro Phe Tyr Thr Thr Leu
 210 215 220
 Asp Arg Leu His His Arg Trp Asn Ile Glu Asn Phe Asn Ile Ile Glu
 225 230 235 240
 Gln His Met Leu Glu Thr Pro Tyr Leu Ser Asn Gln His Thr Ser Arg
 245 250 255
 Asp Ile Leu Ala Leu Ser Ile Arg Asp Phe Ser Ser Ser Gln Phe Thr
 260 265 270
 Tyr Gln Gln Glu Leu Gln His Leu Glu Ser Trp Val Lys Glu Cys Arg

275 280 285
 Leu Asp Gln Leu Gln Phe Ala Arg Gln Lys Leu Ala Tyr Phe Tyr Leu
 290 295 300
 Ser Ala Ala Gly Thr Met Phe Ser Pro Glu Leu Ser Asp Ala Arg Thr
 305 310 315 320
 Leu Trp Ala Lys Asn Gly Val Leu Thr Thr Ile Val Asp Asp Phe Phe
 325 330 335
 Asp Val Ala Gly Ser Lys Glu Glu Leu Glu Asn Leu Val Met Leu Val
 340 345 350
 Glu Met Trp Asp Glu His His Lys Val Glu Phe Tyr Ser Glu Gln Val
 355 360 365
 Glu Ile Ile Phe Ser Ser Ile Tyr Asp Ser Val Asn Gln Leu Gly Glu
 370 375 380
 Lys Ala Ser Leu Val Gln Asp Arg Ser Ile Thr Lys His Leu Val Glu
 385 390 395 400
 Ile Trp Leu Asp Leu Leu Lys Ser Met Met Thr Glu Val Glu Trp Arg
 405 410 415
 Leu Ser Lys Tyr Val Pro Thr Glu Lys Glu Tyr Met Ile Asn Ala Ser
 420 425 430
 Leu Ile Phe Gly Leu Gly Pro Ile Val Leu Pro Ala Leu Tyr Phe Val
 435 440 445
 Gly Pro Lys Ile Ser Glu Ser Ile Val Lys Asp Pro Glu Tyr Asp Glu
 450 455 460
 Leu Phe Lys Leu Met Ser Thr Cys Gly Arg Leu Leu Asn Asp Val Gln
 465 470 475 480
 Thr Phe Glu Arg Glu Tyr Asn Glu Gly Lys Leu Asn Ser Val Ser Leu
 485 490 495
 Leu Val Leu His Gly Gly Pro Met Ser Ile Ser Asp Ala Lys Arg Lys
 500 505 510
 Leu Gln Lys Pro Ile Asp Thr Cys Arg Arg Asp Leu Leu Ser Leu Val
 515 520 525
 Leu Arg Glu Glu Ser Val Val Pro Arg Pro Cys Lys Glu Leu Phe Trp
 530 535 540
 Lys Met Cys Lys Val Cys Tyr Phe Phe Tyr Ser Thr Thr Asp Gly Phe
 545 550 555 560

Ser Ser Gln Val Glu Arg Ala Lys Glu Val Asp Ala Val Ile Asn Glu
565 570 575

Pro Leu Lys Leu Gln Gly Ser His Thr Leu Val Ser Asp Val
580 585 590

<210> 135

<211> 743

<212> PRT

<213> Populus trichocarpa

<400> 135

Met Gln Asn Phe His Gly Thr Lys Glu Arg Ile Lys Lys Met Phe Asp
1 5 10 15

Lys Ile Glu Leu Ser Val Ser Ser Tyr Asp Thr Ala Trp Val Ala Met
20 25 30

Val Pro Ser Pro Asp Cys Pro Glu Thr Pro Cys Phe Pro Glu Cys Thr
35 40 45

Lys Trp Ile Leu Glu Asn Gln Leu Gly Asp Gly Ser Trp Ser Leu Pro
50 55 60

His Gly Asn Pro Leu Leu Val Lys Asp Ala Leu Ser Ser Thr Leu Ala
65 70 75 80

Cys Ile Leu Ala Leu Lys Arg Trp Gly Ile Gly Glu Glu Gln Ile Asn
85 90 95

Lys Gly Leu Arg Phe Ile Glu Leu Asn Ser Ala Ser Val Thr Asp Asn
100 105 110

Glu Gln His Lys Pro Ile Gly Phe Asp Ile Ile Phe Pro Gly Met Ile
115 120 125

Glu Tyr Ala Ile Asp Leu Asp Leu Asn Leu Pro Leu Lys Pro Thr Asp
130 135 140

Ile Asn Ser Met Leu His Arg Arg Ala Leu Glu Leu Thr Ser Gly Gly
145 150 155 160

Gly Lys Asn Leu Glu Gly Arg Arg Ala Tyr Leu Ala Tyr Val Ser Glu
165 170 175

Gly Ile Gly Lys Leu Gln Asp Trp Glu Met Ala Met Lys Tyr Gln Arg
180 185 190

Lys Asn Gly Ser Leu Phe Asn Ser Pro Ser Thr Thr Ala Ala Ala Phe
195 200 205

Ile His Ile Gln Asp Ala Glu Cys Leu His Tyr Ile Arg Ser Leu Leu

210
 Gln Lys Phe Gly Asn Ala Val Pro Thr Ile Tyr Pro Leu Asp Ile Tyr
 225 230 235 240
 Ala Arg Leu Ser Met Val Asp Ala Leu Glu Arg Leu Gly Ile Asp Arg
 245 250 255
 His Phe Arg Lys Glu Arg Lys Phe Val Leu Asp Glu Thr Tyr Arg Phe
 260 265 270
 Trp Leu Gln Gly Glu Glu Glu Ile Phe Ser Asp Asn Ala Thr Cys Ala
 275 280 285
 Leu Ala Phe Arg Ile Leu Arg Leu Asn Gly Tyr Asp Val Ser Leu Glu
 290 295 300
 Asp His Phe Ser Asn Ser Leu Gly Gly Tyr Leu Lys Asp Ser Gly Ala
 305 310 315 320
 Ala Leu Glu Leu Tyr Arg Ala Leu Gln Leu Ser Tyr Pro Asp Glu Ser
 325 330 335
 Leu Leu Glu Lys Gln Asn Ser Arg Thr Ser Tyr Phe Leu Lys Gln Gly
 340 345 350
 Leu Ser Asn Val Ser Leu Cys Gly Asp Arg Leu Arg Lys Asn Ile Ile
 355 360 365
 Gly Glu Val His Asp Ala Leu Asn Phe Ser Asp His Ala Asn Leu Gln
 370 375 380
 Arg Leu Ala Ile Arg Arg Arg Ile Lys His Tyr Ala Thr Asp Asp Thr
 385 390 395 400
 Arg Ile Leu Lys Thr Ser Tyr Arg Cys Ser Thr Ile Gly Asn Gln Asp
 405 410 415
 Phe Leu Lys Leu Ala Val Glu Asp Phe Asn Ile Cys Gln Ser Ile Gln
 420 425 430
 Arg Glu Glu Phe Lys His Ile Glu Arg Trp Val Val Glu Arg Arg Leu
 435 440 445
 Asp Lys Leu Lys Phe Ala Arg Gln Lys Glu Ala Tyr Cys Tyr Phe Ser
 450 455 460
 Ala Ala Ala Thr Leu Phe Ala Pro Glu Leu Ser Asp Ala Arg Met Ser
 465 470 475 480
 Trp Ala Lys Asn Gly Val Leu Thr Thr Val Val Asp Asp Phe Phe Asp
 485 490 495

Val Gly Gly Ser Glu Glu Glu Leu Val Asn Leu Ile Glu Leu Ile Glu
 500 505 510
 Arg Trp Asp Val Asn Gly Ser Ala Asp Phe Cys Ser Glu Glu Val Glu
 515 520 525
 Ile Ile Tyr Ser Ala Ile His Ser Thr Ile Ser Glu Ile Gly Asp Lys
 530 535 540
 Ser Phe Gly Trp Gln Gly Arg Asp Val Lys Ser Gln Val Ile Lys Ile
 545 550 555 560
 Trp Leu Asp Leu Leu Lys Ser Met Leu Thr Glu Ala Gln Trp Ser Ser
 565 570 575
 Asn Lys Ser Val Pro Thr Leu Asp Glu Tyr Met Thr Thr Ala His Val
 580 585 590
 Ser Phe Ala Leu Gly Pro Ile Val Leu Pro Ala Leu Tyr Phe Val Gly
 595 600 605
 Pro Lys Leu Ser Glu Glu Val Ala Gly His Pro Glu Leu Leu Asn Leu
 610 615 620
 Tyr Lys Val Thr Ser Thr Cys Gly Arg Leu Leu Asn Asp Trp Arg Ser
 625 630 635 640
 Phe Lys Arg Glu Ser Glu Glu Gly Lys Leu Asn Ala Ile Ser Leu Tyr
 645 650 655
 Met Ile His Ser Gly Gly Ala Ser Thr Glu Glu Glu Thr Ile Glu His
 660 665 670
 Phe Lys Gly Leu Ile Asp Ser Gln Arg Arg Gln Leu Leu Gln Leu Val
 675 680 685
 Leu Gln Glu Lys Asp Ser Ile Ile Pro Arg Pro Cys Lys Asp Leu Phe
 690 695 700
 Trp Asn Met Ile Lys Leu Leu His Thr Phe Tyr Met Lys Asp Asp Gly
 705 710 715 720
 Phe Thr Ser Asn Glu Met Arg Asn Val Val Lys Ala Ile Ile Asn Glu
 725 730 735
 Pro Ile Ser Leu Asp Glu Leu
 740

<210> 136
 <211> 983
 <212> PRT

<213> Phomopsis amygdali

<400> 136

Met Glu Phe Asp Glu Pro Leu Val Asp Glu Ala Arg Ser Leu Val Gln
 1 5 10 15

Arg Thr Leu Gln Asp Tyr Asp Asp Arg Tyr Gly Phe Gly Thr Met Ser
 20 25 30

Cys Ala Ala Tyr Asp Thr Ala Trp Val Ser Leu Val Thr Lys Thr Val
 35 40 45

Asp Gly Arg Lys Gln Trp Leu Phe Pro Glu Cys Phe Glu Phe Leu Leu
 50 55 60

Glu Thr Gln Ser Asp Ala Gly Gly Trp Glu Ile Gly Asn Ser Ala Pro
 65 70 75 80

Ile Asp Gly Ile Leu Asn Thr Ala Ala Ser Leu Leu Ala Leu Lys Arg
 85 90 95

His Val Gln Thr Glu Gln Ile Ile Gln Pro Gln His Asp His Lys Asp
 100 105 110

Leu Ala Gly Arg Ala Glu Arg Ala Ala Ala Ser Leu Arg Ala Gln Leu
 115 120 125

Ala Ala Leu Asp Val Ser Thr Thr Glu His Val Gly Phe Glu Ile Ile
 130 135 140

Val Pro Ala Met Leu Asp Pro Leu Glu Ala Glu Asp Pro Ser Leu Val
 145 150 155 160

Phe Asp Phe Pro Ala Arg Lys Pro Leu Met Lys Ile His Asp Ala Lys
 165 170 175

Met Ser Arg Phe Arg Pro Glu Tyr Leu Tyr Gly Lys Gln Pro Met Thr
 180 185 190

Ala Leu His Ser Leu Glu Ala Phe Ile Gly Lys Ile Asp Phe Asp Lys
 195 200 205

Val Arg His His Arg Thr His Gly Ser Met Met Gly Ser Pro Ser Ser
 210 215 220

Thr Ala Ala Tyr Leu Met His Ala Ser Gln Trp Asp Gly Asp Ser Glu
 225 230 235 240

Ala Tyr Leu Arg His Val Ile Lys His Ala Ala Gly Gln Gly Thr Gly
 245 250 255

Ala Val Pro Ser Ala Phe Pro Ser Thr His Phe Glu Ser Ser Trp Ile
 260 265 270

Leu Thr Thr Leu Phe Arg Ala Gly Phe Ser Ala Ser His Leu Ala Cys
 275 280 285
 Asp Glu Leu Asn Lys Leu Val Glu Ile Leu Glu Gly Ser Phe Glu Lys
 290 295 300
 Glu Gly Gly Ala Ile Gly Tyr Ala Pro Gly Phe Gln Ala Asp Val Asp
 305 310 315 320
 Asp Thr Ala Lys Thr Ile Ser Thr Leu Ala Val Leu Gly Arg Asp Ala
 325 330 335
 Thr Pro Arg Gln Met Ile Lys Val Phe Glu Ala Asn Thr His Phe Arg
 340 345 350
 Thr Tyr Pro Gly Glu Arg Asp Pro Ser Leu Thr Ala Asn Cys Asn Ala
 355 360 365
 Leu Ser Ala Leu Leu His Gln Pro Asp Ala Ala Met Tyr Gly Ser Gln
 370 375 380
 Ile Gln Lys Ile Thr Lys Phe Val Cys Asp Tyr Trp Trp Lys Ser Asp
 385 390 395 400
 Gly Lys Ile Lys Asp Lys Trp Asn Thr Cys Tyr Leu Tyr Pro Ser Val
 405 410 415
 Leu Leu Val Glu Val Leu Val Asp Leu Val Ser Leu Leu Glu Gln Gly
 420 425 430
 Lys Leu Pro Asp Val Leu Asp Gln Glu Leu Gln Tyr Arg Val Ala Ile
 435 440 445
 Thr Leu Phe Gln Ala Cys Leu Arg Pro Leu Leu Asp Gln Asp Ala Glu
 450 455 460
 Gly Ser Trp Asn Lys Ser Ile Glu Ala Thr Ala Tyr Gly Ile Leu Ile
 465 470 475 480
 Leu Thr Glu Ala Arg Arg Val Cys Phe Phe Asp Arg Leu Ser Glu Pro
 485 490 495
 Leu Asn Glu Ala Ile Arg Arg Gly Ile Ala Phe Ala Asp Ser Met Ser
 500 505 510
 Gly Thr Glu Ala Gln Leu Asn Tyr Ile Trp Ile Glu Lys Val Ser Tyr
 515 520 525
 Ala Pro Ala Leu Leu Thr Lys Ser Tyr Leu Leu Ala Ala Arg Trp Ala
 530 535 540

Ala Lys Ser Pro Leu Gly Ala Ser Val Gly Ser Ser Leu Trp Thr Pro
 545 550 555 560
 Pro Arg Glu Gly Leu Asp Lys His Val Arg Leu Phe His Gln Ala Glu
 565 570 575
 Leu Phe Arg Ser Leu Pro Glu Trp Glu Leu Arg Ala Ser Met Ile Glu
 580 585 590
 Ala Ala Leu Phe Thr Pro Leu Leu Arg Ala His Arg Leu Asp Val Phe
 595 600 605
 Pro Arg Gln Asp Val Gly Glu Asp Lys Tyr Leu Asp Val Val Pro Phe
 610 615 620
 Phe Trp Thr Ala Ala Asn Asn Arg Asp Arg Thr Tyr Ala Ser Thr Leu
 625 630 635 640
 Phe Leu Tyr Asp Met Cys Phe Ile Ala Met Leu Asn Phe Gln Leu Asp
 645 650 655
 Glu Phe Met Glu Ala Thr Ala Gly Ile Leu Phe Arg Asp His Met Asp
 660 665 670
 Asp Leu Arg Gln Leu Ile His Asp Leu Leu Ala Glu Lys Thr Ser Pro
 675 680 685
 Lys Ser Ser Gly Arg Ser Ser Gln Gly Thr Lys Asp Ala Asp Ser Gly
 690 695 700
 Ile Glu Glu Asp Val Ser Met Ser Asp Ser Ala Ser Asp Ser Gln Asp
 705 710 715 720
 Arg Ser Pro Glu Tyr Asp Leu Val Phe Ser Ala Leu Ser Thr Phe Thr
 725 730 735
 Lys His Val Leu Gln His Pro Ser Ile Gln Ser Ala Ser Val Trp Asp
 740 745 750
 Arg Lys Leu Leu Ala Arg Glu Met Lys Ala Tyr Leu Leu Ala His Ile
 755 760 765
 Gln Gln Ala Glu Asp Ser Thr Pro Leu Ser Glu Leu Lys Asp Val Pro
 770 775 780
 Gln Lys Thr Asp Val Thr Arg Val Ser Thr Ser Thr Thr Thr Phe Phe
 785 790 795 800
 Asn Trp Val Arg Thr Thr Ser Ala Asp His Ile Ser Cys Pro Tyr Ser
 805 810 815

Phe His Phe Val Ala Cys His Leu Gly Ala Ala Leu Ser Pro Lys Gly
820 825 830

Ser Asn Gly Asp Cys Tyr Pro Ser Ala Gly Glu Lys Phe Leu Ala Ala
835 840 845

Ala Val Cys Arg His Leu Ala Thr Met Cys Arg Met Tyr Asn Asp Leu
850 855 860

Gly Ser Ala Glu Arg Asp Ser Asp Glu Gly Asn Leu Asn Ser Leu Asp
865 870 875 880

Phe Pro Glu Phe Ala Asp Ser Ala Gly Asn Gly Gly Ile Glu Ile Gln
885 890 895

Lys Ala Ala Leu Leu Arg Leu Ala Glu Phe Glu Arg Asp Ser Tyr Leu
900 905 910

Glu Ala Phe Arg Arg Leu Gln Asp Glu Ser Asn Arg Val His Gly Pro
915 920 925

Ala Gly Gly Asp Glu Ala Arg Leu Ser Arg Arg Arg Met Ala Ile Leu
930 935 940

Glu Phe Phe Ala Gln Gln Val Asp Leu Tyr Gly Gln Val Tyr Val Ile
945 950 955 960

Arg Asp Ile Ser Ala Arg Ile Pro Lys Asn Glu Val Glu Lys Lys Arg
965 970 975

Lys Leu Asp Asp Ala Phe Asn
980

<210> 137

<211> 881

<212> PRT

<213> Physcomitrella patens

<400> 137

Met Ala Ser Ser Thr Leu Ile Gln Asn Arg Ser Cys Gly Val Thr Ser
1 5 10 15

Ser Met Ser Ser Phe Gln Ile Phe Arg Gly Gln Pro Leu Arg Phe Pro
20 25 30

Gly Thr Arg Thr Pro Ala Ala Val Gln Cys Leu Lys Lys Arg Arg Cys
35 40 45

Leu Arg Pro Thr Glu Ser Val Leu Glu Ser Ser Pro Gly Ser Gly Ser
50 55 60

Tyr Arg Ile Val Thr Gly Pro Ser Gly Ile Asn Pro Ser Ser Asn Gly
65 70 75 80

His Leu Gln Glu Gly Ser Leu Thr His Arg Leu Pro Ile Pro Met Glu
 85 90 95
 Lys Ser Ile Asp Asn Phe Gln Ser Thr Leu Tyr Val Ser Asp Ile Trp
 100 105 110
 Ser Glu Thr Leu Gln Arg Thr Glu Cys Leu Leu Gln Val Thr Glu Asn
 115 120 125
 Val Gln Met Asn Glu Trp Ile Glu Glu Ile Arg Met Tyr Phe Arg Asn
 130 135 140
 Met Thr Leu Gly Glu Ile Ser Met Ser Pro Tyr Asp Thr Ala Trp Val
 145 150 155 160
 Ala Arg Val Pro Ala Leu Asp Gly Ser His Gly Pro Gln Phe His Arg
 165 170 175
 Ser Leu Gln Trp Ile Ile Asp Asn Gln Leu Pro Asp Gly Asp Trp Gly
 180 185 190
 Glu Pro Ser Leu Phe Leu Gly Tyr Asp Arg Val Cys Asn Thr Leu Ala
 195 200 205
 Cys Val Ile Ala Leu Lys Thr Trp Gly Val Gly Ala Gln Asn Val Glu
 210 215 220
 Arg Gly Ile Gln Phe Leu Gln Ser Asn Ile Tyr Lys Met Glu Glu Asp
 225 230 235 240
 Asp Ala Asn His Met Pro Ile Gly Phe Glu Ile Val Phe Pro Ala Met
 245 250 255
 Met Glu Asp Ala Lys Ala Leu Gly Leu Asp Leu Pro Tyr Asp Ala Thr
 260 265 270
 Ile Leu Gln Gln Ile Ser Ala Glu Arg Glu Lys Lys Met Lys Lys Ile
 275 280 285
 Pro Met Ala Met Val Tyr Lys Tyr Pro Thr Thr Leu Leu His Ser Leu
 290 295 300
 Glu Gly Leu His Arg Glu Val Asp Trp Asn Lys Leu Leu Gln Leu Gln
 305 310 315 320
 Ser Glu Asn Gly Ser Phe Leu Tyr Ser Pro Ala Ser Thr Ala Cys Ala
 325 330 335
 Leu Met Tyr Thr Lys Asp Val Lys Cys Phe Asp Tyr Leu Asn Gln Leu
 340 345 350

Leu Ile Lys Phe Asp His Ala Cys Pro Asn Val Tyr Pro Val Asp Leu
 355 360 365
 Phe Glu Arg Leu Trp Met Val Asp Arg Leu Gln Arg Leu Gly Ile Ser
 370 375 380
 Arg Tyr Phe Glu Arg Glu Ile Arg Asp Cys Leu Gln Tyr Val Tyr Arg
 385 390 395 400
 Tyr Trp Lys Asp Cys Gly Ile Gly Trp Ala Ser Asn Ser Ser Val Gln
 405 410 415
 Asp Val Asp Asp Thr Ala Met Ala Phe Arg Leu Leu Arg Thr His Gly
 420 425 430
 Phe Asp Val Lys Glu Asp Cys Phe Arg Gln Phe Phe Lys Asp Gly Glu
 435 440 445
 Phe Phe Cys Phe Ala Gly Gln Ser Ser Gln Ala Val Thr Gly Met Phe
 450 455 460
 Asn Leu Ser Arg Ala Ser Gln Thr Leu Phe Pro Gly Glu Ser Leu Leu
 465 470 475 480
 Lys Lys Ala Arg Thr Phe Ser Arg Asn Phe Leu Arg Thr Lys His Glu
 485 490 495
 Asn Asn Glu Cys Phe Asp Lys Trp Ile Ile Thr Lys Asp Leu Ala Gly
 500 505 510
 Glu Val Glu Tyr Asn Leu Thr Phe Pro Trp Tyr Ala Ser Leu Pro Arg
 515 520 525
 Leu Glu His Arg Thr Tyr Leu Asp Gln Tyr Gly Ile Asp Asp Ile Trp
 530 535 540
 Ile Gly Lys Ser Leu Tyr Lys Met Pro Ala Val Thr Asn Glu Val Phe
 545 550 555 560
 Leu Lys Leu Ala Lys Ala Asp Phe Asn Met Cys Gln Ala Leu His Lys
 565 570 575
 Lys Glu Leu Glu Gln Val Ile Lys Trp Asn Ala Ser Cys Gln Phe Arg
 580 585 590
 Asp Leu Glu Phe Ala Arg Gln Lys Ser Val Glu Cys Tyr Phe Ala Gly
 595 600 605
 Ala Ala Thr Met Phe Glu Pro Glu Met Val Gln Ala Arg Leu Val Trp
 610 615 620
 Ala Arg Cys Cys Val Leu Thr Thr Val Leu Asp Asp Tyr Phe Asp His

625 630 640

Gly Thr Pro Val Glu Glu Leu Arg Val Phe Val Gln Ala Val Arg Thr
645 650 655

Trp Asn Pro Glu Leu Ile Asn Gly Leu Pro Glu Gln Ala Lys Ile Leu
660 665 670

Phe Met Gly Leu Tyr Lys Thr Val Asn Thr Ile Ala Glu Glu Ala Phe
675 680 685

Met Ala Gln Lys Arg Asp Val His His His Leu Lys His Tyr Trp Asp
690 695 700

Lys Leu Ile Thr Ser Ala Leu Lys Glu Ala Glu Trp Ala Glu Ser Gly
705 710 715 720

Tyr Val Pro Thr Phe Asp Glu Tyr Met Glu Val Ala Glu Ile Ser Val
725 730 735

Ala Leu Glu Pro Ile Val Cys Ser Thr Leu Phe Phe Ala Gly His Arg
740 745 750

Leu Asp Glu Asp Val Leu Asp Ser Tyr Asp Tyr His Leu Val Met His
755 760 765

Leu Val Asn Arg Val Gly Arg Ile Leu Asn Asp Ile Gln Gly Met Lys
770 775 780

Arg Glu Ala Ser Gln Gly Lys Ile Ser Ser Val Gln Ile Tyr Met Glu
785 790 795 800

Glu His Pro Ser Val Pro Ser Glu Ala Met Ala Ile Ala His Leu Gln
805 810 815

Glu Leu Val Asp Asn Ser Met Gln Gln Leu Thr Tyr Glu Val Leu Arg
820 825 830

Phe Thr Ala Val Pro Lys Ser Cys Lys Arg Ile His Leu Asn Met Ala
835 840 845

Lys Ile Met His Ala Phe Tyr Lys Asp Thr Asp Gly Phe Ser Ser Leu
850 855 860

Thr Ala Met Thr Gly Phe Val Lys Lys Val Leu Phe Glu Pro Val Pro
865 870 875 880

Glu

<210> 138
<211> 513

<212> PRT

<213> Stevia rebaudiana

<400> 138

Met Asp Ala Val Thr Gly Leu Leu Thr Val Pro Ala Thr Ala Ile Thr
 1 5 10 15

Ile Gly Gly Thr Ala Val Ala Leu Ala Val Ala Leu Ile Phe Trp Tyr
 20 25 30

Leu Lys Ser Tyr Thr Ser Ala Arg Arg Ser Gln Ser Asn His Leu Pro
 35 40 45

Arg Val Pro Glu Val Pro Gly Val Pro Leu Leu Gly Asn Leu Leu Gln
 50 55 60

Leu Lys Glu Lys Lys Pro Tyr Met Thr Phe Thr Arg Trp Ala Ala Thr
 65 70 75 80

Tyr Gly Pro Ile Tyr Ser Ile Lys Thr Gly Ala Thr Ser Met Val Val
 85 90 95

Val Ser Ser Asn Glu Ile Ala Lys Glu Ala Leu Val Thr Arg Phe Gln
 100 105 110

Ser Ile Ser Thr Arg Asn Leu Ser Lys Ala Leu Lys Val Leu Thr Ala
 115 120 125

Asp Lys Thr Met Val Ala Met Ser Asp Tyr Asp Asp Tyr His Lys Thr
 130 135 140

Val Lys Arg His Ile Leu Thr Ala Val Leu Gly Pro Asn Ala Gln Lys
 145 150 155 160

Lys His Arg Ile His Arg Asp Ile Met Met Asp Asn Ile Ser Thr Gln
 165 170 175

Leu His Glu Phe Val Lys Asn Asn Pro Glu Gln Glu Glu Val Asp Leu
 180 185 190

Arg Lys Ile Phe Gln Ser Glu Leu Phe Gly Leu Ala Met Arg Gln Ala
 195 200 205

Leu Gly Lys Asp Val Glu Ser Leu Tyr Val Glu Asp Leu Lys Ile Thr
 210 215 220

Met Asn Arg Asp Glu Ile Phe Gln Val Leu Val Val Asp Pro Met Met
 225 230 235 240

Gly Ala Ile Asp Val Asp Trp Arg Asp Phe Phe Pro Tyr Leu Lys Trp
 245 250 255

Val Pro Asn Lys Lys Phe Glu Asn Thr Ile Gln Gln Met Tyr Ile Arg

260

265

270

Arg Glu Ala Val Met Lys Ser Leu Ile Lys Glu His Lys Lys Arg Ile
 275 280 285

Ala Ser Gly Glu Lys Leu Asn Ser Tyr Ile Asp Tyr Leu Leu Ser Glu
 290 295 300

Ala Gln Thr Leu Thr Asp Gln Gln Leu Leu Met Ser Leu Trp Glu Pro
 305 310 315 320

Ile Ile Glu Ser Ser Asp Thr Thr Met Val Thr Thr Glu Trp Ala Met
 325 330 335

Tyr Glu Leu Ala Lys Asn Pro Lys Leu Gln Asp Arg Leu Tyr Arg Asp
 340 345 350

Ile Lys Ser Val Cys Gly Ser Glu Lys Ile Thr Glu Glu His Leu Ser
 355 360 365

Gln Leu Pro Tyr Ile Thr Ala Ile Phe His Glu Thr Leu Arg Arg His
 370 375 380

Ser Pro Val Pro Ile Ile Pro Leu Arg His Val His Glu Asp Thr Val
 385 390 395 400

Leu Gly Gly Tyr His Val Pro Ala Gly Thr Glu Leu Ala Val Asn Ile
 405 410 415

Tyr Gly Cys Asn Met Asp Lys Asn Val Trp Glu Asn Pro Glu Glu Trp
 420 425 430

Asn Pro Glu Arg Phe Met Lys Glu Asn Glu Thr Ile Asp Phe Gln Lys
 435 440 445

Thr Met Ala Phe Gly Gly Gly Lys Arg Val Cys Ala Gly Ser Leu Gln
 450 455 460

Ala Leu Leu Thr Ala Ser Ile Gly Ile Gly Arg Met Val Gln Glu Phe
 465 470 475 480

Glu Trp Lys Leu Lys Asp Met Thr Gln Glu Glu Val Asn Thr Ile Gly
 485 490 495

Leu Thr Thr Gln Met Leu Arg Pro Leu Arg Ala Ile Ile Lys Pro Arg
 500 505 510

Ile

<210> 139
 <211> 509

<212> PRT

<213> Arabidopsis thaliana

<400> 139

Met Ala Phe Phe Ser Met Ile Ser Ile Leu Leu Gly Phe Val Ile Ser
 1 5 10 15

Ser Phe Ile Phe Ile Phe Phe Phe Lys Lys Leu Leu Ser Phe Ser Arg
 20 25 30

Lys Asn Met Ser Glu Val Ser Thr Leu Pro Ser Val Pro Val Val Pro
 35 40 45

Gly Phe Pro Val Ile Gly Asn Leu Leu Gln Leu Lys Glu Lys Lys Pro
 50 55 60

His Lys Thr Phe Thr Arg Trp Ser Glu Ile Tyr Gly Pro Ile Tyr Ser
 65 70 75 80

Ile Lys Met Gly Ser Ser Ser Leu Ile Val Leu Asn Ser Thr Glu Thr
 85 90 95

Ala Lys Glu Ala Met Val Thr Arg Phe Ser Ser Ile Ser Thr Arg Lys
 100 105 110

Leu Ser Asn Ala Leu Thr Val Leu Thr Cys Asp Lys Ser Met Val Ala
 115 120 125

Thr Ser Asp Tyr Asp Asp Phe His Lys Leu Val Lys Arg Cys Leu Leu
 130 135 140

Asn Gly Leu Leu Gly Ala Asn Ala Gln Lys Arg Lys Arg His Tyr Arg
 145 150 155 160

Asp Ala Leu Ile Glu Asn Val Ser Ser Lys Leu His Ala His Ala Arg
 165 170 175

Asp His Pro Gln Glu Pro Val Asn Phe Arg Ala Ile Phe Glu His Glu
 180 185 190

Leu Phe Gly Val Ala Leu Lys Gln Ala Phe Gly Lys Asp Val Glu Ser
 195 200 205

Ile Tyr Val Lys Glu Leu Gly Val Thr Leu Ser Lys Asp Glu Ile Phe
 210 215 220

Lys Val Leu Val His Asp Met Met Glu Gly Ala Ile Asp Val Asp Trp
 225 230 235 240

Arg Asp Phe Phe Pro Tyr Leu Lys Trp Ile Pro Asn Lys Ser Phe Glu
 245 250 255

Ala Arg Ile Gln Gln Lys His Lys Arg Arg Leu Ala Val Met Asn Ala

260

265

270

Leu Ile Gln Asp Arg Leu Lys Gln Asn Gly Ser Glu Ser Asp Asp Asp
 275 280 285

Cys Tyr Leu Asn Phe Leu Met Ser Glu Ala Lys Thr Leu Thr Lys Glu
 290 295 300

Gln Ile Ala Ile Leu Val Trp Glu Thr Ile Ile Glu Thr Ala Asp Thr
 305 310 315 320

Thr Leu Val Thr Thr Glu Trp Ala Ile Tyr Glu Leu Ala Lys His Pro
 325 330 335

Ser Val Gln Asp Arg Leu Cys Lys Glu Ile Gln Asn Val Cys Gly Gly
 340 345 350

Glu Lys Phe Lys Glu Glu Gln Leu Ser Gln Val Pro Tyr Leu Asn Gly
 355 360 365

Val Phe His Glu Thr Leu Arg Lys Tyr Ser Pro Ala Pro Leu Val Pro
 370 375 380

Ile Arg Tyr Ala His Glu Asp Thr Gln Ile Gly Gly Tyr His Val Pro
 385 390 395 400

Ala Gly Ser Glu Ile Ala Ile Asn Ile Tyr Gly Cys Asn Met Asp Lys
 405 410 415

Lys Arg Trp Glu Arg Pro Glu Asp Trp Trp Pro Glu Arg Phe Leu Asp
 420 425 430

Asp Gly Lys Tyr Glu Thr Ser Asp Leu His Lys Thr Met Ala Phe Gly
 435 440 445

Ala Gly Lys Arg Val Cys Ala Gly Ala Leu Gln Ala Ser Leu Met Ala
 450 455 460

Gly Ile Ala Ile Gly Arg Leu Val Gln Glu Phe Glu Trp Lys Leu Arg
 465 470 475 480

Asp Gly Glu Glu Glu Asn Val Asp Thr Tyr Gly Leu Thr Ser Gln Lys
 485 490 495

Leu Tyr Pro Leu Met Ala Ile Ile Asn Pro Arg Arg Ser
 500 505

<210> 140

<211> 525

<212> PRT

<213> Gibberella fujikoroi

<400> 140

Met Ser Lys Ser Asn Ser Met Asn Ser Thr Ser His Glu Thr Leu Phe
 1 5 10 15
 Gln Gln Leu Val Leu Gly Leu Asp Arg Met Pro Leu Met Asp Val His
 20 25 30
 Trp Leu Ile Tyr Val Ala Phe Gly Ala Trp Leu Cys Ser Tyr Val Ile
 35 40 45
 His Val Leu Ser Ser Ser Ser Thr Val Lys Val Pro Val Val Gly Tyr
 50 55 60
 Arg Ser Val Phe Glu Pro Thr Trp Leu Leu Arg Leu Arg Phe Val Trp
 65 70 75 80
 Glu Gly Gly Ser Ile Ile Gly Gln Gly Tyr Asn Lys Phe Lys Asp Ser
 85 90 95
 Ile Phe Gln Val Arg Lys Leu Gly Thr Asp Ile Val Ile Ile Pro Pro
 100 105 110
 Asn Tyr Ile Asp Glu Val Arg Lys Leu Ser Gln Asp Lys Thr Arg Ser
 115 120 125
 Val Glu Pro Phe Ile Asn Asp Phe Ala Gly Gln Tyr Thr Arg Gly Met
 130 135 140
 Val Phe Leu Gln Ser Asp Leu Gln Asn Arg Val Ile Gln Gln Arg Leu
 145 150 155 160
 Thr Pro Lys Leu Val Ser Leu Thr Lys Val Met Lys Glu Glu Leu Asp
 165 170 175
 Tyr Ala Leu Thr Lys Glu Met Pro Asp Met Lys Asn Asp Glu Trp Val
 180 185 190
 Glu Val Asp Ile Ser Ser Ile Met Val Arg Leu Ile Ser Arg Ile Ser
 195 200 205
 Ala Arg Val Phe Leu Gly Pro Glu His Cys Arg Asn Gln Glu Trp Leu
 210 215 220
 Thr Thr Thr Ala Glu Tyr Ser Glu Ser Leu Phe Ile Thr Gly Phe Ile
 225 230 235 240
 Leu Arg Val Val Pro His Ile Leu Arg Pro Phe Ile Ala Pro Leu Leu
 245 250 255
 Pro Ser Tyr Arg Thr Leu Leu Arg Asn Val Ser Ser Gly Arg Arg Val
 260 265 270
 Ile Gly Asp Ile Ile Arg Ser Gln Gln Gly Asp Gly Asn Glu Asp Ile

275 280 285
 Leu Ser Trp Met Arg Asp Ala Ala Thr Gly Glu Glu Lys Gln Ile Asp
 290 295 300
 Asn Ile Ala Gln Arg Met Leu Ile Leu Ser Leu Ala Ser Ile His Thr
 305 310 315 320
 Thr Ala Met Thr Met Thr His Ala Met Tyr Asp Leu Cys Ala Cys Pro
 325 330 335
 Glu Tyr Ile Glu Pro Leu Arg Asp Glu Val Lys Ser Val Val Gly Ala
 340 345 350
 Ser Gly Trp Asp Lys Thr Ala Leu Asn Arg Phe His Lys Leu Asp Ser
 355 360 365
 Phe Leu Lys Glu Ser Gln Arg Phe Asn Pro Val Phe Leu Leu Thr Phe
 370 375 380
 Asn Arg Ile Tyr His Gln Ser Met Thr Leu Ser Asp Gly Thr Asn Ile
 385 390 395 400
 Pro Ser Gly Thr Arg Ile Ala Val Pro Ser His Ala Met Leu Gln Asp
 405 410 415
 Ser Ala His Val Pro Gly Pro Thr Pro Pro Thr Glu Phe Asp Gly Phe
 420 425 430
 Arg Tyr Ser Lys Ile Arg Ser Asp Ser Asn Tyr Ala Gln Lys Tyr Leu
 435 440 445
 Phe Ser Met Thr Asp Ser Ser Asn Met Ala Phe Gly Tyr Gly Lys Tyr
 450 455 460
 Ala Cys Pro Gly Arg Phe Tyr Ala Ser Asn Glu Met Lys Leu Thr Leu
 465 470 475 480
 Ala Ile Leu Leu Leu Gln Phe Glu Phe Lys Leu Pro Asp Gly Lys Gly
 485 490 495
 Arg Pro Arg Asn Ile Thr Ile Asp Ser Asp Met Ile Pro Asp Pro Arg
 500 505 510
 Ala Arg Leu Cys Val Arg Lys Arg Ser Leu Arg Asp Glu
 515 520 525

<210> 141

<211> 499

<212> PRT

<213> Trametes versicolor

<400> 141

Met Glu Asp Pro Thr Val Leu Tyr Ala Cys Leu Ala Ile Ala Val Ala
 1 5 10 15
 Thr Phe Val Val Arg Trp Tyr Arg Asp Pro Leu Arg Ser Ile Pro Thr
 20 25 30
 Val Gly Gly Ser Asp Leu Pro Ile Leu Ser Tyr Ile Gly Ala Leu Arg
 35 40 45
 Trp Thr Arg Arg Gly Arg Glu Ile Leu Gln Glu Gly Tyr Asp Gly Tyr
 50 55 60
 Arg Gly Ser Thr Phe Lys Ile Ala Met Leu Asp Arg Trp Ile Val Ile
 65 70 75 80
 Ala Asn Gly Pro Lys Leu Ala Asp Glu Val Arg Arg Arg Pro Asp Glu
 85 90 95
 Glu Leu Asn Phe Met Asp Gly Leu Gly Ala Phe Val Gln Thr Lys Tyr
 100 105 110
 Thr Leu Gly Glu Ala Ile His Asn Asp Pro Tyr His Val Asp Ile Ile
 115 120 125
 Arg Glu Lys Leu Thr Arg Gly Leu Pro Ala Val Leu Pro Asp Val Ile
 130 135 140
 Glu Glu Leu Thr Leu Ala Val Arg Gln Tyr Ile Pro Thr Glu Gly Asp
 145 150 155 160
 Glu Trp Val Ser Val Asn Cys Ser Lys Ala Ala Arg Asp Ile Val Ala
 165 170 175
 Arg Ala Ser Asn Arg Val Phe Val Gly Leu Pro Ala Cys Arg Asn Gln
 180 185 190
 Gly Tyr Leu Asp Leu Ala Ile Asp Phe Thr Leu Ser Val Val Lys Asp
 195 200 205
 Arg Ala Ile Ile Asn Met Phe Pro Glu Leu Leu Lys Pro Ile Val Gly
 210 215 220
 Arg Val Val Gly Asn Ala Thr Arg Asn Val Arg Arg Ala Val Pro Phe
 225 230 235 240
 Val Ala Pro Leu Val Glu Glu Arg Arg Arg Leu Met Glu Glu Tyr Gly
 245 250 255
 Glu Asp Trp Ser Glu Lys Pro Asn Asp Met Leu Gln Trp Ile Met Asp
 260 265 270
 Glu Ala Ala Ser Arg Asp Ser Ser Val Lys Ala Ile Ala Glu Arg Leu

275 280 285
 Leu Met Val Asn Phe Ala Ala Ile His Thr Ser Ser Asn Thr Ile Thr
 290 295 300
 His Ala Leu Tyr His Leu Ala Glu Met Pro Glu Thr Leu Gln Pro Leu
 305 310 315 320
 Arg Glu Glu Ile Glu Pro Leu Val Lys Glu Glu Gly Trp Thr Lys Ala
 325 330 335
 Ala Met Gly Lys Met Trp Trp Leu Asp Ser Phe Leu Arg Glu Ser Gln
 340 345 350
 Arg Tyr Asn Gly Ile Asn Ile Val Ser Leu Thr Arg Met Ala Asp Lys
 355 360 365
 Asp Ile Thr Leu Ser Asp Gly Thr Phe Leu Pro Lys Gly Thr Leu Val
 370 375 380
 Ala Val Pro Ala Tyr Ser Thr His Arg Asp Asp Ala Val Tyr Ala Asp
 385 390 395 400
 Ala Leu Val Phe Asp Pro Phe Arg Phe Ser Arg Met Arg Ala Arg Glu
 405 410
 Gly Glu Gly Thr Lys His Gln Phe Val Asn Thr Ser Val Glu Tyr Val
 420 425 430
 Pro Phe Gly His Gly Lys His Ala Cys Pro Gly Arg Phe Phe Ala Ala
 435 440 445
 Asn Glu Leu Lys Ala Met Leu Ala Tyr Ile Val Leu Asn Tyr Asp Val
 450 455 460
 Lys Leu Pro Gly Asp Gly Lys Arg Pro Leu Asn Met Tyr Trp Gly Pro
 465 470 475 480
 Thr Val Leu Pro Ala Pro Ala Gly Gln Val Leu Phe Arg Lys Arg Gln
 485 490 495
 Val Ser Leu

<210> 142

<211> 525

<212> PRT

<213> Stevia rebaudiana

<400> 142

Met Gly Leu Phe Pro Leu Glu Asp Ser Tyr Ala Leu Val Phe Glu Gly
 1 5 10 15

Leu Ala Ile Thr Leu Ala Leu Tyr Tyr Leu Leu Ser Phe Ile Tyr Lys
 20 25 30
 Thr Ser Lys Lys Thr Cys Thr Pro Pro Lys Ala Ser Gly Glu Ile Ile
 35 40 45
 Pro Ile Thr Gly Ile Ile Leu Asn Leu Leu Ser Gly Ser Ser Gly Leu
 50 55 60
 Pro Ile Ile Leu Ala Leu Ala Ser Leu Ala Asp Arg Cys Gly Pro Ile
 65 70 75 80
 Phe Thr Ile Arg Leu Gly Ile Arg Arg Val Leu Val Val Ser Asn Trp
 85 90 95
 Glu Ile Ala Lys Glu Ile Phe Thr Thr His Asp Leu Ile Val Ser Asn
 100 105 110
 Arg Pro Lys Tyr Leu Ala Ala Lys Ile Leu Gly Phe Asn Tyr Val Ser
 115 120 125
 Phe Ser Phe Ala Pro Tyr Gly Pro Tyr Trp Val Gly Ile Arg Lys Ile
 130 135 140
 Ile Ala Thr Lys Leu Met Ser Ser Ser Arg Leu Gln Lys Leu Gln Phe
 145 150 155 160
 Val Arg Val Phe Glu Leu Glu Asn Ser Met Lys Ser Ile Arg Glu Ser
 165 170 175
 Trp Lys Glu Lys Lys Asp Glu Glu Gly Lys Val Leu Val Glu Met Lys
 180 185 190
 Lys Trp Phe Trp Glu Leu Asn Met Asn Ile Val Leu Arg Thr Val Ala
 195 200 205
 Gly Lys Gln Tyr Thr Gly Thr Val Asp Asp Ala Asp Ala Lys Arg Ile
 210 215 220
 Ser Glu Leu Phe Arg Glu Trp Phe His Tyr Thr Gly Arg Phe Val Val
 225 230 235 240
 Gly Asp Ala Phe Pro Phe Leu Gly Trp Leu Asp Leu Gly Gly Tyr Lys
 245 250 255
 Lys Thr Met Glu Leu Val Ala Ser Arg Leu Asp Ser Met Val Ser Lys
 260 265 270
 Trp Leu Asp Glu His Arg Lys Lys Gln Ala Asn Asp Asp Lys Lys Glu
 275 280 285
 Asp Met Asp Phe Met Asp Ile Met Ile Ser Met Thr Glu Ala Asn Ser

290

295

Pro Leu Glu Gly Tyr Gly Thr Asp Thr Ile Ile Lys Thr Thr Cys Met
305 310 315 320

Thr Leu Ile Val Ser Gly Val Asp Thr Thr Ser Ile Val Leu Thr Trp
325 330 335

Ala Leu Ser Leu Leu Asn Asn Arg Asp Thr Leu Lys Lys Ala Gln
340 345 350

Glu Glu Leu Asp Met Cys Val Gly Lys Gly Arg Gln Val Asn Glu Ser
355 360 365

Asp Leu Val Asn Leu Ile Tyr Leu Glu Ala Val Leu Lys Glu Ala Leu
370 375 380

Arg Leu Tyr Pro Ala Ala Phe Leu Gly Gly Pro Arg Ala Phe Leu Glu
385 390 395 400

Asp Cys Thr Val Ala Gly Tyr Arg Ile Pro Lys Gly Thr Cys Leu Leu
405 410 415

Ile Asn Met Trp Lys Leu His Arg Asp Pro Asn Ile Trp Ser Asp Pro
420 425 430

Cys Glu Phe Lys Pro Glu Arg Phe Leu Thr Pro Asn Gln Lys Asp Val
435 440 445

Asp Val Ile Gly Met Asp Phe Glu Leu Ile Pro Phe Gly Ala Gly Arg
450 455 460

Arg Tyr Cys Pro Gly Thr Arg Leu Ala Leu Gln Met Leu His Ile Val
465 470 475 480

Leu Ala Thr Leu Leu Gln Asn Phe Glu Met Ser Thr Pro Asn Asp Ala
485 490 495

Pro Val Asp Met Thr Ala Ser Val Gly Met Thr Asn Ala Lys Ala Ser
500 505 510

Pro Leu Glu Val Leu Leu Ser Pro Arg Val Lys Trp Ser
515 520 525

<210> 143
<211> 476
<212> PRT
<213> Stevia rebaudiana

<400> 143
Met Ile Gln Val Leu Thr Pro Ile Leu Leu Phe Leu Ile Phe Phe Val
1 5 10 15

Phe Trp Lys Val Tyr Lys His Gln Lys Thr Lys Ile Asn Leu Pro Pro
 20 25 30
 Gly Ser Phe Gly Trp Pro Phe Leu Gly Glu Thr Leu Ala Leu Leu Arg
 35 40 45
 Ala Gly Trp Asp Ser Glu Pro Glu Arg Phe Val Arg Glu Arg Ile Lys
 50 55 60
 Lys His Gly Ser Pro Leu Val Phe Lys Thr Ser Leu Phe Gly Asp Arg
 65 70 75 80
 Phe Ala Val Leu Cys Gly Pro Ala Gly Asn Lys Phe Leu Phe Cys Asn
 85 90 95
 Glu Asn Lys Leu Val Ala Ser Trp Trp Pro Val Pro Val Arg Lys Leu
 100 105 110
 Phe Gly Lys Ser Leu Leu Thr Ile Arg Gly Asp Glu Ala Lys Trp Met
 115 120 125
 Arg Lys Met Leu Leu Ser Tyr Leu Gly Pro Asp Ala Phe Ala Thr His
 130 135 140
 Tyr Ala Val Thr Met Asp Val Val Thr Arg Arg His Ile Asp Val His
 145 150 155 160
 Trp Arg Gly Lys Glu Glu Val Asn Val Phe Gln Thr Val Lys Leu Tyr
 165 170 175
 Ala Phe Glu Leu Ala Cys Arg Leu Phe Met Asn Leu Asp Asp Pro Asn
 180 185 190
 His Ile Ala Lys Leu Gly Ser Leu Phe Asn Ile Phe Leu Lys Gly Ile
 195 200 205
 Ile Glu Leu Pro Ile Asp Val Pro Gly Thr Arg Phe Tyr Ser Ser Lys
 210 215 220
 Lys Ala Ala Ala Ala Ile Arg Ile Glu Leu Lys Lys Leu Ile Lys Ala
 225 230 235 240
 Arg Lys Leu Glu Leu Lys Glu Gly Lys Ala Ser Ser Ser Gln Asp Leu
 245 250 255
 Leu Ser His Leu Leu Thr Ser Pro Asp Glu Asn Gly Met Phe Leu Thr
 260 265 270
 Glu Glu Glu Ile Val Asp Asn Ile Leu Leu Leu Leu Phe Ala Gly His
 275 280 285
 Asp Thr Ser Ala Leu Ser Ile Thr Leu Leu Met Lys Thr Leu Gly Glu

290 295 300

His Ser Asp Val Tyr Asp Lys Val Leu Lys Glu Gln Leu Glu Ile Ser
305 310 315 320

Lys Thr Lys Glu Ala Trp Glu Ser Leu Lys Trp Glu Asp Ile Gln Lys
325 330 335

Met Lys Tyr Ser Trp Ser Val Ile Cys Glu Val Met Arg Leu Asn Pro
340 345 350

Pro Val Ile Gly Thr Tyr Arg Glu Ala Leu Val Asp Ile Asp Tyr Ala
355 360 365

Gly Tyr Thr Ile Pro Lys Gly Trp Lys Leu His Trp Ser Ala Val Ser
370 375 380

Thr Gln Arg Asp Glu Ala Asn Phe Glu Asp Val Thr Arg Phe Asp Pro
385 390 395 400

Ser Arg Phe Glu Gly Ala Gly Pro Thr Pro Phe Thr Phe Val Pro Phe
405 410 415

Gly Gly Gly Pro Arg Met Cys Leu Gly Lys Glu Phe Ala Arg Leu Glu
420 425 430

Val Leu Ala Phe Leu His Asn Ile Val Thr Asn Phe Lys Trp Asp Leu
435 440 445

Leu Ile Pro Asp Glu Lys Ile Glu Tyr Asp Pro Met Ala Thr Pro Ala
450 455 460

Lys Gly Leu Pro Ile Arg Leu His Pro His Gln Val
465 470 475

<210> 144
<211> 525
<212> PRT
<213> Arabidopsis thaliana

<400> 144
Met Glu Ser Leu Val Val His Thr Val Asn Ala Ile Trp Cys Ile Val
1 5 10 15

Ile Val Gly Ile Phe Ser Val Gly Tyr His Val Tyr Gly Arg Ala Val
20 25 30

Val Glu Gln Trp Arg Met Arg Arg Ser Leu Lys Leu Gln Gly Val Lys
35 40 45

Gly Pro Pro Pro Ser Ile Phe Asn Gly Asn Val Ser Glu Met Gln Arg
50 55 60

Ile Gln Ser Glu Ala Lys His Cys Ser Gly Asp Asn Ile Ile Ser His
65 70 75 80

Asp Tyr Ser Ser Ser Leu Phe Pro His Phe Asp His Trp Arg Lys Gln
85 90 95

Tyr Gly Arg Ile Tyr Thr Tyr Ser Thr Gly Leu Lys Gln His Leu Tyr
100 105 110

Ile Asn His Pro Glu Met Val Lys Glu Leu Ser Gln Thr Asn Thr Leu
115 120 125

Asn Leu Gly Arg Ile Thr His Ile Thr Lys Arg Leu Asn Pro Ile Leu
130 135 140

Gly Asn Gly Ile Ile Thr Ser Asn Gly Pro His Trp Ala His Gln Arg
145 150 155 160

Arg Ile Ile Ala Tyr Glu Phe Thr His Asp Lys Ile Lys Gly Met Val
165 170 175

Gly Leu Met Val Glu Ser Ala Met Pro Met Leu Asn Lys Trp Glu Glu
180 185 190

Met Val Lys Arg Gly Gly Glu Met Gly Cys Asp Ile Arg Val Asp Glu
195 200 205

Asp Leu Lys Asp Val Ser Ala Asp Val Ile Ala Lys Ala Cys Phe Gly
210 215 220

Ser Ser Phe Ser Lys Gly Lys Ala Ile Phe Ser Met Ile Arg Asp Leu
225 230 235 240

Leu Thr Ala Ile Thr Lys Arg Ser Val Leu Phe Arg Phe Asn Gly Phe
245 250 255

Thr Asp Met Val Phe Gly Ser Lys Lys His Gly Asp Val Asp Ile Asp
260 265 270

Ala Leu Glu Met Glu Leu Glu Ser Ser Ile Trp Glu Thr Val Lys Glu
275 280 285

Arg Glu Ile Glu Cys Lys Asp Thr His Lys Lys Asp Leu Met Gln Leu
290 295 300

Ile Leu Glu Gly Ala Met Arg Ser Cys Asp Gly Asn Leu Trp Asp Lys
305 310 315 320

Ser Ala Tyr Arg Arg Phe Val Val Asp Asn Cys Lys Ser Ile Tyr Phe
325 330 335

Ala Gly His Asp Ser Thr Ala Val Ser Val Ser Trp Cys Leu Met Leu

340

345

350

Leu Ala Leu Asn Pro Ser Trp Gln Val Lys Ile Arg Asp Glu Ile Leu
 355 360 365

Ser Ser Cys Lys Asn Gly Ile Pro Asp Ala Glu Ser Ile Pro Asn Leu
 370 375 380

Lys Thr Val Thr Met Val Ile Gln Glu Thr Met Arg Leu Tyr Pro Pro
 385 390 395 400

Ala Pro Ile Val Gly Arg Glu Ala Ser Lys Asp Ile Arg Leu Gly Asp
 405 410 415

Leu Val Val Pro Lys Gly Val Cys Ile Trp Thr Leu Ile Pro Ala Leu
 420 425 430

His Arg Asp Pro Glu Ile Trp Gly Pro Asp Ala Asn Asp Phe Lys Pro
 435 440 445

Glu Arg Phe Ser Glu Gly Ile Ser Lys Ala Cys Lys Tyr Pro Gln Ser
 450 455 460

Tyr Ile Pro Phe Gly Leu Gly Pro Arg Thr Cys Val Gly Lys Asn Phe
 465 470 475 480

Gly Met Met Glu Val Lys Val Leu Val Ser Leu Ile Val Ser Lys Phe
 485 490 495

Ser Phe Thr Leu Ser Pro Thr Tyr Gln His Ser Pro Ser His Lys Leu
 500 505 510

Leu Val Glu Pro Gln His Gly Val Val Ile Arg Val Val
 515 520 525

<210> 145

<211> 529

<212> PRT

<213> Vitis vinifera

<400> 145

Met Tyr Phe Leu Leu Gln Tyr Leu Asn Ile Thr Thr Val Gly Val Phe
 1 5 10 15

Ala Thr Leu Phe Leu Ser Tyr Cys Leu Leu Leu Trp Arg Ser Arg Ala
 20 25 30

Gly Asn Lys Lys Ile Ala Pro Glu Ala Ala Ala Ala Trp Pro Ile Ile
 35 40 45

Gly His Leu His Leu Leu Ala Gly Gly Ser His Gln Leu Pro His Ile
 50 55 60

Thr Leu Gly Asn Met Ala Asp Lys Tyr Gly Pro Val Phe Thr Ile Arg
 65 70 75 80
 Ile Gly Leu His Arg Ala Val Val Val Ser Ser Trp Glu Met Ala Lys
 85 90 95
 Glu Cys Ser Thr Ala Asn Asp Gln Val Ser Ser Ser Arg Pro Glu Leu
 100 105 110
 Leu Ala Ser Lys Leu Leu Gly Tyr Asn Tyr Ala Met Phe Gly Phe Ser
 115 120 125
 Pro Tyr Gly Ser Tyr Trp Arg Glu Met Arg Lys Ile Ile Ser Leu Glu
 130 135 140
 Leu Leu Ser Asn Ser Arg Leu Glu Leu Leu Lys Asp Val Arg Ala Ser
 145 150 155 160
 Glu Val Val Thr Ser Ile Lys Glu Leu Tyr Lys Leu Trp Ala Glu Lys
 165 170 175
 Lys Asn Glu Ser Gly Leu Val Ser Val Glu Met Lys Gln Trp Phe Gly
 180 185 190
 Asp Leu Thr Leu Asn Val Ile Leu Arg Met Val Ala Gly Lys Arg Tyr
 195 200 205
 Phe Ser Ala Ser Asp Ala Ser Glu Asn Lys Gln Ala Gln Arg Cys Arg
 210 215 220
 Arg Val Phe Arg Glu Phe Phe His Leu Ser Gly Leu Phe Val Val Ala
 225 230 235 240
 Asp Ala Ile Pro Phe Leu Gly Trp Leu Asp Trp Gly Arg His Glu Lys
 245 250 255
 Thr Leu Lys Lys Thr Ala Ile Glu Met Asp Ser Ile Ala Gln Glu Trp
 260 265 270
 Leu Glu Glu His Arg Arg Arg Lys Asp Ser Gly Asp Asp Asn Ser Thr
 275 280 285
 Gln Asp Phe Met Asp Val Met Gln Ser Val Leu Asp Gly Lys Asn Leu
 290 295 300
 Gly Gly Tyr Asp Ala Asp Thr Ile Asn Lys Ala Thr Cys Leu Thr Leu
 305 310 315 320
 Ile Ser Gly Gly Ser Asp Thr Thr Val Val Ser Leu Thr Trp Ala Leu
 325 330 335
 Ser Leu Val Leu Asn Asn Arg Asp Thr Leu Lys Lys Ala Gln Glu Glu

340

345

350

Leu Asp Ile Gln Val Gly Lys Glu Arg Leu Val Asn Glu Gln Asp Ile
 355 360 365

Ser Lys Leu Val Tyr Leu Gln Ala Ile Val Lys Glu Thr Leu Arg Leu
 370 375 380

Tyr Pro Pro Gly Pro Leu Gly Gly Leu Arg Gln Phe Thr Glu Asp Cys
 385 390 395 400

Thr Leu Gly Gly Tyr His Val Ser Lys Gly Thr Arg Leu Ile Met Asn
 405 410 415

Leu Ser Lys Ile Gln Lys Asp Pro Arg Ile Trp Ser Asp Pro Thr Glu
 420 425 430

Phe Gln Pro Glu Arg Phe Leu Thr Thr His Lys Asp Val Asp Pro Arg
 435 440 445

Gly Lys His Phe Glu Phe Ile Pro Phe Gly Ala Gly Arg Arg Ala Cys
 450 455 460

Pro Gly Ile Thr Phe Gly Leu Gln Val Leu His Leu Thr Leu Ala Ser
 465 470 475 480

Phe Leu His Ala Phe Glu Phe Ser Thr Pro Ser Asn Glu Gln Val Asn
 485 490 495

Met Arg Glu Ser Leu Gly Leu Thr Asn Met Lys Ser Thr Pro Leu Glu
 500 505 510

Val Leu Ile Ser Pro Arg Leu Ser Leu Asn Cys Phe Asn Leu Met Lys
 515 520 525

Ile

<210> 146

<211> 479

<212> PRT

<213> *Medicago trunculata*

<400> 146

Met Glu Pro Asn Phe Tyr Leu Ser Leu Leu Leu Phe Val Thr Phe
 1 5 10 15

Ile Ser Leu Ser Leu Phe Phe Ile Phe Tyr Lys Gln Lys Ser Pro Leu
 20 25 30

Asn Leu Pro Pro Gly Lys Met Gly Tyr Pro Ile Ile Gly Glu Ser Leu
 35 40 45

Glu Phe Leu Ser Thr Gly Trp Lys Gly His Pro Glu Lys Phe Ile Phe
 50 55 60
 Asp Arg Met Arg Lys Tyr Ser Ser Glu Leu Phe Lys Thr Ser Ile Val
 65 70 75 80
 Gly Glu Ser Thr Val Val Cys Cys Gly Ala Ala Ser Asn Lys Phe Leu
 85 90 95
 Phe Ser Asn Glu Asn Lys Leu Val Thr Ala Trp Trp Pro Asp Ser Val
 100 105 110
 Asn Lys Ile Phe Pro Thr Thr Ser Leu Asp Ser Asn Leu Lys Glu Glu
 115 120 125
 Ser Ile Lys Met Arg Lys Leu Leu Pro Gln Phe Phe Lys Pro Glu Ala
 130 135 140
 Leu Gln Arg Tyr Val Gly Val Met Asp Val Ile Ala Gln Arg His Phe
 145 150 155 160
 Val Thr His Trp Asp Asn Lys Asn Glu Ile Thr Val Tyr Pro Leu Ala
 165 170 175
 Lys Arg Tyr Thr Phe Leu Leu Ala Cys Arg Leu Phe Met Ser Val Glu
 180 185 190
 Asp Glu Asn His Val Ala Lys Phe Ser Asp Pro Phe Gln Leu Ile Ala
 195 200 205
 Ala Gly Ile Ile Ser Leu Pro Ile Asp Leu Pro Gly Thr Pro Phe Asn
 210 215 220
 Lys Ala Ile Lys Ala Ser Asn Phe Ile Arg Lys Glu Leu Ile Lys Ile
 225 230 235 240
 Ile Lys Gln Arg Arg Val Asp Leu Ala Glu Gly Thr Ala Ser Pro Thr
 245 250 255
 Gln Asp Ile Leu Ser His Met Leu Leu Thr Ser Asp Glu Asn Gly Lys
 260 265 270
 Ser Met Asn Glu Leu Asn Ile Ala Asp Lys Ile Leu Gly Leu Leu Ile
 275 280 285
 Gly Gly His Asp Thr Ala Ser Val Ala Cys Thr Phe Leu Val Lys Tyr
 290 295 300
 Leu Gly Glu Leu Pro His Ile Tyr Asp Lys Val Tyr Gln Glu Gln Met
 305 310 315 320
 Glu Ile Ala Lys Ser Lys Pro Ala Gly Glu Leu Leu Asn Trp Asp Asp

Leu Lys Lys Met Lys Tyr Ser Trp Asn Val Ala Cys Glu Val Met Arg
340 345 350

Leu Ser Pro Pro Leu Gln Gly Gly Phe Arg Glu Ala Ile Thr Asp Phe
355 360 365

Met Phe Asn Gly Phe Ser Ile Pro Lys Gly Trp Lys Leu Tyr Trp Ser
370 375 380

Ala Asn Ser Thr His Lys Asn Ala Glu Cys Phe Pro Met Pro Glu Lys
385 390 395 400

Phe Asp Pro Thr Arg Phe Glu Gly Asn Gly Pro Ala Pro Tyr Thr Phe
405 410 415

Val Pro Phe Gly Gly Gly Pro Arg Met Cys Pro Gly Lys Glu Tyr Ala
420 425 430

Arg Leu Glu Ile Leu Val Phe Met His Asn Leu Val Lys Arg Phe Lys
435 440 445

Trp Glu Lys Val Ile Pro Asp Glu Lys Ile Ile Val Asp Pro Phe Pro
450 455 460

Ile Pro Ala Lys Asp Leu Pro Ile Arg Leu Tyr Pro His Lys Ala
465 470 475

<210> 147
<211> 710
<212> PRT
<213> Stevia rebaudiana

<400> 147
Met Gln Ser Asp Ser Val Lys Val Ser Pro Phe Asp Leu Val Ser Ala
1 5 10 15

Ala Met Asn Gly Lys Ala Met Glu Lys Leu Asn Ala Ser Glu Ser Glu
20 25 30

Asp Pro Thr Thr Leu Pro Ala Leu Lys Met Leu Val Glu Asn Arg Glu
35 40 45

Leu Leu Thr Leu Phe Thr Thr Ser Phe Ala Val Leu Ile Gly Cys Leu
50 55 60

Val Phe Leu Met Trp Arg Arg Ser Ser Ser Lys Lys Leu Val Gln Asp
65 70 75 80

Pro Val Pro Gln Val Ile Val Val Lys Lys Lys Glu Lys Glu Ser Glu
85 90 95

Val Asp Asp Gly Lys Lys Lys Val Ser Ile Phe Tyr Gly Thr Gln Thr
 100 105 110
 Gly Thr Ala Glu Gly Phe Ala Lys Ala Leu Val Glu Glu Ala Lys Val
 115 120 125
 Arg Tyr Glu Lys Thr Ser Phe Lys Val Ile Asp Leu Asp Asp Tyr Ala
 130 135 140
 Ala Asp Asp Asp Glu Tyr Glu Glu Lys Leu Lys Lys Glu Ser Leu Ala
 145 150 155 160
 Phe Phe Phe Leu Ala Thr Tyr Gly Asp Gly Glu Pro Thr Asp Asn Ala
 165 170 175
 Ala Asn Phe Tyr Lys Trp Phe Thr Glu Gly Asp Asp Lys Gly Glu Trp
 180 185 190
 Leu Lys Lys Leu Gln Tyr Gly Val Phe Gly Leu Gly Asn Arg Gln Tyr
 195 200 205
 Glu His Phe Asn Lys Ile Ala Ile Val Val Asp Asp Lys Leu Thr Glu
 210 215 220
 Met Gly Ala Lys Arg Leu Val Pro Val Gly Leu Gly Asp Asp Asp Gln
 225 230 235 240
 Cys Ile Glu Asp Asp Phe Thr Ala Trp Lys Glu Leu Val Trp Pro Glu
 245 250 255
 Leu Asp Gln Leu Leu Arg Asp Glu Asp Asp Thr Ser Val Thr Thr Pro
 260 265 270
 Tyr Thr Ala Ala Val Leu Glu Tyr Arg Val Val Tyr His Asp Lys Pro
 275 280 285
 Ala Asp Ser Tyr Ala Glu Asp Gln Thr His Thr Asn Gly His Val Val
 290 295 300
 His Asp Ala Gln His Pro Ser Arg Ser Asn Val Ala Phe Lys Lys Glu
 305 310 315 320
 Leu His Thr Ser Gln Ser Asp Arg Ser Cys Thr His Leu Glu Phe Asp
 325 330 335
 Ile Ser His Thr Gly Leu Ser Tyr Glu Thr Gly Asp His Val Gly Val
 340 345 350
 Tyr Ser Glu Asn Leu Ser Glu Val Val Asp Glu Ala Leu Lys Leu Leu
 355 360 365
 Gly Leu Ser Pro Asp Thr Tyr Phe Ser Val His Ala Asp Lys Glu Asp

370
 Gly Thr Pro Ile Gly Gly Ala Ser Leu Pro Pro Phe Pro Pro Cys
 385 390 395 400
 Thr Leu Arg Asp Ala Leu Thr Arg Tyr Ala Asp Val Leu Ser Ser Pro
 405 410 415
 Lys Lys Val Ala Leu Leu Ala Leu Ala His Ala Ser Asp Pro Ser
 420 425 430
 Glu Ala Asp Arg Leu Lys Phe Leu Ala Ser Pro Ala Gly Lys Asp Glu
 435 440 445
 Tyr Ala Gln Trp Ile Val Ala Asn Gln Arg Ser Leu Leu Glu Val Met
 450 455 460
 Gln Ser Phe Pro Ser Ala Lys Pro Pro Leu Gly Val Phe Phe Ala Ala
 465 470 475 480
 Val Ala Pro Arg Leu Gln Pro Arg Tyr Tyr Ser Ile Ser Ser Ser Pro
 485 490 495
 Lys Met Ser Pro Asn Arg Ile His Val Thr Cys Ala Leu Val Tyr Glu
 500 505 510
 Thr Thr Pro Ala Gly Arg Ile His Arg Gly Leu Cys Ser Thr Trp Met
 515 520 525
 Lys Asn Ala Val Pro Leu Thr Glu Ser Pro Asp Cys Ser Gln Ala Ser
 530 535 540
 Ile Phe Val Arg Thr Ser Asn Phe Arg Leu Pro Val Asp Pro Lys Val
 545 550 555 560
 Pro Val Ile Met Ile Gly Pro Gly Thr Gly Leu Ala Pro Phe Arg Gly
 565 570 575
 Phe Leu Gln Glu Arg Leu Ala Leu Lys Glu Ser Gly Thr Glu Leu Gly
 580 585 590
 Ser Ser Ile Phe Phe Phe Gly Cys Arg Asn Arg Lys Val Asp Phe Ile
 595 600 605
 Tyr Glu Asp Glu Leu Asn Asn Phe Val Glu Thr Gly Ala Leu Ser Glu
 610 615 620
 Leu Ile Val Ala Phe Ser Arg Glu Gly Thr Ala Lys Glu Tyr Val Gln
 625 630 635 640
 His Lys Met Ser Gln Lys Ala Ser Asp Ile Trp Lys Leu Leu Ser Glu
 645 650 655

Gly Ala Tyr Leu Tyr Val Cys Gly Asp Ala Lys Gly Met Ala Lys Asp
660 665 670

Val His Arg Thr Leu His Thr Ile Val Gln Glu Gln Gly Ser Leu Asp
675 680 685

Ser Ser Lys Ala Glu Leu Tyr Val Lys Asn Leu Gln Met Ser Gly Arg
690 695 700

Tyr Leu Arg Asp Val Trp
705 710

<210> 148

<211> 692

<212> PRT

<213> Arabidopsis thaliana

<400> 148

Met Thr Ser Ala Leu Tyr Ala Ser Asp Leu Phe Lys Gln Leu Lys Ser
1 5 10 15

Ile Met Gly Thr Asp Ser Leu Ser Asp Asp Val Val Leu Val Ile Ala
20 25 30

Thr Thr Ser Leu Ala Leu Val Ala Gly Phe Val Val Leu Leu Trp Lys
35 40 45

Lys Thr Thr Ala Asp Arg Ser Gly Glu Leu Lys Pro Leu Met Ile Pro
50 55 60

Lys Ser Leu Met Ala Lys Asp Glu Asp Asp Asp Leu Asp Leu Gly Ser
65 70 75 80

Gly Lys Thr Arg Val Ser Ile Phe Phe Gly Thr Gln Thr Gly Thr Ala
85 90 95

Glu Gly Phe Ala Lys Ala Leu Ser Glu Glu Ile Lys Ala Arg Tyr Glu
100 105 110

Lys Ala Ala Val Lys Val Ile Asp Leu Asp Asp Tyr Ala Ala Asp Asp
115 120 125

Asp Gln Tyr Glu Glu Lys Leu Lys Lys Glu Thr Leu Ala Phe Phe Cys
130 135 140

Val Ala Thr Tyr Gly Asp Gly Glu Pro Thr Asp Asn Ala Ala Arg Phe
145 150 155 160

Tyr Lys Trp Phe Thr Glu Glu Asn Glu Arg Asp Ile Lys Leu Gln Gln
165 170 175

Leu Ala Tyr Gly Val Phe Ala Leu Gly Asn Arg Gln Tyr Glu His Phe

180	185	190
Asn Lys Ile Gly Ile Val Leu Asp Glu Glu Leu Cys Lys Lys Gly Ala	195	200
Lys Arg Leu Ile Glu Val Gly Leu Gly Asp Asp Asp Gln Ser Ile Glu	210	220
Asp Asp Phe Asn Ala Trp Lys Glu Ser Leu Trp Ser Glu Leu Asp Lys	225	235
Leu Leu Lys Asp Glu Asp Asp Lys Ser Val Ala Thr Pro Tyr Thr Ala	245	255
Val Ile Pro Glu Tyr Arg Val Val Thr His Asp Pro Arg Phe Thr Thr	260	270
Gln Lys Ser Met Glu Ser Asn Val Ala Asn Gly Asn Thr Thr Ile Asp	275	285
Ile His His Pro Cys Arg Val Asp Val Ala Val Gln Lys Glu Leu His	290	300
Thr His Glu Ser Asp Arg Ser Cys Ile His Leu Glu Phe Asp Ile Ser	305	315
Arg Thr Gly Ile Thr Tyr Glu Thr Gly Asp His Val Gly Val Tyr Ala	325	335
Glu Asn His Val Glu Ile Val Glu Glu Ala Gly Lys Leu Leu Gly His	340	350
Ser Leu Asp Leu Val Phe Ser Ile His Ala Asp Lys Glu Asp Gly Ser	355	365
Pro Leu Glu Ser Ala Val Pro Pro Pro Phe Pro Gly Pro Cys Thr Leu	370	380
Gly Thr Gly Leu Ala Arg Tyr Ala Asp Leu Leu Asn Pro Pro Arg Lys	385	395
Ser Ala Leu Val Ala Leu Ala Ala Tyr Ala Thr Glu Pro Ser Glu Ala	405	415
Glu Lys Leu Lys His Leu Thr Ser Pro Asp Gly Lys Asp Glu Tyr Ser	420	430
Gln Trp Ile Val Ala Ser Gln Arg Ser Leu Leu Glu Val Met Ala Ala	435	445
Phe Pro Ser Ala Lys Pro Pro Leu Gly Val Phe Phe Ala Ala Ile Ala	450	460

Pro Arg Leu Gln Pro Arg Tyr Tyr Ser Ile Ser Ser Ser Pro Arg Leu
465 470 475 480

Ala Pro Ser Arg Val His Val Thr Ser Ala Leu Val Tyr Gly Pro Thr
485 490 495

Pro Thr Gly Arg Ile His Lys Gly Val Cys Ser Thr Trp Met Lys Asn
500 505 510

Ala Val Pro Ala Glu Lys Ser His Glu Cys Ser Gly Ala Pro Ile Phe
515 520 525

Ile Arg Ala Ser Asn Phe Lys Leu Pro Ser Asn Pro Ser Thr Pro Ile
530 535 540

Val Met Val Gly Pro Gly Thr Gly Leu Ala Pro Phe Arg Gly Phe Leu
545 550 555 560

Gln Glu Arg Met Ala Leu Lys Glu Asp Gly Glu Glu Leu Gly Ser Ser
565 570 575

Leu Leu Phe Phe Gly Cys Arg Asn Arg Gln Met Asp Phe Ile Tyr Glu
580 585 590

Asp Glu Leu Asn Asn Phe Val Asp Gln Gly Val Ile Ser Glu Leu Ile
595 600 605

Met Ala Phe Ser Arg Glu Gly Ala Gln Lys Glu Tyr Val Gln His Lys
610 615 620

Met Met Glu Lys Ala Ala Gln Val Trp Asp Leu Ile Lys Glu Glu Gly
625 630 635 640

Tyr Leu Tyr Val Cys Gly Asp Ala Lys Gly Met Ala Arg Asp Val His
645 650 655

Arg Thr Leu His Thr Ile Val Gln Glu Gln Glu Gly Val Ser Ser Ser
660 665 670

Glu Ala Glu Ala Ile Val Lys Lys Leu Gln Thr Glu Gly Arg Tyr Leu
675 680 685

Arg Asp Val Trp
690

<210> 149
<211> 713
<212> PRT
<213> Giberella fujikuroi

<400> 149
Met Ala Glu Leu Asp Thr Leu Asp Ile Val Val Leu Gly Val Ile Phe

5

15

20

35

50

65

85

100

115

130

145

165

180

195

210

225

245

260

275

280

285

Ala Lys Asp Arg Asn Cys Leu His Met Glu Ile Asp Ile Ser Gly Ser
 290 295 300
 Asn Leu Lys Tyr Glu Thr Gly Asp His Ile Ala Ile Trp Pro Thr Asn
 305 310 315 320
 Pro Gly Glu Glu Val Asn Lys Phe Leu Asp Ile Leu Asp Leu Ser Gly
 325 330 335
 Lys Gln His Ser Val Val Thr Val Lys Ala Leu Glu Pro Thr Ala Lys
 340 345 350
 Val Pro Phe Pro Asn Pro Thr Thr Tyr Asp Ala Ile Leu Arg Tyr His
 355 360 365
 Leu Glu Ile Cys Ala Pro Val Ser Arg Gln Phe Val Ser Thr Leu Ala
 370 375 380
 Ala Phe Ala Pro Asn Asp Asp Ile Lys Ala Glu Met Asn Arg Leu Gly
 385 390 395 400
 Ser Asp Lys Asp Tyr Phe His Glu Lys Thr Gly Pro His Tyr Tyr Asn
 405 410 415
 Ile Ala Arg Phe Leu Ala Ser Val Ser Lys Gly Glu Lys Trp Thr Lys
 420 425 430
 Ile Pro Phe Ser Ala Phe Ile Glu Gly Leu Thr Lys Leu Gln Pro Arg
 435 440 445
 Tyr Tyr Ser Ile Ser Ser Ser Ser Leu Val Gln Pro Lys Lys Ile Ser
 450 455 460
 Ile Thr Ala Val Val Glu Ser Gln Gln Ile Pro Gly Arg Asp Asp Pro
 465 470 475 480
 Phe Arg Gly Val Ala Thr Asn Tyr Leu Phe Ala Leu Lys Gln Lys Gln
 485 490 495
 Asn Gly Asp Pro Asn Pro Ala Pro Phe Gly Gln Ser Tyr Glu Leu Thr
 500 505 510
 Gly Pro Arg Asn Lys Tyr Asp Gly Ile His Val Pro Val His Val Arg
 515 520 525
 His Ser Asn Phe Lys Leu Pro Ser Asp Pro Gly Lys Pro Ile Ile Met
 530 535 540
 Ile Gly Pro Gly Thr Gly Val Ala Pro Phe Arg Gly Phe Val Gln Glu
 545 550 555 560

Arg Ala Lys Gln Ala Arg Asp Gly Val Glu Val Gly Lys Thr Leu Leu
565 570 575

Phe Phe Gly Cys Arg Lys Ser Thr Glu Asp Phe Met Tyr Gln Lys Glu
580 585 590

Trp Gln Glu Tyr Lys Glu Ala Leu Gly Asp Lys Phe Glu Met Ile Thr
595 600 605

Ala Phe Ser Arg Glu Gly Ser Lys Lys Val Tyr Val Gln His Arg Leu
610 615 620

Lys Glu Arg Ser Lys Glu Val Ser Asp Leu Leu Ser Gln Lys Ala Tyr
625 630 635 640

Phe Tyr Val Cys Gly Asp Ala Ala His Met Ala Arg Glu Val Asn Thr
645 650 655

Val Leu Ala Gln Ile Ile Ala Glu Gly Arg Gly Val Ser Glu Ala Lys
660 665 670

Gly Glu Glu Ile Val Lys Asn Met Arg Ser Ala Asn Gln Tyr Gln Val
675 680 685

Cys Ser Asp Phe Val Thr Leu His Cys Lys Glu Thr Thr Tyr Ala Asn
690 695 700

Ser Glu Leu Gln Glu Asp Val Trp Ser
705 710

<210> 150

<211> 453

<212> PRT

<213> Arabidopsis thaliana

<400> 150

Met Gly Gly Leu Lys Phe His Val Leu Met Tyr Pro Trp Phe Ala Thr
1 5 10 15

Gly His Met Thr Pro Phe Leu Phe Leu Ala Asn Lys Leu Ala Glu Lys
20 25 30

Gly His Thr Val Thr Phe Leu Leu Pro Lys Lys Ser Leu Lys Gln Leu
35 40 45

Glu His Phe Asn Leu Phe Pro His Asn Ile Val Phe Arg Ser Val Thr
50 55 60

Val Pro His Val Asp Gly Leu Pro Val Gly Thr Glu Thr Ala Ser Glu
65 70 75 80

Ile Pro Val Thr Ser Thr Asp Leu Leu Met Ser Ala Met Asp Leu Thr
85 90 95

Arg Asp Gln Val Glu Ala Val Val Arg Ala Val Glu Pro Asp Leu Ile
 100 105 110
 Phe Phe Asp Phe Ala His Trp Ile Pro Glu Val Ala Arg Asp Phe Gly
 115 120 125
 Leu Lys Thr Val Lys Tyr Val Val Val Ser Ala Ser Thr Ile Ala Ser
 130 135 140
 Met Leu Val Pro Gly Gly Glu Leu Gly Val Pro Pro Pro Gly Tyr Pro
 145 150 155 160
 Ser Ser Lys Val Leu Leu Arg Lys Gln Asp Ala Tyr Thr Met Lys Lys
 165 170 175
 Leu Glu Pro Thr Asn Thr Ile Asp Val Gly Pro Asn Leu Leu Glu Arg
 180 185 190
 Val Thr Thr Ser Leu Met Asn Ser Asp Val Ile Ala Ile Arg Thr Ala
 195 200 205
 Arg Glu Ile Glu Gly Asn Phe Cys Asp Tyr Ile Glu Lys His Cys Arg
 210 215 220
 Lys Lys Val Leu Leu Thr Gly Pro Val Phe Pro Glu Pro Asp Lys Thr
 225 230 235 240
 Arg Glu Leu Glu Glu Arg Trp Val Lys Trp Leu Ser Gly Tyr Glu Pro
 245 250 255
 Asp Ser Val Val Phe Cys Ala Leu Gly Ser Gln Val Ile Leu Glu Lys
 260 265 270
 Asp Gln Phe Gln Glu Leu Cys Leu Gly Met Glu Leu Thr Gly Ser Pro
 275 280 285
 Phe Leu Val Ala Val Lys Pro Pro Arg Gly Ser Ser Thr Ile Gln Glu
 290 295 300
 Ala Leu Pro Glu Gly Phe Glu Glu Arg Val Lys Gly Arg Gly Leu Val
 305 310 315 320
 Trp Gly Gly Trp Val Gln Gln Pro Leu Ile Leu Ser His Pro Ser Val
 325 330 335
 Gly Cys Phe Val Ser His Cys Gly Phe Gly Ser Met Trp Glu Ser Leu
 340 345 350
 Leu Ser Asp Cys Gln Ile Val Leu Val Pro Gln Leu Gly Asp Gln Val
 355 360 365

Leu Asn Thr Arg Leu Leu Ser Asp Glu Leu Lys Val Ser Val Glu Val
370 375 380

Ala Arg Glu Glu Thr Gly Trp Phe Ser Lys Glu Ser Leu Cys Asp Ala
385 390 395 400

Val Asn Ser Val Met Lys Arg Asp Ser Glu Leu Gly Asn Leu Val Arg
405 410 415

Lys Asn His Thr Lys Trp Arg Glu Thr Val Ala Ser Pro Gly Leu Met
420 425 430

Thr Gly Tyr Val Asp Ala Phe Val Glu Ser Leu Gln Asp Leu Val Ser
435 440 445

Gly Thr Thr His Asp
450

<210> 151
<211> 1362
<212> DNA
<213> Arabidopsis thaliana

<400> 151
atgggtggtt tgaagtttca tgtacttatg tatccatggt tcgcaacagg ccatatgacc 60
ccgttccttt ttcttgccaa caaattggct gagaaagggtc atacgggtcac tttcttgctt 120
cccaagaaat ctctgaaaca gttggaacat ttcaatctgt ttccacacaa cattgtcttt 180
cgctctgtca ccgtccctca tgtggatggt ctccccgttg gcacagagac agcctctgag 240
atccctgtga catcaactga tctcttgatg tctgctatgg atctcacacg tgatcaagtt 300
gaagctgtgg tccgagccgt tgaaccggac ctgatcttct ttgactttgc tcattggatt 360
ccagaagtag ctagggactt cggccttaag actgtaaagt acgtcgtggt gtctgcatcg 420
actatagcta gtatgcttgt ccaggtggt gagttagggtg ttctccacc gggatatcca 480
tcatcaaagg tgctgcttcg taaacaagat gcttacacta tgaagaaact ggagcctaca 540
aatacaatcg atgtcggacc aaacctcttg gaacgagtca ctacaagtct tatgaactct 600
gatgtcattg cgataaggac agccagagaa atcgaaggaa acttttgcca ctatatagaa 660
aaacattgca ggaaaaagg tctcttgaca ggtccgggtgt tccctgagcc agacaagact 720
agagagctag aggaacgatg ggttaagtgg ctaagtgggt atgaaccaga ctcagtgggtg 780
ttttgtgcac tgggctcaca agtcatttta gagaaagatc aattccaaga actctgctta 840
ggaatggagc taacaggttc accgtttctt gtagcgggtta agccccctag aggctcatca 900
acgattcaag aagcacttcc tgaaggattc gaagagcggg ttaaaggaag aggccttggt 960
tggggaggat gggttcaaca accattgata ttgtctcatc catcagtcgg gtgctttgtg 1020
agccattgtg ggtttggatc aatgtgggag tctttgctga gtgattgtca gatagtctta 1080
gtaccacagt tgggtgatca agtcctgaac acaagattgc tgagtgcga actcaagggtt 1140

tcggttgaag	tggaagaga	ggaaacagga	tggttctcga	aagagagctt	gtgcgatgct	1200
gtcaatagtg	tgatgaaaag	ggacagcgag	ctcgggaacc	tggtgaggaa	gaatcacacc	1260
aagtggaggg	agacagtagc	tagtcctgga	ctaatgactg	gttatgtcga	tgctttcgta	1320
gagtcattgc	aggatcttgt	ctctgggacc	acccatgact	ga		1362