



(51) International Patent Classification:

C12N 15/82 (2006.01)

C12N 9/12 (2006.01)

C12N 5/04 (2006.01)

(21) International Application Number:

PCT/IB2023/051306

(22) International Filing Date:

14 February 2023 (14.02.2023)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

2022900309

14 February 2022 (14.02.2022) AU

(71) Applicants: THE NEW ZEALAND INSTITUTE FOR PLANT AND FOOD RESEARCH LIMITED [NZ/NZ];

Mt Albert Research Centre, 120 Mt Albert Road, Mt Albert, Auckland, 1025 (NZ), WUHAN BOTANICAL GARDEN [CN/CN]; Chinese Academy of Sciences, No 201 Jiufeng 1 Road, East Lake High-Tech Development Zone, Wuhan, Hubei 430074 (CN).

(72) Inventors; and

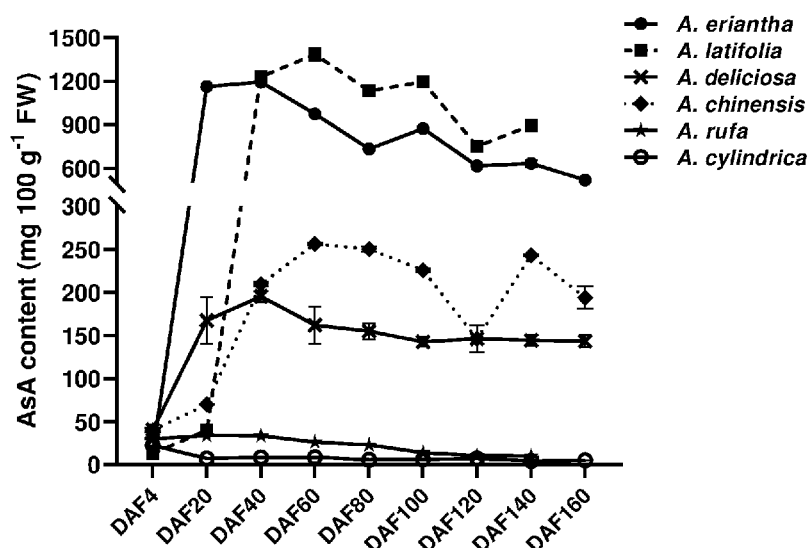
(71) Applicants: **LI, Dawei** [CN/CN]; c/o Wuhan Botanical Garden, Chinese Academy of Sciences, No. 201 Jiufeng 1 Road, East Lake High-Tech Development Zone, Wuhan, Hubei 430074 (CN). **LIU, Xiaoying** [CN/CN]; c/o Wuhan Botanical Garden, Chinese Academy of Sciences, No. 201 Jiufeng 1 Road, East Lake High-Tech Development Zone, Wuhan, Hubei 430074 (CN). **ZHONG, Caihong** [CN/CN]; c/o Wuhan Botanical Garden, Chinese Academy of Sciences, No. 201 Jiufeng 1 Road, East Lake High-Tech Development Zone, Wuhan, Hubei 430074 (CN). **XIE, Xiaodong** [CN/CN]; c/o Wuhan Botanical Garden, Chinese Academy of Sciences, No. 201 Jiufeng 1 Road, East Lake High-Tech Development Zone, Wuhan, Hubei 430074 (CN). **BULLEY, Sean Michael Winsley** [NZ/NZ]; c/o The New Zealand Institute for Plant and Food Research Limited, Mt Albert Research Centre, 120 Mt Albert Road, Mt Albert, Auckland, 1025 (NZ).

(74) Agent: **AJ PARK**; Level 22, Aon Centre, 1 Willis Street, Wellington (NZ).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ,

(54) Title: REGULATION OF GENE EXPRESSION

FIGURE 1



(57) Abstract: The invention provides a method of producing a host cell, plant cell or plant with increased L-ascorbic acid (AsA) content or increased GDP-L-galactose phosphorylase (GPP) translation, production and/or activity, the method comprising transformation of the host cell or plant cell with a polynucleotide encoding a polypeptide that regulates GDP-L-galactose phosphorylase (GPP). The invention also provides host cells, plant cells and plants, genetically modified to contain and/or express the polynucleotides.

CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

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

Published:

- with international search report (Art. 21(3))
- with sequence listing part of description (Rule 5.2(a))
- in black and white; the international application as filed contained color or greyscale and is available for download from PATENTSCOPE

REGULATION OF GENE EXPRESSION

FIELD OF THE INVENTION

[0001] The present invention relates to polynucleotides, genetic constructs and methods for producing plants with altered GDP-L-galactose phosphorylase 3 (GGP) translation, production and/or activity; and/or altered L-ascorbic acid (AsA) content; and plants and plant cells produced from said methods.

BACKGROUND TO THE INVENTION

[0002] L-ascorbic acid (AsA) is the most abundant soluble antioxidant in plants and is also an essential nutrient for humans and a few other animals. AsA contributes significantly to the overall intake of "free radical scavengers" or "anti-oxidative metabolites" in the human diet. Convincing evidence now shows that such metabolites either singly or in combination, benefit health and well-being, acting as anti-cancer forming agents and protecting against coronary heart disease. Eating sufficient amounts of fruits and vegetables to consistently maintain optimum AsA concentrations is still a challenge for people in both developed and developing countries.

[0003] Almost the entire dietary AsA intake in humans is derived from plant products. However, the AsA content of plant tissues is remarkably variable. Whilst leaf AsA content is generally high and relatively uniform in herbaceous and woody plants, a huge and unexplained variability in AsA content is found in non-green edible plant tissues.

[0004] Understanding how AsA biosynthesis is regulated may provide tools to manipulate biosynthesis in plants. Understanding the regulation of gene expression, and the factors/elements controlling such expression also provide valuable tools for genetic manipulation. The enzyme GDP-L-galactose phosphorylase (GGP) is known to play a key role in AsA biosynthesis. The regulation of GGP expression and activity is yet to be fully resolved.

[0005] It would be desirable to increase the AsA content of plants.

[0006] It is an object of the invention to provide new, improved and/or alternative tools for manipulating the expression and content of AsA in plants, and/or to at least provide the public with a useful choice.

[0007] In this specification, where reference has been made to external sources of information, including patent specifications and other documents, this is generally for the purpose of providing a context for discussing the features of the present invention.

Unless stated otherwise, reference to such sources of information is not to be construed, in any jurisdiction, as an admission that such sources of information are prior art or form part of the common general knowledge in the art.

SUMMARY OF THE INVENTION

5 **[0008]** In a first aspect the invention relates to a method of producing a plant cell or plant with increased L-ascorbic acid (AsA) content, the method comprising transforming a plant cell with a polynucleotide encoding a polypeptide with an amino acid sequence of any one of SEQ ID NOs: 1-102 and 348-356, or a variant of the polypeptide having at least about 70% sequence identity to a polypeptide with the amino acid sequence of any one of SEQ ID NOs: 1-102 and 348-356.

15 **[0009]** In a second aspect the invention relates to a method of producing a plant cell or plant having increased GDP-L-galactose phosphorylase 3 (GGP) translation, production and/or activity, the method comprising transforming a plant cell with a polynucleotide encoding a polypeptide with an amino acid sequence of any one of SEQ ID NOs: 1-102 and 348-356, or a variant of the polypeptide having at least about 70% sequence identity to a polypeptide with the amino acid sequence of any one of SEQ ID NOs: 1-102 and 348-356.

20 **[0010]** In a third aspect the invention relates to a method of producing a plant cell or plant with increased L-ascorbic acid (AsA) content, the method comprising transforming a plant cell with a polynucleotide comprising a nucleotide sequence of any one of SEQ ID NOs: 103-108, or a variant thereof having at least about 70% sequence identity to a nucleotide having the nucleotide sequence of any one of SEQ ID NOs: 103-108.

25 **[0011]** In a fourth aspect the invention relates to a method of producing a plant cell or plant having increased GDP-L-galactose phosphorylase 3 (GGP) translation, production and/or activity, the method comprising transforming a plant cell with a polynucleotide comprising a nucleotide sequence of any one of SEQ ID NOs: 103-108, or a variant thereof having at least about 70% sequence identity to a nucleotide having the nucleotide sequence of any one of SEQ ID NOs: 103-108.

30 **[0012]** In a fifth aspect the invention relates to a method of producing a plant cell or plant with increased L-ascorbic acid (AsA) content or increased GDP-L-galactose phosphorylase 3 (GGP) translation, production and/or activity, the method comprising upregulating in the plant cell, or expressing in the plant, a polypeptide with the amino acid sequence of any one of SEQ ID NO: 1-102 and 348-356, or a variant of the

polypeptide having at least about 70% sequence identity to the amino acid sequence of any one of SEQ ID NO: 1-102 and 348-356.

- [0013]** In a sixth aspect the invention relates to a method of producing a plant cell or plant with increased L-ascorbic acid (AsA) content or increased GDP-L-galactose phosphorylase 3 (GGP) translation, production and/or activity, the method comprising upregulating in the plant cell, or expressing in the plant, a polynucleotide comprising a nucleotide sequence selected from any one of the sequences SEQ ID NO: 103-108 or a variant having at least about 70% sequence identity to the nucleotide sequence of any one of SEQ ID NO: 103-108.
- [0014]** In a seventh aspect, the invention relates to an isolated polynucleotide encoding a polypeptide with the amino acid sequence of any one of SEQ ID NOs: 1-4 and 102, or a variant of the polypeptide having at least about 70% sequence identity to the amino acid sequence of any one of SEQ ID NOs: 1-4 and 102, preferably at least about 75% sequence identity.
- [0015]** In an eighth aspect, the invention relates to an isolated polynucleotide encoding a polypeptide with an amino acid sequence of any one of SEQ ID NOs: 109-153 155-157, 159, 162-164, 166-174, 176, 178, 180-183, 185, 186, 188-190, 192-200, 202-204, and 206-209, or a variant of the polypeptide having at least about 70% sequence identity to the amino acid sequence of any one of SEQ ID NOs: 109-153 155-157, 159, 162-164, 166-174, 176, 178, 180-183, 185, 186, 188-190, 192-200, 202-204, and 206-209, preferably at least about 75% sequence identity.
- [0016]** In a ninth aspect the invention relates to a genetic construct comprising a polynucleotide encoding a polypeptide with an amino acid sequence of any one of SEQ ID NOs: 1-4 and 102, or a variant of the polypeptide having at least about 70% sequence identity to the amino acid sequence of any one of SEQ ID NOs: 1-4 and 102, preferably at least about 75% sequence identity.
- [0017]** In a tenth aspect the invention relates to a genetic construct comprising a polynucleotide encoding a polypeptide with an amino acid sequence of any one of SEQ ID NOs: 109-153 155-157, 159, 162-164, 166-174, 176, 178, 180-183, 185, 186, 188-190, 192-200, 202-204, and 206-209, or a variant of the polypeptide having at least about 70% sequence identity to the amino acid sequence of any one of SEQ ID NOs: 109-153 155-157, 159, 162-164, 166-174, 176, 178, 180-183, 185, 186, 188-190, 192-200, 202-204, and 206-209, preferably at least about 75% sequence identity.

[0018] In various embodiments the genetic construct may further comprise a second polynucleotide

- a) encoding a polypeptide with an amino acid sequence of any one of SEQ ID NOs: 109-209, or a variant of the polypeptide having at least about 70% sequence identity to the amino acid sequence of any one of SEQ ID NOs: 109-209, preferably at least about 75% sequence identity; and/or
- b) having a nucleotide sequence of SEQ ID NO: 210, or a variant of the nucleotide having at least about 70% sequence identity to the nucleotide sequence of SEQ ID NO: 210.

[0019] In an eleventh aspect the invention relates to a genetic construct comprising a polynucleotide comprising a nucleotide sequence of any one of SEQ ID NOs: 103-108, or a variant of the polypeptide having at least about 70% sequence identity to the nucleotide sequence of any one of SEQ ID NOs: 103-108, preferably at least about 75% sequence identity.

[0020] In a twelfth aspect the invention relates to a genetic construct comprising a polynucleotide comprising a nucleotide sequence of SEQ ID NO: 210, or a variant of the polypeptide having at least about 70% sequence identity to the nucleotide sequence of SEQ ID NO: 210.

[0021] In a thirteenth aspect the invention relates to a kit comprising a first genetic construct of the ninth or eleventh aspects described herein; and a second genetic construct of the tenth or twelfth aspects described herein.

[0022] In a fourteenth aspect the invention provides a host cell comprising one or more genetic constructs described herein.

[0023] In a fifteenth aspect the invention provides a host cell genetically modified to express a polynucleotide encoding a polypeptide with an amino acid sequence of any one of SEQ ID NOs: 1-102 and 348-356, or a variant of the polypeptide having at least about 70% sequence identity to the amino acid sequence of any one of SEQ ID NOs: 1-102 and 348-356.

[0024] In a sixteenth aspect the invention provides a host cell genetically modified to express a polynucleotide comprising a nucleotide sequence of any one of SEQ ID NOs: 103-108, or a variant thereof having at least about 70% sequence identity to the nucleotide sequence of any one of SEQ ID NOs: 103-108.

- 5 **[0025]** In a seventeenth aspect the invention provides a plant cell genetically modified to express a polynucleotide encoding a polypeptide with the amino acid sequence of any one of SEQ ID NO: 1-102 and 348-356 or a variant of the polypeptide having at least 70% sequence identity to a polypeptide with the amino acid sequence of any one of SEQ ID NO: 1-102 and 348-356.
- [0026]** In an eighteenth aspect the invention provides a plant cell genetically modified to express a polynucleotide comprising a nucleotide sequence selected from any one of SEQ ID NO: 103-108 or a variant thereof having at least 70% sequence identity to the nucleotide sequence of any one of SEQ ID NO: 103-108.
- 10 **[0027]** In a nineteenth aspect the invention provides a plant comprising a plant cell described herein.
- [0028]** In a twentieth aspect the invention provides a method for selecting a plant with altered GDP-L-galactose phosphorylase 3 (GGP) activity, the method comprising testing a plant for altered expression of a polynucleotide encoding a polypeptide with the amino acid sequence of any one of SEQ ID NO: 1-102 and 348-356 or a variant of the polypeptide having at least 70% sequence identity to a polypeptide with the amino acid sequence of any one of SEQ ID NO: 1-102 and 348-356.
- 15 **[0029]** In a further aspect the invention provides a method for selecting a plant with altered GDP-L-galactose phosphorylase 3 (GGP) activity, the method comprising testing a plant for altered expression of a polynucleotide comprising a nucleotide sequence selected from any one of the sequences SEQ ID NO: 103-108 or a variant thereof wherein the variant comprises a sequence that has at least 70% sequence identity to the nucleotide sequence of any one of SEQ ID NO: 103-108.
- 20 **[0030]** In a further aspect the invention provides a method for selecting a plant with altered L-ascorbic acid (AsA) content, the method comprising testing a plant for altered expression of a polynucleotide encoding a polypeptide with the amino acid sequence of any one of SEQ ID NO: 1-102 and 348-356 or a variant of the polypeptide having at least 70% sequence identity to a polypeptide with the amino acid sequence of any one of SEQ ID NO: 1-102 and 348-356.
- 25 **[0031]** In a further aspect the invention provides a method for selecting a plant with altered L-ascorbic acid (AsA) content, the method comprising testing a plant for altered expression of a polynucleotide comprising a nucleotide sequence selected from any one of the sequences SEQ ID NO: 103-108 or a variant thereof comprising a sequence that
- 30

has at least 70% sequence identity to the nucleotide sequence of any one of SEQ ID NO: 103-108.

[0032] In a further aspect the invention provides a plant cell or plant produced by any method described herein.

5 **[0033]** In a further aspect the invention provides a plant cell or plant selected by any method described herein.

[0034] In a further aspect the invention provides a group or population of plants produced by any method described herein.

10 **[0035]** In a further aspect the invention provides a method of producing AsA, the method comprising extracting AsA from a host cell described herein.

[0036] In a further aspect the invention provides a method of producing AsA, the method comprising extracting AsA from a plant cell or plant described herein.

[0037] The following embodiments may relate to any or all of the above aspects.

15 **[0038]** In various embodiments the method may further comprise transforming, or co-transforming with the polynucleotide, a second polynucleotide encoding a polypeptide with an amino acid sequence of any one of SEQ ID NOs: 109-209, or a variant of the polypeptide having at least about 70% sequence identity to the amino acid sequence of any one of SEQ ID NOs: 109-209.

20 **[0039]** In various embodiments the method may further comprise transforming, or co-transforming with the polynucleotide, a second polynucleotide with a nucleotide sequence of SEQ ID NO: 210, or a variant thereof having at least about 70% sequence identity to the nucleotide sequence of SEQ ID NO: 210. In various embodiments the variant has at least about 75, 80, 85, 90, 95, 96, 97, 98, 99 or 100% sequence identity to the nucleotide sequence of SEQ ID NO: 210.

25 **[0040]** In various embodiments the host cell or plant cell may be genetically modified to express a second polynucleotide encoding a polypeptide with an amino acid sequence of any one of SEQ ID NOs: 109-209, or a variant of the polypeptide having at least about 70% sequence identity to the amino acid sequence of any one of SEQ ID NOs: 109-209.

30 **[0041]** In various embodiments the host cell or plant cell may be genetically modified to express a second polynucleotide with a nucleotide sequence of SEQ ID NO:

210, or a variant thereof having at least about 70% sequence identity to the nucleotide sequence of SEQ ID NO: 210.

- [0042]** In various embodiments the polynucleotide may encode a polypeptide with an amino acid sequence of, or may encode a variant of the polypeptide having at least
- 5 about 70, 75, 80, 85, 90, 95, 96, 97, 98, 99 or 100% sequence identity to an amino acid sequence of
- a) any one of SEQ ID Nos:1-102 and 348-356;
 - b) any one of SEQ ID Nos: 1-50, 52-60, 62-83, 85-95, 97-102, 348-350, and 352-354;
 - 10 c) any one of SEQ ID Nos: 1-49, 53-56, 58-60, 62, 64-66, 70-79, 82, 83, 85-87, 89-91, 93-95, 98, 102, 348-350, 352, and 353;
 - d) any one of SEQ ID Nos: 1-49, 53-56, 58, 60, 62, 64, 66, 70, 71, 74-76, 78, 79, 83, 85-87, 89-91, 93-95, 98, 102, 348-350, 352, and 353;
 - 15 e) any one of SEQ ID Nos: 1-49, 53-56, 58, 60, 62, 64, 66, 70, 71, 74-76, 78, 79, 83, 85-87, 89-90, 93, 94, 98, 102, 348-350, 352, and 353;
 - f) any one of SEQ ID Nos: 1-5, 7-10, 12-19, 21-24, 28-31, 33, 35, 36, 38-45, 48, 49, 53, 55, 56, 58, 60, 62, 64, 66, 71, 74-76, 79, 83, 85, 87, 89, 90, 93, 94, 98, 102, 350, 352;
 - 20 g) any one of SEQ ID Nos:1-4, 38, 41-45, 48, 49, 51, 52, 56, 60, 63, 66, 79, 80, 85, 89, 90, 92, 96, 97, 100-102, and 348-356;
 - h) any one of SEQ ID Nos:1-4, 49, 51, 66, 80, 85, 90, 92, 96, 102, and 348-356;
 - i) any one of SEQ ID Nos:1-4 and 102;
 - j) any one of SEQ ID Nos:1-4;
 - k) SEQ ID No:1 or 102;
 - 25 l) SEQ ID No:1; or
 - m) SEQ ID No:102.

[0043] In various embodiments the polynucleotide may have a nucleotide sequence of, or may encode a variant of the polynucleotide having at least about 70, 75, 80, 85, 90, 95, 96, 97, 98, 99 or 100% sequence identity to a nucleotide sequence of:

- a) any one of SEQ ID NOs:103-108,
- 5 b) SEQ ID No: 103 or 105, or
- c) SEQ ID No: 103.

[0044] In various embodiments the polynucleotide, or the second polynucleotide, may encode a polypeptide with an amino acid sequence of, or may encode a variant of the polypeptide having at least about 70, 75, 80, 85, 90, 95, 96, 97, 98, 99 or 100% sequence identity to an amino acid sequence of

- a) any one of SEQ ID Nos: 109-209;
- b) any one of SEQ ID Nos: 109-174, 176-178, and 180-209;
- c) any one of SEQ ID Nos: 109-159, 162-164, 166-174, 176, 178, 180-186, 188-190, 192-204, and 206-209;
- 15 d) any one of SEQ ID Nos: 109-153 155-157, 159, 162-164, 166-174, 176, 178, 180-186, 188-190, 192-204, and 206-209;
- e) any one of SEQ ID Nos: 109-153 155-157, 159, 162-164, 166-174, 176, 178, 180-183, 185, 186, 188-190, 192-200, 202-204, and 206-209;
- f) any one of SEQ ID Nos: 109-114, 125, 126, 142, 177, 179, 187, 191 and 205;
- 20 g) any one of SEQ ID Nos: 109-114 and 179;
- h) any one of SEQ ID Nos: 109-114; or
- i) SEQ ID NO:109.

[0045] In various embodiments the polynucleotide, or the second polynucleotide, may have a nucleotide sequence of, or may encode a variant of the polynucleotide having at least about 70, 75, 80, 85, 90, 95, 96, 97, 98, 99 or 100% sequence identity to a nucleotide sequence of SEQ ID NO:210.

[0046] In various embodiments the polypeptide or polypeptide variant may comprise one or more SANT/MYB domains. In various embodiments the polypeptide or polypeptide variant may comprise two SANT/MYB domains.

[0047] In various embodiments the upregulating may comprise genetic engineering.

5 **[0048]** In various embodiments the transformation comprises stable transformation. In various embodiments, the method comprises stably transforming the plant cell with the polynucleotide or variant thereof. In various embodiments the host cell, or plant cell, is stably transformed and/or stably genetically modified.

10 **[0049]** In various embodiments the method may comprise transforming a plant cell with a genetic construct described herein. In various embodiments the method may comprise stably transforming a plant cell with a genetic construct described herein.

15 **[0050]** In various embodiments the host cell or plant cell comprises an endogenous GGP gene, preferably a functional GGP gene. For example, in some embodiments the host cell or plant cell has been previously transformed with a functional GGP gene. In some embodiments the method comprises transforming the host cell or plant cell with a GGP gene.

20 **[0051]** In some embodiments, the GGP gene has a promoter that comprises one or more MYBS1 binding sites, preferably two or more MYBS1 binding sites. In some embodiments, the method comprises gene-editing the promoter of an endogenous GGP gene to provide a promoter comprising one or more MYBS1 binding sites. In some embodiments, the method comprises modifying an endogenous GGP promoter to provide a GGP promoter comprising one or more MYBS1 binding sites. In some embodiments, the method comprises transforming the plant cell or host cell with a GGP gene. In some
25 **[0051]** In some embodiments, the method comprises transforming the plant cell or host cell with a GGP gene having a promoter that comprises one or more MYBS1 binding sites, preferably two or more MYBS1 binding sites.

30 **[0052]** In various embodiments the upregulating may comprise crossing with a plant that expresses a polypeptide comprising an amino acid sequence having at least 70, 75, 80, 85, 90, 95, 96, 97, 98, or 99% sequence identity to a polypeptide with the amino acid sequence of any one of SEQ ID NO: 1-102 and 348-356 and/or SEQ ID No:109-209.

[0053] According to some embodiments, the upregulating comprises crossing with a plant which expresses a polynucleotide comprising a nucleotide sequence selected from any one of the sequences SEQ ID NO: 103 to 108 and 210.

[0054] In various embodiments the genetic construct may comprise one or more polynucleotides operably linked to a promoter. In various embodiments the promoter is at least one of:

- i) a promoter that is not normally associated with the polynucleotide in nature,
- 5 ii) a promoter derived from a bacterium, a fungus, an insect, a mammal or a virus.
- iii) a bacterial promoter,
- iv) a fungal promoter,
- v) an insect promoter,
- vi) a mammalian promoter, and
- 10 vii) a virus promoter.

[0055] In various embodiments the promoter may be derived from a bacterium, a fungus, an insect, a mammal or a virus.

- [0056]** In various embodiments, the promoter may be a constitutive promoter, or a tissue-specific promoter. In various embodiments, the promoter results in expression in
- 15 leaves and/or roots. Preferably, the promoter results in expression in leaves. More preferably, the promoter is a tissue-specific promoter that results in targeted expression in leaves and/or roots.

- [0057]** In various embodiments the one or more polynucleotides may be regulatable by a compound. In some embodiments the compound is L-ascorbic acid (AsA), or a
- 20 related metabolite.

[0058] In various embodiments the host cell may be a plant cell.

- [0059]** The term "comprising" as used in this specification and claims means "consisting at least in part of". When interpreting statements in this specification and claims which include the term "comprising", other features besides the features prefaced
- 25 by this term in each statement can also be present. Related terms such as "comprise" and "comprised" are to be interpreted in similar manner.

[0060] As used herein the term "and/or" means "and" or "or", or both.

[0061] As used herein the term '(s)' following a noun means the plural and/or singular form of that noun.

[0062] It is intended that reference to a range of numbers disclosed herein (for example, 1 to 10) also incorporates reference to all rational numbers within that range (for example, 1, 1.1, 2, 3, 3.9, 4, 5, 6, 6.5, 7, 8, 9 and 10) and also any range of rational numbers within that range (for example, 2 to 8, 1.5 to 5.5 and 3.1 to 4.7) and, therefore, all sub-ranges of all ranges expressly disclosed herein are hereby expressly disclosed. These are only examples of what is specifically intended and all possible combinations of numerical values between the lowest value and the highest value enumerated are to be considered to be expressly stated in this application in a similar manner.

[0063] This invention may also be said broadly to consist in the parts, elements and features referred to or indicated in the specification of the application, individually or collectively, and any or all combinations of any two or more said parts, elements or features, and where specific integers are mentioned herein which have known equivalents in the art to which this invention relates, such known equivalents are deemed to be incorporated herein as if individually set forth.

[0064] Although the present invention is broadly as defined above, those persons skilled in the art will appreciate that the invention is not limited thereto and that the invention also includes embodiments of which the following description gives examples.

BRIEF DESCRIPTION OF THE DRAWINGS

[0065] The invention will now be described by way of example only and with reference to the drawings in which:

[0066] Figure 1 provides a line graph demonstrating AsA content in six species of *Actinidia* in the days after flowering (DAF).

[0067] Figure 2 provides graphs demonstrating the expression of *GGP* (*GGP1*, *GGP2* and *GGP3*) (lines) and AsA content (bars) in (A) *A. rufa* and (B) *A. eriantha* in the days after flowering (DAF). Data is shown as the mean +/- standard deviation (n=3).

[0068] Figure 3 provides bar graphs demonstrating that over-expression or knock down expression of *AceGGP3* alters AsA concentration in *A. eriantha* fruits. (A) shows the expression abundance of *AceGGP3* and (B) shows the AsA content in *A. eriantha* fruit 7 days after infiltration; and (C) shows the expression abundance of *AceGGP3* and (D) shows the AsA content in *A. eriantha* calli 6 days after infiltration with: EV (empty vector); *OX-AceGGP3* (*AceGGP3* overexpression) or *TRV-AceGGP3* (*AceGGP3*-antisense expression). Data is shown as the mean +/- standard deviation (n=3). Significant

differences were detected by t-test using GraphPad Prism 8 (*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$).

[0069] Figure 4 provides bar graphs showing (A) RT-qPCR analysis of *AceGGP3* and (B) relative AsA content in 5 transgenic lines of *A. eriantha* with *AceGGP3*-overexpressed. WT: wild-type; OE-GGP3#10, OE-GGP3#14, OE-GGP3#20, OE-GGP3#21, OE-GGP3#23 represent the five different transgenic lines, respectively. Data is shown as the mean $\pm$ standard deviation ($n=3$). Significant differences were detected by t-test using GraphPad Prism 8 (*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$).

[0070] Figure 5 provides a bar graph showing the relative AsA content of *A. eriantha calli* in CRISPR/Cas9-induced *AceGGP3* mutants. wt: wild-type; ggp3#2, ggp3#11, ggp3#13 and ggp3#15 represent four gene editing lines, respectively. Data is shown as the mean $\pm$ standard deviation ($n=3$). Significant differences were detected by t-test using GraphPad Prism 8 (*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$).

[0071] Figure 6 provides bar and line graphs showing a correlation between (A) AsA content (bars) and expression of *AceMYBS1* (lines), and (B) the expression of *AceGGP3* and *AceMYBS1* in different developmental stage of *A. eriantha* fruits (every 20 days). Error bars represent $\pm$ -SD ($n=3$).

[0072] Figure 7A provides a bar graph showing the results of dual-luciferase assays in tobacco leaves showing that *AceMYBS1* activates transcription of different length fragments of *AceGGP3* promoters (P2660, P2088, P1606, P1106). The empty vector (EV) is a control. Error bars: $\pm$ SD. Significant differences were detected by t-test (**, $p < 0.01$; ns: no significance). Figure 7B shows the position of two *AceMYBS1* predicted binding targets on the *AceGGP3* promoter and electrophoretic mobility shift assay (EMSA) showing the binding of *AceMYBS1* to the *AceGGP3* promoter. The unlabelled probes were used as competitors. 100x and 300x represent the rates of the competitor.

[0073] Figure 8 provides bar graphs showing the results of RT-qPCR analysis of (A) *AceMYBS1* and (B) *AceGGP3* transcripts, and (C) AsA content of *A. eriantha* fruits (of B) 7 days after infiltration with transient expressed-*AceMYBS1* vectors. EV: empty vector; OX-*AceMYBS1*: *AceMYBS1*-overexpression; TRV-*AceMYBS1*: *AceMYBS1*-antisense expression. Experiments were repeated three times and each experiment contained three to six kiwifruits per genotype. Error bars: $\pm$ -SD. Significant differences were detected by t-test (*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$; ns: no significance).

[0074] Figure 9 provides bar graphs showing the results of RT-qPCR analysis of (A) *AceMYBS1* and (B) *AceGGP3*; and (C) AsA content in *A. eriantha* calli transfected with the

following vectors: EV: empty vector; OX-AceMYBS1: AceMYBS1-overexpression; TRV-AceMYBS1: AceMYBS1-antisense expression; TRV-AceGGP3: AceGGP3-antisense expression; TRV-AceGGP3+OX-AceMYBS1: overexpressed AceMYBS1 in TRV-AceGGP3 background. All experiments were performed in three replicates. Error bars: +/-SD.

- 5 Significant differences were detected by t-test (*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$; ns: no significance).

[0075] Figure 10 provides bar graphs showing the results of RT-qPCR analysis of (A) AceMYBS1 and (B) AceGGP3 in wild-type (WT) and six AceMYBS1-overexpression transgenic kiwifruit (*A. eriantha*) lines: OE-AceMYBS1#1, OE-AceMYBS1#2, OE-AceMYBS1#3, OE-AceMYBS1#4, OE-AceMYBS1#5, OE-AceMYBS1#7.

[0076] Figure 11 provides bar graphs showing the results of (A) bimolecular luminescence complementation (BiLC) assay demonstrating that AceMYBS1 interacts with AceGBF3 *in vivo* and (B) an *in vitro* pull-down assay showing the interaction between AceMYBS1 and AceGBF3. In (A), Agrobacterium clones containing the respective recombinant plasmids were combined at 1:1 (v/v) and infiltrated into *N. benthamiana* leaves. Error bars show the mean +/- SD; Significant differences were detected by t-test: ***, $p < 0.001$; ns: no significance. In (B), AceGBF3-GST protein was incubated with immobilized AceMYBS1-6xHis or 6xHis protein, and immuno-precipitated fractions were detected by Anti-GST antibody.

[0077] Figure 12 provides bar graphs showing the results of (A) dual luciferase assays in tobacco leaves showing the transcription of AceGGP3 activated by AceMYBS1 and AceGBF3 individually or collectively; (B) AsA content and (C) RT-qPCR analysis of the expression level in transiently expressed *A. eriantha* fruits 7 days post-transformation; the expression level in transiently expressed *A. eriantha* calli of (D) AceGGP3 and (E) AsA content; and the expression level in transiently expressed tobacco leaves of (F) NbGGP and (G) AsA content. EV: empty vector; OX-AceGBF3: AceGBF3-overexpression; TRV-AceGBF3: AceGBF3-antisense expression. The experiments were repeated three times and each experiment contain three to six kiwifruits per genotype. All above error bars denoted standard deviation (+/-SD), and performed three technical repeats each experimental group. Significant differences were detected by t-test (*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$; ns: no significance). Different letters above the bars indicated significant difference ($p < 0.05$) as obtain by one-way ANOVA test.

[0078] Figure 13 provides bar graphs showing (A) AceGGP3 expression level as determined by RT-qPCR analysis and (B) AsA content in AceMYBS1 and AceGBF3 transiently co-expressed *A. eriantha* fruits. EV: empty vector; TRV-AceMYBS1: AceMYBS1-antisense expression; OX-AceMYBS1: AceMYBS1-overexpression; TRV-

AceMYBS1+TRV-AceGBF3: co-antisense expression AceMYBS1 and AceGBF3; OX-AceMYBS1+OX-AceGBF3: co-overexpression AceMYBS1 and AceGBF3; OX-AceMYBS1+TRV-AceGBF3: AceMYBS1-overexpression in the AceGBF3-antisense expression background; OX-AceGBF3+TRV-AceMYBS1: AceGBF3-overexpression in the AceMYBS1-antisense expression background. The experiments were repeated three times and each experiment contain three to six kiwifruits per genotype. Significant differences were detected by t-test (*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$; ns: no significance).

[0079] Figure 14 provides bar graphs showing (A) relative AsA content and (B) RT-qPCR analysis of AceGGP3 in wild type (WT) and transgenic *A. eriantha* calli overexpressing AceGBF3. All above error bars denoted standard deviation (+/-SD), and performed three technical repeats each experimental group. Significant differences were detected by t-test (*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$; ns: no significance).

[0080] Figure 15 provides bar graphs showing the mean leaf AsA content (in mg/100g fresh weight) of (A) rice, (B) soybean, and (C) *Arabidopsis thaliana* plants stably transformed to express AceMYBS1. Individual transformed lines are shown as separate bars. Measurements were performed in triplicate, and error bars denote the standard deviation. In (A) and (B), asterisks denote significant difference from wild-type according to Student's T-test ($p < 0.05$). In (C), different letters denote significantly different groupings according to Student's T-test ($p < 0.05$). WT: wild-type control plants.

DETAILED DESCRIPTION OF THE INVENTION

[0081] The present invention, in some embodiments thereof, relates to methods for producing host cells, including plant cells or plants, having increased L-ascorbic acid (AsA) and/or increased GGP translation, production and/or activity.

[0082] The present invention is based on the identification, through genetic and molecular characterisation described herein, of two bZIP transcription factors that are positive regulators of AsA biosynthesis by activating transcription of GGP. An MYBS1-like transcription factor in kiwifruit was shown to bind the promoter of GGP3 and when overexpressed in kiwifruit resulted in significantly increased GGP3 expression and AsA accumulation. Overexpression of a GBF3 bZIP transcription factor also increased AsA content in an additive manner with MYBS1.

[0083] Identification of transcription factors that contribute to the regulation of AsA biosynthesis in plants allows for marker aided selection to be developed to breed plants having higher AsA content, and for high levels of AsA to be produced via biotechnological

means, such as biopharming in crop plants and metabolic engineering of host cells such as yeast.

[0084] Thus, according to one aspect of the invention there is provided a method of producing a plant cell or plant with increased L-ascorbic acid (AsA) content, and/or increased GDP-L-galactose phosphorylase 3 (GGP) translation, production and/or activity, the method comprising transforming a plant cell with a polynucleotide encoding a polypeptide with an amino acid sequence of any one of SEQ ID NOs: 1-102 and 348-356, or a variant of the polypeptide having at least about 70% sequence identity to a polypeptide with the amino acid sequence of any one of SEQ ID NOs: 1-102 and 348-356 (MYBS1-like protein).

[0085] In various embodiments the method may further comprise transforming, or co-transforming with the polynucleotide, a second polynucleotide encoding a polypeptide with an amino acid sequence of any one of SEQ ID NOs: 109-209, or a variant of the polypeptide having at least about 70% sequence identity to the amino acid sequence of any one of SEQ ID NOs: 109-209 (GBF3-like protein).

[0086] The applicants have identified a polynucleotide (SEQ ID No:103) encoding an MYBS1-like transcription factor polypeptide (SEQ ID No:1) and a polynucleotide (SEQ ID No:210) that encodes a GBF3-like transcription factor polypeptide (Seq ID No:109) from *Actinidia eriantha* as described in the examples.

[0087] The applicants have shown that these polypeptides induce expression of GGP3 and increase AsA content in several kiwifruit species and a tobacco species.

[0088] The applicants have also identified

- a) MYBS1-like polypeptide sequences from a number of species that have significant sequence conservation with SEQ ID No:1 and are variants of each other (SEQ ID Nos:2-102 and 348-356), and
- b) GBF3-like polypeptide sequences from a number of species that have significant sequence conservation with SEQ ID No:109 and are variants of each other (SEQ ID Nos:110-209).

[0089] Genetic constructs, vectors and plants containing polynucleotide sequences encoding an MYBS1-like polypeptide (SEQ ID NOs: 1-102 and 348-356) or sequences encoding the polypeptide sequences (SEQ ID NO: 103-108) and/or polynucleotide sequences encoding a GBF3-like polypeptide (SEQ ID Nos:109-209) or sequences encoding the polypeptide sequences (SEQ ID NO: 210) are disclosed herein.

[0090] In certain embodiments, there are provided plants and host cells comprising the genetic constructs and vectors disclosed herein. Preferably the plants and host cells are stably transformed with the genetic constructs and/or vectors.

5 [0091] In some embodiments, there are provided plants altered GGP translation, production and/or activity relative to suitable control plants, and plants altered in AsA content relative to suitable control plants. In some embodiments, there are provided plants with increased GGP translation, production and/or activity and increased AsA. Preferably the plants are stably transformed or modified.

10 [0092] In other embodiments there are provided methods for the production of such plants and methods of selection of such plants.

[0093] Suitable control plants include non-transformed plants of the same species or variety or plants transformed with control constructs.

1. Polynucleotides and fragments

15 [0094] The term "polynucleotide(s)" as used herein, means a single or double-stranded deoxyribonucleotide or ribonucleotide polymer of any length but preferably at least 15 nucleotides, and include as non-limiting examples, coding and non-coding sequences of a gene, sense and antisense sequences complements, exons, introns, genomic DNA, cDNA, pre-mRNA, mRNA, rRNA, siRNA, miRNA, tRNA, ribozymes, recombinant polynucleotides, isolated and purified naturally occurring DNA or RNA
20 sequences, synthetic RNA and DNA sequences, nucleic acid probes, primers and fragments.

[0095] A "fragment" of a polynucleotide sequence provided herein is a subsequence of contiguous nucleotides that is capable of specific hybridization to a target of interest, e.g., a sequence that is at least 15 nucleotides in length. Fragments as herein disclosed
25 comprise 15 nucleotides, preferably at least 20 nucleotides, more preferably at least 30 nucleotides, more preferably at least 50 nucleotides, more preferably at least 50 nucleotides and most preferably at least 60 nucleotides of contiguous nucleotides of a polynucleotide as herein disclosed. A fragment of a polynucleotide sequence can be used in antisense, gene silencing, triple helix or ribozyme technology, or as a primer, a probe,
30 included in a microarray, or used in polynucleotide-based selection methods as herein disclosed.

[0096] The term "primer" refers to a short polynucleotide, usually having a free 3'OH group, that is hybridized to a template and used for priming polymerization of a polynucleotide complementary to the template. Such a primer is preferably at least 5,

more preferably at least 6, more preferably at least 7, more preferably at least 9, more preferably at least 10, more preferably at least 11, more preferably at least 12, more preferably at least 13, more preferably at least 14, more preferably at least 15, more preferably at least 16, more preferably at least 17, more preferably at least 18, more preferably at least 19, more preferably at least 20 nucleotides in length.

[0097] The term “probe” refers to a short polynucleotide that is used to detect a polynucleotide sequence, that is complementary to the probe, in a hybridization-based assay. The probe may consist of a “fragment” of a polynucleotide as defined herein. Preferably such a probe is at least 5, more preferably at least 10, more preferably at least 20, more preferably at least 30, more preferably at least 40, more preferably at least 50, more preferably at least 100, more preferably at least 200, more preferably at least 300, more preferably at least 400 and most preferably at least 500 nucleotides in length.

2. Polypeptides and fragments

[0098] The term “polypeptide”, as used herein, encompasses amino acid chains of any length but preferably at least 5 amino acids, including full-length proteins, in which amino acid residues are linked by covalent peptide bonds. Polypeptides as herein disclosed may be purified natural products, or may be produced partially or wholly using recombinant or synthetic techniques. The term may refer to a polypeptide, an aggregate of a polypeptide such as a dimer or other multimer, a fusion polypeptide, a polypeptide fragment, a polypeptide variant, or derivative thereof.

[0099] A “fragment” of a polypeptide is a subsequence of the polypeptide that performs a function that is required for the biological activity and/or provides three dimensional structure of the polypeptide. The term may refer to a polypeptide, an aggregate of a polypeptide such as a dimer or other multimer, a fusion polypeptide, a polypeptide fragment, a polypeptide variant, or derivative thereof capable of performing the above enzymatic activity.

[00100] The term “isolated” as applied to the polynucleotide or polypeptide sequences disclosed herein is used to refer to sequences that are removed from their natural cellular environment. An isolated molecule may be obtained by any method or combination of methods including biochemical, recombinant, and synthetic techniques.

[00101] The term “recombinant” refers to a polynucleotide sequence that is removed from sequences that surround it in its natural context and/or is recombined with sequences that are not present in its natural context.

[00102] A "recombinant" polypeptide sequence is produced by translation from a "recombinant" polynucleotide sequence.

[00103] The term "derived from" with respect to polynucleotides or polypeptides as disclosed herein being derived from a particular genera or species, means that the polynucleotide or polypeptide has the same sequence as a polynucleotide or polypeptide found naturally in that genera or species. The polynucleotide or polypeptide, derived from a particular genera or species, may therefore be produced synthetically or recombinantly.

3. Variants

[00104] As used herein, the term "variant" refers to polynucleotide or polypeptide sequences different from the specifically identified sequences, wherein one or more nucleotides or amino acid residues is deleted, substituted, or added. Variants may be naturally occurring allelic variants, or non-naturally occurring variants. Variants may be from the same or from other species and may encompass homologues, paralogues and orthologues. Variants described herein can also be created via site-directed mutagenesis of the coding sequence for a polypeptide, or by combining domains from the coding sequences for different naturally-occurring polypeptides ("domain swapping"). Techniques for modifying genes encoding functional polypeptides described herein are known and include, *inter alia*, directed evolution techniques, site-directed mutagenesis techniques and random mutagenesis techniques, and can be useful to increase specific activity of a polypeptide, alter substrate specificity, alter expression levels, alter subcellular location, or modify polypeptide-polypeptide interactions in a desired manner. Such modified polypeptides are considered variants.

[00105] In certain embodiments, variants of the polynucleotides and polypeptides disclosed herein possess biological activities that are the same or similar to those of the polynucleotides or polypeptides disclosed herein. The term "variant" with reference to polynucleotides and polypeptides encompasses all forms of polynucleotides and polypeptides as defined herein.

4. Polynucleotide variants

[00106] Variant polynucleotide sequences preferably exhibit at least 70%, more preferably at least 71%, more preferably at least 72%, more preferably at least 73%, more preferably at least 74%, more preferably at least 75%, more preferably at least 76%, more preferably at least 77%, more preferably at least 78%, more preferably at least 79%, more preferably at least 80%, more preferably at least 81%, more preferably at least 82%, more preferably at least 83%, more preferably at least 84%, more preferably at least 85%, more preferably at least 86%, more preferably at least 87%, more preferably at least 88%, more preferably at least 89%, more preferably at least

90%, more preferably at least 91%, more preferably at least 92%, more preferably at least 93%, more preferably at least 94%, more preferably at least 95%, more preferably at least 96%, more preferably at least 97%, more preferably at least 98%, and most preferably at least 99% identity to a sequence as disclosed herein. Identity is found over a comparison window of at least 20 nucleotide positions, preferably at least 50 nucleotide positions, more preferably at least 100 nucleotide positions, more preferably at least 200 nucleotide positions, more preferably at least 300 nucleotide positions, more preferably at least 400 nucleotide positions, more preferably at least 500 nucleotide positions, and most preferably over the entire length of a polynucleotide disclosed herein.

- 10 **[00107]** Polynucleotide sequence identity can be determined in the following manner. The subject polynucleotide sequence is compared to a candidate polynucleotide sequence using BLASTN (from the BLAST suite of programs, version 2.2.5 [Nov 2002]) in bl2seq (Tatiana A. Tatusova, Thomas L. Madden (1999), "Blast 2 sequences - a new tool for comparing protein and nucleotide sequences", FEMS Microbiol Lett. 174:247-250), which is publicly available from NCBI (<ftp://ftp.ncbi.nih.gov/blast/>). The default parameters of bl2seq are utilized except that filtering of low complexity parts should be turned off.

[00108] The identity of polynucleotide sequences may be examined using the following unix command line parameters:

```
bl2seq -i nucleotideseq1 -j nucleotideseq2 -F F -p blastn
```

- 20 **[00109]** The parameter -F F turns off filtering of low complexity sections. The parameter -p selects the appropriate algorithm for the pair of sequences. The bl2seq program reports sequence identity as both the number and percentage of identical nucleotides in a line "Identities = ".

- 25 **[00110]** Polynucleotide sequence identity may also be calculated over the entire length of the overlap between a candidate and subject polynucleotide sequences using global sequence alignment programs (e.g. Needleman, S. B. and Wunsch, C. D. (1970) J. Mol. Biol. 48, 443-453). A full implementation of the Needleman-Wunsch global alignment algorithm is found in the needle program in the EMBOSS package (Rice, P. Longden, I. and Bleasby, A. EMBOSS: The European Molecular Biology Open Software Suite, Trends in Genetics June 2000, vol 16, No 6. pp.276-277) which can be obtained from <http://www.hgmp.mrc.ac.uk/Software/EMBOSS/>. The European Bioinformatics Institute server also provides the facility to perform EMBOSS-needle global alignments between two sequences online at <http://www.ebi.ac.uk/emboss/align/>.

- 35 **[00111]** Alternatively the GAP program may be used which computes an optimal global alignment of two sequences without penalizing terminal gaps. GAP is described in

the following paper: Huang, X. (1994) On Global Sequence Alignment. Computer Applications in the Biosciences 10, 227-235.

[00112] Another method for calculating polynucleotide % sequence identity is based on aligning sequences to be compared using Clustal X (Jeanmougin *et al.*, 1998, *Trends Biochem. Sci.* **23**, 403-5.)

[00113] Polynucleotide variants disclosed herein also encompass those which exhibit a similarity to one or more of the specifically identified sequences that is likely to preserve the functional equivalence of those sequences and which could not reasonably be expected to have occurred by random chance. Such sequence similarity with respect to polypeptides may be determined using the publicly available bl2seq program from the BLAST suite of programs described *supra*.

[00114] The similarity of polynucleotide sequences may be examined using the following unix command line parameters:

bl2seq -i nucleotideseq1 -j nucleotideseq2 -F F -p tblastx

[00115] The parameter -F F turns off filtering of low complexity sections. The parameter -p selects the appropriate algorithm for the pair of sequences. This program finds regions of similarity between the sequences and for each such region reports an "E value" which is the expected number of times one could expect to see such a match by chance in a database of a fixed reference size containing random sequences. The size of this database is set by default in the bl2seq program. For small E values, much less than one, the E value is approximately the probability of such a random match.

[00116] Variant polynucleotide sequences preferably exhibit an E value of less than 1×10^{-10} more preferably less than 1×10^{-20} , more preferably less than 1×10^{-30} , more preferably less than 1×10^{-40} , more preferably less than 1×10^{-50} , more preferably less than 1×10^{-60} , more preferably less than 1×10^{-70} , more preferably less than 1×10^{-80} , more preferably less than 1×10^{-90} and most preferably less than 1×10^{-100} when compared with any one of the specifically identified sequences.

[00117] Alternatively, variant polynucleotides as disclosed herein hybridize to the specified polynucleotide sequences, or complements thereof under stringent conditions.

[00118] The term "hybridize under stringent conditions", and grammatical equivalents thereof, refers to the ability of a polynucleotide molecule to hybridize to a target polynucleotide molecule (such as a target polynucleotide molecule immobilized on a DNA or RNA blot, such as a Southern blot or Northern blot) under defined conditions of

temperature and salt concentration. The ability to hybridize under stringent hybridization conditions can be determined by initially hybridizing under less stringent conditions then increasing the stringency to the desired stringency.

[00119] With respect to polynucleotide molecules greater than about 100 bases in length, typical stringent hybridization conditions are no more than 25 to 30° C (for example, 10° C) below the melting temperature (T_m) of the native duplex (see generally, Sambrook *et al.*, Eds, 1987, Molecular Cloning, A Laboratory Manual, 2nd Ed. Cold Spring Harbor Press; Ausubel *et al.*, 1987, Current Protocols in Molecular Biology, Greene Publishing,). T_m for polynucleotide molecules greater than about 100 bases can be calculated by the formula $T_m = 81.5 + 0.41\% (G + C - \log (Na^+))$. (Sambrook *et al.*, Eds, 1987, Molecular Cloning, A Laboratory Manual, 2nd Ed. Cold Spring Harbor Press; Bolton and McCarthy, 1962, PNAS 84:1390). Typical stringent conditions for polynucleotide of greater than 100 bases in length would be hybridization conditions such as prewashing in a solution of 6X SSC, 0.2% SDS; hybridizing at 65°C, 6X SSC, 0.2% SDS overnight; followed by two washes of 30 minutes each in 1X SSC, 0.1% SDS at 65° C and two washes of 30 minutes each in 0.2X SSC, 0.1% SDS at 65°C.

[00120] With respect to polynucleotide molecules having a length less than 100 bases, exemplary stringent hybridization conditions are 5 to 10°C below T_m . On average, the T_m of a polynucleotide molecule of length less than 100 bp is reduced by approximately (500/oligonucleotide length)°C.

[00121] With respect to the DNA mimics known as peptide nucleic acids (PNAs) (Nielsen *et al.*, Science. 1991 Dec 6;254(5037):1497-500) T_m values are higher than those for DNA-DNA or DNA-RNA hybrids, and can be calculated using the formula described in Giesen *et al.*, Nucleic Acids Res. 1998 Nov 1;26(21):5004-6. Exemplary stringent hybridization conditions for a DNA-PNA hybrid having a length less than 100 bases are 5 to 10° C below the T_m .

[00122] Variant polynucleotides as disclosed herein also encompass polynucleotides that differ from the sequences as herein disclosed but that, as a consequence of the degeneracy of the genetic code, encode a polypeptide having similar activity to a polypeptide encoded by a polynucleotide of the present invention. A sequence alteration that does not change the amino acid sequence of the polypeptide is a "silent variation". Except for ATG (methionine) and TGG (tryptophan), other codons for the same amino acid may be changed by art recognized techniques, e.g., to optimize codon usage in a particular host organism.

[00123] Polynucleotide sequence alterations resulting in conservative substitutions of one or several amino acids in the encoded polypeptide sequence without significantly altering its biological activity are also included in the invention. A skilled artisan will be aware of methods for making phenotypically silent amino acid substitutions (see, e.g.,
5 Bowie *et al.*, 1990, Science 247, 1306).

[00124] Variant polynucleotides due to silent variations and conservative substitutions in the encoded polypeptide sequence may be determined using the publicly available bl2seq program from the BLAST suite of programs (version 2.2.5 [Nov 2002]) from NCBI (<ftp://ftp.ncbi.nih.gov/blast/>) via the tblastx algorithm as previously
10 described.

[00125] The function of the polypeptide encoded by a variant polynucleotide disclosed herein as modifier of GGP translation, production and/or activity may be assessed for example by expressing such a sequence in a host cell and testing its activity as described herein in the Examples. Function of a variant may also be tested for its
15 ability to alter AsA content in plants, also as described in the Examples section herein.

5. Polypeptide variants

[00126] The term "variant" with reference to polypeptides encompasses naturally occurring, recombinantly and synthetically produced polypeptides. Variant polypeptide sequences preferably exhibit at least 70%, more preferably at least 71%, more
20 preferably at least 72%, more preferably at least 73%, more preferably at least 74%, more preferably at least 75%, more preferably at least 76%, more preferably at least 77%, more preferably at least 78%, more preferably at least 79%, more preferably at least 80%, more preferably at least 81%, more preferably at least 82%, more preferably at least 83%, more preferably at least 84%, more preferably at least 85%, more
25 preferably at least 86%, more preferably at least 87%, more preferably at least 88%, more preferably at least 89%, more preferably at least 90%, more preferably at least 91%, more preferably at least 92%, more preferably at least 93%, more preferably at least 94%, more preferably at least 95%, more preferably at least 96%, more preferably at least 97%, more preferably at least 98%, and most preferably at least 99% identity to
30 a sequences of the present invention. Identity is found over a comparison window of at least 20 amino acid positions, preferably at least 50 amino acid positions, more preferably at least 100 amino acid positions, and most preferably over the entire length of a polypeptide as herein disclosed.

[00127] Polypeptide sequence identity can be determined in the following manner.
35 The subject polypeptide sequence is compared to a candidate polypeptide sequence using BLASTP (from the BLAST suite of programs, version 2.2.5 [Nov 2002]) in bl2seq, which is

publicly available from NCBI (<ftp://ftp.ncbi.nih.gov/blast/>). The default parameters of bl2seq are utilized except that filtering of low complexity regions should be turned off.

[00128] Polypeptide sequence identity may also be calculated over the entire length of the overlap between a candidate and subject polypeptide sequences using global sequence alignment programs. EMBOSS-needle (available at <http://www.ebi.ac.uk/emboss/align/>) and GAP (Huang, X. (1994) On Global Sequence Alignment. *Computer Applications in the Biosciences* 10, 227-235.) as discussed above are also suitable global sequence alignment programs for calculating polypeptide sequence identity.

10 [00129] Another method for calculating polypeptide % sequence identity is based on aligning sequences to be compared using Clustal X (Jeanmougin *et al.*, 1998, *Trends Biochem. Sci.* **23**, 403-5.)

[00130] Polypeptide variants as disclosed herein also encompass those which exhibit a similarity to one or more of the specifically identified sequences that is likely to preserve the functional equivalence of those sequences and which could not reasonably be expected to have occurred by random chance. Such sequence similarity with respect to polypeptides may be determined using the publicly available bl2seq program from the BLAST suite of programs (version 2.2.5 [Nov 2002]) from NCBI (<ftp://ftp.ncbi.nih.gov/blast/>). The similarity of polypeptide sequences may be examined using the following unix command line parameters:

bl2seq -i peptideseq1 -j peptideseq2 -F F -p blastp

[00131] The parameter -F F turns off filtering of low complexity sections. The parameter -p selects the appropriate algorithm for the pair of sequences. This program finds regions of similarity between the sequences and for each such region reports an "E value" which is the expected number of times one could expect to see such a match by chance in a database of a fixed reference size containing random sequences. For small E values, much less than one, this is approximately the probability of such a random match.

[00132] Variant polypeptide sequences preferably exhibit an E value of less than 1×10^{-10} more preferably less than 1×10^{-20} , more preferably less than 1×10^{-30} , more preferably less than 1×10^{-40} , more preferably less than 1×10^{-50} , more preferably less than 1×10^{-60} , more preferably less than 1×10^{-70} , more preferably less than 1×10^{-80} , more preferably less than 1×10^{-90} and most preferably less than 1×10^{-100} when compared with any one of the specifically identified sequences.

[00133] Conservative substitutions of one or several amino acids of a described polypeptide sequence without significantly altering its biological activity are also included in the invention. A skilled artisan will be aware of methods for making phenotypically silent amino acid substitutions (see, e.g., Bowie *et al.*, 1990, *Science* 247, 1306).

5 [00134] The function of a variant polypeptide disclosed herein as modifier of GGP translation, production and/or activity may be assessed for example by expressing such a sequence in a host cell and testing its activity as described herein in the Examples. Function of a variant may also be tested for its ability to alter AsA content in plants, also as described in the Examples section herein.

10 [00135] In various embodiments the polypeptide or polypeptide variant may comprise one or more SANT/MYB domains. A SANT domain is a protein domain that allows many chromatin remodelling proteins to interact with histones (Boyer *et al.*, 2002, *Molecular Cell* 10(4): 935-942; Boyer *et al.*, 2004, *Nature Reviews Molecular Cell Biology* 5(2): 158-163). SANT domains have an acidic predicted isoelectric point (pI), whereas
15 MYB domains have a basic pI (Ko *et al.*, 2008, *Molecular Cancer* 7(1): 77). MYBS1 proteins are predicted to have two SANT/MYB domains. For example, MYBS1 from *Actinidia eriantha* (SEQ ID No:103) has an N-terminal SANT domain with a predicted acidic isoelectric point (pI), and a second SANT/MYB region with a predicted basic pI.

6. Methods for identifying variants

20 *Physical methods*

[00136] Variant polypeptides may be identified using PCR-based methods (Mullis *et al.*, Eds. 1994 *The Polymerase Chain Reaction*, Birkhauser). Typically, the polynucleotide sequence of a primer, useful to amplify variants of polynucleotide molecules as disclosed herein by PCR, may be based on a sequence encoding a conserved region of the
25 corresponding amino acid sequence.

[00137] Alternatively library screening methods, well known to those skilled in the art, may be employed (Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, 2nd Ed. Cold Spring Harbor Press, 1987). When identifying variants of the probe sequence, hybridization and/or wash stringency will typically be reduced relatively to when exact
30 sequence matches are sought.

[00138] Polypeptide variants may also be identified by physical methods, for example by screening expression libraries using antibodies raised against polypeptides disclosed herein (Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, 2nd Ed. Cold Spring Harbor Press, 1987) or by identifying polypeptides from natural sources with the aid of
35 such antibodies.

Computer based methods

[00139] The variant sequences as disclosed herein, including both polynucleotide and polypeptide variants, may also be identified by computer-based methods well-known to those skilled in the art, using public domain sequence alignment algorithms and sequence similarity search tools to search sequence databases (public domain databases include Genbank, EMBL, Swiss-Prot, PIR and others). See, e.g., Nucleic Acids Res. 29: 1-10 and 11-16, 2001 for examples of online resources. Similarity searches retrieve and align target sequences for comparison with a sequence to be analyzed (i.e., a query sequence). Sequence comparison algorithms use scoring matrices to assign an overall score to each of the alignments.

[00140] An exemplary family of programs useful for identifying variants in sequence databases is the BLAST suite of programs (version 2.2.5 [Nov 2002]) including BLASTN, BLASTP, BLASTX, tBLASTN and tBLASTX, which are publicly available from (ftp://ftp.ncbi.nih.gov/blast/) or from the National Center for Biotechnology Information (NCBI), National Library of Medicine, Building 38A, Room 8N805, Bethesda, MD 20894 USA. The NCBI server also provides the facility to use the programs to screen a number of publicly available sequence databases. BLASTN compares a nucleotide query sequence against a nucleotide sequence database. BLASTP compares an amino acid query sequence against a protein sequence database. BLASTX compares a nucleotide query sequence translated in all reading frames against a protein sequence database. tBLASTN compares a protein query sequence against a nucleotide sequence database dynamically translated in all reading frames. tBLASTX compares the six-frame translations of a nucleotide query sequence against the six-frame translations of a nucleotide sequence database. The BLAST programs may be used with default parameters or the parameters may be altered as required to refine the screen.

[00141] The use of the BLAST family of algorithms, including BLASTN, BLASTP, and BLASTX, is described in the publication of Altschul *et al.*, Nucleic Acids Res. 25: 3389-3402, 1997.

[00142] The "hits" to one or more database sequences by a queried sequence produced by BLASTN, BLASTP, BLASTX, tBLASTN, tBLASTX, or a similar algorithm, align and identify similar portions of sequences. The hits are arranged in order of the degree of similarity and the length of sequence overlap. Hits to a database sequence generally represent an overlap over only a fraction of the sequence length of the queried sequence.

[00143] The BLASTN, BLASTP, BLASTX, tBLASTN and tBLASTX algorithms also produce "Expect" values for alignments. The Expect value (E) indicates the number of

hits one can "expect" to see by chance when searching a database of the same size containing random contiguous sequences. The Expect value is used as a significance threshold for determining whether the hit to a database indicates true similarity. For example, an E value of 0.1 assigned to a polynucleotide hit is interpreted as meaning that in a database of the size of the database screened, one might expect to see 0.1 matches over the aligned portion of the sequence with a similar score simply by chance. For sequences having an E value of 0.01 or less over aligned and matched portions, the probability of finding a match by chance in that database is 1% or less using the BLASTN, BLASTP, BLASTX, tBLASTN or tBLASTX algorithm.

10 **[00144]** Multiple sequence alignments of a group of related sequences can be carried out with CLUSTALW (Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) CLUSTALW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, 22:4673-4680) CLUSTAL Omega (Sievers et al., (2011). *Molecular Systems Biology* 7:539, <https://www.ebi.ac.uk/Tools/msa/clustalo/>) or T-COFFEE (Cedric Notredame, Desmond G. Higgins, Jaap Heringa, T-Coffee: A novel method for fast and accurate multiple sequence alignment, *J. Mol. Biol.* (2000) 302: 205-217)) or PILEUP, which uses progressive, pairwise alignments. (Feng and Doolittle, 1987, *J. Mol. Evol.* 25, 351).

20 **[00145]** Pattern recognition software applications are available for finding motifs or signature sequences. For example, MEME (Multiple Em for Motif Elicitation) finds motifs and signature sequences in a set of sequences, and MAST (Motif Alignment and Search Tool) uses these motifs to identify similar or the same motifs in query sequences. The MAST results are provided as a series of alignments with appropriate statistical data and a visual overview of the motifs found. MEME and MAST were developed at the University of California, San Diego.

30 **[00146]** PROSITE (Bairoch and Bucher, 1994, *Nucleic Acids Res.* 22, 3583; Hofmann et al., 1999, *Nucleic Acids Res.* 27, 215) is a method of identifying the functions of uncharacterized proteins translated from genomic or cDNA sequences. The PROSITE database (www.expasy.org/prosite) contains biologically significant patterns and profiles and is designed so that it can be used with appropriate computational tools to assign a new sequence to a known family of proteins or to determine which known domain(s) are present in the sequence (Falquet et al., 2002, *Nucleic Acids Res.* 30, 235). Prosearch is a tool that can search SWISS-PROT and EMBL databases with a given sequence pattern or signature.

[00147] Another example of a protein domain model database is Pfam (Sonnhammer et al., 1997, A comprehensive database of protein families based on seed alignments, *Proteins*, 28: 405-420; Finn et al., 2010, The Pfam protein families database', *Nucl. Acids Res.*, 38: D211-D222). "Pfam" refers to a large collection of protein domains and protein families maintained by the Pfam Consortium and available at several sponsored world wide web sites, including: pfam.xfam.org/ (European Bioinformatics Institute (EMBL-EBI). The latest release of Pfam is Pfam 35.0 (November 2021). Pfam domains and families are identified using multiple sequence alignments and hidden Markov models (HMMs). Pfam-A family or domain assignments, are high quality assignments generated by a curated seed alignment using representative members of a protein family and profile hidden Markov models based on the seed alignment. (Unless otherwise specified, matches of a queried protein to a Pfam domain or family are Pfam-A matches.) All identified sequences belonging to the family are then used to automatically generate a full alignment for the family (Sonnhammer (1998) *Nucleic Acids Research* 26, 320-322; Bateman (2000) *Nucleic Acids Research* 26, 263-266; Bateman (2004) *Nucleic Acids Research* 32, Database Issue, D138-D141; Finn (2006) *Nucleic Acids Research Database Issue* 34, D247-251; Finn (2010) *Nucleic Acids Research Database Issue* 38, D21 1-222). By accessing the Pfam database, for example, using the above-referenced website, protein sequences can be queried against the HMMs using HMMER homology search software {e.g., HMMER2, HMMER3, or a higher version, hmmer.org). Significant matches that identify a queried protein as being in a pfam family (or as having a particular Pfam domain) are those in which the bit score is greater than or equal to the gathering threshold for the Pfam domain. Expectation values (e values) can also be used as a criterion for inclusion of a queried protein in a Pfam or for determining whether a queried protein has a particular Pfam domain, where low e values (much less than 1.0, for example less than 0.1, or less than or equal to 0.01) represent low probabilities that a match is due to chance.

7. Methods for isolating or producing polynucleotides

[00148] The polynucleotide molecules disclosed herein can also be isolated by using a variety of techniques known to those of ordinary skill in the art. By way of example, such polynucleotides can be isolated through use of the polymerase chain reaction (PCR) described in Mullis *et al.*, Eds. 1994 *The Polymerase Chain Reaction*, Birkhauser, incorporated herein by reference. The polynucleotides as herein disclosed can be amplified using primers, as defined herein, derived from the polynucleotide sequences as herein disclosed.

[00149] Further methods for isolating polynucleotides as disclosed herein include use of all, or portions of, the polynucleotides having the sequence set forth herein as

hybridization probes. The technique of hybridizing labelled polynucleotide probes to polynucleotides immobilized on solid supports such as nitrocellulose filters or nylon membranes, can be used to screen the genomic or cDNA libraries. Exemplary hybridization and wash conditions are: hybridization for 20 hours at 65°C in 5.0 X SSC, 0.5% sodium dodecyl sulfate, 1 X Denhardt's solution; washing (three washes of twenty minutes each at 55°C) in 1.0 X SSC, 1% (w/v) sodium dodecyl sulfate, and optionally one wash (for twenty minutes) in 0.5 X SSC, 1% (w/v) sodium dodecyl sulfate, at 60°C. An optional further wash (for twenty minutes) can be conducted under conditions of 0.1 X SSC, 1% (w/v) sodium dodecyl sulfate, at 60°C.

- 10 **[00150]** The polynucleotide fragments as disclosed herein may be produced by techniques well-known in the art such as restriction endonuclease digestion, oligonucleotide synthesis and PCR amplification.

- 15 **[00151]** A partial polynucleotide sequence may be used, in methods well-known in the art to identify the corresponding full-length polynucleotide sequence. Such methods include PCR-based methods, 5'RACE (Frohman MA, 1993, *Methods Enzymol.* 218: 340-56) and hybridization-based method, computer/database-based methods. Further, by way of example, inverse PCR permits acquisition of unknown sequences, flanking the polynucleotide sequences disclosed herein, starting with primers based on a known region (Triglia *et al.*, 1998, *Nucleic Acids Res* 16, 8186, incorporated herein by reference). The method uses several restriction enzymes to generate a suitable fragment in the known region of a gene. The fragment is then circularized by intramolecular ligation and used as a PCR template. Divergent primers are designed from the known region. In order to physically assemble full-length clones, standard molecular biology approaches can be utilized (Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, 20 2nd Ed. Cold Spring Harbor Press, 1987).

- 25 **[00152]** It may be beneficial, when producing a transgenic plant from a particular species, to transform such a plant with a sequence or sequences derived from that species. The benefit may be to alleviate public concerns regarding cross-species transformation in generating transgenic organisms. Additionally when down-regulation of a gene is the desired result, it may be necessary to utilise a sequence identical (or at least highly similar) to that in the plant, for which reduced expression is desired. For these reasons among others, it is desirable to be able to identify and isolate orthologues of a particular gene in several different plant species.

- 30 **[00153]** Variants (including orthologues) may be identified by the methods described herein.

8. Methods for isolating or producing polypeptides

[00154] The polypeptides as disclosed herein, including variant polypeptides, may be prepared using peptide synthesis methods well known in the art such as direct peptide synthesis using solid phase techniques (e.g. Stewart *et al.*, 1969, in Solid-Phase Peptide Synthesis, WH Freeman Co, San Francisco California), or automated synthesis, for example using an Applied Biosystems 431A Peptide Synthesizer (Foster City, California). Mutated forms of the polypeptides may also be produced during such syntheses.

[00155] The polypeptides and variant polypeptides as disclosed herein may also be purified from natural sources using a variety of techniques that are well known in the art (e.g. Deutscher, 1990, Ed, Methods in Enzymology, Vol. 182, *Guide to Protein Purification*,).

[00156] Alternatively the polypeptides and variant polypeptides as disclosed herein may be expressed recombinantly in suitable host cells as disclosed herein and separated from the cells as discussed below.

9. Constructs, vectors and components thereof

[00157] According to one embodiment, the polynucleotides useful in the methods according to some embodiments of the invention may be provided in a nucleic acid construct useful in transforming a plant or host cell. Suitable plant and host cells are described herein.

[00158] The term "genetic construct" refers to a polynucleotide molecule, usually double-stranded DNA, which may have inserted into it another polynucleotide molecule (the insert polynucleotide molecule) such as, but not limited to, a cDNA molecule. A genetic construct may contain the necessary elements that permit transcribing the insert polynucleotide molecule, and, optionally, translating the transcript into a polypeptide.

The insert polynucleotide molecule may be derived from the host cell, or may be derived from a different cell or organism and/or may be a synthetic or recombinant polynucleotide. Once inside the host cell the genetic construct may become integrated in the host chromosomal DNA. The genetic construct may be linked to a vector.

[00159] The term "vector" refers to a polynucleotide molecule, usually double stranded DNA, which is used to transport the genetic construct into a host cell. The vector may be capable of replication in at least one additional host system, such as *E. coli*.

[00160] The term "expression construct" refers to a genetic construct that includes the necessary regulatory elements that permit transcribing the insert polynucleotide

molecule, and, optionally, translating the transcript into a polypeptide. An expression construct typically comprises in a 5' to 3' direction:

- a) a promoter functional in the host cell into which the construct will be transformed,
- 5 b) the polynucleotide to be expressed, and
- c) a terminator functional in the host cell into which the construct will be transformed.

[00161] The term "coding region" or "open reading frame" (ORF) refers to the sense strand of a genomic DNA sequence or a cDNA sequence that is capable of producing a transcription product and/or a polypeptide under the control of appropriate regulatory sequences. The coding sequence is identified by the presence of a 5' translation start codon and a 3' translation stop codon. When inserted into a genetic construct, a "coding sequence" is capable of being expressed when it is operably linked to promoter and terminator sequences.

15 **[00162]** 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.

[00163] "Operably-linked" means that the sequence to be expressed is placed under the control of regulatory elements that include promoters, tissue-specific regulatory elements, temporal regulatory elements, enhancers, repressors and terminators. 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.

25 **[00164]** "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 can 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
5 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.

[00165] 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
10 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 can be present, e.g., introns, enhancers, upstream activation regions, transcription terminators, and inducible elements.

15 [00166] The term "noncoding region" includes to untranslated sequences that are upstream of the translational start site and downstream of the translational stop site. These sequences are also referred to respectively as the 5' UTR and the 3' UTR. These sequences may include elements required for transcription initiation and termination and for regulation of translation efficiency. The term "noncoding" also includes intronic
20 sequences within genomic clones.

[00167] Terminators are sequences, which terminate transcription, and are found in the 3' untranslated ends of genes downstream of the translated sequence. Terminators are important determinants of mRNA stability and in some cases have been found to have spatial regulatory functions.

25 [00168] The term "promoter" refers to nontranscribed cis-regulatory elements upstream of the coding region that regulate gene transcription. Promoters comprise cis-initiator elements which specify the transcription initiation site and conserved boxes such as the TATA box, and motifs that are bound by transcription factors.

[00169] A "transgene" is a polynucleotide that is taken from one organism and
30 introduced into a different organism by transformation. The transgene may be derived from the same species or from a different species as the species of the organism into which the transgene is introduced.

[00170] An "inverted repeat" is a sequence that is repeated, where the second half of the repeat is in the complementary strand, e.g.,

(5')GATCTA.....TAGATC(3')
 (3')CTAGAT.....ATCTAG(5')

[00171] Read-through transcription will produce a transcript that undergoes complementary base-pairing to form a hairpin structure provided that there is a 3-5 bp spacer between the repeated regions.

10. Methods for producing constructs and vectors

[00172] The genetic constructs as disclosed herein comprise one or more polynucleotide sequences as disclosed herein and/or polynucleotides encoding polypeptides as disclosed herein, and may be useful for transforming, for example, bacterial, fungal, insect, mammalian or plant organisms. The genetic constructs disclosed herein are intended to include expression constructs as herein defined.

[00173] Methods for producing and using genetic constructs and vectors are well known in the art and are described generally in Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, 2nd Ed. Cold Spring Harbor Press, 1987; Ausubel *et al.*, Current Protocols in Molecular Biology, Greene Publishing, 1987).

11. Host cells

[00174] In other embodiments, there is provided a host cell which comprises a genetic construct or vector as disclosed herein. In preferred embodiments, the host cell is genetically modified to express one of:

- a) a polynucleotide encoding a polypeptide with the amino acid sequence of any one of SEQ ID NO: 1-102 and 348-356, or a variant of the polypeptide, or
- b) a polynucleotide comprising a nucleotide sequence selected from any one of the sequences SEQ ID NO: 103-108, or a variant thereof,

and/or, one of

- c) a polynucleotide encoding a polypeptide with the amino acid sequence of any one of SEQ ID NO: 109 to 209, or a variant of the polypeptide, or
- d) a polynucleotide comprising a nucleotide sequence of SEQ ID NO: 210, or a variant thereof.

[00175] Host cells comprising genetic constructs, such as expression constructs, as disclosed herein are useful in methods well known in the art (e.g. Sambrook *et al.*, Molecular Cloning : A Laboratory Manual, 2nd Ed. Cold Spring Harbor Press, 1987 ; Ausubel *et al.*, Current Protocols in Molecular Biology, Greene Publishing, 1987) for

recombinant production of polypeptides disclosed herein. Such methods may involve the culture of host cells in an appropriate medium in conditions suitable for or conducive to expression of a polynucleotide or polypeptide disclosed herein. The expressed recombinant polypeptide, which may optionally be secreted into the culture, may then be separated from the medium, host cells or culture medium by methods well known in the art (e.g. Deutscher, Ed, 1990, Methods in Enzymology, Vol 182, Guide to Protein Purification).

12. Methods for producing plant cells and plants comprising constructs and vectors

[00176] In other embodiments there is provided a plant cell which comprises a genetic construct as disclosed herein, and a plant cell modified to alter expression of a polynucleotide or polypeptide as disclosed herein. Plants comprising such cells are also provided.

[00177] Alteration of GGP translation, production and/or activity may be altered in a plant through methods according to some embodiments of the invention. Such methods may involve the transformation of plant cells and plants, with a construct designed to alter expression of a polynucleotide or polypeptide that alters GGP expression, and/or AsA content in such plant cells and plants. Such methods also include the transformation of plant cells and plants with a combination of a construct as disclosed herein and one or more other constructs designed to alter expression of one or more polynucleotides or polypeptides which modulate GGP activity and/or AsA content in such plant cells and plants.

[00178] Methods for transforming plant cells, plants and portions thereof with polypeptides are described in Draper *et al.*, 1988, Plant Genetic Transformation and Gene Expression. A Laboratory Manual. Blackwell Sci. Pub. Oxford, p. 365; Potrykus and Spangenburg, 1995, Gene Transfer to Plants. Springer-Verlag, Berlin.; and Gelvin *et al.*, 1993, Plant Molecular Biol. Manual. Kluwer Acad. Pub. Dordrecht. A review of transgenic plants, including transformation techniques, is provided in Galun and Breiman, 1997, Transgenic Plants. Imperial College Press, London.

[00179] Transformation may be transient or stable, as is known in the art. Transient transformation results in the temporary introduction of nucleic acid into a cell, however the introduced nucleic acid does not integrate into the cell's genome. Stable transformation results in modification of the cell's genome, which will persist and may be passed on to subsequent generations of the cell. Preferably the plant cells and plants are stably transformed.

[00180] In some embodiments, the plant cell to be transformed comprises an endogenous GGP gene, preferably a functional GGP gene. For example, in some embodiments, the plant cell to be transformed has been previously transformed with a functional GGP gene. In other embodiments, the plant cell is co-transformed or subsequently transformed with a functional GGP gene. A functional GGP gene is a gene that encodes and expresses a functional GGP protein. A functional GGP protein is one that is capable of performing one or more functions of a GGP protein, such as catalysing the conversion of GDP-L-galactose to L-galactose 1-phosphate.

[00181] In some embodiments the GGP gene is a GGP3 gene. In some embodiments, the GGP gene encodes a protein with at least about 70% sequence identity to SEQ ID No: 365, preferably at least about 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, at least about 99%, or 100% sequence identity to SEQ ID No: 365.

[00182] In some embodiments, the GGP gene has a promoter that comprises one or more MYBS1 binding sites, preferably two or more MYBS1 binding sites. In some embodiments, the methods described herein may further comprise gene-editing the promoter of an endogenous GGP gene to provide a promoter comprising one or more MYBS1 binding sites. This may be achieved, for example, by modifying an endogenous GGP promoter. Alternatively, this may be achieved by replacing an endogenous promoter with a GGP promoter comprising one or more MYBS1 binding sites, such as SEQ ID No: 345. In some embodiments, the methods described herein may further comprise transforming the plant cell with a GGP gene. In some embodiments, the methods may further comprise transforming the plant cell with a GGP gene having a promoter that comprises one or more MYBS1 binding sites, preferably two or more MYBS1 binding sites.

[00183] An MYBS1 binding site is any nucleotide sequence that an MYBS1 protein is capable of binding to. MYBS1 binding sites may be determined by a number of methods known in the art, for example using bioinformatic prediction (using, for example, JASPAR 2020 (Fornes *et al.*, 2019, *Nucleic Acids Research* **48**(D1): D87-D92)) or using an electrophoretic mobility shift assay (EMSA). Some exemplary methods for identifying MYBS1 binding sites are presented in Example 4.

[00184] Preferably the one or more MYBS1 binding sites comprise the sequence TCTTATC or its reverse complement GATAAGA. In some embodiments, the GGP promoter has at least about 70% sequence identity to SEQ ID No: 345, preferably at least about 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%,

85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, at least about 99%, or 100% sequence identity to SEQ ID No: 345.

Methods for genetic manipulation of plants

[00185] A number of plant transformation strategies are available (e.g. Birch, 1997, Ann Rev Plant Phys Plant Mol Biol, 48, 297, Hellens RP, et al (2000) Plant Mol Biol 42: 819-32, Hellens R et al Plant Meth 1: 13). For example, strategies may be designed to increase expression of a polynucleotide/polypeptide in a plant cell, organ and/or at a particular developmental stage where/when it is normally expressed or to ectopically express a polynucleotide/polypeptide in a cell, tissue, organ and/or at a particular developmental stage which/when it is not normally expressed. The expressed polynucleotide/polypeptide may be derived from the plant species to be transformed or may be derived from a different plant species.

[00186] Genetic constructs for expression of genes in transgenic plants typically include promoters for driving the expression of one or more cloned polynucleotide, terminators and selectable marker sequences to detect presence of the genetic construct in the transformed plant.

[00187] The promoters suitable for use in the constructs as described herein are functional in a cell, tissue or organ of a monocot or dicot plant and include cell-, tissue- and organ-specific promoters, cell cycle specific promoters, temporal promoters, inducible promoters, constitutive promoters that are active in most plant tissues, and recombinant promoters. Choice of promoter will depend upon the temporal and spatial expression of the cloned polynucleotide, so desired. The promoters may be those normally associated with a transgene of interest, or promoters which are derived from genes of other plants, viruses, and plant pathogenic bacteria and fungi. Those skilled in the art will, without undue experimentation, be able to select promoters that are suitable for use in modifying and modulating plant traits using genetic constructs comprising the polynucleotide sequences as herein disclosed. Examples of constitutive plant promoters include the CaMV 35S promoter, the nopaline synthase promoter and the octopine synthase promoter, and the Ubi 1 promoter from maize. Plant promoters which are active in specific tissues, respond to internal developmental signals or external abiotic or biotic stresses are described in the scientific literature. Exemplary promoters are described, e.g., in WO 02/00894, which is herein incorporated by reference.

[00188] In some embodiments, the promoter is active in leaves and/or roots. In some embodiments, the promoter is a tissue-specific promoter that is active in leaves and/or roots. In a preferred embodiment, the promoter is active in leaves. Promoters

that are active in leaves and/or roots, including tissue-specific promoters, are known in the art.

[00189] Exemplary terminators that are commonly used in plant transformation genetic construct include, e.g., the cauliflower mosaic virus (CaMV) 35S terminator, the *Agrobacterium tumefaciens* nopaline synthase or octopine synthase terminators, the *Zea mays* zein gene terminator, the *Oryza sativa* ADP-glucose pyrophosphorylase terminator and the *Solanum tuberosum* PI-II terminator.

[00190] Selectable markers commonly used in plant transformation include the neomycin phosphotransferase II gene (NPT II) which confers kanamycin resistance, the *aadA* gene, which confers spectinomycin and streptomycin resistance, the phosphinothricin acetyl transferase (*bar* gene) for Ignite (AgrEvo) and Basta (Hoechst) resistance, and the hygromycin phosphotransferase gene (*hpt*) for hygromycin resistance.

[00191] Use of genetic constructs comprising reporter genes (coding sequences which express an activity that is foreign to the host, usually an enzymatic activity and/or a visible signal (e.g., luciferase, GUS, GFP) which may be used for promoter expression analysis in plants and plant tissues are also contemplated. The reporter gene literature is reviewed in Herrera-Estrella *et al.*, 1993, Nature 303, 209, and Schrott, 1995, In: Gene Transfer to Plants (Potrykus, T., Spangenberg. Eds) Springer Verlag. Berlin, pp. 325-336.

[00192] The following are representative publications disclosing genetic transformation protocols that can be used to genetically transform the following plant species: Rice (Alam *et al.*, 1999, Plant Cell Rep. 18, 572); apple (Yao *et al.*, 1995, Plant Cell Reports 14, 407-412); maize (US Patent Serial Nos. 5, 177, 010 and 5, 981, 840); wheat (Ortiz *et al.*, 1996, Plant Cell Rep. 15, 1996, 877); tomato (US Patent Serial No. 5, 159, 135); potato (Kumar *et al.*, 1996 Plant J. 9, : 821); cassava (Li *et al.*, 1996 Nat. Biotechnology 14, 736); lettuce (Micheltmore *et al.*, 1987, Plant Cell Rep. 6, 439); tobacco (Horsch *et al.*, 1985, Science 227, 1229); cotton (US Patent Serial Nos. 5, 846, 797 and 5, 004, 863); grasses (US Patent Nos. 5, 187, 073 and 6. 020, 539); peppermint (Niu *et al.*, 1998, Plant Cell Rep. 17, 165); citrus plants (Pena *et al.*, 1995, Plant Sci.104, 183); caraway (Krens *et al.*, 1997, Plant Cell Rep, 17, 39); banana (US Patent Serial No. 5, 792, 935); soybean (US Patent Nos. 5, 416, 011 ; 5, 569, 834 ; 5, 824, 877 ; 5, 563, 04455 and 5, 968, 830); pineapple (US Patent Serial No. 5, 952, 543); poplar (US Patent No. 4, 795, 855); monocots in general (US Patent Nos. 5, 591, 616 and 6, 037, 522); brassica (US Patent Nos. 5, 188, 958 ; 5, 463, 174 and 5, 750, 871); cereals (US Patent No. 6, 074, 877); pear (Matsuda *et al.*, 2005, Plant Cell Rep.

24(1):45-51); Prunus (Ramesh et al., 2006 Plant Cell Rep. 25(8):821-8; Song and Sink 2005 Plant Cell Rep. 2006 ;25(2):117-23; Gonzalez Padilla et al., 2003 Plant Cell Rep.22(1):38-45); strawberry (Oosumi et al., 2006 Planta. 223(6):1219-30; Folta et al., 2006 Planta Apr 14; PMID: 16614818), rose (Li et al., 2003), Rubus (Graham et al., 5 1995 Methods Mol Biol. 1995;44:129-33), tomato (Dan et al., 2006, Plant Cell Reports V25:432-441), apple (Yao et al., 1995, *Plant Cell Rep.* 14, 407-412) and *Actinidia eriantha* (Wang et al., 2006, Plant Cell Rep. 25,5: 425-31). Transformation of other species is also contemplated by the invention. Suitable methods and protocols are available in the scientific literature.

10 **[00193]** In one embodiment, there is provided a method of producing a plant cell or plant with increased L-ascorbic acid (AsA) content or increased GDP-L-galactose phosphorylase 3 (GGP) translation, production and/or activity, the method comprising upregulating in the plant cell or plant expression of one of:

15 a) a polynucleotide encoding a polypeptide with the amino acid sequence of any one of SEQ ID NO: 1-102 and 348-356, or a variant of the polypeptide, or

b) a polynucleotide comprising a nucleotide sequence selected from any one of the sequences SEQ ID NO: 103-108, or a variant thereof,

and/or, one of

20 c) a polynucleotide encoding a polypeptide with the amino acid sequence of any one of SEQ ID NO: 109 to 209, or a variant of the polypeptide, or

d) a polynucleotide comprising a nucleotide sequence of SEQ ID NO: 210, or a variant thereof.

25 **[00194]** Several methods known in the art may be employed to alter expression of a nucleotide and/or polypeptide as herein disclosed. Such methods include but are not limited to Tilling (Till *et al.*, 2003, Methods Mol Biol, 2%, 205), and so called "Deletagene" technology (Li *et al.*, 2001, Plant Journal 27(3), 235)

30 **[00195]** Other methods may involve the use of sequence-specific nucleases that generate targeted double-stranded DNA breaks in genes of interest. Examples of such methods include: zinc finger nucleases (Curtin, et al., 2011, Sander, et al., 2011), transcription activator-like effector nucleases or "TALENs" (Cermak, et al., 2011, Mahfouz, et al., 2011, Li, et al., 2012), and LAGLIDADG homing endonucleases, also termed "meganucleases" (Tzfira, et al., 2012).

[00196] Targeted genome editing using engineered nucleases such as clustered, regularly interspaced, short palindromic repeat (CRISPR) technology, is an important new approach for generating RNA-guided nucleases, such as Cas9, with customizable specificities. Genome editing mediated by these nucleases has been used to rapidly, easily and efficiently modify endogenous genes in a wide variety of biomedically important cell types and in organisms that have traditionally been challenging to manipulate genetically. A modified version of the CRISPR-Cas9 system has been developed to recruit heterologous domains that can regulate endogenous gene expression or label specific genomic loci in living cells (Sander and Joung, 2014). The technique is applicable to fungi (Nodvig, et al., 2015).

[00197] Upregulating expression of a polypeptide in a plant, for example by genome editing, can be achieved by: (i) replacing an endogenous sequence encoding the polypeptide of interest or a regulatory sequence under the control which it is placed, and/or (ii) inserting a new gene encoding the polypeptide of interest in a targeted region of the genome, and/or (iii) introducing point mutations which result in up-regulation of the endogenous gene encoding the polypeptide of interest (e.g., by altering the regulatory sequences such as promoter, enhancers, 5'-UTR and/or 3'-UTR, or mutations in the coding sequence).

[00198] In this manner, an endogenous gene encoding a polypeptide with the amino acid sequence of any one of SEQ ID NO: 1-102 and 348-356 or 109-209 or a variant of the polypeptide, or comprising a nucleotide sequence selected from any one of the sequences SEQ ID NO: 103-108 or 210 or a variant thereof, may be upregulated, resulting in increased AsA content and/or increased GGP translation, production and/or activity.

[00199] Antibodies or fragments thereof, targeted to a particular polypeptide may also be expressed in plants to modulate the activity of that polypeptide (Jobling *et al.*, 2003, Nat. Biotechnol., 21(1), 35). Transposon tagging approaches may also be applied. Additionally peptides interacting with a polypeptide as herein disclosed may be identified through technologies such as phage-display (Dyax Corporation). Such interacting peptides may be expressed in or applied to a plant to affect activity of a polypeptide as herein disclosed. Use of each of the above approaches in alteration of expression of a nucleotide and/or polypeptide as herein disclosed is specifically contemplated.

[00200] The terms "to alter expression of" and "altered expression" of a polynucleotide or polypeptide as herein disclosed, are intended to encompass the situation where genomic DNA corresponding to a polynucleotide as herein disclosed is modified thus leading to altered expression of a polynucleotide or polypeptide as herein

disclosed. Modification of the genomic DNA may be through genetic transformation or other methods known in the art for inducing mutations. The "altered expression" can be related to an increase or decrease in the amount of messenger RNA and/or polypeptide produced and may also result in altered activity of a polypeptide due to alterations in the sequence of a polynucleotide and polypeptide produced.

13. Methods of selecting plants

[00201] Methods are also provided for selecting plants with altered GGP activity or AsA content. Such methods involve testing of plants for altered expression of a polynucleotide or polypeptide as herein disclosed. Such methods may be applied at a young age or early developmental stage when the altered GGP activity or AsA content may not necessarily be easily measurable.

[00202] The expression of a polynucleotide, such as a messenger RNA, is often used as an indicator of expression of a corresponding polypeptide. Exemplary methods for measuring the expression of a polynucleotide include but are not limited to Northern analysis, RT-PCR and dot-blot analysis (Sambrook *et al.*, Molecular Cloning : A Laboratory Manual, 2nd Ed. Cold Spring Harbor Press, 1987). Polynucleotides or portions of the polynucleotides as herein disclosed are thus useful as probes or primers, as herein defined, in methods for the identification of plants with altered levels GGP or AsA. The polynucleotides as herein disclosed may be used as probes in hybridization experiments, or as primers in PCR based experiments, designed to identify such plants.

[00203] Alternatively antibodies may be raised against polypeptides as herein disclosed. Methods for raising and using antibodies are standard in the art (see for example: Antibodies, A Laboratory Manual, Harlow A Lane, Eds, Cold Spring Harbour Laboratory, 1998). Such antibodies may be used in methods to detect altered expression of the polypeptides disclosed herein. Such methods may include ELISA (Kemeny, 1991, A Practical Guide to ELISA, NY Pergamon Press) and Western analysis (Towbin & Gordon, 1994, J Immunol Methods, 72, 313).

[00204] These approaches for analysis of polynucleotide or polypeptide expression and the selection of plants with altered GGP activity or altered AsA content are useful in conventional breeding programs designed to produce varieties with altered GGP activity or AsA content.

14. Plants

[00205] The term "plant" is intended to include a whole plant, any part of a plant, propagules and progeny of a plant.

[00206] The term “propagule” means any part of a plant that may be used in reproduction or propagation, either sexual or asexual, including seeds and cuttings.

[00207] A “transgenic” or “transformed” plant refers to a plant which contains new genetic material as a result of genetic manipulation or transformation. The new genetic material may be derived from a plant of the same species as the resulting transgenic or transformed plant or from a different species. A transformed plant includes a plant which is either stably or transiently transformed with new genetic material. Preferably a transformed plant is stably transformed.

[00208] The plants and plant cells according to some embodiments of the invention may be grown and either self-ed or crossed with a different plant strain and the resulting hybrids, with the desired phenotypic characteristics, may be identified. Two or more generations may be grown to ensure that the subject phenotypic characteristics are stably maintained and inherited. Plants and plant cells resulting from such standard breeding approaches also form an aspect of the present invention.

[00209] The function of a variant polynucleotide disclosed herein as encoding a MYBS1-like or GBF3-like transcription factor may be assessed for example by expressing such a sequence in bacteria and testing activity of the encoded protein as described in the Example section herein.

[00210] Alteration of GGP activity and/or AsA content may also be altered in a plant or plant cell through methods according to some embodiments of the invention. Such methods may involve the transformation of plant cells and plants, with a construct as herein disclosed designed to alter expression of a polynucleotide or polypeptide which modulates GGP activity and/or AsA content in such plant cells and plants. Such methods preferably also include the transformation of plant cells and plants with a combination of the construct as herein disclosed and one or more other constructs designed to alter expression of one or more other polynucleotides or polypeptides which modulate AsA content in such plant cells and plants.

[00211] Any plant is suitable for use in the invention. The L-galactose biosynthetic pathway, which produces AsA is present in all plants. The enzyme GDP-L-galactose phosphorylase (GGP) is critical to the pathway and is also present in all plants. Therefore, the methods of the invention can be used to alter GGP expression and/or increase AsA production in any plant.

[00212] In various embodiments the plant or plant cell is a gymnosperm plant species.

[00213] In a further embodiment the plant or plant cell is an angiosperm plant species.

[00214] In a further embodiment the plant or plant cell is a dicotyledonous plant species.

- 5 **[00215]** Plants and plant cells that are particularly useful in the methods of the invention disclosed herein include all plants and plant cells which belong to the superfamily *Viridiplantae*, in particular monocotyledonous and dicotyledonous plants including a fodder or forage legume, ornamental plant, food crop, tree, or shrub selected from the list comprising *Abrus precatorius*, *Acacia* spp., *Acer* spp., *Actinidia* spp.,
- 10 *Aesculus* spp., *Arachis duranensis*, *Arabidopsis* spp., *Arachis ipaensis*, *Betula* spp., *Brassica* spp., *Buddleja alternifolia*, *Cajanus cajan*, *Camellia sinensis*, *Capsicum* spp., *Carex littledalei*, *Carica papaya*, *Carya illinoensis*, *Castanea mollissima*, *Catharanthus roseus*, *Cephalotus follicularis*, *Chenopodium quinoa*, *Cinnamomum cassia*, *Citrus clementina*, *Citrus sinensis*, *Citrus unshiu*, *Coffea arabica*, *Coronilla varia*, *Corchorus*
- 15 *olitorius*, *Corylus heterophyll*, *Cotoneaster serotina*, *Crataegus* spp., *Cucumis* spp., *Cupressus* spp., *Cyathea dealbata*, *Cydonia oblonga*, *Cryptomeria japonica*, *Cymbopogon* spp., *Dalbergia monetaria*, *Davallia divaricata*, *Desmodium* spp., *Dicksonia squarosa*, *Diheteropogon amplexans*, *Dioclea* spp., *Dolichos* spp., *Dorycnium rectum*, *Durio zibethinus*, *Echinochloa pyramidalis*, *Ehrhartia* spp., *Eleusine coracana*, *Eragrostis* spp.,
- 20 *Erythrina* spp., *Eucalyptus* spp., *Euclea schimperi*, *Eulalia villosa*, *Fagopyrum* spp., *Feijoa sellowiana*, *Fragaria* spp., *Flemingia* spp., *Freycinetia banksii*, *Geranium thunbergii*, *Ginkgo biloba*, *Glycine javanica*, *Glycine max*, *Glycine soja*, *Gliricidia* spp., *Gossypium anomalum*, *Gossypium barbadense*, *Gossypium darwinii*, *Gossypium hirsutum*, *Gossypium mustelinum*, *Gossypium raimondii*, *Gossypium stocksii*, *Grevillea* spp.,
- 25 *Guibourtia coleosperma*, *Hedysarum* spp., *Helianthus annuus*, *Hemarthia altissima*, *Herrania umbratica*, *Heteropogon contortus*, *Hevea brasiliensis*, *Hibiscus syriacus*, *Hordeum vulgare*, *Hyparrhenia rufa*, *Hypericum erectum*, *Hyperthelia dissoluta*, *Indigo incarnata*, *Ipomoea nil*, *Ipomoea triloba*, *Iris* spp., *Jatropha curcas*, *Juglans macrocarpa*, *Juglans regia*, *Juglans macrocarpa* x *Juglans regia*, *Lactuca saligna*, *Lactuca sativa*,
- 30 *Leptarrhenia pyrolifolia*, *Lespedeza* spp., *Lettuca* spp., *Leucaena leucocephala*, *Loudetia simplex*, *Lotonus bainesii*, *Lotus* spp., *Lupinus angustifolius*, *Macadamia integrifolia*, *Macrotyloma axillare*, *Malus* spp., *Manihot esculenta*, *Medicago sativa*, *Metasequoia glyptostroboides*, *Morella rubra*, *Morus notabilis*, *Mucuna pruriens*, *Musa sapientum*, *Nicotiana* spp., *Nelumbo nucifera*, *Nyssa sinensis*, *Onobrychis* spp., *Ornithopus* spp.,
- 35 *Oryza* spp., *Panicum hallii*, *Panicum virgatum*, *Papaver somniferum*, *Peltophorum africanum*, *Pennisetum* spp., *Persea gratissima*, *Petunia* spp., *Phalaenopsis equestris*, *Phaseolus* spp., *Phoenix canariensis*, *Phoenix dactylifera*, *Phormium cookianum*, *Photinia*

- spp., *Picea glauca*, *Pinus* spp., *Pisum sativum*, *Podocarpus totara*, *Pogonarthria fleckii*, *Pogonarthria squarrosa*, *Populus* spp., *Prosopis alba*, *Prosopis cineraria*, *Prunus armeniaca*, *Prunus avium*, *Prunus dulcis*, *Prunus mume*, *Prunus persica*, *Prunus yedoensis* var. *nudiflora*, *Pseudotsuga menziesii*, *Pterolobium stellatum*, *Punica granatum*, *Pyrus* spp., *Quercus* spp., *Rhamnella rubrinervis*, *Rhaphiolepis umbellata*, *Rhodamnia argentea*, *Rhododendron griersonianum*, *Rhododendron simsii*, *Rhododendron williamsianum*, *Rhopalostylis sapida*, *Rhus natalensis*, *Ribes grossularia*, *Ribes* spp., *Ricinus communis*, *Robinia pseudoacacia*, *Rosa* spp., *Rubus* spp., *Salix* spp., *Schyzachyrium sanguineum*, *Sciadopitys verticillata*, *Sequoia sempervirens*, *Sequoiadendron giganteum*, *Solanum chilense*, *Solanum commersonii*, *Solanum lycopersicum*, *Solanum pennellii*, *Solanum tuberosum*, *Sorghum bicolor*, *Spatholobus suberectus*, *Spinacia* spp., *Sporobolus fimbriatus*, *Stiburus alopecuroides*, *Stylosanthos humilis*, *Tadehagi* spp., *Taxodium distichum*, *Telopea speciosissima*, *Trema orientale*, *Tetracentron sinense*, *Themeda triandra*, *Theobroma cacao*, *Trema orientale*, *Trifolium* spp., *Triticum* spp., *Tsuga heterophylla*, *Vaccinium* spp., *Vicia* spp., *Vigna angularis*, *Vigna radiata* var. *radiata*, *Vigna unguiculata*, *Vitis riparia*, *Vitis vinifera*, *Watsonia pyramidata*, *Zantedeschia aethiopica*, *Zea mays*, *Ziziphus jujuba*, amaranth, artichoke, asparagus, broccoli, Brussels sprouts, cabbage, canola, carrot, cauliflower, celery, collard greens, flax, kale, lentil, oilseed rape, okra, onion, potato, rice, soybean, straw, sugar beet, sugar cane, sunflower, tomato, squash and tea, amongst others.

- [00216]** In some embodiments, the plant or plant cell may be a crop plant, such as a food crop or a biofuel crop. In one embodiment, the plant or plant cell may be a cereal, legume, or fruit plant. In one embodiment, the plant or plant cell may be a cereal or legume plant. In one embodiment, the plant or plant cell may be from the family *Poaceae*. In one embodiment, the cereal may be rice, wheat, oats, barley, triticale, rye, finger millet, Sonoran millet, sorghum, or maize. In one embodiment, the cereal may be barley, wheat, rice, or maize. In another embodiment, the legume may be alfalfa or soybeans. In another embodiment, the fruit plant or plant cell may be apple, kiwifruit, or tomato. In one embodiment, the fruit plant or plant cell may be apple or kiwifruit. In one embodiment, the plant or plant cell may be selected from the group comprising kiwifruit, maize, tomato, wheat, barley, tobacco, soybean, rice, apple, cotton, brassicas, and alfalfa. In one embodiment, the plant or plant cell may be selected from the list consisting of *Actinidia arguta*, *Actinidia chinensis*, *Actinidia eriantha*, *Arabidopsis thaliana*, *Glycine max*, *Gossypium hirsutum*, *Hordeum vulgare*, *Malus domestica*, *Medicago sativa*, *Nicotiana benthamiana*, *Nicotiana tabacum*, *Oryza sativa*, *Solanum lycopersicum*, *Triticum aestivum*, and *Zea mays*. In one embodiment, the plant or plant cell may be a rice, soybean, or kiwifruit plant or plant cell. In one embodiment, the plant or plant cell

may be selected from the list consisting of *Actinidia arguta*, *Actinidia chinensis*, *Actinidia eriantha*, *Glycine max*, and *Oryza sativa*.

[00217] In some embodiments, plants or plant cells grown specifically for “biomass” may be used. For example, suitable plants or plant cells include corn, switchgrass, sorghum, miscanthus, sugarcane, poplar, pine, wheat, rice, soy, cotton, barley, turf grass, tobacco, potato, bamboo, rape, sugar beet, sunflower, willow, and eucalyptus. In further embodiments, the plant or plant cell is switchgrass (*Panicum virgatum*), giant reed (*Arundo donax*), reed canarygrass (*Phalaris arundinacea*), *Miscanthus*×*giganteus*, *Miscanthus* sp., *sericea lespedeza* (*Lespedeza cuneata*), millet, ryegrass (*Lolium multiflorum*, *Lolium* sp.), timothy, *Kochia* (*Kochia scoparia*), forage soybeans, alfalfa, clover, sunn hemp, kenaf, bahiagrass, bermudagrass, dallisgrass, pangolagrass, big bluestem, indiagrass, fescue (*Festuca* sp.), *Dactylis* sp., *Brachypodium distachyon*, smooth brome grass, orchardgrass, or Kentucky bluegrass amongst others.

[00218] Alternatively algae and other non-*Viridiplantae* can be used for the methods of some embodiments of the invention.

[00219] In some embodiments the plant or plant cell is a fruit species. In one embodiment, the plant or plant cell is a plant of the Cucurbitaceae family, such as *S. grosvenorii*. In various embodiments the fruit species is selected from the group comprising the following genera: *Actinidia*, *Malus*, *Citrus*, *Fragaria* and *Vaccinium*.

[00220] In a further embodiment the plant or plant cell is selected from the group consisting of *Actinidia eriantha*, *Cucumis sativus*, *Glycine max*, *Solanum lycopersicum*, *Vitis vinifera*, *Arabidopsis thaliana*, *Malus x domestica*, *Medicago truncatula*, *Populus trichocarpa*, *Actinidia arguta*, *Actinidia chinensis*, *Fragaria vulgaris*, *Solanum tuberosum*, and *Zea mays*. In a further embodiment the plant or plant cell is selected from the group consisting of *Actinidia eriantha*, *Cucumis sativus*, *Glycine max*, *Vitis vinifera*, *Arabidopsis thaliana*, *Malus x domestica*, *Medicago truncatula*, *Populus trichocarpa*, *Actinidia arguta*, *Actinidia chinensis*, *Fragaria vulgaris*, and *Zea mays*.

[00221] According to one embodiment, the plant or plant cell is a plant of the *Rosaceae* family, such as but not limited to, apple tree, pear tree, quince tree, apricot tree, plum tree, cherry tree, peach tree, raspberry bush, loquat tree, strawberry plant, almond tree, and ornamental trees and shrubs (e.g. roses, meadowsweets, photinias, firethorns, rowans, and hawthorns).

[00222] In one embodiment the pear is of the genus *Pyrus*. Preferred pear species include: *Pyrus calleryana*, *Pyrus caucasica*, *Pyrus communis*, *Pyrus elaeagnifolia*, *Pyrus*

hybrid cultivar, *Pyrus pyrifolia*, *Pyrus salicifolia*, *Pyrus ussuriensis* and *Pyrus x bretschnideri*.

[00223] In one embodiment the plant or plant cell is of the genus *Malus*. Preferred *Malus* species include: *Malus aldenhamensis*, *Malus angustifolia*, *Malus asiatica*, *Malus*
 5 *baccata*, *Malus coronaria*, *Malus domestica*, *Malus doumeri*, *Malus florentina*, *Malus floribunda*, *Malus fusca*, *Malus halliana*, *Malus honanensis*, *Malus hupehensis*, *Malus ioensis*, *Malus kansuensis*, *Malus mandshurica*, *Malus micromalus*, *Malus niedzwetzkyana*, *Malus ombrophilia*, *Malus orientalis*, *Malus prattii*, *Malus prunifolia*, *Malus pumila*, *Malus sargentii*, *Malus sieboldii*, *Malus sieversii*, *Malus sylvestris*, *Malus toringoides*, *Malus*
 10 *transitoria*, *Malus trilobata*, *Malus tschonoskii*, *Malus x domestica*, *Malus x domestica x Malus sieversii*, *Malus x domestica x Pyrus communis*, *Malus xiaojinensis*, *Malus yunnanensis*, *Malus sp.*, and *Mespilus germanica*. In one embodiment the plant species is *Malus domestica*. In a specific embodiment, the plant is a *Malus domestica*, *Malus trilobata* or *Malus sieboldii*.

15 [00224] In another embodiment, the plant or plant cell is a plant of a *Vitis* species. Exemplary *Vitis* species include, but are not limited to, *Vitis piasezkii maxim* and *Vitis saccharifera makino*.

[00225] In one embodiment the plant or plant cell is a plant from a species selected from a group comprising but not limited to the following genera: *Smilax* (eg *Smilax*
 20 *glyciphylla*), and *Fragaria*.

[00226] In a further embodiment the plant or plant cell is from a vegetable species selected from a group comprising but not limited to the following genera: *Brassica*, *Lycopersicon* and *Solanum*.

[00227] Particularly preferred vegetable plant species are: *Solanum lycopersicum*
 25 (formerly *Lycopersicon esculentum*) and *Solanum tuberosum*.

[00228] In a further embodiment the plant or plant cell is from monocotyledonous species.

[00229] In a further embodiment the plant or plant cell is from a crop species selected from a group comprising but not limited to the following genera: *Glycine*, *Zea*,
 30 *Hordeum* and *Oryza*. Particularly preferred crop plant species are: *Oryza sativa*, *Glycine max* and *Zea mays*.

[00230] In various embodiments, the plant or plant cell is from a food crop or biofuel crop. Species useful for food or biofuel production are known in the art. For example,

species useful for biofuel production may include *Miscanthus* × *giganteus*, *Cenchrus purpureus*, *Cocos nucifera* L., *Jatropha* L., and *Ricinus communis* L.

[00231] In various embodiments, the plant or plant cell is selected from the group consisting of rice, wheat, oats, barley, triticale, rye, finger millet, Sonoran millet, sorghum, maize, banana, *Miscanthus*, elephant grass or Uganda grass, coconut, sugarcane, cotton, sunflower, soybean, flax, sesame, jatropha, sugar beet, alfalfa, forage brassica, oilseed rape, mustard seed, almond, walnut, pecan, macadamia, peanut, and castor bean.

[00232] Plants may be grouped by phylogenetic classification as described in APG IV (The Angiosperm Phylogeny Group *et al.* (2016), An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV, *Botanical Journal of the Linnean Society* **181**(1), 1–20, <https://doi.org/10.1111/boj.12385>), which is incorporated herein by reference.

[00233] In a further embodiment the plant or plant cell is selected from the group consisting of magnoliids, monocots, basal eudicots, Gunnerales, Dilleniales, superrosids, Santalales, Berberidopsidales, Caryophyllales, Cornales, Ericales, Campanulids, Icaciniales, Metteniusales, Garryales, Boraginales, Gentianales, Vahliales, and Lamiales.

[00234] In a further embodiment the plant or plant cell is selected from the group consisting of magnoliids, monocots, basal eudicots, Gunnerales, Dilleniales, superrosids, Santalales, Berberidopsidales, Caryophyllales, Cornales, Ericales, campanulids, Icaciniales, Metteniusales, Garryales, Boraginales, Gentianales, Vahliales, Lamiales, Convolvulaceae, Montiniaceae, Sphenocleaceae, and Hydroleaceae.

[00235] In a further embodiment the plant or plant cell is selected from the group consisting of magnoliids, monocots, basal eudicots, Gunnerales, Dilleniales, Fabids, Geraniales, Myrtales, Crossosomatales, Picramniales, Malvales, Huerteales, Sapindales, Vitales, Saxifragales, Santalales, Berberidopsidales, Caryophyllales, Cornales, Ericales, campanulids, Icaciniales, Metteniusales, Garryales, Boraginales, Gentianales, Vahliales, Lamiales, Convolvulaceae, Montiniaceae, Sphenocleaceae, and Hydroleaceae.

[00236] In a further embodiment the plant or plant cell is selected from the group consisting of magnoliids, monocots, basal eudicots, Gunnerales, Dilleniales, Fabids, Geraniales, Myrtales, Crossosomatales, Picramniales, Malvales, Huerteales, Sapindales, Vitales, Saxifragales, Santalales, Berberidopsidales, Caryophyllales, Cornales, Ericales, campanulids, Icaciniales, Metteniusales, Garryales, Boraginales, Gentianales, Vahliales, Lamiales, Convolvulaceae, Montiniaceae, Sphenocleaceae, Hydroleaceae, Akaniaceae,

Tropaeolaceae, Moringaceae, Caricaceae, Setchellanthaceae, Limnanthaceae, Salvadoraceae, Bataceae, Koeberliniaceae, Emblingiaceae, Pentadiplandraceae, Resedaceae, Gyrostemonaceae, Tovariaceae, Capparaceae, and Cleomaceae.

5 **[00237]** In some embodiments, the plant or plant cell contains an endogenous GGP3 gene, preferably a functional GGP3 gene. In some embodiments, the plant or plant cell has been previously transformed with a functional GGP3 gene. In some embodiments, the plant or plant cell is co-transformed or subsequently transformed with a functional GGP3 gene. In some embodiments, the GGP3 gene has a promoter comprising one or more MYBS1 binding sites, more preferably two or more MYBS1 binding sites. Preferably
10 the MYBS1 binding sites comprise the sequence TCTTATC or its reverse complement GATAAGA.

15. Methods for extracting and measuring AsA from plants

15 **[00238]** Methods are also provided for the production of L-ascorbic acid (AsA) by extraction of AsA from a plant of the invention. AsA may be extracted from plants as follows:

20 **[00239]** Frozen tissue samples are ground to a fine powder in a Cryomill at liquid nitrogen temperature. About 200 mg of frozen powdered tissue is then suspended in 5 volumes of 7% metaphosphoric acid containing 2 mM TCEP (Pierce), vortexed for 20 sec and incubated in a heating block for 2 h at 40°C. TCEP is used in the extraction solution, because it is more effective reducing agent under acidic conditions than DTT, ensuring
25 that all of vitamin C is in the ascorbic acid reduced form. The extract is centrifuged at 4 °C and twenty µL of the supernatant is injected into a Rocket Column And eluted using two solvents A (0.28% o-phosphoric acid, 0.1 mM EDTA and 0.25% methanol) and B (acetonitrile). L-ascorbic acid and other compounds were eluted using a 5-min gradient to 90% B. Standards were run with every batch or 20 samples processed. AsA is
30 calculated from the area under the absorption at 240 nm curve at ~1 minute of elution.

[00240] This method may be up-scaled for larger scale AsA extraction using approaches well-known to those skilled in the art.

30 **[00241]** The above methods should be considered in no way limiting and suitable variations or alternatives will be apparent to those skilled in the art.

EXAMPLE

1. Materials and methods

Plant materials and growth conditions

[00242] Fruit samples were collected from field-grown plants in Hubei, China, during the 2016–2018 growing season. The AsA content in the fruits of 48 *Actinidia* species and 1 cross-population (*A. eriantha* × *A. rufa*) were measured during 2017–2019. Six kiwifruit taxa with >150-fold variation in AsA content were selected to study the changes in AsA during fruit development (*A. eriantha*, *A. latifolia*, *A. chinensis* var. *deliciosa*, *A. chinensis* var. *chinensis*, *A. rufa* and *A. cylindrica*). Tissue culture materials for tobacco (*Nicotiana benthamiana*) and kiwifruit (*A. eriantha*) were grown at 23–25 °C under long-day conditions (16 h of light/8 h of darkness). Transgenic and gene-edited plants were potted and grown in Containment Glasshouse-1 at the Wuhan Botanical Garden, Chinese Academy of Sciences, Hubei, China (14 h of light/10 h of darkness, 18 °C min/30 °C max). Samples collected from individual plants were considered biological replicates.

Illumina RNA-seq and transcriptome analysis

[00243] Fruit RNAs from 20 days after fruiting (DAF20) to DAF120 were extracted separately from *A. eriantha* and *A. rufa*. RNA was extracted using RNeasy Plant Mini Kits (Qiagen, Inc., USA). Library construction and sequencing were performed at BerryGenomics (www.berrygenomics.com) using the Illumina NovaSeq 2000 platform, obtaining paired-end reads of 150 bp and generating > 20 million reads per sample. Sequencing adapters and low-quality raw reads were filtered out using Trimmomatic (DOI: 10.1093/bioinformatics/btu170) with default parameters. The filtered clean reads were mapped to the Hongyang_v3 reference genome (<http://kiwifruitgenome.org/>) by Hisat2 (DOI: 10.1038/s41587-019-0201-4). Read counts and FPKM (fragments per kilobase of exon per million mapped fragments) of annotated genes were calculated by HTSeq (DOI: 10.1093/bioinformatics/btu638.) and Stringtie2 (DOI: 10.1038/nbt.3122), respectively. The differentially expressed genes between the kiwifruit materials with differentiation of AsA in the same development stage were identified by an R package edgeR with a p-value < 0.05, and the Gene Ontology (GO) term enrichment was performed using the Gene Ontology Resource (<http://geneontology.org/>).

RNA extraction and quantitative RT-PCR analysis

[00244] Quantitative reverse transcription (qRT)-PCR was performed following previously described methods (Wang et al., *Plant biotechnology journal* **16**(8): 1424–1433.). Total RNA was isolated with an RNA Extraction Kit (TIANGEN, Beijing, China) and the single-stranded cDNA of all samples was obtained using a one-step gDNA removal

and cDNA Synthesis Supermix Kit (TransGen, Beijing, China). Kiwifruit ACTIN (*Achn107181*) and Protein Phosphatase 2A (*Achn381211*) were used as the reference genes for expression normalisation. The $2^{-\Delta\Delta Ct}$ method (Livak & Schmittgen, 2001) was used to calculate the relative expression of each gene. Primers used for qRT-PCR are provided in Table 1 below. All qPCR analyses were performed with three technical replicates.

AsA measurement

[00245] Measurement of total AsA concentration was performed using High-performance liquid chromatography (HPLC) following a previously described method (Queval & Noctor, 2007 *Anal Biochem* **363**(1): 58-69; Li *et al.*, 2016 *Plant Molecular Biology* **92**(4): 473-482).

Phylogenetic analysis and tree construction

[00246] A phylogenetic tree was constructed using MEGA7.0 (Kumar *et al.*, 2016 *Mol Biol Evol* **33**(7): 1870-1874) software with kiwifruit and other species sequences retrieved from kiwifruit (<http://kiwifruitgenome.org/>) and GenBank databases (Table 2). Genetic distances were calculated using the Jukes-Cantor distance matrix and evolutionary relationships were inferred using the neighbour-joining method with 1000 bootstrap resampling.

Vector construction and kiwifruit transformation

[00247] The coding DNA sequence (CDS) of *AceGGP3*, *AceMYBS1*, *AcrMYBS1* and *AceGBF3* were amplified from cDNA from *A. eriantha* or *A. rufa* and cloned into the overexpression vector (POE-3Flag-DN, from the Wuhan Botanical Garden laboratory, 35S promoter driven expression, and G418 (Geneticin, Invitrogen) or kanamycin selectable markers) to generate 35S::*AceGGP3*, 35S::*AceMYBS1*, 35S::*AcrMYBS1*, 35S::*AceGBF3* respectively.

[00248] The *AceGGP3* and *AceMYBS1* were edited by CRISPR/Cas9 as described previously (Wang *et al.*, 2018. *Plant biotechnology journal* **16**(8): 1424-1433). CRISPR RGEN Tools (http://www.rgenome.net/?tdsourcetag=s_pcqg_aiomsg) was used to select specific sgRNAs that targeted *AceGGP3* and *AceMYBS1* respectively and the sgRNAs were cloned into CRISPR/Cas9 vector to generate Cas9-*AceGGP3* and Cas9-*AceMYBS1* editing vectors. Using *Agrobacterium*-mediated transformation, the recombinant plasmids were transformed into the calli of *A. eriantha* following previously described methods (Akbaş *et al.*, 2009. *Appl Biochem Biotechnol* **158**(2): 470-475; Yuan, 2011 *Proceedings of the 2011 International Conference on Future Computer Science and Application (FCSA 2011)*. 17-20; Wang *et al.*, 2018. *Plant biotechnology journal* **16**(4): 844-855; Wang *et al.*,

2018. *Plant biotechnology journal* **16**(8): 1424-1433). Primers used for vector construction and identification of transgenic lines are listed in Table 1.

Transient expression assay in kiwifruit and tobacco

[00249] Antisense viral vectors *AceGGP3*-TRV2, *AceMYBS1*-TRV2 and *AceGBF3*-TRV2 were obtained by cloning the CDS of *AceGGP3*, *AceMYBS1*, and *AceGBF3* into the TRV2 vector (Yu *et al.*, 2019a) respectively. TRV1 vector was used as an auxiliary plasmid. The TRV1 and TRV2 vectors were transformed into *Agrobacterium tumefaciens* GV3101A respectively and mixed with each other before injection into the *calli* and fruits of *A. eriantha*. The same method was used for transient over expression experiments with 35S driven *AceGGP3*, *AceMYBS1* and *AceGBF3* vectors in *A. eriantha* and *Nicotiana benthamiana* (tobacco) leaves.

Yeast one-hybrid (Y1H) assay

[00250] Yeast one-hybrid assay was performed as previously described (Lin *et al.*, 2007. *Science* **318**(5854): 1302-1305; Cheng *et al.*, 2021. *The Plant cell* **33**(4): 1229-1251.). The promoter sequences of *AceGGP3* (2.6-kb, Table S2) and *AcrGGP3* (2.5-kb, Table S3) were amplified and inserted into the corresponding sites of the reporter plasmid pLacZi (Clontech) to generate *AceGGP3 pro::LacZ* and *AcrGGP3pro::LacZ*. The CDS of *AceMYBS1* and *AcrMYBS1* were separately cloned into pJG4-5 vectors (Clontech) to construct pJG-*AceMYBS1* and pJG-*AcrMYBS1*. *AceGGP3 pro::LacZ* or *AcrGGP3pro::LacZ* were co-transformed with pJG-*AceMYBS1* or pJG-*AcrMYBS1* into yeast strain EGY48 using a high-efficiency yeast transformation method (Gietz & Schiestl, 2007), respectively. The *p53::LacZ*+pJG-p53 was used as positive control and the *AceGGP3pro::LacZ* or *AcrGGP3pro::LacZ* add the pJG4-5 empty vector were the negative control. Transformants were grown on SD/-Trp-Ura dropout medium: 6.7 g.L⁻¹ yeast nitrogen base, 20 g.L⁻¹ Galactose, 10 g.L⁻¹ Raffinose, 2 g.L⁻¹ dropout mix-TRP-URA and 20 g.L⁻¹ agar. After sterilization at 121 °C for 15 min, added 100 mL 10 × BU salt (contain 70 g.L⁻¹ Na₂HPO₄·7H₂O and 30 g.L⁻¹ NaH₂PO₄, pH= 7.0) and 0.08 mg.mL⁻¹ X-gal for colorimetric screening. All yeast strains were incubated at 30 °C for 3 d. Primers are listed in Table 1.

Yeast two-hybrid (Y2H) assay

[00251] Construction of a cDNA library of kiwifruit for yeast two-hybrid experiments were carried out by GeneCreate Biological Engineering Co., Ltd (Wuhan, China) using mRNA from the leaf and fruit. The CDS of *AceMYBS1* was cloned into the pGBKT7 vector (Clontech) as a bait plasmid. Y2H screening assays were performed as described in the BD Matchmaker Two-Hybrid Library Screening Kit user manual (Clontech). For Y2H

assay, AceGBF3 was cloned from cDNA library and inserted into pGADT7 vector (Clontech). The combinations of recombinant plasmids were transferred into yeast strain Y2H gold and grown on SD/-Trp-Leu and SD/-Trp-Leu-His-Ade dropout medium plates supplemented with 0.02 mg.L⁻¹ X- α -gal, at 30 °C for 3 d. The primers are listed in Table

5 1.

Transcriptional activation analysis

[00252] The full-length cDNA of AceMYBS1 and AceGBF3 were PCR amplified and fused into pGBKT7 vector to generated two constructed (BD-AceMYBS1 and BD-AceGBF3) and transformed into the yeast strain AH109. The transcriptional activation analysis was performed as described by (Geng & Liu, 2018. *Journal of experimental botany* **69**(10): 2677-2692).

Dual-luciferase (Dual-LUC) assay

[00253] For the Dual-LUC assay, the full-length promoter of *AceGGP3* (2.6 kb; SEQ ID No: 345) and *AcrGGP3* (2.5 kb; SEQ ID No: 346) or their truncated fragments were cloned into the pGreen0800-Luc vector (Hellens *et al.*, 2005) to obtain reporters: *AceGGP3pro-2660::LUC* (*P₂₆₆₀*), *AceGGP3pro-1106::LUC* (*P₁₁₀₆*), *AceGGP3pro-1606::LUC* (*P₁₆₀₆*), *AceGGP3pro-2088::LUC* (*P₂₀₈₈*), and *AcrGGP3pro::LUC*. The 35S: AceMYBS1, 35S: AcrMYBS1 and 35S: AceGBF3 were transformed into *A. tumefaciens* strain EHA105. Effector and reporters were mixed 5:1 (v/v) then co-infiltrated into 4-weeks-old tobacco leaves as previously (Gao *et al.*, 2020. *Journal of experimental botany* **71**(12): 3560-3574). After 2-3 days at 23 °C, the promoter activities were determined by measuring Firefly Luciferase to Renilla Luciferase (LUC/REN) ratios using a Dual-luciferase Kit (TransGen, Beijing, China) with a Chemiluminescence Imaging System (Clinx, Shanghai, China).

Protein expression and electrophoretic mobility shift assay (EMSA)

[00254] The full-length *AceMYBS1* CDS was inserted into the pET32a vector (Novagen) containing 6 $\times$ His (both fusion in N-terminal and C-terminal) and expressed using *E. coli* strain BL21 (TransGen, Beijing, China) to produce recombinant AceMYBS1-His protein. An *E. coli* strain expressing 6 $\times$ His was used as a negative control. *E. coli* was cultivated 12 hours at 16 °C and then diluted 1: 100 (v/v) into fresh medium and grown on for another 2-3 hours at 37 °C. When cell growth reached the logarithmic phase, IPTG was added to 0.5 mM final concentration and induced protein expression proceeded at 16 °C for 10-12 hours. The fusion protein was purified according to manufacturer instructions using ProteinIso Ni-NAT Resin (TransGen, Beijing, China). The oligonucleotide probes (Table S1) containing the AceMYBS1 binding sequences of *AceGGP3*, which was predicted

by JASPAR²⁰²⁰ (<http://jaspar.genereg.net/>) were synthesized and labelled with biotin. Double strand DNA probes were obtained by annealing two complementary oligonucleotides. The fusion protein AceMYBS1-His was mixed with probes and incubated at room temperature for 20 min. DNA gel mobility shift assay was performed using the

5 EMSA Kit (Beyotime, China) following the manufacturer protocol.

Subcellular localization

[00255] The *AceMYBS1-YFP* or *AceGBF3-YFP* mixed with *NLS-mCherry-RFP* were co-infiltrated into tobacco leaves with infiltration buffer by described previously (Gao *et al.*, 2020. *Journal of experimental botany* **71**(12): 3560-3574). Fluorescence was observed

10 48 h post-infiltration by Confocal Microscopy (Leica TCS-SP8; excitation wavelength with YFP: 510 nm and RFP: 552 nm).

Bimolecular fluorescence complementation assay (BiFC)

[00256] The full-length *AceMYBS1* CDS was fused with C-terminal *YFP* (*cYFP*), and *AceGBF3* was fused with N- terminal *YFP* (*nYFP*). The recombinant vectors or control

15 (empty vectors) were transformed into *A. tumefaciens* strain EHA105 and then co-transformed into onion epidermal cells with infiltration buffer (Zhu *et al.*, 2020. *The Plant cell* **32**(10): 3155-3169). YFP fluorescence was detected by Confocal Microscopy (Leica TCS-SP8; excitation wavelength with YFP: 510 nm and DAPI: 488 nm) 48h after infiltration.

Pull down

[00257] The full-length CDS of *AceGBF3* was cloned it into pGEX-4T vector (GE Healthcare) in which it is fused to *glutathione-S-transferase* (*GST*) sequence. The recombinant vectors were introduced into BL21 cells to produce *AceGBF3-GST* fusion protein, and expressed *AceGBF3-GST* and *AceMYBS1-6×His* fusion protein was purified as

25 described previously (Xu, 2020). Aliquots (20 µL) of *AceGBF3-GST* and *AceMYBS1-His* proteins were mixed with 1 mL binding buffer (50 mM Tris- HCl [pH= 7.5], 100 mM NaCl, 0.25 % Triton-X100- [v/v], 35 mM β-Mercaptoethanol), then 50 µL ProteinIso GST Resin or 50 µL ProteinIso Ni-NAT Resin (TransGen, Beijing, China) was added and the mixture was rotated at 4 °C for 3-4 hours. The samples were washed 4 times with binding buffer,

30 with expressed GST or 6×His used as negative controls. 6×SDS protein loading buffer was added to samples (to 1x final) and samples were denatured by boiling for 10 min before electrophoresis. After electrophoresis the gels analysed by Western blot using anti-His (1:10,000 [v/v], Proteintech, 66005-1-Ig) and anti-GST (1:10,000 (v/v), Proteintech, 66002-2-Ig) antibodies.

Bimolecular luminescence complementation (BiLC) assay

[00258] BiLC was performed as described (Chen *et al.*, 2008. *Plant physiology* **146**(2): 368-376). The full-length of AceMYBS1 and AceGBF3 CDS were insert into pCAMBIA1300-cLUC and pCAMBIA1300-nLUC vector to create AceMYBS1-cLUC and AceGBF3-nLUC constructs, respectively. Agrobacterium cultures harbouring the different constructs were mixed at 1:1 (v/v) and co-transformed into tobacco leaves. Plants were incubated under dark for 12 h and then transferred to light conditions at 25 °C for 48 h. Immediately prior to luciferase activity observation, the transformed tobacco leaves were soaked in 0.15 mg.mL⁻¹ D-Luciferin potassium (Coolaber, China) for 2-3 min and images were captured using a Chemiluminescence Imaging System (Clinx, Shanghai, China).

Statistical analyses

[00259] One-way analysis of variance (ANOVA) was performed using SPSS v20 (IBM Corp., Armonk, NY, USA), and Student's t-test was performed using GraphPad 8.0 software. Significant differences were detected by t-tests. In the figures, the following notations are used: *, P < 0.05; **, P < 0.01; and ***, P < 0.001. In figures, the different letters above the bars denote significance groupings (P < 0.05) as determined by ANOVA, the data represent the mean values, and error bars represent standard deviations.

Accession numbers

[00260] The sequence information for this study was obtained from the Kiwifruit (<http://kiwifruitgenome.org/>), *Nicotiana benthamiana* (https://solgenomics.net/organism/Nicotiana_benthamiana/genome) and National Center for Biotechnology Information (<https://www.ncbi.nlm.nih.gov/>) databases. Additionally, corresponding Acc gene models from *A. chinensis* (genotype Red5) are also referred to and represent PS1.1.69.0 version gene models (Pilkington *et al.*, 2018, *BMC Genomics* **19**). GenBank accession numbers are listed in Table 2.

Table 1: Primers used for qRT-PCR, vector construction and transformation and Yeast-1-Hybrid and Yeast-2-Hybrid.

Primers	Sequence (5'-3')	SEQ ID	Objective
AcrGGP1-qpcrF1	TGTGCTGATGATAGAAGTGGCA	211	qRT-PCR
AcrGGP1-qpcrR1	GGGCATCATTCGCTACTTCATG	212	qRT-PCR
AcrGGP3-qpcrF1	TGAAGCCATCTCTTGTGCTGAT	213	qRT-PCR
AcrGGP3-qpcrR1	TGCTTCTTCACGAGATTGAGGA	214	qRT-PCR
AceGGP1-qpcrF1	TGAAGCCATCTCTTGTGCTGAT	215	qRT-PCR
AceGGP1-qpcrR1	AAGGGCATCATTCGCTACTTCA	216	qRT-PCR
AceGGP3-qpcrF1	TGAAGCCATCTCTTGTGCTGAT	217	qRT-PCR

AceGGP3-qpcrR1	TGCTTCTTCACGAGATTGAGGA	218	qRT-PCR
AcrGGP1-qpcrF2	ATGATAGAAGTGGCAGCACGGCTGA	219	qRT-PCR
AcrGGP1-qpcrR2	ATCATTCGCTACTTCATGAGATTGT	220	qRT-PCR
AcrGGP3-qpcrF2	TGGAAGCGGCAGCACAGCTGA	221	qRT-PCR
AcrGGP3-qpcrR2	ACAGTGGGAGCCTTGTTAAG	222	qRT-PCR
AceGGP1-qpcrF2	ATGATAGAAGCGGCAGCACGGCTGG	223	qRT-PCR
AceGGP1-qpcrR2	ATCATTCGCTACTTCACGAGATTGA	224	qRT-PCR
AceGGP3-qpcrF2	GGGAAGCGGCAGCACAGCTGA	225	qRT-PCR
AceGGP3-qpcrR2	ACTGTGGGAGCCTTTGTTAAT	226	qRT-PCR
AcrGGP2-qpcrF	CCGACGGTTGTTTCCAATTACC	227	qRT-PCR
AcrGGP2-qpcrR	CTGAGGCAATTACGTCCACAAC	228	qRT-PCR
AceGGP2-qpcrF	CCGACGGTTGTTTCCAATTACC	229	qRT-PCR
AceGGP2-qpcrR	CTGAGGCAATTACGTCCACAAC	230	qRT-PCR
AceMYBS1- qpcrF	AATTCAGTGGGCAGAGGAC	231	qRT-PCR
AceMYBS1- qpcrR	CATCCTCAACTAGTAATTGG	232	qRT-PCR
AcrMYBS1- qpcrF	ATGCCATTGCAATTCAGT	233	qRT-PCR
AcrMYBS1- qpcrR	TGCAGCTACATCCTCAACTA	234	qRT-PCR
AceGBF3- qpcrF	GGGACCACCACAGCCTATGATGCCA	235	qRT-PCR
AceGBF3- qpcrR	TGACATTGCTAGCCCATCGAAACCT	236	qRT-PCR
NbGGP3- qpcrF	TTTCGCATCGATCAGGTTCTTCA	237	qRT-PCR
NbGGP3- qpcrR	CTCAAACCTGAAAAGCACTTC	238	qRT-PCR
Nbβ-Tubulin-F	CAAGATGCTACTGCAGACGAG	239	qRT-PCR
Nbβ-Tubulin-R	CTGGAAGTTGTGGTTTTGGC	240	qRT-PCR
PMM- qpcrF	CCATGTGGAGTGTGATGCCAGTGTC	241	qRT-PCR
PMM-qpcrR	CAAAGCAGCAAGTTGGATGCCGTTT	242	qRT-PCR
GMP- qpcrF	GCTCTGGCTAGGGACAACTGAT	243	qRT-PCR
GMP- qpcrR	TGGAATTCGATCATCTCTTTG	244	qRT-PCR
GME- qpcrF	ATGGTCAGCATGAATGAGATGGCCG	245	qRT-PCR
GME-qpcrR	GCTTCTCCTTAATCAGGGTGTGTGTC	246	qRT-PCR
GGP- qpcrF	GAGAGCTTCTTGCTTGC	247	qRT-PCR
GGP-qpcrR	GCACCCAAGCTGTTGTAACC	248	qRT-PCR
GPP- qpcrF	ATCAGAGAAGGGCGAAGGAGACAAT	249	qRT-PCR
GPP-qpcrR	ACGAACCAGGTCAACGGGCGTCTTA	250	qRT-PCR
GLDH- qpcrF	CGGCATCGTTACCTACTACTCTTC	251	qRT-PCR
GLDH-qpcrR	ATCTTCGGGCAATGGGGCGTAGCGG	252	qRT-PCR
GDH- qpcrF	TCAACTTCTTCGACACCTCTCCGTA	253	qRT-PCR
GDH-qpcrR	ACTCTCTCGGCACTGAAATCAAAGC	254	qRT-PCR
Actin-F	TGAGAGATTCCGTTGCCCAGAAGT	255	qRT-PCR
Actin-R	TTCCTTACTCATGCGGTCTGCGAT	256	qRT-PCR
PP2A-F	GCAGCACATAATTCCACAGG	257	qRT-PCR
PP2A-R	TTTCTGAGCCCATAACAGGAG	258	qRT-PCR
POE-AceGGP3-F	GGGACTCTAGAGGATCCATGTTGAAGATCAAGA GGGTTC	259	Over- expression
POE-AceGGP3-R	GTGGTACCCGGGGATCCTCAGTGCTGAACTAGG CATTC	260	Over- expression
POE-AceMYBS1-F	GGGACTCTAGAGGATCC ATGTCAGCTTCAGTGGATT	261	Over- expression
POE-AceMYBS1-R	GTGGTACCCGGGGATCCTTATTGGTGCATTGTTG GGGGTGCC	262	Over- expression
POE-AcrMYBS1-F	GGACTCTAGAGGATCCATGTCAGCTTCAGTGGAT T	263	Over- expression
POE-AcrMYBS1-R	TGGTACCCGGGGATCCTTGGTGCATTGTTGGGG GTGCC	264	Over- expression
POE-AceGBF3-F	GGGACTCTAGAGGATCC ATGGGAAGTAGTGAAGACGTGA	265	over- expression

POE-AceGBF3-R	GTGGTACCCGGGGATCC TCAGTTGACAGAGCCTG	266	Over- expression
POE-SnRK1A-F	GGAAGTCTAGAGGATCC ATGGATGGATCTGGTGGTCAAGGA	267	Over- expression
POE-SnRK1A -R	TGGTACCCGGGGATCC AAGAACCCGGAGCTGAGCAAGAAA	268	Over- expression
TRV2-AceGGP3/1-F	TAAGGTTACCGAATTCTCACTAAAGTTGGACAGG AAG	269	Inhibit- expression
TRV2-AceGGP3/1-R	GCTCGGTACCGGATCCAGCAGCTCAGAGATTTTC A	270	Inhibit- expression
TRV2-AceMYBS1-F	TAAGGTTACCGAATTCGGACAACATAGGGGTGG TGG	271	Inhibit- expression
TRV2-AceMYBS1-R	GCTCGGTACCGGATCCAGGAGCATTTCAAGAAA CTT	272	Inhibit- expression
TRV2-AceGBF3-F	TAAGGTTACCGAATTCAGTTAGGTTGAGCTAAG TC	273	Inhibit- expression
TRV2-AceGBF3-R	GCTCGGTACCGGATCCTGCTGTCAAGGACAATG TTG	274	Inhibit- expression
Crispr-AceGGP3-T1F	GGTCTCTTGCACTCCAGAAGTGTTCATTTCAGTT TCAGAGCTATGC	275	Gene editing
Crispr-AceGGP3-T2R	GGTCTCTAAACCATTCACCTCTTGAAAGCATGC ACCAGCCGGGAA	276	Gene editing
Crispr-AceMYBS1-T1F	GGTCTCTTGCACTCAATAGTGGCGATGCTTGTT TCAGAGCTATGCTGGA	277	Gene editing
Crispr-AceMYBS1-T2R	GGTCTCTAAACCAACATCCTCAACTAGTAATTGC ACCAGCCGGGAATCGA	278	Gene editing
Idcrispr- AceGGP3-F	AACCCAACTTAATTTTCGGGCCTTTC	279	Identification of gene edit
Idcrispr- AceGGP3-R	ACAAACTATAAACTGTAGAGAACG	280	Identification of gene edit
Idcrispr-AceMYBS1-F	GGTCTGGACAAGTTTGGGAA	281	Identification of gene edit
Idcrispr-AceMYBS1-R	CTGGCGTCCCAACAGCTGAT	282	Identification of gene edit
YFP- AceMYBS1-F	TTTACAATTACGGATCCATGTCAGCTTCAGTGGA TT	283	Subcellular localization
YFP- AceMYBS1-R	CCCTTGCCCATGGATCCTTGGTGCATTGTTGGGG GTG	284	Subcellular localization
YFP- AceGBF3-F	TTTACAATTACGGATCCATGGGAAGTAGTGAAGA CGT	285	Subcellular localization
YFP- AceGBF3-R	CCCTTGCCCATGGATCCGTTGACAGAGCCTGGCT	286	Subcellular localization
pLaczi- AceGGP3-F	GGAATTCGAAAAGCTCAATTAAATGAATATA	287	Yeast one hybrid
pLaczi- AceGGP3-R	GGGGTACCCCTCGAACTCAGAAAACGCAAAAA CA	288	Yeast one hybrid
pLaczi- AcrGGP3-F	GGAATTCATACGACTCACTATAGGGCGAATTG	289	Yeast one hybrid
pLaczi- AcrGGP3-R	GGGGTACCCC CTCGAACTCAGAAAACGCAAAAACA	290	Yeast one hybrid
PJG- AceMYBS1-F	ATGCCTCTCCCGAATTCATGTCAGCTTCAGTGGA TT	291	Yeast one hybrid
PJG- AceMYBS1-R	GTCCAAAGCTTCTCGAGTTATTGGTGCATTGTTG GGGGTGCC	292	Yeast one hybrid
PJG- AcrMYBS1-F	ATGCCTCTCCCGAATTCATGTCAGCTTCAGTGGA TT	293	Yeast one hybrid

PJG- AcrMYBS1-R	GTCCAAAGCTTCTCGAGTTATTGGTGCATTGTTG GGGGTGCC	294	Yeast one hybrid
pGreen- AceGGP3- lianF	GGTACCTTGAAAAGCTCAATTAATGAATATA	295	Dual-LUC assay
pGreen- AceGGP3- lianR	AAGCTTGTGACACCGAGCTCGAATTCAAGCT	296	Dual-LUC assay
pGreen-AceGGP3- 2088F	GGTACCTATCCCAAATATTTATTCACTTAG	297	Dual-LUC assay
pGreen-AceGGP3- 1606F	GGTACCCGGATCAGCATTTGTTTTCTTCTTA	298	Dual-LUC assay
pGreen-AceGGP3- 1106F	TATAGGGCGAATTGGTCTGTGCTGACTCGCAT GAAATA	299	Dual-LUC assay
pGreen-AceGGP3- 606F	TATAGGGCGAATTGG TAGAAAGTGAGGTGTTGCGTCAAGA	300	Dual-LUC assay
pGreen-AceGGP3- inR	AGCTTGATATCGAATTGTGACACCGAGCTCGAA TTCAAGCT	301	Dual-LUC assay
pGreen-AcrGGP3- lianF	GGTACCCTAAACCTAACACAAATGGGAAGCA	302	Dual-LUC assay
pGreen-AcrGGP3- lianR	AAGCTTCTCGAACTCAGAAAACGCAAAAACA	303	Dual-LUC assay
pGreen-AcrGGP3- 1919F	GGTACCCTAAACCTAACACAAATGGGAAGCA	304	Dual-LUC assay
pGreen-AcrGGP3- 1439F	GGTACCTCCAATCATCTCACGCCATCCAAGC	305	Dual-LUC assay
pGreen-AcrGGP3- 1009F	TATAGGGCGAATTGG ATCATAGGGTGGTTGGGTTGTTTGG	306	Dual-LUC assay
pGreen-AcrGGP3- 515F	TATAGGGCGAATTGGGCCTCTCCTCACCTCACCC CCAAAG	307	Dual-LUC assay
pGreen-AcrGGP3- inR	AGCTTGATATCGAATCTCGAACTCAGAAAACGCA AAAACA	308	Dual-LUC assay
pGreen-AceMYBS1- F	CGGGGTACCGTGATTGGACCAAGAACTGTTGG CTCTTCTTATTGGAG	309	Dual-LUC assay
pGreen-AceMYBS1- R	CCCAAGCTTATAGCTGGGGATTGGCACGTGTCC GGATTTCGATTGCAGC	310	Dual-LUC assay
AD- AceMYBS1-F	GGAGGCCAGTGAATTC ATGTCAGCTTCAGTGGATTG	311	Yeast two hybrid
AD- AceMYBS1-R	CGAGCTCGATGGATCC TTGGTGCATTGTTGGGGGTG	312	Yeast two hybrid
BD- AceMYBS1-F	CATGGAGGCCGAATTC ATGTCAGCTTCAGTGGATTG	313	Yeast two hybrid
BD- AceMYBS1-R	GCAGGTCGACGGATCC TTGGTGCATTGTTGGGGGTG	314	Yeast two hybrid
AD- AceGBF3-F	GGAGGCCAGTGAATTC ATGGGAAGTAGTGAAGACGT	315	Yeast two hybrid
AD- AceGBF3-R	CGAGCTCGATGGATCC GTTGACAGAGCCTGGCT	316	Yeast two hybrid
BD- AceGBF3-F	CATGGAGGCCGAATTC ATGGGAAGTAGTGAAGACGT	317	Yeast two hybrid
BD- AceGBF3-R	GCAGGTCGACGGATCC GTTGACAGAGCCTGGCT	318	Yeast two hybrid
AD- AceGBF3-NF	GGAGGCCAGTGAATTC ATGGGAAGTAGTGAAGACGT	319	Yeast two hybrid
AD- AceGBF3-NR	CGAGCTCGATGGATCC CATAGCAACTCCATCCACA	320	Yeast two hybrid
YCE- AceMYBS1-F	CGCCACTAGTGGATCCATGTCAGCTTCAGTGGAT TGG	321	BiFC assay

YCE- AceMYBS1-R	TCCCGGGAGCGGTACCTTGGTGCATTGTTGGGG GTG	322	BiFC assay
YNE- AceGBF3-F	CGCCACTAGTGGATCCATGGGAAGTAGTGAAGA CGTG	323	BiFC assay
YNE- AceGBF3-R	TCCCGGGAGCGGTACCTCAGTTGACAGAGCCTG GCT	324	BiFC assay
1300cLUC- AceMYBS1-F	GACGAGCTCGGTACC ATGGGAAGTAGTGAAGACGTG	325	BiLC assay
1300cLUC- AceMYBS1-R	ACGAGATCTGGTCGAC GTTGACAGAGCCTGGCTTGT	326	BiLC assay
1300nLUC- AceGBF3-F	GACGAGCTCGGTACC ATGTCAGCTTCAGTGGATTGG	327	BiLC assay
1300nLUC- AceGBF3-R	ACGAGATCTGGTCGAC TTATTGGTGCATTGTTGGGGGTG	328	BiLC assay
His-AceMYBS1-F	GGCTGATATCGGATCC ATGTCAGCTTCAGTGGATTG	329	EMSA and pull-down
His-AceMYBS1-R	GTGCGGCCGCAAGCTT TTGGTGCATTGTTGGGGGTG	330	EMSA and pull-down
GST- AceGBF3-F	GGTCCGCGTGGATCCATGGGAAGTAGTGAAGA CG	331	Pull-down
GST- AceGBF3-R	CAGTCACGATGAATTCGTTGACAGAGCCTGGCTT G	332	Pull-down
Bio- AceMYBS1A-F	Bio- TAATTAAATAGATAAGAAAAGAGAAAAAGG	333	Probe of EMSA assay
Bio- AceMYBS1A-R	Bio- CCTTTTCTCTTTTCTTATCTATTTAATTA	334	Probe of EMSA assay
Bio- AceMYBS1A- mutF	Bio- TAATTAAATAAAAAAAAAAAGAGAAAAAGG	335	Mutated probe of EMSA assay
Bio- AceMYBS1A- mutR	Bio- CCTTTTCTCTTTTTTTTTTATTTAATTA	336	Mutated probe of EMSA assay
Bio- AceMYBS1B-F	Bio- AAATAAGGGCAATCTTATCATTTATGTCAA	337	Probe of EMSA assay
Bio- AceMYBS1B -R	Bio- TTGACATAAATGATAAGATTGCCCTTATTT	338	Probe of EMSA assay
Bio- AceMYBS1B - mutF	Bio- AAATAAGGGCAAAAAAAAAAATTTATGTCAA	339	Mutated probe of EMSA assay
Bio- AceMYBS1B- mutR	Bio- TTGACATAAATTTTTTTTTTGCCCTTATTT	340	Mutated probe of EMSA assay
AceMYBS1A-F	TAATTAAATAGATAAGAAAAGAGAAAAAGG	341	Competing probe of EMSA assay
AceMYBS1A-R	CCTTTTCTCTTTTCTTATCTATTTAATTA	342	Competing probe of EMSA assay
AceMYBS1B-F	AAATAAGGGCAATCTTATCATTTATGTCAA	343	Competing probe of EMSA assay
AceMYBS1B -R	TTGACATAAATGATAAGATTGCCCTTATTT	344	Competing probe of EMSA assay

Table 2: GenBank accession numbers.

Gene name	Gene ID	Species
CaMYBS1	KAG7023245.1	<i>Cucurbita argyrosperma</i> subsp. <i>argyrosperma</i>
MtMYBS1	XP_013458057.1	<i>Medicago truncatula</i>
BhMYBS1	XP_038904415.1	<i>Benincasa hispida</i>
JcMYBS1	XP_012092968.1	<i>Jatropha curcas</i>
ZmMYBS1-1	XP_035817471.1	<i>Zea mays</i>
ZmMYBS1-2	XP_008655472.1	<i>Zea mays</i>
PduMYBS1	XP_034216346.1	<i>Prunus dulcis</i>
SpMYBS1	XP_015067500.1	<i>Solanum pennellii</i>
SIMYBS1	XP_004235740.1	<i>Solanum lycopersicum</i>
ZjMYBS1	XP_015883030.1	<i>Ziziphus jujuba</i>
CisiMYBS1	XP_006485754.1	<i>Citrus sinensis</i>
PtMYBS1	XP_006376751.1	<i>Populus trichocarpa</i>
CcMYBS1	XP_006440920.1	<i>Citrus clementina</i>
McMYBS1	XP_022134696.1	<i>Momordica charantia</i>
PavMYBS1	XP_021810789.1	<i>Prunus avium</i>
HuMYBS1	XP_021292289.1	<i>Herrania umbratica</i>
CnMYBS1	KAG1346991.1	<i>Cocos nucifera</i>
GhMYBS1	XP_016668580.2	<i>Gossypium hirsutum</i>
RchMYBS1	XP_024170989.1	<i>Rosa chinensis</i>
AtsMYBS1-1	XP_020182102.2	<i>Aegilops tauschii</i> subsp. <i>Stragulata</i>
AtsMYBS1-2	XP_020153631.2	<i>Aegilops tauschii</i> subsp. <i>Stragulata</i>
AtsMYBS1-3	XP_020147245.1	<i>Aegilops tauschii</i> subsp. <i>Stragulata</i>
EgrMYBS1	XP_018721495.1	<i>Eucalyptus grandis</i>
PdaMYBS1	XP_008777770.2	<i>Phoenix dactylifera</i>
TwMYBS1	XP_038680313.1	<i>Tripterygium wilfordii</i>
BrMYBS1-1	XP_009107320.1	<i>Brassica rapa</i>
BrMYBS1-2	XP_009147920.1	<i>Brassica rapa</i>
StMYBS1	KAF7806327.1	<i>Senna tora</i>
HaMYBS1	XP_022017137.1	<i>Helianthus annuus</i>
TtMYBS1	KAF5200833.1	<i>Thalictrum thalictroides</i>
PaMYBS1	XP_034914820.1	<i>Populus alba</i>
VriMYBS1	XP_034672935.1	<i>Vitis riparia</i>
CsaMYBS1-1	XP_030492585.1	<i>Cannabis sativa</i>
CsaMYBS1-2	XP_004138233.1	<i>Cucumis sativus</i>
PgMYBS1	XP_031371370.1	<i>Punica granatum</i>
PvMYBS1	XP_031288186.1	<i>Pistacia vera</i>
ItMYBS1	XP_031105045.1	<i>Ipomoea triloba</i>
QlMYBS1	XP_030937185.1	<i>Quercus lobata</i>
RaMYBS1	XP_030521694.1	<i>Rhodamnia argentea</i>
SoMYBS1	XP_030439880.1	<i>Syzygium oleosum</i>
AhMYBS1-1	XP_025644418.1	<i>Arachis hypogaea</i>

AhMYBS1-2	XP_025616452.1	<i>Arachis hypogaea</i>
EguMYBS1	XP_010923684.1	<i>Elaeis guineensis</i>
PalMYBS1-1	XP_028753871.1	<i>Prosopis alba</i>
PalMYBS1-2	XP_028753870.1	<i>Prosopis alba</i>
DcMYBS1	XP_020701896.1	<i>Dendrobium catenatum</i>
CasiMYBS1	XP_028110714.1	<i>Camellia sinensis</i>
GsMYBS1	RZC07495.1	<i>Glycine soja</i>
ApMYBS1	XP_027331692.1	<i>Abrus precatorius</i>
CaMYBS1	XP_004508571.1	<i>Cicer arietinum</i>
OsMYBS1-1	XP_025882005.1	<i>Oryza sativa</i> Japonica Group
OsMYBS1-2	XP_015627636.1	<i>Oryza sativa</i> Japonica Group
RcoMYBS1	XP_002511974.1	<i>Ricinus communis</i>
BdMYBS1-1	XP_010236797.1	<i>Brachypodium distachyon</i>
BdMYBS1-2	XP_003565794.1	<i>Brachypodium distachyon</i>
MnMYBS1	XP_010105364.1	<i>Morus notabilis</i>
EsMYBS1	XP_006393319.1	<i>Eutrema salsugineum</i>
QsMYBS1	XP_023913108.1	<i>Quercus suber</i>
LsMYBS1	XP_023738581.1	<i>Lactuca sativa</i>
CrMYBS1-1	XP_006305410.1	<i>Capsella rubella</i>
CrMYBS1-2	XP_006298960.1	<i>Capsella rubella</i>
SiMYBS1	XP_004968851.1	<i>Setaria italica</i>
VraMYBS1	XP_014507110.1	<i>Vigna radiata</i> var. <i>radiata</i>
BnMYBS1	XP_013715778.1	<i>Brassica napus</i>
HbMYBS1	XP_021648857.1	<i>Hevea brasiliensis</i>
MeMYBS1	XP_021611780.1	<i>Manihot esculenta</i>
SbMYBS1-1	XP_002447620.2	<i>Sorghum bicolor</i>
SbMYBS1-2	XP_002455466.1	<i>Sorghum bicolor</i>
AdMYBS1	XP_015944219.1	<i>Arachis duranensis</i>
AiMYBS1-1	XP_016189009.1	<i>Arachis ipaensis</i>
AiMYBS1-2	XP_002891472.1	<i>Arabidopsis lyrata</i> subsp. <i>lyrata</i>
SinMYBS1	XP_011073055.1	<i>Sesamum indicum</i>
FcMYBS1	BAG74460.1	<i>Fagus crenata</i>
DeMYBS1-1	KAF8730393.1	<i>Digitaria exilis</i>
DeMYBS1-2	KAF8670787.1	<i>Digitaria exilis</i>
AsuGBF3.1	KAG7644487.1	<i>Arabidopsis suecica</i>
AsuGBF3.2	KAG7644488.1	<i>Arabidopsis suecica</i>
AthGBF3.1	NP_001323456.1	<i>Arabidopsis thaliana</i>
AthGBF3.2	NP_001323457.1	<i>Arabidopsis thaliana</i>
EcoGBF3	ACC77654.1	<i>Eleusine coracana</i>
VviGBF3	RVX04237.1	<i>Vitis vinifera</i>
AlyGBF3	XP_002880229.1	<i>Arabidopsis lyrata</i> subsp. <i>lyrata</i>
CroGBF3	AAK14790.1	<i>Catharanthus roseus</i>
AthGBF3.3	CAA45358.1	<i>Arabidopsis thaliana</i>
AthGBF3.4	NP_174494.2	<i>Arabidopsis thaliana</i>
AthGBF3.5	AAA90947.1	<i>Arabidopsis thaliana</i>
AthGBF3.6	NP_850248.2	<i>Arabidopsis thaliana</i>

AthGBF3.7	NP_171893.1	<i>Arabidopsis thaliana</i>
BraGBF3	KAG5395572.1	<i>Brassica rapa</i> subsp. <i>trilocularis</i>
DexGBF3.1	KAF8695903.1	<i>Digitaria exilis</i>
DexGBF3.2	KAF8675825.1	<i>Digitaria exilis</i>
DexGBF3.3	KAF8641874.1	<i>Digitaria exilis</i>
CsaGBF3.1	KAF4389315.1	<i>Cannabis sativa</i>
CsaGBF3.2	KAF4354785.1	<i>Cannabis sativa</i>
AthGBF3.8	NP_171893.1	<i>Arabidopsis thaliana</i>
TurGBF3	EMS64924.1	<i>Triticum urartu</i>
GarGBF3	XP_017635172.1	<i>Gossypium arboreum</i>
MnoGBF3	EXC02957.1	<i>Morus notabilis</i>
AceGGP3	DTZ79_29g10040/Actinidia32270	<i>Actinidia eriantha</i> /A. <i>chinensis</i>
AeGGP1	DTZ79_17g07300/Actinidia05074	<i>Actinidia eriantha</i> /A. <i>chinensis</i>
AeGGP2	DTZ79_27g04660/Actinidia36370	<i>Actinidia eriantha</i> /A. <i>chinensis</i>
AceMYBS1	DTZ79_16g09490/Actinidia31027	<i>Actinidia eriantha</i> /A. <i>chinensis</i>
AceGBF3	DTZ79_15g00300/Actinidia27344	<i>Actinidia eriantha</i> /A. <i>chinensis</i>
PMM	DTZ79_08g09640	<i>Actinidia eriantha</i>
GMP	DTZ79_03g03560	<i>Actinidia eriantha</i>
GME	DTZ79_24g13050	<i>Actinidia eriantha</i>

2. GGP3 regulates AsA biosynthesis in kiwifruit

[00261] To determine the AsA variation among members of the *Actinidia* genus, the AsA content in the fruits of 48 species was determined by HPLC. The concentration varied tremendously among the different species, ranging from 4.4 to 1185 mg.100 g⁻¹ FW. Species with low (0–30 mg 100.g⁻¹ FW), moderate (30–200 mg 100.g⁻¹ FW) and high vitamin C (200–1200 mg.100 g⁻¹ FW) contents constituted 29.1%, 56.3% and 14.6% of the species, respectively. Changes in AsA contents during the growing season were determined in six *Actinidia* species representing low, moderate and high AsA concentrations (Figure 1). AsA content in *Actinidia* species with high or moderate fruit AsA content accumulated AsA rapidly after fertilization, peaking at 60 days after flowering (DAF60) sampling point, before decreasing as the fruit progressed towards maturity, whereas that of low-AsA content species stayed low level during the whole growing season. AsA biosynthesis during early fruit development is thus the main reason for AsA accumulation among members of the *Actinidia* genus, and differences in AsA synthesis at this stage might constitute the main determinant of AsA variation among *Actinidia* species.

[00262] To identify potential AsA regulatory genes, transcriptome sequencing of the fruits at DAF20, DAF40, DAF60 and DAF120 was performed for *A. eriantha* and *A. rufa*, whose fruits presented a more than 30-fold difference in AsA content. A total of 24415 differentially expressed genes (DEGs) were identified by pairwise comparisons, and Gene

Ontology (GO) enrichment analysis revealed that these DEGs were significantly enriched in biological processes related to the response to catalytic activity and biosynthetic processes (data not shown). The transcript *Actinidia32270*, encoded by *GGP*, was identified in *A. eriantha* on the basis of the increased transcript abundance in the L-galactose pathway. Three *GGP* homologous genes were further subjected to RT-qPCR and correlation analysis of AsA contents in the fruits of *A. eriantha*, *A. rufa* and their hybrids at different developmental stages was conducted (Figure 2, A and B). *GGP3* in *A. eriantha* (*AceGGP3*) was most highly expressed and correlated with high AsA content in the fruits.

- 10 [00263] *GGP3* expression was also highly correlated with AsA concentration in the *A. eriantha* × *A. rufa* hybrid (data not shown), and the expression of *AceGGP3* allele derived from *A. eriantha* was significantly higher than the *A. rufa* derived allele *AcrGGP3* (data not shown).

3. *AceGGP3* overexpression leads to a sharp accumulation of AsA in kiwifruit

- 15 [00264] Transient expression assays were performed to confirm the function of *AceGGP3* in AsA accumulation using 35S driven *AceGGP3* overexpression and silencing constructs in on-vine kiwifruit and *calli* in tissue culture. At 7 days after infiltration *AceGGP3* expression and AsA content was significantly higher infiltrated fruit than in the control fruit (Figure 3A and B). Moreover, silencing of *AceGGP3* in fruiting plants led to an
20 approximately 24.0% decrease in AsA compared with that in fruits infiltrated with bacteria harbouring empty vectors (Figure 3B). Similar results were obtained in transient assays of *A. eriantha calli*, in which the overexpression of *AceGGP3* significantly increased the AsA content (Figure 3C and D).

- [00265] To further analyse *AceGGP3* function, transgenic kiwifruit lines were
25 generated by *Agrobacterium* transformation of *calli*. Five independent transgenic lines of *A. eriantha* overexpressing *AceGGP3* were obtained. Compared to wild-type (WT) plants, the transgenic plants exhibited varying levels of *AceGGP3* gene expression – from 3.1- to 8.1-fold (Figure 4A), and AsA contents increased by 6.3-, 20.0-, 22.7-, 16.7- and 14.1-fold, respectively (Figure 4B).

- 30 [00266] *AceGGP3* was then mutated in kiwifruit via the CRISPR/Cas9 system, and the two targeted sites were located on the first and second exons of *AceGGP3*. In total, four G418-resistant lines containing three types of homozygous mutant *AceGGP3* genes (*ggp3*#2, a 9 bp deletion; *ggp3*#11 and #15, a 1 bp insertions; *ggp3*#13, a 2 bp insertion) were selected. Three mutations resulted in frameshift mutations or amino acid
35 deletions that induced premature termination or truncation of the predicted protein,

causing a loss of the *AceGGP3* gene function. The AsA content in the fruits of the *ggp3* mutants decreased by 32.2%, 5.11% (not statistically significant), 18.2% and 15.9% compared with that of the WT (Figure 5). The 1 bp insertion in *gg3#11* was predicted to cause a nonsense frameshift but AsA levels remains effectively unchanged. However, a scan of ORFs in this mutated CDS showed that translation initiation from the 4th ATG codon downstream would produce a truncated version of GGP missing 53 amino acids from the N-terminus, and the unchanged AsA content in *gg3#11* suggests this variant has functional GGP activity. Collectively, the results show that up- or downregulated expression of *AceGGP3* significantly affects AsA accumulation in *A. eriantha*.

4. **AceMYBS1 acts as a transcriptional activator of *AceGGP3***

[00267] Through transcriptome analysis, a transcript, *Actinidia31027* (*Actinidia chinensis* (Hong yang) protein v3; Actinidia39811: *Actinidia chinensis* (Hong yang) protein v3; Acc18653.1:LG16-18681898..18684251 [e=0]), whose expression is strongly correlated with *Actinidia32270* (*GGP3*; SEQ ID No: 347) expression, was identified (SEQ ID No:204), and RT-qPCR analysis also showed its expression was highly correlated with both *AceGGP* expression and AsA content in *A. eriantha* (Figure 6A and B). A phylogenetic tree comprising the sequences of amino acids of Actinidia31027 and MYBs of other plant species showed that Actinidia31027 was most closely related to TwMYBS1, as they shared 86% sequence identity, whereas it was most distant from CcMYB5 (not shown). Therefore, this protein has been designated as AceMYBS1. Multiple alignments performed on the SMART website (<http://smart.embl-heidelberg.de/>) (Schultz *et al.*, 1998, *Proceedings of the National Academy of Sciences of the United States of America* **95**(11): 5857-5864; Letunic *et al.*, 2020, *Nucleic Acids Research* **49**(D1): D458-D460) showed that Actinidia31027 contains two SANT/MYB regions. A SANT domain is a protein domain that allows many chromatin remodelling proteins to interact with histones (Boyer *et al.*, 2002, *Molecular Cell* **10**(4): 935-942; Boyer *et al.*, 2004, *Nature Reviews Molecular Cell Biology* **5**(2): 158-163), and because SANT domains share many similarities with MYB DNA binding domains and they are often conflated together. SANT and MYB domains can be distinguished by the predicted isoelectric point (pI) of the domain peptide, with histone interacting SANT domains having acidic pIs, and MYB DNA binding domains having basic pI (Ko *et al.*, 2008, *Molecular Cancer* **7**(1): 77). The N-terminal SANT of Actinidia31027 has a predicted acidic isoelectric point (pI), and the second SANT/MYB region has a predicted basic pI (data not shown). Therefore this predicts that AceMYBS1 has one histone interacting domain at its N-terminus side and a MYB DNA binding domain in the middle of the protein.

[00268] To reveal the mechanism by which *AceGGP3* is regulated by AceMYBS1, a Y1H assay was conducted. All the yeast cells grew well on SD/-Trp/-Ura media, whereas

only the positive control and bait vector *AceGGP3pro::LacZ* co-transformed with the prey vector pJG-AceMYBS1 or pJG-AcrMYBS1 had blue cells on media supplemented with X-gal (data not shown). Bioinformatics predictions were performed with JASPAR 2020 (Fornes *et al.*, 2019, *Nucleic Acids Research* **48**(D1): D87-D92). Among many predicted *MYBS1* *cis*-elements, two main *cis*-elements were identified within the *AceGGP3* promoter (-2455 bp, TCTTATC; -1354 bp, TCTTATC). Four 5'-truncated *AceGGP3* promoters (designated *P₁₁₀₆*, *P₁₆₀₆*, *P₂₀₈₈*, and *P₂₆₆₀*) were amplified from *A. eriantha*, and the transcriptional activity was assessed with a dual-LUC reporter system (Figure 7). The activities of *P₁₆₀₆* and *P₂₆₆₀* were higher than those of *P₁₁₀₆* and *P₂₀₈₈*, indicating that *cis*-elements are likely located between positions -1106–1606 bp and -2088–2660 bp, which is the same interval as the predicted position. However, regardless of which 5'-truncated *AceGGP3* promoter was used to drive the reporter gene, *AceMYBS1* co-transfected tobacco presented a much higher relative luciferase LUC than did *AceGGP3* co-transfected tobacco, suggesting that AceMYBS1 also binds to other regions of the *AceGGP3* promoter (Figure 7A). To verify the binding specificity of AceMYBS1 to these motifs, we carried out EMSAs and found that AceMYBS1-HIS fusion proteins could bind DNA probes containing the motifs, whereas non-labelled competing probes effectively reduced the binding ability of AceMYBS1 in a dose-dependent manner, and mutation of the core sequence abolished the binding (Figure 7B). The experiments above were repeated in *A. rufa* (has low AsA) and it was found that AcrMYBS1 could not bind to the AcrGGP3 promoter to regulate its expression (data not shown). The *A. rufa* GGP3 promoter (SEQ ID No: 346) has a number of polymorphisms (deletions/insertions) compared to that of *A. eriantha* (SEQ ID No: 345), including putative MYBS1 binding sites. Without wishing to be bound by theory, the investigator's hypothesis is that as the AcrMYBS1 protein is almost identical to AceMYBS1 (97.7% identity at the amino acid level), the lack of binding is likely due to a defective AcrGGP3 promoter, rather than a non-functional MYBS1 protein. It may therefore be possible to increase AsA content in *A. rufa* by replacing or modifying the AcrGGP3 promoter, either alone or in combination with overexpression of *MYBS1*.

[00269] To explore the AceMYBS1 expression profiles of *A. eriantha*, RT-qPCR was performed to assess transcript accumulation in the fruits. Consistent with AsA content, AceMYBS1 was highly expressed in the early fruit development stage and was positively correlated with the expression of *AceGGP3* (Figure 6A and B). A transient expression assay in *A. eriantha* fruits confirmed that the overexpression of AceMYBS1 increased the expression of *AceGGP3* by 1.4-fold (Figures 8A and B), leading to a 1.7-fold increase in the AsA content in the fruits (Figure 8C). In contrast, the expression level of *AceGGP3* in the TRV-AceMYBS1 fruits decreased by 30%. As expected, overexpression of AceMYBS1 in *A. eriantha* calli equally led to increased expression of *AceGGP3* (1.8-fold) and content of AsA (2.36-fold), and compared with empty vector-infected calli, TRV-AceMYBS1-

infected calli exhibited significantly decreased *AceGGP3* expression and AsA content (Figures 9A-C). When *AceGGP3* was suppressed by TRV-*AceGGP3*, AsA content did not increase, even when *AceMYBS1* was overexpressed in *A. eriantha* calli and transgenic lines, suggesting that *AceMYBS1* is an upstream regulatory gene of *AceGGP3*.

5 **[00270]** To gain further understanding of the regulatory roles of *AceMYBS1*, six stable overexpression transgenic kiwifruit lines were generated. RT-qPCR confirmed that the six independent transgenic lines accumulated high levels of *AceMYBS1* transcripts (>32- to 55-fold; Figure 10A), which successfully increased the expression levels of *AceGGP3* by 2.5- to 4-fold (Figure 10B). AsA content increased by 1.4-, 1.6-, 2.0-, 1.5-, 1.7- and 1.8- fold that of controls, respectively (data not shown). Using the CRISPR/Cas9 system, *AceMYBS1* was further mutated in *A. eriantha*. The two targeted sites were located within the first and third exons. In total, five independent *AceMYBS1* mutants (*mybs1#12*, a 1 bp deletion; *mybs1#28*, a 26 bp deletion; *mybs1#30*, a 3 bp deletion; *mybs1#41*, a 1 bp insertion; *mybs1#53*, a 4 bp deletion) were identified (data not shown). The AsA contents of the *AceMYBS1*-edited lines decreased by 33.2%, 23.1%, 9.8%, 40.4% and 40.0%, respectively (data not shown). For *mybs1#30* which only had a 9.8% reduction in AsA content; while the 3 bp deletion of ATG was not in-frame, the effect was only slight because the result was effectively synonymous due to degeneracy and the only difference to WT was the loss of an aspartic acid at position 205 (just over 20 amino acids downstream from MYB DNA binding domain). Taken together, these findings support that *AceMYBS1* is a positive factor that modulates AsA synthesis.

5. ***AceGBF3* functions additively with *AceMYBS1* to upregulate the expression of *AceGGP3* and increase the synthesis of AsA in kiwifruit**

[00271] To identify *AceMYBS1*-interacting proteins a yeast two-hybrid (Y2H) screen was conducted. Approximately 85 yeast transformants were screened, and 8 positive clones were identified as containing the same cDNA as its full-length sequence. This sequence encodes a bZIP protein (*Actinidia27344*; SEQ ID No:109), whose expression was highly correlated with *AceGGP3* expression (SEQ ID No:347); this protein was then designated as *AceGBF3* based on bioinformatic analysis (data not shown). Further domain mapping analysis revealed that the N-terminal region (*AceGBF3*-N, amino acids 1–151) but not the C-terminal half (*AceGBF3*-C, amino acids 152–299) of *AceGBF3* interacted with *AceMYBS1* (data not shown). An additional transcriptional activity assay was then performed. Yeast cells transformed with BD-*AceMYBS1* or BD-*AceGBF3* but not the empty pGBKT7 vector grew normally and turned blue on SD/-T/-H/-A/+X- α -Gal media (data not shown). Subcellular localization assays were performed in tobacco leaves transformed with a nuclear marker to visualize the subcellular locations of *AceMYBS1* and *AceGBF3*. Fluorescence of the *AceMYBS1* and *AceGBF3*-fused YFP proteins

was detected only in the nucleus and perfectly merged with nuclear markers (data not shown).

[00272] To further explore potential interactions between the AceGBF3 and AceMYBS1 proteins, three different methods were employed. Firstly, a BiFC experiment
 5 was selected as a method for an *in vivo* assay using a cell biology approach. AceMYBS1 and AceGBF3 proteins were fused to the C-terminus of yellow fluorescent protein (YFP) (AceMYBS1-cYFP) and the N-terminus of YFP (AceGBF3-nYFP), respectively, and then transiently transformed into onion epidermal cells. The YFP signal localized to the nucleus showing close interaction between AceMYBS1 and AceGBF3 as well as nuclear localization
 10 (data not shown). To confirm this interaction AceMYBS1 was fused to the C-terminal half of LUC (AceMYBS1-cLUC) and AceGBF3 was fused to the N-terminal half of LUC (AceGBF3-nLUC) and the constructs were transiently expressed in tobacco leaves. Only leaves co-transformed with AceMYBS1-cLUC and AceGBF3-nLUC produced a strong LUC signal (Figure 11A). Finally, AceMYBS1-AceGBF3 physical interactions were confirmed via
 15 an *in vitro* pulldown assay using recombinant purified proteins. The AceGBF3-glutathione S-transferase (GST) fusion protein was precipitated with AceMYBS1-6× His but not with 6× His alone when GST resin was used (Figure 11B).

[00273] To test the hypothesis that AceGBF3 and AceMYBS1 co-regulate AsA synthesis, we performed dual-LUC, overexpression and virus-induced gene silencing
 20 experiments. These showed that AceMYBS1 and AceMYBS1 plus AceGBF3 but not AceGBF3 alone were capable of inducing the expression of AceGGP3 (of which increased by 3.74- and 6.67-fold, respectively) (Figure 12A). Transient overexpression of AceGBF3 in different plant materials, including kiwifruit (Figure 12B, C), *calli* (Figure 12D, E) and tobacco leaves (Figure 12, F, H), consistently promoted the accumulation of AsA and
 25 expression of GGP3. Furthermore, co-overexpression of AceGBF3 together with AceMYBS1 promoted the accumulation of AsA additively (Figure 13A, B).

[00274] When AceMYBS1 was silenced in *A. eriantha* fruits, both the expression level of AceGGP3 and the content of AsA were notably reduced or unchanged, regardless of whether AceGBF3 was overexpressed or suppressed, confirming that AceGBF3 has to
 30 interact with AceMYBS1 to regulate the transcription of. Moreover, five transgenic *A. eriantha* lines constitutively overexpressing AceGBF3 were generated, and their AceGGP3 transcript levels and AsA contents increased by 2.11–3.38-fold and 1.22–1.78-fold, respectively, consistent with previous results (Figure 14A, B). The average AsA content in *calli* (which was 2-fold higher than that in WT *calli*) of the AceGBF3 and AceMYBS1 co-
 35 expression lines was significantly higher than that in the *calli* of the AceGBF3 (1.4-fold) and AceMYBS1 (1.6-fold) lines alone (Figure 14B). Taken together, these results show

that AceGBF3 interacts with AceMYBS1 to co-regulate *AceGGP3* expression to form an AceGBF3-AceMYBS1-AceGGP3 regulatory network involved in the synthesis and metabolism of AsA in kiwifruit.

6. *AceMYBS1* domain structure and genetics

5 [00275] The *AceMYBS1* identified herein is located on chr16 and its *A. chinensis* and *A. eriantha* isoforms share 98% amino sequence identity and show no obvious perturbation in the two predicted SANT domains (data not shown). In addition to *AceMYBS1* there is another highly similar MYBS1-like gene sharing 87% amino acid identity on chr26, which lies inside the recently identified *A. eriantha* AsA supergene QTL interval (McCallum *et al.*, 2019, *Plants (Basel, Switzerland)* **8**(7): 237). Alignment of *A. chinensis* chromosome 26 and *AceMYBS1* (chr16) protein sequences revealed the chromosome 16 and 26 alleles shared highly similar SANT/MYB domains and mainly differed by a 12 amino acid deletion in the chr26 allele in an unstructured region in the C-terminal third of the protein (data not shown). Unstructured regions are typically associated with protein-protein binding interactions (Kragelund *et al.*, 2012, *Trends in Plant Science* **17**(11): 625-632; Millard *et al.*, 2019, *Nucleic Acids Research* **47**(18): 9592-9608). Protein disorder predictions using ODiNPred (Dass *et al.*, 2020, *Scientific Reports* **10**(1): 14780) also support the SMART domain architecture predictions (data not shown). The high degree of similarity, particularly between both the SANT and MYB domains suggests that the chr26 *AceMYBS1* gene should also regulate *GGP3* expression but it is very lowly expressed in fruit and had little correlation with AsA levels (data not shown).

7. Stable expression of *AceMYBS1* in plants

Materials and methods

25 [00276] A plasmid T-DNA construct as described in Li *et al.* (2022, *New Phytologist* **235**(6): 2497-2497) containing the *AceMYBS1* gene (derived from *Actinidia eriantha*) was transformed using the *Agrobacterium tumefaciens*-mediated method into rice (*Oryza sativa* L.), soybean (*Glycine max*), and Arabidopsis (*Arabidopsis thaliana*). Stably transformed lines were selected by PCR methods. The mature leaf L-ascorbic acid content (AsA) of each of the lines (triplicate plants) was then measured using HPLC as previously described (Li *et al.*, 2022).

Rice (*Oryza sativa* L.)

[00277] AsA content was measured in 11 transformed lines of *Oryza sativa* L. and compared to wild-type non-transformed controls (WT). AsA content was higher in 10 out

of 11 lines ($p < 0.05$), with variation ranging 1.5 to 89-fold that of untransformed WT (Table 3; Figure 15A).

Table 3: Leaf AsA content of stably-transformed rice plants.

Line Number	Mean AsA (mg/100g fresh weight)	AsA content relative to WT	Std deviation	T-test stat
#1	5.84	3.63	3.783	0.195
#2	2.50*	1.55	0.214	0.026
#8	144.26*	89.67	16.778	0.005
#9	5.54*	3.45	0.008	0.000
#16	21.12*	13.12	0.387	0.000
#17	3.45*	2.14	0.653	0.036
#18	3.35*	2.08	0.249	0.009
#19	4.62*	2.87	0.050	0.000
#20	2.42*	1.50	0.158	0.019
#21	2.93*	1.82	0.106	0.001
#26	2.48*	1.54	0.224	0.015
WT	1.61	1.00	0.040	-

* $p < 0.05$.

5 Soybean (*Glycine max*)

[00278] AsA content was measured in 5 transformed lines of *Glycine max* and compared to wild-type non-transformed controls (WT). AsA content was higher in 3 out of 5 lines ($p < 0.05$), with one line having reduced AsA. Variation ranged from 0.17 to 5.7-fold that of untransformed WT (Table 4; Figure 15B).

10 **Table 4: Leaf AsA content of stably-transformed soybean plants.**

Line Number	Mean AsA (mg/100g fresh weight)	AsA content relative to WT	Std deviation	T-test stat
#4	12.46*	1.99	1.287	0.005
#5	1.05*	0.17	0.087	0.035
#6	35.88*	5.74	3.464	0.002
#7	9.23*	1.48	0.618	0.040
#8	6.96	1.11	1.133	0.702
WT	6.25	1.00	1.682	-

* $p < 0.05$.

Arabidopsis thaliana

[00279] The AsA content was measured in 15 transformed lines of *Arabidopsis thaliana* and compared to wild-type non-transformed controls (WT). Two lines displayed small increases (1.21-1.26-fold), while six lines had slightly reduced AsA ($p < 0.05$).

Variation ranged from 0.75 to 1.26-fold that of untransformed WT (Table 5; Figure 15E). This shows that by generating a number of transformed lines and selecting plants that show the highest AsA content, it is possible to produce *Arabidopsis thaliana* plants having increased AsA content compared to wild-type.

5 **Table 5: Leaf AsA content of stably-transformed *Arabidopsis thaliana* plants.**

Line Number	Mean AsA (mg/100g fresh weight)	AsA content relative to WT	Std deviation	T-test stat
#1	1.29*	1.21	0.063	0.020
#2	1.34*	1.26	0.084	0.037
#3	1.16	1.09	0.046	0.102
#4	1.16	1.09	0.034	0.149
#5	0.96	0.91	0.051	0.125
#6	1.15	1.08	0.047	0.176
#8	0.82*	0.77	0.048	0.041
#9	0.96	0.90	0.129	0.217
#10	0.96*	0.90	0.009	0.042
#11	0.83*	0.78	0.027	0.003
#12	0.98	0.92	0.041	0.220
#13	0.80*	0.75	0.010	0.009
#14	0.97	0.91	0.033	0.057
#15	0.95*	0.90	0.059	0.011
#17	0.90*	0.85	0.015	0.024
WT	1.06	1.00	0.042	-

* $p < 0.05$.

Discussion

[00280] This example shows that constitutive over-expression of *AceMYBS1* can greatly increase AsA concentration in plants. Very large increases were observed in some rice and soybean lines (~89 and ~5.7-fold of the wild-type control). This constitutes the largest increases in AsA observed by transgenic approaches to date (Bulley *et al.*, 2012, *Plant Biotechnology Journal* **10**(4): 390-397; Macknight *et al.*, 2017, *Current Opinion in Biotechnology* **44**: 153-160). Furthermore, as rice (monocots) and soybeans (dicots) are only distantly related, this example suggests that overexpression of *MYBS1* genes can increase AsA concentration in a wide variety of plant species.

[00281] While the AsA content of individual lines varied considerably, this is not unexpected. The expression of transgenes in plants is greatly affected by the genomic context into which they are inserted, and this will vary between individual transformed lines. It is common practice to generate many transformants and select the lines that show the greatest effect. Accordingly, it is not necessary for every transformed line to

show an equivalent increase in AsA content. Rather, a combination of transformation and selection can be used to provide plants with greatly increased AsA content, even if most transformed lines show only modest increases.

[00282] Over-expression of *AceMYBS1* in *Arabidopsis* showed mixed results, but in two cases modest increases in AsA were observed. As previously stated, expression of transgenes can vary by genomic context, and this experiment confirms that it is possible to increase AsA content in *Arabidopsis* by over-expressing *MYBS1* genes. Greater increases in AsA content may be possible by screening additional lines. This provides further evidence that the observed effect is broadly applicable to a variety of different plant species.

8. Expression of *MYBS1* from various plant species.

[00283] *MYBS1* genes from various plant species will be stably expressed in *Arabidopsis thaliana*, and the leaf AsA content measured.

[00284] The plant transformation vector pHex2s will be used, with the *MYBS1* transgene being driven by a 35s constitutive promoter. The *MYBS1* transgenes to be used are listed in Table 6.

Table 6: MYBS1 genes.

Gene	Species	SEQ ID No
AceMYBS1	<i>Actinidia eriantha</i>	103
AcrMYBS1	<i>Actinidia rufa</i>	357
MYBS1-like	<i>Actinidia chinensis</i>	108
MYBS1-like	<i>Actinidia eriantha</i>	358
AtMYBS1	<i>Arabidopsis thaliana</i>	359
AtMYBS2	<i>Arabidopsis thaliana</i>	360
TaMYBS1	<i>Triticum aestivum</i>	361
OsMYBS1	<i>Oryza sativa</i> Japonica group	362
GmMYBS1	<i>Glycine max</i>	363
CaMYBS1	<i>Capsicum annum</i>	364

[00285] *Arabidopsis thaliana* (Col0) will be transformed using *Agrobacterium*-mediated transformation using the floral dip method according to standard protocols.

[00286] Transformed lines will be identified and gene expression confirmed using qRT-PCR methods as described in Example 1. Transformed lines will also have their ascorbate content determined as described in Example 1, and will also be tested for increased expression of the MYBS1 target gene GGP (GDP-galactose phosphorylase) using qRT-PCR.

[00287] It is not the intention to limit the scope of the invention to the abovementioned examples only. As would be appreciated by a skilled person in the art, many variations are possible without departing from the scope of the invention as set out in the appended claims.

5 INDUSTRIAL APPLICATION

[00288] The methods and tools described herein are useful for the production of plants having altered AsA content, including crops such as fruits and grains.

CLAIMS

1. A method of producing a plant cell or plant with increased L-ascorbic acid (AsA) content, the method comprising transforming a plant cell with a polynucleotide encoding a polypeptide with an amino acid sequence of any one of SEQ ID NOs: 1-102 and 348-356, or a variant of the polypeptide having at least about 70% sequence identity to a polypeptide with the amino acid sequence of any one of SEQ ID NOs: 1-102 and 348-356.
2. A method of producing a plant cell or plant having increased GDP-L-galactose phosphorylase 3 (GGP) translation, production and/or activity, the method comprising transforming a plant cell with a polynucleotide encoding a polypeptide with an amino acid sequence of any one of SEQ ID NOs: 1-102 and 348-356, or a variant of the polypeptide having at least about 70% sequence identity to a polypeptide with the amino acid sequence of any one of SEQ ID NOs: 1-102 and 348-356.
3. A method of producing a plant cell or plant with increased L-ascorbic acid (AsA) content, the method comprising transforming a plant cell with a polynucleotide comprising a nucleotide sequence of any one of SEQ ID NOs:103-108, or a variant thereof having at least about 70% sequence identity to a nucleotide having the nucleotide sequence of any one of SEQ ID NOs:103-108.
4. A method of producing a plant cell or plant having increased GDP-L-galactose phosphorylase 3 (GGP) translation, production and/or activity, the method comprising transforming a plant cell with a polynucleotide comprising a nucleotide sequence of any one of SEQ ID NOs: 103-108, or a variant thereof having at least about 70% sequence identity to a nucleotide having the nucleotide sequence of any one of SEQ ID NOs: 103-108.
5. A method of producing a plant cell or plant with increased L-ascorbic acid (AsA) content or increased GDP-L-galactose phosphorylase 3 (GGP) translation, production and/or activity, the method comprising upregulating in the plant cell, or expressing in the plant, a polypeptide with the amino acid sequence of any one of SEQ ID NO: 1-102 and 348-356, or a variant of the polypeptide having at least about 70% sequence identity to the amino acid sequence of any one of SEQ ID NO: 1-102 and 348-356.
6. A method of producing a plant cell or plant with increased L-ascorbic acid (AsA) content or increased GDP-L-galactose phosphorylase 3 (GGP) translation, production and/or activity, the method comprising upregulating in the plant cell, or expressing in the plant, a polynucleotide comprising a nucleotide sequence selected from any one of the

sequences SEQ ID NO: 103-108 or a variant having at least about 70% sequence identity to the nucleotide sequence of any one of SEQ ID NO: 103-108.

7. The method of any one of claims 1 to 6, wherein the method further comprises transforming, or co-transforming with the polynucleotide, a second polynucleotide encoding a polypeptide with an amino acid sequence of any one of SEQ ID NOs: 109-209, or a variant of the polypeptide having at least about 70% sequence identity to the amino acid sequence of any one of SEQ ID NOs: 109-209.

8. The method of any one of claims 1 to 4 or 7, wherein the method comprises stably transforming the plant cell with the polynucleotide or variant thereof.

9. The method of any one of claims 1 to 8, wherein the method further comprises transforming, or co-transforming with the polynucleotide, a second polynucleotide having a nucleotide sequence of SEQ ID NO: 210, or a variant thereof having at least about 70% sequence identity to the nucleotide sequence of SEQ ID NO: 210.

10. The method of any one of claims 1 to 9, wherein the plant cell or plant is from a food or biofuel crop.

11. An isolated polynucleotide encoding a polypeptide with the amino acid sequence of any one of SEQ ID NOs: 1-4 and 102, or a variant of the polypeptide having at least about 75% sequence identity to the amino acid sequence of any one of SEQ ID NOs: 1-4 and 102.

12. An isolated polynucleotide encoding a polypeptide with an amino acid sequence of any one of SEQ ID NOs: 109-153 155-157, 159, 162-164, 166-174, 176, 178, 180-183, 185, 186, 188-190, 192-200, 202-204, and 206-209, or a variant of the polypeptide having at least about 75% sequence identity to the amino acid sequence of any one of SEQ ID NOs: 109-153 155-157, 159, 162-164, 166-174, 176, 178, 180-183, 185, 186, 188-190, 192-200, 202-204, and 206-209.

13. A genetic construct comprising a polynucleotide encoding a polypeptide with an amino acid sequence of any one of SEQ ID NOs: 1-4 and 102, or a variant of the polypeptide having at least about 75% sequence identity to the amino acid sequence of any one of SEQ ID NOs: 1-4 and 102.

14. A genetic construct comprising a polynucleotide encoding a polypeptide with an amino acid sequence of any one of SEQ ID NOs: 109-153 155-157, 159, 162-164, 166-174, 176, 178, 180-183, 185, 186, 188-190, 192-200, 202-204, and 206-209, or a variant of the polypeptide having at least about 75% sequence identity to the amino acid

sequence of any one of SEQ ID NOs: 109-153 155-157, 159, 162-164, 166-174, 176, 178, 180-183, 185, 186, 188-190, 192-200, 202-204, and 206-209.

15. A genetic construct comprising a polynucleotide comprising a nucleotide sequence of any one of SEQ ID NOs: 103-108, or a variant of the polypeptide having at least about 75% sequence identity to the nucleotide sequence of any one of SEQ ID NOs: 103-108.

16. A genetic construct comprising a polynucleotide comprising a nucleotide sequence of SEQ ID NO: 210, or a variant of the polypeptide having at least about 70% sequence identity to the nucleotide sequence of SEQ ID NO: 210.

17. The genetic construct of claim 13 or 15, wherein the construct further comprises a second polynucleotide

- a) encoding a polypeptide with an amino acid sequence of any one of SEQ ID NOs: 109-209, or a variant of the polypeptide having at least about 75% sequence identity to the amino acid sequence of any one of SEQ ID NOs: 109-209; and/or
- b) having a nucleotide sequence of SEQ ID NO: 210, or a variant of the polypeptide having at least about 70% sequence identity to the nucleotide sequence of SEQ ID NO: 210.

18. A host cell genetically modified to express a polynucleotide encoding a polypeptide with an amino acid sequence of any one of SEQ ID NOs: 1-102 and 348-356, or a variant of the polypeptide having at least about 70% sequence identity to the amino acid sequence of any one of SEQ ID NOs: 1-102 and 348-356.

19. A host cell genetically modified to express a polynucleotide comprising a nucleotide sequence of any one of SEQ ID NOs: 103-108, or a variant thereof having at least about 70% sequence identity to the nucleotide sequence of any one of SEQ ID NOs: 103-108.

20. A plant cell genetically modified to express a polynucleotide encoding a polypeptide with the amino acid sequence of any one of SEQ ID NO: 1-102 and 348-356 or a variant of the polypeptide having at least 70% sequence identity to a polypeptide with the amino acid sequence of any one of SEQ ID NO: 1-102 and 348-356.

21. A plant cell genetically modified to express a polynucleotide comprising a nucleotide sequence selected from any one of SEQ ID NO: 103-108 or a variant thereof

having at least 70% sequence identity to the nucleotide sequence of any one of SEQ ID NO: 103-108.

22. The host cell or plant cell of any one of claims 18 to 21, wherein the cell is genetically modified to express a second polynucleotide encoding a polypeptide with an amino acid sequence of any one of SEQ ID NOs: 109-209, or a variant of the polypeptide having at least about 70% sequence identity to the amino acid sequence of any one of SEQ ID NOs: 109-209.

23. The host cell or plant cell of any one of claims 18 to 21, wherein the cell is genetically modified to express a second polynucleotide having a nucleotide sequence of SEQ ID NO: 210, or a variant thereof having at least about 70% sequence identity to the nucleotide sequence of SEQ ID NO: 210.

24. The method, isolated polynucleotide, genetic construct, host cell or plant cell of any one of claims 1-11, 13, 15, and 17-23, wherein the polynucleotide encodes a polypeptide with an amino acid sequence of:

- a) any one of SEQ ID Nos:1-102 and 348-356;
- b) any one of SEQ ID Nos:1-4, 38, 41-45, 48, 49, 51, 52, 56, 60, 63, 66, 79, 80, 85, 89, 90, 92, 96, 97, 100-102, and 348-356;
- c) any one of SEQ ID Nos:1-4, 49, 51, 66, 80, 85, 90, 92, 96, 102, and 348-356;
- d) any one of SEQ ID Nos:1-4 and 102;
- e) any one of SEQ ID Nos:1-4; or
- f) SEQ ID No:1 or 102.

25. The method, isolated polynucleotide, genetic construct, host cell or plant cell of any one of claims 1-11, 13, 15, and 17-23, wherein the polynucleotide encodes a variant of the polypeptide having at least about 70, 75, 80, 85, 90, 95, 96, 97, 98, 99 or 100% sequence identity to an amino acid sequence of

- a) any one of SEQ ID Nos:1-102 and 348-356;
- b) any one of SEQ ID Nos:1-4, 38, 41-45, 48, 49, 51, 52, 56, 60, 63, 66, 79, 80, 85, 89, 90, 92, 96, 97, 100-102, and 348-356;

- c) any one of SEQ ID Nos:1-4, 49, 51, 66, 80, 85, 90, 92, 96, 102, and 348-356;
- d) any one of SEQ ID Nos:1-4 and 102;
- e) any one of SEQ ID Nos:1-4; or
- f) SEQ ID No:1 or 102.

26. The method, isolated polynucleotide, genetic construct, host cell or plant cell of any one of claims 1-11, 13, 15, and 17-23, wherein the polynucleotide has a nucleotide sequence of

- a) any one of SEQ ID NOs:103-108,
- b) SEQ ID No: 103 or 105, or
- c) SEQ ID No: 103.

27. The method, isolated polynucleotide, genetic construct, host cell or plant cell of any one of claims 1-11, 13, 15, and 17-23, wherein the polynucleotide encodes a variant having at least about 70, 75, 80, 85, 90, 95, 96, 97, 98, 99 or 100% sequence identity to a nucleotide sequence of

- a) any one of SEQ ID NOs:103-108,
- b) SEQ ID No: 103 or 105, or
- c) SEQ ID No: 103.

28. The method, isolated polynucleotide, genetic construct, host cell or plant cell of any one of claims 8, 12, 17, 22 and 24-27, wherein the polynucleotide or second polynucleotide encodes a polypeptide with an amino acid sequence of:

- a) any one of SEQ ID Nos: 109-209;
- b) any one of SEQ ID Nos: 109-114, 125, 126, 142, 177, 179, 187, 191 and 205;
- c) any one of SEQ ID Nos: 109-114 and 179;
- d) any one of SEQ ID Nos: 109-114; or
- e) SEQ ID NO:109.

29. The method, isolated polynucleotide, genetic construct, host cell or plant cell of any one of claims 8, 12, 17, 22 and 24-27, wherein the polynucleotide or second polynucleotide encodes a variant having at least about 70, 75, 80, 85, 90, 95, 96, 97, 98, 99 or 100% sequence identity to an amino acid sequence of:

- a) any one of SEQ ID Nos: 109-209;
- b) any one of SEQ ID Nos: 109-114, 125, 126, 142, 177, 179, 187, 191 and 205;
- c) any one of SEQ ID Nos: 109-114 and 179;
- d) any one of SEQ ID Nos: 109-114; or
- e) SEQ ID NO:109.

30. The method, isolated polynucleotide, genetic construct, host cell or plant cell of any one of claims 8, 12, 17, 22 and 24-27, wherein the polynucleotide or second polynucleotide has a nucleotide sequence of SEQ ID No:210.

31. The method, isolated polynucleotide, genetic construct, host cell or plant cell of any one of claims 8, 12, 17, 22 and 24-27, wherein the polynucleotide or second polynucleotide encodes a variant having at least about 70, 75, 80, 85, 90, 95, 96, 97, 98, 99 or 100% sequence identity to a nucleotide sequence of SEQ ID No:210.

32. A kit comprising a first genetic construct of any one of claims 13, 15, 24 and 25; and a second genetic construct of any one of claims 14, 16, 28 and 29.

33. A host cell comprising one or more genetic constructs of any one of claims 13 to 17 and 24 to 31.

34. A plant comprising a plant cell of any one of claims 20, 21 and 24 to 31.

35. A method for selecting a plant with altered GDP-L-galactose phosphorylase 3 (GGP) activity, the method comprising testing a plant for altered expression of a polynucleotide encoding a polypeptide with the amino acid sequence of any one of SEQ ID NO: 1-102 and 348-356 or a variant of the polypeptide having at least 70% sequence identity to a polypeptide with the amino acid sequence of any one of SEQ ID NO: 1-102 and 348-356.

36. A method for selecting a plant with altered GDP-L-galactose phosphorylase 3 (GGP) activity, the method comprising testing a plant for altered expression of a

polynucleotide comprising a nucleotide sequence selected from any one of the sequences SEQ ID NO: 103-108 or a variant thereof wherein the variant comprises a sequence that has at least 70% sequence identity to the nucleotide sequence of any one of SEQ ID NO: 103-108.

37. A method for selecting a plant with altered L-ascorbic acid (AsA) content, the method comprising testing a plant for altered expression of a polynucleotide encoding a polypeptide with the amino acid sequence of any one of SEQ ID NO: 1-102 and 348-356 or a variant of the polypeptide having at least 70% sequence identity to a polypeptide with the amino acid sequence of any one of SEQ ID NO: 1-102 and 348-356.

38. A method for selecting a plant with altered L-ascorbic acid (AsA) content, the method comprising testing a plant for altered expression of a polynucleotide comprising a nucleotide sequence selected from any one of the sequences SEQ ID NO: 103-108 or a variant thereof comprising a sequence that has at least 70% sequence identity to the nucleotide sequence of any one of SEQ ID NO: 103-108.

39. A plant cell or plant produced by a method according to any one of claims 1 to 10 and 24-31.

40. A plant cell or plant selected by a method according to any one of claims 35-38.

41. A group or population of plants produced by a method according to any one of claims 1 to 10 and 24-31.

42. A method of producing AsA, the method comprising extracting AsA from a host cell of any one of claims 18, 19, and 22-31.

43. A method of producing AsA, the method comprising extracting AsA from a plant cell or plant of any one of claims 20 to 31.

FIGURE 1

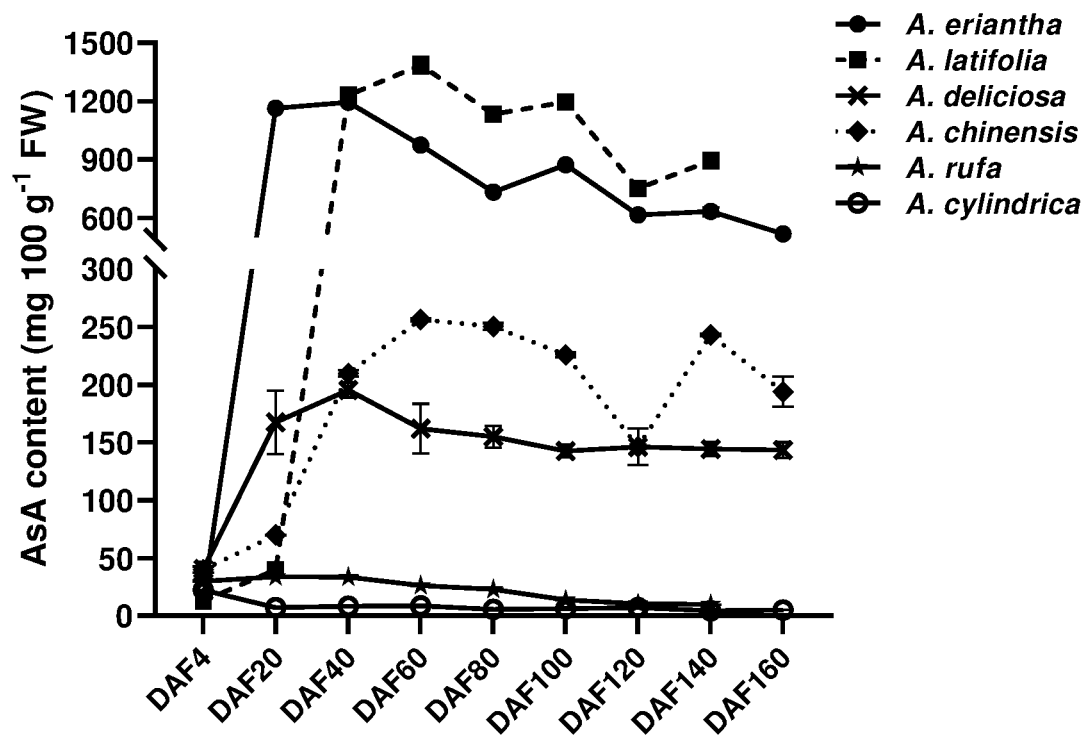
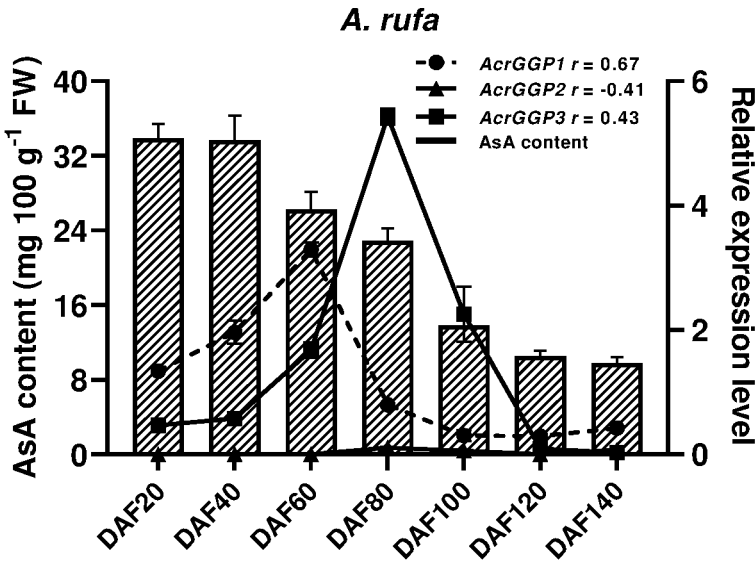


FIGURE 2

A



B

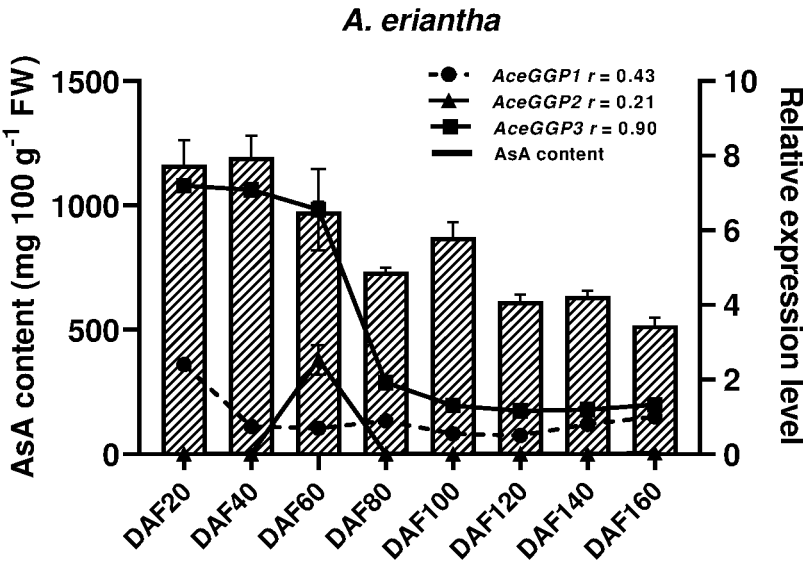


FIGURE 3

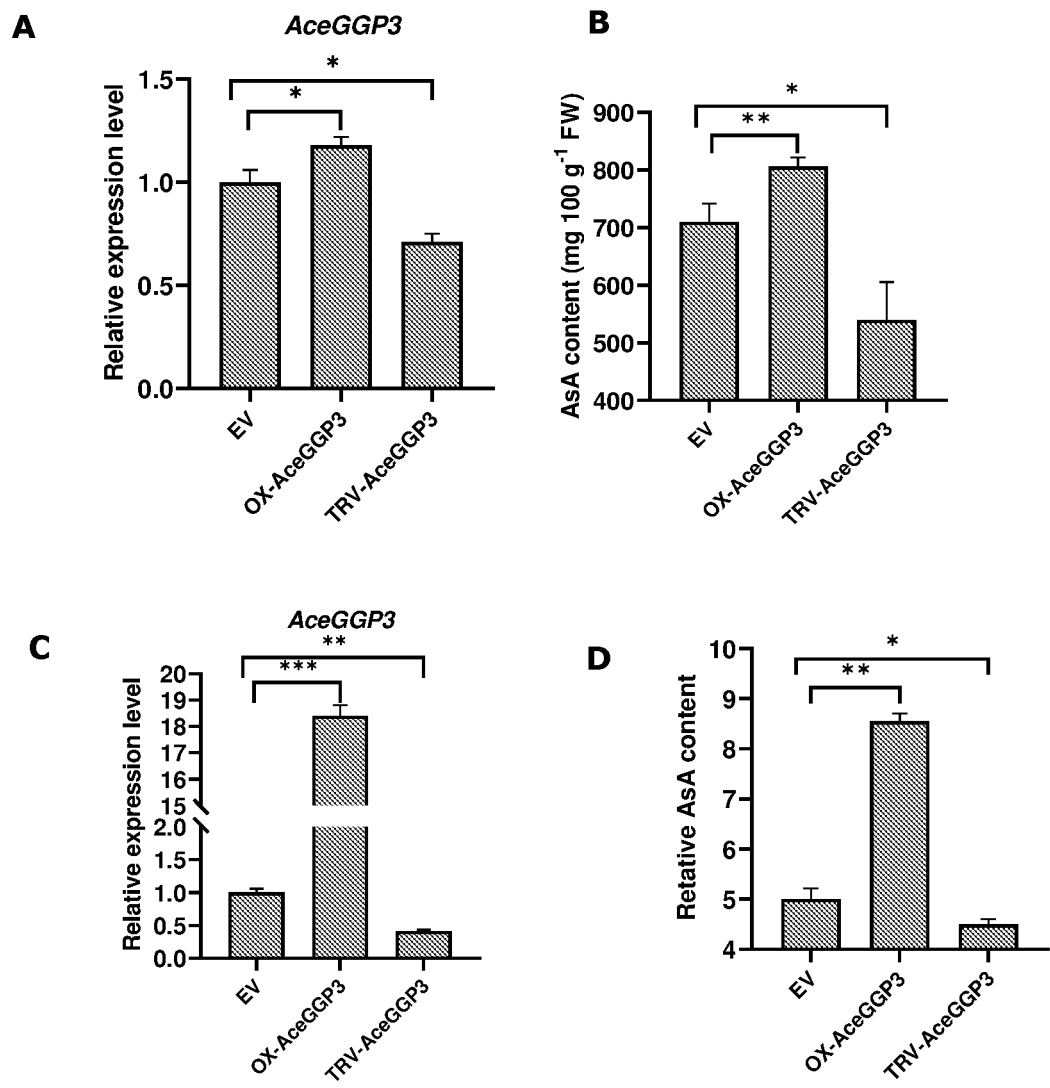


FIGURE 4

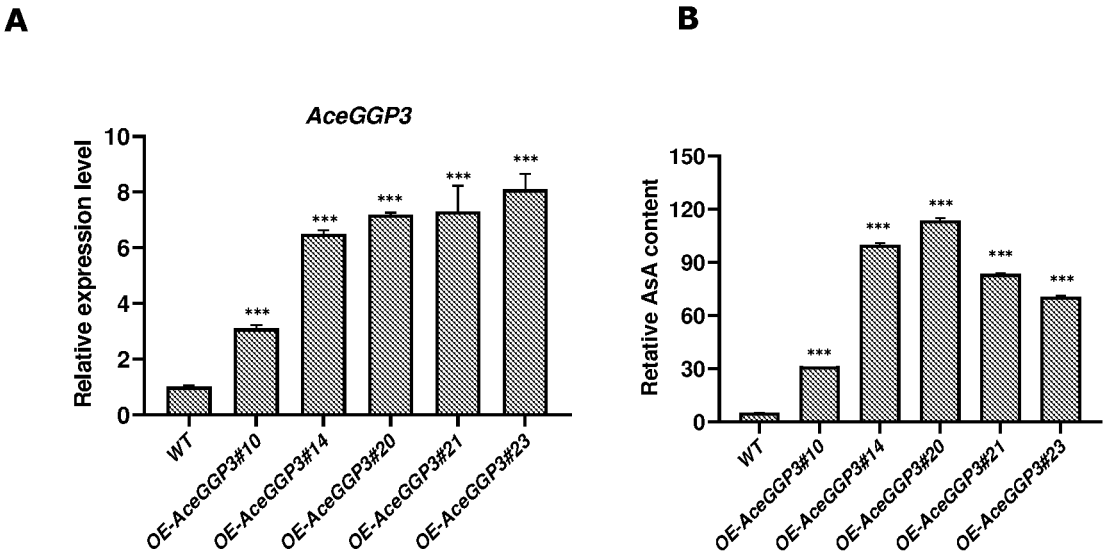


FIGURE 5

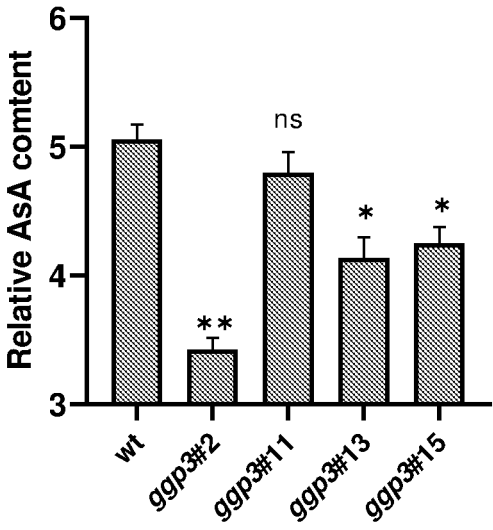
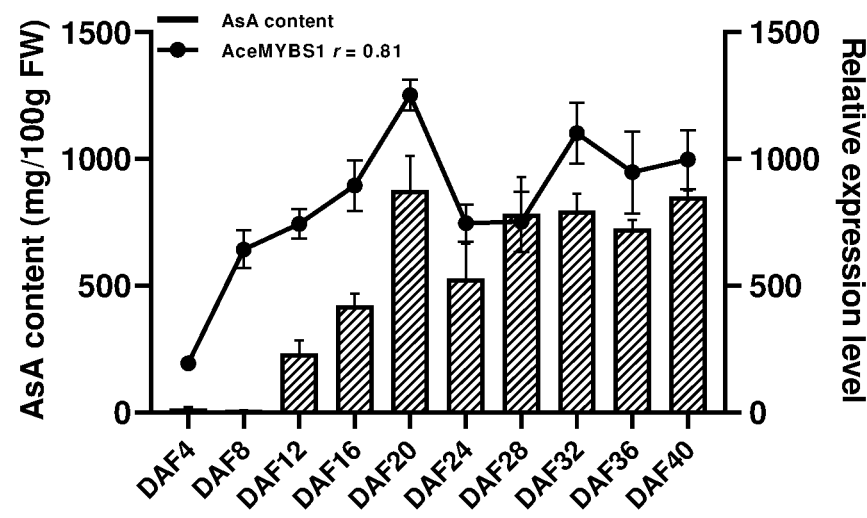


FIGURE 6

A



B

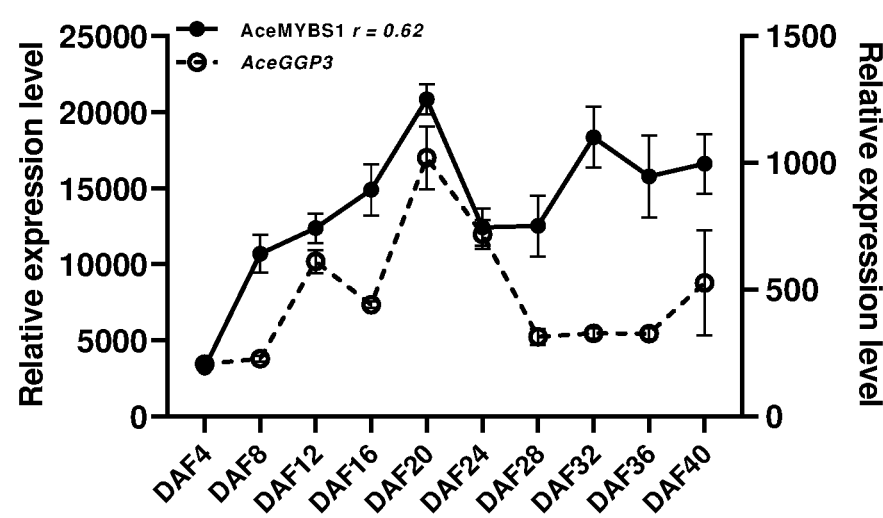
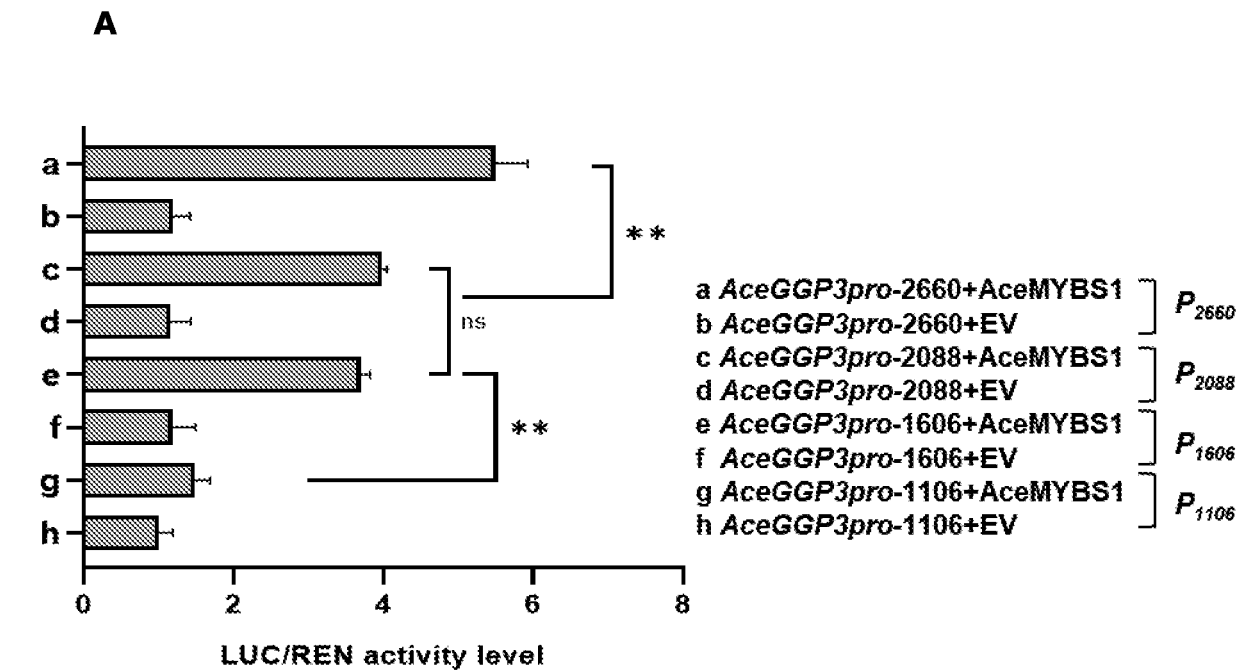


FIGURE 7



B

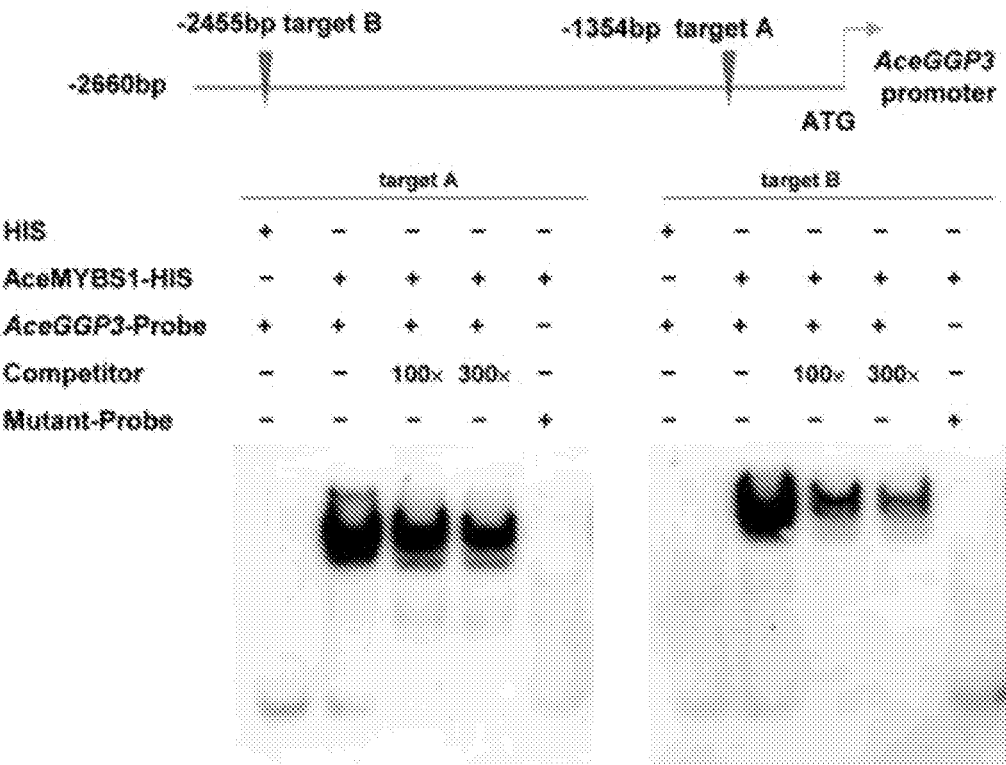


FIGURE 8

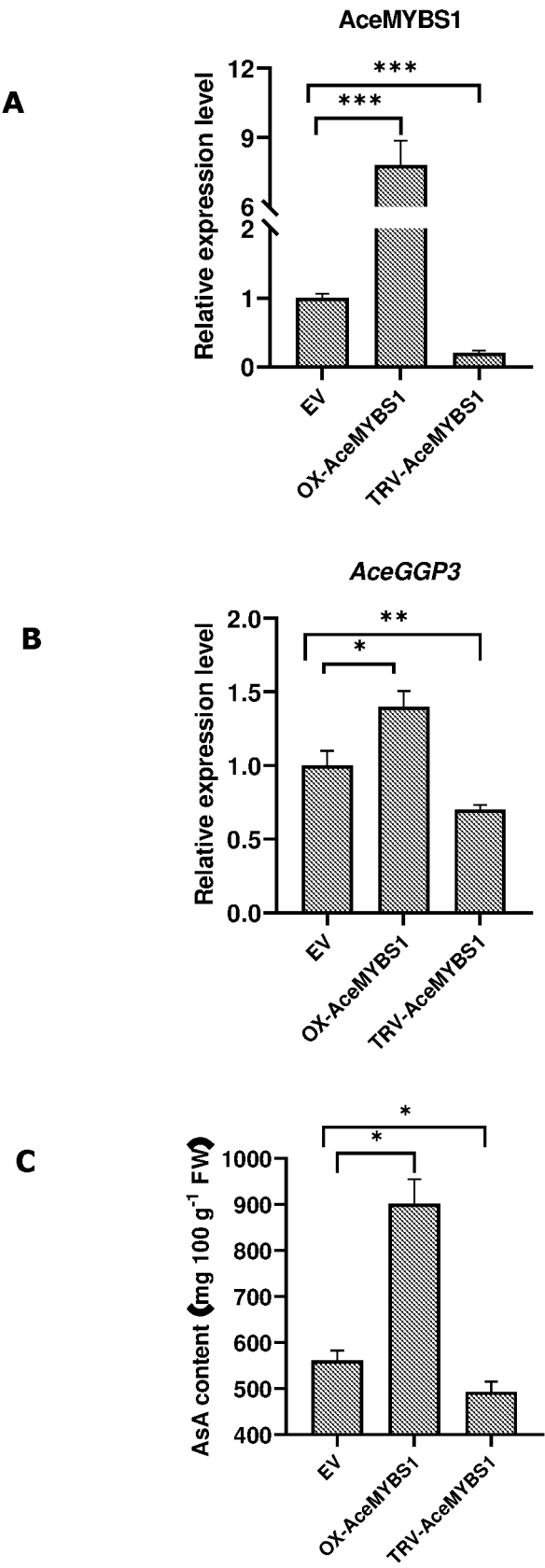


FIGURE 9

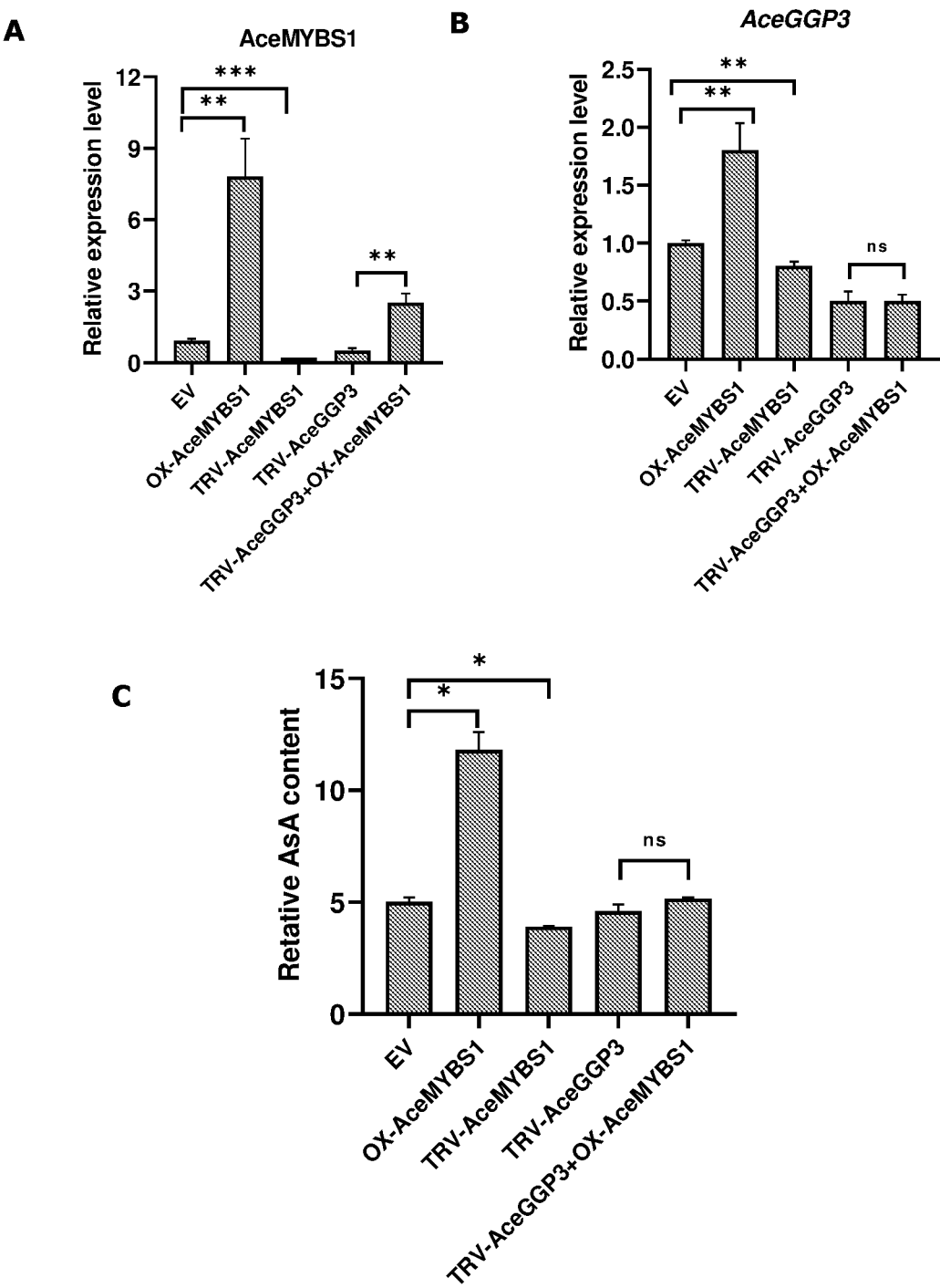


FIGURE 10

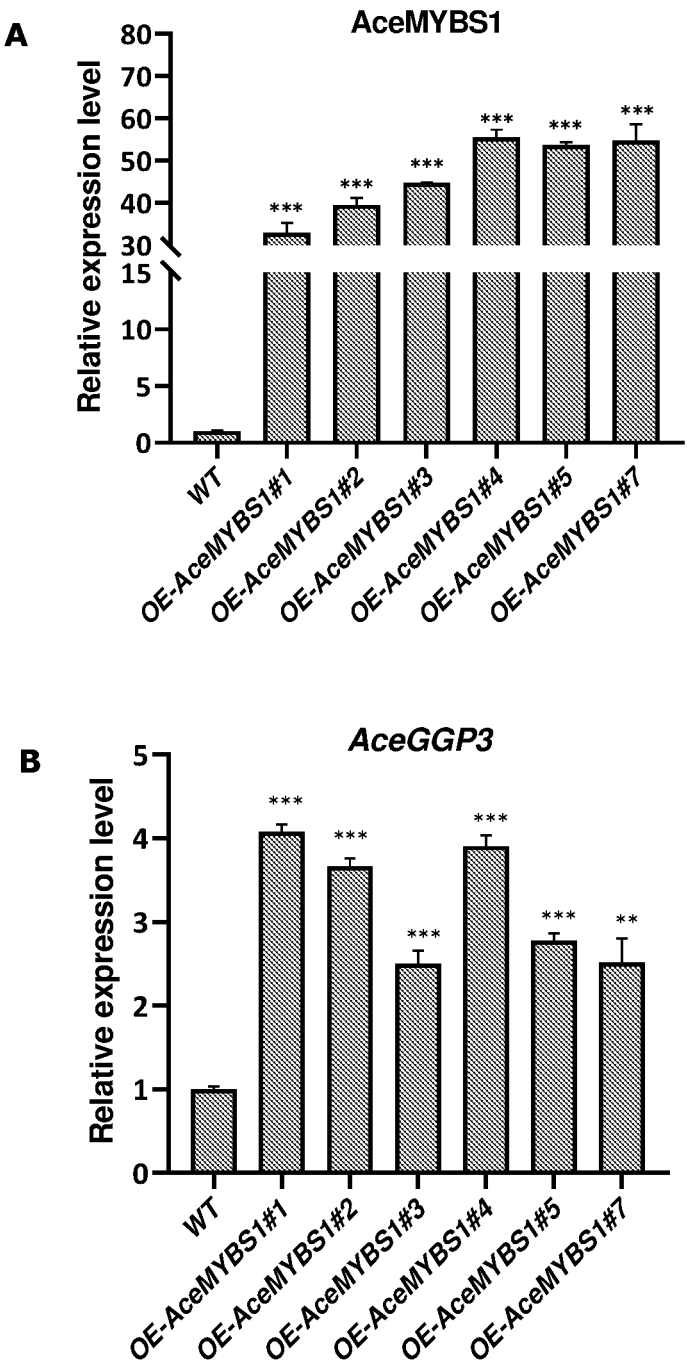


FIGURE 11

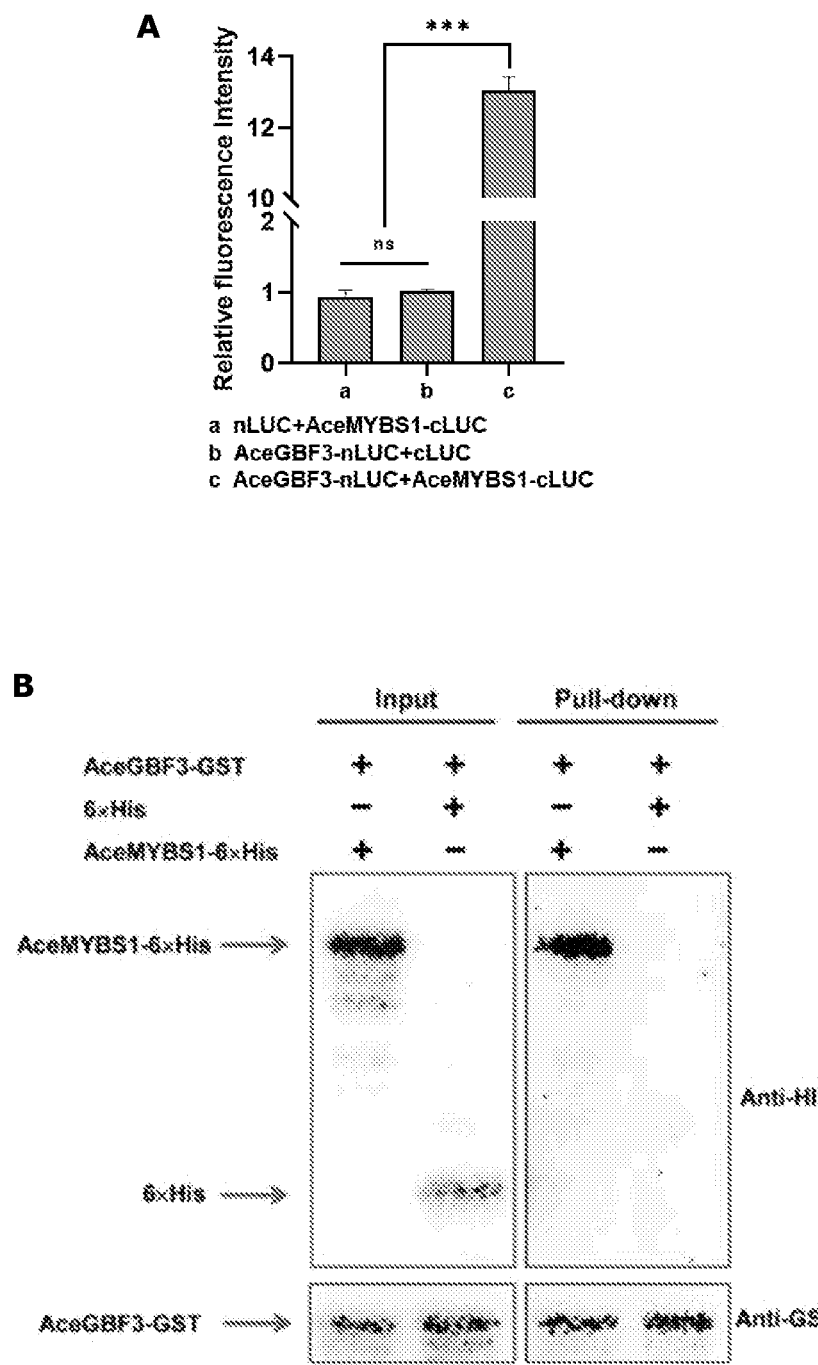
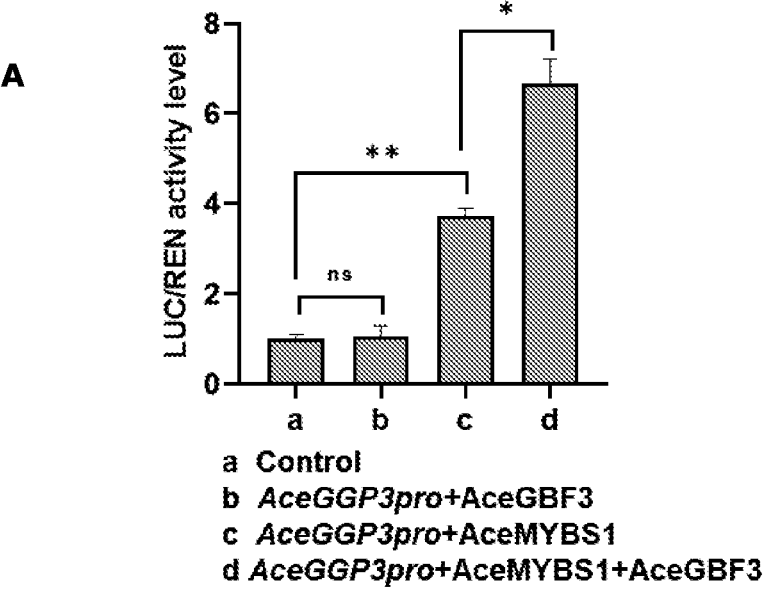
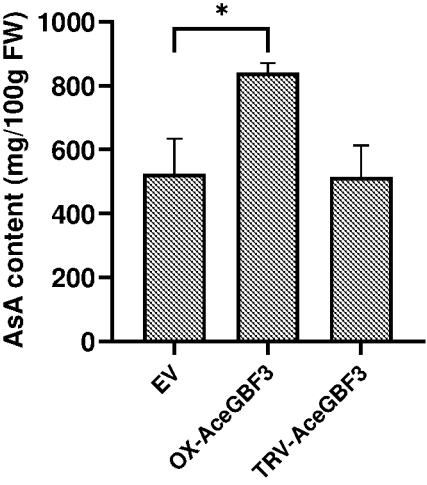


FIGURE 12



B



C

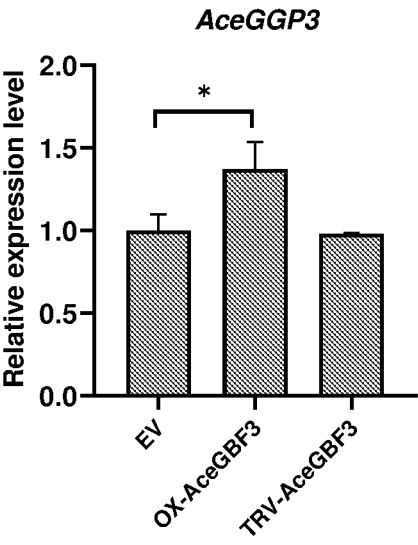


FIGURE 12 continued

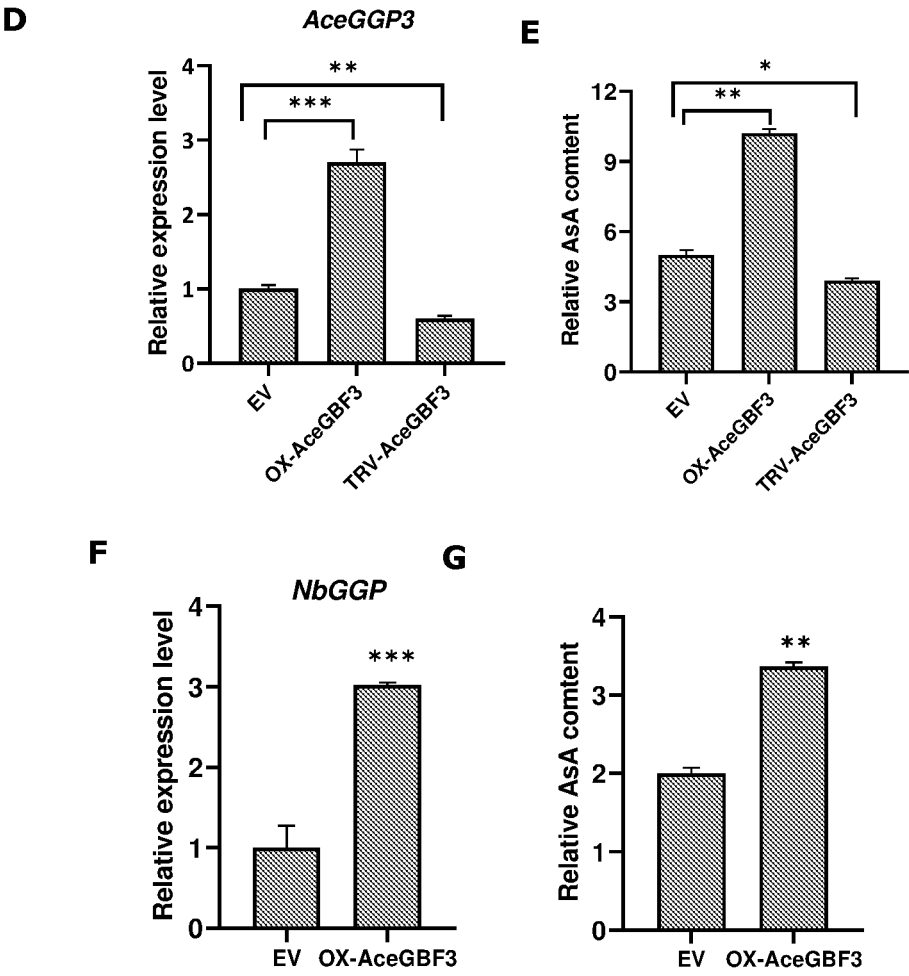


FIGURE 13

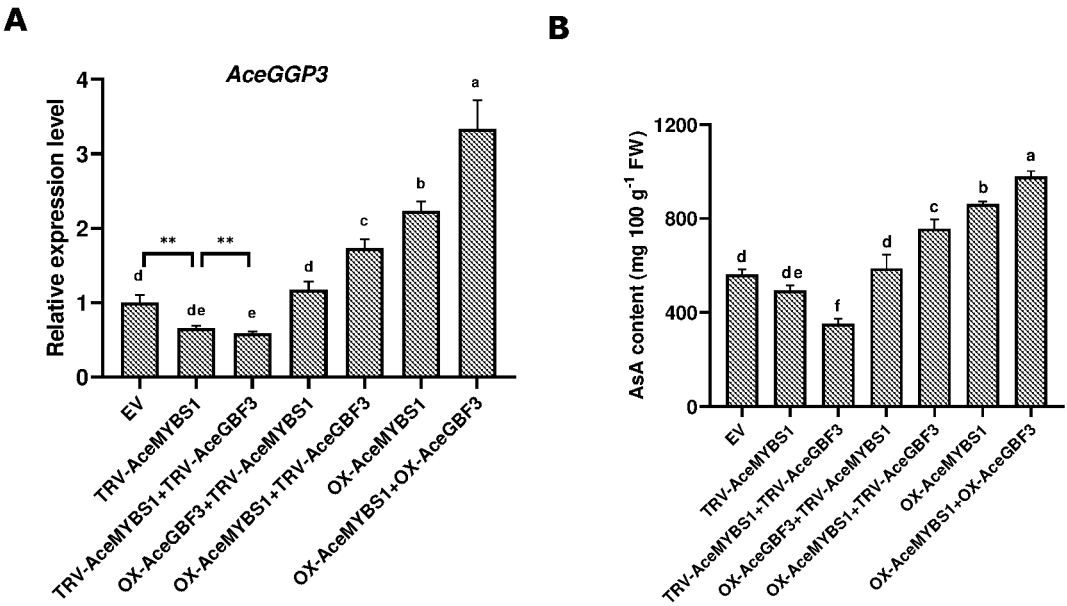


FIGURE 14

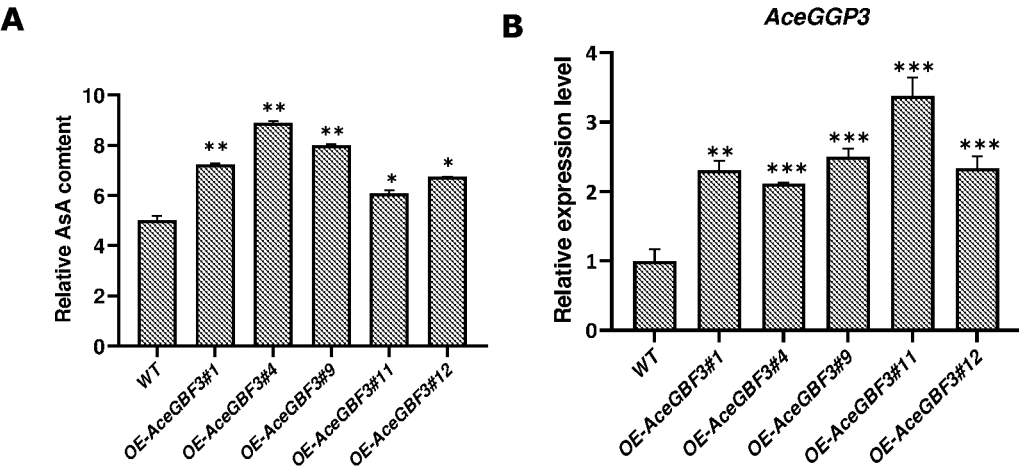
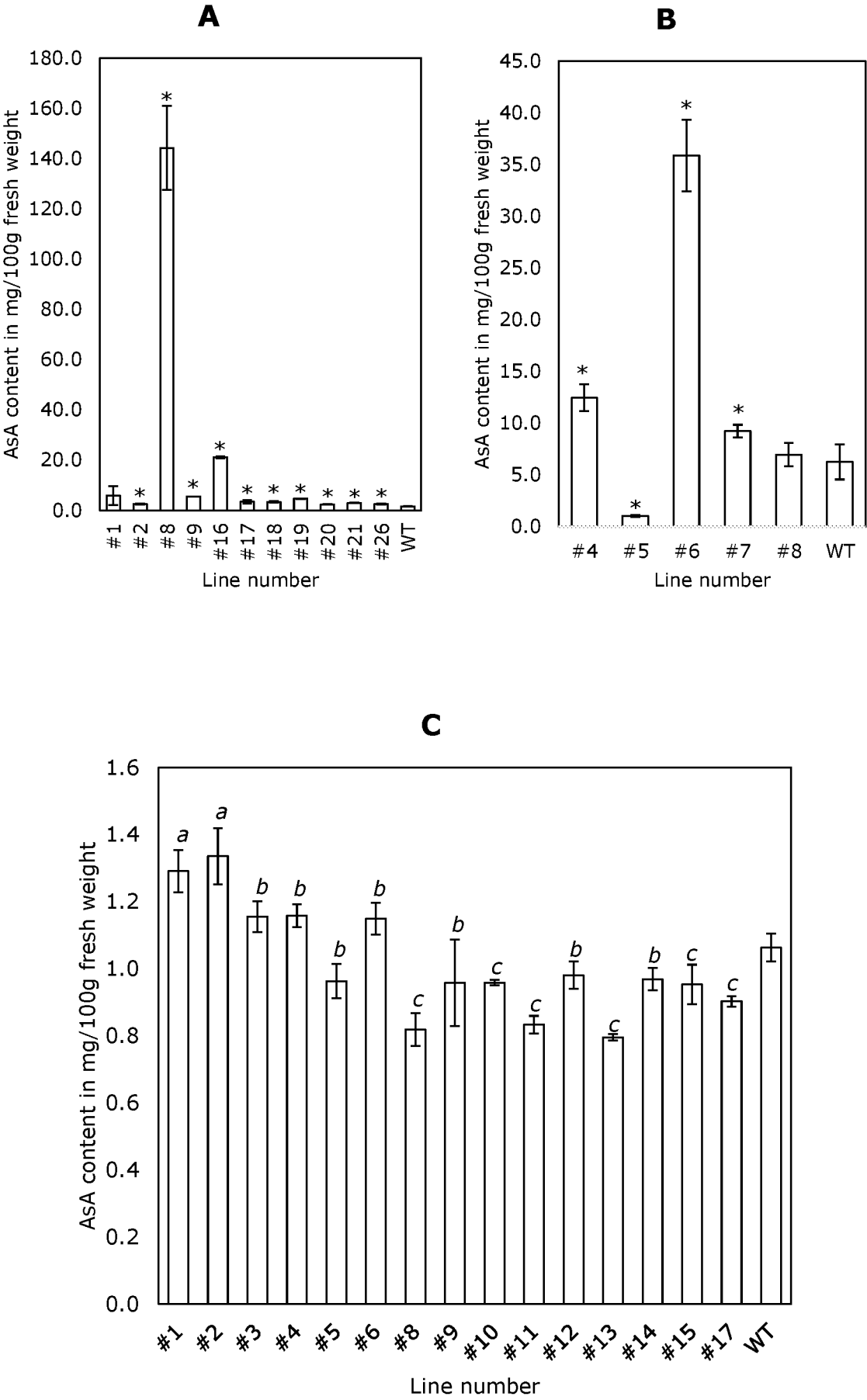


FIGURE 15



Sequence Listing

1	Sequence Listing Information	
1-1	File Name	1010531.xml
1-2	DTD Version	V1_3
1-3	Software Name	WIPO Sequence
1-4	Software Version	2.2.0
1-5	Production Date	2023-02-13
1-6	Original free text language code	
1-7	Non English free text language code	
2	General Information	
2-1	Current application: IP Office	
2-2	Current application: Application number	
2-3	Current application: Filing date	
2-4	Current application: Applicant file reference	1057174
2-5	Earliest priority application: IP Office	AU
2-6	Earliest priority application: Application number	2022900309
2-7	Earliest priority application: Filing date	2022-02-14
2-8en	Applicant name	The New Zealand Institute for Plant and Food Research Limited
2-8	Applicant name: Name Latin	
2-9en	Inventor name	
2-9	Inventor name: Name Latin	
2-10en	Invention title	REGULATION OF GENE EXPRESSION
2-11	Sequence Total Quantity	365

3-1	Sequences	
3-1-1	Sequence Number [ID]	1
3-1-2	Molecule Type	AA
3-1-3	Length	302
3-1-4	Features Location/ Qualifiers	source 1..302 mol_type=protein organism=Actinidia eriantha
3-1-5	NonEnglishQualifier Value Residues	MSASVDWSRE EEKAFENAIA IHWAEDSKDE DHWEKIASLV PTKTMEELKK HYQLLVEDVA 60 AIESGHPVPI SYIGEEASSS SSKDHHGISG VTATASEKRS NCSYGS GFPG LGHDSTGHGG 120 KGGSRAEQER RKGIPWTEEE HRLFLLGLDK FGKGDWRSIS RNFVISRTPT QVASHAQKYF 180 IRLNSINRDR RRSSIHDTIS VNSGDALAHQ PPPITGQQPN PTSTVAAAIG PPPPPGKHRI 240 PIPPPPHMAG LGMYGTPVGH PVTAPSHHMA SAVGTPVMLP PGHHHPYVVP VAYPMAPPTM 300 HQ 302
3-2	Sequences	
3-2-1	Sequence Number [ID]	2
3-2-2	Molecule Type	AA
3-2-3	Length	302
3-2-4	Features Location/ Qualifiers	source 1..302 mol_type=protein organism=Actinidia rufa
3-2-5	NonEnglishQualifier Value Residues	MSASVDWSRE EEKAFENAIA IHWAEDSKDE GHWEKIASLV PTKTMEELKK HYQLLVEDVA 60 AIESGHPVPI SYIGEEASSS SSKDHHGISG VTATASEKRS NCSYGS GFPG LGHDSTGHGG 120 KGGSRAEQER RKGIPWTEEE HRLFLLGLDK FGKGDWRSIS RNFVISRTPT QVASHAQKYF 180 IRLNSINRDR RRSSIHDTIS VNSGDASAHQ PPPITGQQPN PTSTVAAAIG PPPPPGKHRI 240 PVPPPPHMAG LGMYGTPVGH PMAAPSHHMA SAVGTPVMLP PGHHHPYVVP VAYPMAPQTM 300 HQ 302
3-3	Sequences	
3-3-1	Sequence Number [ID]	3
3-3-2	Molecule Type	AA
3-3-3	Length	303
3-3-4	Features Location/ Qualifiers	source 1..303 mol_type=protein organism=Actinidia chinensis var. chinensis
3-3-5	NonEnglishQualifier Value Residues	MSASVDWSRE EEKTFENAIA IHWAEDSKD EDHWEKIASL VPTKTMEELK KHYQLLVEDV 60 AAIESGHPVI PSYIGEEASS SSKDHHGIS GVTATASEKR SNCSYGS GFPG GLGHSTGHG 120 GKGGSRAEQE RRGIPWTEE EHRLFLLGLE KFGKGDWRSI SRNFVISRTP TQVASHAQKY 180 FIRLNSINRD RRRSSIHDT SVNSGDASAH QPPPITGQQP NPTSTVATAI GPPPPPGKHR 240 IPVPLSPHMA GLMGYTPVGH HPVAAPSHHM ASAVGTPVML PPGHHHPYV VAYPMAPQT 300 MHQ 303
3-4	Sequences	
3-4-1	Sequence Number [ID]	4
3-4-2	Molecule Type	AA
3-4-3	Length	288
3-4-4	Features Location/ Qualifiers	source 1..288 mol_type=protein organism=Actinidia chinensis var. chinensis
3-4-5	NonEnglishQualifier Value Residues	MSASVDWSRE EEKAFENAIA IHWAEDSKD DDHWTIASL VPTKTIEELK KHYQLLVEDV 60 AAIEAGHPVI PNYIGEEASS SLSKDRHGIS GLTASEKRS CIYGS GFPG GHDSAGHGK 120 GGSRAEQERR KGIPWTEEEH RLFLGLDKF GKGDWRSIS RNFVISRTPT QVASHAQKYFI 180 RLNSINRDR RSSIHDTISV NGGDASAHQ PPIMGQQPNS TSAMAAIGPP PPHMAGLGM 240 GAPVGHVAA SSHMASAVG TPVMLPPGHH HPYVVPVAYP MAPPTMHQ 288
3-5	Sequences	
3-5-1	Sequence Number [ID]	5
3-5-2	Molecule Type	AA
3-5-3	Length	294
3-5-4	Features Location/ Qualifiers	source 1..294 mol_type=protein organism=Morella rubra
3-5-5	NonEnglishQualifier Value Residues	MSSAVVSKE EEKAFENAIA VHWVEDSAE QWEKIASVVP SKSMEELKQH YQMLVEDVRA 60 IEAGHIPLPN YAGEETSSN KDSHGSGAT ASDKKSNCGY GTGFSGLGHD SAGNGGKGGS 120

		RSEQERRKGI PWTEEEHRLF LLGLDKFGKG DWRSISRNFV ISRTPTQVAS HAQKYFIRLN 180 SMNRDRRRSS IHDITSVNNG DVASHQAPIT GQQTNMNISS VAAIGPPVKH RAQAHPGLG 240 MYGAPVGHPV AAPPGLHLASA IGTPVMLPHG HHPHPHPYV VPVAYPMAPP PMHQ 294
3-6	Sequences	
3-6-1	Sequence Number [ID]	6
3-6-2	Molecule Type	AA
3-6-3	Length	294
3-6-4	Features Location/ Qualifiers	source 1..294 mol_type=protein organism=Quercus suber
3-6-5	NonEnglishQualifier Value Residues	MSSGGIWSKE EEKTFENAIA MHWVDEDCKE QWEKIASMVP SKTMDELKQH YQLLVEDVSA 60 IEAGHIPLPN YVGEEASSN KDCHGSSGAM ASDKRSNCGF GSGFPGLGHD STGHGKGKGS 120 RSEQERRKGI PWTEEEHRLF LLGLDKFGKG DWRSISRNFV ISRTPTQVAS HAQKYFIRLN 180 SMNRDRRRSS IHDITSVNNG DVASHQAPIT GQQTNMNPSS VATIGPPVKH RAQHPMPGLG 240 MYGTSVGHPV AAPPVHMASA VGTPVMLPPG HHPHSHPPYV MPVAYPMAPP PMHQ 294
3-7	Sequences	
3-7-1	Sequence Number [ID]	7
3-7-2	Molecule Type	AA
3-7-3	Length	295
3-7-4	Features Location/ Qualifiers	source 1..295 mol_type=protein organism=Juglans microcarpa x Juglans regia
3-7-5	NonEnglishQualifier Value Residues	MTSEVVWSKE EEKAFENAIA MHWVEDDSKE HWETIASMVP SKSMEELKQH YQVLVEDVRA 60 IEAGHIPLPN YLGEEASSN KDCHGSSGAT ASDKRSNCGY GSGFSGLGHD STGHGKGKGS 120 RSEQERRKGI PWTEEEHRLF LLGLDKFGKG DWRSISRNFV ISRTPTQVAS HAQKYFIRLN 180 SMNRDRRRSS IHDITSVNNG DVTSHQAPIT GQQIHMLSS VAAIRPPAKH RAPPHPGLG 240 MYGMPVGHPV AAPPGHMASA VGTPVMLPPG PHPHPHPHY VVPVAYPMAP PPMHQ 295
3-8	Sequences	
3-8-1	Sequence Number [ID]	8
3-8-2	Molecule Type	AA
3-8-3	Length	289
3-8-4	Features Location/ Qualifiers	source 1..289 mol_type=protein organism=Nyssa sinensis
3-8-5	NonEnglishQualifier Value Residues	MSASPSTSVV WSREEEKAFE NAIALHWAED SKEQWDKIAS MVPSKTIEEL KQHYQLLVED 60 VSAIEAGHIP LPNYIGEEAS PSSKEHHSFS GVTGTDKRSN CSYGSGFSEL AHDSNGQGGK 120 GGSRSDQERR KGIPWTEEEH RLFLGLDKF GKGDWRSISR NFVISRTPTQ VASHAQKYFI 180 RLNSMNRDRR RSSIHDITSV NNGDVSAIT GQQTQNPNS AAAIGSVVKH RTQPNVPGLG 240 MYGAQVGHPV AGPGHMASAV GTPVMLPPGH HPPYVVPVAY PMAPPTQHQ 289
3-9	Sequences	
3-9-1	Sequence Number [ID]	9
3-9-2	Molecule Type	AA
3-9-3	Length	298
3-9-4	Features Location/ Qualifiers	source 1..298 mol_type=protein organism=Rhododendron simsii
3-9-5	NonEnglishQualifier Value Residues	MSSASIEWSR EEEKAFENAI AIHWSDDSK E QNWSKIASMV PNKTVEELKQ HYELLLEDVI 60 AIEAGKVPVP KYVGEEASSS STKDFHGSCG GDTSSEKRSN CGFGSGYSGL GQDPNGHGGK 120 GGGSRAEQER RKGPWTEEE HRLFLGLDK FGKGDWRSIS RNFVISRTPT QVASHAQKYF 180 IRLNSMNRDR RSSIHDITS INGGDPSTQP PPITGHQTNA AANPTNVA AI GPPPIKHRIP 240 PPNMAGLGM GAPMGHPVAA GPSHMASAVG TPVMLTPGHH HPYVPLAYP MAPPTMHQ 298
3-10	Sequences	
3-10-1	Sequence Number [ID]	10
3-10-2	Molecule Type	AA
3-10-3	Length	297
3-10-4	Features Location/ Qualifiers	source 1..297 mol_type=protein organism=Carya illinoensis
3-10-5	NonEnglishQualifier Value Residues	MSSVVWSKE EDKAFENAIA VHWVEDSNE HWEKIASMVP SKSMDELKQH YQMLADDVRA 60 IEEGHIPLPS YVGQEATSSS KDNHGLSGSS DSNKRSNCGH GSKFSGLGND SAGHGKGKGS 120

		RSEQERRKGI PWTEEEHRLF LLGLDKFGKG DWRSISRNFV ISRTPTQVAS HAQKYFIRLN 180 SMNRDRRRSS IHDITSLNNG DVTSHQAPIT GQQNNMMSS AATIGPPMKH RAPAPLHMPG 240 LGMYGTPVGH PVAAPPGHMA SAVGTPVMLP PGHHPHSHPP PYVVPVAYPM PPPPMHQ 297
3-11	Sequences	
3-11-1	Sequence Number [ID]	11
3-11-2	Molecule Type	AA
3-11-3	Length	291
3-11-4	Features Location/ Qualifiers	source 1..291 mol_type=protein organism=Juglans regia
3-11-5	NonEnglishQualifier Value Residues	MTSVVVWSKG EEKAFENAIA MHWVEDSKE HWETIASMVP SKSMEELKQH YQVLVEDVRA 60 IEAGHIPLPN YLGEEASSN KDCHGSSGAT ASDKRSNCGY GSGFSGLGHD STGHGGKGGG 120 RSEQERRKGI PWTEEEHRLF LLGLDKFGKG DWRSISRNFV ISRTPTQVAS HAQKYFIRLN 180 SMNRDRRRSS IHDITSVNNG DVTSHQAPIT GQQIHMLSS VAAIGPPAKH RAPPHPGLG 240 MYGMPVGHPV AAPPGHMAS VGTVPMLTPG PHPHPYVVPV AYPMAPPPMH Q 291
3-12	Sequences	
3-12-1	Sequence Number [ID]	12
3-12-2	Molecule Type	AA
3-12-3	Length	294
3-12-4	Features Location/ Qualifiers	source 1..294 mol_type=protein organism=Nyssa sinensis
3-12-5	NonEnglishQualifier Value Residues	MSASTSTSVV WSREEEKAFE NAIALHWAED SKEQWEKIAS MVPSKTIEEL KQHYQLLVED 60 VSAIEAGDIP LPNYIGEEAS SSSTKDHHSF SGLTTPDKRS NCSFGSRFSG LAHDSTGQGA 120 KGGSRSDQER RKGIPWTEEE HRLFLGLDK FGKGDWRSIS RNFVISRTPT QVASHAQKYF 180 IRLNSMNRDR RRSSIHDITS LNNGDVVPIT GQQQSNPSPS NAVAAGPPV KHRSHPNMP 240 SLGMYGAPIG HPVAAAAPGH MASAVGTPVM LPPGHHPYV VPVAYPMPPP AMHH 294
3-13	Sequences	
3-13-1	Sequence Number [ID]	13
3-13-2	Molecule Type	AA
3-13-3	Length	353
3-13-4	Features Location/ Qualifiers	source 1..353 mol_type=protein organism=Carya illinoensis
3-13-5	NonEnglishQualifier Value Residues	MPTSLILHT FRSTSYYPH LLRLWLCLLF SNFTTKELSL QNTFSCSKLD GCKAKTMSSV 60 VVWSKEEDKA FENAIHVHW DEDSNEHWEK IASMVPSKSM DELKQHYQML ADDVRAIEEG 120 HIPLPSYVGQ EATSSSKDNH GLSGSSDSNK RSNCGHGSKF SGLGNDGAGH GKGKGSRSEQ 180 ERRKGIPWTE EEHRLFLGLL DKFGKGDWRS ISRNFVISRT PTQVASHAQK YFIRLNSMNR 240 DRRRSSIHDI TSLNNGDVTSHQAPITGQQN NMNMSSAATI GPPMKHRAPA PLHMPGLGMY 300 GTPVGHPVAA PPGHMASAVG TPVMLPPGHH PSHPPPYVVPV VAYPMPPP MHQ 353
3-14	Sequences	
3-14-1	Sequence Number [ID]	14
3-14-2	Molecule Type	AA
3-14-3	Length	296
3-14-4	Features Location/ Qualifiers	source 1..296 mol_type=protein organism=Juglans regia
3-14-5	NonEnglishQualifier Value Residues	MSSVIVWSKE EDKAFENAIA THWVEDSNE HWEKIASMVP SKSMDELKQH YQMLVDDVRA 60 IEEGHIPLPN YVQGEATSSS KDNHGLSGAT DSNKRSNCGY GSRFSGLGND SAGHGGKGGG 120 RSEQERRKGI PWTEEEHRLF LLGLDKFGKG DWRSISRNFV ISRTPTQVAS HAQKYFIRLN 180 SMNRDRRRSS IHDITSLNNG DVTSHQAPIT GQQNNMMSS VAAIGPPVKH RAPAPLHMPG 240 LGMYGTPVGH PIAAPGHMAS AVGTPIMLPP GHHPHSHPPP YVVPVAYPMR PPPMHQ 296
3-15	Sequences	
3-15-1	Sequence Number [ID]	15
3-15-2	Molecule Type	AA
3-15-3	Length	353
3-15-4	Features Location/ Qualifiers	source 1..353 mol_type=protein organism=Carya illinoensis
	NonEnglishQualifier Value	

3-15-5	Residues	MPTSLILHT FRSTSYYPH LLRLWLCLF SNFTTKELSL QNTFSCSKLD GCKAKTMSSV 60 VVWSKEEDKA FENAIAVHWV DEDSNEHWEK IASMVPSKSM DELKQHYQML ADDVRAIEEG 120 HIPLPSYVGQ EAASSSKDNH GLSGSSDSNK RSNCGHGSKF SGLGND SAGH GGKGGSRSEQ 180 ERRKGIPWTE EEHRLFLLGL DKFGKGDWRS ISRN FVISRT PTQVASHAQK YFIRLNSMNR 240 DRRRSSIHI TSLNNGDVTS HQAPITGQQN NMNMSSAATI GPPMKHRAPA PLHMPGLGMY 300 GTPVGHPVAA PPGHMASAVG TPVMLPPGHH PSHPPPYVV PVAYPMPPPP MHQ 353
3-16	Sequences	
3-16-1	Sequence Number [ID]	16
3-16-2	Molecule Type	AA
3-16-3	Length	353
3-16-4	Features Location/ Qualifiers	source 1..353 mol_type=protein organism=Carya illinoensis
3-16-5	NonEnglishQualifier Value Residues	MPTSLILHT FRSTSYYPH LLRLWLCLF SNFTTKELSL QNTFSCSKLD GCKAKTMSSV 60 VVWSKEEDKA FENAIAVHWV DEDSNEHWEK IASMVPSKSM DELKQHYQML ADDVRAIEEG 120 HIPLPSYVGQ EATSSSKDNH GLSGSSDSNK RSNCGHGSKF SGLGND SAGH GGKGGSRSEQ 180 ERRKGIPWTE EEHRLFLLGL DKFGKGDWRS ISRN FVISRT PTQVASHAQK YFIRLNSMNR 240 DRRRSSIHI TSLNNGDVTS HQAPITGQQN NMNMSSAATI GPPMKHRAPA PLHMPGLGMY 300 GTPVGHPVAA PPGHMASAVG TPVMLPPGHH PSHPPPYVV PVAYPMPPPP MHQ 353
3-17	Sequences	
3-17-1	Sequence Number [ID]	17
3-17-2	Molecule Type	AA
3-17-3	Length	299
3-17-4	Features Location/ Qualifiers	source 1..299 mol_type=protein organism=Hevea brasiliensis
3-17-5	NonEnglishQualifier Value Residues	MESAISWSRE EEKAFENAI THWTKEEEDS KELWEKIASL VPSKSIQELK LHYQLLEDV 60 SAIEAGNIPL PNYVGEEATS SAKDSHGSSG AVATDKRLNC GYGSGFFGLG HNSSGHGGKG 120 GSRTDQERKK GIPWTEEEHR LFLLGLDKFG KGDWRSISRNFVISRTPTQV ASHAQKYFIR 180 LNSMNRDRRR SSIHDITNVN NGDVSSHQAP ITGQQANTNP AGAVATAAIG SSVKDRVQPH 240 MPSLGMYGTP VGHPVAAPP G HMVSAVGTPV MLPPGHHHP HPPYVVPVAY PMAPPQTMH 299
3-18	Sequences	
3-18-1	Sequence Number [ID]	18
3-18-2	Molecule Type	AA
3-18-3	Length	298
3-18-4	Features Location/ Qualifiers	source 1..298 mol_type=protein organism=Vaccinium darrowii
3-18-5	NonEnglishQualifier Value Residues	MSSASVEWSR EEEKAFENAI ATHWTADSND QNWSKFASMV PNKTIEELKQ HYEVLVEDVT 60 AIEAGNVPIPI KYVGEEASSS STKDFHGSCG GVTSTEKRLN CGFGSGYSGL GQDPNGHGGK 120 GGGSRAEQER RKGIPWTEEE HRLFLLGLEK FGKGDWRSIS RN FVISRTPTQV QVASHAQKYF 180 IRLNSMNRDR RRSSIHDITS INGGDPSTQP PPITGHQTNA VANPTHVA AV GQPTIKHRIP 240 PPNMAGMGMY GAPMGHPVAA GPSHMASAVG TPVMLPPGHH HPYVVPVAYP MAPPTMHQ 298
3-19	Sequences	
3-19-1	Sequence Number [ID]	19
3-19-2	Molecule Type	AA
3-19-3	Length	317
3-19-4	Features Location/ Qualifiers	source 1..317 mol_type=protein organism=Camellia sinensis
3-19-5	NonEnglishQualifier Value Residues	MSSSSSASAI VVWCKEEKA FENAIAMHWS EDDSSKEQWD KIASMVPSKT IEELKQHYQL 60 LVDDVAAIEA GHIPVPSYVG EEASSSSSKD HHHGFHGV TN TAVDKRSSCN YGSGFSGLAA 120 VAHDSAGHSG GGKGGSRSEQ ERRKGIPWTE EEHRLFLLGL DKFGKGDWRS ISRN FVISRT 180 PTQVASHAQK YFIRLNSMNR DRRRSSIHI TSVNGGGGSG DVPIQPPIT GQQTNPANPT 240 TTSIAIGPPPV KHRAPTNPMP LGMYGGAPVG HPVAAPGHMG HMGSAVGTPV MLPPGHHHP 300 PYVVPVAYPM APPTMHQ 317
3-20	Sequences	
3-20-1	Sequence Number [ID]	20
3-20-2	Molecule Type	AA
3-20-3	Length	291

3-20-4	Features Location/ Qualifiers	source 1..291 mol_type=protein organism=Carya illinoensis
3-20-5	NonEnglishQualifier Value Residues	MTSVAVWSKE EEKAFENAIA MHWVDEDSKE HWEKIASMVP SKSVEELKQH YQMLVEDVRA 60 IEAGHIPLPN YLGEETSSSN KDCHGSSGAT ASDKRSNCGY GTGFSGLGHD SAGHGKGKGS 120 RSEQERRKGI PWTEEEHRLF LLGLDKFGKG DWRSISRNFV ISRTPTQVAS HAQKYFIRLN 180 SMNRDRRRSS IHDITSVNNG DVTSHQAPIT GQQIHMLSS VAAIGPPAKH RTPLHMPGMY 240 GTPVGHPVVA PPGHMASAVG TPVMLPPGPH PHPHPYVVPV AYPMAPPPMH Q 291
3-21	Sequences	
3-21-1	Sequence Number [ID]	21
3-21-2	Molecule Type	AA
3-21-3	Length	291
3-21-4	Features Location/ Qualifiers	source 1..291 mol_type=protein organism=Carya illinoensis
3-21-5	NonEnglishQualifier Value Residues	MTSVAVWSKE EEKAFENAIA MHWVDEDSKE HWEKIASMVP SKSVEELKQH YQMLVEDVRA 60 IEAGHIPLPN YLGEETSSSN KDCHGSSGAT ASDKRSNCGY GTGFSGLGHD SAGHGKGKGS 120 RSEQERRKGI PWTEEEHRLF LLGLHKFGKG DWRSISRNFV ISRTPTQVAS HAQKYFIRLN 180 SMNRDRRRSS IHDITSVNNG DVTSHQAPIT GQQIHMLSS VAAIGPPAKH RTPLHMPGMY 240 GTPVGHPVAA PPGHMASAVG TPVMLPPGPH PHPHPYFVPV AYPMAPPPMH Q 291
3-22	Sequences	
3-22-1	Sequence Number [ID]	22
3-22-2	Molecule Type	AA
3-22-3	Length	298
3-22-4	Features Location/ Qualifiers	source 1..298 mol_type=protein organism=Juglans microcarpa x Juglans regia
3-22-5	NonEnglishQualifier Value Residues	MSSVIVWSKE EDKAFENAIA THWVDEDSNE HWEKIASMVP SKSMDELKQH YQMLVDDVRA 60 IEEGHIPLPN YVGQEVTSST KDNHGLSRAT DSNKRSNCGY GSRFSGLGND SAGHGKGKGS 120 LRSEQERRKG IPWTEEEHRL FLLGLDKFGK GDWRSISRNF VISRTPTQVA SHAQKYFIRL 180 NSMNRDRRRS SIHDITSLNN GDVTSHKVPI TGQQNNMNMS SAATIGPPVK HRAAPLHMP 240 GLGMYGTPVG HPVAAPPGHM ASAVGTPVML PPGHHPHSH PYPVVPVAYP MPPPPVHQ 298
3-23	Sequences	
3-23-1	Sequence Number [ID]	23
3-23-2	Molecule Type	AA
3-23-3	Length	387
3-23-4	Features Location/ Qualifiers	source 1..387 mol_type=protein organism=Camellia sinensis var. sinensis
3-23-5	NonEnglishQualifier Value Residues	MWSSFPLIHL KPILILHSIS PLLPLLIRI SPPLPLPLS LFLQPKNKN FPQQNPSTYS 60 PLSHTQFIKN MSSSSASAI VVWCKEEKA FENAIAMHWS EDDSSKEQWD KIASMVPSKT 120 IEELKQHYQL LVDDVAAIEA GHIPVPSYVG EEASSSSKD HHHGFHGVN TAVDKRSSCN 180 YGSFSGGLAA VAHDSAGHSG GKGGSRSEQ ERRKGIPWTE EEHRLFLLGL DKFGKGDWRS 240 ISRNFVISRT PTQVASHAQK YFIRLNSMNR DRRRSSIHI TSVNGGGSG DVPIQQPPIT 300 GQQTNPANPT TTAIGPPP V KHRAPTNMPG LGMYGGAPVG HPVAAPGHMG HMGSAGVTPV 360 MLPPGHHHP PYVVPVAYPM APPTMHQ 387
3-24	Sequences	
3-24-1	Sequence Number [ID]	24
3-24-2	Molecule Type	AA
3-24-3	Length	317
3-24-4	Features Location/ Qualifiers	source 1..317 mol_type=protein organism=Camellia sinensis
3-24-5	NonEnglishQualifier Value Residues	MSSSSASAI VVWCKEEKA FENAIAMHWS EDDSSKEQWD KIASMVPSKT IEELKQHYQL 60 LVDDVAAIEA GHIPVPSYVG EEASSSSKD HHHGFHGVN TAVDKRSSCN YGSFSGGLAA 120 VAHDSAGHSG GKGGSRSEQ ERRKGIPWTE EEHRLFLLGL DKFGKGDWRS ISRNFVISRT 180 PTQVASHAQK YFIRLNSMNR DRRRSSIHI TSVNGGGSG DVPIQQPPIT GQQTNPANPT 240 TTAIGPPP V KHRAPPNMP LGMYGGAPVG HPVAAPGHMG HMGSAGVTPV MLPPGHHHP 300 PYVVPVAYPM APPTMHQ 317
3-25	Sequences	

3-25-1	Sequence Number [ID]	25
3-25-2	Molecule Type	AA
3-25-3	Length	297
3-25-4	Features Location/ Qualifiers	source 1..297 mol_type=protein organism=Ricinus communis
3-25-5	NonEnglishQualifier Value Residues	MESATITWSR EEDIAFENAI ATHWIEDDSE EQWEKIASMV PSRNIEELKQ HYRLLEDVD 60 AIEAGNVPLP NYVGEETTSS SSKDSHGFSG AVTTDKRLNC GFGSGFMGLG PNSSGHGKG 120 GSRADQERRK GIPWTEEEHR LFLLGLDKFG KGDWRSISRNV FVLSRTPTQV ASHAQKYFIR 180 LNSMNRDRRR SSIHDITSVN NGEVSSHQAP ITGQQGNTNP AAGAPAMGSA VKHRAQHHP 240 SIGMYGAPVG HPVAAPPGHM ASAVGTPVML PPPGHHPPHP YVLPVAYPMP PPQTMHQ 297
3-26	Sequences	
3-26-1	Sequence Number [ID]	26
3-26-2	Molecule Type	AA
3-26-3	Length	291
3-26-4	Features Location/ Qualifiers	source 1..291 mol_type=protein organism=Carya illinoensis
3-26-5	NonEnglishQualifier Value Residues	MTSVAVWSKE EEKAFENAI MHWVEDSKE HWEKIASMVP SKSVEELKQH YQMLVEDVRA 60 IEAGHIPLPN YLGEETSSSN KDCHGSSGAT ASDKRSNCGY GTGFSGLGHD SAGHGKGKGS 120 RSEQERRKGI PWTEEEHRLF LLGLDKFGKG DWRSISRNV ISRTPTQVAS HAQKYFIRLN 180 SMNRDRRRSS IHDITSVNNG DVTSHQAPIT GQQIHMNLSS VAAIGPPAKH RTPLHMPGMY 240 GTPVGHPVAA PPGHMASAVE TPVMLPPGPH PHPHPYVVPV AYPMAPPPMH Q 291
3-27	Sequences	
3-27-1	Sequence Number [ID]	27
3-27-2	Molecule Type	AA
3-27-3	Length	294
3-27-4	Features Location/ Qualifiers	source 1..294 mol_type=protein organism=Quercus lobata
3-27-5	NonEnglishQualifier Value Residues	MSSEIWSKE EEKTFENAI MHWVEDCKE QWEKIASMVP SKTMDCLKQH YQLLEDVSA 60 IEAGHIPLPN YVGEASSSN KDCHGFSGAM ASDKRSNCGF GSGFSGLGHD STGHGKGKGS 120 RSEQERRKGI PWTEEEHRLF LLGLDKFGKG DWRSISRNV ISRTPTQVAS HAQKYFIRLN 180 SMNRDRRRSS IHDITSVNNG DVASHQAPIT GQQTNMNPSS VATIGPPVKH RAQPHMPGLG 240 MYGTSVGHPV AAPPVHMASA VGTPVMLPPG HPHSHPPYV MPVAYPMAPP PMHQ 294
3-28	Sequences	
3-28-1	Sequence Number [ID]	28
3-28-2	Molecule Type	AA
3-28-3	Length	290
3-28-4	Features Location/ Qualifiers	source 1..290 mol_type=protein organism=Capsicum baccatum
3-28-5	NonEnglishQualifier Value Residues	MSPSDWSREE EKAFENAIAM HWVENCEQQW EKFGMLVPTK TIDELKQHYQ LLVDDVTAIE 60 AGQVIPNYK GEEASSSTK EGNLGSAGS VDRRSNCGYS NGFSGTNHDP VGHGKGKNSR 120 SEQERRKGIP WTEEEHRLF LGLEKFGKGD WRSISRNVFI SRTPTQVASH AQKYFIRLNS 180 MNRDRRRSSI HDITSINNGD VSTHQAPITG QQVNTNPSNA AALGPSVKHR TQPNMHGMGM 240 YGAPMGHPVA APSSHMASAV GTPVMMPHGH PPYVLPVAYP IPPPPPPMHQ 290
3-29	Sequences	
3-29-1	Sequence Number [ID]	29
3-29-2	Molecule Type	AA
3-29-3	Length	290
3-29-4	Features Location/ Qualifiers	source 1..290 mol_type=protein organism=Capsicum annuum
3-29-5	NonEnglishQualifier Value Residues	MSPSDWSREE EKAFENAIAM HWVENCEQQW EKFGMLVPTK TIDELKQHYQ LLVDDVTAIE 60 AGQVIPNYK GEEASSSTK EGNLGSAGS VDRRSNCGYS NGFSGTNHDP IGHGKGKNSR 120 SEQERRKGIP WTEEEHRLF LGLEKFGKGD WRSISRNVFI SRTPTQVASH AQKYFIRLNS 180 MNRDRRRSSI HDITSINNGD VSTHQAPITG QQVNTNPSNA AALGPSVKHR TQPNMHGMGM 240 YGAPMGHPVA APSSHMASAV GTPVMMPHGH PPYVLPVAYP IPPPPPPMHQ 290
3-30	Sequences	

3-30-1	Sequence Number [ID]	30
3-30-2	Molecule Type	AA
3-30-3	Length	296
3-30-4	Features Location/ Qualifiers	source 1..296 mol_type=protein organism=Jatropha curcas
3-30-5	NonEnglishQualifier Value Residues	MASAISWSKE ENKAFENAIA TYGVEDDDSK EQWEKVALMV PSKSIEELKQ HYQLLVEDVS 60 AIEAGNIPLP TYVGEEATSS SSKDSHGLSG ALAADKRSNC GYGSRFLGLD NSSSHGGKGG 120 SSSDQERRKG IPWTEEEHRL FLLGLDKFGK GDWRSISRNF VISRTPTQVA SHAQKYFIRL 180 NSMNRDRRRS SIHDITSVNN GNVSSQQAPI TGQQANTNPA STAAAMGHGV KHRVQPHMPS 240 LGMYGAPMGH PVAAPPGHMA SAVGTPVMLP PPGHHPPPPY VVPVAYPMAP PQTMHQ 296
3-31	Sequences	
3-31-1	Sequence Number [ID]	31
3-31-2	Molecule Type	AA
3-31-3	Length	291
3-31-4	Features Location/ Qualifiers	source 1..291 mol_type=protein organism=Nicotiana glauca
3-31-5	NonEnglishQualifier Value Residues	MSVSELWSKE EEKAFENAIA MHWVEDSEQQ WEKFAAVVPT KTIDEIKRHY QLLVDDVAIA 60 ESGHVPLPNY IGEESSSSSI KETNLGHS GS SDRRSNCGYV NGFSGPTHDP IGHSGKGNK 120 AEQERRKGIP WTEEEHRLFL LGLDKFGKGD WRSISRNFVI SRTPTQVASH AQKYFIRLNS 180 MNRDRRRSSI HDITSINGGD ISTHQAPITG QQPNNPNSNP AALGPNIKHR NQPNMHGFGM 240 YGAPMGHPVA APPTMASAV GTPVMLPHGH HPPYVLPVAY PLPPHPPPPM H 291
3-32	Sequences	
3-32-1	Sequence Number [ID]	32
3-32-2	Molecule Type	AA
3-32-3	Length	294
3-32-4	Features Location/ Qualifiers	source 1..294 mol_type=protein organism=Castanea mollissima
3-32-5	NonEnglishQualifier Value Residues	MSSGIIWSKE EEKTFENAIA MHWVDEDCKE QWEKIASMVP SKTMDCLKQH YQLLVEDVSA 60 IEAGHIPLPN YVGEEASSSN KDCHGFSGAM ASDKRSNCGF GGGFSLGHD STGHGKGG 120 RSEQERRKGI PWTEDEHRLF LLGLDKFGKG DWRSISRNFV ISRTPTQVAS HAQKYFIRLN 180 SMNRDRRRSS IHDITSVNNG DVASHQAPIT GQQTNNMPSS VATIGPPVKH RTQPHMPGLG 240 MYGTSVGHP I AAPVHMASA VGTPVMLPPG HHPHSHPPYV MPVAYPMAPP PMHQ 294
3-33	Sequences	
3-33-1	Sequence Number [ID]	33
3-33-2	Molecule Type	AA
3-33-3	Length	292
3-33-4	Features Location/ Qualifiers	source 1..292 mol_type=protein organism=Cephalotus follicularis
3-33-5	NonEnglishQualifier Value Residues	MSSGIAWTRE EEKAFKAIA MHWIEDDSKE QWEKIASMVP SKSFEELKHH YQILVEDVRA 60 IEDGHIPIPN YTGEATSSST KDFHGFSGAV TWDKRSNCGY GSGFSRLSHD SSSHREKQGL 120 RSDQERRKGI PWTEEEHRLF LLGLDKFGKG DWRSISRNFV ISRTPTQVAS HAQKYFIRLN 180 SMNRDRRRSS IHDITSINNG DVSSQDITIT GQEANANPTG APAMAQSGKH RPHPPHMRGL 240 GMYGAPVGHP VAGPGHMASA VGTPVMLPPG HHLHPSYLVV VAYPMAPPTL QQ 292
3-34	Sequences	
3-34-1	Sequence Number [ID]	34
3-34-2	Molecule Type	AA
3-34-3	Length	296
3-34-4	Features Location/ Qualifiers	source 1..296 mol_type=protein organism=Theobroma cacao
3-34-5	NonEnglishQualifier Value Residues	MAGTGTWSRE VEKAFENAIA MHWTEEGSEE QWEKIASMVP SKSLEELKQH YQLLVEDVSA 60 IEAGQVPLPS YTGEATSSV AKDFHGSSGA AAAAAPDKRS SSGYNGFSG LSHDSCGHGG 120 KGSSRSDQER RKGIPWTEEE HRLFLGLDK FGKGDWRSIS RNFVISRTPT QVASHAQKYF 180 IRLNSMNRDR RRSSIHDITS VNNGDTSHQA PITGQQANTN SPAAAAAMGP SVKHRAQPHM 240 PGLMGYAPV GRPVAAPGHM ASAVGTPVML PPGHHHHPPP YIVPVAYPMA PPPMHQ 296
3-35	Sequences	

3-35-1	Sequence Number [ID]	35
3-35-2	Molecule Type	AA
3-35-3	Length	298
3-35-4	Features Location/ Qualifiers	source 1..298 mol_type=protein organism=Ipomoea nil
3-35-5	NonEnglishQualifier Value Residues	MPAVVWSREE EKAFENAIAM HWVEDSKEQW NKIASMVPSK SVDELKQHYQ LLVEDVCAIE 60 AGHVPVPTYI GEEPPTSTHG NSAITSSDKR SNAAAAAYGT GFSGAATHDS AAPGGGKGGS 120 RSDQERRKGI PWTEEEHRLF LLGLDKFGKG DWRSISRNFV ISRTPTQVAS HAQKYFIRLN 180 SMNRDRRRSS IHDITSVNGG DVAAAHQPPI TGQQVNPTPS NVAAALGPAA KHRNQVNPBG 240 GLGMYGAPLG HPVPAPPCHI ASAVGTPVML PPGVGHHPY MLPLAYPIAP PQPPNMHQ 298
3-36	Sequences	
3-36-1	Sequence Number [ID]	36
3-36-2	Molecule Type	AA
3-36-3	Length	298
3-36-4	Features Location/ Qualifiers	source 1..298 mol_type=protein organism=Ipomoea triloba
3-36-5	NonEnglishQualifier Value Residues	MPQAVVWSRE EEKAFENAI MHWVEDSKEQ WNKIASMVPS KSVDELKQHY QLLVEDVCAI 60 EAGHVPVPTY IGEEPPTSTH ANSAITSSDR RSNASAAAYGT GFSGAATHDS AAPAAGKGGS 120 RSDQERRKGI PWTEEEHRLF LLGLDKFGKG DWRSISRNFV ISRTPTQVAS HAQKYFIRLN 180 SMNRDRRRSS IHDITSVNGG DVAAAHQPPI TGQQVNPTPS NVAAALGPAA KHRNQANPHG 240 GLGMYGAPLG HPVPAPPCHI ASAVGTPVML PPGVGHHPY MLPLAYPIAP PQPPNMHQ 298
3-37	Sequences	
3-37-1	Sequence Number [ID]	37
3-37-2	Molecule Type	AA
3-37-3	Length	292
3-37-4	Features Location/ Qualifiers	source 1..292 mol_type=protein organism=Corchorus olitorius
3-37-5	NonEnglishQualifier Value Residues	MAAAGIVTWS REEEKAFENA IAMHWIEEES NEQWEKIASM VPSKSVEELK QHYQLLEDV 60 NAIEAGQVPL PSYTGEPTS SAKDFHGFSG TSSSDKRSNS GYGNGFSGLN HDSNGHGGKG 120 SSRSDQERRK GIPWTEEEH LFLGLDKFG KGDWRSISRNFV ISRTPTQV ASHAQKYFIR 180 LNSMNRDRRR SSIHDITSVN NGETSHQAPI TGQQSNTNPS GAAAMGPSVK HRAQPHMPGL 240 GMYGAPVGRP VAAPGHMASA VGTPVMLPPG HPPPPYIVP VAYPMAPPPM HQ 292
3-38	Sequences	
3-38-1	Sequence Number [ID]	38
3-38-2	Molecule Type	AA
3-38-3	Length	297
3-38-4	Features Location/ Qualifiers	source 1..297 mol_type=protein organism=Solanum commersonii
3-38-5	NonEnglishQualifier Value Residues	MSPSDSSSVV WSREEEKAFE NAIALHWVED SEQQWEQFAS MVPTKTIDEL KEHYQLLEDV 60 VAAIEAGQVP IPNYKGEEAS SSSTKEVNLG YPGSVDGRRS NCGYSNGFSG TTHDPIGHGG 120 KGNSRSEQRK RKGIPWTEEE HRLFLGLDK FGKGDWRSIS RNFVISRTPT QVASHAQKYF 180 IRLNSMNRDR RRSSIHDITS INNGGDVSTN QAPITGQQVN PNPSNPAALG PSVKHRTQPN 240 MHSMGMYGAP MGHPVSAPPS HLASAVGTPV MLPHGHHPY VLPLAYPIPP PPPPMHQ 297
3-39	Sequences	
3-39-1	Sequence Number [ID]	39
3-39-2	Molecule Type	AA
3-39-3	Length	298
3-39-4	Features Location/ Qualifiers	source 1..298 mol_type=protein organism=Rhododendron williamsianum
3-39-5	NonEnglishQualifier Value Residues	MSPASIEWSR EEEKAFENAI AIHWTDSDKE QNWSKIASMV PNKTVEELKQ HYELLLEDVI 60 AIEAGKVPVP KYVGEEASSS STKDFHGS CGVTSTEKRSN CGLSGYSGL GQDPNGHGGK 120 GGGSRAEQER RKGIPWTEEE HRLFLGLDK FGKGDWRSIS RNFVISRTPT QVASHAQKYF 180 IRLNSMNRDR RRSSIHDITS INGGDPSTQP LPITGHQTNA AANPTNMAAI GPPPIKHRIP 240 PPNMAGLGMY GAPMGHPVAA GPSHMASAVG TPVMLTPGHH HPYVPLAYP MAPPTMHQ 298
3-40	Sequences	

3-40-1	Sequence Number [ID]	40
3-40-2	Molecule Type	AA
3-40-3	Length	298
3-40-4	Features Location/ Qualifiers	source 1..298 mol_type=protein organism=Rhododendron griersonianum
3-40-5	NonEnglishQualifier Value Residues	MSPASIEWSR EEEKAFENAI AIHWTDSDKE QNWSKIASMV PNKTVEELKQ HYEELLLEDVI 60 AIEAGKVPVP KYVGEEASSS STKDFHGS CGVTSTEKRSN CGLGSGYSGL GQDPNGHGGK 120 GGGSRAEQER RKGIPWTEEE HRLFLLGLDK FGKGDWRSIS RNFVISRTPT QVASHAQKYF 180 IRLNSMNRDR RRSSIHDTITS INGGDPSTQP LPITGHQTNA AANPTNMAAI GPPPIKHRIP 240 PPNMAGLGMV GAPMGHPVAA GPSHMASAVG TPVMLTPGHH HPYVVPLPYP MAPPTMHQ 298
3-41	Sequences	
3-41-1	Sequence Number [ID]	41
3-41-2	Molecule Type	AA
3-41-3	Length	297
3-41-4	Features Location/ Qualifiers	source 1..297 mol_type=protein organism=Solanum tuberosum
3-41-5	NonEnglishQualifier Value Residues	MSPSDSSSVV WSREEEKAFE NAIALHWVED SEQQWEQFAS MVPTKTIDEL KEHYQLLLED 60 VAAIEAGQVP IPNYKGEEAS SSSTKEVNLG YPGSVDGRRS NCGYSNGFSG TTHDPIGHGG 120 KGNRSRSEQER RKGIPWTEEE HRLFLLGLDK FGKGDWRSIS RNFVISRTPT QVASHAQKYF 180 IRLNSMNRDR RRSSIHDTITS INNGGDVSTH QAPITGQQVN PNPSNPAALG PSVKHRTQPN 240 MHSMGMYGAP MGHPVAAPPS HLASAVGTPV MLPHGHHPY VLPLAYPIPP PPPPMHQ 297
3-42	Sequences	
3-42-1	Sequence Number [ID]	42
3-42-2	Molecule Type	AA
3-42-3	Length	347
3-42-4	Features Location/ Qualifiers	source 1..347 mol_type=protein organism=Solanum tuberosum
3-42-5	NonEnglishQualifier Value Residues	MSPSDSSSVV WSREEEKAFE NAIALHWVED SEQQWEQFAS MVPTKTIDEL KEHYQLLLED 60 VAAIEAGQVP IPNYKGEEAS SSSTKEVNLG YPGSVDGRRS NCGYSNGFSG TTHDPIGHGG 120 KGNRSRSEQER RKGIPWTEEE HRLFLLGLDK FGKGDWRSIS RNFVISRTPT QVASHAQKYF 180 IRLNSMNRDR RRSSIHDTITS INNGGDVSTH QAPITGQQVN PNPSNPAALG PSVKHRTQPN 240 MHSMGMYGAP MGHPVAAPPS HLASAVGTPV MLPHGHHPY VLPLAYPIPP PPPPMHLED 300 SLAGKNRGSV LNAKSTNREQ WHMIVGKRTN EGSVEKKKRG EPSLGNL 347
3-43	Sequences	
3-43-1	Sequence Number [ID]	43
3-43-2	Molecule Type	AA
3-43-3	Length	293
3-43-4	Features Location/ Qualifiers	source 1..293 mol_type=protein organism=Nicotiana attenuata
3-43-5	NonEnglishQualifier Value Residues	MSTSTSSLMW SREEEKAFEN AIAIHWIEDS KLQWQKIASM VPSKTIQELK QHYQLLVDDV 60 TAIEAGNVPI PLYTGEETSS STKDSSHGYS GISNGRSNCS YVNGFSGSVH DQIGQGKGKS 120 SKSEQERRKG IPWTEEEHRL FLLGLDKFGK GDWRSISRNY VISRTPTQVA SHAQKYFIRL 180 NSMNRDRRRS SIHDVTSVNN GEVATNQTPV TGQQTNPSPS NAVGPNLGQR NQPNMHGLSM 240 YGAPMGHPVA APPGHVSAV GTPIMLPPGH HPPYVLPVGY HMPHPPPPM HHQ 293
3-44	Sequences	
3-44-1	Sequence Number [ID]	44
3-44-2	Molecule Type	AA
3-44-3	Length	293
3-44-4	Features Location/ Qualifiers	source 1..293 mol_type=protein organism=Nicotiana tomentosiformis
3-44-5	NonEnglishQualifier Value Residues	MSASTSSLVW SREEEKAFEN AIAIHWIEDS KLQWQKIASM VPSKTIIEELK HHYQLLVDDV 60 TAIEAGNVPI PIYTGEETSS STKDSNHGYS GLSTGRSNCS YVNGFSASVH DQIGQGKGKS 120 SKSEQERRKG IPWTEEEHRL FLLGLDKFGK GDWRSISRNY VISRTPTQVA SHAQKYFIRL 180 NSMNRDRRRS SIHDITSVNN GEVATNQTPV TGQQTNPSPS NAVGPSVGQR NQPNMHGLGM 240 YGAPMGHPVA APPSSHVSAV GTPVMLPPGH HPHYVLPVGY HMPHPPPPM HHQ 293

3-45	Sequences	
3-45-1	Sequence Number [ID]	45
3-45-2	Molecule Type	AA
3-45-3	Length	297
3-45-4	Features Location/ Qualifiers	source 1..297 mol_type=protein organism=Solanum tuberosum
3-45-5	NonEnglishQualifier Value Residues	MSPSDSSSVV WSREEEKAFE NAIALHWVED SEQQWEQFAS MVPTKTIDEL KEHYQLLLED 60 VAAIEAGQVP IPNYKGEEAS SSSTKEVNLG YPGSVDGRRS NCGYSNGFSG TTHDPIGHGG 120 KGNRSRSEQR RKGIPWTEEE HRLFLLGLDK FGKGDWRSIS RNFVISRTPT QVASHAQKYF 180 IRLNSMNRDR RRSSIHDTIS INNGGDVSTH QAPITGQQVN PNPSNPAALG PSVKHRTQPN 240 MHSMGMYAP MGHPVAAPPN HLASAVGTPV MLPHGHHPPY VLPLAYPIPP PPPPMHQ 297
3-46	Sequences	
3-46-1	Sequence Number [ID]	46
3-46-2	Molecule Type	AA
3-46-3	Length	296
3-46-4	Features Location/ Qualifiers	source 1..296 mol_type=protein organism=Theobroma cacao
3-46-5	NonEnglishQualifier Value Residues	MAGTGTWSRE VEKAFENAIA MHWTEEGSEE QWEKIASMVP SKSLEELKQH YQLLVEDVSA 60 IEAGQVPLPS YTGEATSSA AKDFHGSAGA AAAAAPDKRS SSGYGNGFSG LSHDSCGHGG 120 KGSSRSQER RKGIPWTEEE HRLFLLGLDK FGKGDWRSIS RNFVISRTPT QVASHAQKYF 180 IRLNSMNRDR RRSSIHDTIS VNNGDTSHQA PITGQQANTN SPAAAAAMGP SVKHRAQPHM 240 PGLGMYGAPV GRPVAAPGHM ASAVGTPVML PPGHHHHPPP YIVPVAYPMA PPPMHQ 296
3-47	Sequences	
3-47-1	Sequence Number [ID]	47
3-47-2	Molecule Type	AA
3-47-3	Length	297
3-47-4	Features Location/ Qualifiers	source 1..297 mol_type=protein organism=Herrania umbratica
3-47-5	NonEnglishQualifier Value Residues	MAGTGTWSRE EEKAFENAIA MHWTEEGSEE QWEKIASMVP SKSLEELKQH YQLLVEDVGA 60 IEAGQVPLPS YTGEATSSA AKDFHGFSGA AAAAAPDKRS SSGYGNGFSG LSHDSCGHGG 120 KGSSRSQER RKGIPWTEEE HRLFLLGLDK FGKGDWRSIS RNFVISRTPT QVASHAQKYF 180 IRLNSMNRDR RRSSIHDTIS VNNGDTSHQA PITGQQANAN SPAAAAAMGP SVKHRAQPHM 240 PGLGMYGAPI GRPVAAPGHM ASAVGTPVML PPGHHHHPPP PYMPVAYPM APPPMHQ 297
3-48	Sequences	
3-48-1	Sequence Number [ID]	48
3-48-2	Molecule Type	AA
3-48-3	Length	295
3-48-4	Features Location/ Qualifiers	source 1..295 mol_type=protein organism=Nicotiana attenuata
3-48-5	NonEnglishQualifier Value Residues	MSVSDLP SKL WSKEEEKAFE NAIAMHWVED SEQQWEKFAA VVPTKTIDEI KRHYQLLVDD 60 VAAIDAGHVP LPNYIGEES SSSNKETNLG YSGSVDRRSN CGYTNGFSGP THDPIGHS GK 120 GNSKAEQERR KGIPWTEEEH RLFLGLDKF GKGDWRSISR NFVISRTPTQ VASHAQKYFI 180 RLNSMNRDRR RSSIHDTISI NGGDVSTHQA PITGQQPNPN PSNPAALGPN IKHRNQPNMH 240 GYSMYGAPMG HPVAAPPSHM ASAVGTPVML PHGHHPPYVL PVAYPIPPH PPPMH 295
3-49	Sequences	
3-49-1	Sequence Number [ID]	49
3-49-2	Molecule Type	AA
3-49-3	Length	296
3-49-4	Features Location/ Qualifiers	source 1..296 mol_type=protein organism=Nicotiana tabacum
3-49-5	NonEnglishQualifier Value Residues	MSVSDLPSEV WSKEEEKAFE NAIAMHWVED SEQQWEKFAA VVPTKTIDEI KRHYQLLVDD 60 VAAIDAGHVP LPNYIGEES SSSNKETNLG YSGSVDRRSN CGYVNGFSGP THDPIGHS GK 120 GNSKAEQERR KGIPWTEEEH RLFLGLDKF GKGDWRSISR NFVISRTPTQ VASHAQKYFI 180 RLNSMNRDRR RSSIHDTISI NGGDVSIHQ PITGQQPNPN PSNPAALGPN IKHRNQPNMH 240 GFSMYGAPMG HPVAAPPSHM ASAVGTPVML PHGHHPPYVL PVAYPIPPH PPPMHQ 296

3-50	Sequences	
3-50-1	Sequence Number [ID]	50
3-50-2	Molecule Type	AA
3-50-3	Length	295
3-50-4	Features Location/ Qualifiers	source 1..295 mol_type=protein organism=Populus deltoides
3-50-5	NonEnglishQualifier Value Residues	MASVVTWTRD EEKTFENAIA MHWIDEDSNE QWEKIASMVP SKSLEELKLH YKILVEDVCA 60 IEAGNVPIPN YEGEEAASST KDLHGLSGTM TTVKKLNCAY GSGFVGLGHE SSGHGGKGAS 120 RSEQERKKGI PWTEEEHRLF LLGLDKFGKG DWRSISRNFV ISRTPTQVAS HAQKYFIRLN 180 SMNRDRRRSS IHDITSLNNG DVSSHQAPIT GQQVNTSPAG PAPAMGPPVK HRTEAHMPGL 240 AMYGPPPLGHP VAPPPGHMAS AVGTPVMLPP PGHHPHPYV VPVAYPTAPP KTMHQ 295
3-51	Sequences	
3-51-1	Sequence Number [ID]	51
3-51-2	Molecule Type	AA
3-51-3	Length	295
3-51-4	Features Location/ Qualifiers	source 1..295 mol_type=protein organism=Gossypium hirsutum
3-51-5	NonEnglishQualifier Value Residues	MAGSVSWSRE EEKAFENAIA MHWIDKEECE EQWEKIASTV PTKSLEELKL HYELLVEDVT 60 AIEAGHVPLP CYKGEEPSSS AKDYFHGPSM APNSDRRSNS GYGNGFSGLT LDSTGHGGKQ 120 SSRSDQERRK GIPWTEEEHR LFLLGLDKFG KGDWRSISRN FVISRTPTQV ASHAQKYFIR 180 LNSMNRDRRR SSIHDITSVN NGDTS HQAPI TGQQANTNSP GAAVMGQSVK HRAQPHLPGL 240 GMYGAPVGRP IAAAPGHIGS AVGTPVMLPP AHHPHSPPY IVPVAYPMAP PPMHQ 295
3-52	Sequences	
3-52-1	Sequence Number [ID]	52
3-52-2	Molecule Type	AA
3-52-3	Length	295
3-52-4	Features Location/ Qualifiers	source 1..295 mol_type=protein organism=Gossypium mustelinum
3-52-5	NonEnglishQualifier Value Residues	MAGSVSWSRE EEKAFENAIA MHWIEKEECK EQWEKIASTV PTKSLEELKL HYELLVEDVT 60 AIEAGHVPLP CYKGEEPSSS AKDHFHGPSM APNSDRRSNS GYGNGFSGLT LDSIGHGGKQ 120 SSRSDQERRK GIPWTEEEHR LFLLGLDKFG KGDWRSISRN FVISRTPTQV ASHAQKYFIR 180 LNSMNRDRRR SSIHDITSVN NGDTS HQAPI TGQQANTNSP GAAVMGQSVK HRAQPHLPGL 240 GMYGAPVGRP IAAAPGHIGS AVGTPVMLPP AHHPHSPPY IVPVAYPMAP PPMHQ 295
3-53	Sequences	
3-53-1	Sequence Number [ID]	53
3-53-2	Molecule Type	AA
3-53-3	Length	290
3-53-4	Features Location/ Qualifiers	source 1..290 mol_type=protein organism=Hevea brasiliensis
3-53-5	NonEnglishQualifier Value Residues	METSITWSRE EEKAFENAIA THWTEEDEES KELWEKIASM VPGKSIQELK QHYQLLVEDV 60 CAIEAGNIPL PNYGGEEATS STKDSHGNSG AVARDKRLNS GYGSGVLGLG HDSSGHGGKG 120 RSRSDQERRK GIPWTEEEHR LFLLGLNKFG KGDWRSISRN FVISRTPTQV ASHAQKYFIR 180 LNSMNRDRRR SSIHDITSVN NGDVSPHQAP ITGQQADTGA AAAIGTVVKL RVQPHMTGLG 240 MYGAPVGHPV APPGHVASAV GTPVMLPPGH HPSYVVPVPY PMAPPQTMHQ 290
3-54	Sequences	
3-54-1	Sequence Number [ID]	54
3-54-2	Molecule Type	AA
3-54-3	Length	307
3-54-4	Features Location/ Qualifiers	source 1..307 mol_type=protein organism=Arachis ipaensis
3-54-5	NonEnglishQualifier Value Residues	MSSSSGTIWS YEEKAFENA IAMHWIEEDS KEHWEKIAEE VPTKTIEELK QHYQLLVEDV 60 TAIEAGQIPF PNYTSEETTS SSKDLHGSSK SSNSDKRSNC NYGSGFSGLG HDSSTSHNGK 120 GSLSRSEQE RRGIPWTEE EHRLFLGLD KFGKGDWRSI SRNFVISRTP TQVASHAQKY 180 FIRLNSMNRD RRRSSIHDIT SVNNGDVANN QAPITGQHGG TTISSSTIAM GQSMKHRAQP 240

		HHHHHHIPAG LGMYGAPVGH PVAAPHGHMA SAVGTPVMLP PGPHPHPHHP PYVVPLAYPM 300 APPTMHQ 307
3-55	Sequences	
3-55-1	Sequence Number [ID]	55
3-55-2	Molecule Type	AA
3-55-3	Length	310
3-55-4	Features Location/ Qualifiers	source 1..310 mol_type=protein organism=Hevea brasiliensis
3-55-5	NonEnglishQualifier Value Residues	MHHPFNIFSL LSFIPDLSN METSITWSRE EEKAFENAIA THWTEEDEES KELWEKIASM 60 VPGKSIQELK QHYQLLVEDV CAIEAGNIPL PNYGGEEATS STKDSHGNSG AVARDKRLNS 120 GYGSGVLGLG HDSSGHGKG RSRSDQERRK GIPWTEEEHR LFLLGLNKF KGDWRSISR 180 FVISRTPTQV ASHAQKYFIR LNSMNRDRR SSIHDITSVN NGDVSPHQAP ITGQQADTGA 240 AAAIGTVVKL RVQPHMTGLG MYGAPVGHV APPGHVASAV GTPVMLPPGH HPSYVVPV 300 PMAPPQTMHQ 310
3-56	Sequences	
3-56-1	Sequence Number [ID]	56
3-56-2	Molecule Type	AA
3-56-3	Length	296
3-56-4	Features Location/ Qualifiers	source 1..296 mol_type=protein organism=Nicotiana tomentosiformis
3-56-5	NonEnglishQualifier Value Residues	MSVSDLPSEV WSKEEEKAFE NAIAMHWVED SEQQWEKFAA VVPTKTIDEI KRHYQLLVDD 60 VAAIDAGHVP LPNYIGEES SSSNKETNLG YSGSVDRRSN SGYVNGFSGP THDPIGHSGK 120 GNSKAEQERR KGIPWTEEEH RLFLGLDKF GKGWRSISR NFVISRTPTQ VASHAQKYFI 180 RLNSMNRDRR RSSIHDITSI NGGDVSIHQ PITGQQPNPN PSNPAALGPN IKHRNQPNMH 240 GFGMYGAPMG HPVAAPPSHM ASAVGTPVML PHGHHPPYVL PVAYPIPPH PPPMHQ 296
3-57	Sequences	
3-57-1	Sequence Number [ID]	57
3-57-2	Molecule Type	AA
3-57-3	Length	279
3-57-4	Features Location/ Qualifiers	source 1..279 mol_type=protein organism=Populus euphratica
3-57-5	NonEnglishQualifier Value Residues	MASVGTWTRD EEKTFENAIA MHWIDEDSNE QWEKIASMVP SKSLEELKLH YKILVEDVCA 60 IEAGNVPIPN YEGEEAASST KDLHGLSGTM TTVKKLNCGY GSGFVGLGHE SSGHSGKGAS 120 RSEQERKKGI PWTEEEHRLF LLGLDKFGKG DWRSISRNFV ISRTPTQVAS HAQKYFIRLN 180 SMNRDRRRSS IHDITSLNNG DVSSHQAPIT GQQVNTSPAG SAPAMGPPVK HRTQAHMPGL 240 AMYGPPLGHP VAHPPVHMAS AVGTPVMLPP PVHHPHPPT 279
3-58	Sequences	
3-58-1	Sequence Number [ID]	58
3-58-2	Molecule Type	AA
3-58-3	Length	300
3-58-4	Features Location/ Qualifiers	source 1..300 mol_type=protein organism=Prosopis alba
3-58-5	NonEnglishQualifier Value Residues	MSSSSGIVWS YEEKAFENA IAMHWVEEDS KEHWKIASA VPSKSMEELK QHYRLLLEDV 60 SAIEAGHIPL PNYTGGEAT SSSKELHASS SKASTSDKRS NCNFRSGFSG VGHESGHAG 120 KGGSSKSSEQ ERRKGIPWTE EEHRLFLLGL DKFGKGDWRS ISRNFMVMSRT PTQVASHAQK 180 YFIRLNSMNR DRRRSSIHD I TSVNNGDVAT NQAPITGQHA NTNPSSTMV AQSMKHRPHQ 240 PHIPAGIGMY GAPVGHVAA PPSHMASAVG TPVMMPPGPH PHPPYVVPMA YPMASPTMHR 300
3-59	Sequences	
3-59-1	Sequence Number [ID]	59
3-59-2	Molecule Type	AA
3-59-3	Length	306
3-59-4	Features Location/ Qualifiers	source 1..306 mol_type=protein organism=Abrus precatorius
3-59-5	NonEnglishQualifier Value Residues	MSSSGTIWSY EEEKAFENAI ALHWIEEASK DQWEKIASAV PSKSIEEVKQ HYQVLVEDVS 60 AIEAGHISFP NYATEETTSS NKDFHGSSKA STSDKRSNCN YSGFGSGLGH DSTTHTGKGG 120

		LSRSSEQERR KGIPWTEEEH RLFLGLDKF GKGDWRSISR NFVISRTPTQ VASHAQKYFI 180 RLNSMNRDRR RSSIHDITSV NNGDVVSNQA PITGQHTSTI SSSTMAVGQS MKHRVQPHIP 240 AGLGMYPGPV GHPVAAPPGH MASAVGTPVM LPPGPHPHPH AHPHAHPHP YVVPPLAYPMA 300 PPTMHQ 306
3-60	Sequences	
3-60-1	Sequence Number [ID]	60
3-60-2	Molecule Type	AA
3-60-3	Length	295
3-60-4	Features Location/ Qualifiers	source 1..295 mol_type=protein organism=Nicotiana glauca
3-60-5	NonEnglishQualifier Value Residues	MSTSTSTSSL MWSREEEKAF ENAIAIHWIE DSKLQWQKIA SMVPSKTIEE LKQHYQLLVD 60 DVTDIAGNV PIPIYTGEET SSSTKDSHSG YSGISTGRSN CGYVNGFSAS VHDQIGQGGK 120 ASSKSEQERR KGIPWTEEEH RLFLGLDKF GKGDWRSISR NYVISRTPTQ VASHAQKYFI 180 RLNSMNRDRR RSSIHDITSV NNGEVATNQT PITGQQTNPV PSNAVGPNVG QRNQPMMHGL 240 SMYGAPMGHP VAAPPGHHVS AVGTPVMLPP GHHPYVLPV GYHMPHPLP PMHHQ 295
3-61	Sequences	
3-61-1	Sequence Number [ID]	61
3-61-2	Molecule Type	AA
3-61-3	Length	293
3-61-4	Features Location/ Qualifiers	source 1..293 mol_type=protein organism=Populus trichocarpa
3-61-5	NonEnglishQualifier Value Residues	MASVGTWTRD EEKTFENAIA MHWIEDSNE QWEKIASMVP SKSLEELKLH YKILVEDVCA 60 IEAGNVPIPN YEGEEAASST KDLHGLSGTM TTVKKLNCGY GSGFVGLGHE SSGHGGKGAS 120 RSEQERKKGI PWTEEEHRLF LLGLDKFGKG DWRSISRNFV ISRTPTQVAS HAQKYFIRLN 180 SMNRDRRRSS IHDITSLNNG DVSSHQAPIT GQQANTSPAG PAPAMGPPVK HRTQAHPGL 240 AMYGPPLGHP VAPPPGHMAS AVGTPVMLPP PGHHPHPYV VPVAYPTAPP KTA 293
3-62	Sequences	
3-62-1	Sequence Number [ID]	62
3-62-2	Molecule Type	AA
3-62-3	Length	297
3-62-4	Features Location/ Qualifiers	source 1..297 mol_type=protein organism=Ziziphus jujuba
3-62-5	NonEnglishQualifier Value Residues	MSSSTVWSKE EEKAFENAIA MHWIEDEKES KEQWEKIASM VPNKNLEELK QHYQMLVDDV 60 NAIEAGHIPV PNYVGDETFS LNKDSHGSSG SAASEKRLNC GHGSGFSGLG QDPSGHGGKG 120 GSRSDQERRK GIPWTEEEH RLFLGLEKFG KGDWRSISRNFV FVISRTPTQV ASHAQKYFIR 180 LNSMNRDRRR SSIHDITSVN NADVTPHQAP ITGQQTANP SSAGAAIAPP VKHRGQPHMP 240 GIGMYGAPLG HPVAAPPGHM GSALGTPVML PPGHHPHPHP PYVVPVAYPM APPTMHQ 297
3-63	Sequences	
3-63-1	Sequence Number [ID]	63
3-63-2	Molecule Type	AA
3-63-3	Length	295
3-63-4	Features Location/ Qualifiers	source 1..295 mol_type=protein organism=Gossypium stocksii
3-63-5	NonEnglishQualifier Value Residues	MAGSVSWSRE EEKAFENAIA MHWIEKKECE EQWAKIASTV PTKSLEELKL HYELLVEDVT 60 AIEAGHVPLP CYKGEEPSST AKDHFHGPSM APNSDRRSNS GYGNGFSGST LDSTGHGGKQ 120 SSRSDQERRK GIPWTEEEH RLFLGLDKFG KGDWRSISRNFV FVISRTPTQV ASHAQKYFIR 180 LNSMNRDRRR SSIHDITSVS NGDTSHQAPI TGQQTNTNSP GAAVMGQSVK HSAQPHLPGL 240 GMYGAPVGRP IAAAPGHIGS AVGTPVMLPP AHHPHSPPPY IVPVAYPMAP PPMHQ 295
3-64	Sequences	
3-64-1	Sequence Number [ID]	64
3-64-2	Molecule Type	AA
3-64-3	Length	307
3-64-4	Features Location/ Qualifiers	source 1..307 mol_type=protein organism=Prunus dulcis
	NonEnglishQualifier Value	

3-64-5	Residues	MSSASVWNKE EDKAFENAIA MHWIDEESEE QWEKIAELVP SKSMEELKQH YQMLVDDVSA 60 IEAGHTPLPN YAAAEETSS SKDTAARASS GASASDKRLN CGHGGGFSAL AHDTSGHGGK 120 GGSRSQERK KGIPWTEEEH RLFLGLLEKF GKGDWRSISR NFVISRTPTQ VASHAQKYFI 180 RLNSMNRDRR RSSIHDITSV NNGDVSSHQP PITGQQNTY APSAAAAGAT TIGVGPQSVK 240 HRAHQPHMAG LGMYGAPMGH PVSAPPGHM GSAVGTPVML PPGHHPHTHP PYVVPVAYPM 300 AHPTMHQ 307
3-65	Sequences	
3-65-1	Sequence Number [ID]	65
3-65-2	Molecule Type	AA
3-65-3	Length	299
3-65-4	Features Location/ Qualifiers	source 1..299 mol_type=protein organism=Spatholobus suberectus
3-65-5	NonEnglishQualifier Value Residues	MSSSGTIWSY EEEKAFENAI AMHWIEEASK EQWEKIASAV PSKSMEEVKQ HYQVLVEDVI 60 AIEAGHISFP NYASEETTSS NKDFHGSSKA TTSDKRSNCN YGSGFSGLGH DSTTHSGKGK 120 LSRSSEQERR KGIPWTEEEH RLFLGLLDFK GKGDWRSISR NFVISRTPTQ VASHAQKYFI 180 RLNSMNRDRR RSSIHDITSV NNGDVASNQA PITGQHSSTI SSSTMGVGGQ SLKHRVQGHI 240 PAGLGMYGTP VGHPVAAPPG HMASAVGTPV MLPHGPHPHP HPPYVLPLAY PMAPPTMHQ 299
3-66	Sequences	
3-66-1	Sequence Number [ID]	66
3-66-2	Molecule Type	AA
3-66-3	Length	307
3-66-4	Features Location/ Qualifiers	source 1..307 mol_type=protein organism=Malus domestica
3-66-5	NonEnglishQualifier Value Residues	MSSSSVWSKE EDKAFENAIA MHWIDENSEE QWEKIAELVP SKSMEELKQH YQMLVDDVGA 60 IEAGRVSPPN YTADETANTV SSSKDSGHRA SSSGASADK RLNCGHGGGF LGSVHDSASH 120 GGKGGSRADQ ERKKGIPWTE EEHRLFLGLL DKFGKGDWRS ISRNVISRT PTQVASHAQK 180 YFIRLNSMNR DRRRSSIHD I TSVNHGDVSS HQQPPITGQQ TSTYAPTAAT TIGVGVGPQS 240 VKHRPQPHMA GLGMYGAPMG HPVSAPPGHM ASAVGTPVML PPGHHPHAHP PYVVPVAYPM 300 AHPTMHQ 307
3-67	Sequences	
3-67-1	Sequence Number [ID]	67
3-67-2	Molecule Type	AA
3-67-3	Length	309
3-67-4	Features Location/ Qualifiers	source 1..309 mol_type=protein organism=Phaseolus vulgaris
3-67-5	NonEnglishQualifier Value Residues	MSSSGTLWSY EEEKAFENAI AMHWIEEASK EQWEKIASAV PSKSMEEVKQ HYQVLVEDVS 60 AIEAGHVPFP NYAYEETTSS NKDFHGSSKA TSSDKRSNCN YGSGFSGLGH DSTTHSSGKG 120 GLRSSEQER RKIPWTEEE HRLFLGLLDFK FGKGDWRSISR RNFVISRTPT QVASHAQKYF 180 IRLNSMNRDR RSSIHDITS VNNGDVASSQ APITGQHSST VSSSTMGVGG SLKHRVQGHI 240 PPGLGMYGTP VGHPVAAPPG HMASAVGTPV MLPPGPHPHP HPHHPHAHP HPPYVLPLAY 300 PMAPPPMHQ 309
3-68	Sequences	
3-68-1	Sequence Number [ID]	68
3-68-2	Molecule Type	AA
3-68-3	Length	295
3-68-4	Features Location/ Qualifiers	source 1..295 mol_type=protein organism=Populus alba
3-68-5	NonEnglishQualifier Value Residues	MASVVTWTRD EEKTFENAIA MHWIDEDSNE QWEKIASMVP SKSLEELKLH YKILVEDVCA 60 IEAGNVPIPN YEGEEAASST KDLHGLSGTM TTVKKLNCGF GSGFVGLGHE SSGHGGKGAS 120 RSEQERKKGI PWTEEEHRLF LLGLDKFGKG DWRSISRNFV ISRTPTQVAS HAQKYFIRLN 180 SMNRDRRRSS IHDITSLNNG DVSSHQAPIT GQQVNTSPAG PAPAMGPPVK HRTQAHMPGL 240 AMYGPPLGHP VAPPPGHMAS VVGTPVMLPP PGHHPHPYV VPVAYPTAPP KTMHQ 295
3-69	Sequences	
3-69-1	Sequence Number [ID]	69
3-69-2	Molecule Type	AA
3-69-3	Length	295

3-69-4	Features Location/ Qualifiers	source 1..295 mol_type=protein organism=Populus tomentosa
3-69-5	NonEnglishQualifier Value Residues	MASVVTWTRD EEKTFENAIA MHWIEDSNE QWEKIASMVP SKSLEELKLH YKILVEDVCA 60 IEAGNVPIPN YEGEEAASST KDLHGLSGTM TTVKKLNCGF GSGFVGLGHE SSGHGGKGAS 120 RSEQERKKGI PWTEEEHRLF LLGLDKFGKG DWRSISRNFV ISRTPTQVAS HAQKYFIRLN 180 SMNRDRRRSS IHDITSLNNG DVSSHQTPTIT GQQVNTSPAG PAPAMGPPVK HRTQAHMPGL 240 AMYGPPPLGHP VAPPPGHMAS VVGTPVMLPP PGHHPHPYV VPVAYPTAPP KTMHQ 295
3-70	Sequences	
3-70-1	Sequence Number [ID]	70
3-70-2	Molecule Type	AA
3-70-3	Length	292
3-70-4	Features Location/ Qualifiers	source 1..292 mol_type=protein organism=Durio zibethinus
3-70-5	NonEnglishQualifier Value Residues	MAGSVTWSRE EEKAFENALA MHWIEDSEEQ WEKMASMVPS KSLEELKQHY QKLVEDVSAI 60 EAGKVPLPCY TGEETPSPK DVHGLSGAAA ATAASDKRSN SGYGNFGSGL SHDSSGHGGK 120 GSSKSDQERR KGIPWTEEEH RLFLLGLDKF GKGDWRSISR NFVISRTPTQ VASHAQKYFI 180 RLNSMNRDRR RSSIHDITSV NNGDTSHQAP ITGQQANTNS SGAASMGQSV KQRAQPHLP 240 LGMYGAPVRN PVAAPGHLGS AVGTPLMLPP GHHPYIVP VAYPMAPPPM HQ 292
3-71	Sequences	
3-71-1	Sequence Number [ID]	71
3-71-2	Molecule Type	AA
3-71-3	Length	307
3-71-4	Features Location/ Qualifiers	source 1..307 mol_type=protein organism=Prunus armeniaca
3-71-5	NonEnglishQualifier Value Residues	MSSASVWNKE EDKAFENAIA MHWIDEESEE QWEKLAELVP SKSMEELKQH YQMLVDDVSA 60 IEAGHTPLPN YAAEEATSS SKDTAARASS GASASDKRLN CGHGGGFSAL AHDTSGHGGK 120 GGSRSDQERK KGIPWTEEEH RLFLLGLEKF GKGDWRSISR NFVISRTPTQ VASHAQKYFI 180 RLNSMNRDRR RSSIHDITSV NNGDVSSHQP PITGQNTY APSAAAAGAT TIGVGPQSVK 240 HRAHQPHMAG LGMYGAPMGH PVSAPPGHM GSAVGTPVML PPGHHPHTHP PYVVPVAYPM 300 AHPTMHQ 307
3-72	Sequences	
3-72-1	Sequence Number [ID]	72
3-72-2	Molecule Type	AA
3-72-3	Length	295
3-72-4	Features Location/ Qualifiers	source 1..295 mol_type=protein organism=Vigna radiata var. radiata
3-72-5	NonEnglishQualifier Value Residues	MSSSGSLWNY EEEKAFENAI AMHWIEEASN EQWEKIASAV PSKSMEEVKQ HYQILVEDVS 60 AIEAGHIPFP NYTSEETTSS NKDFHGSSKA TSSDKRSNCN YSTGFSGLGH DSTTHSTGKG 120 ALRSSEQER RKIPWTEEE HRLFLLGLDK FGKGDWRSIS RNVISRTPT QVASHAQKYF 180 IRLNSMNRDR RRSSIHDITS VNNGDVASNQ APITGQHSST VSSSTMVGQG SLKHRVQGHI 240 APGLGMYGTP VGHPVAAPPG HMASAVGTPV MLPPGPHPPY VLPLAYPMAP PTMHQ 295
3-73	Sequences	
3-73-1	Sequence Number [ID]	73
3-73-2	Molecule Type	AA
3-73-3	Length	301
3-73-4	Features Location/ Qualifiers	source 1..301 mol_type=protein organism=Vigna angularis
3-73-5	NonEnglishQualifier Value Residues	MSSSGSLWNY EEEKAFENAI AMHWIEEASN EQWEKIASAV PSKSMEEVKQ HYQVLVEDVS 60 AIEAGHIPFP NYTSEETTSS NKDFHGSSKA TSSDKRSNCN YSTGFSGLGH DSTTHSTGKG 120 ALRSSEQER RKIPWTEEE HRLFLLGLDK FGKGDWRSIS RNVISRTPT QVASHAQKYF 180 IRLNSMNRDR RRSSIHDITS VNNGDVASNQ APITGQHSST VSSSTMVGQG SLKHRVQGHI 240 APGLGMYGTP VGHPVAAPPG HMASAVGTPV MLPPGPHPHA HPPHYVLPL AYPMAPPPMH 300 Q 301
3-74	Sequences	
3-74-1	Sequence Number [ID]	74

3-74-2	Molecule Type	AA
3-74-3	Length	307
3-74-4	Features Location/ Qualifiers	source 1..307 mol_type=protein organism=Prunus avium
3-74-5	NonEnglishQualifier Value Residues	MSSASVWNKE EDKAFENAIA MHWIDEESEE QWEKIAELVP SKSMEELKQH YQMLVDDVSA 60 IEAGHTPLPN YAAAEETSS SKDTAARASS GASASDKRLN CGHGGGFSAL AHDTSGHGGK 120 GGSRSQERK KGIPWTEEEH RLFLGLLEKF GKGDWRSISR NFVISRTPTQ VASHAQKYFI 180 RLNSMNRDRR RSSIHDITSV NNGDVSSHQP PITGQQTNTY APSAAAAGAT TIGVGPQSVK 240 HRAHQPHMAG LGMYGAPMGH PVSAPPGHM GSAMGTPVML PPGHHPHSH PYVVPVAYPM 300 AHPTMHQ 307
3-75	Sequences	
3-75-1	Sequence Number [ID]	75
3-75-2	Molecule Type	AA
3-75-3	Length	307
3-75-4	Features Location/ Qualifiers	source 1..307 mol_type=protein organism=Prunus persica
3-75-5	NonEnglishQualifier Value Residues	MSSASVWNKE EDKAFENAIA MHWIDEESEE QWEKIAELVP SKSMEELKQH YQMLVDDVSA 60 IEAGHTPLPN YAAAEETSS SKDTAARASS GASASDKRLN CGHGGGFSAL AHDTSGHGGK 120 GGSRSQERK KGIPWTEDEH RLFLGLLEKF GKGDWRSISR NFVISRTPTQ VASHAQKYFI 180 RLNSMNRDRR RSSIHDITSV NNGDVSSHQP PITGQQTNTY APSAAAAGAT TIGVGPQSVK 240 HRAHQPHMAG LGMYGAPMGH PVSAPPGHM GSAVGTPVML PPGHHPHPTH PYVVPVAYPM 300 AHPTMHQ 307
3-76	Sequences	
3-76-1	Sequence Number [ID]	76
3-76-2	Molecule Type	AA
3-76-3	Length	307
3-76-4	Features Location/ Qualifiers	source 1..307 mol_type=protein organism=Prunus mume
3-76-5	NonEnglishQualifier Value Residues	MSSASVWNKE EDKAFENAIA MHWIDEESEE QWEKLAELVP SKSMEELKQH YQMLVDDVSA 60 IEAGHTPLPN YAAAEETSS SKDTAARASS GASAPDKRLN CGHGGGFSAL AHDTSGHGGK 120 GGSRSQERK KGIPWTEEEH RLFLGLLEKF GKGDWRSISR NFVISRTPTQ VASHAQKYFI 180 RLNSMNRDRR RSSIHDITSV NNGDVSSHQP PITGQQTNTY APSAAAAGAT TIGVGPQSAK 240 HRAHQPHMAG LGMYGAPMGH PVSAPPGHM GSAVGTPVML PPGHHPHPTH PYVVPVAYPM 300 AHPTMHQ 307
3-77	Sequences	
3-77-1	Sequence Number [ID]	77
3-77-2	Molecule Type	AA
3-77-3	Length	306
3-77-4	Features Location/ Qualifiers	source 1..306 mol_type=protein organism=Cajanus cajan
3-77-5	NonEnglishQualifier Value Residues	MSSSGTLWSY EEEKAFENAI AMHWIEEASK EQWEKIASAV PSKSMEEVKQ HYQILVEDVS 60 AIEAGHISFP NYASEETTSS NKDFHGSSKA TNSDKRSNCS FSGGFSGLGH DSTTHSGKGG 120 LSRSSEQERR KGIPWTEEEH RLFLGLLDKF GKGDWRSISR NFVISRTPTQ VASHAQKYFI 180 RLNSMNRDRR RSSIHDITSV NNGDVANNQA PITGQHSSTI SPSTMGVGQS MKHRVQGHIP 240 AGLGMYGTPV GHPVAAPPGH MASAVGTPVM LPPGPHPHPH AHPHPPHPP YVLPLAYPMA 300 PPTMHQ 306
3-78	Sequences	
3-78-1	Sequence Number [ID]	78
3-78-2	Molecule Type	AA
3-78-3	Length	293
3-78-4	Features Location/ Qualifiers	source 1..293 mol_type=protein organism=Hibiscus syriacus
3-78-5	NonEnglishQualifier Value Residues	MAASVVSRE EDKAFENAIA MHWTEEEPE DRWEKISSMV PSRSVEELKQ HYEILVEDVT 60 AIEAGHVPLP CYRGEELSSS SKDNFHGVPT AANSRRSNP SYRNGFSGLT VDSSGHRGKG 120 SSRSQERRK GIPWTEEEH RLFLGLLDKFG KGDWRSISRN FVISRTPTQV ASHAQKYFIR 180

		LNSMNRDRRR SSIHDITSVN NGDTS HQAPI TGQQGNANTP GGVALGQSVK HRAQPHLPGL 240 GMYGAPVGRP IAAPGHMGSS VGTPVMLPPG HHPQSPSYIV PVAYPMAPPT MHQ 293
3-79	Sequences	
3-79-1	Sequence Number [ID]	79
3-79-2	Molecule Type	AA
3-79-3	Length	299
3-79-4	Features Location/ Qualifiers	source 1..299 mol_type=protein organism=Solanum pennellii
3-79-5	NonEnglishQualifier Value Residues	MSPSDSSSSS VVWSREEEKA FENAI AQHWV EDSEQQWEQF ASMVPTKTID ELKEHYQLLL 60 EDVAAIEAGQ VPIPNYKGEE ASSSSTKEVN LGYTGSVDGR RSNCGYANGF SGTNHDPIGH 120 GGKGNRSSEQ ERRKGIPWTE EEHRLFLGL DKFGKGDWRS ISRNFVISRT PTQVASHAQK 180 YFIRLNSMNR DRRRSSIHDI TSINNGGDVS THQAPITGQQ VNPNPSPNAA LGPSVKHRTQ 240 PNMHSMGYG APMGHPVAAP PSHLASAVGT PVMLPHGHPH PYVLPLAYPI PPPPPPMHQ 299
3-80	Sequences	
3-80-1	Sequence Number [ID]	80
3-80-2	Molecule Type	AA
3-80-3	Length	302
3-80-4	Features Location/ Qualifiers	source 1..302 mol_type=protein organism=Glycine max
3-80-5	NonEnglishQualifier Value Residues	MSSSGTIWSY EEEKAFENAI AMHWIEESSK EQWEKIASAV PSKSMEDVKQ HYQVLVEDVS 60 AIEAGHISFP NYAASEEITS SNKDFHGSSK ATSSDKRSNC NYGRGFSGLG HDSTTHSSGK 120 GGLSRSSSEQ RRGKIPWTEE EHRLFLGLLE KFGKGDWRSI SRNFVISRTP TQVASHAQKY 180 FIRLNSMNRD RRRSSIHDIT SVNNGDVANN QAPITGQHSS TIPSNTMGVG QSLKHRVQGH 240 IPPGLGMYGT PVGHPVAAPP GHMASAVGTP VMLPPGPHPH AHPHPYVLP LAYPMAPPTM 300 HQ 302
3-81	Sequences	
3-81-1	Sequence Number [ID]	81
3-81-2	Molecule Type	AA
3-81-3	Length	296
3-81-4	Features Location/ Qualifiers	source 1..296 mol_type=protein organism=Vitis riparia
3-81-5	NonEnglishQualifier Value Residues	MSSAVPVWSR EEEKAFENAI AMHWTEDCKE VWDKIASMVP GKSVDDELKQH YQFLVEDVNA 60 IEAGHIPLPN YAADEASSSS VKDHHALPSA TSDKRSNCGF GGGFSGLGHD SAVQGGKGGG 120 RSEQERRKGI PWTEEEHRLF LLGLDKFGKG DWRSISRNFV ISRTPTQVAS HAQKYFIRLN 180 SMNRDRRRSS IHDITSVNNG DVSTPQAPIT GQQGNGNPAS SAAAGVGPPL KHHRAQPSMP 240 AGALGMYGTP MGHVPVAAPP HMASAVGTPV MLPPGPHAPP YVVPVAYPMA PPPMHQ 296
3-82	Sequences	
3-82-1	Sequence Number [ID]	82
3-82-2	Molecule Type	AA
3-82-3	Length	311
3-82-4	Features Location/ Qualifiers	source 1..311 mol_type=protein organism=Vigna unguiculata
3-82-5	NonEnglishQualifier Value Residues	MSSSGTLWSY EEEKAFENAI AMHWIEEASK EQWEKIASAV PTKSMEEVKQ HYQVLVEDVS 60 AIEAGHIPFP NYASEETTSS NKDFHASSKA TNSDKRSNCN FGTGFSGLGH DSTTHSSGKG 120 ALRSSEQER RKGIPWTEEE HRLFLGLLDK FGKGDWRSIS RNFVISRTPT QVASHAQKYF 180 IRLNSMNRDR RRSSIHDITS VNNGDVASNQ APITGQHSST VSSSTMVGVG SLKHRVQGHI 240 PPGLGMYGTP VGHPVAAPP HMASAVGTPV MLPPGPHPHA HAHAPHHPH HPHPPYVLPL 300 AYPMAPPTMH Q 311
3-83	Sequences	
3-83-1	Sequence Number [ID]	83
3-83-2	Molecule Type	AA
3-83-3	Length	293
3-83-4	Features Location/ Qualifiers	source 1..293 mol_type=protein organism=Manihot esculenta
	NonEnglishQualifier Value	

3-83-5	Residues	MESEISWSRE EEKAFENAIA KHWTEEEDES EELWEMIASL VPTKSVQELK LHYQLLVEDV 60 NAIEAGNIPL PNYVGEETSS SAKDSHGISG AVAAEKRLNC GYGSGFLGLG HNSSGHGGKG 120 GSRSDQERKK GIPWTEEEHR LFLGLDKFG KGDWRSISRN FVISRTPTQV ASHAQKYFIR 180 LNSMNRDRRR SSIHDITSVN NGDVSSNQPP ITGQQANSNP AAAAAMGSKV KHSVQPHMPG 240 LGIYGAPVGH PVAAPPGHMG SAVGTPVMLP HGHHPYVVP VAYPVAPPQT MHQ 293
3-84	Sequences	
3-84-1	Sequence Number [ID]	84
3-84-2	Molecule Type	AA
3-84-3	Length	296
3-84-4	Features Location/ Qualifiers	source 1..296 mol_type=protein organism=Vitis vinifera
	NonEnglishQualifier Value	
3-84-5	Residues	MSSEVPVWSR EEEKAFENAI AMHWTECKE VWDKIASMVP GKSVDCLKQH YQFLVEDVNA 60 IEAGHIPLPN YAADEASSSS VKDHALPSA TSDKRSNCGF GGGFSGLGHD SAVQGGKGGG 120 RSEQERRKGI PWTEEEHRLF LLGLDKFGKG DWRSISRNFV ISRTPTQVAS HAQKYFIRLN 180 SMNRDRRRSS IHDITSVNNG DVSTPQAPIT GQQGNGNPAS AAAAGVGPPL KHHRAQPSMP 240 AGALGMYGTP MGHVPVAAPPG HMASAVGTPV MLPPGPHAPP YVVPVAYPMA PPPMHQ 296
3-85	Sequences	
3-85-1	Sequence Number [ID]	85
3-85-2	Molecule Type	AA
3-85-3	Length	307
3-85-4	Features Location/ Qualifiers	source 1..307 mol_type=protein organism=Malus domestica
	NonEnglishQualifier Value	
3-85-5	Residues	MSSSSVWSKE EDKAFENAIA MHWIDENSEE QWEKIAELVP SKSMEELKQH YQMLVDDVGA 60 IEAGRVSPPN YTADEAANTV SSKDSGHRA SSSGTSASDK RLNCGHGGGF LGSGHDSASH 120 GGKGGSRADQ ERKKGIPWTE EEHRLFLGLL DKFGKGDWRS ISRNFVISRT PTQVASHAQK 180 YFIRLNSMNR DRRRSSIHDI TSVNHGDVSS HQQPPITGQQ TSTYAPTAAT TIGVGVPQS 240 VKHGPQPHMA GLGMYGAPMG HPVSAPPGHM ASAVGTPVML PPGHHPHAHP PYVVPVAYPM 300 AHPTMHQ 307
3-86	Sequences	
3-86-1	Sequence Number [ID]	86
3-86-2	Molecule Type	AA
3-86-3	Length	311
3-86-4	Features Location/ Qualifiers	source 1..311 mol_type=protein organism=Arachis duranensis
	NonEnglishQualifier Value	
3-86-5	Residues	MSSSSGTIWS YEEKAFENA IAMHWIEEDS KEQWEKIGEE VPSKTIEELK QHYQLLVEDV 60 TAIEAGQIPF PNYTSEETTS SSKDLHGSSK SSNSDKRSNC NYGSGFSGLG HDSSTSHNGK 120 GSLRSSEQE RRGIPWTEE EHRLFLGLD KFGKGDWRSI SRNFVISRTP TVASHAQKY 180 FIRLNSMNRD RRRSSIHDI SVNNGDVANN QAPITGQHG TTISSSTIAM GQSMKHRAQP 240 HHHHHHIPAG LGMYGAPVGH PVAAPHGHMA SAVGTPVMLP PGPHPHPHPH PHHPPYVVP 300 AYPMAPPTMH Q 311
3-87	Sequences	
3-87-1	Sequence Number [ID]	87
3-87-2	Molecule Type	AA
3-87-3	Length	302
3-87-4	Features Location/ Qualifiers	source 1..302 mol_type=protein organism=Citrus unshiu
	NonEnglishQualifier Value	
3-87-5	Residues	MSSAATWSRE EDKAFENAIA MHCIEEDDSK EEQQQWQKI ASMVPSRSVE ELKQHYQILL 60 EDVNAIEAGN IPLPNYVGED RATSSTKDFH GPSTAADNRS NGVYSGGFSG LSHDPSGHGS 120 KGSSKSEQR RKGIPWTEEE HRLFLGLDK FGKGDWRSIS RNFVISRTP TVASHAQKYF 180 IRLNSMNRDR RRSSIHDI TS VNNGDVSAH QAPITGQQQA NTNPPGAATL PMEASVKHRS 240 QPHMPGLGMY GAPVGHPIAA LGHMASAVGT PVMLPPGHHH HPHPPYAVPV AYPMAPPTM 300 HH 302
3-88	Sequences	
3-88-1	Sequence Number [ID]	88
3-88-2	Molecule Type	AA
3-88-3	Length	296

3-88-4	Features Location/ Qualifiers	source 1..296 mol_type=protein organism=Vitis vinifera
3-88-5	NonEnglishQualifier Value Residues	MSSAVPVWSR EEEKAFENAI AMHWTDDCKE VWDKIASMVP GKSVDDELKQH YQFLVEDVNA 60 IEAGHIPLPN YAADEASSSS VKDHHAGPSA TSDKRSNCGF GGGFSGLGHD STVQGGKGGG 120 RSEQERRKGI PWTEEEHRLF LLGLDKFGKG DWRSISRNFV ISRTPTQVAS HAQKYFIRLN 180 SMNRDRRRSS IHDITSVNNG DVSTPQAPIT GQQGNGNPAS AAAAGVGPPL KHHRAQPSMP 240 AGALGMYGTP MGHPPVAAPP HMASAVGTPV MLPPGPHAPP YVVPVAYPMA PPPPMHQ 296
3-89	Sequences	
3-89-1	Sequence Number [ID]	89
3-89-2	Molecule Type	AA
3-89-3	Length	299
3-89-4	Features Location/ Qualifiers	source 1..299 mol_type=protein organism=Solanum chilense
3-89-5	NonEnglishQualifier Value Residues	MSPSDSSSSS VVWSREEEKA FENAIQHWV EDSEQQWEQF ASMVPTKTID ELKEHYQLLL 60 EDVAAIEAGQ VPIPNYKGEE ASSSSTKEVN LGYPGSVDGR RSNCGYANGF SGTNHDPIGH 120 GGKGNRSRSEQ ERRKGIPWTE EEHRLFLGLL DKFGKGDWRS ISRNFBVISRT PTQVASHAQK 180 YFIRLNSMNR DRRRSSIHDI TSINNGGDVS THQAPITGQQ VNPNPSPNPA LGPSVKHRTQ 240 PNMHSMGMYG APMGHPVAAP PSHLASAVGT PVMLPHGHHS PYVLPLAYPI PPPPPPMHQ 299
3-90	Sequences	
3-90-1	Sequence Number [ID]	90
3-90-2	Molecule Type	AA
3-90-3	Length	299
3-90-4	Features Location/ Qualifiers	source 1..299 mol_type=protein organism=Solanum lycopersicum
3-90-5	NonEnglishQualifier Value Residues	MSPSDSSSSS VVWSREEEKA FENAIQHWV EDSEQQWEQF ASMVPTKTID ELKEHYQLLL 60 EDVAAIEAGQ VPIPNYKGEE ASSSSTKEVN LGYPGSVDGR RSNCGYANGF SGTNHDPIGH 120 GGKGNRSRSEQ ERRKGIPWTE EEHRLFLGLL DKFGKGDWRS ISRNFBVISRT PTQVASHAQK 180 YFIRLNSMNR DRRRSSIHDI TSINNGGDVS THQAPITGQQ VNPNPSPNPA LGPSVKHRTQ 240 PNMHSMGMYG APMGHPVAAP PSHLASAVGT PVMLPHGHHS PYVLPLAYPI PPPPPPMHQ 299
3-91	Sequences	
3-91-1	Sequence Number [ID]	91
3-91-2	Molecule Type	AA
3-91-3	Length	314
3-91-4	Features Location/ Qualifiers	source 1..314 mol_type=protein organism=Mucuna pruriens
3-91-5	NonEnglishQualifier Value Residues	MSSSGTIWSY EEEKAFENAI AMHWIEEDSM EQWEKIASAV PSKSMEEVKQ HYELLVEDVR 60 AIEAGHISFP NYAYEETTSS NKDFHGSSKA TNSDKRSNCN YGSGFSGLGH DSTTHSGKG 120 GLRSSEQER RKGIPWTEEE HRLFLGLLDK FGKGDWRSIS RNFVISRTPT QVASHAQKYF 180 IRLNSMNRDR RRSSIHDI TSINNGDVASNQ APITGQHSST ISPSTMGVQG SLKHRVQGHI 240 PAGLGMYGTP VGHPVAAPP HMASAVGTPV MLPPGPHPHP HPHPHPHPHP HSHPHHPHYV 300 LPLAYPMAPP TMHQ 314
3-92	Sequences	
3-92-1	Sequence Number [ID]	92
3-92-2	Molecule Type	AA
3-92-3	Length	307
3-92-4	Features Location/ Qualifiers	source 1..307 mol_type=protein organism=Glycine max
3-92-5	NonEnglishQualifier Value Residues	MSSSGTIWSY DEEKAFENAI AMHWIEESSK EQWEKIASAV PSKSMEEVKQ HYQVLVEDVS 60 AIEAGHISFP NYASDETTSS NKDFHGFSSA TSSDKRSNCN YGSGFSGLGL DSTTHSGKG 120 GLRSSEQER RKGIPWTEEE HRLFLGLLDK FGKGDWRSIS RNFVISRTPT QVASHAQKYF 180 IRLNSMNRDR RRSSIHDI TSINNGDVASSQ APITGLHSST ISSNTMGVQG SLKHRVQGHI 240 PPGLGMYGTP VGHPVAAPP HMASAVGTPV MLPPGPHPHP HPHHHAHPHP PYVLPLAYPM 300 APPTMHQ 307
3-93	Sequences	
3-93-1	Sequence Number [ID]	93

3-93-2	Molecule Type	AA
3-93-3	Length	302
3-93-4	Features Location/ Qualifiers	source 1..302 mol_type=protein organism=Citrus sinensis
3-93-5	NonEnglishQualifier Value Residues	MSSAATWSRE EDKAFENAI MHCIEEDDSK EEQQQQWQKI ASMVPSRSVE ELKQHYQILL 60 EDVNAIEAGN IPLPNYVGED RATSSTKDFH GPSTAADNRS NGVYGSFGFSG LSHDPSGHGS 120 KGSSKSEQER RKGIPWTEEE HRLFLLGLDK FGKGDWRSIS RNFVISRTPT QVASHAQKYF 180 IRLNSMNRDR RRSSIHDTIS VNNGDVSAHH QAPITGQQQA NTNPPGAATL PMEASVKHRS 240 QPHLPGLGMY GAPVGHPHIAA LGHMASAVGT PVMLPPGHPH HPHPPYAVPV AYPMAPPPTM 300 HH 302
3-94	Sequences	
3-94-1	Sequence Number [ID]	94
3-94-2	Molecule Type	AA
3-94-3	Length	302
3-94-4	Features Location/ Qualifiers	source 1..302 mol_type=protein organism=Citrus clementina
3-94-5	NonEnglishQualifier Value Residues	MSSAATWSRE EDKAFENAI MHCIEEDDSK EEQQQQWQKI ASMVPSRSVE ELKQHYQILL 60 EDVNAIEAGN IPLPNYVGED RATSSTKDFH GPSTAADNRS NGVYGSFGFSG LSHDPSGHGS 120 KGSSKSEQER RKGIPWTEEE HRLFLLGLDK FGKGDWRSIS RNFVISRTPT QVASHAQKYF 180 IRLNSMNRDR RRSSIHDTIS VNNGDVSAHH QAPITGQQQA NTNPPGAATL PMEASVKHRS 240 QPHMPGLGMY GAPVGHPHIAA LGHMASAVGT PVMLPPGHPH HPHPPYAVPV AYPMAPPPTM 300 HQ 302
3-95	Sequences	
3-95-1	Sequence Number [ID]	95
3-95-2	Molecule Type	AA
3-95-3	Length	306
3-95-4	Features Location/ Qualifiers	source 1..306 mol_type=protein organism=Lupinus angustifolius
3-95-5	NonEnglishQualifier Value Residues	MSSSGTIWSY EEEKTFENAI AMHWTEESK EQWEKIASAV PSKSMEEVKQ HYKVLVEDVS 60 AIEAGHIPFP NYAADETKSS SKDFHGSSKA SSSDKRSNCN FGSGFSGGLGH DSASHSSGKG 120 GLSRSEQER RKGIPWTEEE HRLFLLGLDK FGKGDWRSIS RNFVISRTPT QVASHAQKYF 180 IRLNSMNRDR RRSSIHDTIS INNRDVASNQ VPITGQHSST MPSSTVAVGQ SVKHRAHQPH 240 HMHAGLGYG APVGHPVAAP HGHMASAVGT PVMLPPGPHH PHPHAHPHS YVVPLAYPMA 300 PPTMHQ 306
3-96	Sequences	
3-96-1	Sequence Number [ID]	96
3-96-2	Molecule Type	AA
3-96-3	Length	307
3-96-4	Features Location/ Qualifiers	source 1..307 mol_type=protein organism=Glycine max
3-96-5	NonEnglishQualifier Value Residues	MSSSGTIWSY DEEKAFENAI AMHWIEESSK EQWEKIASAV PSKSMEEVKQ HYQVLVEDVS 60 AIEAGHISFP NYASDETTSS NKDFHGSSKA TSSDKRSNCN YGSGFSGGLGL DSTTHSSGKG 120 GLSRSEQER RKGIPWTEEE HRLFLLGLDK FGKGDWRSIS RNFVISRTPT QVASHAQKYF 180 IRLNSMNRDR RRSSIHDTIS VNNGDVASSQ APITGLHSST ISSNTMGVGQ SLKHRVQGHI 240 PPGLGMYGTP VGHPVAAPPV HMASAVGTPV MLPPGPHPHP HPHHHAHPHP PYVPLAYPMA 300 APPTMHQ 307
3-97	Sequences	
3-97-1	Sequence Number [ID]	97
3-97-2	Molecule Type	AA
3-97-3	Length	307
3-97-4	Features Location/ Qualifiers	source 1..307 mol_type=protein organism=Glycine soja
3-97-5	NonEnglishQualifier Value Residues	MSSSGTIWSY DEEKAFENAI AMHWIEESSK EQWEKIASAV PSKSMEEVKQ HYQVLVEDVS 60 AIEAGHISFP NYASDETTSS NKDFHGSSKA TSSDKRSNCN YGSGFSGGLGL DSTTHSSGKG 120 GLSRSEQER RKGIPWTEEE HRLFLLGLEK FGKGDWRSIS RNFVISRTPT QVASHAQKYF 180

		IRLNSMNRDR RRSSIHDTITS VNNGDVASSQ APITGLHSST ISSNTMGVGQ SLKHRVQGGHI 240 PPGLGMYGTP VGHPVAAPPG HMASAVGTPV MLPPGPHPHP HPHHHAHPHP PYVLPLAYPM 300 APPTMHQ 307
3-98	Sequences	
3-98-1	Sequence Number [ID]	98
3-98-2	Molecule Type	AA
3-98-3	Length	299
3-98-4	Features Location/ Qualifiers	source 1..299 mol_type=protein organism=Rhamnella rubrinervis
3-98-5	NonEnglishQualifier Value Residues	MSSTAGSTVW SKDEEKAFEN AIAMHWVDED CEQWEKIAS MVPTKTLEEL KQHYQILVDD 60 VNAIEAGHIP VPNYVGEEAS SLNKDSQGPS GSATSEKRLN CGHGGGFSGV GQDPAAHGGK 120 GGSRSQDERR KGIPWTEEEH RLFLGLDKF GKGDWRSISR NFVISRTPTQ VASHAQKYFI 180 RLNSMNRDRR RSSIHDITSV NNADVSSHQA PITGQQATT VSSPGAVIAP PVKHRGQPHM 240 PGLGMYGAPL GHPVAAPPGH HMGSALGTPV MLPPGHHPHP HPPYVVPVAY PMAPPTMHQ 299
3-99	Sequences	
3-99-1	Sequence Number [ID]	99
3-99-2	Molecule Type	AA
3-99-3	Length	293
3-99-4	Features Location/ Qualifiers	source 1..293 mol_type=protein organism=Populus tomentosa
3-99-5	NonEnglishQualifier Value Residues	MASVVTWTRD EEKTFENAIA MHWIEDSNE QWEKIASMVP SKSLEELKLH YKILVEDVCA 60 IEAGNVPIPN YEGEEAASST KDLHGLSGTM TTVKKLNCGS GFVGLGHESS GHGGKGASRS 120 EQERKKGIPW TEEHRLFL LLDKFGKGDW RSISRNFVIS RTPTQVASHA QKYFIRLNSM 180 NRDRRRSSIH DITSLNNGDV SSHQAPITGQ QVNTSPAGPA PAMGPPVKHR TQAHMPGLAM 240 YGPPLGHPVA PPPGHMASVV GTPVMLPPP HHPHPYVVP VAYPTAPPKT MHQ 293
3-100	Sequences	
3-100-1	Sequence Number [ID]	100
3-100-2	Molecule Type	AA
3-100-3	Length	295
3-100-4	Features Location/ Qualifiers	source 1..295 mol_type=protein organism=Gossypium anomalum
3-100-5	NonEnglishQualifier Value Residues	MAGSVSWSRE EEKAFENAIA MHWIEKEECE EQWEKIASTV PTKSLEELKL HYELLVEDVT 60 AIEAGHVPLP CYKGEEPSSS AKDHFHGPSM APNSDRRSNS GYGNGFSGLT LDSTGHGGKQ 120 SSRSDQERRK GIPWTEEEH RLFLGLDKFG KGDWRSISR FVISRTPTQV ASHAQKYFIR 180 LNSMNRDRR RSSIHDITSVN NGDTS HQAPI TGQQANTNSP GAAVMGQSVK HRAQPHLPGL 240 GMYGAPVGRP IAAAPGHIGS AVGTPVMLPP AHHPHSPPY IVPVAYPMAP PPMHQ 295
3-101	Sequences	
3-101-1	Sequence Number [ID]	101
3-101-2	Molecule Type	AA
3-101-3	Length	295
3-101-4	Features Location/ Qualifiers	source 1..295 mol_type=protein organism=Gossypium raimondii
3-101-5	NonEnglishQualifier Value Residues	MAGSVSWSRE EEKAFENAIA MHWIEKEECE EQWEKIASTV PTKSLEELKL HYELLVEDVT 60 AIEAGHVPLP CYKGEEPSSS AKDYFHGPSM APNSDRRSNS GYGNGFSGLT LDSTGHGGKQ 120 SSRSDQERRK GIPWTEEEH RLFLGLDKFG KGDWRSISR FVISRTPTQV ASHAQKYFIR 180 LNSMNRDRR RSSIHDITSVN NGDTS HQAPI TGQQANTNSP GAAVMGQSVK HRAQPHLPGL 240 GMYGAPVGRP IAAAPGHIGS AVGTPVMLPP AHHPHSPPY IVPVAYPMAP PPMHQ 295
3-102	Sequences	
3-102-1	Sequence Number [ID]	102
3-102-2	Molecule Type	AA
3-102-3	Length	311
3-102-4	Features Location/ Qualifiers	source 1..311 mol_type=protein organism=Actinidia eriantha
3-102-5	NonEnglishQualifier Value Residues	MSDWSREEEK AFENAI AIHW AEEDSKEDDH WPKIASLVPT KTIEELKKHY QLLVEDVA AI 60 EAGHVPIPNY IGEEASSSSS KDRHGISGVT ASEKRSNCNY GSGFFPGFGH DYPAVPWRKE 120

		GPGLSKRDEK GFHGQKKNIG LLLGNAMRQL LVDNALGALN NLSKLFLLGL DKFGKGDWRS 180 ISRNFBVISRT PTQVASHAQK YFIRLNSINR DRRRSSIHDI TSVNGGDASA HQPPPIITGQQ 240 PNSTSAMAAI GPPQPHMAGL GMYGAPVGHP VAASSHHMAS AVGTPVMLPP GHHPYVVPV 300 AYPMAPPTMH Q 311
3-103	Sequences	
3-103-1	Sequence Number [ID]	103
3-103-2	Molecule Type	DNA
3-103-3	Length	909
3-103-4	Features Location/ Qualifiers	source 1..909 mol_type=other DNA organism=Actinidia eriantha
3-103-5	NonEnglishQualifier Value Residues	atgtcagctt cagtggattg gaggagagaa gaagagaaa cttttgagaa tgccattgca 60 attcactggg cagaggactc caaagatgag gaccattggg aaaagattgc ttcattgggt 120 cctactaaaa ccatggaaga attgaagaag cattaccaat tactagttaga ggatgttgct 180 gcaatcgaat cgggacacgt gccaatcccc agctatatg gggaggaagc atcatcatca 240 tcaagcaagg accaccatgg catttcaggg gtcacagcca cagccagtga gaagagatcc 300 aattgcagct atggaagcgg gtttccgggt ttgggtcatg actcgacagg tcatggaggg 360 aaggggagggt ccagggtgca gcaagagaga cgaaaaggga ttccatggac agaagaagaa 420 cataggctgt ttctgcttgg tctggacaag ttgggaaag gggattggag gaggatttca 480 agaaactttg tgatatcaag gactccaaca caagtggcca gtcattgcca gaaatacttc 540 attagattga actccatcaa cagagataga aggagatcaa gcattcatga catcaccagc 600 gtcaatagtg gcgatgcttt ggcacaccaa ccgccgccca tccattggcca gcagccaaat 660 ccaactagca cagtagcggc agccattgga ccaccaccgc cacctggaaa gcacaggatt 720 cccatccac cgccaccgca tatggccggc ttgggcatgt atggaacgcc ggtaggggat 780 ccggtgactg ctccgagtca tcatatggca tcagctgttg ggacgccagt tatgtctcct 840 cctggacacc accacccta tgttgtcca gtagcatatc caatggcacc ccaacaatat 900 caccaataa 909
3-104	Sequences	
3-104-1	Sequence Number [ID]	104
3-104-2	Molecule Type	DNA
3-104-3	Length	885
3-104-4	Features Location/ Qualifiers	source 1..885 mol_type=other DNA organism=Actinidia chinensis
3-104-5	NonEnglishQualifier Value Residues	atggaaaacc ttgcgaaag aatgccatgc aattcactgg gcagaggagg actactaaga 60 tggagggacc attgggaaaa gattgcttca ttggttccta ctaaaacat ggaagaattg 120 aagaaacatt accaattact agttgaggat gttgtgcaa tcgaatccgg acacgtgcca 180 atccccagct atattgggga ggaagcatca tcatcatcaa gcaaggacca ccatggcatt 240 tcaggggtca cagccacagc cagtgagaag agatccaatt gcagctatgg aagcgggttt 300 ccgggtttgg gtcatgactc gacaggatcat ggagggaagg gagggtccag ggctgagcaa 360 gagagacgaa aagggtattc atggacagaa gaagaacata ggctgtttct gcttggtctg 420 gaaaagtttg ggaaagggga ttggaggagc atttcaagaa actttgtgat atcaaggact 480 ccaacacaag tggccagcca tgcccagaaa tacttcatta gattgaactc catcaacaga 540 gatagaagga gatcaagcat ccatgacatc accagcgtca atagtggcga tgcttcggcg 600 caccaaccac cgccatcac tggccagcag ccaaatccaa ctgacagat agcgacagcc 660 attggaccac caccgccacc tggaagcac aggattcccg tcccactgtc accacatatg 720 gcaggcttgg gcatgtatgg aacaccggta gggcatccgg tggctgtctc gagtcatcat 780 atggcgtag ctgttgggac gccagttatg ctctctctg gacaccacca cccttatgtt 840 gtcccagtag catatccaat ggcaccccaa acaatgcacc aataa 885
3-105	Sequences	
3-105-1	Sequence Number [ID]	105
3-105-2	Molecule Type	DNA
3-105-3	Length	936
3-105-4	Features Location/ Qualifiers	source 1..936 mol_type=other DNA organism=Actinidia eriantha
3-105-5	NonEnglishQualifier Value Residues	atgtcagatt ggagcagaga agaagagaaa gcttttgaga atgccattgc aattcactgg 60 gcagaggaag actccaaaga agacgaccat tggccaaaga ttgcttcgtt ggcttcctact 120 aaaaccattg aagaattgaa gaaacattat caattactag ttgaggatgt tgctgcaatc 180 gaagccggac acgtgccat ccctaattat attggagagg aagcatcatc atcgtaagc 240 aaggaccgcc atggcatttc aggggtgaca gccagtgaga agagatccaa ttgcaactat 300 ggaagtgggt ttttccggg ttttggtcat gactatccgg cggttccatg gaggaaggag 360

		ggtccagggc tgagcaagag agacgaaaag ggattccatg gacagaagaa gaacatagggc 420 ttgctcttgg gcaatgctat gcgccaatta ctagttgata atgccctagg agcattgaat 480 aacctctcca aactgttttt gcttggtctg gacaagtttg ggaaagggga ttggagaagc 540 atttcaagga actttgtgat atcaaggact ccaacacaag tggctagcca tgcccagaaa 600 tacttcatca gattgaactc catcaacaga gatagaagga gatcaagcat ccatgacatc 660 accagcgta atggcgggtg tgcctcgca caccaaccgc cgcccatcac tggccagcag 720 ccgaattcga ctagtgcaat ggagccatt ggaccacccc aaccgcacat ggccggttg 780 ggcatgtatg gggcgccagt agggcatccg gtggctgctt cgagtcacat tatggcgctc 840 gccgttggga caccggttat gcttcctcct ggacaccacc acccctatgt tgtcccagta 900 gcatatccaa tggcaccccc aacaatgcac caataa 936
3-106	Sequences	
3-106-1	Sequence Number [ID]	106
3-106-2	Molecule Type	DNA
3-106-3	Length	867
3-106-4	Features Location/ Qualifiers	source 1..867 mol_type=other DNA organism=Actinidia chinensis
3-106-5	NonEnglishQualifier Value Residues	atgtcagctt cagtggattg gagcagagaa gaagagaaa cttttgagaa tgccattgca 60 attcactggg cagaggaaga ctccaaagat gacgaccatt ggacaagat tgcttcattg 120 gttcctacta aaaccattga agaattgaag aaacattatc aattactagt tgaggatggt 180 actgcaatcg aagccggaca tgtgccatc cctaattata ttggagagga agcatcatca 240 tcgttaagca aggaccgcca tggcatttca ggggtgacag ccagtgagaa gagatccaat 300 tgcattctat gaagtgggtt tccgggtttt ggtcatgact cggcagggtc tggaggggaa 360 ggaggggtcca gggctgagca agagagacga aaagggtatt catggacaga agaagaacat 420 aggcgtgttt tgcttggctt ggacaagtgt gggaaagggg attggagaag catttcaagg 480 aactttgtga tatcaaggac tccaacacaa gtggctagcc atgccagaa atacttcac 540 agattgaatt ccatcaacag agatagaagg agatcaagca tccatgacat caccagcgtc 600 aatggcggag atgcttcggc acaccaaccg ccgccatca tgggccagca gccgaattcg 660 actagggcaa tggcagccat tggaccaccc ccaccgcaca tggccggttt gggcatgtat 720 ggggcggcag tagggcatcc ggtggctgct tcgagtcac atatggcatc ggctgttggg 780 acaccggtta tgcttcctcc tggacaccac caccctatg ttgtccctgt agcatatcca 840 atggcacccc caacaatgca ccaataa 867
3-107	Sequences	
3-107-1	Sequence Number [ID]	107
3-107-2	Molecule Type	DNA
3-107-3	Length	912
3-107-4	Features Location/ Qualifiers	source 1..912 mol_type=other DNA organism=Actinidia chinensis
3-107-5	NonEnglishQualifier Value Residues	atgtcagctt cagtggattg gagcagagaa gaagagaaa cttttgagaa tgccattgca 60 attcactggg cagaggagga ctccaaagat gaggaccatt gggaaaagat tgcttcattg 120 gttcctacta aaaccatgga agaattgaag aaacattacc aattactagt tgaggatggt 180 gctgcaatcg aatccggaca cgtgccaatc ccagctata ttggggagga agcatcatca 240 tcatcaagca aggaccacca tggcatttca ggggtcacag ccacagccag tgagaagaga 300 tccaattgca gctatggaag cgggtttccg ggtttgggtc atgactgcac aggtcatgga 360 gggaagggag ggtccagggc tgagcaagag agacgaaaag ggattccatg gacagaagaa 420 gaacataggg tgtttctgct tggcttgtaa aagtttggga aaggggattg gaggagcatt 480 tcaagaaact ttgtgatatc aaggactcca acacaagtgg ccagccatgc ccagaaatac 540 ttcattagat tgaactccat caacagagat agaaggagat caagcatcca tgacatcacc 600 agcgtcaata gtggcgatgc ttcggcgcac caaccaccgc ccatcactgg ccagcagcca 660 aatccaacta gcacagtagc gacagccatt ggaccaccac cgccacttgg aaagcacagg 720 attcccgtcc cactgtcacc acatatggca ggcttgggca tgtatggaac gccggtaggg 780 catccggtgg ctgctccgag tcatcatatg gcgtcagctg ttgggacgcc agttatgctt 840 cctcctggac accaccacc ttatgttgtc ccagtagcat atccaatggc accccaaaca 900 atgcaccaat aa 912
3-108	Sequences	
3-108-1	Sequence Number [ID]	108
3-108-2	Molecule Type	DNA
3-108-3	Length	867
3-108-4	Features Location/ Qualifiers	source 1..867 mol_type=other DNA organism=Actinidia chinensis
	NonEnglishQualifier Value	

3-108-5	Residues	atgtcagctt cagtggattg gaggagagaa gaagagaaag cttttgagaa tgccattgca 60 attcactggg cagaggaaga ctccaaagat gacgaccatt ggacaagat tgcttcattg 120 gttcctacta aaaccattga agaattgaag aaacattatc aattactagt tgaggatggt 180 gctgcaatcg aagccggaca tgtgcccac cctaattata ttggagagga agcatcatca 240 tcgttaagca aggaccgcca tggcatttca gggttgacag ccagtgagaa gagatccaat 300 tgcattctatg gaagtgggtt tccgggtttt ggtcatgact cggcagggtca tggaggggaag 360 ggaggggtcca gggctgagca agagagacga aaagggattc catggacaga agaagaacat 420 aggctgtttt tgcttggctt ggacaagttt gggaaagggg attggagaag catttcaagg 480 aactttgtga tatcaaggac tccaacacaa gtggctagcc atgcccagaa atacttcac 540 agattgaatt ccatcaacag agatagaagg agatcaagca tccatgacat caccagcgtc 600 aatggcggag atgcttcggc acaccaaccg ccaccatca tgggccagca gccgaattcg 660 actagtgcaa tggcagccat tggaccaccc ccaccgcaca tggccgggtt gggcatgtat 720 ggggcgccag tagggcatcc ggtgggtgct tcgagtcac atattggcatc ggctgttggg 780 acaccggtta tgcttcctcc tggacaccac caccctatg ttgtccctgt agcatatcca 840 atggcacccc caacaatgca ccaataa 867
3-109	Sequences	
3-109-1	Sequence Number [ID]	109
3-109-2	Molecule Type	AA
3-109-3	Length	299
3-109-4	Features Location/ Qualifiers	source 1..299 mol_type=protein organism=Actinidia eriantha
3-109-5	NonEnglishQualifier Value Residues	MGSEDVKSP KSAKVSPSKQ EHTNIHLYPE WPNIQAYYGP GVPVPQSYLN SAAVAGHVP 60 PHLWSPQSL MPYGLPYST IYSPGVVYAH PAVPPDSDQE TAVFSAAMMS CPAVTVIQVA 120 TPVSTQQISN SSDITEQLR KLKLSVDGVA MCAVKDNVDV DGEVSHGAS QSAEGGIGAS 180 SNATDCNYGD QRQLKRVRRK QANREFARRS RLRKQAETEE LMERYETLKV ENIALKSEM 240 QLTEGSEKLR VENAALMEKL SVADVTHHGE MVLDEIEPDL AQPNCSEDGF TLNKP GSVN 299
3-110	Sequences	
3-110-1	Sequence Number [ID]	110
3-110-2	Molecule Type	AA
3-110-3	Length	284
3-110-4	Features Location/ Qualifiers	source 1..284 mol_type=protein organism=Actinidia chinensis var. chinensis
3-110-5	NonEnglishQualifier Value Residues	MGSEDVKSP KSAKVSPSKQ AYYGPGVPVP QSYLNSAVVA GHVPHPLWS SPQTLMPPYG 60 LPYSTIYSPG VVYAHPAVPP DDSQETAVFS AAMMSCPAVT VIQVATPVST QQISNSSDIT 120 DQGLRKKLKS VDGVMCAVK DNVVDVGEVS VHGSQSAEG GIGASSNATD CNYGDQRQLK 180 RVRRKQANRE SARRSRLRKQ AETEELMERY ETLKVENIAL KSEMNLTEG SEKL RVENAA 240 LMEKLSVADV THHGEMVLDE IEPDLAQPN SEDGFTLNKP GPVN 284
3-111	Sequences	
3-111-1	Sequence Number [ID]	111
3-111-2	Molecule Type	AA
3-111-3	Length	265
3-111-4	Features Location/ Qualifiers	source 1..265 mol_type=protein organism=Actinidia chinensis var. chinensis
3-111-5	NonEnglishQualifier Value Residues	MGSEDMKSP KSAKVSPSNQ EHTNIHLYPD WANIQAYYGP GVPVPPSYLN PAVMAGHVP 60 PHLWSPQTL MPYGVYPYST IYSPGVVYAH PAVPLVAAPV STQQLSNSSD ITDQGLRKKL 120 KSVDGVAMRD VKNNVDVDGE GLVHEASQSA ECGIRASSNA SDCNYGDQRQ HKQERRKQAN 180 RESAKRSRLR KQAETEELME RYETLKVENI ALKSEMNLQM EGSEKL RVES AALMEKLNVA 240 DVTHKEEMVL DEIEPDLAQ NCSEE 265
3-112	Sequences	
3-112-1	Sequence Number [ID]	112
3-112-2	Molecule Type	AA
3-112-3	Length	329
3-112-4	Features Location/ Qualifiers	source 1..329 mol_type=protein organism=Actinidia rufa
3-112-5	NonEnglishQualifier Value Residues	MVGCVVSQAS GLVRMDVAYY GPGVPVPQSY LNSAAVAGHV PHPLWSPQ DDSQETAVFS 60 AAMMSCPAVT VIQVATPVST QQISNSSDIT DQGLRKKLKS VDGVMCAVK DNVVDVGEVS 120 VHGSQSAEG GIGASSNATD WDANMTARSV SGIIGAPENA GRKPVDTPNG KAKTSTTCAA 180

		TPAGVELPLE VVLRCLAAV FFCFLPPVYF KWFYMEHIND QRQLKRVRRK QANRESARRS 240 RLRKQAETEE LMERYETLKV ENIALKSEM N QLTEGSEKLR VENAALMEKL SVADVTHHGE 300 MVLDEIEPDL AQPNCSEDGF TLNKPGSVN 329
3-113	Sequences	
3-113-1	Sequence Number [ID]	113
3-113-2	Molecule Type	AA
3-113-3	Length	370
3-113-4	Features Location/ Qualifiers	source 1..370 mol_type=protein organism=Actinidia rufa
3-113-5	NonEnglishQualifier Value Residues	MGSSEEMKSP KSAKVSPSNQ EHTNIHLYPD WANIQAYYGP GVPVPPSYLN SAVMAGPVPH 60 PHSWSPPTL MPSYGVYST IYSPGVVYAH PAVPLDSDQE TTLFSANMMS YPAVTVIQVA 120 APVSTQQLSN SSDITDQGLR KKLKSDVGVA MRDVKNVNDV DGEGLVHGAS QSAEGGIRAS 180 SNASNYNDVK VHTQARPVQ GDANVTARSV LGIIVAPENA ARQPVDTDTN NGKAKTIKTC 240 DAATSGVGLP LEVRIQDQRQ QKQERRKQAN RESARRSRLR KQAETEELME RYETLKVENI 300 ALKSEMNLTL EGSEKLRVEN DALMEKLNVA DVTHKEEMVL DEIEPDLAQR NCSEDGVQLN 360 KSGPVNGNKG 370
3-114	Sequences	
3-114-1	Sequence Number [ID]	114
3-114-2	Molecule Type	AA
3-114-3	Length	325
3-114-4	Features Location/ Qualifiers	source 1..325 mol_type=protein organism=Actinidia chinensis var. chinensis
3-114-5	NonEnglishQualifier Value Residues	MGSSEDMKSP KSAKVSPSNQ EHTNIHLYPD WANIQAYYGP GVPVPPSYLN PAVMAGHVPH 60 PHLWSPPTL MPSYGVYST IYSPGVVYAH PAVPLVAAPV STQQLSNSSD ITDQGLRKKL 120 KSDVGAMRD VKNNVDVDGE GLVHEASQSA ECGIRASSNA SDCNYGARPV TQGDANVTAR 180 SVLGIIVAPE NAARQPVDTL NGKAKTIKTC AAATSGVGLP LEVQIQDQRQ HKQERRKQAN 240 RESAKRSRLR KQAETEELME RYETLKVENI ALKSEMNLML EGSEKLRVES AALMEKLNVA 300 DVTHKEEMVL DEIEPDLAQP NCSEE 325
3-115	Sequences	
3-115-1	Sequence Number [ID]	115
3-115-2	Molecule Type	AA
3-115-3	Length	261
3-115-4	Features Location/ Qualifiers	source 1..261 mol_type=protein organism=Camellia sinensis
3-115-5	NonEnglishQualifier Value Residues	FAVFFLLFLH SRFFFCFFL QFYFPFCHLL LYVMGRSEDM ISPKSEKPCP KSEKASPPKQ 60 AAYGPGVPIP PSYLNSSVMA CHVPHPHVWA PPQHVMPPYG VPYSAIYSPG VIATPVSAQM 120 LSTSSYNMDQ GSIKKLRRLD GVTQVVNDN AEVDGEGSAH GASQSIDVSS DASDGNTRDE 180 RQLKREKRKH ANKQSARRSR LRKQAETEEL MKRYEIVKVE NITLKSEMKQ LMESEKLRV 240 ENVALMENLN GSEVAHPERR V 261
3-116	Sequences	
3-116-1	Sequence Number [ID]	116
3-116-2	Molecule Type	AA
3-116-3	Length	376
3-116-4	Features Location/ Qualifiers	source 1..376 mol_type=protein organism=Telopea speciosissima
3-116-5	NonEnglishQualifier Value Residues	MGNKTVTPK SEKASSPGQE QTNVHAYTDW ANIQAYYGP VIPPPYFNSA VASGHAPPPY 60 MWGHPQLMP PYGAPYAIYS HGGVYAHPAV PYGSHGHGIP PSPVLSEFMV ATPLSVEPPA 120 KSSTSKDRGL MKKLKGFDEL AVSIGNKAE NAVTGSAGHM SQSEECGTEG SSDGSDGNTG 180 GDKSNRKRNS EGSPSTGKDG KLDPVPVGEA NVTPGVVPA CDALPSELWV QDERELKRER 240 RKQSNRESAR RSRLRKQAE EELALKVETM SAENIALRSE LNRLMENSEK LRLENVAVLME 300 KLKNAQLEEE GEMAPNNMES ETAPHDSTEN LLSRVNNSSS VTSEQEDSEP NENSNSSTKL 360 HQLLETSPRA DAVAAG 376
3-117	Sequences	
3-117-1	Sequence Number [ID]	117
3-117-2	Molecule Type	AA
3-117-3	Length	358

3-117-4	Features Location/ Qualifiers	source 1..358 mol_type=protein organism=Nelumbo nucifera
3-117-5	NonEnglishQualifier Value Residues	MGNNEAATPP KSEKNSSPAQ EQTNIHPYPD WAAMQAYYGP GVTLPPPYFN PAVTSGHAPH 60 PYMWGPPQPL MPYPGAPYAA IYPHGGVYAH PSITYGSHPH GPGFPPSPVV SEVMVATPVS 120 VEPPAKSSGS KDRDLMKKLR GLGNGNAEGA AEGATSGLSL SAECTGDGSS DGSDGNTSGG 180 DKTERKRSSD SILSAGSDGK VETPAISMPS AQDERELKRQ RRRQSNRESA RRSRLRKQAE 240 TDELVMKVES LTAENIALKS EISRLTQSSE KMRLNAALM EKLKDAELGQ TTEVVANDTE 300 AREDPPVSTE NFLSRVNNPA SVSIDDRRDT EIHKNSNSGA KLRQLLESNT RTDAVAAG 358
3-118	Sequences	
3-118-1	Sequence Number [ID]	118
3-118-2	Molecule Type	AA
3-118-3	Length	377
3-118-4	Features Location/ Qualifiers	source 1..377 mol_type=protein organism=Macadamia integrifolia
3-118-5	NonEnglishQualifier Value Residues	MGNNKAGAPK SEKASSPVQE QTNVHAYTDW ANNIQAYYGP GVIPPPYFNS AVASGHAPPP 60 YMWGHPQTLN PPYGAPYAIY SHGGVYAHPG VPYGSHGHGI PPSPVVSEVM VATPLSVEPP 120 AKSSTSKDRG LMKKLKGFDDG LAVSLGNGKD ENAVRGSAGH MSQSEEYGTD GSSDGSNGNT 180 GGDKSNRKRN SEGTPSTGKD GKLDPAVPGE ANVTPGVVVP ACDALPSELW IQDERELKRE 240 RRQSNRESA RRSRLRKQAE TEELALRVET MSAENIALRS ELNRLMENSE KLRLNAVLM 300 EKLKNTQPEE AGKMAPDNME SETAPPDSTE NLLSRVNNSS SVTSEQGDSE PNENSNSSTK 360 LHQLLESSPR ADAVAAG 377
3-119	Sequences	
3-119-1	Sequence Number [ID]	119
3-119-2	Molecule Type	AA
3-119-3	Length	333
3-119-4	Features Location/ Qualifiers	source 1..333 mol_type=protein organism=Camellia sinensis
3-119-5	NonEnglishQualifier Value Residues	MATSKDFPSL KPTLIDTKSF RKPGLIHSKH LPLCFWNFYF PFCHELLLYYM GRSEDMISPK 60 SEKPCPKSEK ASPPKQAYYG PGVPIPPSYL NSSVMACHVP HPHVWAPPQH VMPPYGVYPS 120 AIYSPGVVCT HRTVPIIATP VSAQMLSTSS YNMDQGSIKK LKRLDGVTKQ VVNDNAEVDG 180 EGSAHGASQR MEVYHVSEHA YQYAQQSLRK SDSMGKLSTD NNVKVHKQAR SVPEKDANVT 240 SRWISRTNVA PEDVAGKLDE RQLKREKRKH ANKQSARRSR LRKQAETEEL MKRYEIVKVE 300 NITLKSEMKQ LMEDESEKLRV ENVALMEIIR GAI 333
3-120	Sequences	
3-120-1	Sequence Number [ID]	120
3-120-2	Molecule Type	AA
3-120-3	Length	371
3-120-4	Features Location/ Qualifiers	source 1..371 mol_type=protein organism=Telopea speciosissima
3-120-5	NonEnglishQualifier Value Residues	MGNNKTVTPK SEKASSPGQE QTNVHAYTDW ANIQAYYGP VIPPPYFNSA VASGHAPPPY 60 MWGHPQLMP PYGAPYAIYS HGGVYAHPAV PYGSHGHGIP PSPVVATPLS VEPPAKSSTS 120 KDRGLMKKLK GFDGLAVSIG NGKAENAVTG SAHGMSQSEE CGTEGSSDGS DGNTGGDKSN 180 RKRNSEGSPS TGKDGKLDPV PVGEANVTPG VVVPACDALP SELWVQDERE LKRERRKQSN 240 RESARRSRLR KQAETEELAL KVETMSAENI ALRSELNRLM ENSEKLRLN AVLMEKLNKNA 300 QLEEEGEMAP NNMESETAPH DSTENLLSRV NNSSSVTSEQ EDSEPNNSN SSTKLHQLLE 360 TSPRAVAAG G 371
3-121	Sequences	
3-121-1	Sequence Number [ID]	121
3-121-2	Molecule Type	AA
3-121-3	Length	372
3-121-4	Features Location/ Qualifiers	source 1..372 mol_type=protein organism=Macadamia integrifolia
3-121-5	NonEnglishQualifier Value Residues	MGNNKAGAPK SEKASSPVQE QTNVHAYTDW ANNIQAYYGP GVIPPPYFNS AVASGHAPPP 60 YMWGHPQTLN PPYGAPYAIY SHGGVYAHPG VPYGSHGHGI PPSPVVATPL SVEPPAKSST 120 SKDRGLMKKL KGFDDLAVSL GNGKDENAVER GSAHGMSQSE EYGTDGSSDG SDGNTGGDKS 180

		NRKRNSEGTP STGKDGKLDP APVGEANVTP GVVVPACDAL PSELWQDER ELKRERRKQS 240 NRESARRSRL RKQAETEELA LRVETMSAEN IALRSELNRL MENSEKLRL NAVLMEKLN 300 TQPEEAGKMA PDNMESETAP PDSTENLLSR VNNSSSVTSE QGDSEPNENS NSSTKLHQLL 360 ESSPRADAVA AG 372
3-122	Sequences	
3-122-1	Sequence Number [ID]	122
3-122-2	Molecule Type	AA
3-122-3	Length	325
3-122-4	Features Location/ Qualifiers	source 1..325 mol_type=protein organism=Prunus persica
3-122-5	NonEnglishQualifier Value Residues	MGNDDDGKSA KSEKSSSPLT ADQTNHSNQT NVHVYPDWAA MQAYYGPRVA LPPYYNSAVA 60 SGHAPHPYMW GPPQPMMPY GTPYAAMYSH GYAHPAVPLG SHGQAVPSSP SAATPLNMET 120 PTKSSGNADR GLMKKLKRFD ALAMSIGNSN VGDAEGGAEH SVSQSLGTEG SSEGSDGNTA 180 GNERELKRER RKQSNRESAR RSRLRKQAE EELARKVDAL SAENVALKSE ISRLTENSST 240 LKLENATLMD KLNTRVGRM EEIMMNIDDK RVQPFSTENL LSRVNNSGSI ARDAEKEGDM 300 YEKNSGAKLH QLLDASPRAD AVAAG 325
3-123	Sequences	
3-123-1	Sequence Number [ID]	123
3-123-2	Molecule Type	AA
3-123-3	Length	371
3-123-4	Features Location/ Qualifiers	source 1..371 mol_type=protein organism=Nyssa sinensis
3-123-5	NonEnglishQualifier Value Residues	MASKEDKTAP KSEKAYPSEQ EQTDIHLFPD WAAIQPYYPG GVPLPPPYFN FAVPAGHFSS 60 PHVWNLQQSL MPPYGVYPYST IYSHGVVYAH PAVPLVAAPV STELPKSSE NIDQCLTKKL 120 KRSDGLAVPV GNGGAEVDAE GSVHEASHSV EHSIDASSHG SDGNTGAHQS QRKNDMSGIL 180 PTDDDVKVHM QASPLLEGEA NITSGRVSGI TAAPENVVGK SVESPIGETE ASAACGASTA 240 AAGLPFEVRD ERQLKRERRK QANRESAKRS RLRKQVETEE LLARYESLKM ENLALKSEIK 300 QLTQESDKVR LENAALMEKL SDAQLGQPRK TVPDQIGNNV DQPHGTLCK LTKLGSVSR 360 ENEECKTLRN S 371
3-124	Sequences	
3-124-1	Sequence Number [ID]	124
3-124-2	Molecule Type	AA
3-124-3	Length	363
3-124-4	Features Location/ Qualifiers	source 1..363 mol_type=protein organism=Corylus heterophylla
3-124-5	NonEnglishQualifier Value Residues	MGNNEEGKSS KPEKSSSPAT PDQTNQTNTQ NIHVYPDWAS MQAYYGPRVA LPPYYNSTVG 60 SGHAPHPYMW GPAQPMMPY GAPYAAIYPH GGVAHPGVP LGSHPHGQGI PLSPAAVTPL 120 SIETPTKSGN TDRGLMKKLK EFDGLAMSIG NGTAEIAEGG AEHRLSESLE TVITPASAAG 180 NLVVSSGMAT PNMNTKTTTN SVPLPCAVLP PEAWVQNERE QKRERRKQSN RESARRSRLR 240 KQAETEELAH KVDLSNAENV AIKSELERLT ENSEKLRLN ATLMEKLKNA KLGQTVEISL 300 NSIDDKRALP VSTENLSRV NNSGSDRSI EEKSDMYENN SNSGAKLHQL LDASPRADAV 360 AAG 363
3-125	Sequences	
3-125-1	Sequence Number [ID]	125
3-125-2	Molecule Type	AA
3-125-3	Length	324
3-125-4	Features Location/ Qualifiers	source 1..324 mol_type=protein organism=Gossypium barbadense
3-125-5	NonEnglishQualifier Value Residues	MGNNDEGKSK SDKSSSPVPT DQANIHVYPD WAAMQAYYGP RVTMPPPYNS AVASGHAPPP 60 YMWGPPQPM PPYGAPYAAI YSHGGVYAH AVPITAAPVE TPTKSSGNTE RGSTKKMKGF 120 DGLAMSIGNG TAENDEGAAE PRLSQAETE GSTDGSNGT TGLEMVKLRQ SLMQLPWQRN 180 PSSMNAKSSP TNVPCGVMP EVWMQNEREL KRERRKQSNR ESARRSRLRK QAETEELARK 240 VESLTAENAT LRSEINQLTE KSEKLRLNA TLVEGLKNAQ LGHTQENITN KNEDKEGEM 300 EKKSGAKLHQ LLDASPRTD VAAS 324
3-126	Sequences	
3-126-1	Sequence Number [ID]	126
3-126-2	Molecule Type	AA

3-126-3	Length	325
3-126-4	Features Location/ Qualifiers	source 1..325 mol_type=protein organism=Gossypium darwinii
3-126-5	NonEnglishQualifier Value Residues	MGNNDEGKSK SDKSSSPVPT DQANIHVYPD WAAMQAYYGP RVTMPPPYNS AVASGHAPPP 60 YMWGPPQPMM PPYGAPYAAI YSHGGVYAHF AVPITAAPVE TPTKSSGNTG RGSTKKMKGF 120 DGLAMSIGNG TAENDEGAEE PRLSQAETE GSTDGSNGT TGQLEMVKLR QSLMQLPWQR 180 NPSSMNAKSS PTNVPCGVMP PEVWMQNERE LKRERRKQSN RESARRSRLR KQAETEELAR 240 KVESLTAENA TLRSEINQLT EKSEKLRLN ATLVEGLKNA QLGTQENIT NKNEDKEGEM 300 YEKKSGAKLH QLLDASPRD AVAAS 325
3-127	Sequences	
3-127-1	Sequence Number [ID]	127
3-127-2	Molecule Type	AA
3-127-3	Length	357
3-127-4	Features Location/ Qualifiers	source 1..357 mol_type=protein organism=Rhododendron simsii
3-127-5	NonEnglishQualifier Value Residues	MGKSDDTRPS KPEKPSSPPQ DQPNVHVYPD WAAMQAYYGP RVPVPPYFNS AGASGHAPHP 60 YMWGPPQHMM PPYGPYAAI YAHGGVYAHF GVPLAGTFPM ETPTKSSGHT DQGTVKKLKG 120 FDGLAMSIGN SNANSADGGA NHVQSQSLET ETSSDTSNGN TAGAGRVKRK SHEKTPTLGG 180 DGNTEQNSP DVPPGVNKG PAVMPNEAW LQSERELKRE RRKQSNRESA RRSRLRRQAE 240 AEELSVKVES LNADNLTKS EISQLTEISE KLKLENASLM DKMKNAKGGR TDAVISTKVS 300 EKRVPGGTEN LLSRVNNSGS VDRSDEGNEM YEKSSKSGAK LHQLLDASPR ADAVAAR 357
3-128	Sequences	
3-128-1	Sequence Number [ID]	128
3-128-2	Molecule Type	AA
3-128-3	Length	326
3-128-4	Features Location/ Qualifiers	source 1..326 mol_type=protein organism=Papaver somniferum
3-128-5	NonEnglishQualifier Value Residues	MGSDETGTSP KSEKSSSSSA LTRQEQQNIP AYPDWPSLQA YYGAGVPVPP PYFNPALTGH 60 PPHPYIWGPP QHMMPPYGSP YTAIYPSGGI YGHPSVSLGS YGHGIPLSA TGETLVATPL 120 SIATPVKFPS SQDLGMKKKV SNTDGGGSQS AGTALEVVPV HGTHPAHVPA VLNRDGVTS 180 HPDDHELKKQ RRKQSNRESA RRSRLRKQAE SEELTLKVES LNVENMALKS EMRRLAEDSE 240 KLRLENTSML EKLKNKQATE AASNEVEVAQ GANTENFLTR VSDSGSLCMH PHQESKTSTN 300 ENANSGTANL HQLLEASPR DAVAVG 326
3-129	Sequences	
3-129-1	Sequence Number [ID]	129
3-129-2	Molecule Type	AA
3-129-3	Length	328
3-129-4	Features Location/ Qualifiers	source 1..328 mol_type=protein organism=Papaver somniferum
3-129-5	NonEnglishQualifier Value Residues	MGSDETGTSP KSEKSSSSSA LTRQEQQNIP AYPDWPSLQA YYGAGVPVPP PYFNPALTGH 60 PPHPYIWGPP QHMMPPYGSP YTAIYPSGGI YGHPSVSLGS YGHGIPLSA TGETLVATPL 120 SIATPVKFPS SQDLGMKKKV SNTDGGGSQS AGTALEVVPV HGTHPAHVPA VLNRDGVTS 180 HPVLDDHELK KQRRKQSNRE SARRSRLRKQ AESEELTLKV ESLNENMAL KSEMRRLAED 240 SEKLRLENTS LMEKLKNKQA TEAASNEVEV AQGANTENFL TRVSDSGSLC MHPHQESKTS 300 TNENANSHTA NLHQLLEASP RTDAVAVG 328
3-130	Sequences	
3-130-1	Sequence Number [ID]	130
3-130-2	Molecule Type	AA
3-130-3	Length	335
3-130-4	Features Location/ Qualifiers	source 1..335 mol_type=protein organism=Papaver somniferum
3-130-5	NonEnglishQualifier Value Residues	MGSDETGTSP KSEKSSSSSA LCIFLTRQTR QEQQNIPAYP DWPSLQAYYG AGVPVPPPYF 60 NPALTGHPPH PYIWGPPQHM MPPYGSPTA IYPSGGIYGH PSVSLGSYGH GIPLSPATGE 120 TLVATPLSIA TPVKFPSSQD LGMKKKVSND TGGGSQSAGT ALEVVP SHGT HPAHVPAVLN 180 RDGVTSEHPV LDDHELKKQR RKQSNRESAR RSRLRKQAES EELTLKVESL NVENMALKSE 240

		MRRLAEDSEK LRLENTSLME KLKNKQATEA ASNEVEVAQG ANTENFLTRV SDGSLCMHP 300 HQESKTSTNE NANSGTANLH QLEASPRTD AVAVG 335
3-131	Sequences	
3-131-1	Sequence Number [ID]	131
3-131-2	Molecule Type	AA
3-131-3	Length	325
3-131-4	Features Location/ Qualifiers	source 1..325 mol_type=protein organism=Papaver somniferum
3-131-5	NonEnglishQualifier Value Residues	MGSDETGTPS KSEKSSSSSA LEQNNIPAYP DWPSLQAYYG AGVPVPPPYF NPALTGHPPH 60 PYIWGPPQHM MPPYGSPYTA IYPSGGIYGH PSVSLGSYGH GIPLSPATGE TLVATPLSIA 120 TPVKFPSSQD LGMKKKVSND TGGGSQSAGT ALEVVP SHGT HPAHVPAVLN RDGVTSEHPV 180 LDDHELKKQR RKQSNRESAR RSRLRKQAES EELTLKVESL NVENMALKSE MRRLAEDSEK 240 LRLENTSLME KLKNKQATEA ASNEVEVAQG ANTENFLTRV SDGSLCMHP HQESKTSTNE 300 NANSGTANLH QLEASPRTD AVAVG 325
3-132	Sequences	
3-132-1	Sequence Number [ID]	132
3-132-2	Molecule Type	AA
3-132-3	Length	352
3-132-4	Features Location/ Qualifiers	source 1..352 mol_type=protein organism=Papaver somniferum
3-132-5	NonEnglishQualifier Value Residues	MGSSEAETTP SKSEKESSTP TQEQTSM PAY PNWAAMQAYY PPGMAMPPPP YYNPAVGS GH 60 PPPPFMWGAP QGMPPPYGSP YGPIYAHGGV YAHPSVAPGA HSQGPTVSAS PVGGEATVAT 120 PLSVTTPAKS SSNKEQGLMK KLKGF DGLVV TLGNGDANGG SGDM SHSGDY KTDGSSD GSD 180 GNTARKQQPR RKRSCDDMTS TGAEVSKHEG VPSDLWKSE RELKRERRKQ SNRESARRSR 240 LRKQAE TEEL ATKVEALNVE NMSLKSEISR L TEKSEKLRL ENAAIMEK LK NAQSSNMVIE 300 EAAQPISTEN FLARVNNNSS QQESEVQENS NSG TKLHQLL ESSPRT DAVA AG 352
3-133	Sequences	
3-133-1	Sequence Number [ID]	133
3-133-2	Molecule Type	AA
3-133-3	Length	352
3-133-4	Features Location/ Qualifiers	source 1..352 mol_type=protein organism=Papaver somniferum
3-133-5	NonEnglishQualifier Value Residues	MGSSEAETTP SKSEKESSTP TQEQTSM PVY PNWAAMQAYY PPGMAMPPPP YYNPAVGS GH 60 PPPPFMWGAP QGMPPPYGSP YGPIYAHGGV YAHPSVAPGA HSQGP AVTAS PVGGEATVAT 120 PLSVTTPAKS SSNKEQGLMK KLKGF DGLVV TLGNGDANGG SGDM SHSGDY KTDGSSD GSD 180 GNTARKQQPR RKRSCDDMTS TGAEVSKHEG VPSDLWKSE RELKRERRKQ SNRESARRSR 240 LRKQAE TEEL ATKVEALNVE NMSLKSEISR LAEKSEKLRL ENAAIMEK LK NAQSSNMVIE 300 EATQPISTEN FLARVNNNSS QQESEVQENS NSG TKLHQLL ESSPRT DAVA AG 352
3-134	Sequences	
3-134-1	Sequence Number [ID]	134
3-134-2	Molecule Type	AA
3-134-3	Length	426
3-134-4	Features Location/ Qualifiers	source 1..426 mol_type=protein organism=Phoenix dactylifera
3-134-5	NonEnglishQualifier Value Residues	MDTCIVTWQL ARLPLHPQAS WPSSLPSTHP PLLASLFFSR SACVRINGAR IEAGVWVIYN 60 STFMGNDEAG TPPKSDKASS SMQEKPTICP YPDWTTMQAY YGPGVAMPHP FFSTSVAPGH 120 APNPYLWGPQ PLMPPPF GMP YAAIYPPAGV YSHPPMLMT TPLSMEMPAK SPSNKDKGLM 180 KKLKRFDGLA VSAGNGDAES RGGDRNGSSQ SEDDSVEGLS DSGDGNNAGG NKTRGKQSSK 240 DIRGPGDGKV DTQTNAAHGV ETSASSKMSL GVTVATADIS RKDERELKRE RRKQSNRESA 300 RRSRLRKQAE TEELAMKVET LGAENMALRS EISRLSEKSD NLRLNSTLM EKLKTSQSSQ 360 AEEMSAAKME TEGVPSVVAE NFLSIIDNSS SISASAQQQD EACENS SSKF HQLLDSKPRT 420 DALAAS 426
3-135	Sequences	
3-135-1	Sequence Number [ID]	135
3-135-2	Molecule Type	AA
3-135-3	Length	274

3-135-4	Features Location/ Qualifiers	source 1..274 mol_type=protein organism=Ipomoea triloba
3-135-5	NonEnglishQualifier Value Residues	MGKDEENNSP KCDLGSQPAQ QFPVPSSYLS TSLPTGHVPQ IHMWPSQQPF MPPYGMPPYM 60 IYPQGIVYAH PAVSLAAVTS SVELVPKPSE AIDQNSAKRL KGHNGLAVPI TNDNFEVNAE 120 GSVHGTSQSD DVATSTDESN EGMEDEPKTVE QPSLNRTPEI SFARNAVRFPP IEVSVQDERQ 180 LKREKRKQAN RESARRSRLR KQAEYEEVVR RYESLNKDNT NLKSEMKNLK EDMEKLRLEN 240 ATLMEKLEST QRPQPGEAVP NDPNVKLEDS CNKK 274
3-136	Sequences	
3-136-1	Sequence Number [ID]	136
3-136-2	Molecule Type	AA
3-136-3	Length	351
3-136-4	Features Location/ Qualifiers	source 1..351 mol_type=protein organism=Punica granatum
3-136-5	NonEnglishQualifier Value Residues	MGNREDAKSS KTEKPSTPPA SSDQPHVHVY PDWAAAQAYY GPRALALPSYY NSAVPPGYAP 60 HPYIWGPPQH MIPPYGAPFA AAYSHGGVYA HPATPITPTP MSGETPSKSS GDASRGLMCK 120 LKSFDFGLMP ITNGNVEAAE NGTKPRVSES TETEGTSDGS DGNTSGADQI RRKRMVDGTP 180 PAAVAKPEGG AAPDVNLKVP SETWLLNERE LKRERRKQSN RESARRSRLR KQAEAEELAN 240 RVGYLTDENA ALTAELNLLM EDSNKLRRN AMLMEKLKHL EPEGTGEEGS DTEVEEVLPL 300 TTENLLSKVD NAGSRETSTE EEGAYEQNLS AGSKLRQLLA AKPRADAVAA T 351
3-137	Sequences	
3-137-1	Sequence Number [ID]	137
3-137-2	Molecule Type	AA
3-137-3	Length	346
3-137-4	Features Location/ Qualifiers	source 1..346 mol_type=protein organism=Populus alba
3-137-5	NonEnglishQualifier Value Residues	MGNNEEGKSS ASDKSSPAQQ DQTSIHVYPD WAAIQAYYGS RVALPPYYS NSAVSGHAPHP 60 YMWGPPQPM ATYGAPYAAI YSHGGVYAH AAVPIGSHPH GPGVLSSPAA NTPLSAETPT 120 KSSGNTDQGL LKKLKGFDGL AMSIGNGDAE SAEGGSRLPQ SMETEGSSDG SDGNTARGKK 180 RSREGTPTAV VDPGMTKALE LRNPSSVNVA KTNPATIPQP GAVLPSEAWS PNDRELKRR 240 RKQSNRESAR RSRLRKQAEA EELAHKVETL TTVNVALKSE IDQFTEKSQK LRLENAALE 300 KLKNAQQGPA QEMTLNIDE QRNSNPGLAK HQLLDASPR DAVAAG 346
3-138	Sequences	
3-138-1	Sequence Number [ID]	138
3-138-2	Molecule Type	AA
3-138-3	Length	352
3-138-4	Features Location/ Qualifiers	source 1..352 mol_type=protein organism=Punica granatum
3-138-5	NonEnglishQualifier Value Residues	MGNREDAKSS KTEKPSTPPA SSDQPHVHVY PDWAAAQAYY GPRALALPSYY NSAVPPGYAP 60 HPYIWGPPQH MIPPYGAPFA AAYSHGGVYA HPATPITPTP MSGETPSKSS GDASRGLMCK 120 LKSFDFGLMP ITNGNVEAAE NGTKPRVSES STETEGTSDG SDGNTSGADQ IRRKRMVDGT 180 PPAAVAKPEG GAAPDVNLKV PSETWLLNER ELKRERRKQS NRESARRSRL RKQAEAEELA 240 NRVGYLTDEN AALTAELNLL MEDSNKLRR NAMLMEKLKH LEPEGTGEEG SDTEVEEVL 300 LTENLLSKV DNAGSRETST EEGAYEQNL SAGSKLRQLL AAKPRADAVA AT 352
3-139	Sequences	
3-139-1	Sequence Number [ID]	139
3-139-2	Molecule Type	AA
3-139-3	Length	357
3-139-4	Features Location/ Qualifiers	source 1..357 mol_type=protein organism=Rhododendron griersonianum
3-139-5	NonEnglishQualifier Value Residues	MGKSDDTRPS KPEKPSSPPQ DQPNVHVYPD WAAVQAYYGP RVPVPPYFNS AAASAHAPHP 60 YMWGPPQHMM PPYGPYAAI YAHGGVYAH GVPLAGTPFM ETPTKSSGHT DQGTVKKLKG 120 FDGLAMSIGN SNANSADGGA NHVQSQSVET ETSSDTSNGN TAGAGRVRRK SHEKTPTLGG 180 DGKTETQNSP DVPPGVNKAG PAVMPPNEAW LQNERELKRE RRKQSNRESA RRSRLRRQAE 240 AEELSVKVES LNADNMTLKS EISRLTDISE KLKLENASLM DKMKNAGGGR TDAVISTKVS 300 EKRVPGGTEN LLSRVNNSGS VDRSDEGNEM YEKSSKSGAK LHQLLDASPR ADAVAAR 357

3-140	Sequences	
3-140-1	Sequence Number [ID]	140
3-140-2	Molecule Type	AA
3-140-3	Length	345
3-140-4	Features Location/ Qualifiers	source 1..345 mol_type=protein organism=Populus trichocarpa
3-140-5	NonEnglishQualifier Value Residues	MGNNEEGKSS ASDKSSPAQQ DQTSIHVYPD WAAIQAYYGS RVALPPYYNS GVASGHAPHP 60 YMWGPPQPM ATYGAPYAAI YSHGGVYAH AVPIGSHPHG PGVLSSPAAH TPLSAETPTK 120 SSGNTDQGLM KKLKGF DGLA MSIGNDAES AEGGSRLPQS METEGSSDGS DGNTARGKKR 180 SREGTPTVVV DPGMTKALEL RNPPSVNVAK TNPATIPQPG AMLPSEAWSP NDRELKRERR 240 KQSNRESARR SRLRKQAEAE ELAHKVETLT TVNMTLKSEI DQFTEKSQKL RLENAALTEK 300 LKNAQLGPTQ EIILNNIDEQ RSSNSGAKLH QLMDASPRAD AVAAG 345
3-141	Sequences	
3-141-1	Sequence Number [ID]	141
3-141-2	Molecule Type	AA
3-141-3	Length	275
3-141-4	Features Location/ Qualifiers	source 1..275 mol_type=protein organism=Ipomoea nil
3-141-5	NonEnglishQualifier Value Residues	MGKDEENNSP KCDLGSQPAQ QFPVPSSYLS TSLPTGHVPQ IHMWPPQQPF MPPYGMYPYM 60 MYPQGIYVAH PAVTLAAVTS SVELVPKPSE AIDQNSAKRL KWHNGLAVPI TNDNFEVNAE 120 GSVHGTSQSD DIATSTDESN EGMEDEPKTVK EQPSLNRTPE ICFARNAVRF PIEVSVQDER 180 QLKREKRKQA NRESARRSRL RKQAEYEEVV RSYDSLKNKD VNLKSEMKNL KEDMEKLRLE 240 NATLMEKLES TQPRQPEEAV PNDPKAKHED SCKKK 275
3-142	Sequences	
3-142-1	Sequence Number [ID]	142
3-142-2	Molecule Type	AA
3-142-3	Length	340
3-142-4	Features Location/ Qualifiers	source 1..340 mol_type=protein organism=Gossypium barbadense
3-142-5	NonEnglishQualifier Value Residues	MGNDEGKSK SDKSSSPVPT DQANIHVYPD WAAMQAYYGG DGKTEAKSNA VAAEVTTTI 60 SPKPIGTVLS PGMNTNMGWP PQPMPPYGA PYATIIYSHG VYAHPAVPIT AAPVEAPTKS 120 SGNTERGSTK KMKGFDGLAM SIGNGTAEND EGAAEPRLSQ SVETEGSTDG SDGNTTGLFY 180 WTGTLEMIKT ALELRNPSSM NAKSSPTNVP CGVMPPEVWM QNERELKRER RKQSNRESAR 240 RSRLRKQAET EELARKVESL TAENATLRSE INQLTEKSEK LRLLENATLVE GLKNAQLGHT 300 QENITNKNED KEGEMYEKKS GAKLHQLLDA SPRTDAVAAS 340
3-143	Sequences	
3-143-1	Sequence Number [ID]	143
3-143-2	Molecule Type	AA
3-143-3	Length	296
3-143-4	Features Location/ Qualifiers	source 1..296 mol_type=protein organism=Ipomoea nil
3-143-5	NonEnglishQualifier Value Residues	MENKILHRQR DGSVTLSEKH TMGKDEENNS PKCDLGSQPA QQFPVPSSYL STSLPTGHVP 60 QIHMWPPQQP FMPPYGMYP MYPQGIYVA HPAVTLAAVT SSVELVPKPS EAIDQNSAKR 120 LKWHNGLAVP ITNDNFEVNA EGSVHGTSQS DDIATSTDES NEGMEDEPKTV KEQPSLNRT 180 EICFARNAVR FPIEVSVQDE RQLKREKRKQ ANRESARRSR LRKQAEYEEV VRSYDSLND 240 NVNLKSEMKN LKEDMEKLRL ENATLMEKLE STQPRQPEEA VPNDPKAKHE DSCKKK 296
3-144	Sequences	
3-144-1	Sequence Number [ID]	144
3-144-2	Molecule Type	AA
3-144-3	Length	160
3-144-4	Features Location/ Qualifiers	source 1..160 mol_type=protein organism=Camellia sinensis
3-144-5	NonEnglishQualifier Value Residues	MGISED MISP KSEKPCPKSE KASPPKQDQT NIHLYPDWAT IQAYYGPGVP IPPSYLNSAT 60 MACHVPHPHV WAPPQHVMPP YGVPYSAIYS PGVVCTHPTV PIATPVSQA MLSTSSYNMD 120 QGSIKKLKRL DGVTKQVVND NAEVDGEGSA HRASQRHASKV 160

3-145	Sequences	
3-145-1	Sequence Number [ID]	145
3-145-2	Molecule Type	AA
3-145-3	Length	345
3-145-4	Features Location/ Qualifiers	source 1..345 mol_type=protein organism=Populus deltoides
3-145-5	NonEnglishQualifier Value Residues	MGNNEEGKSS ASDKSSPTQQ DQTSIHVYPD WAAIQAYYGS RVALPPYYNS GVASGHAPHP 60 YMWGPPQPM ATYGAPYAAI YSHGGVYAH AIPIGSHPHG PGVLSSPAAH TPLSAETPTK 120 SSGNTDQGLM KKLKGFGLA MSIGNDAES AEGGSRLPQS METEGSSDGS DGNTARGKKR 180 SRDGTPTVVV DPGMTKALEL RNPPSVNVAK TNPATIPQPG AMLPSEAWSP NDRELRERR 240 KQSNRESARR SRLRKQAEAE ELAHKVEMLT TVNMTLKSEI DQFTEKSQKL RLENAALTEK 300 LKNAQLGPTQ EMILSNIDEQ RSSNSGAKLH QLMDASPRAD AVAAG 345
3-146	Sequences	
3-146-1	Sequence Number [ID]	146
3-146-2	Molecule Type	AA
3-146-3	Length	347
3-146-4	Features Location/ Qualifiers	source 1..347 mol_type=protein organism=Populus alba
3-146-5	NonEnglishQualifier Value Residues	MGNNEEGKSS ASDKSSPAQQ DQTSIHVYPD WAAIQAYYGS RVALPPYYNS GVASGHAPHP 60 YMWGPPQPM ATYGAPYAAI YSHGGVYAH AAVPIGSHPH GPGVLSSPAA NTPLSAETPT 120 KSSGNTDQGL LKKLKGFGLA AMSIGNDAE SAEGGSRLPQ SMETEGSSDG SDGNTARGKK 180 RSREGTPTAV VDPGMTKALE LRNPPSVNVA KTNPATIPQP GAVLPSEAWS PNDRELRERR 240 RKQSNRESAR SRLRKQAEA EELAHKVETL TTVNVALKSE IDQFTEKSQK LRLENAALTQ 300 EKLKNAQQGP AQEMTLNNID EQRNSNPQAK LHQLLDASPR ADAVAAG 347
3-147	Sequences	
3-147-1	Sequence Number [ID]	147
3-147-2	Molecule Type	AA
3-147-3	Length	346
3-147-4	Features Location/ Qualifiers	source 1..346 mol_type=protein organism=Populus trichocarpa
3-147-5	NonEnglishQualifier Value Residues	MGNNEEGKSS ASDKSSPAQQ DQTSIHVYPD WAAIQAYYGS RVALPPYYNS GVASGHAPHP 60 YMWGPPQPM ATYGAPYAAI YSHGGVYAH AVPIGSHPHG PGVLSSPAAH TPLSAETPTK 120 SSGNTDQGLM KKLKGFGLA MSIGNDAES AEGGSRLPQS METEGSSDGS DGNTARGKKR 180 SREGTPTVVV DPGMTKALEL RNPPSVNVAK TNPATIPQPG AMLPSEAWSP NDRELRERR 240 KQSNRESARR SRLRKQAEAE ELAHKVETLT TVNMTLKSEI DQFTEKSQKL RLENAALTQE 300 KLKNAQLGPT QEIILNNIDE QRSSNSGAKL HQLMDASPR DAVAAG 346
3-148	Sequences	
3-148-1	Sequence Number [ID]	148
3-148-2	Molecule Type	AA
3-148-3	Length	424
3-148-4	Features Location/ Qualifiers	source 1..424 mol_type=protein organism=Hevea brasiliensis
3-148-5	NonEnglishQualifier Value Residues	MGNNEEGKSS KPDKSTSPAP VDQNNIHVYP DWAAMQAYYG PRVALPPYYN SAMASGHAPH 60 PYMWGPPQPM MPPYGAPYAA VYSHGGVYAH PAVPIGSHPH VPGVPSPPAA ATPVSVETPT 120 KSGNTDQGL MKKLKRFGL AMSIGNVSSN AESAEGGDHE LSQSVETEGS SDGSDGNTAG 180 GRKRSREGKP TIGGEMKTEM QASIVPTGEV NGSSDKVLGA AVASAGVSGK SPGSVVSPGT 240 STALELRNPP LNAAKTSATS IPQPCQVLPS EAWLQNEREL KRERRKQSNR ESARRSRLRK 300 QAETEELAHK VESLTSENLA LKSEMQLTE KSDKLRLENT ALLEKLKNSQ LGRTQEIILN 360 SSEEQRASAV GTENLLSRVN NSASVDRSTR EEGELYERNS NSGAKLHQLL GASPRARADA 420 VAAS 424
3-149	Sequences	
3-149-1	Sequence Number [ID]	149
3-149-2	Molecule Type	AA
3-149-3	Length	353
3-149-4	Features Location/ Qualifiers	source 1..353 mol_type=protein organism=Carica papaya

3-149-5	NonEnglishQualifier Value Residues	MGNSEEGKSS KSDKSSSPAP ADQTNIHVYP DWAAMQAYYS PRVPIPPYYN SAMP SGHPPH 60 PYMWGPQPM PYGAPYAAI YSHGGVY AHP AVSLGSHSQG HGVPLSPATA TPLSIETPTK 120 SSGITDRSLM KKLKGF DGLA MSIGNGNTDS AEGAGEHGLS KSIDTEGSSD GSDGSTSGVD 180 QTKRKRREH TPTAGGKNQG QSGSVPAGDP AYNKVLGTAG GPANIIGNAS GTAETEELAR 240 KVECLTAENV GLKSEIGQLT DKSDKLRREN VALMEKLKNA QLV RTEEMGL NNVDDKRAAP 300 VNTENLLSRV NNSGSHDRST EQEGEVYEKK SNSGAKLHQL LGASPRADAV AAG 353
3-150	Sequences	
3-150-1	Sequence Number [ID]	150
3-150-2	Molecule Type	AA
3-150-3	Length	315
3-150-4	Features Location/ Qualifiers	source 1..315 mol_type=protein organism=Brassica carinata
3-150-5	NonEnglishQualifier Value Residues	MGNSEEPKPT KSDKPSSSPA DQTN AHVYPD WAAMLAYYGP NVAMPPYYNS AMAASGHPPP 60 YMWNHQHMM PYGTPYHGGG VYAHPGIPMA NNFKQSDQGD LLFHNQAGTP LNIDIPTKST 120 GNTDNGLMKK LKEFDGLAMS LGIGNPPENN GDDERKRSRN SSETEDGKRH LVQCSSFRSV 180 SPTSGDNGVK PVQRAIVSPD DERELKRERR KQSNRESARR SRLRKQAE TE ELARKVEALT 240 AENMALRSEL NQLNEKSDKL RANATLLDK LKSSEPEKIV SIVKNSGTGD NESNSTSKLH 300 QLLDTKPRAD AVAAG 315
3-151	Sequences	
3-151-1	Sequence Number [ID]	151
3-151-2	Molecule Type	AA
3-151-3	Length	345
3-151-4	Features Location/ Qualifiers	source 1..345 mol_type=protein organism=Rhododendron griersonianum
3-151-5	NonEnglishQualifier Value Residues	MGSSEGIDTP KPEKASPSKQ DLKMVYPFAL AGTIKHSSLS QLEILFNNLI CFQSYAPEI 60 PIPPSYLNSA IMAGHAPHHD AWAQPQTAPV STQLFLSESM DEGITKMLDR LDRVDMQIVE 120 GTHEVDGEGS INGASQSAEG DASSNRSDSH TSQSQGESTS MGKQNADSDV EVNTQANSTA 180 EGDANVASID LTGITVAPEK PVDTSNGKSE TGITRSASFA VAGLPLKASL QE QKQLKREK 240 RKQANRESAC RSRLKKQALD ELMKRYETVK VENTALKSEI NQLTEEVEKL KAENCALMEK 300 LNGTEVTHPR EMVLDETRPD SSTQPNCSED GFILNNSVSF VIGIN 345
3-152	Sequences	
3-152-1	Sequence Number [ID]	152
3-152-2	Molecule Type	AA
3-152-3	Length	311
3-152-4	Features Location/ Qualifiers	source 1..311 mol_type=protein organism=Populus trichocarpa
3-152-5	NonEnglishQualifier Value Residues	MGNNEEGKSS ASDKSSPAQQ DQTSIHVYPD WAAIQAYYGS RVALPPYYNS GVASGHAPHP 60 YMWGPPQPM ATYGAPYAAI YSHGGVY AHP AVPIGSHPHG PGVLSSPAAH TPLSAETPTK 120 SSGNTDQGLM KKLKGF DGLA MSINGDAES AEGGSRLPQS METEGSSDGS DGNTARGKKR 180 SREGTPTVVV DPGMTKALEL RNPPSVNVAK TNPATIPQPG AMLPSEAWSP NDRELKRERR 240 KQSNRESARR SRLRKQAEAE ELAHKVETLT TVNMTLKSEI DQFTEKSQKL RLENAALTSG 300 SISRRNSRMR N 311
3-153	Sequences	
3-153-1	Sequence Number [ID]	153
3-153-2	Molecule Type	AA
3-153-3	Length	294
3-153-4	Features Location/ Qualifiers	source 1..294 mol_type=protein organism=Chenopodium quinoa
3-153-5	NonEnglishQualifier Value Residues	MGNSEGEKPT RSEKAASPPR EQSSVHVYPD WAAVQAYYGS GALPPPYFNP AATSGHAVPP 60 YMWAPPQMM PYGSPYAMP HGGMYGPSAA PICMQIPAPL SLETPTKSSG KDSGSGKVTK 120 ESDVHAASGN GNANAESAGD AAAGVRGQSL STDCGTEGSC ERSYDSSARV KTHPLLDERE 180 LKREKRKQSN RESARRSRLR KQAEEMEELGK KVDNLNVENL ALRSEINKLV EDSEKLRS EN 240 SSLMDKLKIM QGQSDemdSE RTENQETSHI NTENLLSRVN NPVSNGVHEQ SEEK 294
3-154	Sequences	
3-154-1	Sequence Number [ID]	154
3-154-2	Molecule Type	AA

3-154-3	Length	353
3-154-4	Features Location/ Qualifiers	source 1..353 mol_type=protein organism=Panicum hallii
3-154-5	NonEnglishQualifier Value Residues	MAHDEAVATQ KTGNTTSSPK DQPAPSPYPD WSTMQAYYGP GVLPPITYFAP AIAPGHPPPY 60 MWGPQPLMPH PFGTPYAAMY PHGAAYPHPL VPMVSNPLSV EPTKSANSKE KSSNKKLKEI 120 DQTAVCAGSG NSKRTMSSSE DYSAEGSSDV NDQKVNKTSR KRSSVDGPGA ETSAAARIED 180 VANSRAMGAT ISPSPGVIVP PHTGGPTDLS IKDERELKRE KRKQSNRESA RRSRLRKQAE 240 TEELATQVES LTAENTSLRS EIGRLTESSE KLRLENSALM VKLKDPELPA PAEPSPNKAA 300 ASSTSPRPAA ENFLSMIDST ASAPGVS RHT EHGEPEKLRQL LDSNPSTDVA AVS 353
3-155	Sequences	
3-155-1	Sequence Number [ID]	155
3-155-2	Molecule Type	AA
3-155-3	Length	353
3-155-4	Features Location/ Qualifiers	source 1..353 mol_type=protein organism=Lactuca sativa
3-155-5	NonEnglishQualifier Value Residues	MANCEEAKNC KPEVTSSSPP QEQQEQQQQQ TSHVYPDWGA MQAYYGRRMA IPHYFNSSVT 60 SGHTPPPYMW GPLQHMMPPY AAMYPHAGVY AHPGVTVAAS PMHVDSPA KS SGNSDRGLIK 120 KLKGF DGLAM SIGNENGISH SGETEGSSEG SDGNTTEGAK SGQKRSRGS STSSEIGKTD 180 DQLPSTNENG SLKKVKSITV APPNVTMFQN ERELKRERRK QSNRESARRS RLRKQAE AEE 240 LGRVESLTN ENLTLKSEIN RLTVNSSNLK LQNAKLMEKL KKSQIEQDTE DPRLDKKGLT 300 LSTANLLSRV DNGSGTGTDT DTD TDTDTDT VGSGATLRQL LDASPRADAV AAG 353
3-156	Sequences	
3-156-1	Sequence Number [ID]	156
3-156-2	Molecule Type	AA
3-156-3	Length	429
3-156-4	Features Location/ Qualifiers	source 1..429 mol_type=protein organism=Telopea speciosissima
3-156-5	NonEnglishQualifier Value Residues	MGNNEADTPS KSEKASSPAQ EQTNVHPCPE WANIQAYYGP GVTLP PPYFN SAVTTGHAPP 60 PYMWGPPQPL MPPYGVPIAI YSHGGVY AHP AVPYGSHGHG IPPSPVVSEV MVATPFGVEP 120 PAKSSTSKER GLMKKLKGF D GLAVSIGNDN AENALGGAIH RMSQSGESGT EGSSDGS DGN 180 TGGDKSHRKR NCESTPSTGK DGKFD PAPVG EANVTPGSTM GAAMAPVSVQ GISVGT DPLA 240 SMISGLELTA SPNGNVKPNS GTVPPSPGVV VPARDGLPSE LWIQDERELK RERRKQSNRE 300 SARRSRLRKQ AETEELAVTV ETL SAENIAL RSELNQLTEN SQKLRL ENAA LMEKLKDAHL 360 EVAEDMAPNN MESETAPPDS TVNLLSRVNN LGSVRSEQRD SEPHENSNSS TKFHQLLESS 420 PRADAIAAG 429
3-157	Sequences	
3-157-1	Sequence Number [ID]	157
3-157-2	Molecule Type	AA
3-157-3	Length	307
3-157-4	Features Location/ Qualifiers	source 1..307 mol_type=protein organism=Phalaenopsis equestris
3-157-5	NonEnglishQualifier Value Residues	MLGANKTVTS CQSEKSSLIV QEQLSVHPYL DWASMQAYFG PGVPI PAPIYF SAAVAPGHSP 60 HPCMWNPQPM LAPFGGPLTA IYPRGAYPHP SVIHITSAMN PEVPAILKTD KEKDFVKNNK 120 VLDKLATCIG NSNVKGAGCN ENRSTEREAR EKGINPLKPY QTQEDRELKQ ERRKQSNRES 180 ARRSRLRRQA ETEELAIKVG TLIAENSTLR SEISRLTKSS DKLKLENSAL MEKLRMVQLS 240 REENIISDSV ADEEVPSIMV KNLLSRIDNC NPDSGANQME YETYEDSSGK LHQLLDSNPK 300 SDSLTAS 307
3-158	Sequences	
3-158-1	Sequence Number [ID]	158
3-158-2	Molecule Type	AA
3-158-3	Length	370
3-158-4	Features Location/ Qualifiers	source 1..370 mol_type=protein organism=Rhodamnia argentea
3-158-5	NonEnglishQualifier Value Residues	MGNNEETKSS KSEKPSSPVT PQDQTGIHVV PDWAAMHAYY GPRVALPPYY NSAVSSGHGP 60 HPYMWGPPQP MPPPYGPPYA AIYSHGGVYG HPAIPITPTP LAAETPKKSS PNSDNGLVKK 120

		LKSFEGLAMS IGSADADSAD DASDKRSSQS AYTGGSSDGD QSGADKAGRK RSREGISSIG 180 DGKSEVQAKG AGKVDAASEN VSDANPVSIL QPGAVVPSES WLQNERELKR ERRKQSNRES 240 ARRSRLRKQA ETEELAKKVD ALTADNVALK SEISRLTENS DKLRLNTTL MERLGNARGG 300 EKEVESLGKF NDNGLLSDKT ENLLSRVNNS GAIDRRSEDE GEIYERKSNA GAKLHQLLDS 360 KPRTDAVAAG 370
3-159	Sequences	
3-159-1	Sequence Number [ID]	159
3-159-2	Molecule Type	AA
3-159-3	Length	301
3-159-4	Features Location/ Qualifiers	source 1..301 mol_type=protein organism=Papaver somniferum
3-159-5	NonEnglishQualifier Value Residues	MGSDETGTPS KSEKSSSSSA LTRQEQNIP AYPDWPSLQA YYGAGVPVPP PYFNPALTGH 60 PPHPYIWGPP QGSYGHGIPL SPATGETLVA TPLSIATPVK FPSSQDLGMK KKVSNDTGGG 120 SQSAGTALEV VPSHGTHPAH VPAVLNRDGV TSEHPVLDDH ELKKQRRKQS NRESARRSRL 180 RKQAESEELT LKVESLNVEN MALKSEMRRL AEDSEKLRL NTSLMEKLN KQATEAASNE 240 VEVAQGANTE NFLTRVSDSG SLCMHPHQS KTSTNENANS GTANLHQLLE ASPRTDAVAV 300 G 301
3-160	Sequences	
3-160-1	Sequence Number [ID]	160
3-160-2	Molecule Type	AA
3-160-3	Length	359
3-160-4	Features Location/ Qualifiers	source 1..359 mol_type=protein organism=Panicum virgatum
3-160-5	NonEnglishQualifier Value Residues	MAHDEAVATQ KTGTASSPK DQPAPSPYPD WSSMQAYYGP GVLPPPTYFAP AIAPGHPPPY 60 MWGPQPLMPH PFGTPYAAMY AHGTAYPHPL VPMVSNPLSV EPTKSANSKE KSSNKKLKEI 120 DRTAVSAGSG NSKRAMSSSE DYSAEGSSDV NDQKVNKTSR KLSSVDGPAI LPHHCFPAPV 180 VKTSTTNVAN SRAMGGSISP YPGVIVPHT GGPADLSIKD ERELKREKRK QSNRESARRS 240 RLRKQVETEE LATQVESLTA ENTSRLSEIG RLTESSEQLR LENSALMVKL KDPEVPAPAE 300 PSPKATASSP SPRPTAENFL SMIDGTTAP GASRHEHGK PKLRQLLDSN PSTDVA AIS 359
3-161	Sequences	
3-161-1	Sequence Number [ID]	161
3-161-2	Molecule Type	AA
3-161-3	Length	362
3-161-4	Features Location/ Qualifiers	source 1..362 mol_type=protein organism=Panicum virgatum
3-161-5	NonEnglishQualifier Value Residues	MKLWLPRRLE IQHLLSDYI YLQDQPAPSP YPDWSSMQAY YGPGVLPPTY FAPAIAPGHP 60 PPYMWGPQPL MHPFPGTPYA AMYAHGTAYP HPLVPMVSNP LSVEPTKSAN SKEKSSNKKL 120 KEIDRTAVSA GSGNSKRAMS SSSEDYSAEGS SDVNDQKVNK TSRKLSSVDG PAILPHHCFP 180 APVVKTTTN VANSRAMGGS ISPYPGVIVP PHTGGPADLS IKDERELKRE KRKQSNRESA 240 RRSRLRKQVE TEELATQVES LTAENTSLRS EIGRLTESSE QLRENSALM VKLKDPEVPA 300 PAEPSPKATA SSPSPRPTAE NFLSMIDGTT SAPGASRHL HGKPKLRQLL DSNPSTDVAA 360 IS 362
3-162	Sequences	
3-162-1	Sequence Number [ID]	162
3-162-2	Molecule Type	AA
3-162-3	Length	142
3-162-4	Features Location/ Qualifiers	source 1..142 mol_type=protein organism=Camellia sinensis
3-162-5	NonEnglishQualifier Value Residues	MTMNSENGVR EICEEGMGMI FAEDNEALMS MAYYGPGVPI PPSYLNSTVM ACHVPRSHVW 60 APPQHVMPPY GVPYSIIYSP GVVYAHPTVP IVVTPVNAQM LSTSSDNMDQ GLTKKLKRLD 120 GVAMQVVNDN AEVDGEASAH GA 142
3-163	Sequences	
3-163-1	Sequence Number [ID]	163
3-163-2	Molecule Type	AA
3-163-3	Length	422
3-163-4	Features Location/ Qualifiers	source 1..422 mol_type=protein

3-163-5	NonEnglishQualifier Value Residues	organism=Telopea speciosissima MGNNEADTPS KSEKASSPAQ EQTNVHPCPE WANIQAYYGP GVTLPPTYFN SAVTTGHAPP 60 PYMWGPPQPL MPPYGVPIAI YSHGGVYAHF AVPYGSHGHG IPPSPVVSEV MVATPFGVEP 120 PAKSSTSKER GLMKKLKGF D GLAVSIGNDN AENALGGAIH RMSQSGESGT EGSSDGDKSH 180 RKRNCSTPS TGKDGKFDPA PVGEANVTPG STMGAAMAPV SVQGISVGT D PLASMISGLE 240 LTASPNGNVK PNSGTVPPSP GVVVPARDGL PSELWIQDER ELKRERRKQS NRESARRSRL 300 RKQAETEELA VTVETLSAEN IALRSELNQL TENSQKLRL E NAALMEKLDK AHLEVAEDMA 360 PNNMESETAP PDSTVNLLSR VNNLGSVRSE QRDSEPHENS NSSTKFHQLL ESSPRADAIA 420 AG 422
3-164	Sequences	
3-164-1	Sequence Number [ID]	164
3-164-2	Molecule Type	AA
3-164-3	Length	380
3-164-4	Features Location/ Qualifiers	source 1..380 mol_type=protein organism=Helianthus annuus
3-164-5	NonEnglishQualifier Value Residues	MGNCEETKTC KPEKSSSPPP PPAPEHPQAG VHAYHDWAAM QAYYGPRMAV PPYFNSAVAP 60 GHAPPPYMWG PPQHMMPPYA AMYPHGGVYP HPGVPIASPM CIDSPAKSPG NSDRGLMKKL 120 KGF DGLAMSI GNGNDNGISH SGETEGSSEG SDGNTTEGGQ NNRKRSCEGS PNAPAEVGKT 180 EPLSGQFFPT ESNEASKKVT ALTVTPYNAS DCELKKSPTV MASTTLAMVP GEALLQNERE 240 LKRERRKQSN RESARRSRLR KQAEAEELAI KVEALTSENL NLKSEINRLT DTSERLKLQN 300 AKLMEQLKTT QEQDTE DPR SDPKGLSLST ANLLSRVNNG SSSPSGGEGE VYESNNNQKS 360 GAKLRQLLDA SPRTDAAVAG 380
3-165	Sequences	
3-165-1	Sequence Number [ID]	165
3-165-2	Molecule Type	AA
3-165-3	Length	360
3-165-4	Features Location/ Qualifiers	source 1..360 mol_type=protein organism=Panicum virgatum
3-165-5	NonEnglishQualifier Value Residues	MAHDES VATQ KTGTASSPK DQPAPSPYPD WSTMQAYYGP GVLPPTYFAP AIAPGHPPAY 60 MWGPQPLMPH PFGTPYAAMY PHGAAYPHPL VPMVSNPLSV EPTKSATSKD KSSNKKLKEI 120 DRTPVSAGSG NSKRTMSSSE DYSAEGSSDV NDQKVNKTSR KLSSVDGPAI LPQHCF PAPV 180 IKPSATNV DN SRAMGASIS SPGVIVPPHT GGPTDLSIKD ERELKREKRK QSNRESARRS 240 KLKQAETEE MATQVESLTA ENTS LRSEIG RLTVSSEKLR LENSALMVKL KDPEVPAPTD 300 PSPNKA AASS PSPRPAAENF LSMIDGTTSA PGVSWHTEHG EPKLRQLLDS NPSTDVA AVS 360
3-166	Sequences	
3-166-1	Sequence Number [ID]	166
3-166-2	Molecule Type	AA
3-166-3	Length	429
3-166-4	Features Location/ Qualifiers	source 1..429 mol_type=protein organism=Macadamia integrifolia
3-166-5	NonEnglishQualifier Value Residues	MGNNEAGTPS KSEKASSPAQ EQTIIHAYPE WPSVQAYYGP GVTLPPTYFN SAVATGHAPP 60 PYMWGPPQAL MPPYGAPYAI YSHGGVYGHP AVPYGSHGQG IPPSPVVSEV MVATPLSVEP 120 PAKSSTRKDR SLMKKLKGF D GLAVSIGNAT DKNATGGSIH WMSQSGEYGT EGSSDGS DGN 180 TGGDKSNRKR NCEGTPSTGK DGKFD PAPVG EANVTPGST M GAAMAPVSVP GKSVGTAPLT 240 AMISGLELST SPIGNVKTNS GTVPPSPGVV VPARDGLPSE PWIQDERELK RERRKQSNRE 300 SARRSRLRKQ AETEELAVRV ENLSAENIAL RSEINQLIEN SQKLRL ENAA LMEKLD AHL 360 RVEGD MAPSN IEFETTPPDS TENLLSRVNN LGSVKSEHRD SEPHENSNSR TKFYQLLESN 420 PRADAAVAG 429
3-167	Sequences	
3-167-1	Sequence Number [ID]	167
3-167-2	Molecule Type	AA
3-167-3	Length	582
3-167-4	Features Location/ Qualifiers	source 1..582 mol_type=protein organism=Tetracentron sinense
3-167-5	NonEnglishQualifier Value Residues	MWLLAWRPSP SSSPSLQKGI FSLPFMKEGK IECRRSGNRF WFRLGFLLLQ ILSFVLKIVE 60 SRVFLRK SFL FCVSQLKRRI FLDIPQDL SR FSDLKLQALT FSVWILTMGN NEAGTPSTPE 120

		KASSPAQEHT NIHPYPDWAA IQAYYGPGGT LPPPYFNSTV TSGHAPHPYM WAPSQPLIPP 180 YGVPTYTAIYS HGGVYAHPAV PVVATPLSIE TPAKSSSKKG QGLMKVKVGY NGSVAVSIGN 240 NAENAAAGGSV HGLSQSAECG TEGSNDGSDG NTAGVRIPSY DLCLLSSIFP STLIYLLVKQ 300 KIANPKYGLS SSVRHTKPGG KGFVAVHLQS SLSFDKNGKV NTQASLKPGG EANAASSMAL 360 GVTLAPANVA GKLGPVITP SMTTMELSG SPNVKRRTAA SVPPTAGVVL PPETWQDER 420 ELKRQRRKQS NRESARKSRL RKQTESEEIA MKVESLNVEN MALRSEINRI TENAEKLRLE 480 NALLMVL SLS LSHHTLSIH PILYEKL NNA QLGQAEEMVS DKNEAEALR INTENYLSRV 540 NNSDSVDRNV QQESENLENS NSGTLKHQLL QSSPRADAVA AG 582
3-168	Sequences	
3-168-1	Sequence Number [ID]	168
3-168-2	Molecule Type	AA
3-168-3	Length	613
3-168-4	Features Location/ Qualifiers	source 1..613 mol_type=protein organism=Buddleja alternifolia
3-168-5	NonEnglishQualifier Value Residues	MSAIDENIES SSWQTPENVD YHGRNRILSA LSERLYDVYH TFTSAKELWD ALEADYGIDD 60 AGIVRFNAST FTRYIMVDEK SINEQIHEFQ DLLKPLEEGG SKFSEDYKVS CLIDKLPPSW 120 SNFARGLRHQ QGSLTSLQVF KSLRVEEQHR INVKQEGLKA KANLVEHKRN HNQNRFKPRG 180 NRFKMNHNKWW KKDYRHNNNN NNRFKTLNYK DSKGKEKIEK PCFVCGRTNH MAKDCFYKKT 240 EKFNRSRFTK PNNQEVNMLV NGEPSISHLA EDQYSSLNWA VRIMGTGDEE NPPKPEKSSD 300 QNNIHMYPNW ATMQDYGYPR FAVPPYLSA VASPHAPLPY MWAPPQSVMP PYGAPYAAFY 360 ANGGLYTHPG VPNAGGPLSM DTSVKSSGIT NGGFMKKLKE FDELAIPLRN RNGDNAECGI 420 DNRQSESKET EGSSDGNGNV TGVVSTSSFA GLSVKKRSLQ GSPNSVEDGK AKFPGPLTPN 480 DALLWNEPQL KREKRKQSNR ESAKRSRLRR QVRLLLSEIT SPSNNFAFSF VQAEAKELAT 540 KVQTLTSENM TLKSEMKNLL ESSKKLKLEN ATLMEKLKDA QLAQTEEVNL HKIDDLRLKP 600 VGTANLLARV ARK 613
3-169	Sequences	
3-169-1	Sequence Number [ID]	169
3-169-2	Molecule Type	AA
3-169-3	Length	266
3-169-4	Features Location/ Qualifiers	source 1..266 mol_type=protein organism=Hevea brasiliensis
3-169-5	NonEnglishQualifier Value Residues	MGTGEESTPA RPSKPAPSTQ EIPTTPSYPD WSNMQAYYG AGATPPFFAS TVPSPTPHPY 60 LWGSQHPLIP PYGTAVPYPA LYPAGGVYAH PNMTMTPNESA QTNTELEGKT ADGKDRASAK 120 KSKGTSAGKA GECVKAASGS GNDGSSQSAE SGSNASSDAS DENQQDEREL KRQKRKQSNR 180 ESARRSRLRK QAECEELQAR VENLTNDNHN LKEELQRISG ECEKLSEND SIKEELTRLY 240 GPDAVAKLEQ SNHSSVVQSL GDEGNS 266
3-170	Sequences	
3-170-1	Sequence Number [ID]	170
3-170-2	Molecule Type	AA
3-170-3	Length	484
3-170-4	Features Location/ Qualifiers	source 1..484 mol_type=protein organism=Tetracentron sinense
3-170-5	NonEnglishQualifier Value Residues	MGNNEAETPS KSEKASSPTQ EQANIHPYPD WAAIQAYYGP GVTLPPPYFN SAVTSGHAPH 60 PYMWAPHQTM MPPYGVYPYAA IYSHGGVYAH PAVPLVKQYD ETCIMEDGRK KLIFLAHDAH 120 EEKLDLTIVP VMQVATPLSI ETPAKSSGNK DQGLMKKLKG FNGFSVSGN GNSESAAGGS 180 VGGLSQRFAV LHHSSHLLIP FSGSKLGGEL SSSLTVDMVH NLKLYSIHAE CGTEGSSDGS 240 DGNTAGAYKT RRRKRVCGGTP SIGKDGKVD EASLIPGGEV NAASSMSLGV GKSVGTVITP 300 SMTTALELSC SPNGKAKTAA CVPPIVGAVV PPEAWVQDER ELKRQRRKQS NRESARRSRL 360 RKQAESEELA IKVETLNAEN MTLRSEINQL AENSEKLRLE NAALMEKLKN AQLGQVQEMD 420 SDKFEAEMDP PISTENFLSR VNNNSDSVGR NVQHQSENLE NSKSGTKLHQ LLESSPRADP 480 VAAG 484
3-171	Sequences	
3-171-1	Sequence Number [ID]	171
3-171-2	Molecule Type	AA
3-171-3	Length	424
3-171-4	Features Location/ Qualifiers	source 1..424 mol_type=protein organism=Macadamia integrifolia
	NonEnglishQualifier Value	

3-171-5	Residues	MGNNEAGTPS KSEKASSPAQ EQTIIHAYPE WPSVQAYYGP GVTLPPPYFN SAVATGHAPP 60 PYMWGPPQAL MPPYGAPYAI YSHGGVYGHF AVPYGSHGQG IPPSPVVATP LSVEPPAKSS 120 TRKDRSLMKK LKGFGLAVS IGNATDKNAT GGSIHWMSQS GEYGTGESSD GSDGNTGGDK 180 SNRKRNCEGT PSTGKDGKFD PAPVGEANVT PGSTMGAAMA PVSVPKGSVG TAPLTAMISG 240 LELSTSPIGN VKTNSGTVPP SPGVVVPARD GLPSEPWIQD ERELKRERRK QSNRESARRS 300 RLRKQAETEE LAVRVENLSA ENIALRSEIN QLIENSQKLR LENAALMEKL KDAHLRVEGD 360 MAPSNIEFET TPPDSTENLL SRVNNLGSVK SEHRDSEPHF NSNSRTKFYQ LLESNPRADA 420 VAAG 424
3-172	Sequences	
3-172-1	Sequence Number [ID]	172
3-172-2	Molecule Type	AA
3-172-3	Length	410
3-172-4	Features Location/ Qualifiers	source 1..410 mol_type=protein organism=Nelumbo nucifera
3-172-5	NonEnglishQualifier Value Residues	MGNNEAGTPS KSEKASSPAQ EQTSIHGYPD WAAIQAYYGP GVTIPPPYFN SAVASGHPPH 60 PYMWGPPQPL MPPYGAPYAA IYSHGGVYAH PSVSYGSHAH GPGFPPSPVV SEVMVATPLS 120 VEPPAKSLGS KDRGLMKKLR VGNDNAEGAA EGASGGLSQS AECGTGESSD GSDGNTAGGN 180 KSKGKSSNC MLSTGMDGKV ETQASVMPVA QVNIASSGAA VSLGGKTIGI APSPQMAPAL 240 ELRGSPNGKA KASTATVPSE LWLQDERELK RQRRKQSNRE SARRSRLRKQ AETDELAMKV 300 ESLSAENIAL RTEINRLTEN SEKLKLENAA LMEKLKNAEP GETTEMVSNN TETHEDASIS 360 TENFLSKVNN SGSVSIDDPD DTEIHENSNS GAKLHQLLES SPRTDVTAVG 410
3-173	Sequences	
3-173-1	Sequence Number [ID]	173
3-173-2	Molecule Type	AA
3-173-3	Length	469
3-173-4	Features Location/ Qualifiers	source 1..469 mol_type=protein organism=Tetrahymena sinensis
3-173-5	NonEnglishQualifier Value Residues	MGNNEAGKPS KSKASSPAQ KQTNVHLYPD WAAIQHYYPG GVTLPPPYFN SAVTSSHAPH 60 PYMWVPPQPM IPPYGTPTYAA IYSHGGVYAH PAVPLGLNPL GHGIPSSSAV SEALVATPLS 120 METPTKSSGN TDRGLMKKLR GFDGLAVSIG NGNAESTVGG AGHGMSQRLT SELYERGYSG 180 ECGTGESSDG SDRNTSADQT WRKSCEANPS IGNDSKIISH VCPPIPGGEVN TASSRALGVT 240 VASASIAQGS IGTIPKPSIS TALELSGSPS VKAKTDATSV PPSPVAVVPS DVVWQDEREL 300 KRERRKQSNR ESARRSRLRK QSESEELALK VESLNVENMA LKSEINRLTE DSEKRLANA 360 ALLVLTLSLS LSIYIFLFSF LIDCHYNKYL GKLNKQLGL VEETVTNNID DQGAPPITGE 420 NFLSGVNNSG FVNRIPKQEG GAYENTSNS TKLHQLLESN PRADAVAAG 469
3-174	Sequences	
3-174-1	Sequence Number [ID]	174
3-174-2	Molecule Type	AA
3-174-3	Length	465
3-174-4	Features Location/ Qualifiers	source 1..465 mol_type=protein organism=Tetrahymena speciosissima
3-174-5	NonEnglishQualifier Value Residues	MGNNEADTPS KSEKASSPAQ EQTNVHPCPE WANIQAYYGP GVTLPPPYFN SAVTTGHAPP 60 PYMWGPPQPL MPPYGVPIAI YSHGGVYAH AVPYGSHGQG IPPSPVVSEV MVATPFGVEP 120 PAKSSTSKER GLMKKLGKFD GLAVSIGNDN AENALGGAIH RMSQRFVTQL LFISRNWNIL 180 GLLVSWECNN LCILCPVPYT CGESGTEGSS DSGDNTGGD KSHRKRNCES TPSTGKDGKF 240 DPAPVGEANV TPGSTMGAAM APVSVQGISV GTDPLASMIS GLELTASPNG NVKPNSTGTV 300 PSPGVVVPAR DGLPSELWIQ DERELKRERR KQSNRESARR SRLRKQAETE ELAVTVETLS 360 AENIALRSEL NQLTENSQKL RLENAALMEK LKDAHLEVAE DMAPNNMESE TAPPDSTVNL 420 LSRVNNLGSV RSEQRDSEPH ENSNSSTKFH QLLESSPRAD AIAAG 465
3-175	Sequences	
3-175-1	Sequence Number [ID]	175
3-175-2	Molecule Type	AA
3-175-3	Length	426
3-175-4	Features Location/ Qualifiers	source 1..426 mol_type=protein organism=Catharanthus roseus
3-175-5	NonEnglishQualifier Value Residues	MGSSEETKSS KPEKSSSPAP EQSNVHVYPD WAAMQAYYGP RVAVPPYFSS AVASGHPPH 60 YMWGPPQPM PPYGTPTYAAI YAHGGVYTHP GVPLGSHANA HAGATSPGAT EAIAASPLSI 120

		DTPTKSSANG SQGLMNKLRG FDGLAMSIGN GNTDSADGGT DHGISQSGDT EGSSDGSNGT 180 TSKAGQKNKK RSREGTPAND RERKSLTPSS PSAAVNTNGS SEKAMRASKV PAAATEKVMG 240 AVLSPNMTTA SELRNPSAAN AKTSPAKVSQ SCSSLPGETW LQNERELKRE RRRQSNRESA 300 RRSRLRKQAE TEELAKKVQT LTAENMTLRS EINKLTENSE HLRHESALLD KLKNARVMQA 360 GEMNKYDELH RQPTGTADLL ARVNNSGSTD KSNEEGGGDV FENRNSGTKL HQLLDASPR 420 DAVAAG 426
3-176	Sequences	
3-176-1	Sequence Number [ID]	176
3-176-2	Molecule Type	AA
3-176-3	Length	411
3-176-4	Features Location/ Qualifiers	source 1..411 mol_type=protein organism=Nelumbo nucifera
3-176-5	NonEnglishQualifier Value Residues	MGNNEAATPP KSEKNSSPAQ EQTNIHPYPD WAAMQAYYGP GVTLP PPYFN PAVTSGHAPH 60 PYMWGPPQPL MPPYGAPYAA IYPHGGVYAH PSITYGSHPH GPGFPPSPVV SEVMVATPVS 120 VEPPAKSSGS KDRDLMKKLR GLGNGNAEGA AEGATSGLSP SAECGTGSS DGSDGNTSGG 180 DKTERKRSSD SILSAGSDGK VETPAISMPS AQVNAASSGG AISVAGKTVG IAPSPGMTPI 240 FDFRGLRNGK ANTNTVIAPP ELWVQDEREL KRQRRKQSNR ESARRSRLRK QAETDELVMK 300 VESLTAENIA LKSEISRLTQ SSEKMRLENA ALMEKLDKAE LGQTTEVVAN DTEAREDPPV 360 STENFLSRVN NPASVSIDDR RDTEIHKNSN SGAKLRQLE SNTRTDAVAA G 411
3-177	Sequences	
3-177-1	Sequence Number [ID]	177
3-177-2	Molecule Type	AA
3-177-3	Length	427
3-177-4	Features Location/ Qualifiers	source 1..427 mol_type=protein organism=Nicotiana tomentosiformis
3-177-5	NonEnglishQualifier Value Residues	MGNSEDGKSC KPEKSSSTAP DQSNIHMPYP WAAMQAYYGP RVAVPPYVNS PVAPGQAPHP 60 CMWGPLQPM PPYGIPIYAGI YAHGGVYAH GPVIVSRPQA HVMTSSPAVS QTMDAASLSM 120 DPSAKTSGDT NQGLMSKLGK SDGLGMSIGN CSVDNGDGTG HGPSQSDSGQ TEGSSDGSNI 180 HTAEVGEKSK KRSRETPNT SGDGSRTSR SPQPREVNGA TKKETSIAFN PGNIAEKVVG 240 TVFSPTMTT LELRNPVGTL VKASPTNVSR ISPAVPGEAW LQNEREMKRE KRKQSNRESA 300 RRSRLRKQGE AEELAIRVQS LTSENGLGKS EINNFTENSA KLKLENSALM ERLQNKQRGQ 360 AEEVTLGKIG DKRLQPVSTA DLLARVNNSG PLDRTNKDDE IHENNTSGAK LHQLLDASHR 420 TDAVAAR 427
3-178	Sequences	
3-178-1	Sequence Number [ID]	178
3-178-2	Molecule Type	AA
3-178-3	Length	405
3-178-4	Features Location/ Qualifiers	source 1..405 mol_type=protein organism=Nyssa sinensis
3-178-5	NonEnglishQualifier Value Residues	MGNSEEGKPC KPKKPSSPAP PPDQTNFHY PDWAAMQAYY GPRVAVPPYF NSAVASGHAP 60 HPYMWGPPQA MPPYGAPYA AIYAHGGVYV HPGVPLAATP SSMESPAKSS GNTDRGLMKK 120 LKGYDGLGMS IGNGNADSGE GGLDHGVSQS AETGGSSDGS DGNTLGVCQN GRKRSREGTP 180 AVAGGDDKRE TQNSPVPAGK LNGASNKVMG VSVTPASGAG EVVGTVLSPS MTTALELRNT 240 SGPNPKTST SVPQPCVMPG EAWVQNDRLV KRERRKQSNR ESARRSRLRK QAEAGELAMR 300 VESLAAENTT LKSEINRLTQ NSKKLKLENA TLTEKLKNAQ LGQTKEISLD KVDDKRTQPL 360 STENLLSRVN NFGSADRDE NNSNSGAKFR QLLDTSPTD AVAAG 405
3-179	Sequences	
3-179-1	Sequence Number [ID]	179
3-179-2	Molecule Type	AA
3-179-3	Length	427
3-179-4	Features Location/ Qualifiers	source 1..427 mol_type=protein organism=Nicotiana tabacum
3-179-5	NonEnglishQualifier Value Residues	MGNSEDGKSC KPEKSSSTAP DQSNIHVYPD WAAMQAYYGP RVAVPPYVNS PVAPGQAPHP 60 CMWGPLQPM PPYGIPIYAGI YAHGGVYAH GPVIVSRPQA HVMTSSPAVS QTMDAASLSM 120 DPSAKTSGDT NQGLMSKLGK SDGLGMSIGN CSVDNGDGTG HGPSQSDSGQ TEGSSDGSNI 180 HTAEVGEKSK KRSRETPNT SGDGSRTSR SPQPREVNGA TKKETSIAFN PGNIAEKVVG 240 TVFSPTMTT LELRNPVGTL VKASPTNVSR ISPAVPGEAW LQNEREMKRE KRKQSNRESA 300

		RRSRLRKQGE AEELAIRVQS LTSENGLKLS EINNFTENSA KLKLENSALM ERLQNKQRGQ 360 AEEVTLGKIG DKRLQPVSTA DLLARVNNSG PLDRTNKKDE IHENNTSGAK LHQLLDASHR 420 TDAVAAR 427
3-180	Sequences	
3-180-1	Sequence Number [ID]	180
3-180-2	Molecule Type	AA
3-180-3	Length	347
3-180-4	Features Location/ Qualifiers	source 1..347 mol_type=protein organism=Carica papaya
3-180-5	NonEnglishQualifier Value Residues	MGNKEEGKSS KTDKLSSPTL PDQTNIHMY P DWAAMQAYYS PTVPLPPYYF SAMP SGHPPH 60 PYMWGPRPMM PPYGSPYAAL RNTHEGVYAR PAASLNAVPL RIETPTKSSG IMDCSLMKKL 120 KGF DGLAMSI GNGNANGPEG ACEHGLSKSV DTEGFSDGID GSTLGVDQTK RKRSQEETPT 180 AEGENQAQTG SITAGETTS A SNKVLGTAGG PANIIGNAAG TVLSPGMSTV LTLRNSSTWL 240 RKQAETEELA HRVECLTAEN VGLKSKIGQL TDKSDKLRE NAALMEKLKI AQLVRTVEMG 300 LNNMDDKRAP PVSTEQSEV YEKKSNSGAK LHQLSDIPR ADAVSAG 347
3-181	Sequences	
3-181-1	Sequence Number [ID]	181
3-181-2	Molecule Type	AA
3-181-3	Length	422
3-181-4	Features Location/ Qualifiers	source 1..422 mol_type=protein organism=Morus notabilis
3-181-5	NonEnglishQualifier Value Residues	MGNDEEGKSA KSEKSSAPST PDQTNHTNQT NVHVYPDWAA MQAYYGRVPT LPPYYNSAVA 60 SGHAPHPYMW GPPQAMMPY GAPYAAIYSH GSVYAHPAVP VAAHQGQIPS SPAAAPPLSV 120 ETPTKSSGTT EKELVKKLKG FDGLAMSIGN GNAEAVAEQR ASQSTETDGS SDGSDGNSIG 180 ANQSRKRKR EGTPTTVGDE KTEKQAKRVS TSEINAVSDK TLGATVPPSG IPGKLVGSV 240 SPGMTTAL EL RNSNLSTKTS PTATLNNTCA VLPADSWLHN ERELKRERRK QSNRESARRS 300 RLRKQAETEE LGHKVESLTA ENVALQSELN RLTENSKKLR LENAKLEDKL KNVQLGHTEK 360 ILLNDVDKPI STENLLSRVN NSGSIDRTAE EENDMYEKN NSGAKLHQLL DTSRADAVA 420 AG 422
3-182	Sequences	
3-182-1	Sequence Number [ID]	182
3-182-2	Molecule Type	AA
3-182-3	Length	408
3-182-4	Features Location/ Qualifiers	source 1..408 mol_type=protein organism=Nelumbo nucifera
3-182-5	NonEnglishQualifier Value Residues	MGNNEAGTPS KSEKASSPAQ TSIHGYPDWA AIQAYYGPV TIPPYPFNSA VASGHPPHPY 60 MWGPPQLMP PYGAPYAAIY SHGGVYAHPS VSYGSHAHGP GFPPSPVVSE VMVATPLSVE 120 PPAKSLGSKD RGLMKKL RVG NDNAEGAAEG ASGGLSQSAE CGTEGSSDGS DGNTAGGNKS 180 KGKKSNCML STGMDGKVET QASVMPVAQV NIASSGA AVS LGGKTIGIAP SPGMAPALEL 240 RGSPNGKAKA STATVPSELW LQDERELKRQ R R KQSNRESA RRSRLRKQAE TDELAMKVES 300 LSAENIALRT EINRLTENSE KLKLENAALM EKLKNAEPGE TTEMVSNNTE THEDASISTE 360 NFLSKVNNSG SVSIDDPRDT EI HENSNSGA KLHQLLESSP RTD TVVAG 408
3-183	Sequences	
3-183-1	Sequence Number [ID]	183
3-183-2	Molecule Type	AA
3-183-3	Length	398
3-183-4	Features Location/ Qualifiers	source 1..398 mol_type=protein organism=Vitis riparia
3-183-5	NonEnglishQualifier Value Residues	MKGSEAEMSS KSKDISSAE EQTNIHLLYP DWASIQAYYG SRVPL LAPHF NSAVPGSHFP 60 YPYVWAPSQP LIPPYGLPYT AIYSHGGVHA HPAVPLVATP LSKKAPSRSS VDMQGVRRK 120 FKRLDGLAVP VGNVSTEDDA RGSVYEVQS VKHGIDGSTD GSDGNTGAFL PQRNSGSEGI 180 LSTDNDGNFH RFAGSLSEGE AYAASHKVSV NSVAPTNAVAG KSVRPLNRNE EIHAACVTSS 240 TSAGSPFEVC QQDERQLKRE R R KQANRESA KKSRLRKQAE NEELRMRYET LNEENKALKF 300 EISKLTEHLD KVRLENTALR EKLKNKQQL E LQGEMALYKI EADLILPDSP EEVYKLNDSN 360 SLNQNVQMEC GIQENPNSET KLHQLKSN S RTDAIVAG 398
3-184	Sequences	
3-184-1	Sequence Number [ID]	184

3-184-2	Molecule Type	AA
3-184-3	Length	313
3-184-4	Features Location/ Qualifiers	source 1..313 mol_type=protein organism=Brassica carinata
3-184-5	NonEnglishQualifier Value Residues	MGSKEEGKTT QSDKPAQVQA PPPPEQSNVH VYHHDWAAMQ AYYGPRVAIT PQYYNSNGHA 60 PPYIWGSPSP MMAPYGTYP PFCPPGGVYA HPALQMGSQP QGPASQATPV VATPLNLEAH 120 PANSSGNTDQ GSENDSSNG SDGNTTGGEQ SRKKRSREGS PNNDGKPSSQ IVPLLRDENE 180 KQAVTMGTPV MPTNEKEVKR EKRKQSNRES ARRSRLRKQA ETEELSVKVD ALVAENMTLR 240 SKLGQLNDES EKLRLNQAL LDQLKAQASG KTENLISGVD KNNNSVSGSS SSSKNAEQQL 300 LNVSLRTDSV AAS 313
3-185	Sequences	
3-185-1	Sequence Number [ID]	185
3-185-2	Molecule Type	AA
3-185-3	Length	422
3-185-4	Features Location/ Qualifiers	source 1..422 mol_type=protein organism=Hevea brasiliensis
3-185-5	NonEnglishQualifier Value Residues	MGNKEEGKSS KPDKSSSPAP VDHTNIHVYP DWAAMQAYYG PRVALPPYFN SAVASGQAPH 60 PYMWGPPQPM MPPYGAPYAA VYSHGGVYAH PTVPIGLHPN VSGVPSSPAA ATPVSVETPK 120 KSTGNTDRGL MKKLKRFDGL AMSIGNGNTN AESAEGGDHR LSQSAETQGS SDGSDGNTAE 180 GSKRSCGEMP NVGGEMRTEL QVSSVPTGEV NASSHKVLGA AVASASVSGK SEGTAVSPSM 240 STAYELRNPP VNSAKTSVTN IPQPCQVLPS EAWFQNEREL KRERRKQSNR ESARRSRLRK 300 QAETEELAHK VESLTSENLA LKSEINQLTK KSEKLRLENA SLLEKLKNSH LGHTQEILN 360 SSNEQRASAI STEDLLSGVN NSASVDRNPG EEGELYERSS NSGAKLHQLL GASPRADAVA 420 AS 422
3-186	Sequences	
3-186-1	Sequence Number [ID]	186
3-186-2	Molecule Type	AA
3-186-3	Length	446
3-186-4	Features Location/ Qualifiers	source 1..446 mol_type=protein organism=Vitis vinifera
3-186-5	NonEnglishQualifier Value Residues	MKGSEAEMSS KSDKISSSAE EQTNIHLLYP DWASIQAYYG SGVPLLAPHF NSAVPGSHFP 60 YPYVWAPSQS YSPSDTLVSQ PLIPPYGVYP TAIYSHGGVH AHPAVPLVSL PFVRIQDMK 120 QQVFRLHGGs DKLNVDLTIL TVMQVATPLS KKVPSRSSVD MDQGVRRKKFK RLDGLAVPVG 180 NVSTEGDAGG SVYEVSVSVK HGIDGSTDGS DGNTGAFLPQ RNSGSEGILS TDNDGNFHRF 240 AGSLSEGEAY AASHKVSNS VAPTNVAGKS VRPLHRNEEI HAVCVASSTS AGSPFEVCQQ 300 DERQLKRERR KQANRESAKK SRLRKQAENE ELRMRYETLN EENKALKFEI SKLTEHLDKV 360 RENTALREK LKNKQLELQ GEMAPYKIEA DLILPDSPEE VYKLNDSNSL NQNVQMECGI 420 QENPNSETKL HQLLKSNSTRT DAIVAG 446
3-187	Sequences	
3-187-1	Sequence Number [ID]	187
3-187-2	Molecule Type	AA
3-187-3	Length	425
3-187-4	Features Location/ Qualifiers	source 1..425 mol_type=protein organism=Nicotiana tomentosiformis
3-187-5	NonEnglishQualifier Value Residues	MGNSEDGKSC KPEKSSSTAP DQSNIHMPYD WAAMQAYYGP RVAVPPYVNS PVAPGQAPHP 60 CMWGPLQPM MPPYGIPYAGI YAHGGVYAH GVPIVSRPQA HVMTSSPAVS QTMDAASLSM 120 DPSAKTSGDT NQGLMSKLG SDGLGMSIGN CSVDNGDGTD HGPSQSGQTE GSSDGSNIHT 180 AEVGEKSKKR SRETTPTNTSG DGKSRTSSP QPREVNGATK KETSIAFNPG NIAEKVVGTV 240 FSPTMTTTL LRPVGTLLVK ASPTNVSRIS PAVPGEAWLQ NEREMKREKR KQSNRESARR 300 SRLRKQGEAE ELAIRVQSLT SENLGLKSEI NNFTENSAKL KLENSALMER LQNKQRGQAE 360 EVTLGKIGDK RLQPVSTADL LARVNNSGPL DRTNKDDEIH ENNTSGAKLH QLLDASHRTD 420 AVAAR 425
3-188	Sequences	
3-188-1	Sequence Number [ID]	188
3-188-2	Molecule Type	AA
3-188-3	Length	401

3-188-4	Features Location/ Qualifiers	source 1..401 mol_type=protein organism=Carya illinoensis
3-188-5	NonEnglishQualifier Value Residues	MGNNEDGKSS KSEKSSSPVT PDQANQTNQT NIHVYPDWAA MQAYYGPVRT LPPYYNSTVT 60 SGHAPHPYMW GPPQPMMPY GAPYAAIYPH GGVYAHPAVP LGSHSHGQGV PLSPAAPTPL 120 SIETPTKSGN TDRGLMKKLLK EFDGLAMSIG NGTAESTDGG AEHGLSQSSE TDGSSDGS DG 180 NTARPNQTRR KRSRDGTPTG SESKTEMQAS HTSAKEVNVA PNKTS GATVT PGSASGNLVG 240 SVVSPGMATA LELRNTGNIN TKT TTTT SVPQ SCSVLPSEAW MPNERELKRE RRRQSNRESA 300 RRSRMRKQAE TEELARKVDS LTAENTA IKS EIGRLTENS D KLRLNTKLM EKLKNAQLGQ 360 TEEIILNRIE DKRALPVSTE NLLSRVNNSG AVDRSTEEKS D 401
3-189	Sequences	
3-189-1	Sequence Number [ID]	189
3-189-2	Molecule Type	AA
3-189-3	Length	460
3-189-4	Features Location/ Qualifiers	source 1..460 mol_type=protein organism=Telopea speciosissima
3-189-5	NonEnglishQualifier Value Residues	MGNNEADTPS KSEKASSPAQ EQTNVHPCPE WANIQAYYGP VTLPPPYFN SAVTTGHAPP 60 PYMWGPPQPL MPPYGVPIAI YSHGGVY AHP AVPYGSHGHG IPPSPVVATP FGVEPPAKSS 120 TSKERGLMKK LKGF DGLAVS IGNDNAENAL GGAIHRMSQR FVTQLLFISR NWNILGLLVS 180 WECNNLCILC PVPYTCGESG TEGSSDGS DG NTGGDKSHRK RNCESTPSTG KDGKFD PAPV 240 GEANVTPGST MGAAMAPVSV QGISVGT DPL ASMISGLELT ASPNGNVKPN SGTVPSPSPGV 300 VVPARDGLPS ELWIQDEREL KRERRKQSNR ESARRSRLRK QAETEELAVT VETLSAENIA 360 LRSELNQLTE NSQKLRL ENA ALMEKLKDAH LEVAEDMAPN NMESETAPPD STVNLLSRVN 420 NLGSRVSEQR DSEPHENSNS STKFHQLLES SPRADAIAAG 460
3-190	Sequences	
3-190-1	Sequence Number [ID]	190
3-190-2	Molecule Type	AA
3-190-3	Length	463
3-190-4	Features Location/ Qualifiers	source 1..463 mol_type=protein organism=Telopea speciosissima
3-190-5	NonEnglishQualifier Value Residues	MGNNEADTPS KSEKASSPAQ TNVHPCPEWA NIQAYYGPV TLPPPYFN SA VTTGHAPPPY 60 MWGPPQPLMP PYGVPIAIYS HGGVY AHP AVPYGSHGHG IPPSPVSEVMV ATPFGVEPPA 120 KSSTSKERGL MKKLKGF DGL AVSIGNDAE NALGGAIHRM SQRFTVQLLF ISRNWNILGL 180 LVSWECNNLC ILCVPYTCG ESGTEGSSDG SDGNTGGDKS HRKRNCESTP STGKDGKFD P 240 APVGEANVTP GSTMGAAMAP VSVQGISVGT DPLASMISGL ELTASPNGV KPN SGTVPSPS 300 PGVVVPARDG LPSELWIQDE RELKRERRKQ SNRESARRSR LRKQAETEEL AVTVETLSAE 360 NIALRSELNQ LTENSQKLRL ENAALMEKLK DAHLEVAEDM APNNMESETA PPDSTVNLLS 420 RVNNLGSVRS EQRDSEPHEN SNSSTKFHQL LESSPRADAI AAG 463
3-191	Sequences	
3-191-1	Sequence Number [ID]	191
3-191-2	Molecule Type	AA
3-191-3	Length	425
3-191-4	Features Location/ Qualifiers	source 1..425 mol_type=protein organism=Nicotiana tabacum
3-191-5	NonEnglishQualifier Value Residues	MGNSEDGKSC KPEKSSSTAP DQSNIHVYPD WAAMQAYYGP RVAVPPYVNS PVAPGQAPHP 60 CMWGPLQPMMP PYGIPYAGI YAHGGVY AHP GVPIVSRPQA HVMTSSPAVS QTMDAASLSM 120 DPSAKTSGDT NQGLMSKLG SDGLGMSIGN CSVDNGDGT D HGPSQSGQTE GSSDGSNIHT 180 AEVGEKSKKR SRETTPTNTSG DGKSRTRSSP QPREVNGATK KETSIAFNPG NIAEKVVGTV 240 FSPTMTTTL LRPVGT LVK ASPTNVSRIS PAVPGEAWLQ NEREMKREKR KQSNRESARR 300 SRLRKQGEAE ELAIRVQSLT SENLGLKSEI NNFTENSAKL KLENSALMER LQNKQRGQAE 360 EVTLGKIGDK RLQPVSTADL LARVNNSGPL DRTNKDDEIH ENNTSGAKLH QLLDASHRTD 420 AVAAR 425
3-192	Sequences	
3-192-1	Sequence Number [ID]	192
3-192-2	Molecule Type	AA
3-192-3	Length	429
3-192-4	Features Location/ Qualifiers	source 1..429 mol_type=protein

3-192-5	NonEnglishQualifier Value Residues	organism=Carya illinoensis MGNNEDGKSS KSEKSSSPVT PDQANQTNQT NIHVYPDWAA MQAYYGPRVT LPPYYNSTVT 60 SGHAPHPYMW GPPQPMMPY GAPYAAIYPH GGVYAHPAVP LGSHSHGQGV PLSPAAVTPL 120 SIETPTKSGN TDRGLMKKLLK EFDGLAMSIG NGTAESTDGG AEHGLSQSSE TDGSSDGS DG 180 NTARPNQTRR KRSRDGTPTG SESKTEMQAS HTSAKEVNVA PNKTS GATVT PGSASGNLVG 240 SVVSPGMATA LELRNTGNIN TKTITTSVPQ SCSVLPSEAW MPNERELKRE RRRQSNRESA 300 RRSRMRKQAE TEELARKVDS LTAENTA IKS EIGRLTENS D KLRLNTKLM EKLKNAQLGQ 360 TEEIILNRIE DKRALPVSTE NLLSRVNNSG AVDRSTEEKS DMYENSSNSG AKLHQLLDTS 420 PRADAVAAG 429
3-193	Sequences	
3-193-1	Sequence Number [ID]	193
3-193-2	Molecule Type	AA
3-193-3	Length	429
3-193-4	Features Location/ Qualifiers	source 1..429 mol_type=protein organism=Carya illinoensis
3-193-5	NonEnglishQualifier Value Residues	MGNNEDGKSS KSEKSSSPVT PDQANQTNQT NIHVYPDWAA MQAYYGPRVT LPPYYNSTVT 60 SGHAPHPYMW GPPQPMMPY GAPYAAIYPH GGVYAHPAVP LGSHSHGQGV PLSPAAVTPL 120 SIETPTKSGN TDRGLMKKLLK EFDGLAMSIG NGTAESTDGG AEHGLSQSSE TDGSSDGS DG 180 NTARPNQTRR KRSRDGTPTG SEGKTEMQAS HTSAKEVNVA PNKTS GATVT PGSASGNLVG 240 SVVSPGMATA LELRNTGNIN TKTITTSVPQ SCSVLPSEAW MPNERELKRE RRRQSNRESA 300 RRSRMRKQAE TEELARKVDS LTAENTA IKS EIGRLTENS D KLRLNTKLM EKLKSAQLGQ 360 TEEIILNRIE DKRALPVSTE NLLSRVNNSG AVDRSTEEKS DMYENSSNSG AKLHQLLDTS 420 PRADAVAAG 429
3-194	Sequences	
3-194-1	Sequence Number [ID]	194
3-194-2	Molecule Type	AA
3-194-3	Length	424
3-194-4	Features Location/ Qualifiers	source 1..424 mol_type=protein organism=Jatropha curcas
3-194-5	NonEnglishQualifier Value Residues	MGNNEEGKSS KPDKSASPAP MDQTNIHVYP DWAAMQAYYG PRVALPPYYN SAMASGHAPH 60 PYMWGPPQPM MPPYGTPTYA VYSHGGVYAH PAVPIGSHPS APGVQSSPAV GTPLNEMPT 120 KSTGNTDRGL MKKLKGF DGL AMSIGNGNGS GNAESAEGGD RRLSLSAETE GSSDGS DGNT 180 SGGRKRSREG TPTIAGEMRS ETPASSAPSG EVNAYS DKVS GAAVTPAGFS EKSVGAVNSP 240 RMSTALELRN PPVNGAKTNA TNIPQPYAVL PSEAWIQNDR ELKRERRKQS NRESARRSRL 300 RKQAETEELA HKVESLTSEN LALKSEINQL TEKSEKLKHE NAALLEKLKN SKLGQAQETI 360 LNRSEQRAS AVSTENLLSR VNTSASVDRN AGESELYER NSNSGAKLHQ LLGASPRADA 420 VAAG 424
3-195	Sequences	
3-195-1	Sequence Number [ID]	195
3-195-2	Molecule Type	AA
3-195-3	Length	327
3-195-4	Features Location/ Qualifiers	source 1..327 mol_type=protein organism=Lactuca saligna
3-195-5	NonEnglishQualifier Value Residues	MGNCEETKAC KPEKPSSPPE HQQTSVHAYP DWAAMQAYYG HSMVAPPPYFN SGHPPPPYMW 60 APPQHMMPPY AAMYPHYPHP GVPLASPMSI DSPAKSSGNS ERGLMKKLLG FDGLAMSIGN 120 GISHSGETQA SSEGSDGNTA SRKRSREVSP AAAEAGKISV TLMTSATVAM MPNQAQQQNE 180 RELKRERRKQ SNRESARRSR LRKQAEAEEL AIKVDALTSE NLTLKSEINR LTDNSEKLKL 240 QNAKLTEKLK NVQEENDTED TSVSLSTANL LSRVDSCTSE EGGSGNDVEE EEMVYKSSNN 300 NNQNSNATAK LRQLLDASPR ADAVAAG 327
3-196	Sequences	
3-196-1	Sequence Number [ID]	196
3-196-2	Molecule Type	AA
3-196-3	Length	481
3-196-4	Features Location/ Qualifiers	source 1..481 mol_type=protein organism=Vitis vinifera
	NonEnglishQualifier Value	

3-196-5	Residues	MGKSEAEMSS KSKDISSSAE EQTNIHLLYP DWASIQAYYG SGVPLLLAPHF NSAVPGSHFP 60 YPYVWAPSQS YSPSCTLVSQ PLIPPYGVY TAIYSHGGVH AHPAVPLVSL PFVRIVQDMK 120 QQVFRHLHGS DKLNVDLTIL TVMQVATPLS KKVPSRSSVD MDQGVRRKKFK RLDGLAVPVG 180 NVSTEDDAGG SVYEVSQSVK HGIDGSTDGS DGNTGAFLPQ RNRCSGEGILL LLMQIVSCPY 240 PAHNILLSSC QFWFKNGNNG KENLVSDNDG NFHRFAGSLS EGEAYAASHK VSVNTVAPT 300 VAGKSVRPLP RNEEIIAACV ASSTSAGSPF EVCQQDERQL KRERRKQANR ESAKKSRLRK 360 QAENEELRMR YETLNEENKA LKFEISKLTE HLDKVRLENT ALREKLKNKQ QLELQGEMAP 420 YKIEADLILP DSPEEVYKLN DSNSLNQNVQ MECGIQENPN SETKLHQLLK SNSRTDAIVA 480 G 481
3-197	Sequences	
3-197-1	Sequence Number [ID]	197
3-197-2	Molecule Type	AA
3-197-3	Length	424
3-197-4	Features Location/ Qualifiers	source 1..424 mol_type=protein organism=Prunus armeniaca
3-197-5	NonEnglishQualifier Value Residues	MGNDNGKSA KSEKSSSPLT ADQTNHSNQT NIHVYPDWAA MQAYYGPRVA LPPYYNSAVA 60 SGHAPHPYMW GPPQPMMPY GTPYAAMYSH GYAHPAVPLG SHGQAVPSSP SAATPLNMT 120 PTKSSGTTDR GLMKLKRFD ALAMSIGNSN VGDAEGGAEH SVSLSLGTEG SSEGSDGNTA 180 GENQTRRKRS RERTPATAGD GKTDMPQSPV SAKEVNAAACD KVLGATVAAP SVPGKLVGPV 240 VSPGMTMALE FKNSPNLNSK SSSTSVPQSC AVLPPEAWMH NERELKRERR KQSNRESARR 300 SRLRKQAETE ELARKVDALS AENVALKSEI SRLTENSSTL KLENATLMDK LTNTRVGRME 360 EIMMNIDDKR VQPFSTENLL SRVNNSGSIA RDAEKEGDMY EKNSGAKLHQ LLDASPRADA 420 VAAG 424
3-198	Sequences	
3-198-1	Sequence Number [ID]	198
3-198-2	Molecule Type	AA
3-198-3	Length	391
3-198-4	Features Location/ Qualifiers	source 1..391 mol_type=protein organism=Nelumbo nucifera
3-198-5	NonEnglishQualifier Value Residues	MGNNEAGTPS KSEKASSPAQ EQTSIHGYPD WAAIQAYYGP GVTIPPPYFN SAVASGHPPH 60 PYMWGPPQPL MPPYGAPYAA IYSHGGVYAH PSVSYVATPL SVEPPAKSLG SKDRGLMKKL 120 RVGNDNAEGA AEGASGGLSQ SAECGTEGSS DGSDGNTAGG NKSCKKKSSN CMLSTGMDGK 180 VETQASVMPV AQVNIASSGA AVSLGGKTIG IAPSPGMAPA LELRGSPNGK AKASTATVPS 240 ELWLQDEREL KRQRRKQSNR ESARRSRLRK QAETDELAMK VESLSAENIA LRTEINRLTE 300 NSEKLKLENA ALMEKLKNAE PGETTEMVSN NTETHEDASI STENFLSKVN NSGVSIDDP 360 RDEIHNSEN SGAKLHQLLE SSPRTDTVVA G 391
3-199	Sequences	
3-199-1	Sequence Number [ID]	199
3-199-2	Molecule Type	AA
3-199-3	Length	422
3-199-4	Features Location/ Qualifiers	source 1..422 mol_type=protein organism=Trema orientale
3-199-5	NonEnglishQualifier Value Residues	MGNDGEIKCA KTEKSSPPAT PDQTTHTNQS NIHVYPDWAA MQAYYGPRVT LPPYYNSAVA 60 AGHAPHPYMW GPPQPMMPY GAPYAAIYSP GNVYAHPAVP LGTHGQGVPS SPAAVTPMSV 120 ETPTKSSGNT EKELGKKMKG FDALAMSIGN GNSDSGAEQM LSQSMETEGS SDGSYGNTAG 180 ANQSRKRKR EGTPTTAGDE KTEKQVNPVS AKEISGKTLG ATVSTAGVTG KLVGPVVSPG 240 MTTGLELRNS SNLNPKTSPT SLSQACTVLP SDSWLQNERE LKRERRKQSN RESARRSRLR 300 KQAETEELAR KVESLTAENV TLKSELNRLT ENSEKLRLN AKLVDKLKNA QLGHTEEEIIL 360 NNVDKRTQPI STENLLSRVN NSGSIDRTAE EDNDMYEKNS NSGTLHQLL DASPRADAVA 420 AG 422
3-200	Sequences	
3-200-1	Sequence Number [ID]	200
3-200-2	Molecule Type	AA
3-200-3	Length	442
3-200-4	Features Location/ Qualifiers	source 1..442 mol_type=protein organism=Carya illinoensis
	NonEnglishQualifier Value	

3-200-5	Residues	MGNNEDGKSS KSEKSSSPVT PVMNPQSVVP FFWQDQANQT NQTNIHVYPD WAAMQAYYGP 60 RVTLPPTYNS TVTSGHAPHP YMWGPPQPM PPYGAPYAAI YPHGGVYAH AVPLGSHSHG 120 QGVPLSPA AV TPLSIETPTK SGNTRDGLMK KLKEFDGLAM SIGNGTAEEST DGGAEHGLSQ 180 SSETDGSSDG SDGNTARPNQ TRRKRSRDTG PTGSESKTEM QASHTSAKEV NVAPNKTSGA 240 TVTPGSASGN LVGSVSPGM ATALELRNTG NINTKTTTT VPQSCSVLPS EAWMPNEREL 300 KRERRKQSNR ESARRSRMRK QAETEELARK VDSLTAENTA IKSEIGRLTE NSDKLRLENT 360 KLMEKLKNAQ LGQTEEILN RIEDKRALPV STENLLSRVN NSGAVDRSTE EKSDMYENSS 420 NSGAKLHQLL DTSPRAVA AG 442
3-201	Sequences	
3-201-1	Sequence Number [ID]	201
3-201-2	Molecule Type	AA
3-201-3	Length	235
3-201-4	Features Location/ Qualifiers	source 1..235 mol_type=protein organism=Camellia sinensis
	NonEnglishQualifier Value	
3-201-5	Residues	MYSDSNPCI YDGDKLSRL ENLYTRHVMP PYGVPYSAIY SPGVVCTHPT VPIIATPVSA 60 QMLSTSSYNN DQSGIKLKR LDGVTQKVVN DNAEVDGEGS AHRASQSIDV SSDASDCNTR 120 DERQLKREKR KHANKQSARR SRYITMEELE HALKKYNTGD KKTIKEIIAE VDIDNDGRIN 180 YDEFVAMMRK DNPEVLEDMI CWPNRGVLQL GLKNVLSLAT WNHSOSSCFL DWFVL 235
3-202	Sequences	
3-202-1	Sequence Number [ID]	202
3-202-2	Molecule Type	AA
3-202-3	Length	387
3-202-4	Features Location/ Qualifiers	source 1..387 mol_type=protein organism=Carya illinoensis
	NonEnglishQualifier Value	
3-202-5	Residues	MGNNEDGKSS KSEKSSSPVT PDQANQTNTQ NIHVYPDWAA MQAYYGRVPT LPPYYNSTVT 60 SGHAPHPYMW GPPQPMMPY GAPYAAIYPH GGVYAHPAV LAVTPLSIET PTKSGNTDRG 120 LMKKLKEFDG LAMSIGNTA ESTDGGAEHG LSQSSETDGS SDGSDGNTAR PNQTRRKRSR 180 DGTPTGSESK TEMQASHTSA KEVNVAPNKT SGATVTPGSA SGNLVGSVVS PGMATALELR 240 NTGNINTKTT TTSVPQSCSV LPSEAWMPNE RELKRERRKQ SNRESARRSR MRKQAETEEL 300 ARKVDSLTAE NTAIKSEIGR LTENSDKLRL ENTKLMEKLG NAQLGQTEEI ILNRIEDKRA 360 LPVSTENLLS RVNNSGAVDR STEEKSD 387
3-203	Sequences	
3-203-1	Sequence Number [ID]	203
3-203-2	Molecule Type	AA
3-203-3	Length	424
3-203-4	Features Location/ Qualifiers	source 1..424 mol_type=protein organism=Prunus mume
	NonEnglishQualifier Value	
3-203-5	Residues	MGNDNGKSA KSEKSSSPLT ADQTNHSNTQ NIHVYPDWAA MQAYYGRVPA LPPYYNSAVA 60 SGHAPHPYMW GPPQPMMPY GTPYAAMYSH GYAHPAVPLG SHGQAVPSSP SAATPLNMET 120 PTKSSGTTDR GLMKLKRFD ALAMSIGNSN VGDAEGGAEH SVSLSLGTG SSEGSDGNTT 180 GENQTRRKRS RERTPATAGD GKTDMQSPV SAKEVNAACD KVLGATVAAP SVPGLVGPV 240 VSPGMTMALE FKNSPNLSK SSSTSVPQSC AVLPPEAWMH NERELKRERR KQSNRESARR 300 SRLRKQAETE ELARKVDALS AENVALKSEI SRLTENSSTL KLENATLMDK LTNTRVGRME 360 EIMMNIDDKR VQPFSTENLL SRVNNSGSIA RDAEKEGDMY EKNSGAKLHQ LLDASPRADA 420 VAAG 424
3-204	Sequences	
3-204-1	Sequence Number [ID]	204
3-204-2	Molecule Type	AA
3-204-3	Length	424
3-204-4	Features Location/ Qualifiers	source 1..424 mol_type=protein organism=Prunus yedoensis var. nudiflora
	NonEnglishQualifier Value	
3-204-5	Residues	MGNDNGKSA KSEKSSSPLT ADQTNHSNTQ NIHVYPDWAA MQAYYGRVPA LPPYYNSAVA 60 SGHAPHPYMW GPPQPMMPY GTPYAAMYSH GYAHPAVPLG SHGQAVPSSP SAATPLNMET 120 PTKSSGNTDR GLMKLKRFD ALAMSIGNSN VGDAEGGAEH SVSLSLGTG SSEGSDGNTA 180 GENQTRRKRS REGTPATAGD GKTDMQSPV SAKEVNAASD KVLGATVAAP SVPGLVGPV 240 VSPGMTMALE FKNSPNLSK SSSTSVPQSC AVLPPEAWMH NERELKRERR KQSNRESARR 300 SRLRKQAETE ELARKVDALS AENVALKSEI SRLTKNSSTL KIENATLMDK LKNTRVGRME 360

		EIIMNIDDKR VQPFSTENLL SRVNNSGAIA RDAEKEGDMY EKNSGAKLHQ LLDASPRADA 420 VAAG 424
3-205	Sequences	
3-205-1	Sequence Number [ID]	205
3-205-2	Molecule Type	AA
3-205-3	Length	431
3-205-4	Features Location/ Qualifiers	source 1..431 mol_type=protein organism=Nicotiana tomentosiformis
3-205-5	NonEnglishQualifier Value Residues	MGNSEDGKSC KPEKSSSTAP DQSNIHMPYD WAAMQAYYGP RVAVPPYVNS PVAPGQAPHP 60 CMWGPLQPMMPYGYPIYAGI YAHGGVYAHG GVPIVSRPQA HVMTSSPAVS QTMDAASLSM 120 DPSAKTSGDT NQGLMSKLGK SDGLGMSIGN CSVDNGDGDG HDGSPQRFHVS DSGQTEGSSD 180 GSNIHAEVGEKSKRSRET TPNTSGDGKS RTRSSPQPRE VNGATKKETS IAFNPGNIAE 240 KVVGTVFSPMTTTLRLNP VGTLVKASPT NVSRISPAVP GEAWLQNERE MKREKRKQSN 300 RESARRSRLR KQGEAEELAI RVQSLTSENGLKSEINNFT ENSAKLKLEN SALMERLQNK 360 QRGQAEEVTL GKIGDKRLQP VSTADLLARV NNSGPLDRTN KDDEIHENNT SGAKLHQLLD 420 ASHRTDAVAA R 431
3-206	Sequences	
3-206-1	Sequence Number [ID]	206
3-206-2	Molecule Type	AA
3-206-3	Length	415
3-206-4	Features Location/ Qualifiers	source 1..415 mol_type=protein organism=Carya illinoensis
3-206-5	NonEnglishQualifier Value Residues	MGNNEDEGKSS KSEKSSSPVT PDQANQTNQT NIHVYPDWAA MQAYYGPVRT LPPYYNSTVT 60 SGHAPHPYMW GPPQPMMPY GAPPYAAIYPH GGVYAHPAVP LAVTPLSIET PTKSGNTDRG 120 LMKKLKEFDG LAMSIGNGTA ESTDGGAEHG LSQSSETDGS SDGSDGNTAR PNQTRRKRSR 180 DGTPTGSESK TEMQASHTSA KEVNVAPNKT SGATVTPGSA SGNLVGSVVS PGMATALELR 240 NTGNINTKTT TTSVPQSCSV LPSEAWMPNE RELKRERRKQ SNRESARRSR MRKQAEETEEL 300 ARKVDLSLTAETAIKSEIGRLTENSCLKRL ENTKLMEKLNK NAQLGQTEEI ILNRIEDKRA 360 LPVSTENLLS RVNNSGAVDR STEEKSDMYE NSSNSGAKLH QLLEDTSRAD AVAAG 415
3-207	Sequences	
3-207-1	Sequence Number [ID]	207
3-207-2	Molecule Type	AA
3-207-3	Length	415
3-207-4	Features Location/ Qualifiers	source 1..415 mol_type=protein organism=Carya illinoensis
3-207-5	NonEnglishQualifier Value Residues	MGNNEDEGKSS KSEKSSSPVT PDQANQTNQT NIHVYPDWAA MQAYYGPVRT LPPYYNSTVT 60 SGHAPHPYMW GPPQPMMPY GAPPYAAIYPH GGVYAHPAVP LAVTPLSIET PTKSGNTDRG 120 LMKKLKEFDG LAMSIGNGTA ESTDGGAEHG LSQSSETDGS SDGSDGNTAR PNQTRRKRSR 180 DGTPTGSESK TEMQASHTSA KEVNVAPNKT SGATVTPGSA SGNLVGSVVS PGMATALELR 240 NTGNINTKTI TTSVPQSCSV LPSEAWMPNE RELKRERRKQ SNRESARRSR MRKQAEETEEL 300 ARKVDLSLTAETAIKSEIGRLTENSCLKRL ENTKLMEKLNK SAQLGQTEEI ILNRIEDKRA 360 LPVSTENLLS RVNNSGAVDR STEEKSDMYE NSSNSGAKLH QLLEDTSRAD AVAAG 415
3-208	Sequences	
3-208-1	Sequence Number [ID]	208
3-208-2	Molecule Type	AA
3-208-3	Length	459
3-208-4	Features Location/ Qualifiers	source 1..459 mol_type=protein organism=Telopea speciosissima
3-208-5	NonEnglishQualifier Value Residues	MGNNEADTPS KSEKASSPAQ EQTNVHPCPE WANIQAYYGP GVTLPPPYFN SAVTTGHAPP 60 PYMWGPPQPL MPPYGVPIYI YSHGGVYAHG AVPYGSHGHG IPPSPVATPF GVEPPAKSST 120 SKERGLMKKL KGFDDLAVSI GNDNAENALG GAIHRMSQRF VTQLLFISRN WNILGLLVSW 180 ECNNLCILCP VPYTCGESGT EGSSDGS DGN TGGDKSHRKR NCSTPSTGK DGKFDPAVPG 240 EANVTPGSTM GAAMAPVSVQ GISVGTDLPLA SMISGLELTA SPNGNVKPNS GTVPPSPGVV 300 VPARDGLPSE LWIQDERELK RERRKQSNRE SARRSRLRKQ AETEELAVTV ETLSAENIAL 360 RSELNQLTEN SQLRLLENAA LMEKLDKDAHL EVAEDMAPNN MESETAPPDS TVNLLSRVNN 420 LGSVRSEQRD SEPHENSNS TKFHLLESS PRADIAAG 459
3-209	Sequences	

3-209-1	Sequence Number [ID]	209
3-209-2	Molecule Type	AA
3-209-3	Length	360
3-209-4	Features Location/ Qualifiers	source 1..360 mol_type=protein organism=Carex littledalei
3-209-5	NonEnglishQualifier Value Residues	MGNESTATAS KSDKTSPPPK EQPPIYPYPD WSALQAYYGP GMIPPPYISA PGHAPQPYMW 60 APQPIMPYPYS PFMYPHGGAY SHHPAMPFVP QHMSTESPAK SATEKESKSA MRKHKEVNCF 120 ALSSAANGVT EKASSEESDD RKNVTGVTVV QKRKQSSADE PMSDTKSNKK SSSWKDGEAS 180 GPSKVTSEIG EITVKPLPDL PAYPTASAVS NSTIPSSRNT VTDLWYKDEK EMKREKRKQS 240 NRESARRSRL RKQAETEELT VKVESLTAEN NTLKSEIGRL TQGSEKLGSE NSVLQEKLLK 300 QAGQVMPSSV VVENFLSMID KNANTSANIE AQDNTNNNTT NGKLHQLLDP ARRTDAVAAS 360
3-210	Sequences	
3-210-1	Sequence Number [ID]	210
3-210-2	Molecule Type	DNA
3-210-3	Length	900
3-210-4	Features Location/ Qualifiers	source 1..900 mol_type=other DNA organism=Actinidia eriantha
3-210-5	NonEnglishQualifier Value Residues	atgggaagta gtgaagacgt gaaatctcca aaatcggcaa aagtatcccc atctaaacag 60 gaacatacaa acattcatct ctatcccgaa tggccaaata tacaggcata ttatgggcct 120 ggagtcccg taccacaatc atactgaac tctgcagcag tggccggtca tgttcctcat 180 cctcatttgt ggtccccgcc acagtctttg atgccaccgt atgggttgcc ttactcaaca 240 atttactctc ctggggttgt ttatgctcat cctgcagttc ctctgatga ttacacaggaa 300 accgctgttt tctctgccgc catgatgtct tgccccgcag tgactgtcat ccaggtcgca 360 actcctgtga gtacacaaca gatttccaat tctcagaca tcacggaaca aggctaaga 420 aagaagttga aaagtgtgga tggagttgct atgtgtgctg tcaaggacaa tgttgacgtt 480 gatggtgaag gatcagttca tggatcatca cagagtgcag aaggcggtat tgggtgcttc 540 agtaatgcaa ccgattgtaa ctacggcgat caacggcagc tgaacgagc tagaaggaaa 600 caggctaaca gagaatttgc taggaggtct aggttaagga agcaggctga aactgaggag 660 cttatggaga gatatgaaac cctgaaagtg gagaacattg cccttaaatac tgaaatgaat 720 caattgacgg aaggatcaga aaaactgagg gttgagaatg ctgcattgat ggagaagttg 780 agtgttgacg acgttacaca tcatggagaa atgggtttat atgagattga acctgactta 840 gctcaaccta actgtagtga ggatggattt acgctaaaca agccaggctc tgtcaactga 900
3-211	Sequences	
3-211-1	Sequence Number [ID]	211
3-211-2	Molecule Type	DNA
3-211-3	Length	22
3-211-4	Features Location/ Qualifiers	misc_feature 1..22 note=Synthetic construct source 1..22 mol_type=other DNA organism=synthetic construct
3-211-5	NonEnglishQualifier Value Residues	tgtgctgatg atagaagtg ca 22
3-212	Sequences	
3-212-1	Sequence Number [ID]	212
3-212-2	Molecule Type	DNA
3-212-3	Length	22
3-212-4	Features Location/ Qualifiers	misc_feature 1..22 note=Synthetic construct source 1..22 mol_type=other DNA organism=synthetic construct
3-212-5	NonEnglishQualifier Value Residues	gggcatcatt cgctacttca tg 22
3-213	Sequences	
3-213-1	Sequence Number [ID]	213
3-213-2	Molecule Type	DNA
3-213-3	Length	22
3-213-4	Features Location/ Qualifiers	misc_feature 1..22 note=Synthetic construct source 1..22

		mol_type=other DNA organism=synthetic construct	
3-213-5	NonEnglishQualifier Value Residues	tgaagccatc tcttgtgctg at	22
3-214	Sequences		
3-214-1	Sequence Number [ID]	214	
3-214-2	Molecule Type	DNA	
3-214-3	Length	22	
3-214-4	Features Location/ Qualifiers	misc_feature 1..22 note=Synthetic construct source 1..22 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-214-5	Residues	tgcttcttca cgagattgag ga	22
3-215	Sequences		
3-215-1	Sequence Number [ID]	215	
3-215-2	Molecule Type	DNA	
3-215-3	Length	22	
3-215-4	Features Location/ Qualifiers	misc_feature 1..22 note=Synthetic construct source 1..22 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-215-5	Residues	tgaagccatc tcttgtgctg at	22
3-216	Sequences		
3-216-1	Sequence Number [ID]	216	
3-216-2	Molecule Type	DNA	
3-216-3	Length	22	
3-216-4	Features Location/ Qualifiers	misc_feature 1..22 note=Synthetic construct source 1..22 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-216-5	Residues	aagggcatca ttcgctactt ca	22
3-217	Sequences		
3-217-1	Sequence Number [ID]	217	
3-217-2	Molecule Type	DNA	
3-217-3	Length	22	
3-217-4	Features Location/ Qualifiers	misc_feature 1..22 note=Synthetic construct source 1..22 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-217-5	Residues	tgaagccatc tcttgtgctg at	22
3-218	Sequences		
3-218-1	Sequence Number [ID]	218	
3-218-2	Molecule Type	DNA	
3-218-3	Length	22	
3-218-4	Features Location/ Qualifiers	misc_feature 1..22 note=Synthetic construct source 1..22 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-218-5	Residues	tgcttcttca cgagattgag ga	22
3-219	Sequences		
3-219-1	Sequence Number [ID]	219	
3-219-2	Molecule Type	DNA	
3-219-3	Length	25	
3-219-4	Features Location/ Qualifiers	misc_feature 1..25 note=Synthetic construct	

		source 1..25 mol_type=other DNA organism=synthetic construct	
3-219-5	NonEnglishQualifier Value Residues	atgatagaag tggcagcacg gctga	25
3-220	Sequences		
3-220-1	Sequence Number [ID]	220	
3-220-2	Molecule Type	DNA	
3-220-3	Length	25	
3-220-4	Features Location/ Qualifiers	misc_feature 1..25 note=Synthetic construct source 1..25 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-220-5	Residues	atcattcgct acttcatgag attgt	25
3-221	Sequences		
3-221-1	Sequence Number [ID]	221	
3-221-2	Molecule Type	DNA	
3-221-3	Length	21	
3-221-4	Features Location/ Qualifiers	misc_feature 1..21 note=Synthetic construct source 1..21 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-221-5	Residues	tggaagcggc agcacagctg a	21
3-222	Sequences		
3-222-1	Sequence Number [ID]	222	
3-222-2	Molecule Type	DNA	
3-222-3	Length	21	
3-222-4	Features Location/ Qualifiers	misc_feature 1..21 note=Synthetic construct source 1..21 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-222-5	Residues	acagtgggag ccttggttaa g	21
3-223	Sequences		
3-223-1	Sequence Number [ID]	223	
3-223-2	Molecule Type	DNA	
3-223-3	Length	25	
3-223-4	Features Location/ Qualifiers	misc_feature 1..25 note=Synthetic construct source 1..25 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-223-5	Residues	atgatagaag cggcagcacg gctgg	25
3-224	Sequences		
3-224-1	Sequence Number [ID]	224	
3-224-2	Molecule Type	DNA	
3-224-3	Length	25	
3-224-4	Features Location/ Qualifiers	misc_feature 1..25 note=Synthetic construct source 1..25 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-224-5	Residues	atcattcgct acttcacgag attga	25
3-225	Sequences		
3-225-1	Sequence Number [ID]	225	
3-225-2	Molecule Type	DNA	
3-225-3	Length	21	

3-225-4	Features Location/ Qualifiers	misc_feature 1..21 note=Synthetic construct source 1..21 mol_type=other DNA organism=synthetic construct	
3-225-5	NonEnglishQualifier Value Residues		21
3-226	Sequences		
3-226-1	Sequence Number [ID]	226	
3-226-2	Molecule Type	DNA	
3-226-3	Length	21	
3-226-4	Features Location/ Qualifiers	misc_feature 1..21 note=Synthetic construct source 1..21 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-226-5	Residues	agggaagcggc agcacagctg a	21
3-227	Sequences		
3-227-1	Sequence Number [ID]	227	
3-227-2	Molecule Type	DNA	
3-227-3	Length	22	
3-227-4	Features Location/ Qualifiers	misc_feature 1..22 note=Synthetic construct source 1..22 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-227-5	Residues	ccgacggttg tttccaatta cc	22
3-228	Sequences		
3-228-1	Sequence Number [ID]	228	
3-228-2	Molecule Type	DNA	
3-228-3	Length	22	
3-228-4	Features Location/ Qualifiers	misc_feature 1..22 note=Synthetic construct source 1..22 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-228-5	Residues	ctgaggcaat tacgtccaca ac	22
3-229	Sequences		
3-229-1	Sequence Number [ID]	229	
3-229-2	Molecule Type	DNA	
3-229-3	Length	22	
3-229-4	Features Location/ Qualifiers	misc_feature 1..22 note=Synthetic construct source 1..22 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-229-5	Residues	ccgacggttg tttccaatta cc	22
3-230	Sequences		
3-230-1	Sequence Number [ID]	230	
3-230-2	Molecule Type	DNA	
3-230-3	Length	22	
3-230-4	Features Location/ Qualifiers	misc_feature 1..22 note=Synthetic construct source 1..22 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-230-5	Residues	ctgaggcaat tacgtccaca ac	22
3-231	Sequences		
3-231-1	Sequence Number [ID]	231	
3-231-2	Molecule Type	DNA	

3-231-3	Length	19	
3-231-4	Features Location/ Qualifiers	misc_feature 1..19 note=Synthetic construct source 1..19 mol_type=other DNA organism=synthetic construct	
3-231-5	NonEnglishQualifier Value Residues	aattcactgg gcagaggac	19
3-232	Sequences		
3-232-1	Sequence Number [ID]	232	
3-232-2	Molecule Type	DNA	
3-232-3	Length	20	
3-232-4	Features Location/ Qualifiers	misc_feature 1..20 note=Synthetic construct source 1..20 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-232-5	Residues	catcctcaac tagtaattgg	20
3-233	Sequences		
3-233-1	Sequence Number [ID]	233	
3-233-2	Molecule Type	DNA	
3-233-3	Length	19	
3-233-4	Features Location/ Qualifiers	misc_feature 1..19 note=Synthetic construct source 1..19 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-233-5	Residues	atgccattgc aattcactg	19
3-234	Sequences		
3-234-1	Sequence Number [ID]	234	
3-234-2	Molecule Type	DNA	
3-234-3	Length	20	
3-234-4	Features Location/ Qualifiers	misc_feature 1..20 note=Synthetic construct source 1..20 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-234-5	Residues	tgcagctaca tcctcaacta	20
3-235	Sequences		
3-235-1	Sequence Number [ID]	235	
3-235-2	Molecule Type	DNA	
3-235-3	Length	25	
3-235-4	Features Location/ Qualifiers	misc_feature 1..25 note=Synthetic construct source 1..25 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-235-5	Residues	gggaccacca cagcctatga tgcca	25
3-236	Sequences		
3-236-1	Sequence Number [ID]	236	
3-236-2	Molecule Type	DNA	
3-236-3	Length	25	
3-236-4	Features Location/ Qualifiers	misc_feature 1..25 note=Synthetic construct source 1..25 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-236-5	Residues	tgacattgct agcccatcga aacct	25
3-237	Sequences		
3-237-1	Sequence Number [ID]	237	

3-237-2	Molecule Type	DNA	
3-237-3	Length	23	
3-237-4	Features Location/ Qualifiers	misc_feature 1..23 note=Synthetic construct source 1..23 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-237-5	Residues	tttcgcatcg atcaggttct tca	23
3-238	Sequences		
3-238-1	Sequence Number [ID]	238	
3-238-2	Molecule Type	DNA	
3-238-3	Length	21	
3-238-4	Features Location/ Qualifiers	misc_feature 1..21 note=Synthetic construct source 1..21 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-238-5	Residues	ctcaaacctg aaaagcactt c	21
3-239	Sequences		
3-239-1	Sequence Number [ID]	239	
3-239-2	Molecule Type	DNA	
3-239-3	Length	21	
3-239-4	Features Location/ Qualifiers	misc_feature 1..21 note=Synthetic construct source 1..21 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-239-5	Residues	caagatgcta ctgcagacga g	21
3-240	Sequences		
3-240-1	Sequence Number [ID]	240	
3-240-2	Molecule Type	DNA	
3-240-3	Length	20	
3-240-4	Features Location/ Qualifiers	misc_feature 1..20 note=Synthetic construct source 1..20 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-240-5	Residues	ctggaagttg tggttttggc	20
3-241	Sequences		
3-241-1	Sequence Number [ID]	241	
3-241-2	Molecule Type	DNA	
3-241-3	Length	25	
3-241-4	Features Location/ Qualifiers	misc_feature 1..25 note=Synthetic construct source 1..25 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-241-5	Residues	ccatgtggag tgtgatgcc a gtgtc	25
3-242	Sequences		
3-242-1	Sequence Number [ID]	242	
3-242-2	Molecule Type	DNA	
3-242-3	Length	25	
3-242-4	Features Location/ Qualifiers	misc_feature 1..25 note=Synthetic construct source 1..25 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-242-5	Residues	caaagcagca agttggatgc cgttc	25
3-243	Sequences		

3-243-1	Sequence Number [ID]	243	
3-243-2	Molecule Type	DNA	
3-243-3	Length	23	
3-243-4	Features Location/ Qualifiers	misc_feature 1..23 note=Synthetic construct source 1..23 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-243-5	Residues	gctctggcta gggacaaact gat	23
3-244	Sequences		
3-244-1	Sequence Number [ID]	244	
3-244-2	Molecule Type	DNA	
3-244-3	Length	21	
3-244-4	Features Location/ Qualifiers	misc_feature 1..21 note=Synthetic construct source 1..21 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-244-5	Residues	tggaattcga tcatctcttt g	21
3-245	Sequences		
3-245-1	Sequence Number [ID]	245	
3-245-2	Molecule Type	DNA	
3-245-3	Length	25	
3-245-4	Features Location/ Qualifiers	misc_feature 1..25 note=Synthetic construct source 1..25 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-245-5	Residues	atggtcagca tgaatgagat ggccg	25
3-246	Sequences		
3-246-1	Sequence Number [ID]	246	
3-246-2	Molecule Type	DNA	
3-246-3	Length	25	
3-246-4	Features Location/ Qualifiers	misc_feature 1..25 note=Synthetic construct source 1..25 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-246-5	Residues	gcttctcctt aatcagggtg ttgtc	25
3-247	Sequences		
3-247-1	Sequence Number [ID]	247	
3-247-2	Molecule Type	DNA	
3-247-3	Length	17	
3-247-4	Features Location/ Qualifiers	misc_feature 1..17 note=Synthetic construct source 1..17 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-247-5	Residues	gagagcttct tgcttgc	17
3-248	Sequences		
3-248-1	Sequence Number [ID]	248	
3-248-2	Molecule Type	DNA	
3-248-3	Length	20	
3-248-4	Features Location/ Qualifiers	misc_feature 1..20 note=Synthetic construct source 1..20 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-248-5	Residues	gcaccaagc tgttgaacc	20

3-249	Sequences		
3-249-1	Sequence Number [ID]	249	
3-249-2	Molecule Type	DNA	
3-249-3	Length	25	
3-249-4	Features Location/ Qualifiers	misc_feature 1..25 note=Synthetic construct source 1..25 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-249-5	Residues	atcagagaag ggcgaaggag acaat	25
3-250	Sequences		
3-250-1	Sequence Number [ID]	250	
3-250-2	Molecule Type	DNA	
3-250-3	Length	25	
3-250-4	Features Location/ Qualifiers	misc_feature 1..25 note=Synthetic construct source 1..25 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-250-5	Residues	acgaaccagg tcaacgggcg tctta	25
3-251	Sequences		
3-251-1	Sequence Number [ID]	251	
3-251-2	Molecule Type	DNA	
3-251-3	Length	25	
3-251-4	Features Location/ Qualifiers	misc_feature 1..25 note=Synthetic construct source 1..25 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-251-5	Residues	cggcatcggtt acctactact ccttc	25
3-252	Sequences		
3-252-1	Sequence Number [ID]	252	
3-252-2	Molecule Type	DNA	
3-252-3	Length	25	
3-252-4	Features Location/ Qualifiers	misc_feature 1..25 note=Synthetic construct source 1..25 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-252-5	Residues	atcttcgggc aatggggcgt agcgg	25
3-253	Sequences		
3-253-1	Sequence Number [ID]	253	
3-253-2	Molecule Type	DNA	
3-253-3	Length	25	
3-253-4	Features Location/ Qualifiers	misc_feature 1..25 note=Synthetic construct source 1..25 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-253-5	Residues	tcaacttctt cgacacctct ccgta	25
3-254	Sequences		
3-254-1	Sequence Number [ID]	254	
3-254-2	Molecule Type	DNA	
3-254-3	Length	25	
3-254-4	Features Location/ Qualifiers	misc_feature 1..25 note=Synthetic construct source 1..25 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		

3-254-5	Residues	actctctcgg cactgaaatc aaagc	25
3-255	Sequences		
3-255-1	Sequence Number [ID]	255	
3-255-2	Molecule Type	DNA	
3-255-3	Length	24	
3-255-4	Features Location/ Qualifiers	misc_feature 1..24 note=Synthetic construct source 1..24 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-255-5	Residues	tgagagattc cgttgcccag aagt	24
3-256	Sequences		
3-256-1	Sequence Number [ID]	256	
3-256-2	Molecule Type	DNA	
3-256-3	Length	24	
3-256-4	Features Location/ Qualifiers	misc_feature 1..24 note=Synthetic construct source 1..24 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-256-5	Residues	ttccttactc atgcggtctg cgat	24
3-257	Sequences		
3-257-1	Sequence Number [ID]	257	
3-257-2	Molecule Type	DNA	
3-257-3	Length	20	
3-257-4	Features Location/ Qualifiers	misc_feature 1..20 note=Synthetic construct source 1..20 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-257-5	Residues	gcagcacata attccacagg	20
3-258	Sequences		
3-258-1	Sequence Number [ID]	258	
3-258-2	Molecule Type	DNA	
3-258-3	Length	21	
3-258-4	Features Location/ Qualifiers	misc_feature 1..21 note=Synthetic construct source 1..21 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-258-5	Residues	tttctgagcc cataacagga g	21
3-259	Sequences		
3-259-1	Sequence Number [ID]	259	
3-259-2	Molecule Type	DNA	
3-259-3	Length	39	
3-259-4	Features Location/ Qualifiers	misc_feature 1..39 note=Synthetic construct source 1..39 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-259-5	Residues	gggactctag aggatccatg ttgaagatca agagggttc	39
3-260	Sequences		
3-260-1	Sequence Number [ID]	260	
3-260-2	Molecule Type	DNA	
3-260-3	Length	38	
3-260-4	Features Location/ Qualifiers	misc_feature 1..38 note=Synthetic construct source 1..38 mol_type=other DNA organism=synthetic construct	

3-260-5	NonEnglishQualifier Value Residues	gtggtaccgcg gggatcctca gtgctgaact aggcattc	38
3-261	Sequences		
3-261-1	Sequence Number [ID]	261	
3-261-2	Molecule Type	DNA	
3-261-3	Length	36	
3-261-4	Features Location/ Qualifiers	misc_feature 1..36 note=Synthetic construct source 1..36 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-261-5	Residues	gggactctag aggatccatg tcagcttcag tggatt	36
3-262	Sequences		
3-262-1	Sequence Number [ID]	262	
3-262-2	Molecule Type	DNA	
3-262-3	Length	42	
3-262-4	Features Location/ Qualifiers	misc_feature 1..42 note=Synthetic construct source 1..42 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-262-5	Residues	gtggtaccgcg gggatcctta ttggtgcatt gttgggggtg cc	42
3-263	Sequences		
3-263-1	Sequence Number [ID]	263	
3-263-2	Molecule Type	DNA	
3-263-3	Length	35	
3-263-4	Features Location/ Qualifiers	misc_feature 1..35 note=Synthetic construct source 1..35 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-263-5	Residues	ggactctaga ggatccatgt cagcttcagt ggatt	35
3-264	Sequences		
3-264-1	Sequence Number [ID]	264	
3-264-2	Molecule Type	DNA	
3-264-3	Length	38	
3-264-4	Features Location/ Qualifiers	misc_feature 1..38 note=Synthetic construct source 1..38 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-264-5	Residues	tgttaccgcg gggatccttg tgcattgttg ggggtgcc	38
3-265	Sequences		
3-265-1	Sequence Number [ID]	265	
3-265-2	Molecule Type	DNA	
3-265-3	Length	39	
3-265-4	Features Location/ Qualifiers	misc_feature 1..39 note=Synthetic construct source 1..39 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-265-5	Residues	gggactctag aggatccatg ggaagtagtg aagacgtga	39
3-266	Sequences		
3-266-1	Sequence Number [ID]	266	
3-266-2	Molecule Type	DNA	
3-266-3	Length	34	
3-266-4	Features Location/ Qualifiers	misc_feature 1..34 note=Synthetic construct source 1..34 mol_type=other DNA	

		organism=synthetic construct	
3-266-5	NonEnglishQualifier Value Residues	gtggtaccgc gggatcctca gttgacagag cctg	34
3-267	Sequences		
3-267-1	Sequence Number [ID]	267	
3-267-2	Molecule Type	DNA	
3-267-3	Length	40	
3-267-4	Features Location/ Qualifiers	misc_feature 1..40 note=Synthetic construct source 1..40 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-267-5	Residues	ggactctaga ggatccatgg atggatctgg tggtaagga	40
3-268	Sequences		
3-268-1	Sequence Number [ID]	268	
3-268-2	Molecule Type	DNA	
3-268-3	Length	40	
3-268-4	Features Location/ Qualifiers	misc_feature 1..40 note=Synthetic construct source 1..40 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-268-5	Residues	tggtagccgc ggatccaaga acccgagct gagcaagaa	40
3-269	Sequences		
3-269-1	Sequence Number [ID]	269	
3-269-2	Molecule Type	DNA	
3-269-3	Length	37	
3-269-4	Features Location/ Qualifiers	misc_feature 1..37 note=Synthetic construct source 1..37 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-269-5	Residues	taagggtacc gaattctcac taaagttgga caggaag	37
3-270	Sequences		
3-270-1	Sequence Number [ID]	270	
3-270-2	Molecule Type	DNA	
3-270-3	Length	35	
3-270-4	Features Location/ Qualifiers	misc_feature 1..35 note=Synthetic construct source 1..35 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-270-5	Residues	gctcggtacc ggatccagca gctcagagat ttcca	35
3-271	Sequences		
3-271-1	Sequence Number [ID]	271	
3-271-2	Molecule Type	DNA	
3-271-3	Length	36	
3-271-4	Features Location/ Qualifiers	misc_feature 1..36 note=Synthetic construct source 1..36 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-271-5	Residues	taagggtacc gaattcggac aacatagggg tgggtg	36
3-272	Sequences		
3-272-1	Sequence Number [ID]	272	
3-272-2	Molecule Type	DNA	
3-272-3	Length	36	
3-272-4	Features Location/ Qualifiers	misc_feature 1..36 note=Synthetic construct source 1..36	

		mol_type=other DNA organism=synthetic construct	
3-272-5	NonEnglishQualifier Value Residues	gctcgggtacc ggatccagga gcatttcaag aaactt	36
3-273	Sequences		
3-273-1	Sequence Number [ID]	273	
3-273-2	Molecule Type	DNA	
3-273-3	Length	36	
3-273-4	Features Location/ Qualifiers	misc_feature 1..36 note=Synthetic construct source 1..36 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-273-5	Residues	taagggttacc gaattccagt taggttgagc taagtc	36
3-274	Sequences		
3-274-1	Sequence Number [ID]	274	
3-274-2	Molecule Type	DNA	
3-274-3	Length	36	
3-274-4	Features Location/ Qualifiers	misc_feature 1..36 note=Synthetic construct source 1..36 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-274-5	Residues	gctcgggtacc ggatcctgct gtcaaggaca atgttg	36
3-275	Sequences		
3-275-1	Sequence Number [ID]	275	
3-275-2	Molecule Type	DNA	
3-275-3	Length	46	
3-275-4	Features Location/ Qualifiers	misc_feature 1..46 note=Synthetic construct source 1..46 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-275-5	Residues	ggtctcttgc actccagaag tgttgcattc agtttcagag ctatgc	46
3-276	Sequences		
3-276-1	Sequence Number [ID]	276	
3-276-2	Molecule Type	DNA	
3-276-3	Length	46	
3-276-4	Features Location/ Qualifiers	misc_feature 1..46 note=Synthetic construct source 1..46 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-276-5	Residues	ggtctctaaa ccattcaccc tcttgaaagc atgcaccagc cgggaa	46
3-277	Sequences		
3-277-1	Sequence Number [ID]	277	
3-277-2	Molecule Type	DNA	
3-277-3	Length	50	
3-277-4	Features Location/ Qualifiers	misc_feature 1..50 note=Synthetic construct source 1..50 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-277-5	Residues	ggtctcttgc acgtcaatag tggcgatgct tgtttcagag ctatgctgga	50
3-278	Sequences		
3-278-1	Sequence Number [ID]	278	
3-278-2	Molecule Type	DNA	
3-278-3	Length	50	
3-278-4	Features Location/ Qualifiers	misc_feature 1..50 note=Synthetic construct	

		source 1..50 mol_type=other DNA organism=synthetic construct	
3-278-5	NonEnglishQualifier Value Residues	ggctctctaaa ccaacatcct caactagtaa ttgcaccagc cggaatcga	50
3-279	Sequences		
3-279-1	Sequence Number [ID]	279	
3-279-2	Molecule Type	DNA	
3-279-3	Length	25	
3-279-4	Features Location/ Qualifiers	misc_feature 1..25 note=Synthetic construct source 1..25 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-279-5	Residues	aacccaactt aatttcgggc ctttc	25
3-280	Sequences		
3-280-1	Sequence Number [ID]	280	
3-280-2	Molecule Type	DNA	
3-280-3	Length	25	
3-280-4	Features Location/ Qualifiers	misc_feature 1..25 note=Synthetic construct source 1..25 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-280-5	Residues	acaaactata aaactgtaga gaacg	25
3-281	Sequences		
3-281-1	Sequence Number [ID]	281	
3-281-2	Molecule Type	DNA	
3-281-3	Length	20	
3-281-4	Features Location/ Qualifiers	misc_feature 1..20 note=Synthetic construct source 1..20 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-281-5	Residues	ggctctggaca agtttgggaa	20
3-282	Sequences		
3-282-1	Sequence Number [ID]	282	
3-282-2	Molecule Type	DNA	
3-282-3	Length	20	
3-282-4	Features Location/ Qualifiers	misc_feature 1..20 note=Synthetic construct source 1..20 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-282-5	Residues	ctggcggtccc aacagctgat	20
3-283	Sequences		
3-283-1	Sequence Number [ID]	283	
3-283-2	Molecule Type	DNA	
3-283-3	Length	36	
3-283-4	Features Location/ Qualifiers	misc_feature 1..36 note=Synthetic construct source 1..36 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-283-5	Residues	tttacaatta cggatccatg tcagcttcag tggatt	36
3-284	Sequences		
3-284-1	Sequence Number [ID]	284	
3-284-2	Molecule Type	DNA	
3-284-3	Length	37	

3-284-4	Features Location/ Qualifiers	misc_feature 1..37 note=Synthetic construct source 1..37 mol_type=other DNA organism=synthetic construct	
3-284-5	NonEnglishQualifier Value Residues	cccttgccca tggatccttg gtgcattggtt ggggggtg	37
3-285	Sequences		
3-285-1	Sequence Number [ID]	285	
3-285-2	Molecule Type	DNA	
3-285-3	Length	37	
3-285-4	Features Location/ Qualifiers	misc_feature 1..37 note=Synthetic construct source 1..37 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-285-5	Residues	tttacaatta cggatccatg ggaagtagtg aagacgt	37
3-286	Sequences		
3-286-1	Sequence Number [ID]	286	
3-286-2	Molecule Type	DNA	
3-286-3	Length	34	
3-286-4	Features Location/ Qualifiers	misc_feature 1..34 note=Synthetic construct source 1..34 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-286-5	Residues	cccttgccca tggatccgtt gacagagcct ggct	34
3-287	Sequences		
3-287-1	Sequence Number [ID]	287	
3-287-2	Molecule Type	DNA	
3-287-3	Length	33	
3-287-4	Features Location/ Qualifiers	misc_feature 1..33 note=Synthetic construct source 1..33 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-287-5	Residues	ggaattccga aaagctcaat taaatgaat ata	33
3-288	Sequences		
3-288-1	Sequence Number [ID]	288	
3-288-2	Molecule Type	DNA	
3-288-3	Length	35	
3-288-4	Features Location/ Qualifiers	misc_feature 1..35 note=Synthetic construct source 1..35 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-288-5	Residues	ggggtacccc ctcgaactca gaaaacgcaa aaaca	35
3-289	Sequences		
3-289-1	Sequence Number [ID]	289	
3-289-2	Molecule Type	DNA	
3-289-3	Length	33	
3-289-4	Features Location/ Qualifiers	misc_feature 1..33 note=Synthetic construct source 1..33 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-289-5	Residues	ggaattccat acgactcact atagggcgaa ttg	33
3-290	Sequences		
3-290-1	Sequence Number [ID]	290	
3-290-2	Molecule Type	DNA	

3-290-3	Length	35	
3-290-4	Features Location/ Qualifiers	misc_feature 1..35 note=Synthetic construct source 1..35 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-290-5	Residues	gggggtacccc ctcgaactca gaaaacgcaa aaaca	35
3-291	Sequences		
3-291-1	Sequence Number [ID]	291	
3-291-2	Molecule Type	DNA	
3-291-3	Length	36	
3-291-4	Features Location/ Qualifiers	misc_feature 1..36 note=Synthetic construct source 1..36 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-291-5	Residues	atgcctctcc cgaattcatg tcagcttcag tggatt	36
3-292	Sequences		
3-292-1	Sequence Number [ID]	292	
3-292-2	Molecule Type	DNA	
3-292-3	Length	42	
3-292-4	Features Location/ Qualifiers	misc_feature 1..42 note=Synthetic construct source 1..42 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-292-5	Residues	gtccaaagct tctcgagtta ttggtgcatt gttgggggtg cc	42
3-293	Sequences		
3-293-1	Sequence Number [ID]	293	
3-293-2	Molecule Type	DNA	
3-293-3	Length	36	
3-293-4	Features Location/ Qualifiers	misc_feature 1..36 note=Synthetic construct source 1..36 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-293-5	Residues	atgcctctcc cgaattcatg tcagcttcag tggatt	36
3-294	Sequences		
3-294-1	Sequence Number [ID]	294	
3-294-2	Molecule Type	DNA	
3-294-3	Length	42	
3-294-4	Features Location/ Qualifiers	misc_feature 1..42 note=Synthetic construct source 1..42 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-294-5	Residues	gtccaaagct tctcgagtta ttggtgcatt gttgggggtg cc	42
3-295	Sequences		
3-295-1	Sequence Number [ID]	295	
3-295-2	Molecule Type	DNA	
3-295-3	Length	33	
3-295-4	Features Location/ Qualifiers	misc_feature 1..33 note=Synthetic construct source 1..33 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-295-5	Residues	ggtaccttga aaagctcaat taaatgaat ata	33
3-296	Sequences		
3-296-1	Sequence Number [ID]	296	

3-296-2	Molecule Type	DNA	
3-296-3	Length	32	
3-296-4	Features Location/ Qualifiers	misc_feature 1..32 note=Synthetic construct source 1..32 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-296-5	Residues	aagcttgctg acaccgagct cgaattcaag ct	32
3-297	Sequences		
3-297-1	Sequence Number [ID]	297	
3-297-2	Molecule Type	DNA	
3-297-3	Length	31	
3-297-4	Features Location/ Qualifiers	misc_feature 1..31 note=Synthetic construct source 1..31 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-297-5	Residues	ggtacctatc ccaaaatatt tattcactta g	31
3-298	Sequences		
3-298-1	Sequence Number [ID]	298	
3-298-2	Molecule Type	DNA	
3-298-3	Length	31	
3-298-4	Features Location/ Qualifiers	misc_feature 1..31 note=Synthetic construct source 1..31 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-298-5	Residues	ggtacccgga tcagcatttg ttttcttctt a	31
3-299	Sequences		
3-299-1	Sequence Number [ID]	299	
3-299-2	Molecule Type	DNA	
3-299-3	Length	40	
3-299-4	Features Location/ Qualifiers	misc_feature 1..40 note=Synthetic construct source 1..40 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-299-5	Residues	tatagggcga attggtcctg tcgctgactc gcatgaaata	40
3-300	Sequences		
3-300-1	Sequence Number [ID]	300	
3-300-2	Molecule Type	DNA	
3-300-3	Length	40	
3-300-4	Features Location/ Qualifiers	misc_feature 1..40 note=Synthetic construct source 1..40 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-300-5	Residues	tatagggcga attggtagaa agtgaggtgt tgcgtcaaga	40
3-301	Sequences		
3-301-1	Sequence Number [ID]	301	
3-301-2	Molecule Type	DNA	
3-301-3	Length	42	
3-301-4	Features Location/ Qualifiers	misc_feature 1..42 note=Synthetic construct source 1..42 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-301-5	Residues	agcttgatat cgaattgtcg acaccgagct cgaattcaag ct	42
3-302	Sequences		

3-302-1	Sequence Number [ID]	302	
3-302-2	Molecule Type	DNA	
3-302-3	Length	31	
3-302-4	Features Location/ Qualifiers	misc_feature 1..31 note=Synthetic construct source 1..31 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-302-5	Residues	ggtaccctaa acctaacaca aatgggaagc a	31
3-303	Sequences		
3-303-1	Sequence Number [ID]	303	
3-303-2	Molecule Type	DNA	
3-303-3	Length	31	
3-303-4	Features Location/ Qualifiers	misc_feature 1..31 note=Synthetic construct source 1..31 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-303-5	Residues	aagcttctcg aactcagaaa acgcaaaaac a	31
3-304	Sequences		
3-304-1	Sequence Number [ID]	304	
3-304-2	Molecule Type	DNA	
3-304-3	Length	31	
3-304-4	Features Location/ Qualifiers	misc_feature 1..31 note=Synthetic construct source 1..31 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-304-5	Residues	ggtaccctaa acctaacaca aatgggaagc a	31
3-305	Sequences		
3-305-1	Sequence Number [ID]	305	
3-305-2	Molecule Type	DNA	
3-305-3	Length	31	
3-305-4	Features Location/ Qualifiers	misc_feature 1..31 note=Synthetic construct source 1..31 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-305-5	Residues	ggtacctcca atcatctcac gccatccaag c	31
3-306	Sequences		
3-306-1	Sequence Number [ID]	306	
3-306-2	Molecule Type	DNA	
3-306-3	Length	40	
3-306-4	Features Location/ Qualifiers	misc_feature 1..40 note=Synthetic construct source 1..40 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-306-5	Residues	tatagggcga attggatcat aggggtggtg ggttgtttgg	40
3-307	Sequences		
3-307-1	Sequence Number [ID]	307	
3-307-2	Molecule Type	DNA	
3-307-3	Length	40	
3-307-4	Features Location/ Qualifiers	misc_feature 1..40 note=Synthetic construct source 1..40 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-307-5	Residues	tatagggcga attgggcctc tcctcacctc acccccaaag	40

3-308	Sequences	
3-308-1	Sequence Number [ID]	308
3-308-2	Molecule Type	DNA
3-308-3	Length	40
3-308-4	Features Location/ Qualifiers	misc_feature 1..40 note=Synthetic construct source 1..40 mol_type=other DNA organism=synthetic construct
	NonEnglishQualifier Value	
3-308-5	Residues	agcttgatat cgaatctcga actcagaaaa cgcaaaaaca 40
3-309	Sequences	
3-309-1	Sequence Number [ID]	309
3-309-2	Molecule Type	DNA
3-309-3	Length	48
3-309-4	Features Location/ Qualifiers	misc_feature 1..48 note=Synthetic construct source 1..48 mol_type=other DNA organism=synthetic construct
	NonEnglishQualifier Value	
3-309-5	Residues	cgggggtaccg tgattggacc aagaaactgt tggctcttct tattggag 48
3-310	Sequences	
3-310-1	Sequence Number [ID]	310
3-310-2	Molecule Type	DNA
3-310-3	Length	48
3-310-4	Features Location/ Qualifiers	misc_feature 1..48 note=Synthetic construct source 1..48 mol_type=other DNA organism=synthetic construct
	NonEnglishQualifier Value	
3-310-5	Residues	cccaagctta tagctgggga ttggcacgtg tccggattcg attgcagc 48
3-311	Sequences	
3-311-1	Sequence Number [ID]	311
3-311-2	Molecule Type	DNA
3-311-3	Length	36
3-311-4	Features Location/ Qualifiers	misc_feature 1..36 note=Synthetic construct source 1..36 mol_type=other DNA organism=synthetic construct
	NonEnglishQualifier Value	
3-311-5	Residues	ggaggccagt gaattcatgt cagcttcagt ggattg 36
3-312	Sequences	
3-312-1	Sequence Number [ID]	312
3-312-2	Molecule Type	DNA
3-312-3	Length	36
3-312-4	Features Location/ Qualifiers	misc_feature 1..36 note=Synthetic construct source 1..36 mol_type=other DNA organism=synthetic construct
	NonEnglishQualifier Value	
3-312-5	Residues	cgagctcgat ggatccttgg tgcattgttg ggggtg 36
3-313	Sequences	
3-313-1	Sequence Number [ID]	313
3-313-2	Molecule Type	DNA
3-313-3	Length	36
3-313-4	Features Location/ Qualifiers	misc_feature 1..36 note=Synthetic construct source 1..36 mol_type=other DNA organism=synthetic construct
	NonEnglishQualifier Value	

3-313-5	Residues	catggaggcc gaattcatgt cagcttcagt ggattg	36
3-314	Sequences		
3-314-1	Sequence Number [ID]	314	
3-314-2	Molecule Type	DNA	
3-314-3	Length	36	
3-314-4	Features Location/ Qualifiers	misc_feature 1..36 note=Synthetic construct source 1..36 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-314-5	Residues	gcagggtcgac ggatccttgg tgcattgttg ggggtg	36
3-315	Sequences		
3-315-1	Sequence Number [ID]	315	
3-315-2	Molecule Type	DNA	
3-315-3	Length	36	
3-315-4	Features Location/ Qualifiers	misc_feature 1..36 note=Synthetic construct source 1..36 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-315-5	Residues	ggaggccagt gaattcatgg gaagtagtga agacgt	36
3-316	Sequences		
3-316-1	Sequence Number [ID]	316	
3-316-2	Molecule Type	DNA	
3-316-3	Length	33	
3-316-4	Features Location/ Qualifiers	misc_feature 1..33 note=Synthetic construct source 1..33 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-316-5	Residues	cgagctcgat ggatccgttg acagagcctg gct	33
3-317	Sequences		
3-317-1	Sequence Number [ID]	317	
3-317-2	Molecule Type	DNA	
3-317-3	Length	36	
3-317-4	Features Location/ Qualifiers	misc_feature 1..36 note=Synthetic construct source 1..36 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-317-5	Residues	catggaggcc gaattcatgg gaagtagtga agacgt	36
3-318	Sequences		
3-318-1	Sequence Number [ID]	318	
3-318-2	Molecule Type	DNA	
3-318-3	Length	33	
3-318-4	Features Location/ Qualifiers	misc_feature 1..33 note=Synthetic construct source 1..33 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-318-5	Residues	gcagggtcgac ggatccgttg acagagcctg gct	33
3-319	Sequences		
3-319-1	Sequence Number [ID]	319	
3-319-2	Molecule Type	DNA	
3-319-3	Length	36	
3-319-4	Features Location/ Qualifiers	misc_feature 1..36 note=Synthetic construct source 1..36 mol_type=other DNA organism=synthetic construct	

3-319-5	NonEnglishQualifier Value Residues	ggaggccagt gaattcatgg gaagtagtga agacgt	36
3-320	Sequences		
3-320-1	Sequence Number [ID]	320	
3-320-2	Molecule Type	DNA	
3-320-3	Length	35	
3-320-4	Features Location/ Qualifiers	misc_feature 1..35 note=Synthetic construct source 1..35 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-320-5	Residues	cgagctcgat ggatcccata gcaactccat ccaca	35
3-321	Sequences		
3-321-1	Sequence Number [ID]	321	
3-321-2	Molecule Type	DNA	
3-321-3	Length	37	
3-321-4	Features Location/ Qualifiers	misc_feature 1..37 note=Synthetic construct source 1..37 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-321-5	Residues	cgccactagt ggatccatgt cagcttcagt ggattgg	37
3-322	Sequences		
3-322-1	Sequence Number [ID]	322	
3-322-2	Molecule Type	DNA	
3-322-3	Length	36	
3-322-4	Features Location/ Qualifiers	misc_feature 1..36 note=Synthetic construct source 1..36 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-322-5	Residues	tcccgggagc ggtaccttgg tgcattgttg ggggtg	36
3-323	Sequences		
3-323-1	Sequence Number [ID]	323	
3-323-2	Molecule Type	DNA	
3-323-3	Length	37	
3-323-4	Features Location/ Qualifiers	misc_feature 1..37 note=Synthetic construct source 1..37 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-323-5	Residues	cgccactagt ggatccatgg gaagtagtga agacgtg	37
3-324	Sequences		
3-324-1	Sequence Number [ID]	324	
3-324-2	Molecule Type	DNA	
3-324-3	Length	36	
3-324-4	Features Location/ Qualifiers	misc_feature 1..36 note=Synthetic construct source 1..36 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-324-5	Residues	tcccgggagc ggtacctcag ttgacagagc ctggct	36
3-325	Sequences		
3-325-1	Sequence Number [ID]	325	
3-325-2	Molecule Type	DNA	
3-325-3	Length	36	
3-325-4	Features Location/ Qualifiers	misc_feature 1..36 note=Synthetic construct source 1..36 mol_type=other DNA	

		organism=synthetic construct	
3-325-5	NonEnglishQualifier Value Residues	gacgagctcg gtaccatggg aagtagtgaa gacgtg	36
3-326	Sequences		
3-326-1	Sequence Number [ID]	326	
3-326-2	Molecule Type	DNA	
3-326-3	Length	36	
3-326-4	Features Location/ Qualifiers	misc_feature 1..36 note=Synthetic construct source 1..36 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-326-5	Residues	acgagatctg gtcgacgttg acagagcctg gcttgt	36
3-327	Sequences		
3-327-1	Sequence Number [ID]	327	
3-327-2	Molecule Type	DNA	
3-327-3	Length	36	
3-327-4	Features Location/ Qualifiers	misc_feature 1..36 note=Synthetic construct source 1..36 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-327-5	Residues	gacgagctcg gtaccatgtc agcttcagtg gattgg	36
3-328	Sequences		
3-328-1	Sequence Number [ID]	328	
3-328-2	Molecule Type	DNA	
3-328-3	Length	39	
3-328-4	Features Location/ Qualifiers	misc_feature 1..39 note=Synthetic construct source 1..39 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-328-5	Residues	acgagatctg gtcgacttat tgggtcattg ttgggggtg	39
3-329	Sequences		
3-329-1	Sequence Number [ID]	329	
3-329-2	Molecule Type	DNA	
3-329-3	Length	36	
3-329-4	Features Location/ Qualifiers	misc_feature 1..36 note=Synthetic construct source 1..36 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-329-5	Residues	ggctgatatc ggatccatgt cagcttcagt ggattg	36
3-330	Sequences		
3-330-1	Sequence Number [ID]	330	
3-330-2	Molecule Type	DNA	
3-330-3	Length	36	
3-330-4	Features Location/ Qualifiers	misc_feature 1..36 note=Synthetic construct source 1..36 mol_type=other DNA organism=synthetic construct	
	NonEnglishQualifier Value		
3-330-5	Residues	gtgcggccgc aagcttttgg tgcattgttg ggggtg	36
3-331	Sequences		
3-331-1	Sequence Number [ID]	331	
3-331-2	Molecule Type	DNA	
3-331-3	Length	35	
3-331-4	Features Location/ Qualifiers	misc_feature 1..35 note=Synthetic construct source 1..35	

		mol_type=other DNA organism=synthetic construct	
3-331-5	NonEnglishQualifier Value Residues	ggttccgcgt ggatccatgg gaagtagtga agacg	35
3-332	Sequences		
3-332-1	Sequence Number [ID]	332	
3-332-2	Molecule Type	DNA	
3-332-3	Length	35	
3-332-4	Features Location/ Qualifiers	misc_feature 1..35 note=Synthetic construct source 1..35 mol_type=other DNA organism=synthetic construct	
3-332-5	NonEnglishQualifier Value Residues	cagtcacgat gaattcgttg acagagcctg gcttg	35
3-333	Sequences		
3-333-1	Sequence Number [ID]	333	
3-333-2	Molecule Type	DNA	
3-333-3	Length	30	
3-333-4	Features Location/ Qualifiers	misc_feature 1..30 note=Synthetic construct misc_feature 1 note=5' biotinylation source 1..30 mol_type=other DNA organism=synthetic construct	
3-333-5	NonEnglishQualifier Value Residues	taattaaata gataagaaaa gagaaaaagg	30
3-334	Sequences		
3-334-1	Sequence Number [ID]	334	
3-334-2	Molecule Type	DNA	
3-334-3	Length	30	
3-334-4	Features Location/ Qualifiers	misc_feature 1..30 note=Synthetic construct misc_feature 1 note=5' biotinylation source 1..30 mol_type=other DNA organism=synthetic construct	
3-334-5	NonEnglishQualifier Value Residues	cctttttctc ttttcttadc tatttaatta	30
3-335	Sequences		
3-335-1	Sequence Number [ID]	335	
3-335-2	Molecule Type	DNA	
3-335-3	Length	30	
3-335-4	Features Location/ Qualifiers	misc_feature 1..30 note=Synthetic construct misc_feature 1 note=5' biotinylation source 1..30 mol_type=other DNA organism=synthetic construct	
3-335-5	NonEnglishQualifier Value Residues	taattaaata aaaaaaaaa gagaaaaagg	30
3-336	Sequences		
3-336-1	Sequence Number [ID]	336	
3-336-2	Molecule Type	DNA	
3-336-3	Length	30	
3-336-4	Features Location/ Qualifiers	misc_feature 1..30 note=Synthetic construct misc_feature 1 note=5' biotinylation source 1..30 mol_type=other DNA organism=synthetic construct	

3-336-5	NonEnglishQualifier Value Residues	cctttttctc tttttttttt tatttaatta	30
3-337	Sequences		
3-337-1	Sequence Number [ID]	337	
3-337-2	Molecule Type	DNA	
3-337-3	Length	30	
3-337-4	Features Location/ Qualifiers	misc_feature 1..30 note=Synthetic construct misc_feature 1 note=5' biotinylation source 1..30 mol_type=other DNA organism=synthetic construct	
3-337-5	NonEnglishQualifier Value Residues	aaataagggc aatccttatca tttatgtcaa	30
3-338	Sequences		
3-338-1	Sequence Number [ID]	338	
3-338-2	Molecule Type	DNA	
3-338-3	Length	30	
3-338-4	Features Location/ Qualifiers	misc_feature 1..30 note=Synthetic construct misc_feature 1 note=5' biotinylation source 1..30 mol_type=other DNA organism=synthetic construct	
3-338-5	NonEnglishQualifier Value Residues	ttgacataaa tgataagatt gcccttattt	30
3-339	Sequences		
3-339-1	Sequence Number [ID]	339	
3-339-2	Molecule Type	DNA	
3-339-3	Length	30	
3-339-4	Features Location/ Qualifiers	misc_feature 1..30 note=Synthetic construct misc_feature 1 note=5' biotinylation source 1..30 mol_type=other DNA organism=synthetic construct	
3-339-5	NonEnglishQualifier Value Residues	aaataagggc aaaaaaaaaa tttatgtcaa	30
3-340	Sequences		
3-340-1	Sequence Number [ID]	340	
3-340-2	Molecule Type	DNA	
3-340-3	Length	30	
3-340-4	Features Location/ Qualifiers	misc_feature 1..30 note=Synthetic construct misc_feature 1 note=5' biotinylation source 1..30 mol_type=other DNA organism=synthetic construct	
3-340-5	NonEnglishQualifier Value Residues	ttgacataaa tttttttttt gcccttattt	30
3-341	Sequences		
3-341-1	Sequence Number [ID]	341	
3-341-2	Molecule Type	DNA	
3-341-3	Length	30	
3-341-4	Features Location/ Qualifiers	misc_feature 1..30 note=Synthetic construct source 1..30 mol_type=other DNA organism=synthetic construct	
3-341-5	NonEnglishQualifier Value Residues	taattaaata gataagaaaa gagaaaaagg	30

3-342	Sequences	
3-342-1	Sequence Number [ID]	342
3-342-2	Molecule Type	DNA
3-342-3	Length	30
3-342-4	Features Location/ Qualifiers	misc_feature 1..30 note=Synthetic construct source 1..30 mol_type=other DNA organism=synthetic construct
	NonEnglishQualifier Value	
3-342-5	Residues	cctttttctc ttttcttatac tatttaatta 30
3-343	Sequences	
3-343-1	Sequence Number [ID]	343
3-343-2	Molecule Type	DNA
3-343-3	Length	30
3-343-4	Features Location/ Qualifiers	misc_feature 1..30 note=Synthetic construct source 1..30 mol_type=other DNA organism=synthetic construct
	NonEnglishQualifier Value	
3-343-5	Residues	aaataagggc aatcttatca tttatgtcaa 30
3-344	Sequences	
3-344-1	Sequence Number [ID]	344
3-344-2	Molecule Type	DNA
3-344-3	Length	30
3-344-4	Features Location/ Qualifiers	misc_feature 1..30 note=Synthetic construct source 1..30 mol_type=other DNA organism=synthetic construct
	NonEnglishQualifier Value	
3-344-5	Residues	ttgacataaa tgataagatt gcccttattt 30
3-345	Sequences	
3-345-1	Sequence Number [ID]	345
3-345-2	Molecule Type	DNA
3-345-3	Length	2660
3-345-4	Features Location/ Qualifiers	source 1..2660 mol_type=other DNA organism=Actinidia eriantha
	NonEnglishQualifier Value	
3-345-5	Residues	caattaaat gaatatatat atatttttaa atatcattta tttataaatt atacaattat 60 tattttttta ttaataagta ttttttttaa aaaatgattt aaattattta ttttttaatt 120 aatgttatta aattttaata taaaaataat tgtaatatg agaaaaaat aaaaaataa 180 aaaaaataag ggcaatccta tcatttatgt caagaaatta actctagagg gaatcacctg 240 tgctttaatc tttatcctt tttagcctta ctattcattt aaagagctaa gtaaaatttc 300 caattagtca ttttttttaa taccaaatta gaaatagttt tttgttttta tttggcattt 360 atttttttc ctttattgaa aaaaaatcat aattagggga agccaatgt acctaaattt 420 tgtaattttc ttttctttt tctatttcaa ccacttttaa ggaaaagatt tcaaaaccca 480 tatttgcaa catttgacgg aattgacaaa aattttccat ctaaaaata ctctcttcat 540 cattgtttt tgtcagttt tttattttta gatatccaa aatatttatt cacttagaaa 600 attaactgat aaaataatgc aaactctctt tttgtcctt ctcttttgaa ttgacgtggc 660 acatttatct ttttaattt agataattt gaattattga aaaaaaata aactgttttc 720 caaataatag ttttttagaa ataatgcaa taatagttt ttaactgttt tccaaatata 780 ttttttccaa aaataatgca ttattaggaa taattaaaa atatcttcaa atattaaaa 840 atatattttt gataaaattt aataatatat aaaaaataat gaatgtttt gttcgcataa 900 acaatttctg atataaaata ttttgcggaa aatgttttg ttcgcataaa caatttctga 960 taatataatt tgcaaaacta aacctaacac aaatgggtag catttttgct tctttaaaat 1020 cttggattcc cttaaattaga caaataaatt gggacggatc agcatttggt ttcttcttaa 1080 ttatttctca acgctacaac aaaataatta aatagataag aaaagagaaa aaggaaactt 1140 agaaccacc caacttttaa acattgcagt tgggtccttc cgtacgttac agtggctctc 1200 cacaacgtcc acatgaacca catgggcgtg gtaatacaa cgcacccac tctctctctc 1260 tctctctcaa taattgtctc cattctcagt aaaattacca aggccactcg tcccacagtg 1320 cacaccacgg cggatccaca gccacactca ccaatcacct ctctctctct ctctctagaa 1380 ttattttgtt gctcttggag caacacgtca ctttttgaca cgtggtggtc ggatccaatc 1440

		atctcacgcc atccaagcac tcagtttcat gtgttttgcca cgtcaccaca acaattccac 1500 cacaaaccca ggtaaacaca agactaacag aacctcactc cgttaatgcc atcttctgt 1560 cgctgactcg catgaaatac caccactttt ggaaaccaa cgccagaaaa attactctca 1620 ccaatattct ctatgaacaa agaaattggg ttattattta ttatttaca gaaataaatg 1680 gcaaaccaca atttaaaaag acgtctctgc agcgattttc acctcatttt attttctgag 1740 cttttaggtg tctcatccga aaccgacgcc ttcttattac gcaacttttc actcttcttt 1800 gccttctcag tccgaaatg actattttca gacaacatca tagggtgatt ggggtgttta 1860 gctatgtag tacgaaatct aaaaatttga atttggaag tttatgaata tttcatcgca 1920 tcgagtactg gcggaatgtt cacgggggta acaggatttg aactccgtta tttcttcttt 1980 gagtaaacgg acgtggctga atacacggac aaccaattaa atgggtgatg atgtttcgtg 2040 tggagacca cgcgtagaaa gtgaggtgtt gcgtcaagag catcaataa tttctcctct 2100 ctctctctct ctctctctcc ctcttctcta tctatatatc cccaatctg gcctctcctc 2160 acctacccc caaagtctac acagaatcaa cccttcatct cccgcatagg cctcccaaac 2220 ccacctcttc tccacaatcc agacacacct cgagtggccg gagtttagag agagagagag 2280 agagagagag agagattttc tgcttcgac ggggggtaaa acccggtgtt tgacaagtgt 2340 tagacatcac ggctatatc ggagtttctc gtgccctcgt acatgtccgg tctgtacgac 2400 gcaaggggtg ttagtgcgag agcaaccctt cgccgcacgg cggacgtggc gccttgccgt 2460 ccgaaggcgg tagcccctct gatctctctt tcctcgcggg tggcgggtcac ttcgctttct 2520 ccgtctacta gcttaatagg ttattcttta gttagttagt aaatacgtct tattatagtt 2580 cgtaagatta tcaagatct gtgtcttgat tcgtctttcg ttgctcgagt cttggtgttt 2640 ttgcgttttc tgagttcgag 2660
3-346	Sequences	
3-346-1	Sequence Number [ID]	346
3-346-2	Molecule Type	DNA
3-346-3	Length	2520
3-346-4	Features Location/ Qualifiers	source 1..2520 mol_type=other DNA organism=Actinidia rufa
	NonEnglishQualifier Value	
3-346-5	Residues	aattagaaat agcttttctt ttttatttgg catttatttt tttccattta ttgaaaaaaa 60 tcataaattag ggaaaaacca atgtacctaa attatgtaat tttctttttc tttttccatt 120 tcaaccactt ttaaggaaaa gatttcaaaa cccatatttg caaacatttg acggaattga 180 caaagatttc ccatctaaaa aatactacct tcatcattgt ctcttgtag tttttatttt 240 tagatatccc aaaatatttg ttcacttaaa aaattaattg ataaaaataat gcaaacctctc 300 tttttggtct ttctcttttg aattgacgtg acacatttat ctttttaatt ttagataatt 360 tcgaattatt gaaaaaaaag ttaaaactgtt ttccaaataa tagtttttta aaaaatttat 420 ttggtttata ttttttatt tttttgtata attttttttc aaaaataact tattattatg 480 aataaacta aaaaaatata ttcaaaataa aaaaattatt tttgataaaa ttaataata 540 tataaaaaata aatgaatgta ttgttcgca taaacaattt ctgataaaat aatttgcaaa 600 actaaaccta acacaaatgg gaagcatttt tgcttcttta aaatcttgga ttccctaaat 660 tagacaaata aattaggacg gatcaacatt ttttttttt taattatttc tcaacgcgcc 720 aaaaaaataa ttatatagat aagaaaagag aaaaaggaaac ttgagagtgt agaaccacc 780 cgacttttaa atattgcagt ttggctcttc cgtacgttac agtggctctc cataccgtcc 840 acatgaacca tacgggcgta gtaaatataa cgaacccac tctctctctc tactcactc 900 tctctctctc tctctctcaa taattgtctc cattcgcagt aaaattacca aggccactca 960 tcccacagt cacaccacgg ccgatccaca gccacactca ccaatctcct ctctctctct 1020 ctctagaatt tttttgttgc tcttgatca acacgtcact ttttgacacg tgggtgctcg 1080 atccaatcat ctacgcat ccaagcactc tgtttcatgt gtttgccacg tcaccacaac 1140 cattccacca caaacccaat taaacacaag actaacagaa cctcactccg ttaatgccat 1200 cttccagccg ctgactcgga tgaaatacca ccacttttgg aaactaaaac gtcaggaaaag 1260 attactctca tcaaatattc tctgtgaaca aagaacttgg gttattattt cttatttaca 1320 agaaatgatg gcaaaccaca attcaaaaag gacgtctctg cagcgatttc cacctcatct 1380 cctccctgaa agttctctgt gcctccattt cggggccttt aggcactctg tccggaaccg 1440 ccaatttctt attacgcaac ttttactct tctctgcctt ctcggtccca aaatgactat 1500 tttcagacaa catcataggg tggttgggtt gtttgccat ataggtagca aatctaaaaa 1560 tttgaatttg gaaaggttat gaatatttca tcgcatcgag tactggcgga atgttcacgg 1620 ggtaaacagt atcgagtaat gctatgggta cattcaattt gcactctaaa tgcacacttc 1680 agtgcattta gatggctgg gtgttgctg aaaggattga gtggacccta caccagcca 1740 atctaaatgc accatctcat caggtagtga ttttagagtgc aagttaggtg taccataac 1800 attttcttaa cacgatttga actccgttat ttctttcttg agtaaacgga cgtggctgaa 1860 tagacggaca accaattaaa tgggtgatga tgtttcgtgt ggagcaccac gcgtagaaaag 1920 tgagggtgtt cgtaagagc atcaaaataa ttctcctctc tctctctctc cctccctctt 1980 ctctatctat aaatcccca atctggcctc tcctcacctc acccccaaag tctacacaga 2040 atcaaccctt catctccgc acaggcctcc aaaaccact tcttctccac aatccagagt 2100 cccagacaca cctcgagtgg ccggagtta gagtgagaga gagagagaga gagagatttt 2160

		ctgcttcgat cgggggtaaa acccgggtgtt tgacaagttg tagacatcac ggctatattc 2220 ggagttttctc gtgccctcat acatgtccgg tctgtacgac gcaaggggtt tagtagtcgag 2280 agaaaccctt cgccgcacgg cggacgtggc gccttgccgt ccgaaggcgg tagccctct 2340 gatctcctat tcctcgccgg cggcgggtcac ttcgctttct ccgtctacta gcttaatagg 2400 tttattctta gttagtgaat aaatacgtct tattatagtt cgtaagatta tcgaagatct 2460 gttgcttgat tcgtctttcg tagctcgagt cttggtgttt ttgcgttttc tgagttcgag 2520
3-347	Sequences	
3-347-1	Sequence Number [ID]	347
3-347-2	Molecule Type	DNA
3-347-3	Length	1551
3-347-4	Features Location/ Qualifiers	source 1..1551 mol_type=other DNA organism=Actinidia eriantha
3-347-5	NonEnglishQualifier Value Residues	atgttgaaga tcaagagggt tccgactgtt gtttccaatt tccaaaagga tgaggccgac 60 gacggcgctc gatctggcgg tggttgcggc cgaactgtcc tccagaagtg ttgatttcaa 120 ggggcaaaagc tacctctgta tgctttcaag aggggtgaatg aggttgttgg tgaaaagggt 180 gtgcttgccc tcgacaacga agaggctcct gtgctttct tggattcact tctcctcggg 240 gagtgggagg atcgtgtgca gagaggactc tttcgttacg atgtcactgc ttgcgaaacc 300 aaggttattc cgggagagta tggcttcatt gcgcagctga acgagggtcg tcacctaag 360 aagaggccaa ctgagtttcg tgttgataag gtcctgcagc ccttcgattg gagcaaatc 420 aacttacta aagtggaca ggaagaggtt ctgttccagt ttgaagcaag caacgacaac 480 gaagtccagt tcttcccaaa tgcacctgtt gatgttgaga attctcccag tgttgtggcc 540 atcaatgtta gtcctattga atatggtcac gtacttctca tcccttcgat tcttgaatgc 600 ctgcctcaaa ggattgacag ggagagcttc ttgcttgctc ttcacatggc agcagaagct 660 ggaaaccctg acttccgatt gggttacaac agcttgggtg catttgccac tatcaatcac 720 cttcaatttc aggttatta cttagccgtg ccttcccta tcgagaaggc tcccactagg 780 aagataacta ctctgaatgg caacaacata ccttacaatg tgcttatctc cgattctgga 840 aagcgtatct tcttttacc acagtgttac gctgagaaac aggtcttgg agaagtgaat 900 tctgagcttc tggacacaca agtgaaccgg gcagtgtggg aaatcagcgg acatatggtt 960 ttgaagagga aggaggacta tgaggaggcg tctgaaggaa atgcttggag gctccttgct 1020 gaggtctccc ttcggagga gaggttcgaa gaagtcaagg cattgatctt tgaagccatc 1080 tctgtgctg atgatggaag cggcagcaca gctgagaact tgcttgagga gccagatgac 1140 aatcctcaat ctggtgaaga agcaaatgat gcccttaaca aagggtccca ctgtggtatg 1200 gtttgcgttg ttgttgctgc tctgtgctg gctgaagttg cttctgttat ggcgacttgc 1260 tttcgtaatc gcttttactt gtccttctg ggcacaccac caccagaact aatcggcgta 1320 gtggcatttc cgacacataa gaaaatgaag aaaggagtgc accctcagct gcaatggata 1380 tcctatgtta ccagagtggt ccgattgatg aatatcatga tgaccaagat acaccatggt 1440 ggtaaagttt atcacattaa ggcaaggcga cagatggctc aaagcatcgg gcagattgcc 1500 aagtttaagc gtcgttacga acagaaggaa gaagagaatg atggaaaatg a 1551
3-348	Sequences	
3-348-1	Sequence Number [ID]	348
3-348-2	Molecule Type	AA
3-348-3	Length	303
3-348-4	Features Location/ Qualifiers	source 1..303 mol_type=protein organism=Medicago truncatula
3-348-5	NonEnglishQualifier Value Residues	MSSSGAIWSY EEEKAFENAI AMHWIEKEDS KEQWEKIAS VPNKSMQVK QHYQVLVDDV 60 SAIEEGHVSL PNYANELETI NSSNKDSSKA TTSSDKRSSC NFGSGFSGLG HDSTHHSSGK 120 GGLSRSSSSS EQERRKGIPW TEEHRLFL GLDKFGKGDW RSISRNFVIS RTPTQVASHA 180 QKYFIRLNSM NRDRRSSIH DITSVNNGDV ASNHQAPITG QHGNTLPSNT MGLGQSMKHR 240 VQPHHPPGGL GMYGAPVGHP FVAPPAHMAS AVGTPVMLPH HPHHHPPYVV PLAYPMAPPT 300 MHQ 303
3-349	Sequences	
3-349-1	Sequence Number [ID]	349
3-349-2	Molecule Type	AA
3-349-3	Length	268
3-349-4	Features Location/ Qualifiers	source 1..268 mol_type=protein organism=Triticum aestivum
3-349-5	NonEnglishQualifier Value Residues	MATQRAATAA WTNEEDKAFE NAVASGAPPP LDGVPEEDWF AALAASVPA STEEVRHHYE 60 ALVEDVGAID AGRVPLPRYA GEDSTPAAAA APSKDRREDR KSFDSDSGKG CSKAEQERRK 120 GIPWTEEEHR LFLGLGLEKFG KGDWRSISRN FVISRTPTQV ASHAQKYFIR LNSMNRDRR 180

		SSIHDTISIN NAAQAVPPIT GQAPAALGGP PGGMMKHPPM VGMYYGAPMG HPVAGHMVPA 240 AVGTPVMFPP GPGHPPYAMP VGYPAHQ 268
3-350	Sequences	
3-350-1	Sequence Number [ID]	350
3-350-2	Molecule Type	AA
3-350-3	Length	296
3-350-4	Features Location/ Qualifiers	source 1..296 mol_type=protein organism=Nicotiana tabacum
3-350-5	NonEnglishQualifier Value Residues	MSVSDLPSEV WSKEEEKAFE NAIAMHWVED SEQQWEKFAA VVPTKTIDEI KRHYQLLVDD 60 VAAIDAGHVP LPNYIGEES SSSNKETNLG YSGSVDRRSN CGYVNGFSGP THDPIGHSKG 120 GNSKAEQERR KGIPWTEEEH RLFLGLDKF GKGDWRSISR NFVISRTPTQ VASHAQKYFI 180 RLNSMNRDRR RSSIHDITSI NGGDVSIHQ PITGQQPNPN PSNPAALGPN IKHRNQPNMH 240 GFSMYGAPMG HPVAAPPSHM ASAVGTPVML PHGHHPPYVL PVAYPIPPPH PPPMHQ 296
3-351	Sequences	
3-351-1	Sequence Number [ID]	351
3-351-2	Molecule Type	AA
3-351-3	Length	294
3-351-4	Features Location/ Qualifiers	source 1..294 mol_type=protein organism=Zea mays
3-351-5	NonEnglishQualifier Value Residues	MTTLATDTAA WTREEDKAFE NAVAASAAPP ADGPPDDGWF TALVASVPAR TAEVRRHYE 60 ALVEDVAAIE AGRIPLLRYA GEESAAATPE GSGAAGPSTP KDGGGSGHR REERKGGVGA 120 DAGKSCSKAE QERRKGVPWT EEEHRLFLG LDKFGKGDWR SISRNFBVISR TPTQVASHAQ 180 KYFIRLNSMN RDRRRSSIHD ITSVTAGEVA TAGAPITGGG PTAAGAMPMPG PGGMKHHHPV 240 PPMGMYGHAP MGHVPVAGHMV APAAVGTPVM FPPGHHSPPV LPVGYPAPPT KMHQ 294
3-352	Sequences	
3-352-1	Sequence Number [ID]	352
3-352-2	Molecule Type	AA
3-352-3	Length	319
3-352-4	Features Location/ Qualifiers	source 1..319 mol_type=protein organism=Nicotiana benthamiana
3-352-5	NonEnglishQualifier Value Residues	MSVSDLPSEL WSKEEEKAFE NAIAVHWVED SEQQWEKFAE VVPTKTIDEI KRHYQLLVDD 60 VAAIDAGHVP LPNYIGEES SSSNKETNLG YSGSVDRRSN CGYTNGFSGP THDPIGHSKG 120 GNSKAEQERR KGIPWTEEEH RLFLGLDKF GKGDWRSISR NFVISRTPTQ VASHAQKYFI 180 RLNSMNRDRR RSSIHDITSI NGGDVSTHQA PITGQQPNPN PSNPALGPN IKHRNQPNMH 240 SYGMYGAPMG HPVAAPPSHM ASAVGTPVML PHGHHPPYVL PVAYPIPPPH APSMHYFYVF 300 SGVNQQLFCT TYKGIFLAS 319
3-353	Sequences	
3-353-1	Sequence Number [ID]	353
3-353-2	Molecule Type	AA
3-353-3	Length	272
3-353-4	Features Location/ Qualifiers	source 1..272 mol_type=protein organism=Hordeum vulgare
3-353-5	NonEnglishQualifier Value Residues	MASQRTATAA WTNEEDKAFE NAIAGAAPPP LDGVPEDAWF VALAVSVPAR STEEVRHYE 60 ALVEDVGAID AGRVPLRLYA GEESAAAAA AAAPGKDRRE DRKSFDSDSG KGCSKAEQER 120 RKGPWTEEE HRLFLGLDK FGKGDWRSIS RNFVISRTPT QVASHAQKYF IRLNSMNRDR 180 RRSSIHDITS INNAAQAAPQ QHGPITGQAP AAVLGPIMKH HPGMVGMYGG APMGHPVAGH 240 MVPAAVGTPV MFPPGPGHPP YAMPVGYAP HQ 272
3-354	Sequences	
3-354-1	Sequence Number [ID]	354
3-354-2	Molecule Type	AA
3-354-3	Length	320
3-354-4	Features Location/ Qualifiers	source 1..320 mol_type=protein organism=Oryza sativa
3-354-5	NonEnglishQualifier Value Residues	MAVEEASSSS GGGRGGGGGG GGEEGLSGCG GGWTREQEKA FENALATVGD DEEEGDGLWE 60 KLAEAVEGKT ADEVRRHYEL LVEDVDGIEA GRVPLLVYAG DGGVEEGSAG GKKGSAGGG 120

		GGGGGGGGHG EKGSAKSAEQ ERRKGIWTE DEHRLFLLGL EKYGKGDWRS ISRNFVISRT 180 PTQVASHAQK YFIRLNSMNR ERRRSSIHD I TSVNNGD TSA AQGPITGQPN GPSANPGKSS 240 KQSPQPANAP PGVDAYGT TI GQPVGGPLVS AVGTPVTL PV PAAPHIAYGM HAPVPGAVVP 300 GAPVNMPPMP YMPPTSHG 320
3-355	Sequences	
3-355-1	Sequence Number [ID]	355
3-355-2	Molecule Type	AA
3-355-3	Length	284
3-355-4	Features Location/ Qualifiers	source 1..284 mol_type=protein organism=Oryza sativa
3-355-5	NonEnglishQualifier Value Residues	MTSQAATTTT TAAAAAWTR EDDKAFENAL AASVPGARSA EEVRRHYEAL VEDVAAIDAG 60 RVPLPRYAGE ESAAPPDGAG AAAAASKDGG HRRDERKGGG GGYDGGKSCS KAEQERRKGI 120 PWTEEEHRLF LLGLDKFGKG DWRSISRNFV ISRTPTQVAS HAQKYFIRLN SMNRDRRRSS 180 IHDITSVTAG DQVAAQQGAP ITGHQATGNP AAAALGPPGM KHHHHHHHPGG APPMPMYYSA 240 APMGHPVAGH MVPAAVGTPV VFPPGHAPYV VPVGYPPA KMHQ 284
3-356	Sequences	
3-356-1	Sequence Number [ID]	356
3-356-2	Molecule Type	AA
3-356-3	Length	314
3-356-4	Features Location/ Qualifiers	source 1..314 mol_type=protein organism=Arabidopsis thaliana
3-356-5	NonEnglishQualifier Value Residues	MESVVATWSR EEEKAFENAI ALHCVEEIT EDQWNKMSSM VPSKALEEVK KHYQILLEDV 60 KAIENGQVPL PRYHHRKGLI VDEAAAAATS PANRDSHSSG SSEKPNPGT SGISSNGGR 120 SGGSRAEQER RKGIPWTEEE HRLFLLGLDK FGKGDWRSIS RNFVISRTPT QVASHAQKYF 180 IRLNSMNRDR RRSSIHDIIT VNNQAPAVTG GGQQPQVVKH RPAQPQPQPQ PQQQHHPPPT 240 MAGLGMYGGA PVGQPIIAPP DHMGSAVGTP VMLPPPMGTH HHHHHHHHLGV APYAVPAYPV 300 PPLPQQHPAP STMH 314
3-357	Sequences	
3-357-1	Sequence Number [ID]	357
3-357-2	Molecule Type	DNA
3-357-3	Length	909
3-357-4	Features Location/ Qualifiers	source 1..909 mol_type=other DNA organism=Actinidia rufa
3-357-5	NonEnglishQualifier Value Residues	atgtcagctt cagtggattg gagcagagaa gaagagaaa cttttgagaa tgccattgca 60 attcactggg cagaggactc caaagatgag ggccattggg aaaagattgc ttcattggtt 120 cctactaaaa ccatggaaga attgaagaaa cattaccaat tactagttaga ggatgtagct 180 gcaatcgaat cggacacgt gccaatcccc agctatatgt gggaggaagc atcttcatca 240 tcaagcaagg accaccatgg catttcaggg gtcacagcca cagccagtga gaagagatcc 300 aattgcagct atggaagcgg gtttccgggt ttgggtcatg actcgacagg tcatggaggg 360 aagggagggg ccagggtcga gcaagagaga cgaagaggga ttccatggac tgaagaagaa 420 cataggctgt ttctgcttgg tctggacaag ttgggaaag gggattggag gagcatttca 480 agaaactttg ttatatcaag gactccaaca caagtggcca gccatgccca gaaatacttc 540 attagattga actccatcaa cagagataga aggagatcaa gcatccatga catcaccagc 600 gtcaatagtg gcgatgcttc ggccgaccaa ccgccgccca tcatggcca gcagccaaat 660 ccaactagca cggtagcggc agccattgga ccaccaccgc cacttgaaa gcacaggatt 720 cccgtcccac caccaccgca tatggccggc ttgggcatgt atggaacgcc ggtagggcat 780 ccaatggctg ctccgagtca tcatatggcg tcagctgttg ggacgccagt tatgttctct 840 cctggacacc accacccta tgttgtcca gtagcatatc caatggcacc ccaacaatg 900 caccaataa 909
3-358	Sequences	
3-358-1	Sequence Number [ID]	358
3-358-2	Molecule Type	DNA
3-358-3	Length	936
3-358-4	Features Location/ Qualifiers	source 1..936 mol_type=other DNA organism=Actinidia eriantha
3-358-5	NonEnglishQualifier Value Residues	atgtcagatt ggagcagaga agaagagaaa gcttttgaga atgccattgc aattcactgg 60 gcagaggaag actccaaaga agacgacatc tggccaaaga ttgcttcgtt ggttcctact 120

		aaaaccattg aagaattgaa gaaacattat caattactag ttgaggatgt tgctgcaatc 180 gaagccggac acgtgcccat ccctaattat attggagagg aagcatcatc atcgtcaagc 240 aaggaccgcc atggcatttc aggggtgaca gccagtgaga agagatccaa ttgcaactat 300 ggaagtgggt tttttccggg ttttggtcat gactatccgg cggttccatg gaggaaggag 360 ggtccagggc tgagcaagag agacgaaaag ggattccatg gacagaagaa gaacataggc 420 ttgctcttgg gcaatgctat gcgccaatta ctagttgata atgccctagg agcattgaat 480 aacctctcca aactgttttt gcttggctcg gacaagtttg ggaaagggga ttggagaagc 540 atttcaagga actttgtgat atcaaggact ccaacacaag tggctagcca tgcccagaaa 600 tacttcatca gattgaactc catcaacaga gatagaagga gatcaagcat ccatgacatc 660 accagcgta atggcgggtg tgcttcggca caccaaccgc cgcccatcac tggccagcag 720 ccgaattcga ctagtgcaat ggagccatt ggaccacccc aaccgcacat ggccggtttg 780 ggcatgtatg gggcgccagt agggcatccg gtggctgctt cgagtcacat tatggcgctg 840 gccgttggga caccggttat gcttcctcct ggacaccacc acccctatgt tgtccagta 900 gcataatcaa tggcaccccc aacaatgcac caataa 936
3-359	Sequences	
3-359-1	Sequence Number [ID]	359
3-359-2	Molecule Type	DNA
3-359-3	Length	945
3-359-4	Features Location/ Qualifiers	source 1..945 mol_type=other DNA organism=Arabidopsis thaliana
3-359-5	NonEnglishQualifier Value Residues	atggagagtg tgggtggcaac atggagcaga gaagaagaga aagcattcga gaacgcaatt 60 gcgttgcat gtgtagagga agagataaca gaggatcaat ggaacaaaat gtcgtcaatg 120 gtaccaagca aagccttaga agaagtaaag aaacattacc aaatcctttt agaagatgtc 180 aaagcaatcg agaatggta agttccttta cctcgttatc atcaccgcaa aggtctcatc 240 gtagatgaag cagcagcagc agctacttct cctgccaaac gagactctca ttcctctgga 300 tcatctgaga agaaaccaaa tctggcacc tccgggataa gttagtccaa tggaggaaga 360 agtgggtggc cgagagctga gcaagagaga agaaaaggga ttccatggac tgaagaagag 420 catcggttgt tcttttggg ttggacaag ttgggaaag gagattggag aagcatttca 480 aggaactttg tgatctcaag aactccaaca caagttgcaa gtcagtctca aaaatacttc 540 atcaggctta actcgatgaa ccgagataga aggcgttcta gcattcacga catcaccact 600 gtgaacaatc aagctcctgc ggttacagga ggaggacaac aaccgcaagt ggttaaacat 660 agaccagctc agccacaacc acagccaca cgcgaaccac aacaacatca tcccccaaca 720 atggctggat tagggatgta tgggtggtcg ccagtgggac aaccgatcat cgcaccacct 780 gatcatatgg gttcagctgt tggaaacct gtgatgctt cacctccaat gggaactcat 840 catcatcacc atcaccatca tcttgagtt gctcctatg ctgtaccggc ttatccggta 900 ccgccattac cgcagcaaca tccagctccg tcgacaatgc actga 945
3-360	Sequences	
3-360-1	Sequence Number [ID]	360
3-360-2	Molecule Type	DNA
3-360-3	Length	897
3-360-4	Features Location/ Qualifiers	source 1..897 mol_type=other DNA organism=Arabidopsis thaliana
3-360-5	NonEnglishQualifier Value Residues	atgacggttg aggaagttag tgatggttct gtgtggagta gggaggatga tattgccttt 60 gagagagctc tagccaataa taccgatgaa tcagaggaac ggtgggagaa gattgctgca 120 gacgttccag gcaaaagtgt tgaacagatt aaagaacatt acgagctttt agttgaagat 180 gttactagga ttgaatcagg atgtgtgcct ctctcctgcc atgggtctcc tgaaggatcg 240 aatggccatg ctggtgatga aggagcaagt agtaagaaag gaggtaacag tcatcggtga 300 gagtctaacc aagcaggtaa atcaaagtcc gatcaagaac gacgaaaggg tatcgctggg 360 acagaagatg agcacaggtt atttcttctt ggtttggata agtacgggaa aggtgattgg 420 cgtagcattt ctcgcaactt tgtagtaaca agaacaccga cccaagtgtg gagccatgct 480 caaaagtatt tcattcgtct aaattcaatg aacaaagaca gaaggcgatc aagcattcac 540 gacatcacta gtgttggcaa cgcagatgtc tcaacgccac aaggaccaat cactggtcag 600 aacaacagca ataacaacaa caacaacaac aacaacaaca gttctcctgc tgttgctgga 660 ggaggaaaca aatcagccaa gcaagccgtc tctcaagcac cacctggacc tcctatgtat 720 ggaacaccgc ccataggta gccagcagtt ggaacaccag tgaacctccc agctccacct 780 cacatggctt atggagtcca tgcggctcca gtccctggct cagtggttcc tgggtcgagca 840 atgaacattg gtcaaatgcc gtacaccatg ccgcgtacac caacggctca taggtaa 897
3-361	Sequences	
3-361-1	Sequence Number [ID]	361
3-361-2	Molecule Type	DNA
3-361-3	Length	807

3-361-4	Features Location/ Qualifiers	source 1..807 mol_type=other DNA organism=Triticum aestivum
3-361-5	NonEnglishQualifier Value Residues	atggccactc agagagcagc cacggcgggc tggaccaacg aggaggacaa ggcgttcgag 60 aacgcggttg catccgggtg accgccggcg ctggacggcg tgccggagga ggactgggtc 120 gcggcgctcg cggcgagcgt gccggcgcgg tcgacggagg aggtgcggcg ccactacgag 180 gccctggtgg aggacgtcgg cgcgatcgac gccggcgcgg tccccctccc gcgctacgcc 240 ggggaggact ccaccctgc gccgcgagcg gcgcctagca aggaccggcg cgaggatcgg 300 aagagcttcg actccgactc cggcaaggcg tgctccaagg ccgagcagga gcggcgcaag 360 ggcatcccggt ggacggagga ggagcacagg ctgttcttgc tggggctgga gaagtctggc 420 aagggcgact ggcggagcat ctgcgcgaac ttcgtcatct cccggacgcc cacgcagggtg 480 gccagccacg cgagaagta ctcatccgc ctcaactcca tgaaccggga ccgccgccg 540 tccagcatcc acgacatcac cagcatcaac aacgccgccc aggccgtgcc gcccatcacc 600 ggccaggccc ccgccgcgt agggggggcg cccgggggca tgatgaagca tcccgggatg 660 gtcggcatgt acggaggggc gcccatgggc caccggctcg cggggcacat ggtccccgcc 720 gccgtcgga cgcggtcat gtcccccg ggccggggc acccgcccta cgccatgcc 780 gtcggctacc cggcgccgca tcagtga 807
3-362	Sequences	
3-362-1	Sequence Number [ID]	362
3-362-2	Molecule Type	DNA
3-362-3	Length	957
3-362-4	Features Location/ Qualifiers	source 1..957 mol_type=other DNA organism=Oryza sativa
3-362-5	NonEnglishQualifier Value Residues	atggcgggtg aggaggcgag cagcagcagt ggcggcggtc gtggtggggg tggcggtggg 60 ggtggggagg aggggttgtc cggttgccggc ggtgggtgga cgcgcgagca ggagaaggcg 120 ttcgagaacg cgctggcgac ggtgggggat gacgaggagg aaggggacgg gttgtgggag 180 aagctagcgg aggccgtgga ggggaagacg gccgacgagg tgaggcgga ctacgagctg 240 ctggtggagg acgtcgacgg catcgaggcc ggcgggtgc cgctcctggt gtacgccgc 300 gacggggcg tcgaggagg ctctgcggga ggtgggaaga aggggggttg tgggggagga 360 ggtggaggtg gaggggggca tggggagaag ggtcggcta agtcctctga gcaggagcgc 420 cggaaggga tcgcctggac ggaggacgag cacaggctgt tccttcttgg acttgagaag 480 tacggcaaa ggcactggag gagtatctca agaaactttg tgatctcaag gacaccacc 540 caagtagcta gtcatgcaca gaagtatttt attcgctga actcaatgaa cagagagagg 600 cggcgatcaa gtatacatga cataaccagc gtgaacaatg gagatacatc tgctgctcag 660 gggccaatca caggtcagcc aaatggcca tcagcaaatc ctggaaaatc ctctaagcag 720 tctctacagc cagcaaatgc gcctccaggc gtcgatgctt atggtacgac aattggacag 780 ccagttagtg gtctcttgt gtcgcagtt ggcactcctg ttacacttcc tgttctgct 840 gcacctcata tagcctatgg catgcatgcc cctgtccctg gagctgtagt ccctggtgcc 900 ccagtaaaca tgcctccaat gccctacccc atgccgccac caacatctca tggatga 957
3-363	Sequences	
3-363-1	Sequence Number [ID]	363
3-363-2	Molecule Type	DNA
3-363-3	Length	909
3-363-4	Features Location/ Qualifiers	source 1..909 mol_type=other DNA organism=Glycine max
3-363-5	NonEnglishQualifier Value Residues	atgtcatcaa gtggaaccat ttggagctat gaggaagaaa aagcatttga gaatgccata 60 gctatgcatt ggattgagga atcctcaaaa gagcaatggg agaaaattgc ttcagcagtt 120 cccagcaaaa gcatggaaga tgtgaagcaa cattaccagg ttctagtaga ggatgtaagt 180 gcaatagagg caggtcacat atcattccca aactatgctg ctctgagga aatcacatct 240 tcaaataagg actttcatgg ctcttccaag gccacaagct cagataaaa atcaaatgt 300 aattatggac gtggttttgc tgggttagga catgactcca ccactcatag tagtggtaaa 360 ggaggcttgt caaggctatc agaacaagaa agaagaaaag gaatccatg gactgaagag 420 gaacacaggt tattttact tggcttagag aagtttggaa aaggagattg gagaagcatt 480 tcaaggaact ttgtgatatc taggactccc actcaagtgg caagccatgc acaaaagtac 540 ttcataagg tgaattcaat gaatagagac aggaggagg ctagtatcca tgatatcact 600 agtgtgaaca atggagatgt ggctaataac caagcaccta ttacagggca gcatagtagc 660 acaattcctt caaacacaa ggtgttagga caatccctta agcatagagt tcagggtcac 720 ataccacctg gtttaggcat gtatggaaca ccagttggac atcctgtggc tgctctcca 780 gggcacatgg catctgcagt tggcactcct gtcatgcttc ctctggacc ccaccccat 840

		gctcatcctc atccacctta tgttcttctt cttgcttacc caatggcacc tccaacaatg 900 catcaataa 909
3-364	Sequences	
3-364-1	Sequence Number [ID]	364
3-364-2	Molecule Type	DNA
3-364-3	Length	873
3-364-4	Features Location/ Qualifiers	source 1..873 mol_type=other DNA organism=Capsicum annum
	NonEnglishQualifier Value	
3-364-5	Residues	atgtcacctt cagattggag cagagaagaa gaaaaggctt ttgaaaatgc cattgctatg 60 cactgggttg aaaattgtga acaacaatgg gagaaatttg gtttaatggt ccctacaaaa 120 accattgatg agctgaagca acattaccag ttacttgttg atgatgttac agcaattgaa 180 gcagggcaag ttccaatacc gaattacaaa ggagaagaag cttcttcttc ttcgacaaa 240 gagggaaatc tagggtattc tgcagggtca gtagaccgcc ggtctaattg tggttattca 300 aatggatttt ccgggacgaa tcatgaccg attggccatg gcggaaaagg gaattctagg 360 tctgaacaag agagacgaaa ggggattcca tggactgaag aagaacatag gttatcttg 420 ctagggttag agaaatttgg gaaaggagat tggagaagta tatcgaaggaa tttgtgata 480 tcgagaacac caacacaagt agctagtcac gctcagaagt atttatacg tttgaattcc 540 atgaatagag atagacgaag gtcgagtata catgacatta caagtattaa caatggagac 600 gtatctactc accaagcgcc aattacaggc caacaagtga atacaaatcc atcaaatgca 660 gcagcacttg gaccatctgt taagcacaga actcagcaa atatgcatgg tatgggtatg 720 tacggtgctc ccatgggtca tcccgtcgca gcaccgtcga gccacatggc gtcagctggt 780 ggaactccgg ttatgatgcc tcatggacat ctccttatg tacttcagat tgcatatccc 840 ataccaccgc caccaccacc catgcaccag tga 873
3-365	Sequences	
3-365-1	Sequence Number [ID]	365
3-365-2	Molecule Type	AA
3-365-3	Length	516
3-365-4	Features Location/ Qualifiers	source 1..516 mol_type=protein organism=Actinidia eriantha
	NonEnglishQualifier Value	
3-365-5	Residues	MLKIKRVPTV VSNFQKDEAD DGARSGGCG RNCLQKCCIQ GAKLPLYAFK RVNEVVGEKG 60 VLALDNEEAP VAFLDSLILG EWEDRVQRGL FRYDVTACET KVIPGEYGF I AQLNEGRHLK 120 KRPTEFRVDK VLQPFDGSKF NFKVQGQEEV LFQFEASNDN EVQFFPNAPV DVENSPSVVA 180 INVSPIEYGH VLLIPSILEC LPQRIDRESF LLALHMAAEA GNPYFRLGYN SLGAFATINH 240 LHFQAYYLAV PFPIEKAPTR KITTLNGNNI PYNVLISDSG KRIFLLPQCY AEKQALGEVS 300 SEL LDTQVNP AVWEISGHMV LKRKEDYEEA SEGNAWRLLA EVSLSEERFE EVKALIFEAI 360 SCADDGSGST AENLLEEDD NPQSREEAND ALNKGSHCGM VCVVAAAAA AEVASVMATC 420 FRNRFYLSLL GTPPPELIGV VAFPTHKMKM KGVHPQLQWI SYVTQSGRLM NIMMTKIHV 480 GKVYHIKARR QMAQSIGQIA KFKRRYEQKE EENDGK 516