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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
26 November 2015 (26.11.2015)

WIPO | PCT

(10) International Publication Number
WO 2015/179191 A2

(51) International Patent Classification:
C12N 9/12 (2006.01) *C12N 15/82* (2006.01)

(21) International Application Number:
PCT/US2015/030604

(22) International Filing Date:
13 May 2015 (13.05.2015)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/000,173 19 May 2014 (19.05.2014) US

(71) Applicant: **UNIVERSITY OF FLORIDA RESEARCH FOUNDATION, INC.** [US/US]; 223 Grinter Hall, Gainesville, FL 32611 (US).

(72) Inventors: **HANNAH, Larkin, Curtis**; 223 Grinter Hall, Gainesville, FL 32611 (US). **SHAW, Janine, R.**; 223 Grinter Hall, Gainesville, FL 32611 (US). **BOEHLEIN, Susan, Kim**; 223 Grinter Hall, Gainesville, FL 32611 (US). **STEWART, Jon, Dale**; 223 Grinter Hall, Gainesville, FL 32611 (US). **SULLIVAN, Bradford, T.**; 223 Grinter Hall, Gainesville, FL 32611 (US).

(74) Agent: **HANSON, Robert, E.**; Dentons Us LLP, P.O.Box 061080, Wacker Drive Station, Willis Tower, Chicago, IL 60606-1080 (US).

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

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

Published:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

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

WO 2015/179191 A2

(54) **Title**: METHODS FOR INCREASING GRAIN YIELD

(57) **Abstract**: Compositions and methods for increasing heat resistance or starch biosynthesis in plants are provided herein. Polynucleotides, polypeptides, and expression constructs for expressing mutant AGPase subunit proteins, plants comprising the polynucleotides, polypeptides or expression constructs, and methods of producing transgenic plants are also provided.

METHODS FOR INCREASING GRAIN YIELD

Reference to Related Applications

[0001] This application claims the benefit of United States Provisional Application No. 62/000,173 filed May 19, 2014, which is incorporated herein by reference in its entirety.

Statement of Government Support

[0002] This invention was made with Government support under USDA/NIFA 2010-04228 awarded by the U.S. Department of Agriculture. The Government has certain rights in the invention.

Incorporation of the Sequence Listing

[0003] The sequence listing that is contained in the file named "UFFL048WO_ST25.txt", which is 151 kilobytes (as measured in Microsoft Windows®) and was created on May 8, 2015, is filed herewith by electronic submission and is incorporated by reference herein.

Field of the Invention

[0004] The invention provides compositions and methods for increasing grain yield in cereal crops. More specifically, the invention is related to temperature-stable mutants of maize endosperm ADP-glucose pyrophosphorylase (AGPase), and methods for use thereof for increasing grain yield.

Background of the Invention

[0005] ADP-glucose pyrophosphorylase (AGPase) is a highly regulated enzyme in the starch biosynthesis pathway. *In planta* the enzyme has a complex structure, an $\alpha 2\beta 2$ heterotetramer, that when expressed in the endosperm of seeds is heat labile. The heat lability of AGPase is often linked to grain loss during hot weather, and therefore increasing the thermo-stability of this enzyme is of great agronomical importance. Previously, heat stabile variants of this enzyme have been transformed into maize, potato, wheat, and rice and yield increases in seed number or weight were obtained.

[0006] Conventional directed evolution methods for enhancing protein thermo-stabilities rely on collections of random mutations spread across the entire linear sequence of amino acids. In principle, such a strategy offers the chance to explore the complete sequence space of a given

protein and thereby provide the globally optimal solution. In practice, however, the number of sequence variants at the DNA level (64^{517} for the maize endosperm large subunit) means that the required library sizes vastly outstrip the ability of molecular biology to provide them and practical screening methods to evaluate them. This necessarily means that sequence space exploration will be only partial.

SUMMARY OF THE INVENTION

[0007] In one aspect, the present invention provides a polynucleotide comprising a sequence encoding a variant plant AGPase large subunit polypeptide, said polypeptide comprising a mutation at a position homologous or corresponding to amino acid 96, 161, or 443 of SEQ ID NO: 20 relative to a wild type plant AGPase large subunit polypeptide. In some embodiments, the variant plant AGPase large subunit polypeptide comprises a conserved motif having SEQ ID NO: 21, SEQ ID NO: 22, or SEQ ID NO: 23. In other embodiments, the variant plant AGPase large subunit polypeptide comprises at least 70% identity to SEQ ID NO:70. In further embodiments, the variant plant AGPase large subunit polypeptide comprises mutations in at least two positions homologous to amino acid 96, 161, or 443 of SEQ ID NO: 20 relative to a wild type plant AGPase large subunit polypeptide, or in each position homologous to amino acid 96, 161, and 443 of SEQ ID NO: 20 relative to a wild type plant AGPase large subunit polypeptide.

[0008] In some embodiments, the variant plant AGPase large subunit polypeptide of the invention comprises at least one mutation selected from the group consisting of a glycine or a glutamine at said position 161, an arginine or a glycine at said position 96, and an arginine at said position 443. The variant plant AGPase large subunit polypeptide of the invention may be operably linked to a heterologous promoter functional in plants. In yet other embodiments, the variant plant AGPase large subunit polypeptide comprising said mutation relative to a wild type plant AGPase large subunit polypeptide shown in FIG. 1.

[0009] In another aspect, the present invention provides a recombinant construct comprising a polynucleotide comprising a sequence encoding a variant plant AGPase large subunit polypeptide, said polypeptide comprising a mutation at a position homologous to amino acid 96, 161, or 443 of SEQ ID NO: 20 relative to a wild type plant AGPase large subunit

polypeptide. In some embodiments, the invention provides a polypeptide encoded by said polynucleotide. In other embodiments, the invention provides a plant, seed, cell, or plant part comprising said polynucleotide. In further embodiments, the plant, seed, cell, or plant part comprising said polynucleotide is a monocotyledonous or dicotyledonous plant, seed, cell, or plant part. In yet further embodiments, the plant, seed, cell, or plant part is from a plant selected from the group consisting of rice, wheat, barley, oats, rye, sorghum, maize, sugarcane, pineapple, onion, bananas, coconut, lilies, millet, tomato, potato, sweet potato, pea, strawberry, beet, chickpea, watermelon, muskmelon, cassava, taro, sunflower, flax, and beans.

[0010] In another aspect, the invention provides a method of increasing the resistance or tolerance of a plant to heat stress conditions or increasing starch biosynthesis of a plant comprising expressing in the plant a polynucleotide comprising a sequence encoding a variant plant AGPase large subunit polypeptide, said polypeptide comprising a mutation at a position homologous to amino acid 96, 161, or 443 of SEQ ID NO: 20 relative to a wild type plant AGPase large subunit polypeptide. In some embodiments, the plant is a monocotyledonous or a dicotyledonous plant. In further embodiments, the plant is a plant selected from the group consisting of rice, wheat, barley, oats, rye, sorghum, maize, sugarcane, pineapple, onion, bananas, coconut, lilies, millet, tomato, potato, sweet potato, pea, strawberry, beet, chickpea, watermelon, muskmelon, cassava, taro, sunflower, flax, and beans. In other embodiments, the invention provides a method comprising transforming a plant with a polynucleotide comprising a sequence encoding a variant plant AGPase large subunit polypeptide, said polypeptide comprising a mutation at a position homologous to amino acid 96, 161, or 443 of SEQ ID NO: 20 relative to a wild type plant AGPase large subunit polypeptide, and regenerating a plant therefrom. In a further embodiment, the method comprises crossing a parent plant comprising said polynucleotide with itself or a second plant to obtain the plant in which resistance of a plant to heat stress conditions or starch biosynthesis is increased.

BRIEF DESCRIPTION OF THE DRAWINGS

[0011] FIG. 1: Shows an alignment of plant AGPase large subunit sequences. The alignment includes AGPase sequences from *Zea mays*, Genbank Accession No. M81603 (SEQ ID NO: 40); *Zea mays*, Genbank Accession No. Z38111 (SEQ ID NO: 41); *Zea mays*, Genbank Accession No. NP_001106017 (SEQ ID NO: 42); *Hordeum vulgare*, Genbank Accession No.

X67151 (SEQ ID NO: 43); *Hordeum vulgare*, Genbank Accession No. U66876 (SEQ ID NO: 44); *Sorghum bicolor*, Genbank Accession No. T03445 (SEQ ID NO: 45); *Oryza sativa*, Genbank Accession No. T04156 (SEQ ID NO: 46); *Oryza sativa*, Genbank Accession No. AAK27727 (SEQ ID NO: 47); *Oryza sativa*, Genbank Accession No. U66041 (SEQ ID NO: 48); *Oryza sativa*, Genbank Accession No. D50317 (SEQ ID NO: 49); *Triticum aestivum*, Genbank Accession No. P12299 (SEQ ID NO: 50); *Lycopersicon esculentum*, Genbank Accession No. T07674 (SEQ ID NO: 51); *Lycopersicon esculentum*, Genbank Accession No. T07682 (SEQ ID NO: 52); *Lycopersicon esculentum*, Genbank Accession No. AAC49942 (SEQ ID NO: 53); *Solanum tuberosum*, Genbank Accession No. X76136 (SEQ ID NO: 54); *Solanum tuberosum*, Genbank Accession No. X61187 (SEQ ID NO: 55); *Solanum tuberosum*, Genbank Accession No. P55242 (SEQ ID NO: 56); *Ipomoea batatas*, Genbank Accession No. AFL55396 (SEQ ID NO: 57); *Ipomoea batatas*, Genbank Accession No. AFL55397 (SEQ ID NO: 58); *Ipomoea batatas*, Genbank Accession No. AFL55398 (SEQ ID NO: 59); *Ipomoea batatas*, Genbank Accession No. AFL55399 (SEQ ID NO: 60); *Citrus sinensis*, Genbank Accession No. ACF77017 (SEQ ID NO: 61); *Pisum sativum*, Genbank Accession No. X96766 (SEQ ID NO: 62); *Fragaria x ananassa*, Genbank Accession No. AAS00542 (SEQ ID NO: 63); *Beta vulgaris*, Genbank Accession No. P55233 (SEQ ID NO: 64); *Cicer arietinum*, Genbank Accession No. AF356003 (SEQ ID NO: 65); *Citrullus lanatus*, Genbank Accession No. JE0132 (SEQ ID NO: 66); *Citrullus lanatus*, Genbank Accession No. JE0133 (SEQ ID NO: 67); *Cucumis melo*, Genbank Accession No. AF030383 (SEQ ID NO: 68); *Cucumis melo*, Genbank Accession No. AF030384 (SEQ ID NO: 69). Conserved motifs surrounding mutated amino acids are underlined (SEQ ID NOs: 21-23), and amino acid positions mutated according to the present invention are highlighted.

DETAILED DESCRIPTION

[0012] The present invention describes novel ADP-glucose pyrophosphorylase (AGPase) polypeptides. AGPase enzymes may also be referred to as glucose-1-phosphate adenylyltransferase enzymes. As used herein, "AGPase" refers to an enzyme that catalyzes the conversion of a glucose-1-phosphate molecule and an ATP molecule to produce an ADP-glucose molecule and a diphosphate molecule. In many organisms, functional AGPase enzymes exist as multimers of multiple AGPase subunits. As used herein, "AGPase subunit" refers to a polypeptide that interacts with at least one additional polypeptide to form a functional AGPase

enzyme. In many microorganisms, AGPase enzymes exist as a homotetramer of four identical AGPase subunits. In many higher plants, AGPase enzymes exist as a heterotetramer containing two large subunits and two small subunits. As used herein, "AGPase large subunit" refers to a polypeptide that can interact with other AGPase subunits to form a functional AGPase enzyme and that has a higher molecular weight than one or more of the other subunits that interact to form the AGPase enzyme. In a preferred embodiment, the modified AGPase large subunit of the invention possesses at least 70% identity to SEQ ID NO: 70. In other embodiments, the modified AGPase large subunit of the invention possesses at least 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, or 99% identity to SEQ ID NO: 70.

[0013] Maize endosperm AGPase is highly regulated and temperature sensitive. Thermo-sensitivity of this enzyme is linked to grain loss in a variety of cereal crops. The present invention surprisingly provides variants of AGPase that increase the thermal stability of this heat labile enzyme. Nine amino acid positions in the AGPase large subunit were targeted for mutagenesis based on their conformational mobility. These mutagenic positions were selected using crystal structure atomic displacement parameters (B-factors) obtained from the potato small subunit homotetramer crystal structure. After each round of mutagenesis, iodine staining and antibody capture activity assays at varying temperatures were used to select the best position and amino acid change. Following 3 iterations, the initial detection method of iodine staining was saturated and a heat stable AGPase mutant was obtained. Kinetic studies on the heat stable mutant showed that it also exhibited an increased affinity for the activator, 3-PGA. Both temperature stability and allosteric regulation are of great importance in increasing grain yield.

[0014] The present invention therefore provides methods for improving the thermal stability of AGPase based on identification of conformationally mobile residues. Conformationally mobile residues are possible sites for initiating global protein unfolding. Since unfolding is highly cooperative, eliminating these motions can make a protein less susceptible to denaturation. Surprisingly, the present invention shows that mutagenesis of residues chosen based on a B-factor analysis of a different subunit from a protein made by a different organism resulted in improved properties of temperature stability in maize AGPase.

[0015] The present invention further provides methods which take advantage of the observation that thermo-stabilizing mutations are additive, allowing variations at individual codons to be examined sequentially, rather than simultaneously. Once improved variants at each position have been identified using the methods of the present invention, they can be combined to yield mutants with even better thermal stabilities. The invention further provides AGPase mutants in which not all of the mutations contribute directly to enhanced thermo-stability, but the combination of several mutations providing modest improvements leads to a significant overall change. The effect of several mutations of the present invention on thermal stability of the mutant enzyme may also be additive or synergistic.

[0016] The present invention therefore provides novel heat stable variants of AGPase, including variants with from one to three changes to the coding sequence. The heat stable variants of the present invention, including D161G, surprisingly link heat stability with allosteric regulation. This phenomenon was also observed for a T142F variant that exhibits increased heat stability and a decreased K_a for 3-PGA such that activity was found in the absence of 3-PGA.

[0017] The invention therefore permits increases in crop performance, particularly under high heat growing conditions. In view of increasing concerns regarding rising global temperatures due to climate change, this represents a significant advance to agriculture and the art in general.

I. Nucleic Acids and Proteins

[0018] As used herein, the term "nucleic acid" or "polynucleotide" refers to a single or double-stranded polymer of deoxyribonucleotide bases or ribonucleotide bases read from the 5' to the 3' end. A nucleic acid or polynucleotide may also optionally contain non-naturally occurring or altered nucleotide bases that permit correct read through by a polymerase and do not reduce expression of a polypeptide encoded by that nucleic acid. The terms "nucleotide sequence" or "nucleic acid sequence" refer to both the sense and antisense strands of a nucleic acid as either individual single strands or in the duplex. The term "ribonucleic acid" (RNA) is inclusive of RNAi (inhibitory RNA), dsRNA (double stranded RNA), siRNA (small interfering RNA), mRNA (messenger RNA), miRNA (micro-RNA), tRNA (transfer RNA, whether charged or discharged with a corresponding acylated amino acid), and cRNA (complementary RNA).

The words "nucleic acid fragment," "nucleotide sequence fragment", or more generally "fragment" will be understood by those in the art as a functional term that includes genomic sequences, ribosomal RNA sequences, transfer RNA sequences, messenger RNA sequences, operon sequences, and smaller engineered nucleotide sequences that express or may be adapted to express, proteins, polypeptides or peptides. The nomenclature used herein is that required by Title 37 of the United States Code of Federal Regulations §1.822 and set forth in the tables in WIPO Standard ST.25 (1998), Appendix 2, Tables 1 and 3.

[0019] As used herein, the terms “nucleic acid” and “polynucleotide” refer to a deoxyribonucleotide, ribonucleotide, or a mixed deoxyribonucleotide and ribonucleotide polymer in either single- or double-stranded form, and unless otherwise limited, would encompass known analogs of natural nucleotides that can function in a similar manner as naturally-occurring nucleotides. The polynucleotide sequences include the DNA strand sequence that is transcribed into RNA and the strand sequence that is complementary to the DNA strand that is transcribed. The polynucleotide sequences also include both full-length sequences as well as shorter sequences derived from the full-length sequences. Allelic variations of the exemplified sequences also fall within the scope of the subject invention. The polynucleotide sequence includes both the sense and antisense strands either as individual strands or in the duplex.

[0020] As used herein, the term "recombinant nucleic acid," "recombinant polynucleotide" or “recombinant DNA molecule” refers to a polynucleotide that has been altered from its native state, such as by linkage to one or more other polynucleotide sequences to which the recombinant polynucleotide molecule is not normally linked to in nature. Such molecules may or may not be present, for example, in a host genome or chromosome.

[0021] The present invention further provides polynucleotides that are complementary in sequence to the polynucleotides disclosed herein. Polynucleotides and polypeptides of the invention can be provided in purified or isolated form.

[0022] The subject invention also concerns oligonucleotide probes and primers, such as polymerase chain reaction (PCR) primers, that can hybridize to a coding or non-coding sequence of a polynucleotide of the present invention. Oligonucleotide probes of the invention can be used in methods for detecting and quantitating nucleic acid sequences encoding a mutant AGPase

large subunit polypeptide of the invention. Oligonucleotide primers of the invention can be used in PCR methods and other methods involving nucleic acid amplification. In a preferred embodiment, a probe or primer of the invention can hybridize to a polynucleotide of the invention under stringent conditions. Probes and primers of the invention can optionally comprise a detectable label or reporter molecule, such as fluorescent molecules, enzymes, radioactive moiety (e.g., ^{3}H, ^{35}S, ^{125}I, etc.), and the like. Probes and primers of the invention can be of any suitable length for the method or assay in which they are being employed. Typically, probes and primers of the invention will be 10 to 500 or more nucleotides in length. Probes and primers that are 10 to 20, 21 to 30, 31 to 40, 41 to 50, 51 to 60, 61 to 70, 71 to 80, 81 to 90, 91 to 100 or more nucleotides in length are contemplated within the scope of the invention. Probes and primers of the invention can have complete (100%) nucleotide sequence identity with the polynucleotide sequence, or the sequence identity can be less than 100%. For example, sequence identity between a probe or primer and a sequence can be 99%, 98%, 97%, 96%, 95%, 90%, 85%, 80%, 75%, 70% or any other percentage sequence identity allowing the probe or primer to hybridize under stringent conditions to a nucleotide sequence of a polynucleotide of the invention. In one embodiment, a probe or primer of the invention has 70% or greater, 75% or greater, 80% or greater, 85% or greater, 90% or greater, or 95% to 100% sequence identity with a nucleotide sequence provided herein, including the complement thereof.

[0023] Because of the degeneracy of the genetic code, a variety of different polynucleotide sequences can encode polypeptides or mutant polypeptides disclosed herein. A table showing all possible triplet codons (and where U also replaces T) and the amino acid encoded by each codon is described in Lewin (1985). In addition, it is well within the capability of one of skill in the art to create alternative polynucleotide sequences encoding the same, or essentially the same, mutant polypeptides of the subject invention. These variant or alternative polynucleotide sequences are within the scope of the subject invention. As used herein, references to "essentially the same" sequence refers to sequences which encode amino acid substitutions, deletions, additions, or insertions which do not materially alter the functional activity of the polypeptide encoded by the polynucleotides of the present invention. Allelic variants of the nucleotide sequences encoding a wild type or mutant polypeptide of the invention are also encompassed within the scope of the invention.

[0024] Amino acids can be generally categorized in the following classes: non-polar, uncharged polar, basic, and acidic. Conservative substitutions whereby a mutant AGPase small or large subunit polypeptide of the present invention and/or a wild type or mutant AGPase small or large subunit polypeptide having an amino acid of one class is replaced with another amino acid of the same class fall within the scope of the subject invention so long as the polypeptide having the substitution still retains substantially the same functional activity (*e.g.*, enzymatic and/or increased heat stability of an AGPase enzyme) as the polypeptide that does not have the substitution. Functional activity may be determined according to the methods of Giroux *et al* 1996 *Proc Natl Acad Sci USA* 93: 5824-5829 or as set forth in the Experimental section below. Polynucleotides encoding a mutant AGPase small subunit polypeptide and/or a wild type or mutant AGPase large subunit polypeptide having one or more amino acid substitutions in the sequence are contemplated within the scope of the present invention.

[0025] **Table 1** provides a listing of examples of amino acids belonging to each class.

Table 1. Classes of Amino Acid

Nonpolar	Ala, Val, Leu, Ile, Pro, Met, Phe, Trp
Uncharged Polar	Gly, Ser, Thr, Cys, Tyr, Asn, Gln
Acidic	Asp, Glu
Basic	Lys, Arg, His

[0026] Substitution of amino acids other than those specifically exemplified or naturally present in a wild type or mutant polypeptide and/or AGPase enzyme of the invention are also contemplated within the scope of the present invention. For example, non-natural amino acids can be substituted for the amino acids of a mutant AGPase small or large subunit polypeptide, so long as the mutant polypeptide having the substituted amino acids retains substantially the same functional activity as the mutant polypeptide in which amino acids have not been substituted. Examples of non-natural amino acids include, but are not limited to, ornithine, citrulline, hydroxyproline, homoserine, phenylglycine, taurine, iodotyrosine, 2,4-diaminobutyric acid, a-amino isobutyric acid, 4-aminobutyric acid, 2-amino butyric acid, γ-amino butyric acid, e-amino hexanoic acid, 6-amino hexanoic acid, 2-amino isobutyric acid, 3-amino propionic acid, norleucine, norvaline, sarcosine, homocitrulline, cysteic acid, t-butylglycine, t-butylalanine,

phenylglycine, cyclohexylalanine, β-alanine, fluoro-amino acids, designer amino acids such as β-methyl amino acids, C-methyl amino acids, N-methyl amino acids, and amino acid analogues in general. Non-natural amino acids also include amino acids having derivatized side groups. Furthermore, any of the amino acids in the protein can be of the D (dextrorotary) form or L (levorotary) form. Allelic variants of a protein sequence of a wild type or mutant AGPase small or large subunit polypeptide of the present invention are also encompassed within the scope of the invention.

II. Methods of Modifying Nucleic Acids and Proteins

[0027] The subject invention also concerns variants of the polynucleotides of the present invention that encode functional wild type or mutant AGPase small or large subunit polypeptides of the invention. Variant sequences include those sequences wherein one or more nucleotides of the sequence have been substituted, deleted, and/or inserted. The nucleotides that can be substituted for natural nucleotides of DNA have a base moiety that can include, but is not limited to, inosine, 5-fluorouracil, 5-bromouracil, hypoxanthine, 1-methylguanine, 5-methylcytosine, and tritylated bases. The sugar moiety of the nucleotide in a sequence can also be modified and includes, but is not limited to, arabinose, xylulose, and hexose. In addition, the adenine, cytosine, guanine, thymine, and uracil bases of the nucleotides can be modified with acetyl, methyl, and/or thio groups. Sequences containing nucleotide substitutions, deletions, and/or insertions can be prepared and tested using standard techniques known in the art.

[0028] As used herein, the term "percent sequence identity" or "percent identity" refers to the percentage of identical nucleotides in a linear polynucleotide sequence of a reference ("query") polynucleotide molecule (or its complementary strand) as compared to a test ("subject") polynucleotide molecule (or its complementary strand) when the two sequences are optimally aligned (with appropriate nucleotide insertions, deletions, or gaps totaling less than 20 percent of the reference sequence over the window of comparison). Optimal alignment of sequences for aligning a comparison window are well known to those skilled in the art and may be conducted by tools such as the local homology algorithm of Smith and Waterman, the homology alignment algorithm of Needleman and Wunsch, the search for similarity method of Pearson and Lipman, and preferably by computerized implementations of these algorithms such as GAP, BESTFIT, FASTA, and TFASTA available as part of the GCG® Wisconsin Package®

(Accelrys Inc., Burlington, Mass.). Polynucleotides and polypeptides contemplated within the scope of the subject invention can also be defined in terms of identity and/or similarity ranges with those sequences of the invention specifically exemplified herein. In certain embodiments, the invention provides polypeptide sequences having at least about 70, 80, 85, 90, 95, 99, or 99.5 percent identity to a polypeptide sequence provided herein. In certain embodiments, the invention provides polynucleotide sequences having at least about 70, 80, 85, 90, 95, 99, or 99.5 percent identity to a polynucleotide sequence provided herein, including SEQ ID NO: 19.

[0029] In certain embodiments, the invention provides polynucleotides encoding polypeptides comprising the amino acid sequence provided herein, or a fragment or variant thereof. In certain embodiments, the polynucleotides encode polypeptides comprising a variant of the amino acid sequence provided herein wherein the amino acid homologous or corresponding to position 161 has been mutated, wherein the amino acid at position 96 has been mutated, wherein the amino acid at position 443 has been mutated, or combinations, fragments, or variants thereof. As used herein, an amino acid homologous or corresponding to position 96, 161, or 443 is an amino acid that is aligned with positions 96, 161, or 443 of SEQ ID NO: 20 when using the methods to align polypeptides that are described herein. For example, see FIG. 1. In certain embodiments, the polynucleotides of the invention encode one or more polypeptides comprising a variant wherein the amino acids at positions 161 and 96 have been mutated, wherein the amino acids at positions 96 and 443 have been mutated, or wherein the amino acids at positions 161 and 443 have been mutated, or combinations, fragments, or variants thereof. In certain embodiments, the polynucleotides introduced into the plant encode one or more polypeptides comprising a variant wherein the amino acids at positions 161, 96, and 443 have been mutated, or fragments or variants thereof. The invention further provides polynucleotides encoding polypeptides comprising a variant of the amino acid sequences provided herein wherein the amino acid corresponding to position 161 has been mutated to be a G or Q, wherein the amino acid at position 96 has been mutated to be a G or R, or wherein the amino acid at position 443 has been mutated to be an R, or combinations thereof.

[0030] In certain embodiments, the invention provides polynucleotides encoding a polypeptide comprising a fragment or variant of the amino acid sequence provided herein and comprising a conserved motif, for example, such as SEQ ID NO: 21. In other embodiments,

polynucleotides of the invention encode a fragment or variant of the amino acid sequences provided herein and comprising a conserved motif having SEQ ID NO: 21, wherein position 3 of SEQ ID NO: 21 is an uncharged polar amino acid. In yet other embodiments, the polynucleotides of the invention encode a fragment or variant comprising a conserved motif having SEQ ID NO: 21, wherein position 3 of SEQ ID NO: 21 is G or Q. In certain embodiments, the invention provides polynucleotides encoding a polypeptide comprising a fragment or variant comprising a conserved motif having SEQ ID NO: 22. In other embodiments, polynucleotides of the invention encode a fragment or variant comprising a conserved motif having SEQ ID NO: 22, wherein position 8 of the motif is an uncharged polar amino acid or a basic amino acid. In yet other embodiments, the polynucleotides of the invention encode a fragment or variant comprising a conserved motif having SEQ ID NO: 22, wherein position 8 of the motif is a G or R.

[0031] In yet other embodiments, the invention provides polynucleotides encoding a polypeptide comprising conserved motif having SEQ ID NO: 23. In other embodiments, in the conserved motif having SEQ ID NO: 23, position 7 of the motif is a basic amino acid. In further embodiments, position 7 is an R.

[0032] Fragments and variants of a mutant polypeptide of the present invention can be generated as described herein and tested for the presence of enzymatic and heat stable function using standard techniques known in the art. Thus, an ordinarily skilled artisan can readily prepare and test fragments and variants of a mutant polypeptide of the invention and determine whether the fragment or variant retains functional activity relative to full-length or a non-variant mutant polypeptide. Fragments and variants of mutant polypeptides can be tested for AGPase activity, for example using methods disclosed herein or by other methods well-known in the art.

[0033] The subject invention also concerns isolated mutant AGPase small or large subunit polypeptides. In one embodiment, the mutant AGPase small or large subunit polypeptide is an AGPase small or large subunit polypeptide of *Zea mays*. In a specific embodiment, an AGPase large subunit polypeptide of the invention has an amino acid sequence as shown in the sequence listing or drawings, or functional fragment or variant thereof. An AGPase small or large subunit polypeptide or enzyme of the invention can be purified using standard techniques

known in the art. In one embodiment, a polynucleotide of the invention encoding an AGPase small or large subunit polypeptide is incorporated into a microorganism, such as *E. coli*, and the AGPase small or large subunit polypeptide expressed in the microorganism and then isolated therefrom.

[0034] In certain embodiments, polypeptides of the invention, and functional peptide fragments thereof, can be used to generate antibodies that bind specifically to a polypeptide of the invention, and such antibodies are contemplated within the scope of the invention. The antibodies of the invention can be polyclonal or monoclonal and can be produced and isolated using standard methods known in the art.

[0035] Polypeptide fragments according to the invention typically comprise a contiguous span of at least about 25 and about 515 amino acids of a sequence disclosed herein, including SEQ ID NO: 70. In certain embodiments, polypeptide fragments comprise 25, 50, 95, 150, or 500 amino acids of a sequence provided herein.

[0036] Fragments of a mutant AGPase small or large subunit polypeptide of the invention or an AGPase large subunit polypeptide, as described herein, can be obtained by cleaving the polypeptides of the invention with a proteolytic enzyme (such as trypsin, chymotrypsin, or collagenase) or with a chemical reagent, such as cyanogen bromide (CNBr). Alternatively, polypeptide fragments can be generated in a highly acidic environment, for example at pH 2.5. Polypeptide fragments can also be prepared by chemical synthesis or using host cells transformed with an expression vector comprising a polynucleotide encoding a fragment of an AGPase large subunit polypeptide or a fragment of a mutant AGPase small subunit polypeptide of the invention, for example, a mutant polypeptide that is a fragment of an amino acid sequence provided herein. Fragments of a mutant large or small subunit AGPase polypeptide of the invention also contemplated herein include fragments of the polypeptide wherein all or a part of a transit or signal sequence of the polypeptide is removed.

III. Expression Constructs

[0037] Polynucleotides useful in the present invention can be provided in an expression construct. Expression constructs of the invention generally include regulatory elements that are functional in the intended host cell in which the expression construct is to be expressed. Thus, a

person of ordinary skill in the art can select regulatory elements for use in bacterial host cells, yeast host cells, plant host cells, insect host cells, mammalian host cells, and human host cells. Regulatory elements include promoters, transcription termination sequences, translation termination sequences, enhancers, and polyadenylation elements. As used herein, the term "expression construct" refers to a combination of nucleic acid sequences that provides for transcription of an operably linked nucleic acid sequence. As used herein, the term "operably linked" refers to a juxtaposition of the components described wherein the components are in a relationship that permits them to function in their intended manner. In general, operably linked components are in contiguous relation.

[0038] An expression construct of the invention can comprise a promoter sequence operably linked to a polynucleotide sequence encoding a mutant polypeptide of the invention. Promoters can be incorporated into a polynucleotide using standard techniques known in the art. Multiple copies of promoters or multiple promoters can be used in an expression construct of the invention. In a preferred embodiment, a promoter can be positioned about the same distance from the transcription start site in the expression construct as it is from the transcription start site in its natural genetic environment. Some variation in this distance is permitted without substantial decrease in promoter activity. A transcription start site is typically included in the expression construct.

[0039] If the expression construct is to be provided in or introduced into a plant cell, then plant viral promoters, such as, for example, a cauliflower mosaic virus (CaMV) 35S (including the enhanced CaMV 35S promoter (see, for example U.S. Pat. No. 5,106,739)) or a CaMV 19S promoter or a cassava vein mosaic can be used. Other promoters that can be used for expression constructs in plants include, for example, zein promoters including maize zein promoters, prolifera promoter, Ap3 promoter, heat shock promoters, T-DNA 1′- or 2′-promoter of *A. tumefaciens*, polygalacturonase promoter, chalcone synthase A (CHS-A) promoter from petunia, tobacco PR-1a promoter, ubiquitin promoter, actin promoter, alcA gene promoter, pin2 promoter (Xu *et al.*, 1993), maize WipI promoter, maize trpA gene promoter (U.S. Pat. No. 5,625,136), maize CDPK gene promoter, and RUBISCO SSU promoter (U.S. Pat. No. 5,034,322) can also be used. Tissue-specific promoters, for example fruit-specific promoters, such as the E8 promoter of tomato (accession number: AF515784; Good et al. (1994)) can be used. Fruit-

specific promoters such as flower organ-specific promoters can be used with an expression construct of the present invention for expressing a polynucleotide of the invention in the flower organ of a plant. Examples of flower organ-specific promoters include any of the promoter sequences described in U.S. Pat. Nos. 6,462,185; 5,639,948; and 5,589,610. Seed-specific promoters such as the promoter from a β-phaseolin gene (for example, of kidney bean) or a glycinin gene (for example, of soybean), and others, can also be used. Endosperm-specific promoters include, but are not limited to, MEG1 (EPO application No. EP1528104) and those described by Wu *et al.* (1998), Furtado *et al.* (2001), and Hwang *et al.* (2002). Root-specific promoters, such as any of the promoter sequences described in U.S. Pat. No. 6,455,760 or U.S. Pat. No. 6,696,623, or in published U.S. patent application Nos. 2004/0078841; 2004/0067506; 2004/0019934; 2003/0177536; 2003/0084486; or 2004/0123349, can be used with an expression construct of the invention. Constitutive promoters (such as the CaMV, ubiquitin, actin, or NOS promoter), developmentally-regulated promoters, and inducible promoters (such as those promoters than can be induced by heat, light, hormones, or chemicals) are also contemplated for use with polynucleotide expression constructs of the invention.

[0040] Expression constructs of the invention may optionally contain a transcription termination sequence, a translation termination sequence, a sequence encoding a signal peptide, and/or enhancer elements. Transcription termination regions can typically be obtained from the 3′ untranslated region of a eukaryotic or viral gene sequence. Transcription termination sequences can be positioned downstream of a coding sequence to provide for efficient termination. A signal peptide sequence is a short amino acid sequence typically present at the amino terminus of a protein that is responsible for the relocation of an operably linked mature polypeptide to a wide range of post-translational cellular destinations, ranging from a specific organelle compartment to sites of protein action and the extracellular environment. Targeting gene products to an intended cellular and/or extracellular destination through the use of an operably linked signal peptide sequence is contemplated for use with the polypeptides of the invention. Classical enhancers are cis-acting elements that increase gene transcription and can also be included in the expression construct. Classical enhancer elements are known in the art, and include, but are not limited to, the CaMV 35S enhancer element, cytomegalovirus (CMV) early promoter enhancer element, and the SV40 enhancer element. Intron-mediated enhancer elements that enhance gene expression are also known in the art. These elements must be present

within the transcribed region and are orientation dependent. Examples include the maize shrunken-1 enhancer element (Clancy and Hannah, 2002).

[0041] DNA sequences which direct polyadenylation of mRNA transcribed from the expression construct can also be included in the expression construct, and include, but are not limited to, an octopine synthase or nopaline synthase signal. The expression constructs of the invention can also include a polynucleotide sequence that directs transposition of other genes, *i.e.*, a transposon.

IV. AGPase Subunit Expression in Plant Cells

[0042] The subject invention also concerns cells transformed with a polynucleotide of the present invention encoding a mutant AGPase small or large subunit polypeptide of the invention. In one embodiment, the cell is transformed with a polynucleotide sequence comprising a sequence encoding the amino acid sequences provided herein, or a functional fragment or variant thereof, or a fragment or variant thereof having AGPase activity. In a specific embodiment, the cell is transformed with a polynucleotide sequence encoding a variant provided herein. In one embodiment, the polynucleotide sequence is provided in an expression construct of the invention. The transformed cell can be a prokaryotic cell, for example, a bacterial cell such as *E. coli* or *B. subtilis*, or the transformed cell can be a eukaryotic cell, for example, a plant cell, and a yeast cell. Plant cells include, but are not limited to, dicotyledonous and monocotyledonous cells. Plant cells may be cells from rice, wheat, barley, oats, rye, sorghum, maize, sugarcane, pineapple, onion, bananas, coconut, lilies, millet, tomato, potato, sweet potato, pea, strawberry, beet, chickpea, watermelon, muskmelon, cassava, taro, sunflower, flax, and beans. In one embodiment, the plant cell is a cell from a *Zea mays* plant.

[0043] Techniques for transforming plant cells with a gene are known in the art and include, for example, *Agrobacterium* infection, biolistic methods, electroporation, calcium chloride treatment, PEG-mediated transformation, etc. Transformed cells can be selected, redifferentiated, and grown into plants that contain and express a polynucleotide of the invention using standard methods known in the art. The seeds and other plant tissue and progeny of any transformed or transgenic plant cells or plants of the invention are also included within the scope of the present invention.

[0044] The subject invention also concerns methods for expressing the polynucleotides of the invention in plant cells. These methods may include methods that alter the native polynucleotides present in the genome of the plant cells of interest. Methods to alter the polynucleotides of a genome of interest are known in the art and include meganucleases designed against the plant genomic sequence of interest (D'Halluin et al 2013 *Plant Biotechnol J* 11: 933-941); CRISPR-Cas9, TALENs, and other technologies for precise editing of genomes (Feng, *et al. Cell Research* 23:1229-1232, 2013, Podevin, *et al.* 2013 *Trends Biotechnology* 31: 375-383, Wei *et al.*, 2013 *J Gen Genomics* 40 : 281-289, Zhang et al 2013, WO 2013/026740); Cre-lox site-specific recombination (Dale *et al.* (1995) *Plant J* 7:649-659; Lyznik, *et al.* (2007) *Transgenic Plant J* 1:1-9; FLP-FRT recombination (Li *et al.* (2009) *Plant Physiol* 151:1087-1095); Bxb1-mediated integration (Yau *et al. Plant J* (2011) 701:147-166); zinc-finger mediated integration (Wright *et al.* (2005) *Plant J* 44:693-705); Cai et al. (2009) *Plant Mol Biol* 69:699-709); and homologous recombination (Lieberman-Lazarovich and Levy (2011) *Methods Mol Biol* 701: 51-65); Puchta, H. (2002) *Plant Mol Biol* 48:173-182). These and other similar techniques may be used to precisely edit the genome of a plant cell to produce a modified polynucleotide encoding an AGPase subunit polypeptide that contains a mutation of the present invention. Alternatively, random genomic mutations may be induced through the application of appropriate treatments including ethyl methanesulfonate (EMS) treatment (Qu et al 2014 Methods Mol Biol 1062: 225-239) to plant cells, sodium azide or methylnitrosourea (MNU) treatment (Sikora et al 2011 *Int J Plant Genomics* 2011 doi:10.1155/2011/314829) to plant cells, or exposure of plant cells to gamma rays or to fast neutron irradiation (Kodym and Afza 2003 *Methods Mol Biol* 236: 189-204) followed by the use of molecular methods to identify plant cells containing the appropriate mutations such that a variant AGPase subunit of the present invention is encoded by the genome of the plant cell.

[0045] The subject invention also concerns methods for producing a plant that exhibits increased heat stability relative to a wild type plant, wherein a polynucleotide encoding a mutant AGPase small or large subunit polypeptide of the present invention is introduced into a plant cell and the polypeptide(s) encoded by the polynucleotide(s) is expressed. In one embodiment, the plant cell comprises non-mutant genes encoding a wild type AGPase small or large subunit polypeptide. In another embodiment, the plant cell comprises at least one polynucleotide

encoding a mutant AGPase large subunit polypeptide. In a further embodiment, a polynucleotide encoding a mutant AGPase large subunit polypeptide is also introduced into a plant cell along with the polynucleotide encoding the mutant AGPase small subunit polypeptide. In one embodiment, the polynucleotide or polynucleotides is incorporated into the genome of the plant cell and a plant is grown from the plant cell. In a preferred embodiment, the plant grown from the plant cell stably expresses the incorporated polynucleotide or polynucleotides. In another embodiment, the polynucleotide or polynucleotides are expressed from self-replicating vectors derived from plant viruses. These vectors contain the sequences necessary for the vector to replicate within a plant cell as well as the sequences necessary to provide expression of the polynucleotides of interest. Vectors derived from plant viruses have been described, for example in Jung *et al* 2014 *Biotechnol Prog* 30: 905-915; Pflieger *et al* 2014 *BMC Plant Biol* 14: 232; Gu *et al* 2014 *Plant Biotechnol J* 12: 638-649; Huang *et al* 2010 *Plant Biotechnol J* 8: 783-795; Gleba *et al* 2014 *Curr Top Microbiol Immunol* 375: 155-192; and Lindbo 2007 *Plant Physiol* 145: 1232-1240. Self-replicating vectors derived from plant viruses may contain polynucleotides designed to upregulate or to downregulate an endogenous plant gene in the plant of interest, or may be used to express a novel heterologous gene. Expression from these vectors may be confined to the cells that were initially infected, or may be systemic throughout the infected plant. In a preferred embodiment, a polynucleotide encoding the variant AGPase subunit of the invention may be expressed from a virally-derived vector.

[0046] The subject invention also concerns methods for increasing starch synthesis in a plant or plant tissue (such as a plant seed or endosperm tissue). In one embodiment, a method of the invention comprises introducing one or more polynucleotides of the present invention into a plant. In certain embodiments, the polynucleotides introduced into the plant encode one or more polypeptides comprising the amino acid sequences provided herein, or a fragment or variant thereof. In certain embodiments, the polynucleotides introduced into the plant encode one or more polypeptides comprising the amino acid sequence of SEQ ID NO: 20 wherein the amino acid at position 161 has been mutated, wherein the amino acid at position 96 has been mutated, wherein the amino acid at position 443 has been mutated, or combinations, fragments, or variants thereof. In certain embodiments, the amino acids at positions 161 and 96 have been mutated, the amino acids at positions 96 and 443 have been mutated, or the amino acids at positions 161 and 443 have been mutated, or the amino acids at positions 161, 96, and 443 have been mutated, or

fragments or variants thereof. In some embodiments, the amino acid at position 161 has been mutated to be a G or Q, the amino acid at position 96 has been mutated to be a G or R, or the amino acid at position 443 has been mutated to be an R, including all possible combinations thereof.

[0047] In certain embodiments, the polynucleotides introduced into the plant encode one or more polypeptides comprising a fragment or variant comprising a conserved motif having SEQ ID NO: 21. In other embodiments, position 3 of SEQ ID NO: 21 is an uncharged polar amino acid. In yet other embodiments, position 3 is G or Q.

[0048] In certain embodiments, the polynucleotides introduced into the plant encode one or more polypeptides provided herein and comprising a conserved motif having SEQ ID NO: 22. In other embodiments, position 8 of the motif is an uncharged polar amino acid or a basic amino acid. In yet other embodiments, position 8 of the motif is a G or R.

[0049] In yet other embodiments, the polynucleotides introduced into the plant encode one or more polypeptides comprising a conserved motif having SEQ ID NO: 23. In other embodiments, position 7 of the motif is a basic amino acid. In further embodiments, position 7 of the motif is an R.

[0050] In one embodiment, the polynucleotide is stably incorporated into the genome of the plant or plant tissue. The polynucleotide can comprise regulatory elements, such as a promoter and/or enhancer sequences, that provide for increased expression of the polynucleotide and/or the polypeptide encoded thereby. In a specific embodiment, the promoter sequence is one that provides for constitutive or tissue-specific (e.g., endosperm) expression. Plants or plant tissues containing the polynucleotide, or progeny of the plants, optionally can be screened for increased expression of a polynucleotide or polypeptide of the invention. In one embodiment, multiple copies of one or more polynucleotides of the invention are introduced into a plant or plant tissue and stably incorporated into the genome of the plant. In one embodiment, a polynucleotide of the invention is provided in an expression construct as described herein.

[0051] The invention further provides plants comprising the polynucleotides, polypeptides, and expression constructs disclosed herein. Plants of the present invention may be

monocots or dicots, and may include, for example, rice, wheat, barley, oats, rye, sorghum, maize, sugarcane, pineapple, onion, bananas, coconut, lily, turf grasses, and millet plants.

[0052] The following examples are included to demonstrate examples of certain preferred embodiments of the invention. It should be appreciated by those of skill in the art that the techniques disclosed in the examples that follow represent approaches the inventors have found function well in the practice of the invention, and thus can be considered to constitute examples of preferred modes for its practice. However, those of skill in the art should, in light of the present disclosure, appreciate that many changes can be made in the specific embodiments that are disclosed and still obtain a like or similar result without departing from the spirit and scope of the invention.

Example 1

V. Plasmid preparation

[0053] The small and large subunit plasmids of maize endosperm AGPase (pMONc*Bt2* and pMONc*Sh2*, respectively) (Giroux *et al.,* 1996) were expressed in *E. coli* AC70R1-504 cells (Iglesias *et al.,* 1993) as previously described (Boehlein *et al.,* 2008).

B-factor analysis of the potato small subunit homotetramer x-ray crystal structure

[0054] The coordinate file for 1YP2 was manually edited to create four separate files, one for each of the four monomer chains. Each was analyzed individually by the B-FITTER program, which ranked the residues in decreasing order of average B-factor (calculated as a mean value for each residue). To avoid artifacts caused by crystal packing forces, a residue was considered to have an abnormally high B-factor only if the value was ≥55 in each of the four monomers. These criteria yielded nine residues in the structure of 1YB2. The corresponding amino acids in both the small and large subunit of maize endosperm AGPase were determined by using the previously-published alignment and the residues are shown in **Table 3**.

[0055] The ISM (Iterative Saturation Mutagenesis) mutants were prepared by QuikChange site-directed mutagenesis (Agilent) using pMONc*Sh2* as template in the first iteration. pMONc*Sh2*Q96G, pMONc*Sh2*Q96R, pMONc*Sh2*Q96G;D161G,

pMONc*Sh2*Q96R;D161G, pMONc*Sh2*Q96G;D161G:A443R, and pMONc*Sh2*Q96R;D161G:A443R were used as templates in subsequent mutagenesis reactions.

VI. Library screening

[0056] AC70R1-504 transformants were grown overnight in 96-deepwell plates, then 3 μL aliquots of each culture were plated on square Petri dishes containing 2% (w/v) glucose (48 clones / plate, which includes one wild-type (*Bt2*/*Sh2*) and MP-TI/*Sh2* control per plate). One plate was incubated at 37°, the other at 42°C for 7.5 hr. Plates were cooled to room temperature before being placed into a glass chamber containing iodine crystals. After staining, plates were immediately photographed on a light table (to ensure even illumination) and the color intensity of each colony (integrated over its entire surface) was evaluated using the ImageJ program. Staining was calculated as a percentage relative to the wild type control on that particular plate and positives were defined as colonies with more intense color than wild-type.

[0057] After the initial screening and selection of the first mutant, subsequent iterations used the selected mutant as template in individual QuikChange site-directed mutagenesis reactions. The resulting DNAs from mutagenesis at each site were pooled, used to transform AC70R1-504 cells already containing the pMONc*Bt2* plasmid and grown on LB + 2% glucose + 50 μg/ml Kan + 75 μg/ml Spec plates at 42 °C overnight. Darkest staining colonies on each plate were selected. The selections were respotted on two identical LB + Kan + Spec plates. One of these plates was incubated at 37 °C and the other at 42 °C overnight. The resulting colony spots were exposed to iodine vapors. DNA was prepared for the darkest staining 42 °C grown selections and sequencing was performed.

VII. Plate capture assay

[0058] Following the first iteration, dark staining colonies were screened for heat stability using a plate capture assay. This assay utilized a monoclonal antibody specific for the large subunit of AGPase that pulled down both AGPase subunits and retained activity. Thus, large samples of clarified lysates could be analyzed for AGPase activity in this manner. Initial heat stability was assessed by growing each mutant enzyme at 37 °C in 25 ml cultures to an OD of 0.7-1.0 and induced for 2-4 hours at room temperature. Cultures were then pelleted and stored at -80 °C. Pellets were resuspended in 1 ml of glycerol buffer (0.1M HEPES pH 7.4, 5mM

$MgCl_2$, 10mM $KH_2PO_4^-$, 5mM EDTA, 5% glycerol) and disrupted by sonication. Cell debris was removed by centrifugation at 14000 rpm for 10 minutes. Protein was adjusted to 2.5mg/ml (Biorad protein reagent) and stored at -80 °C until use. Preparations made in this way were stable for several days.

[0059] A high binding 96 well plate (96 well EIA/RIA plate, Costar) was prepared for the plate capture assay by coating the appropriate number of wells with 200 µl of Antibody Ab 12.1 (diluted 1:1000 in 1X phosphate buffered saline, PBS). The plate was incubated at 4 °C with shaking overnight then washed three times with PBST (1X PBS with 0.05% Tween20) and once with 1X PBS. Clarified lysates (200 µl of 5 mg/ml) were loaded on the Ab coated plate and incubated for 1 hour at 4 °C with shaking. After 1 hour the plate was washed 4X with cold glycerol buffer and placed on ice. Assay mix (250 µl) was added to each well of the plate and incubated at 55 °C for 0, 10 and 60 min. Reactions were terminated by removing 200 µl of assay mix from plate at the appropriate time and placed in cuvettes. Development mix (300 µl) was added to the cuvettes and then measured A340. Activity was determined by calculating the activity between 10 and 60 min. Initial rates (0-10 min) were discarded, as the reaction mixes needed time to reach 55 °C. Reaction rates between 10-60 min were shown to be linear with time.

VIII. Determination of apparent kinetic constants, K_m, K_{cat} and K_a

[0060] An endpoint assay was used to determine the rate of the AGPase reaction by coupling the formation of PPi to a decrease in NADH concentration (Boehlein *et al.,* 2008). Reaction mixtures (300 µl) contained the following components when saturating: 50 mM HEPES pH 7.4, 15 mM $MgCl_2$, 2.0 mM ATP, and 2.0 mM G-1-P and 5 mM 3-PGA. The Michaelis constants (K_m) were obtained by varying one substrate in the presence of saturating concentrations of the second substrate. The activation constants were obtained by varying the activator at fixed levels of both substrates. All reactions were performed at 37 °C and mixes were prewarmed before starting the reaction with enzyme. Reactions were terminated by boiling the reaction mixes for 2 minutes followed by development with 200 µl of coupling reagent (25 mM imidazole pH 7.4, 4 mM $MgCl_2$, 1 mM EDTA, 0.2 mM NADH, 0.725U aldolase, 0.4U triose phosphate isomerase, 0.6U glycerophosphate dehydrogenase, 1 mM fructose 6-phosphate and 0.8 µg purified PPi-PFK per reaction). The absorbance at 340 nm was monitored after 30

min and the amount of PPi produced was determined from a standard curve using PPi in complete reaction mixtures lacking AGPase. Blank samples contained complete reaction mixtures without enzyme and the change in absorbance between the blank and the reaction was used to calculate the amount of PPi produced. Reactions were linear with time and enzyme concentration.

IX. Activity at 55 °C

[0061] Activity measurements at increased temperature were performed by looking at the rate of the reaction over a period of 10 minutes. Each time course contained an assay mix (saturating substrates and activator) of 1300 µl and was started with the addition of enzyme. At appropriate times, 300 µl of the reaction was removed and boiled for 2 minutes followed by development (see above). A rate at 55 °C was calculated for those enzymes that were linear with time from 2.5-10 minutes.

X. Thermodynamic stability of purified maize AGPase and ISM mutants

[0062] Thermodynamic stability of the purified AGPase and ISM mutants was assessed by pre-treating 10 µl of the enzyme preparation (36 ng of enzyme with 0.5 mg/ml BSA) at 37 °C for 0-7.5 min and then returning them to ice. The activity of the enzyme was then determined at 37 °C, in the presence of saturating concentrations of substrates and activator, as described above. The percentage of activity remaining was calculated by dividing the activity of the treated sample from the untreated sample. The ½ time of inactivation ($T_{1/2}$) was calculated from plots of logarithmic plots of % activity versus time whereby the slope was equal to –k/(2.3). $T_{1/2}$ was calculated from the equation $k = 0.693/t_{1/2}$.

XI. Determination of kinetic constants in the absence of 3-PGA

[0063] In the absence of 3-PGA, the ATP and G-1-P K_m's were determined by varying ATP from 0.25-7.5 mM at fixed concentrations of G-1-P, (5.0-30 mM). Reactions were performed for 10 min and terminated by boiling. Blank mixtures contained everything except ATP.

XII. Pi inhibition of GGR

Table 2. PPi inhibition as carried out in the presence or absence of 3-PGA.

Varying substrate	Activator (3-PGA) 5 mM	3-PGA Concentration when varied	Pi concentration when varied	Pi Concentration when held constant
ATP	-	[a]0.1-3	10-30	0.4
G-1-P	-	5-30	5-50	0.2
ATP	+	0.025-1.5	0.5-25	0.25
G-1-P	+	0.025-1.5	1-25	5

[a] all units are in mM

XIII. Data analysis

[0064] The apparent Michaelis constant and maximum velocity of each reaction was calculated from a V vs. S plots using nonlinear regression analysis and the following equation, $v = V_{max}S/(K_m+S)$, where v is the measured velocity, V_{max} is the maximum velocity, S is the substrate concentration. The activation constant (K_a) was obtained using the following equation: $V=V_{min}+V_{max}*(X)\ /(K_a+X)$, where V_{min} is the velocity in the absence of activator and V_{max} is the change in activity from V_{min} to the total velocity. X is the activator concentration, K_a is the activation constant. All linear regression was carried out using the software program Prism (Graph Pad, San Diego CA).

[0065] Kinetic data for the determination of the K_m value in the absence of 3-PGA for D161G and GGR and were fitted to equations using the nonlinear equations using GraphPad Prism 4.0c. Initial velocity data were fit to Eq. (1), which describes a sequential mechanism where v is the measured reaction velocity, V is the maximal velocity, A and B are the concentrations of substrates (ATP and G-1-P), K_a, and K_b are the corresponding Michaelis–Menten constants, and K_{ia} is the dissociation constant for substrate A. The concentration of ATP was varied from 0.25 - 7.5 for ATP and 5.0 - 30 mM for G-1-P.

$$(1) \qquad v = VAB/(K_{ia}K_{mb}+K_{mb}A+K_{ma}B+AB$$

XIV. Pi inhibition data

[0066] Phosphate inhibition data were fitted to equations 1 and 2, which correspond to noncompetitive inhibition (NC; Eq. 2) or linear mixed type inhibition (L-MT; Eq. 3) using GraphPad Prism software. v is the measured velocity, V_m is the maximum velocity, S is the substrate concentration, K_i is the inhibition constant, K_s is the dissociation constant for the ES complex, I is the inhibitor concentration, α is the factor by which K_i changes when the inhibitor is present.

$$(2) \qquad v = V_m\,(S/K_s)/(1 + S/K_s + I/K_i + SI/K_iK_s) \qquad \text{NC}$$

$$(3) \qquad v = V_m\,(S/K_s)/(1 + S/K_s + I/K_i + SI/\alpha K_iK_s) \qquad \text{L-MT}$$

Example 2

B-factor analysis

[0067] Crystallographic B-factors indicate uncertainties in atomic positions, which often correlate with conformational disorder. Since no three-dimensional structure of a heterotetrameric plant AGPase is currently available, the potato small subunit homotetramer crystal structure (Jin *et al.,* 2005) was used in the methods of the present invention. One monomer of this homotetramer is illustrated in Boehlein *et al* 2015 *Arch Biochem Biophys* 568: 28-37 with the relative B-factors indicated both by chain width (wider = higher B-factor) and color (red = highest B-factor). B-FITTER (Reetz *et al.*, 2007) was used to rank each residue on the basis of their average B-factor in each of the four homotetramer chains. To identify those residues most likely to affect protein structure, those residues with B-factors $\geq$ 55 in each of the four monomers were selected. The positions of these residues in the potato homotetramer crystal structure are shown with their counterparts in the maize large subunits in **Table 3**.

Table 3. Maize endosperm AGPase residues suspected to have unusually high relative B-factors by analogy with the potato small subunit homotetramer x-ray crystal structure (Jin *et al.* 2005) and results of site-saturation mutagenesis libraries at these positions.

Potato ss residue	Residue mean B-factor[a]	Maize ls equivalent	Q_{codon} value; Predicted number of amino acids	Number of positives at 37ºC / 42ºC
Arg 33	62.36	Gln 96	0.85; 18 - 19	9 / 4
Glu 99	74.38	Asp 161	0.80; 17 - 18	2 / 9
Asp 218	71.25	Ile 285	0.80; 17 - 18	2 / 0
Lys 222	68.41	Gln 289	0.81; 17 - 18	1 / 3
Arg 307	65.10	Pro 372	0.83; 18 - 19	2 / 12
Asp 375	74.12	Glu 440	0.87; 19 - 20	10 / 2
Arg 378	67.55	Ala 443	0.85; 18 - 19	2 / 2
Lys 379	67.03	Ser 444	0.87; 19 - 20	9 / 8
Leu 380	66.24	Lys 445	0.84; 18 - 19	9 / 4

[a]The average mean B-factor for all residues in the potato small subunit homotetramer was 34.1.

[0068] All nine of the conformationally disordered residues in the potato small subunit occurred on surface loops. Given the high degree of sequence conservation between the potato and maize AGPase sequences, it is likely that the overall structures are highly similar. Conformationally mobile residues in the potato structure were therefore evaluated in the maize counterpart as being possibly disordered.

Example 3

Site-saturation library construction and evaluation

[0069] Nine random replacement libraries were constructed for the maize large subunit residues identified by B-factor analysis. A polymerase chain reaction (PCR)-based methodology was used (Sullivan *et al.*, 2013) to provide high-quality libraries with little contamination by the parent sequence and undetectable levels of concatameric primer inserts which would otherwise dilute the library and add to the screening burden. An offset primer design strategy made the PCR amplifications exponential, which enhanced transformation yield. This contrasts with conventional primer design methods using primers that overlap completely and result in linear, rather than exponential, amplifications that yield correspondingly smaller quantities of DNA for

Escherichia coli (*E. coli*) transformation. Primer mixtures encompassed an NNK mixture of bases at the targeted codon, wherein N is an equal representation of A, C, G and T, K is an equal representation of G and T and M is an equal representation of A and C. This 32-member mixture encodes all 20 canonical amino acids as well as one stop codon, ensuring that all amino acid replacements were interrogated at the targeted positions of the AGPase large subunit. Primer sequences are shown in **Table 4**.

Table 4. Primers used in site-directed mutagenesis in the first iteration.

Mutation	Direction	SEQ ID	Primer sequence
Q96X	For	1	GCA CTG GAT CTN NKC TCT TTC CTC TGA CAA GCA CAA GAG C
Q96X	Rev	2	AGA GGA AAG AGM NNA GAT CCA GTG CCT CCG CCC AAA ATG A
D161X	For	3	TCA ACT TTG CTN NKG GAT CTG TAC AGG TAT TAG CGG CTA C
D161X	Rev	4	TGT ACA GAT CCM NNA GCA AAG TTG ATC CCG CCT TCA AGG T
I285X	For	5	TGA GCT ATG CTN NKG ATG ATG CAC AGA AAT ATC CAT ACC T
I285X	Rev	6	TGT GCA TCA TCM NNA GCA TAG CTC AGG AAG TTG GTC TCA A
Q289X	For	7	TAG ATG ATG CAN NKA AAT ATC CAT ACC TTG CAT CAA TGG G
Q289X	Rev	8	TAT GGA TAT TTM NNT GCA TCA TCT ATA GCA TAG CTC AGG A
P372X	For	9	ATT TTT ACG ATN NKA AAA CAC CTT TCT TCA CTG CAC CCC G
P372X	Rev	10	AAA GGT GTT TTM NNA TCG TAA AAA TCA AAC TTG GAA GGC T
E440X	For	11	TCT ATG AAA CTN NKG AAG AAG CTT CAA AGC TAC TGT TAG C
E440X	Rev	12	GAA GCT TCT TCM NNA GTT TCA TAG ATG TCC GCT CCC ATC A
A443X	For	13	CTG AAG AAG AAN NKT CAA AGC TAC TGT TAG CTG GGA AGG T
A443X	Rev	14	AGT AGC TTT GAM NNT TCT TCT TCA GTT TCA TAG ATG TCC G
S444X	For	15	AAG AAG AAG CTN NKA AGC TAC TGT TAG CTG GGA AGG TCC C
S444X	Rev	16	AAC AGT AGC TTM NNA GCT TCT TCT TCA GTT TCA TAG ATG T
K445X	For	17	AAG AAG CTT CAN NKC TAC TGT TAG CTG GGA AGG TCC CAG T
K445X	Rev	18	GCT AAC AGT AGM NNT GAA GCT TCT TCT TCA GTT TCA TAG A

N is an equal representation of A, C, G and T; K is an equal representation of G and T; M is an equal representation of A and C.

[0070] The nucleotide diversities at the targeted positions were evaluated for each library by sequencing the collection of pooled plasmid DNAs obtained after the initial PCR product had been used to transform an *E. coli* host. As expected, the fluorescence sequencing chromatograms showed unique DNA sequences at all positions except for those targeted for randomization. The relative peak heights for each base (which correspond to the relative

nucleotide compositions of the mixed plasmid population) were calculated for each of the targeted bases at the appropriate locations on the chromatograms. These were used to calculate values for Q_{codon} and used to estimate the number of amino acids likely to be present at each randomized position (Sullivan *et al.,* 2013). A Q_{codon} value of 1.0 represents perfect randomization while 0.0 indicates no sequence diversity. Libraries predicted to contain fewer than 17 of the 20 possible amino acids on the basis of low Q_{codon} values were rejected and re-constructed until suitably diverse populations were obtained.

[0071] Once site-saturation libraries had been constructed and validated for all nine targeted positions, the plasmid mixtures were used to transform *E. coli* AC70R1-504 cells (a *glgC*$^-$ mutant unable to accumulate glycogen due to loss of the host AGPase) so that the thermal stabilities of the AGPase variants could be assessed on the basis of glycogen accumulation at elevated growth temperatures. The AC70R1-504 cells had previously been transformed with a second plasmid that encoded the wild-type maize endosperm AGPase small subunit, allowing expression of heterotetramers. A number of assay parameters were optimized in preliminary studies by methods known in the art including growth times, glucose concentrations present in the growth medium, growth temperature, etc.

[0072] We evaluated 92 randomly-chosen clones from each library in order to have a 95% chance of covering all variants present. Transformants were grown in the presence of glucose and the accumulated glycogen was revealed by incubating plates in the presence of iodine vapor. Two examples each of wild-type AGPase and a known thermostable variant (MP-TI) were included as controls. Since the darkness of iodine staining correlates with the amount of accumulated glycogen, digital photographs of plates were analyzed quantitatively by the ImageJ program (Image Processing and Analysis in Java; National Institute of Mental Health, Bethesda, Maryland, USA) to provide relative color density for each colony. These were used to calculate a relative color density for each library colony after dividing by the intensity of the positive control on the same plate. These assays were carried out at both 37ºC and 42ºC to provide thermal stabilities under challenging or extremely challenging temperatures. The numbers of clones showing significantly greater staining than the wild-type controls are collected in **Table 3**. These positives were examined by DNA sequencing and several variants were studied in detail following large-scale expression and purification. Following this first round of

mutagenesis, Q96G, S444R, D161G, D161Q, E440A and E440R were purified and their ability to catalyze the reaction at 55°C analyzed. This was to be sure that the dark staining at 42°C correlated with activity of the purified enzyme at increased temperatures. As seen in **Table 5**, only changes at the 161 position conferred increased activity at 55°C. The change to glutamine offered some stability, while the change to glycine significantly increased the ability of the enzyme to operate at a much higher temperature (**Table 5**).

Table 5. Activity at 37 °C and 55 °C of purified first iteration mutants.

Enzyme	Activity 37 °C*	Activity 55 °C*	%Activity 55 °C/37 °C
Wt	15.2	0	-
Q96G	6.7	1.35	20.2
D161G	13.1	7.8	60
D161Q	10.6	1.0	9.4
E440A	9.3	0.1	1.1
A443R	7.9	0	0
S444R	27.8	0.4	1.4

*Activity is presented in μmol/min/mg

[0073] The templates used for the next iteration of saturation mutagenesis were the double mutants Q96G;D161G (GG) and Q96R;D161G (RG). The Q96 mutants were chosen because they had the darkest colony staining. The D161G change was selected because of its increased 55°C activity.

[0074] For the second iteration of saturation mutagenesis, QuikChange mutagenesis (Agilent) was performed at the remaining seven sites. Primer sequences are shown in **Table 6**. Three of the RG new mutant sites (I285, Q289, and P372) were transformed into AC70R1-504 cells containing the wile-type small subunit (*Bt2*). Colonies staining darker than the RG colonies were selected at the P372, I285 and Q289 locations. The DNA from seven dark staining RG;I285NNK colonies and nine RG;Q289NNK colonies was sequenced to determine the highest probability of change, but results were scattered, thus no particular amino acid change was dominant at these positions. In addition, no dark staining colonies were seen at position P372. Next, E440, A443, S444 and K445 sites were individually mutated, the DNA of each mutagenesis reaction was extracted, and combined for transformation into *Bt2* competent cells.

The darkest staining colonies were then re-plated and the DNA was sequenced from the twenty darkest staining colonies on these plates. The sequencing results showed that out of 20 colonies, several changes occurred one time, (A443V, S444Q, S444R, S444G), one change occurred four times (S444E) and one change appeared 12 times (A443R). The selection technique demonstrated that while discretion in picking darker stained colonies by eye was not perfect, the approach worked considerably well if multiple colonies were screened. Thus, the mutant Q96R;D161G;A443R (RGR) was selected for the next iteration.

Table 6. Primers used in site-directed mutagenesis after the first iteration.

Mutation	Direction	SEQ ID	Primer sequence
D161G	For	24	GCGGGATCAACTTTGCTGGTGGGTCTGTACAGGTATTAGCGG
D161G	Rev	25	CCGCTAATACCTGTACAGACCCACCAGCAAAGTTGATCCCGC
I285X	For	26	GAGACCAACTTCCTGTCCTATGCTNNKGATGATGCACAGAAATATCC
I285X	Rev	27	GGATATTTCTGTGCATCATCMNNAGCATAGGACAGGAAGTTGGTCTC
Q289X	For	28	GCTATGCTATAGATGATGCTNNKAAATATCCATACCTTGCATC
Q289X	Rev	29	GATGCAAGGTATGGATATTTMNNAGCATCATCTATAGCATAGC
P372X	For	30	CCTTCCAAGTTTGATTTTTACGACNNKAAAACACCTTTCTTCACTGCACCC
P372X	Rev	31	GGGTGCAGTGAAGAAAGGTGTTTTMNNGTCGTAAAAATCAAACTTGGAAGG
E440X	For	32	GGAGCGGACATCTATGAAACTNNKGAAGAGGCTTCAAAGCTACTGTTAGC
E440X	Rev	33	GCTAACAGTAGCTTTGAAGCCTCTTCMNNAGTTTCATAGATGTCCGCTCC
A443X	For	34	CGGACATCTATGAAACTGAAGAAGAANNKTCAAAGCTACTGTTAGCTGG
A443X	Rev	35	CCAGCTAACAGTAGCTTTGAMNNTTCTTCTTCAGTTTCATAGATGTCCG
S444X	For	36	CATCTATGAAACTGAAGAAGAAGCGNNKAAGCTACTGTTAGCTGGGAAGG
S444X	Rev	37	CCTTCCCAGCTAACAGTAGCTTMNNCGCTTCTTCTTCAGTTTCATAGATG
K445X	For	38	CATCTATGAAACTGAAGAAGAAGCGTCANNKCTACTGTTAGCTGGGAAGGTC
K445X	Rev	39	GACCTTCCCAGCTAACAGTAGMNNTGACGCTTCTTCTTCAGTTTCATAGATG

N is an equal representation of A, C, G and T; K is an equal representation of G and T; M is an equal representation of A and C.

[0075] A similar technique was used with GG as the template. QuikChange saturation mutagenesis was performed at each of the seven sites and the resulting mutant DNAs were combined. AC70R1-504 cells containing *Bt2* were transformed with the combined mutant DNAs, grown overnight at 42°C on glucose containing LB plates, and ensuing colonies were exposed to iodine vapors. Here, 36 dark staining colonies were selected and re-plated on an LB-glucose plate grown at 42 °C. DNA was prepared from the six darkest staining streaks and

sequenced. Mutations occurred at only 2 of the 7 positions, 2 of which resulted in S444E and the remaining was A443R. A443R was selected since this mutation was obtained 2/3 of the time. Thus, the two clones moving on to the next round of iterations were RGR and GGR.

[0076] The fourth iteration of saturation mutagenesis on both templates was performed using the same protocol. Following staining, six GGR;NNK colonies were selected based on their staining and the DNA sequenced. The results of the sequencing data indicated that the differences in staining were no longer visible, thus out of 6 colonies sequenced, 4 resulted in template, 1 had a change of Q289A and one had a change of I285V. Similarly when RGR was used as a template, 3 RGR;NNK clones were sequenced (1 template and 2 S444R). Since the staining technique was no longer able to discriminate between the template and the additional iteration, no additional iterations were performed.

[0077] The activity of each mutant, single to quadruple mutants, at 55°C was then determined using an antibody capture plate assay. This assay was developed so that large numbers of samples could be rapidly purified and assayed. The antibody used has a unique property in that when AGPase is bound to it, it retains activity; however, the complex Ab:AGPase is more heat stable than the AGPase in solution. Plate assays revealed that while the wild-type (wt) enzyme had activity at 55 °C when bound to the antibody, it was not linear with time, and decreased significantly after the first 30 minutes of incubation. Both double mutants, GG and RG, had activity linear with time for 50 minutes. The plate assay was also used to evaluate the activity of the bound AGPase in the presence and absence of 3-PGA. Here it is shown that, both double mutants had appreciable activity in the absence of 3-PGA (**Table 7**), an unexpected result that links the heat stability and allosteric regulatory properties together.

Table 7. Plate assay of measured velocity (v) vs. time at 55 °C and activity in the presence and absence of 3-PGA.

Mutant	55° Activity (+3-PGA)			37° Activity (+3-PGA) 30 min	37° Activity (-3-PGA) 50 min
	10-30 min	30-50 min	10-50 min		
Wt	**88**	**44**	**132**	**71**	**4**
GG	175	143	318	76	56
GGR	110	109	219	65	56
GGRA	111	151	262	86	52
GGRV	113	114	227	78	40
RG	107	139	246	74	95
RGR	114	141	255	96	115
RGRR	121	143	264	43	67
MP	**97**	**132**	**229**	**76**	**75**

GG = Q96G, D161G; GGR = Q96G, D161G, A443R; GGRA = Q96G, D161G, A443R, Q289A; GGRV = Q96G, D161G, A443R, I285V; RG = Q96R, D161G; RGR = Q96R, D161G, A443R; RGRR = Q96R, D161G, A443R, S444R

Example 4

Kinetic characterization of the iterative saturation mutants

[0078] Each mutant, comprising 1, 2, 3 or 4 changes, was grown in a large scale culture, expressed and purified according to standard procedures (Boehlein *et al.,* 2005). The K_m for ATP and G-1-P, in the presence of the activator, 3-PGA, was determined by varying one substrate while the co-substrate was held constant and saturating. As shown in **Table 8**, in the presence of 3-PGA, the K_m for both ATP and G-1-P for each of the mutants is similar to wt maize, with the largest variation in K_m being approximately two-fold. Thus, the effects of these mutations on the kinetic constants are minimal. Next, the activation constant, K_a, was determined for each of the mutant enzymes. Here a decrease in the K_a for each of the mutants of approximately 3-6 fold was observed. This property appears to stem from the D161G change, and includes all enzymes which carry this mutation (**Table 8**). To be certain the effect is the result of the D161G mutation, and not the result of other single mutant changes, Q96R and Q96G were purified and characterized. As shown in **Table 8**, the kinetic constants for Q96G and Q96R are more closely related to wt maize, exhibiting less than a 2-fold change in K_a for 3-PGA. In addition, there was little fluctuation in the K_m values for ATP and G-1-P. Thus it appears that

the increased affinity for 3-PGA can be attributed to the D161G variant alone and the effect is retained when additional mutations are pyramided with D161G.

Table 8. Kinetic constants of ISM mutants.

Enzyme	ATP K_m	G-1-P K_m	3-PGA K_a	V_{max}	V_{min}
D161G	0.044	0.060	0.039	18.6	1.1
RG	0.063	0.038	0.034	17.8	3.2
RGR	0.062	0.028	0.039	13.7	2.0
RGRR	0.038	0.025	0.039	12.8	3.8
GG	0.031	0.021	0.040	13.3	0
GGR	0.057	0.038	0.056	31.4	4.1
GGRA	0.057	0.037	0.044	24.8	4.2
GGRV	0.046	0.024	0.062	19.1	1.7
Q96R	0.072	0.032	0.13	16.0	**0**
Q96G	0.055	0.064	0.19	19.5	0
Wt	**0.069**	**0.039**	**0.22**	**22.6**	0
T142F	0.2	0.04	0.03	17.8	5.1
SH2-E	0.10	0.03	0.07	16.6	4.2

RG = Q96R, D161G; RGR = Q96R, D161G, A443R; RGRR = Q96R, D161G, A443R, S444R; GG = Q96G, D161G; GGR = Q96G, D161G, A443R; GGRA = Q96G, D161G, A443R, Q289A; GGRV = Q96G, D161G, A443R, I285V

[0079] Next, the activity in the absence of 3-PGA was determined. **Table 8** shows both the activity in the presence and absence of the activator. Each of the ISM mutants have a substantial amount of activity in the absence of 3-PGA, with the exception of GG (**Table 8**). This trait appears to also stem from the D161G mutation, since the Q96R and Q96G single mutants do not share this property. As several mutations are pyramided together, an additive effect of this 3-PGA independent activity is seen. Interestingly, the additional mutation of A443R to GG in the third iteration, resurrects the activity in the absence of 3-PGA and dramatically increases the V_{max} in the presence of 3-PGA three-fold (13.3 μmol/min/mg (GG) to almost 31.4 μmol/min/mg (GGR) **Table 8**). This is the only ISM mutant which yielded an increase in the catalytic rate.

Example 5

Kinetic and thermodynamic stability

[0080] While properties that include increased catalytic activity and activator-independent activities are of great agricultural importance, the main goal here was to create a variant that would remain active at increased temperatures. Two different methods were used to screen temperature effects on the purified ISM mutants. First, the kinetic stability of the enzyme at 55 °C was examined. This temperature rendered the wt maize enzyme completely inactive, while other previously constructed variants had substantial activity. Second, the thermodynamic stability of the enzymes was studied by pre-incubation of the naked enzymes (in the absence of substrates, activators and products) for various times at 37 °C, followed by activity measurements at 37 °C.

[0081] **Table 9** displays the results of these 2 experiments. First, it is shown that D161G retains approximately 50% of its activity when assayed at 55°C for 15 minutes (**Table 9**). While the D161G mutant has considerable activity at 55 °C, its activity at 37 °C is slightly lower (84%) than the wild-type activity. Since all the other mutations stemmed from D161G, the activity in each subsequent mutation would be predicted to have this partial activity, unless the sites interact. This in fact was the case for each of the changes, except for the triple mutant, GGR. The addition of A443R to the GG mutant yielded an enzyme with considerably higher activity at 55 °C and surpassed the activity of the wt enzyme at 37 °C. While some mutants render a higher percentage of activity, when comparing 55°C/37°C the triple mutant has by far the highest activity in both circumstances.

Table 9. Velocity of ISM mutants at 37 °C and 55 °C.

Enzyme	V 37 °C	V 55 °C	55 °C/37 °C	T1/2
D161G	16.0	8.9	56	4.4
RG	17.7	14.2	80	5.2
RGR	11.2	6.6	58	5.0
RGRR	9.2	8.5	92	6.8
GG	12.1	8.0	66	7.2
GGR	**26.4**	**21.7**	**82**	**5.7**
GGRA	19.2	15.4	80	7.2
GGRV	12.6	11.6	94	5.5
Wt	**19.2**	**0**	**0**	**6.0**
T142F	21.9	6.0	28	4.6
SH2-E	22.4	24.3	111	3.5
MP + SH2-E	18.4	23.6	128	46

RG = Q96R, D161G; RGR = Q96R, D161G, A443R; RGRR = Q96R, D161G, A443R, S444R; GG = Q96G, D161G; GGR = Q96G, D161G, A443R; GGRA = Q96G, D161G, A443R, Q289A; GGRV = Q96G, D161G, A443R, I285V

Table 10. Kinetic constants in the absence of 3-PGA

Enzyme	ATP K_m	G-1-P K_m	k_{cat}	K_{ia}
[a]**Maize-wt**	[c] 4.0	2.8	35.2	ND
D161G	0.043	6.8	48.8	1.4
GGR	0.10	8.74	72.4	0.72
[b] **T142F**	1.73	0.09	53.8	1.5
[b] **SH2-E**	2.53	0.03	67.6	5.4

V_{max} to k_{cat} conversion, multiply v_{max} (µmol) by 3.5

K_m and K_{ia} values are presented in mM; k_{cat} values are presented in s^{-1}

[a] data from Boehlein *et al.*, 2010

[b] data from Boehlein *et al.*, 2013

[c] ATP K_m could not be determined because G-1-P was not saturating. Since G-1-P saturation was not reached, value reported is the apparent K_m, in the presence of 15 mM G-1-P.

[0082] The thermodynamic stability was also determined for each of the mutants by incubating the enzymes for 1-15 minutes at 37°C, returning them to ice, and then assaying according to standard protocols. Control samples were maintained on ice, and the half-life of inactivation was determined from a plot of the log % activity vs. time. Interestingly, although all of the ISM mutants had a substantial activity at 55 °C, none of the mutants had an increase in their half-life when compared to the wt maize enzyme (**Table 9**).

Example 6

Selection of mutants for further evaluation

[0083] Based on the kinetic results (**Table 10**) and catalytic properties at 55°C (**Table 10**) both D161G and GGR were selected for a more detailed kinetic analysis. D161G was selected because this single mutation was the foundation for every other mutation, and had superior qualities, including a lower K_a for 3-PGA, activity in the absence of 3-PGA and activity at 55°C when compared to the maize wt enzyme. The GGR mutant was primarily chosen based on its increased catalytic rate at 37 °C and high performance at 55 °C. Its high activity in the absence of 3-PGA was also paramount. The first kinetic parameters determined were the Michaelis constants, the catalytic rate (k_{cat}) and the dissociation constant (K_{ia}) for the E-ATP complex, in the absence of 3-PGA. Here, ATP was varied at several fixed concentrations of G-1-P, in the absence of the activator, 3-PGA. For the wt maize enzyme, this activator primarily responds by reducing the K_m for both substrates, with little effect on the reaction rate (Boehlein *et al.,* 2013). Here it is shown that the single mutation of D161G retains a high G-1-P K_m, and has a catalytic rate similar to the wt maize enzyme in the absence of the activator (**Table 10**). Interestingly, the K_m for ATP is significantly decreased in this mutant when compared to wild-type (~100X) and is identical to its K_m value in the presence of 3-PGA (0.04mM). The kinetic mechanism established ATP as the first substrate to bind, thus the activator is no longer necessary for this first step. The GGR enzyme also retains a high K_m for G-1-P in the absence of

an activator, similar to the wt maize enzyme, but its catalytic rate is approximately double (**Table 10**). The ATP K_m is once again comparable to its value in the activated form, alleviating the need for an activator to bind the first substrate. Thus, 3-PGA acts on both of these mutants by increasing the binding of G-1-P with little effect on the binding of ATP. This is in contrast to the single mutation of T142F (Boehlein *et al.,* 2013), a recently constructed mutation created based on evolutionary constraints, where the ATP K_m remains high, and the G-1-P K_m is equal to its 3-PGA activated value. This variant also conditions both heat stability and high activity in the absence of an activator.

[0084] Next, the kinetics of the GGR mutant were further investigated by following Pi inhibition in the presence and absence of 3-PGA. Experiments were performed by keeping both substrates constant and non-saturating at several fixed concentrations of Pi. Inhibition of the wt maize AGPase by Pi is quite complex. In the absence of an activator, Pi was found to be a partial mixed type inhibitor with respect to both ATP and G-1-P (**Table 11**). Under the assay conditions, the calculated K_m for both substrates are very high, greater than 20mM, and the Ki for Pi is also high, in the 3-5mM range. Interestingly, the inhibition is incomplete, thus even at high concentrations, Pi does not completely inhibit the wild-type enzyme. The partiality of inhibition obtained is calculated from a β factor <1 (**Table 11**). Additionally, Pi interacts with the substrates, in such a way that makes their binding better. This is reflected by an α value <1 when either substrate is varied, (**Table 11**).

Table 11. Pi inhibition in the absence of 3-PGA

Enzyme	Varied Substrate	[b]Pattern	K_m	K_i	α	β	V_{max}
[a] Maize	ATP	P-MT	20.4 +/- 10.1	3.4 +/-1.1	0.063+/- 0.034	0.52+/- 0.17	6.7 +/- 2.2
	G-1-P	P-MT	22.2+/-9.0	5.1 +/- 2.1	0.095 +/- 0.065	0.92 +/- 0.31	4.2 +/- 1.1
GGR	ATP	NC	0.54 +/-0.057	113.0 +/- 27			7.1 +/-0.3
	G-1-P		13.1 +/- 0.96	52.7 +/- 2.9			16.0 +/- 0.55
SH2E	ATP	NC	0.97+/-0.081	20.9+/-1.6			11.0+/-1.4
	G-1-P		0.20+/-0.019	0.055 +/- 0.013	0.85 +/- 0.17	0.44 +/- 0.03	5.6 +/- 0.15

ATP, 0.5mM; G-1-P, 5mM

[a] data from Boehlein *et al.*, 2010

[b] data from Boehlein *et al.*, 2013

[0085] A similar examination of the Pi inhibition was then followed with GGR in the absence of 3-PGA. This investigation showed that inhibition was quite simple, with Pi exhibiting a non-competitive pattern with respect to each substrate, thus the substrate and Pi can be bound to the GGR mutant at any given time, and the ESI complex cannot go on to form product. Furthermore, the inhibition by phosphate is complete, and Pi does not affect the binding of either substrate. Under the conditions of this experiment, the Km for G-1-P is extremely high, similar to the wt AGPase, but the ATP Km is very low, approximately 40-fold lower than wt AGPase. While the inhibition by Pi is complete, the Ki value with respect to both substrates (50-100 mM) is extremely high, approximately 10-20 times higher than the normal concentration of Pi in the endosperm tissue of maize (~6.5 mM).

[0086] Inhibition by Pi was then followed in the presence of the activator 3-PGA. Here again, the wt AGPase displayed a partial mixed type inhibition pattern with K_i values 0.5-1 mM (**Table 12**). The interaction factor, α, of >1 indicates that ATP binding in the presence of 3-PGA becomes more difficult as the concentration of Pi rises. The GGR mutant displayed a

linear mixed type pattern with an interaction factor of 2.5 when varying ATP, thus in the presence of 3-PGA, Pi has a negative effect on ATP binding, and Pi inhibition is complete. A simple non-competitive inhibition pattern was seen when varying G-1-P. The K_i for Pi was still much higher (10-20 mM) than the wt AGPase, thus this variant would be more recalcitrant to Pi inhibition in the endosperm. This is similar to what is seen for the evolutionary mutant SH2-E where simple inhibition patterns were also observed.

Table 12. Pi inhibition in the presence of 3-PGA

Enzyme	Varied Substrate	[a]Pattern	[b] K_m	K_i	α	β	V_{max}
Maize	ATP	P-MT	0.11 +/- 0.0057	0.37 +/-0.053	4.61 +/-0.70	0.83+/-0.037	31.3 +/- 0.4
	G-1-P	P-MT	0.091+/- 0.0042	1.00 +/- 0.18	0.90 +/- 0.012	0.52 +/- 0.023	24.9 +/- 0.3
GGR	ATP	L-MT	0.11 +/- 0.005	10.7 +/- 1.7	2.5 +/- 0.66		21.9 +/- 3.6
	G-1-P	NC	0.085 +/- 0.003	20.7 +/- 0.80			26.1 +/- 0.30
SH2E	ATP	NC	0.21 +/- 0.01	7.8 +/- 0.21			15.1 +/-0.33
	G-1-P	NC	0.063 +/- 0.003	6.62 +/- 0.39			10.0 +/- 1.6

ATP and G-1-P when held constant were 0.4mM and 0.2mM respectively

[a] data from Boehlein *et al.*, 2010

[b] data from Boehlein *et al.*, 2013

Example 7

Transformation of Plant Cells and Field Studies

[0087] pIPK27-MCSBAR expression vectors comprising maize 27 kd zein promoters operably linked to the coding regions of the variant *Sh2-ISM* sequences and NOS terminator sequences were constructed. Expression vectors also included an herbicide resistance gene as a selectable marker. Plasmids were amplified in *E. coli*, and transformed into *Agrobacterium tumefaciens* cells. Transformed *A. tumefaciens* cells were used to transform Hi II maize cells. A total of 12 independent events were generated that gave rise to female fertile plants. These plants were pollinated using B73 maize plants.

[0088] Progeny plants comprising the transgene were selected by painting a small strip of a leaf with herbicide one week following pollination. After 7 days, leaf strips were scored for resistance or sensitivity to the herbicide. Resulting mature ears were scored for total weight, seed number, and total seed weight. Temperatures on each day following pollination were recorded using an on-site weather station. Plants were evaluated for yield, including ear weight.

[0089] Approximately thirty-five plants derived from each line were evaluated for ear weight, with approximately 50% herbicide-sensitive and 50% herbicide-tolerant plants. The ear weight of herbicide-resistant plants was divided by the ear weight of herbicide-sensitive plants in order to determine the effect of transgene expression on yield, as herbicide-resistant plants are expected to contain the transgene. The average temperature over the four days following pollination is shown for each event. The ear weight ratios shown in **Table 13** were determined.

Table 13. Ear Weight Ratios of herbicide-resistant plants divided by herbicide-sensitive plants

Event	Ear Weight Ratio	Temp (°F)
7	1.24	96
10	1.03	96
25	1.01	96
11	1	96
6	0.91	96
16	0.91	96
45	0.66	96

Example 8

Determination of Transgene Expression Levels

[0090] Events 7, 10, 25, and 11 were self-pollinated. The resulting seeds were planted in soil and grown in a greenhouse setting. The plants that grew from these seeds were cultivated and self-pollinated. Developing seeds were harvested 12 days after pollination and RNA was extracted from the developing embryos using Qiazol reagent and the 96 well RNA extraction kit from Qiagen. RNA was converted to cDNA using M-MuLV Reverse Transcriptase. cDNA was amplified using primers specific to the ism-2 transcript (SEQ ID NOs:71 and 72). The products of the RT-PCR reactions were visualized on an agarose gel stained with GelRed (VWR). The resulting gel image showed a clear signal at the correct size that was amplified from the RNA of events 7, 10, and 11, but not from event 25.

[0091] Following these qualitative RT-PCR experiments, quantitative RT-PCR (qRT-PCR) experiments were performed with the cDNA derived from the developing embryos of the plants described above. The primers used for amplification were SEQ ID NOs: 71 and 72. These qRT-PCR experiments produced the data shown in **Table 14**. As shown in this table, events 10-9, 10-5, and 11 accumulate transcript from the transgene at approximately 5.4, 7.7, and 20.8-times higher levels than event 7.

Table 14. qRT-PCR quantification of transgene expression levels

Event	Fold Over Control
E7	0.13 ± 0.02
E10-9	0.7 ± 0.2
E10-5	1 ± 0.2
E11	2.7 ± 0.2
E25	Run Error
WT Control	0.0004 ± 0.0001

Example 9

Determination of Yield Components in Greenhouse-Grown Maize

[0092] The plants that were used for the RT-PCR and qRT-PCR experiments described above were allowed to mature, and the resulting seeds were harvested in order to determine the yield components associated with these plants. Temperatures were controlled for the growth of these plants such that the average high temperature in the four days following pollination was 84°F and the average low temperature in this same period was 74°F. Table 16 shows the yield data from these greenhouse-grown plants. This table shows that event 25, which did not show detectable transgene expression in RT-PCR and qRT-PCR experiments, had higher yield than event 11, with the highest expression levels, but lower yield than events 7 or 10, which both showed lower expression levels than event 11. **Table 15** shows that the higher yields observed with events 7 and 10 resulted primarily from an increase in the number of seeds produced. Events 7 and 10 produced an average of 263 and 353 kernels per ear, respectively, while event 25 produced an average of 173 kernels per ear and event 11 produced an average of 92 kernels per ear. Ear weight, ear length, kernel count, and total seed weight differed significantly among the four events tested. Ear circumference and 100 kernel weight did not differ significantly among these events.

Table 15. Yield components of greenhouse-grown maize plants

Event	Ear Weight (g)	Ear Length (cm)	Ear Circumference (cm)	100 Kernel Weight (g)	Kernel Count	Total Seed Weight (g)
25	47.02	8.80	12.60	18.05	186	33.57
	45.12	7.60	13.30	22.06	129	28.46
	62.86	11.10	12.70	22.78	198	45.10
	55.98	7.90	13.60	23.60	178	42.01
AVE	**52.75**	**8.85**	**13.05**	**21.62**	**173**	**37.29**
11	52.68	6.70	14.50	25.84	124	32.04
	41.06	7.40	13.20	23.00	101	23.23
	18.72	4.40	11.00	18.78	50	9.39
AVE	**37.49**	**6.17**	**12.90**	**22.54**	**92**	**21.55**
7	59.04	10.30	12.90	22.30	223	49.73
	77.54	11.50	12.70	19.92	302	60.16
	75.98	9.80	13.30	23.30	263	61.28
AVE	**70.85**	**10.53**	**12.97**	**21.84**	**263**	**57.06**
10	**85.52**	**11.40**	**14.00**	**21.96**	**353**	**77.52**

* * *

[0093] All of the materials and methods disclosed and claimed herein can be made and used without undue experimentation as instructed by the above disclosure. Although the materials and methods of this invention have been described in terms of preferred embodiments and illustrative examples, it will be apparent to those of skill in the art that variations can be applied to the materials and methods described herein without departing from the concept, spirit and scope of this invention. All such similar substitutes and modifications apparent to those skilled in the art are deemed to be within the spirit, scope and concept of this invention as defined by the appended claims.

WHAT IS CLAIMED IS:

Claim 1. A polynucleotide comprising a sequence encoding a variant AGPase subunit polypeptide, said polypeptide comprising at least 70% identity to SEQ ID NO: 70 and comprising a mutation at a position corresponding to amino acid 96, 161, or 443 of SEQ ID NO: 20 relative to a wild type AGPase subunit polypeptide.

Claim 2. The polynucleotide of claim 1, wherein said variant AGPase subunit polypeptide is an AGPase large subunit polypeptide.

Claim 3. The polynucleotide of claim 1, wherein said variant AGPase subunit polypeptide comprises a conserved motif having SEQ ID NO: 21, SEQ ID NO: 22, or SEQ ID NO: 23.

Claim 4. The polynucleotide of claim 1, wherein said polypeptide comprises at least 90% identity to SEQ ID NO:70.

Claim 5. The polynucleotide of claim 1, wherein said polypeptide comprises mutations in at least two positions corresponding to amino acid 96, 161, or 443 of SEQ ID NO: 20 relative to a wild type plant AGPase large subunit polypeptide.

Claim 6. The polynucleotide of claim 1, wherein said polypeptide comprises mutations in positions corresponding to amino acid 96, 161, and 443 of SEQ ID NO: 20 relative to a wild type plant AGPase large subunit polypeptide.

Claim 7. The polynucleotide of claim 1, wherein said polypeptide comprises at least one mutation selected from the group consisting of a glycine or a glutamine at said position 161, an arginine or a glycine at said position 96, and an arginine at said position 443.

Claim 8. The polynucleotide of claim 1, wherein said polynucleotide is operably linked to a promoter functional in plants.

Claim 9. The polynucleotide of claim 1, wherein the polypeptide comprising said mutation relative to a wild type AGPase subunit polypeptide comprises SEQ ID NO: 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, or 70.

Claim 10. A recombinant construct comprising the polynucleotide of claim 1.

Claim 11. A polypeptide encoded by the polynucleotide of claim 1.

Claim 12. A plant, seed, cell, or plant part comprising the polynucleotide of claim 1.

Claim 13. The plant, seed, cell, or plant part of claim 11, defined as a monocotyledonous plant, seed, cell, or plant part.

Claim 14. The plant, seed, cell, or plant part of claim 11, defined as a dicotyledonous plant, seed, cell, or plant part.

Claim 15. The plant, seed, cell, or plant part of claim 11, wherein said plant is selected from the group consisting of rice, wheat, barley, oats, rye, sorghum, maize, sugarcane, pineapple, onion, bananas, coconut, lilies, millet, tomato, potato, sweet potato, pea, strawberry, beet, chickpea, watermelon, muskmelon, cassava, taro, sunflower, flax, and beans.

Claim 16. A method of increasing the resistance of a plant to heat stress conditions or increasing starch biosynthesis of a plant comprising expressing in the plant the polynucleotide of claim 1.

Claim 17. The method of claim 15, wherein the plant is defined as a monocotyledonous plant.

Claim 18. The method of claim 15, wherein the plant is defined as a dicotyledonous plant.

Claim 19. The method of claim 15, wherein the plant is selected from the group consisting of rice, wheat, barley, oats, rye, sorghum, maize, sugarcane, pineapple, onion, bananas, coconut, lilies, millet, tomato, potato, sweet potato, pea, strawberry, beet, chickpea, watermelon, muskmelon, cassava, taro, sunflower, flax, and beans.

Claim 20. The method of claim 15, comprising transforming a plant cell with said polynucleotide and regenerating the plant therefrom.

Claim 21. The method of claim 15, comprising crossing a parent plant comprising said polynucleotide with itself or a second plant to obtain the plant in which resistance of a plant to heat stress conditions or starch biosynthesis is increased.

Common Name	Scientific Name	SEQ ID NO.	GenBank Accession Number
Maize	Zea mays	40	M81603 (Sh2)
		41	Z38111
		42	NP_001106017
Barley	Hordeum vulgare	43	X67151
		44	U66876
Sorghum	Sorghum bicolor	45	T03445
Rice	Oryza sativa	46	T04156
		47	AAK27727
		48	U66041
		49	D50317
Wheat	Triticum aestivum	50	P12299
Tomato	Lycopersicon esculentum	51	T07674
		52	T07682
		53	AAC49942
Potato	Solanum tuberosum	54	X76136
		55	X61187
		56	P55242
Sweet Potato	Ipomoea batatas	57	AFL55396
		58	AFL55397
		59	AFL55398
		60	AFL55399
Sweet orange	Citrus sinensis	61	ACF77017
Pea	Pisum sativum	62	X96766
Strawberry	Fragaria x ananassa	63	AAS00542
Beet	Beta vulgaris	64	P55233
Chickpea	Cicer arietinum	65	AF356003
Watermelon	Citrullus lanatus	66	JE0132
		67	JE0133
Muskmelon	Cucumis melo	68	AF030383
		69	AF030384

FIG. 1

```
Sequence Alignment

Maize_SH2-M81603                  ---MQFA--LALDTNSGPHQIR-SCEGDGIDR-LEKLSIGGRKQEKALRN 43
Sorghum-T03445                    ---MQFS--LASDANSGPHPIRRSCEGGGIDR-LERLSIGGSKQEKALRN 44
Rice-T04156                       ---MQFM--MPLDTNACAQPMRRAGEGAGTERLMERLNIGGMTQEKALRK 45
Rice-AAK27727                     ---MEFM--MPLDTNACAQPMRRAGEGAGTERLMERLNIGGMTQEKALRK 45
Rice-U66041                       ---MQFM--MPLDTNACAQPMRRAGEGAGTERLMERLNIGGMTQEKALRK 45
Maize_Agp1-Z38111                 ---MQFSSVLPLEGKACMSPVRRGSGGYGSER-MRIN-CCSIRRNKALRR 45
Rice-D50317                       ---MQFSSVFPLEGKACVSPIRRGGEGSASDR-LKIGDSSSIKHDRAVRR 46
Barley-X67151                     MSSMQFSSVLPLEGKACVSPVRR--EGSACER-LKIGDSSSIRHERASRR 47
Wheat-P12299                      MSSMQFSSVLPLEGKACISPVRR--EGSASER-LKVGDSSSIRHERASRR 47
Maize_Agpllzm                     ----MGLRVAATAPAPAGVRVLGRGAARVTPRPW------AAVGGRR--- 37
Barley-U66876                     ----MDLRVAAPASVAAAARRGALGCARVRP-----------LQGRRQ-- 33
Sweet_Potato-AFL55399             ----MAVTADGRIALLAARQLREGAAMTVSSCRLSVKFCNGEFMGKKIKL 46
Tomato-T07674                     ----MSVATDVRFALLRN----NPAALTGTNLKI-VKFCNGELMGKKLKY 41
Potato-X76136                     ----------------------------------------MGKKLKY 7
Sweet_Potato-AFL55396             -----MEFCPTLKSSAHLPR---------ETEFFGGRIRGSLNNNVLASK 36
Sweet_Potato-AFL55397             ----MDAYCATLKSTTHLPR---------ESELWG---KRMLKTSVVVNQ 34
Tomato-T07682                     ----MDTCCAAMKSTVHLGRVSTGGFNNGEKEIFGEKIRGSLNNNLRINQ 46
Potato-X61187                     --------------------------------------------------
Watermelon-JE0133                 -MVAMDSCFVSLKSNTHLMKGNWGGLDRCENGFYGEKVRGSFNENAWIKS 49
Muskmelon-AF030383                -MVAMDSCFVSLKSNTQLMKGNWGGLDRCENGFMVEKVRGGFNENVWIKS 49
Citrus_sinensis-ACF77017          ----MDSCCVGLRANTHVVKASKYGSKIGDNALWGERIRGSVSNDGCTKQ 46
Pea-X96766                        ----MASGCVSLKTNTHFPNSKK-------GSFFGERIKGSLKNSSWVTT 39
Fragaria_x_ananassa-AAS00542      ----MDSWCVTLKPNTHLRQPTQAGLCCGANGFLGQRIRESFGNRGWVHG 46
Sweet_Potato-AFL55398             ----MDALCASMRAHPVPVSKGFG---YGDSGLWGEKIRGCSRIKTER-- 41
Tomato-AAC49942                   ----MDALCA---GTAQSVAICN-----QESTFWGQKISGRRLINKGFGV 38
Potato-P55242                     ----MDALCASMKGTAQLVAICN-----QESAFWGEKISGRRLINKGFGV 41
Beet-P55233                       ----MDASAAAINVNAHLTEVGKK---RFLGERISQSLKGKDLR-ALFSR 42
Chickpea-AF356003                 ---------MDLAIGSNYASLRSS---VFLGETLKGNLSTKFLTSPKFSQ 38
Watermelon-JE0132                 --------------------------------------------------
Muskmelon-AF030384                ----------MHKISSQEKNQCFG---FWGDSSLGRNGRWKQIQRNASSR 37
```

FIG. 1 (continued)

```
Maize_SH2-M81603                RCFGG---------RVAATTQCILTSDACP-ETLHSQTQSSRKNYADANR 83
Sorghum-T03445                  RCFGG---------RVAATTQCILTSDACP-ETLHFQTQSSRKSYADANH 84
Rice-T04156                     RCFGD---------GVTGTARCVFTSDADR-DTPHLRTQSSRKNYADASH 85
Rice-AAK27727                   RCFGD---------GVTGTARCVFTSDADR-DTPHLRTQSSRKNYADASH 85
Rice-U66041                     RCFGD---------GVTGTARRVFTSDADR-DTPHLRTQFSRKNYADASH 85
Maize_Agp1-Z38111               MCFSA-------RGAVS-STQCVLTSDAGP-DTLVVRT-SFRRNYADPNE 85
Rice-D50317                     MCLGY-------RGTKN-GAQCVLTSDAGP-DTLHVRT-SFRRNFADPNE 86
Barley-X67151                   MCNGGA------RGPAATGAQCVLTSDASPADTLVLRT-SFRRNYADPNE 90
Wheat-P12299                    MCNGG-------RGPAATGAQCVLTSDASPADTLVLRT-SFRRNYADPNE 89
Maize_Agpllzm                   -RFSVRMS------VATTE-ATTTIAVGASE---DQALEAR-----NSKT 71
Barley-U66876                   CRPSVRVS------VATTESAAAAAAVSASA---DEDAETT-----NPRT 69
Sweet_Potato-AFL55399           RKFQQRNGTKYNVVARPRVSMSLTTDVAGEAKLKDYGMEKT-----DPRT 91
Tomato-T07674                   TKFQLRSN-----VVKPHICMSLTTDIAGEAKLKDLEAKKE-----DART 81
Potato-X76136                   TKFQLRSN-----VVKPNICMSLTTDIAGEAKLKDLERQKKG----DART 48
Sweet_Potato-AFL55396           SRKSLRVDGNKR-KIKPGVAFSVLTRENGTETLTVEAPILER-RRANPKN 84
Sweet_Potato-AFL55397           FGKSLKLERNGR-KIKPGVAFSVLTRETGRETLSVEAPRLER-VRANPKN 82
Tomato-T07682                   LSKSLKLE---K-KIKFGEAYSVITIENDTETVFVDMPRLER-RRANPKD 91
Potato-X61187                   -----------N-KIKPGVAYSVITTENDTQTVFVDMPRLER-RRANPKD 37
Watermelon-JE0133               LKS----EKKAL-KLTPNVAYAVATPNISKQPVSIQVPSIPK-VKANPKN 93
Muskmelon-AF030383              LKY----EKKAL-KLTPNVAYAVT-PNVSKQPMTIQVPTVPK-VKANPKN 92
Citrus_sinensis-ACF77017        LKKSLKAEKRDE-KVKPGVAYAVMTSKHPNEVMTLAPPRLER-RRVDPKN 94
Pea-X96766                      QKK----------IKPASFSAILTSDDPKGSLNLQVPSFLR-LRADPKN 77
Fragaria_x_ananassa-AAS00542    SE----------KTRPGVVSSVV--TTKDFETTLKVPTYHR-PRVDPKN 82
Sweet_Potato-AFL55398           ---HEGMPKKVN-LGVACSILTHDIN---KEHLSFETQHFEEHSQGDPRN 84
Tomato-AAC49942                 RWCKSFTTQQRG-RGVTSAVLTRDIN---KEMLPFENSMFEEQPTADPKA 84
Potato-P55242                   RSCKSFTTQQRG-RNVTPAVLTRDIN---KEMLPFEESMFEEQPTADPKA 87
Beet-P55233                     TESKGRNVNKP---GVAFSVLTSDFNQSVKESLKYEPALFES-PKADPKN 88
Chickpea-AF356003               IHINNLRSFNPR-NGASYSVLTSGINDFEESMTFHEGPYFDT-PKADPKS 86
Watermelon-JE0132               -NSNSSPRSTAR-KLTPGVAYSVLMSEISEVSSTLQAPIFET-PRADPKK 47
Muskmelon-AF030384              NNSDSSSSSRAR-SLHPELLILFSCSEVNEETTTLQAPIFEA-PRADPKK 85
```

FIG. 1 (continued)

```
Maize_SH2-M81603                VSAIILGGGTGSQLFPLTSTRATPAVPVGGCYRLIDIPMSNCFNSGINK- 132
Sorghum-T03445                  VSAIILGGGTGSQLFPLTSTRATPAVPVGGCYRLIDIPMSNCFNSGINK- 133
Rice-T04156                     VSAVILGGGTGVQLFPLTSTRATPAVPVGGCYRLIDIPMSNCFNSGINK- 134
Rice-AAK27727                   VSAVILGGGTGVQLFPLTSTRATPAVPVGGCYRLIDIPMSNCFNSGINK- 134
Rice-U66041                     VSAVILGGGTGVQLFPLTRTRATPAVPVGGCYRLIDIPMSNCFNSGINKN 135
Maize_Agp1-Z38111               VAAVILGGGTGTQLFPLTSTRATPAVPIGGCYRLIDIPMSNCFNSGINK- 134
Rice-D50317                     VAAVILGGGTGTQLFPLTSTRATPAVPIGGCYRLIDIPMSNCFNSGINK- 135
Barley-X67151                   VAAVILGGGTGTQLFPLTSTRATPAVPIGGCYRLIDIPMSNCFNSGINK- 139
Wheat-P12299                    VAAVILGGGTGTQLFPLTSTRATPAVPIGGCYRLIDIPMSNCFNSGINK- 138
Maize_Agpllzm                   VVAVILGGGAGTRLFPLTRRRAKPAVPIGGAYRLIDVPMSNCINSGINK- 120
Barley-U66876                   VVAVILGGGAGTRLFPLTKRRAKPAVPIGGAYRLIDVPMSNCINSGINK- 118
Sweet_Potato-AFL55399           VVAIILGGGAGTRLFPLTKRRAKPAVPIGGAYRLIDVPMSNCINSGINK- 140
Tomato-T07674                   VVAIILGGGGGTRLFPLTKRRAKPAVPIGGAYRLIDVPMSNCINSGINK- 130
Potato-X76136                   VVAIILGGGAGTRLFPLTKRRAKPAVPMGGAYRLIDVPMSNCINSGINK- 97
Sweet_Potato-AFL55396           VAAIILGGGAGTQLFPLTNRAATPAVPLGGCYRLIDIPMSNCINSGVNK- 133
Sweet_Potato-AFL55397           VAAIILGGGAGTQLFPLTNRAATPAVPVGGCYRMIDIPMSNCINSGINK- 131
Tomato-T07682                   VAAVILGGGEGTKLFPLTSRTATPAVPVGGCYRLIDIPMSNCINSAINK- 140
Potato-X61187                   VAAVILGGGEGTKLFPLTSRTATPAVPVGGCYRLIDIPMSNCINSAINK- 86
Watermelon-JE0133               VASIILGGGAGTHLFPLTRRSATPAVPVGGCYRLIDIPMSNCINSGINK- 142
Muskmelon-AF030383              VASIILGGGAGTHLFPLTKRSATPAVPAGGCYRLIDIPMSNCINSGINK- 141
Citrus_sinensis-ACF77017        VAAIILGGGAGTKLFPLTLRAATPAVPVAGCYRLIDIPMSNCINSGINK- 143
Pea-X96766                      VISIVLGGGPGTHLYPLTKRAATPAVPVGGCYRLIDIPMSNCINSGINK- 126
Fragaria_x_ananassa-AAS00542    VASIILGGGAFTQLFPLTRRAATPAVPVGGCYRLIDIPMSNCINSNINK- 131
Sweet_Potato-AFL55398           VASIVLGGGAGTRLFPLTRSRAKPAVPIGGCYRLIDVPMSNCINSGIRK- 133
Tomato-AAC49942                 VASVILGGGVGTRLFPLTSRRAKPAVPIGGCYRLIDVPMSNCINSGIRK- 133
Potato-P55242                   VASVILGGGVGTRLFPLTSRRAKPAVPIGGCYRLIDVPMSNCINSGIRK- 136
Beet-P55233                     VAAIVLGGGAGTRLFPLTSRRAKPAVPIGGCYRLIDVPMSNCINSGIRK- 137
Chickpea-AF356003               VASIILGGGAGTRLFPLTSKRAKPAVPIGGCYRLIDIPMSNCINSGIRK- 135
Watermelon-JE0132               IASIILGGGAGTRLFPLTSQRAKPAVPIGGCYRLIDIPMSNCINSGIEK- 96
Muskmelon-AF030384              VASIILGGGAGTRLFPLTSQRAKPAVPIGGCYRLIDIPMSNCINSGIEK- 134
```

FIG. 1 (continued)

Maize_SH2-M81603	IFVMSQFNSTSLNRHIHRTY-LEGGINFADGSVQVLAATQMPEEPA-GWF	180
Sorghum-T03445	IFVMTQFNSTSLNRHIHRTY-LGGEINFADGSVQVLADTQMPEEPD-GWF	181
Rice-T04156	IFVMTQFNSASLNRHIHHTY-LGGGINFTDGSVQVLAATQMPDEPA-GWF	182
Rice-AAK27727	IFVMTQFNSASLNRHIHHTY-LGGGINFTDGSVQVLAATQMPDEPA-GWF	182
Rice-U66041	IFVMTQFNLTSLNRNIHHTY-LVGGINLTDGSVQVLAATQMPDEPA-GWF	183
Maize_Agp1-Z38111	IFVMTQFNSASLNRHIHRTY-LGGGINFTDGSVEVLAATQMPGEAA-GWF	182
Rice-D50317	IFIMTQFNSASLNRHIHRTY-LGGGINFTDGSVEVLAATQMPGEAA-GWF	183
Barley-X67151	IFVMTQFNSASLNRHIHRTY-LGGGINFTDGSVEVLAATQMPGEAA-GWF	187
Wheat-P12299	IFVMTQFNSASLNRHIHRTY-LGGGINFTDGSVEVLAATQMPGEAA-GWF	186
Maize_Agpllzm	VYILTQFNSQSLNRHLSRAYDFSNGVAIGDGFVEVLAATQRPGTEGKRWF	170
Barley-U66876	VYVLTQFNSASLNRHLFRAYNFSNGVGFGDGFVEVLAATQRPGSEGKRWF	168
Sweet_Potato-AFL55399	VYILTQFNSASLNRHLARAYNFGSGVTFGDGYVEVLAATQTPGEAGKRWF	190
Tomato-T07674	VYILTQFNSASLNRHIARAYNFGNGVTFGDGYVEVLAATQTPGELGKRWF	180
Potato-X76136	VYILTQFNSASLNRHIARAYNFGNGVTFESGYVEVLAATQTPGELGKRWF	147
Sweet_Potato-AFL55396	IFVLTQFNSASLNRHISRTY-FGNGVSFGDGFVEVLAATQTQGETGMKWF	182
Sweet_Potato-AFL55397	IFVLTQFNSASLNRHIARTY-FGNGVSFGDGFVEVLAATQTSGETGMKWF	180
Tomato-T07682	IFVLTQYNSAALNRHIARTY-FGNGVSFGDGFVEVLAATQTPGEAGKKWF	189
Potato-X61187	IFVLTQYNSAPLNRHIARTY-FGNGVSFGDGFVEVLAATQTPGEAGKKWF	135
Watermelon-JE0133	IFVLTQFNSASLNRHISRTY-FGNGVNFGEGFVEVLAATQTSGETGMHWF	191
Muskmelon-AF030383	IFVLTQFNSASLNRHISRTY-FGNGVTFKEGFVEVLAATQTSGESGMYWF	190
Citrus_sinensis-ACF77017	IFVLTQFNSASLNRHIARTY-FGNGTNFGDGFVEVLAATQTPGESGKNWF	192
Pea-X96766	IFVLTQFNSASLNRHIARTY-FGNGVNFGDGFVEVLAATQTPGEAGKKWF	175
Fragaria_x_ananassa-AAS00542	IFVLTQFNSTSLNRHLARTY-FGNGINFGDGFVEVLAATQTSGEAGMDWF	180
Sweet_Potato-AFL55398	IFILTQFNSFSLNRHLARAYGIGNGVNFGDGFVEVLAATQTPGEAGKMWF	183
Tomato-AAC49942	IFILTQFNSFSLNRHLARTYNFGNGVGFGDGFVEVLAATQTPGDAGKMWF	183
Potato-P55242	IFILTQFNSFSLNRHLA-TYNFGNGVGFGDGFVEVLAGTQTPGDGRKMWF	185
Beet-P55233	IFILTQFNSFSLNRHLARTYNFGDGVNFGDGFVEVFAATQTPGESGKKWF	187
Chickpea-AF356003	IFILTQFNSFSLNRHLSRSYNFGNVSTFGEGFVEVLAATQTSGEAGKKWF	185
Watermelon-JE0132	IFVLTQFNSFSLNRHLARIYNFGNGVNFGDGFVEVLAATQTSGETGKKWF	146
Muskmelon-AF030384	ISSN-AVNSFSLNRHLARIYNFGNGVNFGDGFVEVLAATQTSGETGKKWF	183

FIG. 1 (continued)

```
Maize_SH2-M81603                QGTADSIRKFIWVLEDYYSHKSIDNIVILSGDQLYRMNYMELVQKHVEDD 230
Sorghum-T03445                  QGTADSVRKFIWVLEDYYNHKSIEHIVILSGDQLYQMNYMELVQKHVEDN 231
Rice-T04156                     QGTADAIRKFMWILEDHYNQNNIEHVVILCGDQLYRMNYMELVQKHVDDN 232
Rice-AAK27727                   QGTADAIRKFMWILEDHYNQNNIEHVVILCGDQLYRMNYMELVQKHVDDN 232
Rice-U66041                     QGTADAIRKFMWILEDHIHKS-IDNIVILCGDQLYRMNYMELVQKHVDTN 232
Maize_Agp1-Z38111               QGTADAVRKFIWVLEDYYKHKAIEHILILSGDQLYRMDYMELVQKHVDDN 232
Rice-D50317                     QGTADAVRKFIWVLEDYYKHKAIEHILILSGDQLYRMDYMELVQKHVDDN 233
Barley-X67151                   RGTADAVRKFIWVLEDYYKHKSIEHILILSGDQLYRMDYMELVQKHVDDN 237
Wheat-P12299                    RGTADAVRKFIWVLEDYYKNKSIEHILILSGDQLYRMDYMELVQKHVDDN 236
Maize_Agpllzm                   QGTADAVRQFDWLFDD-AKSKDIEDVLILSGDHLYRMDYMDFVQSHRQRG 219
Barley-U66876                   QGTADAVRQFAWLFDD-AKSKDIEDVLILSGDHLYRMDYMDFVQSHRQRD 217
Sweet_Potato-AFL55399           QGTADAVRQFHWLFED-PKSKDIEDVLILSGDHLYRMDYMDFVQSHRQSG 239
Tomato-T07674                   QGTADAVRQFHWLFED-ARSKDIEDVLILSGDHLYRMDYLHFVQSHRQSG 229
Potato-X76136                   QGTAHAVRQFHWLFED-ARSKDIEDVLILSGDHLYRMDYLHFVQSHRQSG 196
Sweet_Potato-AFL55396           QGTADAVRQFTWVFED-AKNKDIDNIVILSGDQLYRMDYMDLVQNHIERN 231
Sweet_Potato-AFL55397           QGPADAVRKFTWVFED-AKNKDIENILILSGDQLYRMDYMDLVQNHLDRN 229
Tomato-T07682                   QGTADAVRKFIWVFED-AKNKNIENILVLSGDHLYRMDYMELVQNHIDRN 238
Potato-X61187                   QGTADAVRKFIWVFED-AKNKNIENIVVLSGDHLYRMDYMELVQNHIDRN 184
Watermelon-JE0133               QGTADAVRQFIWVFED-AKNRNVENILILAGDHMYRMDYMDFVQNHIDRN 240
Muskmelon-AF030383              QGTADAVRQFIWVFED-AKNRNVENILILAGDHMYRMGYMDFVQNHIDRN 239
Citrus_sinensis-ACF77017        QGTADAVRQFTWVFED-AKNRNIENVAILCGDHLYRMDYMDFIQSHVDRD 241
Pea-X96766                      QGTADAVRQFTWIFED-AKNINVENVLILAGDHLYRMDYMDLLQSHVDRN 224
Fragaria_x_ananassa-AAS00542    QGTADAVRQFVWVFED-AKNRNVENILILSGDHLYRMDYMDFVQSHVDSN 229
Sweet_Potato-AFL55398           QGTADAVRQFIWVFED-AKNKNIDNILILSGDHLYRMDYMDFVQRHIDTN 232
Tomato-AAC49942                 QGTADAVRQFIWVFEN-QKNKNVEHIIILSGDHLYRMNYMDFVQKHIDAN 232
Potato-P55242                   Q-AADAVREFIWVFEN-QKNKNVEHIIILSGDHLYRMNYMDFVQKHIDTN 233
Beet-P55233                     QGTADAVRQFFWAFED-SKSKDVEHIVILSGDHLYRMDYMSFWQKHIDTN 236
Chickpea-AF356003               QGTADAVRQFIWVFED-AKTKNVEHILILSGDHLYRMNYMDFVQKHIDTN 234
Watermelon-JE0132               QGTADAVRQFIWLFED-AKTKNVEHTLILSGDHLYRMDYMDFVQRHIDTN 195
Muskmelon-AF030384              QGTADAVRPFIWLFED-AQTKNVEHTLILSGDHLYRMDYMDFVQRHIDTN 232
```

FIG. 1 (continued)

```
Maize_SH2-M81603                ADITISCAPVDESRASKNGLVKIDHTGRVLQFFEKPKGADLNSMRVETNF 280
Sorghum-T03445                  ADITVSCAPVDESRASNNGLVKCDHTGRVLQFFEKPKGADLNSMRVDTNF 281
Rice-T04156                     ADITISCAPIDGSRASDYGLVKFDDSGRVIQFLEKPEGADLESMKVDTSF 282
Rice-AAK27727                   ADITISCAPIDGSRASDYGLVKFDDSGRVIQFLEKPEGADLESMKVDTSF 282
Rice-U66041                     ADITISCAPIDGSRASDYGLVKFDHSGRVIQFLEKPEGADLESM-VDTSF 281
Maize_Agp1-Z38111               ADITLSCAPVGESRASDYGLVKFDSSGRVIQFSEKPKGAALEEMKVDTSF 282
Rice-D50317                     ADITLSCAPVGESRASDYGLVKFDSSGRVIQFSEKPKGTDLEAMKVDTSF 283
Barley-X67151                   ADITLSCAPVGESRASEYGLVKFDSSGRVIQFSEKPKGDDLEAMKVDTSF 287
Wheat-P12299                    ADITLSCAPVGESRASEYGLVKFDSSGRVVQFSEKPKGDDLEAMKVDTSF 286
Maize_Agpllzm                   AGISICCLPIDGSRASDFGLMKIDDTGRVISFSEKPKGDELKAMQVDTTV 269
Barley-U66876                   AGISICCLPIDDSRASDFGLMKIDDTGRVISFSEKPKGADLKAMQVDTTL 267
Sweet_Potato-AFL55399           ADITISSLPIDDRRASDFGLMKIDDKGRVLFFSEKPKGDDLKAMAVDTSV 289
Tomato-T07674                   ADITISSLPIDDSRASDFGLMKIDDTGRVMSFSEKPKGDDLKAMAVDTTV 279
Potato-X76136                   ADITISSLPIDDSRASDFGLMKIDDTGRVMSFSEKPKGDDLKAMAVDTTV 246
Sweet_Potato-AFL55396           SDITLSCATVGDSRASDFGLVKIDRRGRVVQFCEKPKGTDLKAMQVDTTL 281
Sweet_Potato-AFL55397           SDITLSCAPVGDSRAVDFGLVKIDRRGKVVQFQEKPKGADLEAMQVDTTR 279
Tomato-T07682                   ADITLSCAPAEDSRASDFGLVKIDSRGRVVQFAEKPKGFDLKAMQVDTTL 288
Potato-X61187                   ADITLSCAPAEDSRASDFGLVKIDSRGRVVQFAEKPKGFDLKAMQVDTTL 234
Watermelon-JE0133               ADISISCAAVGDSRASDYGLVKIDSRGRIIQFSEKPMGANLSAMRVDTTS 290
Muskmelon-AF030383              ADISISCAAVDDSRASDYGLVKLDSRGRIIQFSEKPKGANLNRMRVDTTS 289
Citrus_sinensis-ACF77017        ADITISCAAVGESRASDYGLVKIDNMGRIAQFAEKPSGANLKAMQVDTSL 291
Pea-X96766                      ADITVSCAAVGDNRASDYGLVKVDDRGNIIQFSEKPKGADLKAMQVDTSR 274
Fragaria_x_ananassa-AAS00542    ADITLSCAVVGDSRASDYGLVKIDSRGKIIQFAEKPRGAGLKAMQSDTTL 279
Sweet_Potato-AFL55398           ADITVSCVPMDDSRASDYGLMKIDGSGRIVHFAEKPKGPALKTMQVDTSL 282
Tomato-AAC49942                 ADITVSCVPMDDGRASDFGLMKIDETGRIIQFAEKPKGPALKVMQVDTSI 282
Potato-P55242                   ADITVSCVPMDDGRASDFGLMKIDETGAIIQFAEKPKGPALKAMQVDTSI 283
Beet-P55233                     ADITVSCIPMDDSRASDYGLMKIDHTGRIVHFAEKPKGSDLTAMQVDTTV 286
Chickpea-AF356003               ADITVSCIPMDDSRASDYGLLKIDGKGRIIQFAEKPKGSELKAMRVDTTL 284
Watermelon-JE0132               ADITVSCIPMDDSRASDYGLMKIDDTGRILHFAEKPKGSDLEAMKVDTTV 245
Muskmelon-AF030384              ADITVSCIPMDDSRASDYGLMKIDDTGRIIHFSEKPKGSDLEEMQVDTAV 282
```

FIG. 1 (continued)

```
Maize_SH2-M81603                LSYAIDDAQKYPYLASMG-IYVFKKDALLDLLKSKYTQLHDFGSEILPRA 329
Sorghum-T03445                  LSYAIGDAQKYQYIASMG-IYVFKKDALLDLLKSKYTQLHDFGSEILPRA 330
Rice-T04156                     LSYAIDDKQKYPYIASMG-IYVLKKDVLLDILKSKYAHLQDFGSEILPRA 331
Rice-AAK27727                   LSYAIDDKQKYPYIASMG-IYVLKKDVLLDILKSKYAHLQDFGSEILPRA 331
Rice-U66041                     LSYAIDDKQKYPYIASMG-IYVLKKDVLLDILKSKYAHLQDFGSEILPRA 330
Maize_Agp1-Z38111               LNFAIDSPAEYPYIASMG-VYVFKRDVLLDLLKSRYAELHDFGSEILPKA 331
Rice-D50317                     LNFAIDDPTKFPYIASMG-VYVFKRDVLLNLLKSRYAELHDFGSEILPRA 332
Barley-X67151                   LNFAIDDPAKYPYIASMG-VYVFKRDVLLNLLKSRYAELHDFGSEILPRA 336
Wheat-P12299                    LNFAIDDPAKYPYIASMG-VYVFKRDVLLNLLKSRYAELHDFGSEILPRA 335
Maize_Agpllzm                   LGLSKEEAENKPYIASMG-IYIFKKDILLNLLRWRFPTANDFGSEIIPAS 318
Barley-U66876                   LGLPKEEAEKKPYIASMG-VYIFKKEILLNLLRWRFPTANDFGSEIIPAA 316
Sweet_Potato-AFL55399           LGLSPEEAKQKPYIASMG-VYVFKKEILLNLLRWRFPTANDFGSEIIPAS 338
Tomato-T07674                   LGLSPEEAKEKPYIASMG-VYVFKKDILLNLLRWRFPTVNDFGSEIIPAS 328
Potato-X76136                   LGLSPEEAKEKPYIASIGKVYVFKKDILLNLLRWRFPTANDFGSEIIPAS 296
Sweet_Potato-AFL55396           LGLPPQDARLNPYIASMG-VYVFKTDVLLRLLRWRYPTSNDFGSEILPAA 330
Sweet_Potato-AFL55397           LGLSPEDAKRNPYIASMG-LYVFRRDLLLNLLRWIYPTANDFGSEIIPAV 328
Tomato-T07682                   VGLSPQDAKKSPYIASMG-VYVFKTDVLLKLLKWSYPTSNDFGSEIIPAA 337
Potato-X61187                   VGLSPQDAKKSPYIASMG-VYVFKTDVLLKLLKWSYPTSNDFGSEIIPAA 283
Watermelon-JE0133               FGLSREESLKSPYIASMG-VYVFKTDILLNLLKWRYPTSNDFGSEIIPAA 339
Muskmelon-AF030383              FGLSREESLKSPYIGSMG-VYVFKTDVLLNLLKWRYPSSNDFGSEIIPAA 338
Citrus_sinensis-ACF77017        LGFSPQEARKCPYVASMG-VYVFKKDVLLKLLRWRYPTSNDFGSEIIPAA 340
Pea-X96766                      LGLSPQDALKSPYIASMG-VYVFKKDVLLKLLKWRYPTSNDFGSEIIPSA 323
Fragaria_x_ananassa-AAS00542    LGFSPQDALKSPYVASMG-VYVFKTDILLELLKKSYPNSNDFGSEIIPAA 328
Sweet_Potato-AFL55398           LGLSENEAKKYPYIASMG-VYVFRTEVLLNLLRSQYPSCNDFGSEIIPAA 331
Tomato-AAC49942                 LGLSEQEASNFPYIASMG-VYVFKTDVLLKLLKSAYPSCNDFGSEIIPSA 331
Potato-P55242                   LGLSEQEASNFPYIASMG-VYVFKTDVLLNLLKSAYPSCNDFGSEIIPSA 332
Beet-P55233                     LGLSDLEAMSNPYIASMG-VYVFRTDVLMELLNRKYPSSNDFGSEIIPSA 335
Chickpea-AF356003               LGLSPEEAKKQPYIASMG-VYVFRTETLLKLLRSNCSTCNDFGSEIIPSA 333
Watermelon-JE0132               LGLSNQDARKNPYIASMG-VYIFRTDLLLKLLTWSYPSCNDFGSEIIPSA 294
Muskmelon-AF030384              LGLSDEDARKNPYIASMG-VYIFRTDLLLKLLTWSYPACNDFGSEIIPAA 331
```

FIG. 1 (continued)

```
Maize_SH2-M81603                VLD-HSVQ-ACIFTGYWEDVGTIKSFFDANLALTEQPSKFDFYDPKTPFF 377
Sorghum-T03445                  VLE-HNVQ-TCIFMGYWEDVGTIKSFFDANLALTEQPSKFDFYDPKTPFF 378
Rice-T04156                     VLE-HNVK-ACVFTEYWEDIGTIKSFFDANLALTEQPPKFEFYDPKTPFF 379
Rice-AAK27727                   VLE-HNVK-ACVFTEYWEDIGTIKSFFDANLALTEQPPKFEFYDPKTPFF 379
Rice-U66041                     LLE-HNVKVACVFTEYWEDIGTIKSFFDANLALTEQPPKFEFYDPKTPFF 379
Maize_Agp1-Z38111               LHE-HNVQ-AYVFTDYWEDIGTIRSFFDANMALCEQPPKFEFYDPKTPFF 379
Rice-D50317                     LHE-HNVQ-AYVFADYWEDIGTIRSFFDANMALCEQPPKFEFYDPKTPFF 380
Barley-X67151                   LHD-HNVQ-AYVFTDYWEDIGTIRSFFDANMALCEQPPKFEFYDPKTPFF 384
Wheat-P12299                    LHD-HNVQ-AYVFTDYWEDIGTIRSFFDANMALCEQPPKFEFYDPKTPFF 383
Maize_Agpllzm                   AKE-IDVKAY-LFNDYWEDIGTIKSFFEANLALAEQPPRFSFYDADKPMY 366
Barley-U66876                   ARE-INVKAY-LFNDYWEDIGTIKSFFEANLALAEQPSKFSFYDASKPMY 364
Sweet_Potato-AFL55399           ARE-FYIQAY-LFNDYWEDIGTIRSFFEANLALTEHPPRFSFYDATKPIY 386
Tomato-T07674                   TKE-FCVKAYYLFNDYWEDIGTIRSFFEANLALTEHPPRFSFYDATKPIY 377
Potato-X76136                   TKE-FCVKAY-LFNDYWEDIGTIRSFFRANLALTEHPPRFSFYDATKPIY 344
Sweet_Potato-AFL55396           VME-HNVQAY-IFRDYWEDIGTIKSFYDANLALTEEFPKFEFYDPKTPFY 378
Sweet_Potato-AFL55397           ITE-HNVQAY-FFKDYWEDIGTIKTFYDANLALAEEFPKFEFYDPKTPFY 376
Tomato-T07682                   IDD-YNVQAY-IFKDYWEDIGTIKSFYNASLALTQEFPEFQFYDPKTPFY 385
Potato-X61187                   IDD-YNVQAY-IFKDYWEDIGTIKSFYNASLALTQEFPEFQFYDPKTPFY 331
Watermelon-JE0133               VKE-HNVQAY-IFRDYWEDIGSIKTFYDANLALTEEFPKFEFYDPKTPIY 387
Muskmelon-AF030383              IKD-HNVQAF-MFRDYWEDIGTIKTFYDANLALHGNVSKFEFYDPKTPFY 386
Citrus_sinensis-ACF77017        IME-HDVQAY-IFRDYWEDIGTIKSFYEANMALTKESPAFHFYDPKTPFY 388
Pea-X96766                      IRE-HNVQAY-FFGDYWEDIGTIKSFYDANLALTEESPKFEFYDPKTPIF 371
Fragaria_x_ananassa-AAS00542    VEE-RNVQAY-IFIDYWEDIGTIQSFYDANLALTEEFPKFQFYDPKTPFF 376
Sweet_Potato-AFL55398           VKD-HNVQAY-LFSDYWEDIGTVKSFFDANLALTEQPPMFDFNDPKTPFY 379
Tomato-AAC49942                 VKD-HNVQAY-LFNDYWEDIGTVKSFFDANLALTKQPPKFDFNDPKTPFY 379
Potato-P55242                   VKD-HNVQAY-LFNDYWEDIGTVKSFFDANLALTKQPPKFDFNDPKTPFY 380
Beet-P55233                     VGE-SNVQAY-LFNDYWEDIGTIKSFFDSNLALTQQPPKFEFYDPKTPFY 383
Chickpea-AF356003               VNDDHNVQAY-LFNDYWEDIGTIKSFFDANLALTDQPPKFQFYDPNTPFY 382
Watermelon-JE0132               VKD-YKVQAY-LFNDYWEDIGTVKSFFDANLALTEQPPKFEFYDPKTPFY 342
Muskmelon-AF030384              VKD-YKVQAY-LFNDYWEDIGTVKSFFDANLALTEQPPKFEFYDPKTPFY 379
```

FIG. 1 (continued)

```
Maize_SH2-M81603                TAPRCLPPTQLDKCK--MKYAFISDGCLLRECNIEHSVIGVCSRVSSGCE 425
Sorghum-T03445                  TAPRYLPPTQLDKCK--IKDASISDGCLLRECSIEHSVIGVCSRVSYGCE 426
Rice-T04156                     TSPRYLPPARLEKCK--IKDAIISDGCSFSECTIEHSVIGISSRVSIGCE 427
Rice-AAK27727                   TSPRYLPPARLEKCK--IKDAIISDGCSFSECTIEHSVIGISSRVSIGCE 427
Rice-U66041                     TSPRYLPPARLDKCKCKIKDAIISDGCSFSECTIEHSVIGISSRVSSGCE 429
Maize_Agp1-Z38111               TSPRYLPPTKSDKCR--IKDAIISHGCFLRECAIEHSIVGVRSRLNSGCE 427
Rice-D50317                     TSPRYLPPTKSDKCR--IKDAIISHGCLLRECTIGHSIVGVRSRLNSACE 428
Barley-X67151                   TSPRYLPPTKSDKCR--IKEAIISHGCFLRECKIEHSIIGVRSRLNSGSE 432
Wheat-P12299                    TSPRYLPPTKSDKCR--IKEAIISHGCFLRECKIEHSIIGVRSRLNSGSE 431
Maize_Agpllzm                   TSRRNLPPSMVNNSK--ITDSIISHGCFLDNCRIEHSVVGVRSRIGSNVH 414
Barley-U66876                   TSRRNLPPSMISGSK--ITDSIISHGCFLDKCRVEHSVVGIRSRIGSNVH 412
Sweet_Potato-AFL55399           TSRRNLPPSAITNSK--IVDSIISHGSFLSDCFVEHSVVGIRSRINSNVH 434
Tomato-T07674                   TSRRNLPPSAIDNSK--IVDSIVSHGSFLTNCFVEHSVVGIRSRIGTNVH 425
Potato-X76136                   TSRRNLPPSAIDNSK--IVDSIVSHGIFLTNCFVEHSVVGIRSRIGTNVH 392
Sweet_Potato-AFL55396           TSPRFLPPTKIDNCK--IKDAIISHGCFLRECTVEHSIIGERSRLDCGVE 426
Sweet_Potato-AFL55397           TSPRFLPPTKIDNCK--IKDAIISHGCFLRECIVEHSIVGERSRLDFGVE 424
Tomato-T07682                   TSPRFLPPTKIDNCK--IKDAIISHGCFLRDCTVEHSIVGERSRLDCGVE 433
Potato-X61187                   TSPRFLPPTKIDNCK--IKDAIISHGCFLRDCSVEHSIVGERSRLDCGVE 379
Watermelon-JE0133               TSPRFLPPTKIDKCQ--IVDAIISHGCFLRECSVQHSIVGERSRLDYGVE 435
Muskmelon-AF030383              TSPRFLPPTKIDRCQ--IVDAIISHGCFLRECSIQHSIVGERSRLDYGVE 434
Citrus_sinensis-ACF77017        TSPRFLPPTKIDNCR--MKDAIISHGCFLRECTVEHSIVGERSRIDYGVE 436
Pea-X96766                      TSPGFLPPTKIDNSR--VVDAIISHGCFLRDCTIQHSIVGERSRLDYGVE 419
Fragaria_x_ananassa-AAS00542    TSPRFLPPTKIDNSR--VVDAIISHGCFLQECFVQSSIVGERSRLDYGVE 424
Sweet_Potato-AFL55398           TSPRFLPPTKVDKCK--IVDAIISHGCFLRECSVKHSIVGIRSRLDYGVE 427
Tomato-AAC49942                 TSARFLPPTKVDKSR--IVDAIISHGGFLRECNIQHSIVGVRSRLDYGVE 427
Potato-P55242                   TSARFLPPTKVDKSR--IVDAIISHGCFLRECNIQHSIVGVRSRLDYGVE 428
Beet-P55233                     TSARFLPPTKVDRCK--IVDSIVSHGCFLQESSIQHSIVGVRSRLESGVE 431
Chickpea-AF356003               TFPRFLPPTKVEKCK--IVDAIISHGCFLRECSVQHSIVGIRSRLESGVE 430
Watermelon-JE0132               TSPRFLPPSKVEKCR--IVDAIISHGCFLRECSVEHSIVGVRSRLEYGVE 390
Muskmelon-AF030384              TSPRSCPPSKVEKCR--IVDAIISHGCFLRECTVEPLIVGVRSRLEYGVE 427
```

FIG. 1 (continued)

Maize_SH2-M81603	LKDSVMMGADIYETEEEASKLLLAGKVPIGIGRNTKIRNCIIDMNARIGK	475
Sorghum-T03445	LKDCVMMGADIYETEEEASKLLLAGEVPVGIGGNTKIRNCIIDINARIGK	476
Rice-T04156	LKDTMMMGADQYETEEETSKLLFEGKVPIGIGENTKIRNCIIDMNARIGR	477
Rice-AAK27727	LKDTMMMGADQYETEEETSKLLFEGKVPIGIGENTKIRNCIIDMNARIGR	477
Rice-U66041	LK--------IYETEEETSKLLFEGKVPIGIGQNTKIRNCIIDMNARIGR	471
Maize_Agp1-Z38111	LKNTMMMGADLYETEDEISRLLAEGKVPIGVGENTKISNCIIDMNARVGR	477
Rice-D50317	LKNTMMMGADLYETEDEISRLLSEGKVPIGVGENTKINNCIIDMNARVGR	478
Barley-X67151	LKNAMMMGADSYETEDEISRLMSEGKVPIGVGENTKISNCIIDMNARIGR	482
Wheat-P12299	LKNAMMMGADSYETEDEISRLMSEGKVPIGVGENTKISNCIIDMNARIGR	481
Maize_Agpllzm	LKDTVMLGADYYETAVERGELLAEGKVPIGIGENTTIQKCIIDKNARIGK	464
Barley-U66876	LKDTVMLGADFYETDAERGDQLAEGKVPIGIGENTSIQNCIIDKNARIGK	462
Sweet_Potato-AFL55399	LKDTVMLGADYYETGAEIASLLTEGGVPIGIGENSRIKECIIDKNARIGK	484
Tomato-T07674	LKDTVMLGADYYETDAEIASQLAEGKVPLGIGENTRIKECIIDKNARIGK	475
Potato-X76136	LKDTVMLGADYYETDAEIRSQLAEGKVPLGIGENTRIKDCIIDKNARIGK	442
Sweet_Potato-AFL55396	LKDTLMMGADNYETESEIASLLADGKVPIGVGENTKIRNAIIDKNVRIGK	476
Sweet_Potato-AFL55397	LKDTLMMGADYYETESEIASLLADGKVPIGIGHNTKISNCIIDKNVRIGK	474
Tomato-T07682	LKDTFMMGADYYQTESEIASLLAEGKVPIGIGENTKIRKCIIDKNAKIGK	483
Potato-X61187	LKDTFMMGADYYQTESEIASLLAEGKVPIGIGENTKIRKCIIDKNAKIGK	429
Watermelon-JE0133	LKDTIMMGADTYQTEPEIAGLLAEGKVPIGIGRNTKIRNCIIDKNAKIGK	485
Muskmelon-AF030383	LKDTIMMGADNYQTESEITGLLAEGKVPVGIGPNTKIRKCIIDKNAKIGK	484
Citrus_sinensis-ACF77017	LKDTVMLGADYYQTESEIASLLAEGKVPIGVGRNTKIRNCIIDKNVKIGK	486
Pea-X96766	LQDTVMMGADYYQTESEIASLLAEGKVPIGIGRNTKIKNCIIDKNAKIGK	469
Fragaria_x_ananassa-AAS00542	LKDSIMMGADSYQTESEIAALLARGKVPIGIGRNTKIRLCIVDLNAKIGK	474
Sweet_Potato-AFL55398	LEDTMVMGADYYQTESEIASLLATGKVPIGIGTNTKIRNCIIDKNARIGK	477
Tomato-AAC49942	FKDTMMMGADYYQTESEIASLLAEGKVPIGVGPNTKIQKCIIDKNAKIGK	477
Potato-P55242	FKDTMMMGADYYQTECEIASLLAEGKVPIGVGPNTKIQNCIIDKNAKIGK	478
Beet-P55233	FQDTMMMGADYYQTESEIASLLAEGKVPVGVGQNTKIKNCIIDKNAKIGK	481
Chickpea-AF356003	LQDTMMMGADYYQTESEIASLLAEGKVPVGVGENTKIRNCIIDKNARIGR	480
Watermelon-JE0132	LKDTMMMGADYYQTESEIASLLAEGKIPIGIGENTKIRNCIIDKNARIGR	440
Muskmelon-AF030384	LKDTMMMGAYYYQTESEIASLLAEGKIPIGIGENTKIRNCIIDKNAKIGR	477

FIG. 1 (continued)

```
Maize_SH2-M81603                NVVITNSKGIQEADHPEEGYYIRSGIVVILKNAT---INDGSVI 516
Sorghum-T03445                  NVVITNSKGIQEADHPEEGYYIKSGIVVILKNAT---IKDGSVI 517
Rice-T04156                     NVIIANTQGVQESDHPEEGYYIRSGIVVILKNAT---IKDGTVI 518
Rice-AAK27727                   NVIIANTQGVQESDHPEEGYYIRSGIVVILKNAT---IKHGPII 518
Rice-U66041                     NAIIANTQGVQESDHPEEG-YIRSGIVVILKNATNATIKHGTVI 514
Maize_Agp1-Z38111               NVSITNKEGVQEADRPDEGYYIRSGIVVVLKNAT---IKDGTVI 518
Rice-D50317                     NVVITNSEGVQESDRPEEGYYIRSGIVVILKNAT---IKDGKVI 519
Barley-X67151                   DVVISNKEGVQEADRPEEGYYIRSGIVVIQKNAT---IKDGTVV 523
Wheat-P12299                    DVVISNKEGVQEADRPEEGYYIRSGIVVIQKNAT---IKDGTVV 522
Maize_Agpllzm                   KVVISNSEGVDEADRTSEGFYIRSGITVVLKNAI---IADGLVI 505
Barley-U66876                   NVTIANTEGVQESDRTSEGFHIRSGITVVLKNSV---IADGLVI 503
Sweet_Potato-AFL55399           NVVIANSEGIQEADRTSEGFYIRSGVTVILKNST---IPDGLVI 525
Tomato-T07674                   NVVIANSEGVQEADRSSEGFYIRSGITVILKNST---IPDGTVI 516
Potato-X76136                   NVVIANSEGVQEADRSSEGFYMASGITVISKNST---IPDGTVI 483
Sweet_Potato-AFL55396           DVVITNKDGVQESDRPDEGFYIRSGITIIMEKAT---IRDGTVI 517
Sweet_Potato-AFL55397           DVIIANKDGVEEADRPEEGFYIRSGIPVIMEKAV---IKDGTVI 515
Tomato-T07682                   NVSIINKDGVQEADRPEEGFYIRSGIIIIAEKAT---IRDGTVI 524
Potato-X61187                   NVSIINKDGVQEADRPEEGFYIRSGIIIILEKAT---IRDGTVI 470
Watermelon-JE0133               DVVIMNKEGVQEADRPEQGFYIRSGITIILEKAT---IEDGTVI 526
Muskmelon-AF030383              DVIIMNKDGVQEADRPEQGFYIRSGITIVMEKAT---IEDGTVI 525
Citrus_sinensis-ACF77017        DVVIVNKDGVQEADRPELGFYIRSGITIIMEKAT---IEDGMVI 527
Pea-X96766                      EVVIANKEGVQEADRSEDGFYIRSGITIIMEKAT---IEDGTVI 510
Fragaria_x_ananassa-AAS00542    DVIIMNKDGIQEADRPEEGFYIREESLSLWRRE---------- 507
Sweet_Potato-AFL55398           DVVIANKDGVDEADRADEGFYIRSGITIVLKNAT---IRDGTVI 518
Tomato-AAC49942                 DVVILNKQGVEEADRSAEGFYIRSGITVIMKNAT---IKDGTVI 518
Potato-P55242                   DVVILNKEGVEEADRSAEGFYIRSGITVIMKNAT---IKDGTVI 519
Beet-P55233                     DVVIANTDGVEEADRPNEGFYIRSGITIILKNAT---IQDGLVI 522
Chickpea-AF356003               NVIITNADGVEEADRTKEGFYIRSGITAILKNAT---IKDGTVI 521
Watermelon-JE0132               NVVIANSDDVQEADRPEDGFYIRSGITVTLKNAT---IKDGTII 481
Muskmelon-AF030384              NVVIANTDDVQEADRPEEGFYIRSGITVTLKNAT---IKDGTII 518
```

FIG. 1 (continued)