

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
14 December 2006 (14.12.2006)

PCT

(10) International Publication Number
WO 2006/133166 A2

(51) International Patent Classification:

C12N 15/82 (2006.01) A01H 5/00 (2006.01)
C12N 9/10 (2006.01) A01H 5/10 (2006.01)
C12N 15/29 (2006.01)

(21) International Application Number:

PCT/US2006/021881

(22) International Filing Date: 6 June 2006 (06.06.2006)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/595,112 7 June 2005 (07.06.2005) US

(71) Applicant (for all designated States except US): **BASF PLANT SCIENCE GMBH** [—/DE]; Carl-bosch-str. 38, 67056 Ludwigshafen (DE).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **HAERTEL, Heiko** [DE/US]; 24 Chestnut Bluffs Lane, Durham, NC 27713 (US). **OSWALD, Oliver** [DE/DE]; Bexbacher Str. 5, 67063 Ludwigshafen (DE).

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

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: NUCLEIC ACID MOLECULES ENCODING SUCROSE SYNTHASE-LIKE POLYPEPTIDES AND METHODS OF USE

(57) Abstract: Described herein are inventions in the field of genetic engineering of plants, including isolated nucleic acid molecules encoding *Sucrose Synthase*-like (*S US*-like) polypeptides to improve agronomic, horticultural, and quality traits. This invention relates generally to nucleic acid sequences encoding proteins that are related to the presence of seed storage compounds in plants. More specifically, the present invention relates to *SUS*-like nucleic acid sequences encoding sugar and lipid metabolism regulator proteins and the use of these sequences in transgenic plants. In particular, the invention is directed to methods for manipulating sugar-related compounds and for increasing oil level and altering the fatty acid composition in plants and seeds. The invention further relates to methods of using these novel plant polypeptides to stimulate plant growth and/or to increase yield and/or composition of seed storage compounds.



WO 2006/133166 A2

NUCLEIC ACID MOLECULES ENCODING SUCROSE SYNTHASE-LIKE POLYPEPTIDES AND METHODS OF USE

This application claims priority to U.S. provisional application US 60/595112 filed on June 7, 2005, herein incorporated by reference in its entirety.

FIELD OF THE INVENTION

[Para 1] Described herein are inventions in the field of genetic engineering of plants, including isolated nucleic acid molecules encoding Sucrose Synthase-like (SUS-like) polypeptides to improve agronomic, horticultural, and quality traits. This invention relates generally to nucleic acid sequences encoding proteins that are related to the presence of seed storage compounds in plants. More specifically, the present invention relates to SUS-like nucleic acid sequences encoding sugar and lipid metabolism regulator proteins and the use of these sequences in transgenic plants. In particular, the invention is directed to methods for manipulating sugar-related compounds and for increasing oil level and altering the fatty acid composition in plants and seeds. The invention further relates to methods of using these novel plant polypeptides to stimulate plant growth and/or to increase yield and/or composition of seed storage compounds.

BACKGROUND OF THE INVENTION

[Para 2] The study and genetic manipulation of plants has a long history that began even before the famed studies of Gregor Mendel. In perfecting this science, scientists have accomplished modification of particular traits in plants ranging from potato tubers having increased starch content to oil-seed plants such as canola and sunflower having increased or altered fatty acid content. With the increased consumption and use of plant oils, the modification of seed oil content and seed oil levels has become increasingly widespread (e.g. Töpfer et al. 1995, Science 268:681-686). Manipulation of biosynthetic pathways in transgenic plants provides a number of opportunities for molecular biologists and plant biochemists to affect plant metabolism giving rise to the production of specific higher-value products. The seed oil production or composition has been altered in numerous traditional oil-seed plants such as soybean (U.S. Patent No. 5,955,650), canola (U.S. Patent No. 5,955,650), sunflower (U.S. Patent No. 6,084,164), and rapeseed (Töpfer et al. 1995, Science 268:681-686), and non-traditional oil seed plants such as tobacco (Cahoon et al. 1992, Proc. Natl. Acad. Sci. USA 89:11184-11188).

[Para 3] Plant seed oils comprise both neutral and polar lipids (see Table 5). The neutral lipids contain primarily triacylglycerol, which is the main storage lipid that accumulates in oil bodies in seeds. The polar lipids are mainly found in the various membranes of the seed cells, e.g. the endoplasmic reticulum, microsomal membranes and the cell membrane. The neutral and polar lipids contain several common fatty acids (see Table 6) and a range of less common fatty acids. The fatty acid composition of membrane lipids is highly regulated and only a select number of fatty acids are found in membrane lipids. On the other hand, a large number of unusual fatty acids can be incorporated into

the neutral storage lipids in seeds of many plant species (Van de Loo F.J. et al. 1993, *Unusual Fatty Acids in Lipid Metabolism in Plants* pp. 91-126, editor TS Moore Jr. CRC Press; Millar et al. 2000, *Trends Plant Sci.* 5:95-101).

[Para 4] Lipids are synthesized from fatty acids and their synthesis may be divided into two parts: the prokaryotic pathway and the eukaryotic pathway (Browse et al. 1986, *Biochemical J.* 235:25-31; Ohlrogge & Browse 1995, *Plant Cell* 7:957-970). The prokaryotic pathway is located in plastids that are the primary site of fatty acid biosynthesis. Fatty acid synthesis begins with the conversion of acetyl-CoA to malonyl-CoA by acetyl-CoA carboxylase (ACCase). Malonyl-CoA is converted to malonyl-acyl carrier protein (ACP) by the malonyl-CoA:ACP transacylase. The enzyme beta-keto-acyl-ACP-synthase III (KAS III) catalyzes a condensation reaction in which the acyl group from acetyl-CoA is transferred to malonyl-ACP to form 3-ketobutyryl-ACP. In a subsequent series of condensation, reduction and dehydration reactions the nascent fatty acid chain on the ACP cofactor is elongated by the step-by-step addition (condensation) of two carbon atoms donated by malonyl-ACP until a 16- or 18-carbon saturated fatty acid chain is formed. The plastidial delta-9 acyl-ACP desaturase introduces the first unsaturated double bond into the fatty acid. Thioesterases cleave the fatty acids from the ACP cofactor and free fatty acids are exported to the cytoplasm where they participate as fatty acyl-CoA esters in the eukaryotic pathway. In this pathway the fatty acids are esterified by glycerol-3-phosphate acyltransferase and lysophosphatidic acid acyl-transferase to the sn-1 and sn-2 positions of glycerol-3-phosphate, respectively, to yield phosphatidic acid (PA). The PA is the precursor for other polar and neutral lipids, the latter being formed in the Kennedy pathway (Voelker 1996, *Genetic Engineering ed.:Setlow* 18:111-113; Shanklin & Cahoon 1998, *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 49:611-641; Frentzen 1998, *Lipids* 100:161-166; Millar et al. 2000, *Trends Plant Sci.* 5:95-101).

[Para 5] Storage lipids in seeds are synthesized from carbohydrate-derived precursors. Plants have a complete glycolytic pathway in the cytosol (Plaxton 1996, *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 47:185-214) and it has been shown that a complete pathway also exists in the plastids of rapeseeds (Kang & Rawsthorne 1994, *Plant J.* 6:795-805). Sucrose is the primary source of carbon and energy, transported from the leaves into the developing seeds. During the storage phase of seeds, sucrose is converted in the cytosol to provide the metabolic precursors glucose-6-phosphate and pyruvate. These are transported into the plastids and converted into acetyl-CoA that serves as the primary precursor for the synthesis of fatty acids. Acetyl-CoA in the plastids is the central precursor for lipid biosynthesis. Acetyl-CoA can be formed in the plastids by different reactions and the exact contribution of each reaction is still being debated (Ohlrogge & Browse 1995, *Plant Cell* 7:957-970). It is however accepted that a large part of the acetyl-CoA is derived from glucose-6-phosphate and pyruvate that are imported from the cytoplasm into the plastids. Sucrose is produced in the source organs (leaves, or anywhere that photosynthesis occurs) and is transported to the developing seeds that are

also termed sink organs. In the developing seeds, sucrose is the precursor for all the storage compounds, i.e. starch, lipids, and partly the seed storage proteins. Therefore, it is clear that carbohydrate metabolism in which sucrose plays a central role is very important to the accumulation of seed storage compounds.

[Para 6] Storage compounds such as triacylglycerols (seed oil) serve as carbon and energy reserves, which are used during germination and growth of the young seedling. Seed (vegetable) oil is also an essential component of the human diet and a valuable commodity providing feed stocks for the chemical industry.

[Para 7] Although the lipid and fatty acid content and/or composition of seed oil can be modified by the traditional methods of plant breeding, the advent of recombinant DNA technology has allowed for easier manipulation of the seed oil content of a plant, and in some cases, has allowed for the alteration of seed oils in ways that could not be accomplished by breeding alone (see, e.g., Töpfer et al. 1995, *Science* 268:681-686). For example, introduction of a Δ^{12} -hydroxylase nucleic acid sequence into transgenic tobacco resulted in the introduction of a novel fatty acid, ricinoleic acid, into the tobacco seed oil (Van de Loo et al. 1995, *Proc. Natl. Acad. Sci USA* 92:6743-6747). Tobacco plants have also been engineered to produce low levels of petroselinic acid by the introduction and expression of an acyl-ACP desaturase from coriander (Cahoon et al. 1992, *Proc. Natl. Acad. Sci USA* 89:11184-11188).

[Para 8] The modification of seed oil content in plants has significant medical, nutritional and economic ramifications. With regard to the medical ramifications, the long chain fatty acids (C18 and longer) found in many seed oils have been linked to reductions in hypercholesterolemia and other clinical disorders related to coronary heart disease (Brenner 1976, *Adv. Exp. Med. Biol.* 83:85-101). Therefore, consumption of a plant having increased levels of these types of fatty acids may reduce the risk of heart disease. Enhanced levels of seed oil content also increase large-scale production of seed oils and thereby reduce the cost of these oils.

[Para 9] In order to increase or alter the levels of compounds such as seed oils in plants, nucleic acid sequences and proteins regulating lipid and fatty acid metabolism must be identified. As mentioned earlier, several desaturase nucleic acids such as the Δ^6 -desaturase nucleic acid, Δ^{12} -desaturase nucleic acid and acyl-ACP desaturase nucleic acid have been cloned and demonstrated to encode enzymes required for fatty acid synthesis in various plant species. Oleosin nucleic acid sequences from such different species as canola, soybean, carrot, pine and *Arabidopsis thaliana* have also been cloned and determined to encode proteins associated with the phospholipid monolayer membrane of oil bodies in those plants.

[Para 10] It has also been determined that two phytohormones, gibberellic acid (GA) and abscisic acid (ABA), are involved in overall regulatory processes in seed development (e.g. Ritchie & Gilroy 1998, *Plant Physiol.* 116:765-776; Arenas-Huertero et al. 2000, *Genes Dev.* 14:2085-2096). Both the

GA and ABA pathways are affected by okadaic acid, a protein phosphatase inhibitor (Kuo et al. 1996, *Plant Cell*. 8:259-269). The regulation of protein phosphorylation by kinases and phosphatases is accepted as a universal mechanism of cellular control (Cohen 1992, *Trends Biochem. Sci.* 17:408-413). Likewise, the plant hormones ethylene (e.g. Zhou et al. 1998, *Proc. Natl. Acad. Sci. USA* 95:10294-10299; Beaudoin et al. 2000, *Plant Cell* 2000:1103-1115) and auxin (e.g. Colon-Carmona et al. 2000, *Plant Physiol.* 124:1728-1738) are involved in controlling plant development as well.

[Para 11] Sucrose synthase (SUS) is one of the key enzymes involved in sucrose synthesis/metabolism, especially in non-photosynthetic tissues. This enzyme catalyses the reversible conversion of sucrose and UDP into UDP-glucose and fructose. Under normal growth conditions, Sus activity has been linked to many important plant processes, e.g. phloem loading/unloading (Martin, T., Frommer, W., Salanoubat, M. and Willmitzer, L. (1993) *Plant J.* 4, 367-377; Nolte, K. D. and Koch, K. E. (1993) *Plant Physiol.* 101, 899-905; Fu, H. and Park, W. D. (1995) *Plant Cell* 7, 1369-1385), and nodule function (Van Ghelue, M., Ribeiro, A., Solheim, B., Akkermans, A. D. L., Bisseling, T. and Pawlowski, K. (1996) *Mol. Gen. Genet.* 250, 437-446). Two SUS genes have been reported in *Arabidopsis*: SUS1 was previously identified as being responsive to anoxia and cold treatment (Martin, T., Frommer, W., Salanoubat, M. and Willmitzer, L. (1993) *Plant J.* 4, 367-377), whereas virtually no physiological information was available on SUS2 (Chopra, S., Del-Favero, J., Dolferus, R. and Jacobs, M. (1992) *Plant Mol. Biol.* 18, 131-134). SUS1 and SUS2 have found to be differentially regulated by environmental stresses via distinct ABA-independent sensing/transduction pathways (Dejardin, A., Sokolov, L.N. and Kleczkowski, L.A. (1999) *Biochem. J.* 344, 503-509).

[Para 12] Although several compounds are known that generally affect plant and seed development, there is a clear need to specifically identify factors that are more specific for the developmental regulation of storage compound accumulation and to identify genes which have the capacity to confer altered or increased oil production to its host plant and to other plant species. This invention discloses nucleic acid sequences from *Arabidopsis thaliana*, *Brassica napus*, *Glycine max*, *Zea mays*, *Oryza sativa*, *Triticum aestivum* or *Physcomitrella patens*. These nucleic acid sequences can be used to alter or increase the levels of seed storage compounds such as proteins, sugars and oils, in plants, including transgenic plants, such as canola, linseed, soybean, sunflower, maize, oat, rye, barley, wheat, rice, pepper, tagetes, cotton, oil palm, coconut palm, flax, castor and peanut, which are oilseed plants containing high amounts of lipid compounds.

SUMMARY OF THE INVENTION

[Para 13] The present invention provides novel isolated nucleic acid and amino acid sequences associated with the metabolism of seed storage compounds in plants, in particular with sequences that are SUS-like.

The present invention provides an isolated polypeptide, preferably a Lipid Metabolism Protein (LMP), or a biologically active portion thereof, comprising one, two, three or all amino acid sequences as described by SEQ ID NO: 39, SEQ ID NO: 40, SEQ ID NO: 41 or SEQ ID NO: 42.

The LMP protein of the present invention or a biologically active portion thereof are defined as polypeptides that are able to participate in the metabolism of compounds necessary for the production of seed storage compounds in plants, construction of cellular membranes in microorganisms or plants, or in the transport of molecules across these membranes. Regulatory proteins, such as DNA binding proteins, transcription factors, kinases, phosphatases, or protein members of metabolic pathways such as the lipid, starch and protein biosynthetic pathways, or membrane transport systems, may play a role in the biosynthesis of seed storage compounds. Examples of such activities are described herein (see putative annotations in Table 7). Examples of LMP-encoding nucleic acid sequences are set forth in Appendix A.

Preferably the isolated polypeptide of the present invention, preferably a functional LMP polypeptide, comprises two amino acid sequences selected from the groups consisting of the amino acid sequences as described by SEQ ID NO: 39, SEQ ID NO: 40, SEQ ID NO: 41 and SEQ ID NO: 42 as listed in the following table.

polypeptide comprising	SEQ ID NO: 39	SEQ ID NO: 40	SEQ ID NO: 41	SEQ ID NO: 42
SEQ ID NO: 39	x	x	x	x
SEQ ID NO: 40		x	x	x
SEQ ID NO: 41			x	x
SEQ ID NO: 42				x

Even more preferably the isolated polypeptide of the present invention, preferably a LMP polypeptide, comprises three amino acid sequences selected from the group consisting of the amino acid sequences as described by SEQ ID NO: 39, SEQ ID NO: 40, SEQ ID NO: 41 and SEQ ID NO: 42, e.g. a polypeptide comprising the amino acid sequences as described by SEQ ID NO: 39, 40, 41, a polypeptide comprising the amino acid sequences as described by SEQ ID NO: 39, 40, 42, a polypeptide comprising the amino acid sequences as described by SEQ ID NO: 39, 41, 42, a polypeptide comprising the amino acid sequences as described by SEQ ID NO: 40, 41, 42.

In a further preferred embodiment the isolated polypeptide of the present invention, preferably a functional LMP polypeptide comprises all four amino acid sequences as described by SEQ ID NO: 39, SEQ ID NO: 40, SEQ ID NO: 41 and SEQ ID NO: 42.

A further object of the present invention is an isolated nucleic acid sequence encoding a protein comprising one or more amino acid sequence as described by SEQ ID NO: 39, SEQ ID NO: 40, SEQ ID NO: 41 or SEQ ID NO: 42. A further object of the present invention is an isolated polypeptide encoded by said nucleic acid sequence.

The present invention provides furthermore an isolated nucleic acid preferably encoding for a LMP protein, comprising a polynucleotide sequence selected from the group consisting of:

- a. a polynucleotide sequence as described by SEQ ID NO: 1; SEQ ID NO: 3; SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35 or SEQ ID NO: 37;
- b. a polynucleotide sequence encoding a polypeptide as described by SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36 or SEQ ID NO: 38;
- c. a polynucleotide sequence having at least 70% sequence identity with the nucleic acid of a) or b) above;
- d. a polynucleotide sequence that is complementary to the nucleic acid of a) or b) above; and
- e. a polynucleotide sequence that hybridizes under stringent conditions to nucleic acid of a) or b) above.

The SEQ ID NO: 1 to 38 refer to the sequence identifier as used in the sequence listing according to WIPO Standard ST. 25. The sequence identifiers SEQ ID NO: 1 to 38 do not refer to the sequence identifier as used in Appendix A. Table 7 provides a concordance of the sequence identifier as used in Appendix A and the sequence listing according to the WIPO Standard ST. 25.

The present inventions provides furthermore an isolated polypeptide encoded by such polynucleotide sequences. In a preferred embodiment the isolated nucleic acid of the present invention encodes a polypeptide that functions as a modulator of a seed storage compound in microorganisms or plants. In a further preferred embodiment the isolated polypeptide of the present invention functions as a modulator of a seed storage compound in microorganisms or plants.

The present invention provides an expression vector containing a nucleic acid of the present invention, wherein the nucleic acid is operatively linked to a promoter selected from the group consisting of a seed-specific promoter, a root-specific promoter, and a non-tissue-specific promoter.

A further object of the present invention is a method of producing a transgenic plant having a modified level of a seed storage compound weight percentage compared to the empty vector control comprising,

- a. a first step of introduction into a plant cell of an expression vector containing a nucleic acid, and
- b. a further step of generating from the plant cell the transgenic plant, wherein the nucleic acid encodes a polypeptide that functions as a modulator of a seed storage compound in the plant, and wherein the nucleic acid comprises a polynucleotide of the present invention.

In a preferred embodiment the method of the present invention the nucleic acid comprises a polynucleotide sequence having at least 90% sequence identity with the polynucleotide sequence of a) or b) of Claim 1. In a further preferred embodiment the total seed oil content weight percentage is increased in the transgenic plant as compared to an empty vector control.

The present invention provides a method of modulating the level of a seed storage compound weight percentage in a plant comprising, modifying the expression of a nucleic acid in the plant, comprising

- a. a first step of introduction into a plant cell of an expression vector comprising a nucleic acid, and
- b. a further step of generating from the plant cell the transgenic plant, wherein the nucleic acid encodes a polypeptide that functions as a modulator of a seed storage compound in the plant wherein the nucleic acid comprises a polynucleotide sequence of the present invention.

In a further preferred embodiment of the method of the present invention, the total seed oil content weight percentage is increased in the transgenic plant as compared to an empty vector control.

The present invention furthermore provides transgenic plants made by a method of the present invention, preferably with an increase in the total seed oil content weight percentage in the transgenic plant as compared to an empty vector control. Preferably the transgenic plant of the present invention is selected from the group consisting of rapeseed, canola, linseed, soybean, sunflower, maize, oat, rye, barley, wheat, pepper, tagetes, cotton, oil palm, coconut palm, flax, castor, sugarbeet, rice and peanut. A further object of the present invention is a seed produced by the transgenic plant of the present invention, wherein the plant expresses the polypeptide that functions as a modulator of a seed storage compound and wherein the plant is true breeding for a modified level of seed storage compound weight percentage as compared to an empty vector control.

[Para 14] The present invention also provides an isolated nucleic acid from *Arabidopsis thaliana*, *Brassica napus*, *Glycine max*, *Zea mays*, *Oryza sativa*, *Triticum aestivum* or *Physcomitrella patens* encoding a Lipid Metabolism Protein (LMP), or a portion thereof. These sequences may be used to modify or increase lipids and fatty acids, cofactors and enzymes in microorganisms and plants.

[Para 15] *Arabidopsis* plants are known to produce considerable amounts of fatty acids like linoleic and linolenic acid (see, e.g., Table 6) and for their close similarity in many aspects (gene homology etc.) to the oil crop plant *Brassica*. Therefore, nucleic acid molecules originating from a plant like *Arabidopsis thaliana*, *Brassica napus*, *Glycine max*, *Zea mays*, *Oryza sativa*, *Triticum aestivum* or *Physcomitrella patens* or related organisms are especially suited to modify the lipid and fatty acid metabolism in a host, especially in microorganisms and plants. Furthermore, nucleic acids from the plant *Arabidopsis thaliana*, *Brassica napus*, *Glycine max*, *Zea mays*, *Oryza sativa*, *Triticum aestivum* or *Physcomitrella patens* or related organisms can be used to identify those DNA sequences and enzymes in other species which are useful to modify the biosynthesis of precursor molecules of fatty acids in the respective organisms.

[Para 16] The present invention further provides an isolated nucleic acid comprising a fragment of at least 15 nucleotides of a nucleic acid from a plant (*Arabidopsis thaliana*, *Brassica napus*, *Glycine max*, *Zea mays*, *Oryza sativa*, *Triticum aestivum*) or a moss (*Physcomitrella patens*) encoding a Lipid Metabolism Protein (LMP), or a portion thereof.

[Para 17] Also provided by the present invention are polypeptides encoded by the nucleic acids, and heterologous polypeptides comprising polypeptides encoded by the nucleic acids, and antibodies to those polypeptides.

[Para 18] The terms "heterologous nucleic acid sequence" or "heterologous DNA" are used interchangeably to refer to a nucleotide sequence, which is ligated to, or is manipulated to become ligated to, a nucleic acid sequence to which it is not ligated in nature, or to which it is ligated at a different location in nature. Heterologous DNA is not endogenous to the cell into which it is introduced, but has been obtained from another cell. Generally, although not necessarily, such heterologous DNA encodes RNA and proteins that are not normally produced by the cell into which it is expressed. A promoter, transcription regulating sequence or other genetic element is considered to be "heterologous" in relation to another sequence (e.g., encoding a marker sequence or an agronomically relevant trait) if said two sequences are not combined or differently operably linked their natural environment. Preferably, said sequences are not operably linked in their natural environment (i.e. come from different genes). Most preferably, said regulatory sequence is covalently joined and adjacent to a nucleic acid to which it is not adjacent in its natural environment.

[Para 19] Additionally, the present invention relates to and provides the use of LMP nucleic acids in the production of transgenic plants having a modified level, by e.g. 1, 2,5, 5, 7,5, 10, 12,5, 15, 17,5, 20, 22,5 or 25 % by weight or more, preferably by 5 % by weight or more, more preferably by 7,5 % by weight or more and even more preferably by 10 % by weight or more as compared to an empty vector control or composition, by e.g. 1, 2,5, 5, 7,5, 10, 12,5, 15, 17,5, 20, 22,5 or 25 % by weight or more, preferably by 5 % by weight or more, more preferably by 7,5 % by weight or more and even more preferably by 10 % by weight or more as compared to an empty vector control of a seed storage compound, in an preferred embodiment of seed oil.

[Para 20] The percent increases of a seed storage compound are generally determined compared to an empty vector control. An empty vector control is a transgenic plant, which has been transformed with the same vector or construct as a transgenic plant according to the present invention except for such a vector or construct lacking the nucleic acid sequences of the present inventions, preferably the nucleic acid sequences as disclosed in Appendix A. An empty vector control is shown for example in example 15.

[Para 21] In regard to an altered composition, the present invention can be used to, for example, increase the percentage of oleic acid relative to other plant oils. A method of producing a transgenic plant with a modified level or composition of a seed storage compound includes the steps of trans-

forming a plant cell with an expression vector comprising a LMP nucleic acid, and generating a plant with a modified level or composition of the seed storage compound from the plant cell. In a preferred embodiment, the plant is an oil producing species selected from the group consisting of canola, linseed, soybean, sunflower, maize, oat, rye, barley, wheat, rice, pepper, tagetes, cotton, oil palm, coconut palm, flax, castor, and peanut, for example.

[Para 22] According to the present invention, the compositions and methods described herein can be used to alter the composition of a LMP in a transgenic plant and to increase or decrease the level of a LMP in a transgenic plant comprising increasing or decreasing the expression of a LMP nucleic acid in the plant. Increased or decreased expression of the LMP nucleic acid can be achieved through transgenic overexpression, cosuppression approaches, antisense approaches, and in vivo mutagenesis of the LMP nucleic acid. The present invention can also be used to increase or decrease the level of a lipid in a seed oil, to increase or decrease the level of a fatty acid in a seed oil, or to increase or decrease the level of a starch in a seed or plant.

[Para 23] More specifically, the present invention includes and provides a method for increasing total oil content in a seeds comprising, by e.g. 1, 2,5, 5, 7,5, 10, 12,5, 15, 17,5, 20, 22,5 or 25 % by weight or more, preferably by 5 % by weight or more, more preferably by 7,5 % by weight or more and even more preferably by 10 % by weight or more as compared to an empty vector control: transforming a plant with a nucleic acid construct that comprises as operably linked components, a promoter and nucleic acid sequences capable of modulating the level of SUS-like mRNA or SUS-like protein, and growing the plant. Furthermore, the present invention includes and provides a method for increasing the level of oleic acid in a seed comprising: transforming a plant with a nucleic acid construct that comprises as operably linked components, a promoter, a structural nucleic acid sequence capable of increasing the level of oleic acid, and growing the plant.

[Para 24] The term "transgenic" or "recombinant" when used in reference to a cell or an organism (e.g., with regard to a barley plant or plant cell) refers to a cell or organism which contains a transgene, or whose genome has been altered by the introduction of a transgene. A transgenic organism or tissue may comprise one or more transgenic cells. Preferably, the organism or tissue is substantially consisting of transgenic cells (i.e., more than 80%, preferably 90%, more preferably 95%, most preferably 99% of the cells in said organism or tissue are transgenic). The term "transgene" as used herein refers to any nucleic acid sequence, which is introduced into the genome of a cell or which has been manipulated by experimental manipulations by man. Preferably, said sequence is resulting in a genome which is different from a naturally occurring organism (e.g., said sequence, if endogenous to said organism, is introduced into a location different from its natural location, or its copy number is increased or decreased). A transgene may be an "endogenous DNA sequence", "an "exogenous DNA sequence" (e.g., a foreign gene), or a "heterologous DNA sequence". The term "endogenous DNA sequence" refers to a nucleotide sequence, which is naturally found in the cell into which it is intro-

duced so long as it does not contain some modification (e.g., a point mutation, the presence of a selectable marker gene, etc.) relative to the naturally-occurring sequence.

[Para 25] The term "wild-type", "natural" or of "natural origin" means with respect to an organism, polypeptide, or nucleic acid sequence, that said organism is naturally occurring or available in at least one naturally occurring organism which is not changed, mutated, or otherwise manipulated by man.

[Para 26] Also included herein is a seed produced by a transgenic plant transformed by a LMP DNA sequence, wherein the seed contains the LMP DNA sequence and wherein the plant is true breeding for a modified level of a seed storage compound. The present invention additionally includes a seed oil produced by the aforementioned seed.

[Para 27] Further provided by the present invention are vectors comprising the nucleic acids, host cells containing the vectors, and descendent plant materials produced by transforming a plant cell with the nucleic acids and/or vectors.

[Para 28] According to the present invention, the compounds, compositions, and methods described herein can be used to increase or decrease the relative percentages of a lipid in a seed oil, increase or decrease the level of a lipid in a seed oil, or to increase or decrease the level of a fatty acid in a seed oil, or to increase or decrease the level of a starch or other carbohydrate in a seed or plant, or to increase or decrease the level of proteins in a seed or plant. The manipulations described herein can also be used to improve seed germination and growth of the young seedlings and plants and to enhance plant yield of seed storage compounds.

[Para 29] It is further provided a method of producing a higher or lower than normal or typical level of storage compound in a transgenic plant expressing a LMP nucleic acid from *Arabidopsis thaliana*, *Brassica napus*, *Glycine max*, *Zea mays*, *Oryza sativa*, *Triticum aestivum* or *Physcomitrella patens* in the transgenic plant, wherein the transgenic plant is *Arabidopsis thaliana*, *Brassica napus*, *Glycine max*, *Oryza sativa*, *Zea mays*, *Triticum aestivum*, *Hordeum vulgare*, *Helianthus annuus* or *Beta vulgaris* or a species different from *Arabidopsis thaliana*, *Brassica napus*, *Glycine max*, *Zea mays*, *Oryza sativa*, *Triticum aestivum* or *Physcomitrella patens*. Also included herein are compositions and methods of the modification of the efficiency of production of a seed storage compound. As used herein, where the phrase *Arabidopsis thaliana*, *Brassica napus*, *Glycine max*, *Oryza sativa*, *Zea mays*, *Triticum aestivum*, *Hordeum vulgare*, *Helianthus annuus* or *Beta vulgaris* is used, this also means *Arabidopsis thaliana* and/or *Brassica napus* and/or *Glycine max* and/or *Oryza sativa* and/or *Zea mays* and/or *Triticum aestivum* and/or *Hordeum vulgare* and/or *Helianthus annuus* and/or *Beta vulgaris*.

[Para 30] Accordingly, it is an object of the present invention to provide novel isolated LMP nucleic acids and isolated LMP amino acid sequences from *Arabidopsis thaliana*, *Brassica napus*, *Glycine max*, *Zea mays*, *Oryza sativa*, *Triticum aestivum* or *Physcomitrella patens* as well as active fragments, analogs, and orthologs thereof. Those active fragments, analogs, and orthologs can also be

from different plant species as one skilled in the art will appreciate that other plant species will also contain those or related nucleic acids.

[Para 31] It is another object of the present invention to provide transgenic plants having modified levels of seed storage compounds, and in particular, modified levels of a lipid, a fatty acid or a sugar. In a preferred embodiment the transgenic plants of the present invention have an increased seed oil content by e.g. 1, 2,5, 5, 7,5, 10, 12,5, 15, 17,5, 20, 22,5 or 25 % by weight or more, preferably by 5 % by weight or more, more preferably by 7,5 % by weight or more and even more preferably by 10 % by weight or more as compared to an empty vector control.

[Para 32] In a further preferred embodiment of the method of the present invention, the nucleic acid sequence as described by SEQ ID NO: 5, 9, 17, 21, 37 as well as active fragments, analogs, and orthologs thereof are overexpressed, preferably using a constitutive promoter, preferably a USP promoter, in plants in order have an increased seed oil content, preferably by 5 % by weight or more, further preferably by 5 % by weight or more, more preferably by 7,5 % by weight or more and even more preferably by 10 % by weight or more as compared to an empty vector control. A further object of the present invention are vectors containing of a nucleic acid sequence as described by SEQ ID NO: 5, 9, 17, 21, 37 as well as active fragments, analogs, and orthologs thereof, preferably operably linked to a constitutive promoter, preferably a USP promoter. A further object of the present invention are the plants obtained by the overexpression of the nucleic acid sequence as described by SEQ ID NO: 5, 9, 17, 21, 37 as well as active fragments, analogs, and orthologs thereof preferably operably linked to a constitutive promoter, preferably a USP promoter and showing an increased seed oil content, preferably by 5 % by weight or more, further preferably by 5 % by weight or more, more preferably by 7,5 % by weight or more and even more preferably by 10 % by weight or more as compared to the empty vector control.

[Para 33] The polynucleotides and polypeptides of the present invention, including agonists and/or fragments thereof, have also uses that include modulating plant growth, and potentially plant yield, preferably increasing plant growth under adverse conditions (drought, cold, light, UV). In addition, antagonists of the present invention may have uses that include modulating plant growth and/or yield, through preferably increasing plant growth and yield. In yet another embodiment, over-expression polypeptides of the present invention using a constitutive promoter may be useful for increasing plant yield under stress conditions (drought, light, cold, UV) by modulating light utilization efficiency. Moreover, polynucleotides and polypeptides of the present invention may improve seed germination and seed dormancy and, hence, improve plant growth and/or yield of seed storage compounds.

[Para 34] The isolated nucleic acid molecules of the present invention may further comprise an operably linked promoter or partial promoter region. The promoter can be a constitutive promoter, an inducible promoter, a tissue-specific promoter, or a combination thereof. The constitutive promoter can be, for example, the superpromoter (Ni et al., Plant J. 7:661-676, 1995; US5955646). The tissue-

specific promoter can, for example, be active in vegetative tissue or reproductive tissue, or any other specific tissue type within the plant or plant part, including seeds. Further, the tissue-specific promoter active in vegetative tissue can be a root-specific, shoot-specific, meristem-specific, or leaf-specific promoter. The isolated nucleic acid molecule of the present invention can still further comprise a 5' non-translated sequence, 3' non-translated sequence, introns, or combinations thereof.

[Para 35] The present invention also provides a method for increasing or decreasing the number and/or size of one or more plant organs of a plant by expressing an isolated nucleic acid from *Arabidopsis thaliana*, *Brassica napus*, *Glycine max*, *Zea mays*, *Oryza sativa*, *Triticum aestivum* or *Physcomitrella patens* encoding a Lipid Metabolism Protein (LMP), or a portion thereof. More specifically, seed size and/or seed number and/or weight might be manipulated. Moreover, root length can be increased. Longer roots can alleviate not only the effects of water depletion from soil but also improve plant anchorage/standability thus reducing lodging. Also, longer roots have the ability to cover a larger volume of soil and improve nutrient uptake. All of these advantages of altered root architecture have the potential to increase crop yield. Additionally, the number and size of leaves might be increased by the nucleic acid sequences provided in this application. This will have the advantage of improving photosynthetic light utilization efficiency by increasing photosynthetic light capture capacity and photosynthetic efficiency.

[Para 36] It is a further object of the present invention to provide methods for producing such aforementioned transgenic plants.

[Para 37] It is another object of the present invention to provide seeds and seed oils from such aforementioned transgenic plants.

[Para 38] These and other objects, features and advantages of the present invention will become apparent after a review of the following detailed description of the disclosed embodiments and the appended claims.

BRIEF DESCRIPTION OF THE DRAWINGS

[Para 39] The invention can be more fully understood from the following detailed description and the accompanying drawings and sequence listing which form a part of this application.

[Para 40] Figures 1A-C. Seq ID 1-3 - Nucleic acid sequence, open reading frame of the nucleic acid sequence, and the amino acid sequence of the open reading frame of the *Arabidopsis thaliana* gene AtSUS-2a.

[Para 41] Figures 2A-C. Seq ID 4-6 - Nucleic acid sequence, open reading frame of the nucleic acid sequence, and the amino acid sequence of the open reading frame of the *Arabidopsis thaliana* gene AtSUS-2b.

[Para 42] Figures 3A-C. Seq ID 7-9 - Nucleic acid sequence, open reading frame of the nucleic acid sequence, and the amino acid sequence of the open reading frame of the *Brassica napus* gene BnSUS-2277.

[Para 43] Figures 4A-C. Seq ID 10-12 - Nucleic acid sequence, open reading frame of the nucleic acid sequence, and the amino acid sequence of the open reading frame of the Glycine max gene GmSUS-2025.

[Para 44] Figures 5A-C. Seq ID 13-15 - Nucleic acid sequence, open reading frame of the nucleic acid sequence, and the amino acid sequence of the open reading frame of the Glycine max gene GmSUS-4819.

[Para 45] Figures 6A-C. Seq ID 16-18 - Nucleic acid sequence, open reading frame of the nucleic acid sequence, and the amino acid sequence of the open reading frame of the Zea mays gene ZmSUS-7691.

[Para 46] Figures 7A-C. Seq ID 19-21 - Nucleic acid sequence, open reading frame of the nucleic acid sequence, and the amino acid sequence of the open reading frame of the Zea mays gene ZmSUS-8057.

[Para 47] Figures 8A-C. Seq ID 22-24 Nucleic acid sequence, open reading frame of the nucleic acid sequence, and the amino acid sequence of the open reading frame of the Triticum aestivum gene TaSUS-4775.

[Para 48] Figures 9A-C. Seq ID 25-27 - Nucleic acid sequence, open reading frame of the nucleic acid sequence, and the amino acid sequence of the open reading frame of the Oryza sativa gene Os-SUS-9471.

[Para 49] Figures 10A-C. Seq ID 28-30 - Nucleic acid sequence, open reading frame of the nucleic acid sequence, and the amino acid sequence of the open reading frame of the Physcomitrella patens gene PpSUS-2.

[Para 50] Figure 11. T2, T3, and T4 total seed fatty acid data obtained with AtSUS-2b driven by a USP promoter (The genetic background of the transformed lines is Columbia-2; C24 represents a non-transformed high fatty acid content seed control [Columbia-24]; each circle represents the data obtained with 5 mg bulked seeds of one individual plant).

DETAILED DESCRIPTION OF THE INVENTION

[Para 51] The present invention may be understood more readily by reference to the following detailed description of the preferred embodiments of the invention and the Examples included therein.

[Para 52] Before the present compounds, compositions, and methods are disclosed and described, it is to be understood that this invention is not limited to specific nucleic acids, specific polypeptides, specific cell types, specific host cells, specific conditions, or specific methods, etc., as such may, of course, vary, and the numerous modifications and variations therein will be apparent to those skilled in the art. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only and is not intended to be limiting. As used in the specification and in the claims, "a" or "an" can mean one or more, depending upon the context in which it is used. Thus, for example, reference to "a cell" can mean that at least one cell can be utilized.

[Para 53] The present invention is based, in part, on the isolation and characterization of nucleic acid molecules encoding SUS-like LMPs from plants including *Arabidopsis thaliana*, canola (*Brassica napus*), soybean (*Glycine max*), maize (*Zea mays*), wheat (*Triticum aestivum*), rice (*Oryza sativa*) and the moss *Physcomitrella patens* and other related crop species like maize, barley, linseed, sugar beat or sunflower.

[Para 54] In accordance with the purpose(s) of this invention, as embodied and broadly described herein, this invention, in one aspect, provides an isolated nucleic acid from a plant including *Arabidopsis thaliana*, canola (*Brassica napus*), soybean (*Glycine max*), maize (*Zea mays*), wheat (*Triticum aestivum*), rice (*Oryza sativa*) and the moss *Physcomitrella patens* encoding a Lipid Metabolism Protein (LMP), or a portion thereof.

[Para 55] One aspect of the invention pertains to isolated nucleic acid molecules that encode LMP polypeptides or biologically active portions thereof, as well as nucleic acid fragments sufficient for use as hybridization probes or primers for the identification or amplification of an LMP-encoding nucleic acid (e.g., LMP DNA). As used herein, the term "nucleic acid molecule" is intended to include DNA molecules (e.g., cDNA or genomic DNA) and RNA molecules (e.g., mRNA) and analogs of the DNA or RNA generated using nucleotide analogs. This term also encompasses untranslated sequence located at both the 3' and 5' ends of the coding region of a gene: at least about 1000 nucleotides of sequence upstream from the 5' end of the coding region and at least about 200 nucleotides of sequence downstream from the 3' end of the coding region of the gene. The nucleic acid molecule can be single-stranded or double-stranded, but preferably is double-stranded DNA. An "isolated" nucleic acid molecule is one which is substantially separated from other nucleic acid molecules which are present in the natural source of the nucleic acid. Preferably, an "isolated" nucleic acid is substantially free of sequences that naturally flank the nucleic acid (i.e., sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated LMP nucleic acid molecule can contain less than about 5kb, 4kb, 3kb, 2kb, 1kb, 0.5kb, or 0.1kb of nucleotide sequences which naturally flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived (e.g., a *Arabidopsis thaliana*, *Brassica napus*, *Glycine max*, *Zea mays*, *Oryza sativa*, *Triticum aestivum* or *Physcomitrella patens* cell). Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular material, or culture medium when produced by recombinant techniques, or chemical precursors or other chemicals when chemically synthesized.

[Para 56] A nucleic acid molecule of the present invention, e.g., a nucleic acid molecule having a nucleotide sequence of Appendix A, or a portion thereof, can be isolated using standard molecular biology techniques and the sequence information provided herein. For example, an *Arabidopsis thaliana*, *Brassica napus*, *Glycine max*, *Zea mays*, *Oryza sativa*, *Triticum aestivum* or *Physcomitrella patens* LMP cDNA can be isolated from an *Arabidopsis thaliana*, *Brassica napus*, *Glycine max*, *Zea*

mays, *Oryza sativa*, *Triticum aestivum* or *Physcomitrella patens* library using all or portion of one of the sequences of Appendix A as a hybridization probe and standard hybridization techniques (e.g., as described in Sambrook et al. 1989, *Molecular Cloning: A Laboratory Manual*. 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY). Moreover, a nucleic acid molecule encompassing all or a portion of one of the sequences of Appendix A can be isolated by the polymerase chain reaction using oligonucleotide primers designed based upon this sequence (e.g., a nucleic acid molecule encompassing all or a portion of one of the sequences of Appendix A can be isolated by the polymerase chain reaction using oligonucleotide primers designed based upon this same sequence of Appendix A). For example, mRNA can be isolated from plant cells (e.g., by the guanidinium-thiocyanate extraction procedure of Chirgwin et al. 1979, *Biochemistry* 18:5294-5299) and cDNA can be prepared using reverse transcriptase (e.g., Moloney MLV reverse transcriptase, available from Gibco/BRL, Bethesda, MD; or AMV reverse transcriptase, available from Seikagaku America, Inc., St. Petersburg, FL). Synthetic oligonucleotide primers for polymerase chain reaction amplification can be designed based upon one of the nucleotide sequences shown in Appendix A. A nucleic acid of the invention can be amplified using cDNA or, alternatively, genomic DNA, as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques. The nucleic acid so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis. Furthermore, oligonucleotides corresponding to a LMP nucleotide sequence can be prepared by standard synthetic techniques, e.g., using an automated DNA synthesizer.

[Para 57] In a preferred embodiment, an isolated nucleic acid of the invention comprises one of the nucleotide sequences shown in Appendix A. The sequences of Appendix A correspond to the *Arabidopsis thaliana*, *Brassica napus*, *Glycine max*, *Zea mays*, *Oryza sativa*, *Triticum aestivum* or *Physcomitrella patens* LMP cDNAs of the invention. These cDNAs comprise sequences encoding LMPs (i.e., the “coding region,” indicated in Appendix A), as well as 5' untranslated sequences and 3' untranslated sequences. Alternatively, the nucleic acid molecules can comprise only the coding region of any of the sequences in Appendix A or can contain whole genomic fragments isolated from genomic DNA.

[Para 58] For the purposes of this application, it will be understood that each of the sequences set forth in Appendix A has an identifying entry number (e.g., BnSUS-2277). Each of these sequences may generally comprise three parts: a 5' upstream region, a coding region, and a downstream region. A coding region of these sequences is indicated as “ORF position” (Table 7).

[Para 59] In another preferred embodiment, an isolated nucleic acid molecule of the invention comprises a nucleic acid molecule, which is a complement of one of the nucleotide sequences shown in Appendix A, or a portion thereof. A nucleic acid molecule which is complementary to one of the nucleotide sequences shown in Appendix A is one which is sufficiently complementary to one of the

nucleotide sequences shown in Appendix A such that it can hybridize to one of the nucleotide sequences shown in Appendix A, thereby forming a stable duplex.

[Para 60] In still another preferred embodiment, an isolated nucleic acid molecule of the invention comprises a nucleotide sequence which is at least about 50-60%, preferably at least about 60-70%, more preferably at least about 70-80%, 80-90%, or 90-95%, and even more preferably at least about 95%, 96%, 97%, 98%, 99% or more homologous to a nucleotide sequence shown in Appendix A, or a portion thereof. In an additional preferred embodiment, an isolated nucleic acid molecule of the invention comprises a nucleotide sequence which hybridizes, e.g., hybridizes under stringent conditions, to one of the nucleotide sequences shown in Appendix A, or a portion thereof. These hybridization conditions include washing with a solution having a salt concentration of about 0.02 molar at pH 7 at about 60°C.

[Para 61] Moreover, the nucleic acid molecule of the invention can comprise only a portion of the coding region of one of the sequences in Appendix A, for example a fragment which can be used as a probe or primer or a fragment encoding a biologically active portion of a LMP. The nucleotide sequences determined from the cloning of the LMP genes from *Arabidopsis thaliana*, *Brassica napus*, *Glycine max*, *Zea mays*, *Oryza sativa*, *Triticum aestivum* or *Physcomitrella patens* allows for the generation of probes and primers designed for use in identifying and/or cloning LMP homologues in other cell types and organisms, as well as LMP homologues from other plants or related species. Therefore this invention also provides compounds comprising the nucleic acids disclosed herein, or fragments thereof. These compounds include the nucleic acids attached to a moiety. These moieties include, but are not limited to, detection moieties, hybridization moieties, purification moieties, delivery moieties, reaction moieties, binding moieties, and the like. The probe/primer typically comprises substantially purified oligonucleotide. The oligonucleotide typically comprises a region of nucleotide sequence that hybridizes under stringent conditions to at least about 12, preferably about 25, more preferably about 40, 50, or 75 consecutive nucleotides of a sense strand of one of the sequences set forth in Appendix A, an anti-sense sequence of one of the sequences set forth in Appendix A, or naturally occurring mutants thereof. Primers based on a nucleotide sequence of Appendix A can be used in PCR reactions to clone LMP homologues. Probes based on the LMP nucleotide sequences can be used to detect transcripts or genomic sequences encoding the same or homologous proteins. In preferred embodiments, the probe further comprises a label group attached thereto, e.g. the label group can be a radioisotope, a fluorescent compound, an enzyme, or an enzyme co-factor. Such probes can be used as a part of a genomic marker test kit for identifying cells which express a LMP, such as by measuring a level of a LMP-encoding nucleic acid in a sample of cells, e.g., detecting LMP mRNA levels or determining whether a genomic LMP gene has been mutated or deleted.

[Para 62] In one embodiment, the nucleic acid molecule of the invention encodes a protein or portion thereof which includes an amino acid sequence which is sufficiently homologous to an amino

acid encoded by a sequence of Appendix A such that the protein or portion thereof maintains the same or a similar function as the wild-type protein. As used herein, the language “sufficiently homologous” refers to proteins or portions thereof which have amino acid sequences which include a minimum number of identical or equivalent (e.g., an amino acid residue which has a similar side chain as an amino acid residue in one of the ORFs of a sequence of Appendix A) amino acid residues to an amino acid sequence such that the protein or portion thereof is able to participate in the metabolism of compounds necessary for the production of seed storage compounds in plants, construction of cellular membranes in microorganisms or plants, or in the transport of molecules across these membranes. Regulatory proteins, such as DNA binding proteins, transcription factors, kinases, phosphatases, or protein members of metabolic pathways such as the lipid, starch and protein biosynthetic pathways, or membrane transport systems, may play a role in the biosynthesis of seed storage compounds. Examples of such activities are described herein (see putative annotations in Table 7). Examples of LMP-encoding nucleic acid sequences are set forth in Appendix A.

[Para 63] As altered or increased sugar and/or fatty acid production is a general trait wished to be inherited into a wide variety of plants like maize, wheat, rye, oat, triticale, rice, barley, soybean, peanut, cotton, canola, manihot, pepper, sunflower, sugar beet and tagetes, solanaceous plants like potato, tobacco, eggplant, and tomato, Vicia species, pea, alfalfa, bushy plants (coffee, cacao, tea), Salix species, trees (oil palm, coconut) and perennial grasses and forage crops, these crop plants are also preferred target plants for genetic engineering as one further embodiment of the present invention.

[Para 64] Portions of proteins encoded by the LMP nucleic acid molecules of the invention are preferably biologically active portions of one of the LMPs. As used herein, the term “biologically active portion of a LMP” is intended to include a portion, e.g., a domain/ motif, of a LMP that participates in the metabolism of compounds necessary for the biosynthesis of seed storage lipids, or the construction of cellular membranes in microorganisms or plants, or in the transport of molecules across these membranes, or has an activity as set forth in Table 7. To determine whether a LMP or a biologically active portion thereof can participate in the metabolism of compounds necessary for the production of seed storage compounds and cellular membranes, an assay of enzymatic activity may be performed. Such assay methods are well known to those skilled in the art, and as described in Example 14 of the Exemplification.

[Para 65] Biologically active portions of a LMP include peptides comprising amino acid sequences derived from the amino acid sequence of a LMP (e.g., an amino acid sequence encoded by a nucleic acid of Appendix A or the amino acid sequence of a protein homologous to a LMP, which include fewer amino acids than a full length LMP or the full length protein which is homologous to a LMP) and exhibit at least one activity of a LMP. Typically, biologically active portions (peptides, e.g., peptides which are, for example, 5, 10, 15, 20, 30, 35, 36, 37, 38, 39, 40, 50, 100 or more amino acids in length) comprise a domain or motif with at least one activity of a LMP. Moreover, other biologically

active portions, in which other regions of the protein are deleted, can be prepared by recombinant techniques and evaluated for one or more of the activities described herein. Preferably, the biologically active portions of a LMP include one or more selected domains/motifs or portions thereof having biological activity.

[Para 66] Additional nucleic acid fragments encoding biologically active portions of a LMP can be prepared by isolating a portion of one of the sequences, expressing the encoded portion of the LMP or peptide (e.g., by recombinant expression in vitro) and assessing the activity of the encoded portion of the LMP or peptide.

[Para 67] The invention further encompasses nucleic acid molecules that differ from one of the nucleotide sequences shown in Appendix A (and portions thereof) due to degeneracy of the genetic code and thus encode the same LMP as that encoded by the nucleotide sequences shown in Appendix A. In a further embodiment, the nucleic acid molecule of the invention encodes a full length protein which is substantially homologous to an amino acid sequence of a polypeptide encoded by an open reading frame shown in Appendix A. In one embodiment, the full-length nucleic acid or protein or fragment of the nucleic acid or protein is from *Arabidopsis thaliana*, *Brassica napus*, *Glycine max*, *Zea mays*, *Oryza sativa*, *Triticum aestivum* or *Physcomitrella patens*.

[Para 68] In addition to the *Arabidopsis thaliana*, *Brassica napus*, *Glycine max*, *Zea mays*, *Oryza sativa*, *Triticum aestivum* or *Physcomitrella patens* LMP nucleotide sequences shown in Appendix A, it will be appreciated by those skilled in the art that DNA sequence polymorphisms that lead to changes in the amino acid sequences of LMPs may exist within a population (e.g., the *Arabidopsis thaliana*, *Brassica napus*, *Glycine max*, *Zea mays*, *Oryza sativa*, *Triticum aestivum* or *Physcomitrella patens* population). Such genetic polymorphism in the LMP gene may exist among individuals within a population due to natural variation. As used herein, the terms "gene" and "recombinant gene" refer to nucleic acid molecules comprising an open reading frame encoding a LMP, preferably a *Arabidopsis thaliana*, *Brassica napus*, *Glycine max*, *Zea mays*, *Oryza sativa*, *Triticum aestivum* or *Physcomitrella patens* LMP. Such natural variations can typically result in 1-40% variance in the nucleotide sequence of the LMP gene. Any and all such nucleotide variations and resulting amino acid polymorphisms in LMP that are the result of natural variation and that do not alter the functional activity of LMPs are intended to be within the scope of the invention.

[Para 69] Nucleic acid molecules corresponding to natural variants and non- *Arabidopsis thaliana*, *Brassica napus*, *Glycine max*, *Zea mays*, *Oryza sativa*, *Triticum aestivum* or *Physcomitrella patens* orthologs of the *Arabidopsis thaliana*, *Brassica napus*, *Glycine max*, *Zea mays*, *Oryza sativa*, *Triticum aestivum* or *Physcomitrella patens* LMP cDNA of the invention can be isolated based on their homology to *Arabidopsis thaliana*, *Brassica napus*, *Glycine max*, *Zea mays*, *Oryza sativa*, *Triticum aestivum* or *Physcomitrella patens* LMP nucleic acid disclosed herein using the *Arabidopsis thaliana*, *Brassica napus*, *Glycine max*, *Zea mays*, *Oryza sativa*, *Triticum aestivum* or *Physcomitrella patens* cDNA, or a

portion thereof, as a hybridization probe according to standard hybridization techniques under stringent hybridization conditions. As used herein, the term "orthologs" refers to two nucleic acids from different species, but that have evolved from a common ancestral gene by speciation. Normally, orthologs encode proteins having the same or similar functions. Accordingly, in another embodiment, an isolated nucleic acid molecule of the invention is at least 15 nucleotides in length and hybridizes under stringent conditions to the nucleic acid molecule comprising a nucleotide sequence of Appendix A. In other embodiments, the nucleic acid is at least 30, 50, 100, 250, or more nucleotides in length. As used herein, the term "hybridizes under stringent conditions" is intended to describe conditions for hybridization and washing under which nucleotide sequences at least 60% homologous to each other typically remain hybridized to each other. Preferably, the conditions are such that sequences at least about 65%, more preferably at least about 70%, and even more preferably at least about 75% or more homologous to each other typically remain hybridized to each other. Such stringent conditions are known to those skilled in the art and can be found in Current Protocols in Molecular Biology, John Wiley & Sons, N.Y., 1989: 6.3.1-6.3.6.

[Para 70] For the purposes of the invention hybridization means preferably hybridization under conditions equivalent to hybridization in 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO₄, 1 mM EDTA at 50°C with washing in 2 X SSC, 0.1% SDS at 50°C, more desirably in 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO₄, 1 mM EDTA at 50°C with washing in 1 X SSC, 0.1% SDS at 50°C, more desirably still in 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO₄, 1 mM EDTA at 50°C with washing in 0.5 X SSC, 0.1% SDS at 50°C, preferably in 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO₄, 1 mM EDTA at 50°C with washing in 0.1 X SSC, 0.1% SDS at 50°C, more preferably in 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO₄, 1 mM EDTA at 50°C with washing in 0.1 X SSC, 0.1% SDS at 65°C to a nucleic acid comprising 50 to 200 or more consecutive nucleotides.

[Para 71] A further preferred, non-limiting example of stringent hybridization conditions are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2 X SSC, 0.1% SDS at 50-65°C. Preferably, an isolated nucleic acid molecule of the invention that hybridizes under stringent conditions to a sequence of Appendix A corresponds to a naturally occurring nucleic acid molecule. As used herein, a "naturally-occurring" nucleic acid molecule refers to an RNA or DNA molecule having a nucleotide sequence that occurs in nature (e.g., encodes a natural protein). In one embodiment, the nucleic acid encodes a natural *Arabidopsis thaliana*, *Brassica napus*, *Glycine max*, *Zea mays*, *Oryza sativa*, *Triticum aestivum* or *Physcomitrella patens* LMP.

[Para 72] In addition to naturally-occurring variants of the LMP sequence that may exist in the population, the skilled artisan will further appreciate that changes can be introduced by mutation into a nucleotide sequence of Appendix A, thereby leading to changes in the amino acid sequence of the encoded LMP, without altering the functional ability of the LMP. For example, nucleotide substitu-

tions leading to amino acid substitutions at “non-essential” amino acid residues can be made in a sequence of Appendix A. A “non-essential” amino acid residue is a residue that can be altered from the wild-type sequence of one of the LMPs (Appendix A) without altering the activity of said LMP, whereas an “essential” amino acid residue is required for LMP activity. Other amino acid residues, however, (e.g., those that are not conserved or only semi-conserved in the domain having LMP activity) may not be essential for activity and thus are likely to be amenable to alteration without altering LMP activity.

[Para 73] Accordingly, another aspect of the invention pertains to nucleic acid molecules encoding LMPs that contain changes in amino acid residues that are not essential for LMP activity. Such LMPs differ in amino acid sequence from a sequence yet retain at least one of the LMP activities described herein. In one embodiment, the isolated nucleic acid molecule comprises a nucleotide sequence encoding a protein, wherein the protein comprises an amino acid sequence at least about 50% homologous to an amino acid sequence encoded by a nucleic acid of Appendix A and is capable of participation in the metabolism of compounds necessary for the production of seed storage compounds in *Arabidopsis thaliana*, *Brassica napus*, *Glycine max*, *Zea mays*, *Oryza sativa*, *Triticum aestivum* or *Physcomitrella patens*, or cellular membranes, or has one or more activities set forth in Table 7. Preferably, the protein encoded by the nucleic acid molecule is at least about 50-60% homologous to one of the sequences encoded by a nucleic acid of Appendix A, more preferably at least about 60-70% homologous to one of the sequences encoded by a nucleic acid of Appendix A, even more preferably at least about 70-80%, 80-90%, 90-95% homologous to one of the sequences encoded by a nucleic acid of Appendix A, and most preferably at least about 96%, 97%, 98%, or 99% homologous to one of the sequences encoded by a nucleic acid of Appendix A.

[Para 74] To determine the percent homology of two amino acid sequences (e.g., one of the sequences encoded by a nucleic acid of Appendix A and a mutant form thereof) or of two nucleic acids, the sequences are aligned for optimal comparison purposes (e.g., gaps can be introduced in the sequence of one protein or nucleic acid for optimal alignment with the other protein or nucleic acid). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in one sequence (e.g., one of the sequences encoded by a nucleic acid of Appendix A) is occupied by the same amino acid residue or nucleotide as the corresponding position in the other sequence (e.g., a mutant form of the sequence selected from the polypeptide encoded by a nucleic acid of Appendix A), then the molecules are homologous at that position (i.e., as used herein amino acid or nucleic acid “homology” is equivalent to amino acid or nucleic acid “identity”). The percent homology between the two sequences is a function of the number of identical positions shared by the sequences (i.e., % homology = numbers of identical positions/total numbers of positions x 100). The sequence identity can be generally based on any one of the full length sequences of Appendix A as 100 %.

[Para 75] For the purposes of the invention, the percent sequence identity between two nucleic acid or polypeptide sequences is determined using the Vector NTI 7.0 (PC) software package (InforMax, 7600 Wisconsin Ave., Bethesda, MD 20814). A gap-opening penalty of 15 and a gap extension penalty of 6.66 are used for determining the percent identity of two nucleic acids. A gap-opening penalty of 10 and a gap extension penalty of 0.1 are used for determining the percent identity of two polypeptides. All other parameters are set at the default settings. For purposes of a multiple alignment (Clustal W algorithm), the gap-opening penalty is 10, and the gap extension penalty is 0.05 with blosum62 matrix. It is to be understood that for the purposes of determining sequence identity when comparing a DNA sequence to an RNA sequence, a thymidine nucleotide sequence is equivalent to an uracil nucleotide.

[Para 76] An isolated nucleic acid molecule encoding a LMP homologous to a protein sequence encoded by a nucleic acid of Appendix A can be created by introducing one or more nucleotide substitutions, additions or deletions into a nucleotide sequence of Appendix A such that one or more amino acid substitutions, additions, or deletions are introduced into the encoded protein. Mutations can be introduced into one of the sequences of Appendix A by standard techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis. Preferably, conservative amino acid substitutions are made at one or more predicted non-essential amino acid residues. A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art. These families include amino acids with basic side chains (e.g., lysine, arginine, histidine), acidic side chains (e.g., aspartic acid, glutamic acid), uncharged polar side chains (e.g., glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (e.g., alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (e.g., threonine, valine, isoleucine) and aromatic side chains (e.g., tyrosine, phenylalanine, tryptophan, histidine). Thus, a predicted non-essential amino acid residue in a LMP is preferably replaced with another amino acid residue from the same side chain family. Alternatively, in another embodiment, mutations can be introduced randomly along all or part of a LMP coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for a LMP activity described herein to identify mutants that retain LMP activity. Following mutagenesis of one of the sequences of Appendix A, the encoded protein can be expressed recombinantly and the activity of the protein can be determined using, for example, assays described herein (see Examples 11-13 of the Exemplification).

[Para 77] LMPs are preferably produced by recombinant DNA techniques. For example, a nucleic acid molecule encoding the protein is cloned into an expression vector (as described above), the expression vector is introduced into a host cell (as described herein), and the LMP is expressed in the host cell. The LMP can then be isolated from the cells by an appropriate purification scheme using standard protein purification techniques. Alternative to recombinant expression, a LMP or peptide

thereof can be synthesized chemically using standard peptide synthesis techniques. Moreover, native LMP can be isolated from cells, for example using an anti-LMP antibody, which can be produced by standard techniques utilizing a LMP or fragment thereof of this invention.

[Para 78] The invention also provides LMP chimeric or fusion proteins. As used herein, a LMP “chimeric protein” or “fusion protein” comprises a LMP polypeptide operatively linked to a non-LMP polypeptide. An “LMP polypeptide” refers to a polypeptide having an amino acid sequence corresponding to a LMP, whereas a “non-LMP polypeptide” refers to a polypeptide having an amino acid sequence corresponding to a protein which is not substantially homologous to the LMP, e.g., a protein which is different from the LMP and which is derived from the same or a different organism. Within the fusion protein, the term “operatively linked” is intended to indicate that the LMP polypeptide and the non-LMP polypeptide are fused to each other so that both sequences fulfill the proposed function attributed to the sequence used. The non-LMP polypeptide can be fused to the N-terminus or C-terminus of the LMP polypeptide. For example, in one embodiment, the fusion protein is a GST-LMP (glutathione S-transferase) fusion protein in which the LMP sequences are fused to the C-terminus of the GST sequences. Such fusion proteins can facilitate the purification of recombinant LMPs. In another embodiment, the fusion protein is a LMP containing a heterologous signal sequence at its N-terminus. In certain host cells (e.g., mammalian host cells), expression and/or secretion of a LMP can be increased through use of a heterologous signal sequence.

[Para 79] Preferably, a LMP chimeric or fusion protein of the invention is produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, for example by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers that give rise to complementary overhangs between two consecutive gene fragments, which can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for example, Current Protocols in Molecular Biology, eds. Ausubel et al., John Wiley & Sons: 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (e.g., a GST polypeptide). An LMP-encoding nucleic acid can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the LMP.

[Para 80] In addition to the nucleic acid molecules encoding LMPs described above, another aspect of the invention pertains to isolated nucleic acid molecules that are antisense thereto. An “antisense” nucleic acid comprises a nucleotide sequence that is complementary to a “sense” nucleic acid encoding a protein, e.g., complementary to the coding strand of a double-stranded cDNA molecule or com-

plementary to an mRNA sequence. Accordingly, an antisense nucleic acid can hydrogen bond to a sense nucleic acid. The antisense nucleic acid can be complementary to an entire LMP coding strand, or to only a portion thereof. In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence encoding a LMP. The term "coding region" refers to the region of the nucleotide sequence comprising codons that are translated into amino acid residues (e.g., the entire coding region of BnSUS-2277 comprises nucleotides 93-2507). In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence encoding LMP. The term "noncoding region" refers to 5' and 3' sequences that flank the coding region that are not translated into amino acids (i.e., also referred to as 5' and 3' untranslated regions).

[Para 81] Given the coding strand sequences encoding LMP disclosed herein (e.g., the sequences set forth in Appendix A), antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of LMP mRNA, but more preferably is an oligonucleotide that is antisense to only a portion of the coding or noncoding region of LMP mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of LMP mRNA. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45, or 50 nucleotides in length. An antisense or sense nucleic acid of the invention can be constructed using chemical synthesis and enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (e.g., an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, e.g., phosphorothioate derivatives and acridine substituted nucleotides can be used. Examples of modified nucleotides which can be used to generate the antisense nucleic acid include 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxymethyl) uracil, 5-carboxymethylamino-methyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydro-uracil, beta-D-galactosylqueosine, inosine, N-6-isopentenyladenine, 1-methyl-guanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methyl-cytosine, N-6-adenine, 7-methylguanine, 5-methyl-aminomethyluracil, 5-methoxyamino-methyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyl-uracil, 5-methoxyuracil, 2-methylthio-N-6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diamino-purine. Alternatively, the antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (i.e., RNA transcribed from the

inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

[Para 82] In another variation of the antisense technology, a double-strand interfering RNA construct can be used to cause a down-regulation of the LMP mRNA level and LMP activity in transgenic plants. This requires transforming the plants with a chimeric construct containing a portion of the LMP sequence in the sense orientation fused to the antisense sequence of the same portion of the LMP sequence. A DNA linker region of variable length can be used to separate the sense and antisense fragments of LMP sequences in the construct.

[Para 83] The antisense nucleic acid molecules of the invention are typically administered to a cell or generated in situ such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a LMP to thereby inhibit expression of the protein, e.g., by inhibiting transcription and/or translation. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule which binds to DNA duplexes, through specific interactions in the major groove of the double helix. The antisense molecule can be modified such that it specifically binds to a receptor or an antigen expressed on a selected cell surface, e.g., by linking the antisense nucleic acid molecule to a peptide or an antibody which binds to a cell surface receptor or antigen. The antisense nucleic acid molecule can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of the antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong prokaryotic, viral, or eukaryotic including plant promoters are preferred.

[Para 84] In yet another embodiment, the antisense nucleic acid molecule of the invention is an anomeric nucleic acid molecule. An anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual units, the strands run parallel to each other (Gaultier et al. 1987, *Nucleic Acids Res.* 15:6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o-methyl-ribonucleotide (Inoue et al. 1987, *Nucleic Acids Res.* 15:6131-6148) or a chimeric RNA-DNA analogue (Inoue et al. 1987, *FEBS Lett.* 215:327-330).

[Para 85] In still another embodiment, an antisense nucleic acid of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity, which are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (e.g., hammerhead ribozymes (described in Haselhoff & Gerlach 1988, *Nature* 334:585-591)) can be used to catalytically cleave LMP mRNA transcripts to thereby inhibit translation of LMP mRNA. A ribozyme having specificity for a LMP-encoding nucleic acid can be designed based upon the nucleotide sequence of a LMP cDNA disclosed herein (i.e., Bn01 in Appendix A) or on the basis of a heterologous sequence to be isolated according to methods taught in this invention. For example, a derivative of a *Tetrahymena* L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a LMP-encoding mRNA

(see, e.g., Cech et al., U.S. Patent No. 4,987,071 and Cech et al., U.S. Patent No. 5,116,742). Alternatively, LMP mRNA can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules (see, e.g., Bartel, D. & Szostak J.W. 1993, *Science* 261:1411-1418).

[Para 86] Alternatively, LMP gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region of a LMP nucleotide sequence (e.g., a LMP promoter and/or enhancers) to form triple helical structures that prevent transcription of a LMP gene in target cells (See generally, Helene C. 1991, *Anticancer Drug Des.* 6:569-84; Helene C. et al. 1992, *Ann. N.Y. Acad. Sci.* 660:27-36; and Maher, L.J. 1992, *Bioassays* 14:807-15).

[Para 87] Another aspect of the invention pertains to vectors, preferably expression vectors, containing a nucleic acid encoding a LMP (or a portion thereof). As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid," which refers to a circular double stranded DNA loop into which additional DNA segments can be ligated. Another type of vector is a viral vector, wherein additional DNA segments can be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (e.g., bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (e.g., non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively linked. Such vectors are referred to herein as "expression vectors." In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. In the present specification, "plasmid" and "vector" can be used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors, such as viral vectors (e.g., replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

[Para 88] The recombinant expression vectors of the invention comprise a nucleic acid of the invention in a form suitable for expression of the nucleic acid in a host cell, which means that the recombinant expression vectors include one or more regulatory sequences, selected on the basis of the host cells to be used for expression, which is operatively linked to the nucleic acid sequence to be expressed. Within a recombinant expression vector, "operably linked" is intended to mean that the nucleotide sequence of interest is linked to the regulatory sequence(s) in a manner which allows for expression of the nucleotide sequence and both sequences are fused to each other so that each fulfills its proposed function (e.g., in an in vitro transcription/translation system or in a host cell when the vector is introduced into the host cell). The term "regulatory sequence" is intended to include promoters, enhancers, and other expression control elements (e.g., polyadenylation signals). Such regulatory sequences are described, for example, in Goeddel; *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990) or see: Gruber and Crosby, in: *Methods in*

Plant Molecular Biology and Biotechnology, CRC Press, Boca Raton, Florida, eds.: Glick & Thompson, Chapter 7, 89-108 including the references therein. Regulatory sequences include those that direct constitutive expression of a nucleotide sequence in many types of host cell and those that direct expression of the nucleotide sequence only in certain host cells or under certain conditions. It will be appreciated by those skilled in the art that the design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of protein desired, etc. The expression vectors of the invention can be introduced into host cells to thereby produce proteins or peptides, including fusion proteins or peptides, encoded by nucleic acids as described herein (e.g., LMPs, mutant forms of LMPs, fusion proteins, etc.).

[Para 89] The recombinant expression vectors of the invention can be designed for expression of LMPs in prokaryotic or eukaryotic cells. For example, LMP genes can be expressed in bacterial cells, insect cells (using baculovirus expression vectors), yeast and other fungal cells (see Romanos M.A. et al. 1992, Foreign gene expression in yeast: a review, *Yeast* 8:423-488; van den Hondel, C.A.M.J.J. et al. 1991, Heterologous gene expression in filamentous fungi, in: *More Gene Manipulations in Fungi*, Bennet & Lasure, eds., p. 396-428:Academic Press: San Diego; and van den Hondel & Punt 1991, Gene transfer systems and vector development for filamentous fungi, in: *Applied Molecular Genetics of Fungi*, Peberdy et al., eds., p. 1-28, Cambridge University Press: Cambridge), algae (Falciatore et al. 1999, *Marine Biotechnology* 1:239-251), ciliates of the types: Holotrichia, Peritrichia, Spirotrichia, Suctoria, Tetrahymena, Paramecium, Colpidium, Glaucoma, Platyophrya, Potomacus, Pseudocohnilembus, Euplotes, Engelmaniella, and Stylonychia, especially of the genus Stylonychia lemnae with vectors following a transformation method as described in WO 98/01572 and multicellular plant cells (see Schmidt & Willmitzer 1988, High efficiency *Agrobacterium tumefaciens*-mediated transformation of *Arabidopsis thaliana* leaf and cotyledon plants, *Plant Cell Rep.*:583-586); *Plant Molecular Biology and Biotechnology*, C Press, Boca Raton, Florida, chapter 6/7, S.71-119 (1993); White, Jenes et al., *Techniques for Gene Transfer*, in: *Transgenic Plants*, Vol. 1, Engineering and Utilization, eds.: Kung and Wu, Academic Press 1993, 128-43; Potrykus 1991, *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 42:205-225 (and references cited therein) or mammalian cells. Suitable host cells are discussed further in Goeddel, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA 1990). Alternatively, the recombinant expression vector can be transcribed and translated in vitro, for example using T7 promoter regulatory sequences and T7 polymerase.

[Para 90] Expression of proteins in prokaryotes is most often carried out with vectors containing constitutive or inducible promoters directing the expression of either fusion or non-fusion proteins. Fusion vectors add a number of amino acids to a protein encoded therein, usually to the amino terminus of the recombinant protein but also to the C-terminus or fused within suitable regions in the proteins. Such fusion vectors typically serve one or more of the following purposes: 1) to increase expression of recombinant protein; 2) to increase the solubility of the recombinant protein; and 3) to aid

in the purification of the recombinant protein by acting as a ligand in affinity purification. Often, in fusion expression vectors, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant protein to enable separation of the recombinant protein from the fusion moiety subsequent to purification of the fusion protein. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin, and enterokinase.

[Para 91] Typical fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith & Johnson 1988, Gene 67:31-40), pMAL (New England Biolabs, Beverly, MA) and pRIT5 (Pharmacia, Piscataway, NJ) which fuse glutathione S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein. In one embodiment, the coding sequence of the LMP is cloned into a pGEX expression vector to create a vector encoding a fusion protein comprising, from the N-terminus to the C-terminus, GST-thrombin cleavage site-X protein. The fusion protein can be purified by affinity chromatography using glutathione-agarose resin. Recombinant LMP unfused to GST can be recovered by cleavage of the fusion protein with thrombin.

[Para 92] Examples of suitable inducible non-fusion *E. coli* expression vectors include pTrc (Amann et al. 1988, Gene 69:301-315) and pET 11d (Studier et al. 1990, Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, California 60-89). Target gene expression from the pTrc vector relies on host RNA polymerase transcription from a hybrid trp-lac fusion promoter. Target gene expression from the pET 11d vector relies on transcription from a T7 gn10-lac fusion promoter mediated by a coexpressed viral RNA polymerase (T7 gn1). This viral polymerase is supplied by host strains BL21 (DE3) or HMS174 (DE3) from a resident prophage harboring a T7 gn1 gene under the transcriptional control of the lacUV 5 promoter.

[Para 93] One strategy to maximize recombinant protein expression is to express the protein in a host bacteria with an impaired capacity to proteolytically cleave the recombinant protein (Gottesman S. 1990, Gene Expression Technology: Methods in Enzymology 185:119-128, Academic Press, San Diego, California). Another strategy is to alter the nucleic acid sequence of the nucleic acid to be inserted into an expression vector so that the individual codons for each amino acid are those preferentially utilized in the bacterium chosen for expression (Wada et al. 1992, Nucleic Acids Res. 20:2111-2118). Such alteration of nucleic acid sequences of the invention can be carried out by standard DNA synthesis techniques.

[Para 94] In another embodiment, the LMP expression vector is a yeast expression vector. Examples of vectors for expression in yeast *S. cerevisiae* include pYepSec1 (Baldari et al. 1987, Embo J. 6:229-234), pMFa (Kurjan & Herskowitz 1982, Cell 30:933-943), pJRY88 (Schultz et al. 1987, Gene 54:113-123), and pYES2 (Invitrogen Corporation, San Diego, CA). Vectors and methods for the construction of vectors appropriate for use in other fungi, such as the filamentous fungi, include those detailed in: van den Hondel & Punt 1991, "Gene transfer systems and vector development for fila-

mentous fungi,” in: *Applied Molecular Genetics of Fungi*, Peberdy et al., eds., p. 1-28, Cambridge University Press: Cambridge.

[Para 95] Alternatively, the LMPs of the invention can be expressed in insect cells using baculovirus expression vectors. Baculovirus vectors available for expression of proteins in cultured insect cells (e.g., Sf 9 cells) include the pAc series (Smith et al. 1983, *Mol. Cell Biol.* 3:2156-2165) and the pVL series (Lucklow & Summers 1989, *Virology* 170:31-39).

[Para 96] In yet another embodiment, a nucleic acid of the invention is expressed in mammalian cells using a mammalian expression vector. Examples of mammalian expression vectors include pCDM8 (Seed 1987, *Nature* 329:840) and pMT2PC (Kaufman et al. 1987, *EMBO J.* 6:187-195). When used in mammalian cells, the expression vector's control functions are often provided by viral regulatory elements. For example, commonly used promoters are derived from polyoma, Adenovirus 2, cytomegalovirus, and Simian Virus 40. For other suitable expression systems for both prokaryotic and eukaryotic cells see chapters 16 and 17 of Sambrook, Fritsh and Maniatis, *Molecular Cloning: A Laboratory Manual*. 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

[Para 97] In another embodiment, the LMPs of the invention may be expressed in uni-cellular plant cells (such as algae, see Falciatore et al. (1999, *Marine Biotechnology* 1:239-251 and references therein) and plant cells from higher plants (e.g., the spermatophytes, such as crop plants). Examples of plant expression vectors include those detailed in: Becker, Kemper, Schell and Masterson (1992, “New plant binary vectors with selectable markers located proximal to the left border,” *Plant Mol. Biol.* 20:1195-1197) and Bevan (1984, “Binary *Agrobacterium* vectors for plant transformation,” *Nucleic Acids Res.* 12:8711-8721; “Vectors for Gene Transfer in Higher Plants”; in: *Transgenic Plants*, Vol. 1, Engineering and Utilization, eds.: Kung und R. Wu, Academic Press, 1993, S. 15-38).

[Para 98] A plant expression cassette preferably contains regulatory sequences capable to drive gene expression in plant cells and which are operably linked so that each sequence can fulfil its function such as termination of transcription, including polyadenylation signals. Preferred polyadenylation signals are those originating from *Agrobacterium tumefaciens* t-DNA such as the gene 3 known as octopine synthase of the Ti-plasmid pTiACH5 (Gielen et al. 1984, *EMBO J.* 3:835) or functional equivalents thereof but also all other terminators functionally active in plants are suitable.

[Para 99] As plant gene expression is very often not limited on transcriptional levels a plant expression cassette preferably contains other operably linked sequences like translational enhancers such as the overdrive-sequence containing the 5'-untranslated leader sequence from tobacco mosaic virus enhancing the protein per RNA ratio (Gallie et al. 1987, *Nucleic Acids Res.* 15:8693-8711).

[Para 100] Plant gene expression has to be operably linked to an appropriate promoter conferring gene expression in a timely, cell or tissue specific manner. Preferred are promoters driving constitutive expression (Benfey et al. 1989, *EMBO J.* 8:2195-2202) like those derived from plant viruses like

the 35S CAMV (Franck et al. 1980, Cell 21:285-294), the 19S CaMV (see also US 5,352,605 and WO 84/02913) or plant promoters like those from Rubisco small subunit described in US 4,962,028. Even more preferred are seed-specific promoters driving expression of LMP proteins during all or selected stages of seed development. Seed-specific plant promoters are known to those of ordinary skill in the art and are identified and characterized using seed-specific mRNA libraries and expression profiling techniques. Seed-specific promoters include the napin-gene promoter from rapeseed (US 5,608,152), the USP-promoter from *Vicia faba* (Baeumlein et al. 1991, Mol. Gen. Genetics 225:459-67), the oleosin-promoter from *Arabidopsis* (WO 98/45461), the phaseolin-promoter from *Phaseolus vulgaris* (US 5,504,200), the Bce4-promoter from *Brassica* (WO9113980) or the legumin B4 promoter (LeB4; Baeumlein et al. 1992, Plant J. 2:233-239) as well as promoters conferring seed specific expression in monocot plants like maize, barley, wheat, rye, rice etc. Suitable promoters to note are the *lpt2* or *lpt1*-gene promoter from barley (WO 95/15389 and WO 95/23230) or those described in WO 99/16890 (promoters from the barley hordein-gene, the rice glutelin gene, the rice oryzin gene, the rice prolamin gene, the wheat gliadin gene, wheat glutelin gene, the maize zein gene, the oat glutelin gene, the *Sorghum kasirin*-gene, and the rye secalin gene).

[Para 101] Plant gene expression can also be facilitated via an inducible promoter (for a review see Gatz 1997, Annu. Rev. Plant Physiol. Plant Mol. Biol. 48:89-108). Chemically inducible promoters are especially suitable if gene expression is desired in a time specific manner. Examples for such promoters are a salicylic acid inducible promoter (WO 95/19443), a tetracycline inducible promoter (Gatz et al. 1992, Plant J. 2:397-404), and an ethanol inducible promoter (WO 93/21334).

[Para 102] Promoters responding to biotic or abiotic stress conditions are also suitable promoters such as the pathogen inducible PRP1-gene promoter (Ward et al., 1993, Plant. Mol. Biol. 22:361-366), the heat inducible hsp80-promoter from tomato (US 5,187,267), cold inducible alpha-amylase promoter from potato (WO 96/12814) or the wound-inducible pinII-promoter (EP 375091).

[Para 103] Other preferred sequences for use in plant gene expression cassettes are targeting-sequences necessary to direct the gene-product in its appropriate cell compartment (for review see Kermode 1996, Crit. Rev. Plant Sci. 15:285-423 and references cited therein) such as the vacuole, the nucleus, all types of plastids like amyloplasts, chloroplasts, chromoplasts, the extracellular space, mitochondria, the endoplasmic reticulum, oil bodies, peroxisomes and other compartments of plant cells. Also especially suited are promoters that confer plastid-specific gene expression, as plastids are the compartment where precursors and some end products of lipid biosynthesis are synthesized. Suitable promoters such as the viral RNA-polymerase promoter are described in WO 95/16783 and WO 97/06250 and the *clpP*-promoter from *Arabidopsis* described in WO 99/46394.

[Para 104] The invention further provides a recombinant expression vector comprising a DNA molecule of the invention cloned into the expression vector in an antisense orientation. That is, the DNA molecule is operatively linked to a regulatory sequence in a manner that allows for expression (by

transcription of the DNA molecule) of an RNA molecule that is antisense to LMP mRNA. Regulatory sequences operatively linked to a nucleic acid cloned in the antisense orientation can be chosen which direct the continuous expression of the antisense RNA molecule in a variety of cell types, for instance viral promoters and/or enhancers, or regulatory sequences can be chosen which direct constitutive, tissue specific or cell type specific expression of antisense RNA. The antisense expression vector can be in the form of a recombinant plasmid, phagemid or attenuated virus in which antisense nucleic acids are produced under the control of a high efficiency regulatory region, the activity of which can be determined by the cell type into which the vector is introduced. For a discussion of the regulation of gene expression using antisense genes see Weintraub et al. (1986, Antisense RNA as a molecular tool for genetic analysis, Reviews - Trends in Genetics, Vol. 1) and Mol et al. (1990, FEBS Lett. 268:427-430).

[Para 105] Another aspect of the invention pertains to host cells into which a recombinant expression vector of the invention has been introduced. The terms "host cell" and "recombinant host cell" are used interchangeably herein. It is to be understood that such terms refer not only to the particular subject cell but also to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein. A host cell can be any prokaryotic or eukaryotic cell. For example, a LMP can be expressed in bacterial cells, insect cells, fungal cells, mammalian cells (such as Chinese hamster ovary cells (CHO) or COS cells), algae, ciliates, or plant cells. Other suitable host cells are known to those skilled in the art.

[Para 106] Vector DNA can be introduced into prokaryotic or eukaryotic cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection," "conjugation," and "transduction" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid (e.g., DNA) into a host cell, including calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, lipofection, natural competence, chemical-mediated transfer, or electroporation. Suitable methods for transforming or transfecting host cells including plant cells can be found in Sambrook et al. (1989, Molecular Cloning: A Laboratory Manual, 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY) and other laboratory manuals such as Methods in Molecular Biology 1995, Vol. 44, Agrobacterium protocols, ed: Gartland and Davey, Humana Press, Totowa, New Jersey.

[Para 107] For stable transfection of mammalian and plant cells, it is known that, depending upon the expression vector and transfection technique used, only a small fraction of cells may integrate the foreign DNA into their genome. In order to identify and select these integrants, a gene that encodes a selectable marker (e.g., resistance to antibiotics) is generally introduced into the host cells along with the gene of interest. Preferred selectable markers include those that confer resistance to drugs, such as

G418, hygromycin, kanamycin and methotrexate or in plants that confer resistance towards an herbicide such as glyphosate or glufosinate. A nucleic acid encoding a selectable marker can be introduced into a host cell on the same vector as that encoding a LMP or can be introduced on a separate vector. Cells stably transfected with the introduced nucleic acid can be identified by, for example, drug selection (e.g., cells that have incorporated the selectable marker gene will survive, while the other cells die).

[Para 108] To create a homologous recombinant microorganism, a vector is prepared which contains at least a portion of a LMP gene into which a deletion, addition or substitution has been introduced to thereby alter, e.g., functionally disrupt, the LMP gene. Preferably, this LMP gene is an *Arabidopsis thaliana*, *Brassica napus*, *Glycine max*, *Zea mays*, *Oryza sativa*, *Triticum aestivum* or *Physcomitrella patens* LMP gene, but it can be a homologue from a related plant or even from a mammalian, yeast, or insect source. In a preferred embodiment, the vector is designed such that, upon homologous recombination, the endogenous LMP gene is functionally disrupted (i.e., no longer encodes a functional protein; also referred to as a knock-out vector). Alternatively, the vector can be designed such that, upon homologous recombination, the endogenous LMP gene is mutated or otherwise altered but still encodes functional protein (e.g., the upstream regulatory region can be altered to thereby alter the expression of the endogenous LMP). To create a point mutation via homologous recombination, DNA-RNA hybrids can be used in a technique known as chimeraplasty (Cole-Strauss et al. 1999, *Nucleic Acids Res.* 27:1323-1330 and Kmiec 1999, *American Scientist* 87:240-247). Homologous recombination procedures in *Arabidopsis thaliana* or other crops are also well known in the art and are contemplated for use herein.

[Para 109] In a homologous recombination vector, the altered portion of the LMP gene is flanked at its 5' and 3' ends by additional nucleic acid of the LMP gene to allow for homologous recombination to occur between the exogenous LMP gene carried by the vector and an endogenous LMP gene in a microorganism or plant. The additional flanking LMP nucleic acid is of sufficient length for successful homologous recombination with the endogenous gene. Typically, several hundreds of base pairs up to kilobases of flanking DNA (both at the 5' and 3' ends) are included in the vector (see e.g., Thomas & Capecchi 1987, *Cell* 51:503, for a description of homologous recombination vectors). The vector is introduced into a microorganism or plant cell (e.g., via polyethyleneglycol mediated DNA). Cells in which the introduced LMP gene has homologously recombined with the endogenous LMP gene are selected using art-known techniques.

[Para 110] In another embodiment, recombinant microorganisms can be produced which contain selected systems which allow for regulated expression of the introduced gene. For example, inclusion of a LMP gene on a vector placing it under control of the lac operon permits expression of the LMP gene only in the presence of IPTG. Such regulatory systems are well known in the art.

[Para 111] A host cell of the invention, such as a prokaryotic or eukaryotic host cell in culture can be used to produce (i.e., express) a LMP. Accordingly, the invention further provides methods for producing LMPs using the host cells of the invention. In one embodiment, the method comprises culturing a host cell of the invention (into which a recombinant expression vector encoding a LMP has been introduced, or which contains a wild-type or altered LMP gene in its genome) in a suitable medium until LMP is produced. In another embodiment, the method further comprises isolating LMPs from the medium or the host cell.

[Para 112] Another aspect of the invention pertains to isolated LMPs, and biologically active portions thereof. An “isolated” or “purified” protein or biologically active portion thereof is substantially free of cellular material when produced by recombinant DNA techniques, or chemical precursors or other chemicals when chemically synthesized. The language “substantially free of cellular material” includes preparations of LMP in which the protein is separated from cellular components of the cells in which it is naturally or recombinantly produced. In one embodiment, the language “substantially free of cellular material” includes preparations of LMP having less than about 30% (by dry weight) of non-LMP (also referred to herein as a “contaminating protein”), more preferably less than about 20% of non-LMP, still more preferably less than about 10% of non-LMP, and most preferably less than about 5% non-LMP. When the LMP or biologically active portion thereof is recombinantly produced, it is also preferably substantially free of culture medium, i.e., culture medium represents less than about 20%, more preferably less than about 10%, and most preferably less than about 5% of the volume of the protein preparation. The language “substantially free of chemical precursors or other chemicals” includes preparations of LMP in which the protein is separated from chemical precursors or other chemicals that are involved in the synthesis of the protein. In one embodiment, the language “substantially free of chemical precursors or other chemicals” includes preparations of LMP having less than about 30% (by dry weight) of chemical precursors or non-LMP chemicals, more preferably less than about 20% chemical precursors or non-LMP chemicals, still more preferably less than about 10% chemical precursors or non-LMP chemicals, and most preferably less than about 5% chemical precursors or non-LMP chemicals. In preferred embodiments, isolated proteins or biologically active portions thereof lack contaminating proteins from the same organism from which the LMP is derived. Typically, such proteins are produced by recombinant expression of, for example, an *Arabidopsis thaliana*, *Brassica napus*, *Glycine max*, *Zea mays*, *Oryza sativa*, *Triticum aestivum* or *Physcomitrella patens* LMP in other plants than *Arabidopsis thaliana*, *Brassica napus*, *Glycine max*, *Zea mays*, *Oryza sativa*, *Triticum aestivum* or the moss *Physcomitrella patens* or microorganisms, algae or fungi.

[Para 113] An isolated LMP or a portion thereof of the invention can participate in the metabolism of compounds necessary for the production of seed storage compounds in *Arabidopsis thaliana*, *Brassica napus*, *Glycine max*, *Zea mays*, *Oryza sativa*, *Triticum aestivum*, or *Physcomitrella patens* or of cellular membranes, or has one or more of the activities set forth in Table 7. In preferred embodiments,

the protein or portion thereof comprises an amino acid sequence which is sufficiently homologous to an amino acid sequence encoded by a nucleic acid of Appendix A such that the protein or portion thereof maintains the ability to participate in the metabolism of compounds necessary for the construction of cellular membranes in *Arabidopsis thaliana*, *Brassica napus*, *Glycine max*, *Zea mays*, *Oryza sativa*, *Triticum aestivum* or *Physcomitrella patens*, or in the transport of molecules across these membranes. The portion of the protein is preferably a biologically active portion as described herein. In another preferred embodiment, a LMP of the invention has an amino acid sequence encoded by a nucleic acid of Appendix A. In yet another preferred embodiment, the LMP has an amino acid sequence which is encoded by a nucleotide sequence which hybridizes, e.g., hybridizes under stringent conditions, to a nucleotide sequence of Appendix A. In still another preferred embodiment, the LMP has an amino acid sequence which is encoded by a nucleotide sequence that is at least about 50-60%, preferably at least about 60-70%, more preferably at least about 70-80%, 80-90%, 90-95%, and even more preferably at least about 96%, 97%, 98%, 99% or more homologous to one of the amino acid sequences encoded by a nucleic acid of Appendix A. The preferred LMPs of the present invention also preferably possess at least one of the LMP activities described herein. For example, a preferred LMP of the present invention includes an amino acid sequence encoded by a nucleotide sequence which hybridizes, e.g., hybridizes under stringent conditions, to a nucleotide sequence of Appendix A, and which can participate in the metabolism of compounds necessary for the construction of cellular membranes in *Arabidopsis thaliana*, *Brassica napus*, *Glycine max*, *Zea mays*, *Oryza sativa*, *Triticum aestivum* or *Physcomitrella patens*, or in the transport of molecules across these membranes, or which has one or more of the activities set forth in Table 7.

[Para 114] In other embodiments, the LMP is substantially homologous to an amino acid sequence encoded by a nucleic acid of Appendix A and retains the functional activity of the protein of one of the sequences encoded by a nucleic acid of Appendix A yet differs in amino acid sequence due to natural variation or mutagenesis, as described in detail above. Accordingly, in another embodiment, the LMP is a protein which comprises an amino acid sequence which is at least about 50-60%, preferably at least about 60-70%, and more preferably at least about 70-80, 80-90, 90-95%, and most preferably at least about 96%, 97%, 98%, 99% or more homologous to an entire amino acid sequence and which has at least one of the LMP activities described herein. In another embodiment, the invention pertains to a full *Arabidopsis thaliana*, *Brassica napus*, *Glycine max*, *Zea mays*, *Oryza sativa*, *Triticum aestivum*, or *Physcomitrella patens* protein which is substantially homologous to an entire amino acid sequence encoded by a nucleic acid of Appendix A.

[Para 115] Dominant negative mutations or trans-dominant suppression can be used to reduce the activity of a LMP in transgenics seeds in order to change the levels of seed storage compounds. To achieve this a mutation that abolishes the activity of the LMP is created and the inactive non-functional LMP gene is overexpressed in the transgenic plant. The inactive trans-dominant LMP pro-

tein competes with the active endogenous LMP protein for substrate or interactions with other proteins and dilutes out the activity of the active LMP. In this way the biological activity of the LMP is reduced without actually modifying the expression of the endogenous LMP gene. This strategy was used by Pontier et al to modulate the activity of plant transcription factors (Pontier D, Miao ZH, Lam E, *Plant J* 2001 Sep. 27(6): 529-38, Trans-dominant suppression of plant TGA factors reveals their negative and positive roles in plant defense responses).

[Para 116] Homologues of the LMP can be generated by mutagenesis, e.g., discrete point mutation or truncation of the LMP. As used herein, the term "homologue" refers to a variant form of the LMP that acts as an agonist or antagonist of the activity of the LMP. An agonist of the LMP can retain substantially the same, or a subset, of the biological activities of the LMP. An antagonist of the LMP can inhibit one or more of the activities of the naturally occurring form of the LMP, by, for example, competitively binding to a downstream or upstream member of the cell membrane component metabolic cascade which includes the LMP, or by binding to a LMP which mediates transport of compounds across such membranes, thereby preventing translocation from taking place.

[Para 117] In an alternative embodiment, homologues of the LMP can be identified by screening combinatorial libraries of mutants, e.g., truncation mutants, of the LMP for LMP agonist or antagonist activity. In one embodiment, a variegated library of LMP variants is generated by combinatorial mutagenesis at the nucleic acid level and is encoded by a variegated gene library. A variegated library of LMP variants can be produced by, for example, enzymatically ligating a mixture of synthetic oligonucleotides into gene sequences such that a degenerate set of potential LMP sequences is expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (e.g., for phage display) containing the set of LMP sequences therein. There are a variety of methods that can be used to produce libraries of potential LMP homologues from a degenerate oligonucleotide sequence. Chemical synthesis of a degenerate gene sequence can be performed in an automatic DNA synthesizer, and the synthetic gene then ligated into an appropriate expression vector. Use of a degenerate set of genes allows for the provision, in one mixture, of all of the sequences encoding the desired set of potential LMP sequences. Methods for synthesizing degenerate oligonucleotides are known in the art (see, e.g., Narang 1983, *Tetrahedron* 39:3; Itakura et al. 1984, *Annu. Rev. Biochem.* 53:323; Itakura et al. 1984, *Science* 198:1056; Ike et al. 1983, *Nucleic Acids Res.* 11:477).

[Para 118] In addition, libraries of fragments of the LMP coding sequences can be used to generate a variegated population of LMP fragments for screening and subsequent selection of homologues of a LMP. In one embodiment, a library of coding sequence fragments can be generated by treating a double stranded PCR fragment of a LMP coding sequence with a nuclease under conditions wherein nicking occurs only about once per molecule, denaturing the double stranded DNA, renaturing the DNA to form double stranded DNA which can include sense/antisense pairs from different nicked products, removing single stranded portions from reformed duplexes by treatment with S1 nuclease,

and ligating the resulting fragment library into an expression vector. By this method, an expression library can be derived which encodes N-terminal, C-terminal and internal fragments of various sizes of the LMP.

[Para 119] Several techniques are known in the art for screening gene products of combinatorial libraries made by point mutations or truncation, and for screening cDNA libraries for gene products having a selected property. Such techniques are adaptable for rapid screening of the gene libraries generated by the combinatorial mutagenesis of LMP homologues. The most widely used techniques, which are amenable to high through-put analysis, for screening large gene libraries typically include cloning the gene library into replicable expression vectors, transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates isolation of the vector encoding the gene whose product was detected. Recursive ensemble mutagenesis (REM), a new technique that enhances the frequency of functional mutants in the libraries, can be used in combination with the screening assays to identify LMP homologues (Arkin & Yourvan 1992, Proc. Natl. Acad. Sci. USA 89:7811-7815; Delgrave et al. 1993, Protein Engineering 6:327-331).

[Para 120] In another embodiment, cell based assays can be exploited to analyze a variegated LMP library, using methods well known in the art.

[Para 121] The nucleic acid molecules, proteins, protein homologues, fusion proteins, primers, vectors, and host cells described herein can be used in one or more of the following methods: identification of *Arabidopsis thaliana*, *Brassica napus*, *Glycine max*, *Zea mays*, *Oryza sativa*, *Triticum aestivum*, or *Physcomitrella patens* and related organisms; mapping of genomes of organisms related to *Arabidopsis thaliana*, *Brassica napus*, *Glycine max*, *Zea mays*, *Oryza sativa*, *Triticum aestivum* or *Physcomitrella patens*; identification and localization of *Arabidopsis thaliana*, *Brassica napus*, *Glycine max*, *Zea mays*, *Oryza sativa*, *Triticum aestivum* or *Physcomitrella patens* sequences of interest; evolutionary studies; determination of LMP regions required for function; modulation of a LMP activity; modulation of the metabolism of one or more cell functions; modulation of the transmembrane transport of one or more compounds; and modulation of seed storage compound accumulation.

[Para 122] The plant *Arabidopsis thaliana* represents one member of higher (or seed) plants. It is related to other plants such as *Brassica napus*, *Glycine max*, *Zea mays*, *Oryza sativa*, or *Triticum aestivum* which require light to drive photosynthesis and growth. Plants like *Arabidopsis thaliana*, *Brassica napus*, *Glycine max*, *Zea mays*, *Oryza sativa*, *Triticum aestivum* or the moss *Physcomitrella patens* share a high degree of homology on the DNA sequence and polypeptide level, allowing the use of heterologous screening of DNA molecules with probes evolving from other plants or organisms, thus enabling the derivation of a consensus sequence suitable for heterologous screening or functional annotation and prediction of gene functions in third species. The ability to identify such functions can therefore have significant relevance, e.g., prediction of substrate specificity of enzymes. Further,

these nucleic acid molecules may serve as reference points for the mapping of *Arabidopsis* genomes, or of genomes of related organisms.

[Para 123] The LMP nucleic acid molecules of the invention have a variety of uses. First, the nucleic acid and protein molecules of the invention may serve as markers for specific regions of the genome. This has utility not only in the mapping of the genome, but also for functional studies of *Arabidopsis thaliana*, *Brassica napus*, *Glycine max*, *Zea mays*, *Oryza sativa*, *Triticum aestivum* or *Physcomitrella patens* proteins. For example, to identify the region of the genome to which a particular *Arabidopsis thaliana*, *Brassica napus*, *Glycine max*, *Zea mays*, *Oryza sativa*, *Triticum aestivum* or *Physcomitrella patens* DNA-binding protein binds, the *Arabidopsis thaliana*, *Brassica napus*, *Glycine max*, *Zea mays*, *Oryza sativa*, *Triticum aestivum*, or *Physcomitrella patens* genome could be digested, and the fragments incubated with the DNA-binding protein. Those which bind the protein may be additionally probed with the nucleic acid molecules of the invention, preferably with readily detectable labels; binding of such a nucleic acid molecule to the genome fragment enables the localization of the fragment to the genome map of *Arabidopsis thaliana*, *Brassica napus*, *Glycine max*, *Zea mays*, *Oryza sativa*, *Triticum aestivum*, or *Physcomitrella patens*, and, when performed multiple times with different enzymes, facilitates a rapid determination of the nucleic acid sequence to which the protein binds. Further, the nucleic acid molecules of the invention may be sufficiently homologous to the sequences of related species such that these nucleic acid molecules may serve as markers for the construction of a genomic map in related plants.

[Para 124] The LMP nucleic acid molecules of the invention are also useful for evolutionary and protein structural studies. The metabolic and transport processes in which the molecules of the invention participate are utilized by a wide variety of prokaryotic and eukaryotic cells; by comparing the sequences of the nucleic acid molecules of the present invention to those encoding similar enzymes from other organisms, the evolutionary relatedness of the organisms can be assessed. Similarly, such a comparison permits an assessment of which regions of the sequence are conserved and which are not, which may aid in determining those regions of the protein which are essential for the functioning of the enzyme. This type of determination is of value for protein engineering studies and may give an indication of what the protein can tolerate in terms of mutagenesis without losing function.

[Para 125] Manipulation of the LMP nucleic acid molecules of the invention may result in the production of LMPs having functional differences from the wild-type LMPs. These proteins may be improved in efficiency or activity, may be present in greater numbers in the cell than is usual, or may be decreased in efficiency or activity.

[Para 126] There are a number of mechanisms by which the alteration of a LMP of the invention may directly affect the accumulation and/or composition of seed storage compounds. In the case of plants expressing LMPs, increased transport can lead to altered accumulation of compounds and/or solute partitioning within the plant tissue and organs which ultimately could be used to affect the ac-

cumulation of one or more seed storage compounds during seed development. An example is provided by Mitsukawa et al. (1997, Proc. Natl. Acad. Sci. USA 94:7098-7102), where overexpression of an Arabidopsis high-affinity phosphate transporter gene in tobacco cultured cells enhanced cell growth under phosphate-limited conditions. Phosphate availability also affects significantly the production of sugars and metabolic intermediates (Hurry et al. 2000, Plant J. 24:383-396) and the lipid composition in leaves and roots (Härtel et al. 2000, Proc. Natl. Acad. Sci. USA 97:10649-10654). Likewise, the activity of the plant ACCase has been demonstrated to be regulated by phosphorylation (Savage & Ohlrogge 1999, Plant J. 18:521-527) and alterations in the activity of the kinases and phosphatases (LMPs) that act on the ACCase could lead to increased or decreased levels of seed lipid accumulation. Moreover, the presence of lipid kinase activities in chloroplast envelope membranes suggests that signal transduction pathways and/or membrane protein regulation occur in envelopes (see, e.g., Müller et al. 2000, J. Biol. Chem. 275:19475-19481 and literature cited therein). The ABI1 and ABI2 genes encode two protein serine/threonine phosphatases 2C, which are regulators in abscisic acid signaling pathway, and thereby in early and late seed development (e.g. Merlot et al. 2001, Plant J. 25:295-303). For more examples see also the section "Background of the Invention."

[Para 127] The present invention also provides antibodies that specifically bind to an LMP-polypeptide, or a portion thereof, as encoded by a nucleic acid disclosed herein or as described herein.

[Para 128] Antibodies can be made by many well-known methods (see, e.g. Harlow and Lane, "Antibodies; A Laboratory Manual." Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1988). Briefly, purified antigen can be injected into an animal in an amount and in intervals sufficient to elicit an immune response. Antibodies can either be purified directly, or spleen cells can be obtained from the animal. The cells can then fused with an immortal cell line and screened for antibody secretion. The antibodies can be used to screen nucleic acid clone libraries for cells secreting the antigen. Those positive clones can then be sequenced (see, for example, Kelly et al. 1992, Bio/Technology 10:163-167; Bebbington et al. 1992, Bio/Technology 10:169-175).

[Para 129] The phrase "selectively binds" with the polypeptide refers to a binding reaction, which is determinative of the presence of the protein in a heterogeneous population of proteins and other biologics. Thus, under designated immunoassay conditions, the specified antibodies bound to a particular protein do not bind in a significant amount to other proteins present in the sample. Selective binding to an antibody under such conditions may require an antibody that is selected for its specificity for a particular protein. A variety of immunoassay formats may be used to select antibodies that selectively bind with a particular protein. For example, solid-phase ELISA immuno-assays are routinely used to select antibodies selectively immunoreactive with a protein. See Harlow and Lane "Antibodies, A Laboratory Manual," Cold Spring Harbor Publications, New York (1988), for a description of immunoassay formats and conditions that could be used to determine selective binding.

[Para 130] In some instances, it is desirable to prepare monoclonal antibodies from various hosts. A description of techniques for preparing such monoclonal antibodies may be found in Stites et al., editors, "Basic and Clinical Immunology," (Lange Medical Publications, Los Altos, Calif., Fourth Edition) and references cited therein, and in Harlow and Lane ("Antibodies, A Laboratory Manual," Cold Spring Harbor Publications, New York, 1988).

[Para 131] Throughout this application, various publications are referenced. The disclosures of all of these publications and those references cited within those publications in their entireties are hereby incorporated by reference into this application in order to more fully describe the state of the art to which this invention pertains.

[Para 132] It will be apparent to those skilled in the art that various modifications and variations can be made in the present invention without departing from the scope or spirit of the invention. Other embodiments of the invention will be apparent to those skilled in the art from consideration of the specification and practice of the invention disclosed herein. It is intended that the specification and Examples be considered as exemplary only, with a true scope and spirit of the invention being indicated by the claims included herein.

EXAMPLES

[Para 133] Example 1: General Processes - a) General Cloning Processes.

[Para 134] Cloning processes such as, for example, restriction cleavages, agarose gel electrophoresis, purification of DNA fragments, transfer of nucleic acids to nitrocellulose and nylon membranes, linkage of DNA fragments, transformation of *Escherichia coli* and yeast cells, growth of bacteria and sequence analysis of recombinant DNA were carried out as described in Sambrook et al. (1989, Cold Spring Harbor Laboratory Press: ISBN 0-87969-309-6) or Kaiser, Michaelis and Mitchell (1994, "Methods in Yeast Genetics," Cold Spring Harbor Laboratory Press: ISBN 0-87969-451-3).

[Para 135] General Processes - b) Chemicals.

[Para 136] The chemicals used were obtained, if not mentioned otherwise in the text, in p.a. quality from the companies Fluka (Neu-Ulm), Merck (Darmstadt), Roth (Karlsruhe), Serva (Heidelberg) and Sigma (Deisenhofen). Solutions were prepared using purified, pyrogen-free water, designated as H₂O in the following text, from a Milli-Q water system water purification plant (Millipore, Eschborn). Restriction endonucleases, DNA-modifying enzymes and molecular biology kits were obtained from the companies AGS (Heidelberg), Amersham (Braunschweig), Biometra (Göttingen), Boehringer (Mannheim), Genomed (Bad Oeynhausen), New England Biolabs (Schwalbach/ Taunus), Novagen (Madison, Wisconsin, USA), Perkin-Elmer (Weiterstadt), Pharmacia (Freiburg), Qiagen (Hilden) and Stratagene (Amsterdam, Netherlands). They were used, if not mentioned otherwise, according to the manufacturer's instructions.

[Para 137] General Processes - c) Plant Material and Growth: *Arabidopsis* plants.

[Para 138] *Arabidopsis thaliana* cv Columbia were grown on plates with half-strength MS medium (Murashige & Skoog, 1962, *Physiol. Plant.* 15, 473-497), pH 6.2, 2% sucrose and 0.8% agar. Seeds were sterilized for 20 minutes in 20% bleach 0.5% triton X100 and rinsed 6 times with excess sterile water. Wild type *Arabidopsis* seeds were preincubated for three days in the dark at 4°C before placing them into an incubator (AR-75, Percival Scientific, Boone, IA) at a photon flux density of 60-80 $\mu\text{mol m}^{-2} \text{s}^{-1}$ and a light period of 16 hours (22°C), and a dark period of 8 hours (18°C). Plants were either grown as described above or on soil under standard conditions as described in Focks & Benning (1998, *Plant Physiol.* 118:91-101).

[Para 139] *Brassica napus*. *Brassica napus* varieties AC Excel and Cresor were used for this study to create cDNA libraries. Seed, seed pod, flower, leaf, stem and root tissues were collected from plants that were in some cases dark-, salt-, heat -and drought-treated. However, this study focused on the use of seed and seed pod tissues for cDNA libraries. Plants were tagged to harvest seeds collected 60 - 75 days after planting from two time points: 1-15 days and 15 -25 days after anthesis. Plants have been grown in Metromix (Scotts, Marysville, OH) at 71°F under a 14 hr photoperiod. Six seed and seed pod tissues of interest in this study were collected to create the following cDNA libraries: Immature seeds, mature seeds, immature seed pods, mature seed pods, night-harvested seed pods and Cresor variety (high erucic acid) seeds. Tissue samples were collected within specified time points for each developing tissue and multiple samples within a time frame pooled together for eventual extraction of total RNA. Samples from immature seeds were taken between 1-25 days after anthesis (daa), mature seeds between 25-50 daa, immature seed pods between 1-15 daa, mature seed pods between 15-50 daa, night-harvested seed pods between 1-50 daa and Cresor seeds 5-25 daa.

[Para 140] *Glycine max*. *Glycine max* cv. Resnick was used for this study to create cDNA libraries. Seed, seed pod, flower, leaf, stem, and root tissues were collected from plants that were in some cases dark-, salt-, heat- and drought-treated. In some cases plants have been nematode infected as well. However, this study focused on the use of seed and seed pod tissues for cDNA libraries. Plants were tagged to harvest seeds at the set days after anthesis: 5-15, 15-25, 25-35, and 33-50.

[Para 141] *Zea mays*. *Zea mays* hybrid B73 x Mo17 and B73 inbred (the female inbred parent of the hybrid B73 x Mo17) were used to generate cDNA libraries. Fruit or Seed (Fertilized ovules/young kernels at stage 1 and 9 d post pollination; kernels at milk stage [R3, early starch production], 23 d post pollination; kernels at early dough stage (R4), developing starch grains and well-formed embryo present, 30 d post pollination of field-grown plants; very young kernels at blister stage [R2, watery endosperm]; kernels at early dent stage (R5), endosperm becoming firm, 36 d post pollination; B73 inbreds, kernels at 9 and 19 d post pollination), flowers (tassel development: from 6 cm tassel (V10) up to and including anthesis, 44 to 70 dap; ear development: ear shoots from 2 cm (V13) up to and including silking (unpollinated), 51 to 70 dap), leaves/shoot/rosettes (mixed ages, all prior to seed-fill; includes leaves of a) 3-leaf plants(V3), b) 6-leaf plants (V6), and c) an older source leaf (3rd from the

ground), just before tassel emergence in the field), stem (located underground of 2 to 5-leaf plants; roots and most leaf tissue removed, 13 to 29 dap of field-grown plants; Stem tissue near the ear at tassel emergence and during seed-fill (milk stage), 56 to 84 dap, field-grown plants) and root tissues (from young to mid-age plants: from seedlings, 6-leaf plants, and 9-leaf plants; 12 to 35 dap) were collected from plants.

[Para 142] *Oryza sativa*. *Oryza sativa* ssp. Japonica cv. Nippon-barre was used for this study to create cDNA libraries. Seed, seed pod, flower, leaf, stem and root tissues were collected from plants that were in some cases dark-, salt-, heat- and drought-treated. This study focused on the use of seed embryo tissues for cDNA libraries. Embryo and endosperm were collected separately in case endosperm tissue might interfere with RNA extraction. Plants have been grown in the greenhouse on Wisconsin soil (has high organic matter) at 85°F under a 14-h photoperiod. Rice embryos were dissected out of the developing seeds.

[Para 143] *Triticum aestivum*. *Triticum aestivum* cv. Galeon was used for this study to create cDNA libraries. Seed, flower, fruits, leaf, stem and root tissues were collected from plants that were in some cases dark-, salt-, heat-, and drought-treated. Plants have been grown in the greenhouse in metromix under a 12-h photoperiod at 72°F during the day period and 65°F during the night period.

[Para 144] *Physcomitrella patens*. For this study, plants of the species *Physcomitrella patens* (Hedw.) B.S.G. from the collection of the genetic studies section of the University of Hamburg were used. They originate from the strain 16/14 collected by H.L.K. Whitehouse in Gransden Wood, Huntingdonshire (England), which was subcultured from a spore by Engel (1968, Am. J. Bot. 55: 438-446). Proliferation of the plants was carried out by means of spores and by means of regeneration of the gametophytes. The protonema developed from the haploid spore as a chloroplast-rich chloronema and chloroplast-low caulonema, on which buds formed after approximately 12 days. These grew to give gametophores bearing antheridia and archegonia. After fertilization, the diploid sporophyte with a short seta and the spore capsule resulted, in which the meiospores mature.

[Para 145] Culturing was carried out in a climatic chamber at an air temperature of 25°C and light intensity of 55 $\mu\text{mol m}^{-2} \text{s}^{-1}$ (white light; Philips TL 65W/25 fluorescent tube) and a light/dark change of 16/8 hours.

[Para 146] Example 2: Total DNA Isolation from Plants. The details for the isolation of total DNA relate to the working up of 1g fresh weight of plant material.

[Para 147] CTAB buffer: 2% (w/v) N-cethyl-N,N,N-trimethylammonium bromide (CTAB); 100 mM Tris HCl pH 8.0; 1.4 M NaCl; 20 mM EDTA. N-Laurylsarcosine buffer: 10% (w/v) N-laurylsarcosine; 100 mM Tris HCl pH 8.0; 20 mM EDTA.

[Para 148] The plant material was triturated under liquid nitrogen in a mortar to give a fine powder and transferred to 2ml Eppendorf vessels. The frozen plant material was then covered with a layer of 1 ml of decomposition buffer (1 ml CTAB buffer, 100 μl of N-laurylsarcosine buffer, 20 μl of β -

mercaptoethanol and 10µl of proteinase K solution, 10 mg/ml) and incubated at 60°C for one hour with continuous shaking. The homogenate obtained was distributed into two Eppendorf vessels (2 ml) and extracted twice by shaking with the same volume of chloroform/isoamyl alcohol (24:1). For phase separation, centrifugation was carried out at 8000g and RT for 15 min in each case. The DNA was then precipitated at -70°C for 30 min using ice-cold isopropanol. The precipitated DNA was sedimented at 4°C and 10,000g for 30 min and resuspended in 180µl of TE buffer (Sambrook et al. 1989, Cold Spring Harbor Laboratory Press: ISBN 0-87969-309-6). For further purification, the DNA was treated with NaCl (1.2 M final concentration) and precipitated again at -70°C for 30 min using twice the volume of absolute ethanol. After a washing step with 70% ethanol, the DNA was dried and subsequently taken up in 50 µl of H₂O + RNase (50 mg/ml final concentration). The DNA was dissolved overnight at 4°C and the RNase digestion was subsequently carried out at 37°C for 1 h. Storage of the DNA took place at 4°C.

[Para 149] Example 3: Isolation of Total RNA and poly-(A)+ RNA from Plants - *Arabidopsis thaliana*. For the investigation of transcripts, both total RNA and poly-(A)+ RNA were isolated.

[Para 150] RNA is isolated from siliques of *Arabidopsis* plants according to the following procedure:

RNA preparation from *Arabidopsis* seeds – “hot” extraction:

1. Buffers, enzymes and solution
 - 2M KCl
 - Proteinase K
 - Phenol (for RNA)
 - Chloroform:Isoamylalcohol
(Phenol:chloroform 1:1; pH adjusted for RNA)
 - 4 M LiCl, DEPC-treated
 - DEPC-treated water
 - 3M NaOAc, pH 5, DEPC-treated
 - Isopropanol
 - 70% ethanol (made up with DEPC-treated water)
 - Resuspension buffer: 0.5% SDS, 10 mM Tris pH 7.5, 1mM EDTA made up with DEPC-treated water as this solution can not be DEPC-treated
 - Extraction Buffer:
 - 0.2M Na Borate
 - 30 mM EDTA
 - 30 mM EGTA
 - 1% SDS (250µl of 10% SDS-solution for 2.5ml buffer)
 - 1% Deoxycholate (25mg for 2,5ml buffer)
 - 2% PVPP (insoluble - 50mg for 2.5ml buffer)

2% PVP 40K (50mg for 2.5ml buffer)

10 mM DTT

100mM β -Mercaptoethanol (fresh, handle under fume hood - use 35 μ l of 14.3M solution for 5ml buffer)

2. Extraction

[Para 151] Heat extraction buffer up to 80°C. Grind tissue in liquid nitrogen-cooled mortar, transfer tissue powder to 1.5ml tube. Tissue should be kept frozen until buffer is added so transfer the sample with pre-cooled spatula and keep the tube in liquid nitrogen all time. Add 350 μ l preheated extraction buffer (here for 100mg tissue, buffer volume can be as much as 500 μ l for bigger samples) to tube, vortex and heat tube to 80°C for ~1 min. Keep then on ice. Vortex sample, grind additionally with electric mortar.

3. Digestion

[Para 152] Add Proteinase K (0.15mg/100mg tissue), vortex and keep at 37°C for one hour.

[Para 153] First Purification. Add 27 μ l 2M KCl. Chill on ice for 10 min. Centrifuge at 12.000 rpm for 10 minutes at room temperature. Transfer supernatant to fresh, RNAase-free tube and do one phenol extraction, followed by a chloroform:isoamylalcohol extraction. Add 1 vol. isopropanol to supernatant and chill on ice for 10 min. Pellet RNA by centrifugation (7000 rpm for 10 min at RT). Resolve pellet in 1ml 4M LiCl by 10 to 15min vortexing. Pellet RNA by 5min centrifugation.

[Para 154] Second Purification. Resuspend pellet in 500 μ l Resuspension buffer. Add 500 μ l phenol and vortex. Add 250 μ l chloroform:isoamylalcohol and vortex. Spin for 5 min. and transfer supernatant to fresh tube. Repeat chloform:isoamylalcohol extraction until interface is clear. Transfer supernatant to fresh tube and add 1/10 vol 3M NaOAc, pH 5 and 600 μ l isopropanol. Keep at -20 for 20 min or longer. Pellet RNA by 10 min centrifugation. Wash pellet once with 70% ethanol. Remove all remaining alcohol before resolving pellet with 15 to 20 μ l DEPC-water. Determine quantity and quality by measuring the absorbance of a 1:200 dilution at 260 and 280nm. 40 μ g RNA/ml = 1OD₂₆₀

[Para 155] RNA from wild-type of Arabidopsis is isolated as described (Hosein, 2001, Plant Mol. Biol. Rep., 19, 65a-65e; Ruuska, S.A., Girke, T., Benning, C., & Ohlrogge, J.B., 2002, Plant Cell, 14, 1191-1206).

[Para 156] The mRNA is prepared from total RNA, using the Amersham Pharmacia Biotech mRNA purification kit, which utilizes oligo(dT)-cellulose columns.

[Para 157] Isolation of Poly-(A)+ RNA was isolated using Dyna BeadsR (Dyna, Oslo, Norway) following the instructions of the manufacturer's protocol. After determination of the concentration of the RNA or of the poly(A)+ RNA, the RNA was precipitated by addition of 1/10 volumes of 3 M sodium acetate pH 4.6 and 2 volumes of ethanol and stored at -70°C.

[Para 158] *Physcomitrella patens*, *Brassica napus*, *Glycine max*, *Zea mays*, *Oryza sativa* and *Triticum aestivum*. *Physcomitrella patens* was either modified in liquid culture using Knop medium according to Reski & Abel (1985, *Planta* 165: 354-358) or cultured on Knop solid medium using 1% oxoid agar (Unipath, Basingstoke, England). The protonemas used for RNA and DNA isolation were cultured in aerated liquid cultures. The protonemas were comminuted every 9 days and transferred to fresh culture medium.

[Para 159] *Brassica napus* and *Glycine max* seeds were separated from pods to create homogeneous materials for seed and seed pod cDNA libraries. Tissues were ground into fine powder under liquid N₂ using a mortar and pestle and transferred to a 50 ml tube. Tissue samples were stored at -80°C until extractions could be performed.

[Para 160] In the case of *Oryza sativa*, 5K - 10K embryos and endosperm were isolated through dissection. Tissues were placed in small tubes or petri dishes on ice during dissection. Containers were placed on dry ice, then stored at -80°C.

[Para 161] In the case of *Triticum aestivum*, seed germination samples of Galeon wheat seeds were planted at a depth of 2" in metromix in a 20" x 12" flat. The soil was soaked liberally with water and then watered twice daily. 3-4 days later when the coleoptiles were ~1 cm, the seedlings were washed with water and blotted. To create flower cDNA libraries an equal number of heads are collected at 30%, 60%, and 100% head emergence from the sheath on each of two days. There were no anthers showing yet. In order to generate seed tissue cDNA libraries grains were either watery ripe or in milk stage depending on the position of grains in the head; for later seed developmental stages Only the seed heads were harvested. For the root libraries, only roots were harvested. Plants had one main stem and three strong tillers. Plants were grown in pots, the medium was washed off, and the roots were saved for this sample. Plants were untreated.

[Para 162] In the case of *Zea mays*, tissues were ground into fine powder under liquid N₂ using a mortar and pestle and transferred to a 50ml tube. Tissue samples were stored at -80°C until extractions could be performed.

[Para 163] Total RNA was extracted from tissues using RNeasy Maxi kit (Qiagen) according to manufacture's protocol and mRNA was processed from total RNA using Oligotex mRNA Purification System kit (Qiagen), also according to manufacture's protocol. mRNA was sent to Hyseq Pharmaceuticals Incorporated (Sunnyville, CA) for further processing of mRNA from each tissue type into cDNA libraries and for use in their proprietary processes in which similar inserts in plasmids are clustered based on hybridization patterns.

[Para 164] Example 4: cDNA Library Construction. For cDNA library construction, first strand synthesis was achieved using Murine Leukemia Virus reverse transcriptase (Roche, Mannheim, Germany) and oligo-d(T)-primers, second strand synthesis by incubation with DNA polymerase I, Klenow enzyme and RNaseH digestion at 12°C (2 h), 16°C (1 h) and 22°C (1 h). The reaction was

stopped by incubation at 65°C (10 min) and subsequently transferred to ice. Double stranded DNA molecules were blunted by T4-DNA-polymerase (Roche, Mannheim) at 37°C (30 min). Nucleotides were removed by phenol/chloroform extraction and Sephadex G50 spin columns. EcoRI adapters (Pharmacia, Freiburg, Germany) were ligated to the cDNA ends by T4-DNA-ligase (Roche, 12°C, overnight) and phosphorylated by incubation with polynucleotide kinase (Roche, 37°C, 30 min). This mixture was subjected to separation on a low melting agarose gel. DNA molecules larger than 300 base pairs were eluted from the gel, phenol extracted, concentrated on Elutip-D-columns (Schleicher and Schuell, Dassel, Germany) and were ligated to vector arms and packed into lambda ZAPII phages or lambda ZAP-Express phages using the Gigapack Gold Kit (Stratagene, Amsterdam, Netherlands) using material and following the instructions of the manufacturer.

[Para 165] *Arabidopsis thaliana*, *Brassica napus*, *Glycine max*, *Zea mays*, *Oryza sativa* and *Triticum aestivum* cDNA libraries were generated at Hyseq Pharmaceuticals Incorporated (Sunnyville, CA). No amplification steps were used in the library production to retain expression information. Hyseq's genomic approach involves grouping the genes into clusters and then sequencing representative members from each cluster. cDNA libraries were generated from oligo dT column purified mRNA. Colonies from transformation of the cDNA library into E.coli were randomly picked and the cDNA insert were amplified by PCR and spotted on nylon membranes. A set of ³³P radiolabeled oligonucleotides were hybridized to the clones and the resulting hybridization pattern determined to which cluster a particular clone belonged. cDNA clones and their DNA sequences were obtained for use in overexpression in transgenic plants and in other molecular biology processes described herein.

[Para 166] For *Physcomitrella patens* cDNA library construction first strand synthesis was achieved using Murine Leukemia Virus reverse transcriptase (Roche, Mannheim, Germany) and oligo-d(T)-primers, second strand synthesis by incubation with DNA polymerase I, Klenow enzyme and RNaseH digestion at 12°C (2 h), 16°C (1 h) and 22°C (1 h). The reaction was stopped by incubation at 65°C (10 min) and subsequently transferred to ice. Double stranded DNA molecules were blunted by T4-DNA-polymerase (Roche, Mannheim) at 37°C (30 min). Nucleotides were removed by phenol/chloroform extraction and use of Sephadex G50 spin columns. EcoRI adapters (Pharmacia, Freiburg, Germany) were ligated to the cDNA ends by T4-DNA-ligase (Roche, 12°C, overnight) and phosphorylated by incubation with polynucleotide kinase (Roche, 37°C, 30 min). This mixture was subjected to separation on a low melting agarose gel. DNA molecules larger than 300 basepairs were eluted from the gel, phenol extracted, concentrated on Elutip-D-columns (Schleicher and Schuell, Dassel, Germany) and were ligated to vector arms and packed into lambda ZAPII phages or lambda ZAP-Express phages using the Gigapack Gold Kit (Stratagene, Amsterdam, Netherlands) using material and following the instructions of the manufacturer.

[Para 167] Example 5: Identification of LMP Genes of Interest that Are SUS-like. *Arabidopsis thaliana*, *Brassica napus*, *Glycine max*, *Zea mays*, *Oryza sativa*, *Triticum aestivum* and *Physcomitrella*

patens. This example illustrates how cDNA clones encoding SUS-like polypeptides of *Arabidopsis thaliana*, *Brassica napus*, *Glycine max*, *Zea mays*, *Oryza sativa*, *Triticum aestivum* and *Physcomitrella patens* were identified and isolated.

[Para 168] In order to identify *Arabidopsis thaliana* SUS-like genes in propriety databases, a similarity analysis using BLAST software (Basic Local Alignment Search Tool, version 2.2.6, Altschul et al. 1997, Nucleic Acid Res. 25: 3389-3402) was carry out. The default settings were used except for e-value cut-off (1e-10) and all protein searches were done using the BLOSUM62 matrix. The amino acid sequence of the *Arabidopsis* SUS polypeptide (AtSUS-2a) was used as a query to search and align DNA databases from *Arabidopsis thaliana*, *Brassica napus*, *Glycine max*, *Zea mays*, *Oryza sativa*, *Triticum aestivum* and *Physcomitrella patens* that were translated in all six reading frames, using the TBLASTN algorithm. Such similarity analysis of the BPS in-house databases resulted in the identification of numerous ESTs and cDNA contigs.

[Para 169] RNA expression profile data obtained from the Hyseq clustering process were used to determine organ-specificity. Clones showing a greater expression in seed libraries compared to the other tissue libraries were selected as LMP candidate genes. The *Arabidopsis thaliana*, *Brassica napus*, *Glycine max*, *Zea mays*, *Oryza sativa*, *Triticum aestivum* and *Physcomitrella patens* clones were selected for overexpression in *Arabidopsis*.

[Para 170] Example 6: Cloning of full-length cDNAs and orthologs of identified LMP genes. Clones corresponding to full-length sequences and partial cDNAs from *Arabidopsis thaliana*, *Brassica napus*, *Glycine max*, *Zea mays*, *Oryza sativa*, *Triticum aestivum* or *Physcomitrella patens* that were identified either in *Physcomitrella patens* EstMax (see also WO 02/074977 A2 for details) or in Hyseq databases are isolated by RACE PCR using the SMART RACE cDNA amplification kit from Clontech allowing both 5'- and 3' rapid amplification of cDNA ends (RACE). The isolation of cDNAs and the RACE PCR protocol used are based on the manufacturer's conditions. The RACE product fragments are extracted from agarose gels with a QIAquick Gel Extraction Kit (Qiagen) and ligated into the TOPO pCR 2.1 vector (Invitrogen) following manufacturer's instructions. Recombinant vectors are transformed into TOP10 cells (Invitrogen) using standard conditions (Sambrook et al. 1989). Transformed cells are grown overnight at 37°C on LB agar containing 50µg/ml kanamycin and spread with 40µl of a 40mg/ml stock solution of X-gal in dimethylformamide for blue-white selection. Single white colonies are selected and used to inoculate 3ml of liquid LB containing 50µg/ml kanamycin and grown overnight at 37°C. Plasmid DNA is extracted using the QIAprep Spin Miniprep Kit (Qiagen) following manufacturer's instructions. Subsequent analyses of clones and restriction mapping are performed according to standard molecular biology techniques (Sambrook et al. 1989).

[Para 171] Clones of *Arabidopsis thaliana*, *Brassica napus*, *Glycine max*, *Zea mays*, *Triticum aestivum*, *Oryza sativa* or *Physcomitrella patens* genes obtained from Hyseq were sequenced at using a ABI 377 slab gel sequencer and BigDye Terminator Ready Reaction kits (PE Biosystems, Foster City,

CA). Gene specific primers were designed using these sequences and genes were amplified from the plasmid supplied from Hyseq using touch-down PCR. In some cases, primers were designed to add an "AACA" Kozak-like sequence just upstream of the gene start codon and two bases downstream were, in some cases, changed to GC to facilitate increased gene expression levels (Chandrashekhara et al., 1997, Plant Molecular Biology 35:993-1001). PCR reaction cycles were: 94°C, 5 min; 9 cycles of 94°C, 1 min, 65°C, 1 min, 72°C, 4 min, and in which the anneal temperature was lowered by 1°C each cycle; 20 cycles of 94°C, 1 min, 55°C, 1 min, 72°C, 4 min; and the PCR cycle was ended with 72°C, 10 min. Amplified PCR products were gel purified from 1% agarose gels using GenElute -EtBr spin columns (Sigma) and after standard enzymatic digestion, were ligated into the plant binary vector pBPS-GB1 for transformation into *Arabidopsis thaliana* or other crops. The binary vector was amplified by overnight growth in *E. coli* DH5 in LB media and appropriate antibiotic and plasmid was prepared for downstream steps using Qiagen MiniPrep DNA preparation kit. The insert was verified throughout the various cloning steps by determining its size through restriction digest and inserts were sequenced in parallel to plant transformations to ensure the expected gene was used in *Arabidopsis thaliana* or other crop transformation.

[Para 172] Gene sequences can be used to identify homologous or heterologous genes (orthologs, the same LMP gene from another plant) from cDNA or genomic libraries. This can be done by designing PCR primers to conserved sequence regions identified by multiple sequence alignments. Orthologs are often identified by designing degenerate primers to full-length or partial sequences of genes of interest.

[Para 173] Homologous genes (e. g. full-length cDNA clones) can be isolated via nucleic acid hybridization using for example cDNA libraries: Depending on the abundance of the gene of interest, 100,000 up to 1,000,000 recombinant bacteriophages are plated and transferred to nylon membranes. After denaturation with alkali, DNA is immobilized on the membrane by, e.g., UV crosslinking. Hybridization is carried out at high stringency conditions. Aqueous solution hybridization and washing is performed at an ionic strength of 1 M NaCl and a temperature of 68°C. Hybridization probes are generated by, e.g., radioactive (³²P) nick transcription labeling (High Prime, Roche, Mannheim, Germany). Signals are detected by autoradiography.

[Para 174] Partially homologous or heterologous genes that are related but not identical can be identified in a procedure analogous to the above-described procedure using low stringency hybridization and washing conditions. For aqueous hybridization, the ionic strength is normally kept at 1 M NaCl while the temperature is progressively lowered from 68 to 42°C.

[Para 175] Isolation of gene sequences with homologies (or sequence identity/similarity) only in a distinct domain of (for example 10-20 amino acids) can be carried out by using synthetic radio labeled oligonucleotide probes. Radio labeled oligonucleotides are prepared by phosphorylation of the 5' end of two complementary oligonucleotides with T4 polynucleotide kinase. The complementary oligonu-

cleotides are annealed and ligated to form concatemers. The double stranded concatemers are then radiolabeled by for example nick transcription. Hybridization is normally performed at low stringency conditions using high oligonucleotide concentrations.

ligonucleotide hybridization solution:

x SSC

.01 M sodium phosphate

mM EDTA (pH 8)

.5% SDS

00 µg/ml denaturated salmon sperm DNA

.1% nonfat dried milk

[Para 176] During hybridization, temperature is lowered stepwise to 5-10°C below the estimated oligonucleotide T_m or down to room temperature followed by washing steps and autoradiography. Washing is performed with low stringency such as 3 washing steps using 4x SSC. Further details are described by Sambrook et al. (1989, "Molecular Cloning: A Laboratory Manual," Cold Spring Harbor Laboratory Press) or Ausubel et al. (1994, "Current Protocols in Molecular Biology," John Wiley & Sons).

[Para 177] Example 7: Identification of Genes of Interest by Screening Expression Libraries with Antibodies. c-DNA clones can be used to produce recombinant protein for example in E. coli (e. g. Qiagen QIAexpress pQE system). Recombinant proteins are then normally affinity purified via Ni-NTA affinity chromatography (Qiagen). Recombinant proteins can be used to produce specific antibodies for example by using standard techniques for rabbit immunization. Antibodies are affinity purified using a Ni-NTA column saturated with the recombinant antigen as described by Gu et al. (1994, BioTechniques 17:257-262). The antibody can then be used to screen expression cDNA libraries to identify homologous or heterologous genes via an immunological screening (Sambrook et al. 1989, "Molecular Cloning: A Laboratory Manual," Cold Spring Harbor Laboratory Press or Ausubel et al. 1994, "Current Protocols in Molecular Biology," John Wiley & Sons).

[Para 178] Example 8: Northern-Hybridization. For RNA hybridization, 20µg of total RNA or 1µg of poly-(A)+ RNA is separated by gel electrophoresis in 1.25% agarose gels using formaldehyde as described in Amasino (1986, Anal. Biochem. 152:304), transferred by capillary attraction using 10 x SSC to positively charged nylon membranes (Hybond N+, Amersham, Braunschweig), immobilized by UV light and pre-hybridized for 3 hours at 68°C using hybridization buffer (10% dextran sulfate w/v, 1 M NaCl, 1% SDS, 100µg/ml of herring sperm DNA). The labeling of the DNA probe with the Highprime DNA labeling kit (Roche, Mannheim, Germany) is carried out during the pre-hybridization using alpha-32P dCTP (Amersham, Braunschweig, Germany). Hybridization is carried out after addition of the labeled DNA probe in the same buffer at 68°C overnight. The washing steps

are carried out twice for 15 min using 2 x SSC and twice for 30 min using 1 x SSC, 1% SDS at 68°C. The exposure of the sealed filters is carried out at -70°C for a period of 1 day to 14 days.

[Para 179] Example 9: DNA Sequencing and Computational Functional Analysis. cDNA libraries can be used for DNA sequencing according to standard methods, in particular by the chain termination method using the ABI PRISM Big Dye Terminator Cycle Sequencing Ready Reaction Kit (Perkin-Elmer, Weiterstadt, Germany). Random sequencing can be carried out subsequent to preparative plasmid recovery from cDNA libraries via in vivo mass excision, retransformation, and subsequent plating of DH10B on agar plates (material and protocol details from Stratagene, Amsterdam, Netherlands). Plasmid DNA can be prepared from overnight grown *E. coli* cultures grown in Luria-Broth medium containing ampicillin (see Sambrook et al. (1989, Cold Spring Harbor Laboratory Press: ISBN 0-87969-309-6) on a Qiagen DNA preparation robot (Qiagen, Hilden) according to the manufacturer's protocols). Sequences can be processed and annotated using the software package EST-MAX commercially provided by Bio-Max (Munich, Germany). The program incorporates bioinformatics methods important for functional and structural characterization of protein sequences. For reference see <http://pedant.mips.biochem.mpg.de>.

[Para 180] The most important algorithms incorporated in EST-MAX are: FASTA: Very sensitive protein sequence database searches with estimates of statistical significance (Pearson W.R. 1990, Rapid and sensitive sequence comparison with FASTP and FASTA. *Methods Enzymol.* 183:63-98). BLAST: Very sensitive protein sequence database searches with estimates of statistical significance (Altschul S.F., Gish W., Miller W., Myers E.W., and Lipman D.J. Basic local alignment search tool, *J. Mol. Biol.* 215:403-410). PREDATOR: High-accuracy secondary structure prediction from single and multiple sequences. (Frishman & Argos 1997, 75% accuracy in protein secondary structure prediction. *Proteins* 27:329-335). CLUSTALW: Multiple sequence alignment (Thompson, J.D., Higgins, D.G. and Gibson, T.J. 1994, CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice, *Nucleic Acids Res.* 22:4673-4680). TMAP: Transmembrane region prediction from multiply aligned sequences (Persson B. & Argos P. 1994, Prediction of transmembrane segments in proteins utilizing multiple sequence alignments, *J. Mol. Biol.* 237:182-192). ALOM2: Transmembrane region prediction from single sequences (Klein P., Kanehisa M., and DeLisi C. 1984, Prediction of protein function from sequence properties: A discriminant analysis of a database. *Biochim. Biophys. Acta* 787:221-226. Version 2 by Dr. K. Nakai). PROSEARCH: Detection of PROSITE protein sequence patterns. Kolakowski L.F. Jr., Leunissen J.A.M. and Smith J.E. 1992, ProSearch: fast searching of protein sequences with regular expression patterns related to protein structure and function. *Biotechniques* 13:919-921). BLIMPS: Similarity searches against a database of ungapped blocks (Wallace & Henikoff 1992, PATMAT: A searching and extraction program for sequence, pattern and block queries and databases, *CABIOS* 8:249-254. Written by Bill Alford).

[Para 181] Example 10: Plasmids for Plant Transformation. For plant transformation binary vectors such as pBinAR can be used (Höfgen & Willmitzer 1990, Plant Sci. 66:221-230). Construction of the binary vectors can be performed by ligation of the cDNA in sense or antisense orientation into the T-DNA. 5' to the cDNA a plant promoter activates transcription of the cDNA. A polyadenylation sequence is located 3' to the cDNA. Tissue-specific expression can be achieved by using a tissue specific promoter. For example, seed-specific expression can be achieved by cloning the napin or LeB4 or USP promoter 5' to the cDNA. Also any other seed specific promoter element can be used. For constitutive expression within the whole plant the CaMV 35S promoter can be used. The expressed protein can be targeted to a cellular compartment using a signal peptide, for example for plastids, mitochondria or endoplasmic reticulum (Kermode 1996, Crit. Rev. Plant Sci. 15:285-423). The signal peptide is cloned 5' in frame to the cDNA to achieve subcellular localization of the fusion protein.

[Para 182] Further examples for plant binary vectors are the pBPS-GB1, pSUN2-GW or pBPS-GB047 vectors into which the LMP gene candidates are cloned. These binary vectors contain an antibiotic resistance gene driven under the control of the AtAct2-I promoter and a USP seed-specific promoter or a constitutive promoter in front of the candidate gene with the NOSpA terminator or the OCS terminator. Partial or full-length LMP cDNA are cloned into the multiple cloning site of the plant binary vector in sense or antisense orientation behind the USP seed-specific or PtxA promoters. The recombinant vector containing the gene of interest is transformed into Top10 cells (Invitrogen) using standard conditions. Transformed cells are selected for on LB agar containing 50µg/ml kanamycin grown overnight at 37°C. Plasmid DNA is extracted using the QIAprep Spin Miniprep Kit (Qiagen) following manufacturer's instructions. Analysis of subsequent clones and restriction mapping is performed according to standard molecular biology techniques (Sambrook et al. 1989, Molecular Cloning, A Laboratory Manual. 2nd Edition. Cold Spring Harbor Laboratory Press. Cold Spring Harbor, NY).

[Para 183] Example 11: Agrobacterium Mediated Plant Transformation. Agrobacterium mediated plant transformation with the LMP nucleic acids described herein can be performed using standard transformation and regeneration techniques (Gelvin, Stanton B. & Schilperoort R.A, Plant Molecular Biology Manual, 2nd ed. Kluwer Academic Publ., Dordrecht 1995 in Sect., Ringbuc Zentrale Signatur:BT11-P; Glick, Bernard R. and Thompson, John E. Methods in Plant Molecular Biology and Biotechnology, S. 360, CRC Press, Boca Raton 1993). For example, Agrobacterium mediated transformation can be performed using the GV3 (pMP90) (Koncz & Schell, 1986, Mol. Gen. Genet. 204:383-396) or LBA4404 (Clontech) Agrobacterium tumefaciens strain.

[Para 184] Arabidopsis thaliana can be grown and transformed according to standard conditions (Bechtold 1993, Acad. Sci. Paris. 316:1194-1199; Bent et al. 1994, Science 265:1856-1860). Additionally, rapeseed can be transformed with the LMR nucleic acids of the present invention via cotyledon or hypocotyl transformation (Moloney et al. 1989, Plant Cell Report 8:238-242; De Block et al.

1989, *Plant Physiol.* 91:694-701). Use of antibiotic for *Agrobacterium* and plant selection depends on the binary vector and the *Agrobacterium* strain used for transformation. Rapeseed selection is normally performed using a selectable plant marker. Additionally, *Agrobacterium* mediated gene transfer to flax can be performed using, for example, a technique described by Mlynarova et al. (1994, *Plant Cell Report* 13:282-285).

[Para 185] The *Arabidopsis* SUS or SUS-like genes were cloned into a binary vector and expressed either under a constitutive promoter like the superpromoter (Stanton B. Gelvin, USP# 5,428,147 and USP#5,217,903) or seed-specific promoters like USP (unknown seed protein) from *Vicia faba* (Baeumlein et al. 1991, *Mol. Gen. Genetics* 225:459-67), or the legumin B4 promoter (LeB4; Baeumlein et al. 1992, *Plant J.* 2:233-239) as well as promoters conferring seed-specific expression in monocot plants like maize, barley, wheat, rye, rice etc. were used.

[Para 186] Transformation of soybean can be performed using for example a technique described in EP 0424 047, U.S. Patent No. 5,322,783 (Pioneer Hi-Bred International) or in EP 0397 687, U.S. Patent No. 5,376,543, or U.S. Patent No. 5,169,770 (University Toledo), or by any of a number of other transformation procedures known in the art. Soybean seeds are surface sterilized with 70% ethanol for 4 minutes at room temperature with continuous shaking, followed by 20% (v/v) CLOROX supplemented with 0.05% (v/v) TWEEN for 20 minutes with continuous shaking. Then the seeds are rinsed 4 times with distilled water and placed on moistened sterile filter paper in a Petri dish at room temperature for 6 to 39 hours. The seed coats are peeled off, and cotyledons are detached from the embryo axis. The embryo axis is examined to make sure that the meristematic region is not damaged. The excised embryo axes are collected in a half-open sterile Petri dish and air-dried to a moisture content less than 20% (fresh weight) in a sealed Petri dish until further use.

[Para 187] The method of plant transformation is also applicable to *Brassica napus* and other crops. In particular, seeds of canola are surface sterilized with 70% ethanol for 4 minutes at room temperature with continuous shaking, followed by 20% (v/v) CLOROX supplemented with 0.05 % (v/v) TWEEN for 20 minutes, at room temperature with continuous shaking. Then, the seeds are rinsed 4 times with distilled water and placed on moistened sterile filter paper in a Petri dish at room temperature for 18 hours. The seed coats are removed and the seeds are air dried overnight in a half-open sterile Petri dish. During this period, the seeds lose approximately 85% of their water content. The seeds are then stored at room temperature in a sealed Petri dish until further use.

[Para 188] *Agrobacterium tumefaciens* culture is prepared from a single colony in LB solid medium plus appropriate antibiotics (e.g. 100mg/l streptomycin, 50mg/l kanamycin) followed by growth of the single colony in liquid LB medium to an optical density at 600 nm of 0.8. Then, the bacteria culture is pelleted at 7000 rpm for 7 minutes at room temperature, and re-suspended in MS (Murashige & Skoog 1962, *Physiol. Plant.* 15:473-497) medium supplemented with 100 mM acetosyringone. Bacteria cultures are incubated in this pre-induction medium for 2 hours at room temperature before use.

The axis of soybean zygotic seed embryos at approximately 44% moisture content are imbibed for 2 hours at room temperature with the pre-induced *Agrobacterium* suspension culture. (The imbibition of dry embryos with a culture of *Agrobacterium* is also applicable to maize embryo axes). The embryos are removed from the imbibition culture and are transferred to Petri dishes containing solid MS medium supplemented with 2% sucrose and incubated for 2 days, in the dark at room temperature. Alternatively, the embryos are placed on top of moistened (liquid MS medium) sterile filter paper in a Petri dish and incubated under the same conditions described above. After this period, the embryos are transferred to either solid or liquid MS medium supplemented with 500mg/l carbenicillin or 300mg/l cefotaxime to kill the *agrobacteria*. The liquid medium is used to moisten the sterile filter paper. The embryos are incubated during 4 weeks at 25°C, under 440 $\mu\text{mol m}^{-2}\text{s}^{-1}$ and 12 hours photoperiod. Once the seedlings have produced roots, they are transferred to sterile metromix soil. The medium of the in vitro plants is washed off before transferring the plants to soil. The plants are kept under a plastic cover for 1 week to favor the acclimatization process. Then the plants are transferred to a growth room where they are incubated at 25°C, under 440 $\mu\text{mol m}^{-2}\text{s}^{-1}$ light intensity and 12 h photoperiod for about 80 days.

[Para 189] Samples of the primary transgenic plants (T_0) are analyzed by PCR to confirm the presence of T-DNA. These results are confirmed by Southern hybridization wherein DNA is electrophoresed on a 1% agarose gel and transferred to a positively charged nylon membrane (Roche Diagnostics). The PCR DIG Probe Synthesis Kit (Roche Diagnostics) is used to prepare a digoxigenin-labeled probe by PCR as recommended by the manufacturer.

[Para 190] As an example for monocot transformation, the construction of ptxA promoter (PF55368-2 US, Song H. et al., 2004) in combination with maize Ubiquitin intron and SUS or SUS-like nucleic acid molecules is described. The PtxA-SUS ortholog gene construct in pUC is digested with PacI and XmaI. pBPSMM348 is digested with PacI and XmaI to isolate maize Ubiquitin intron (ZmUbi intron) followed by electrophoresis and the QIAEX II Gel Extraction Kit (cat# 20021). The ZmUbi intron is ligated into the PtxA-SUS or SUS-like nucleic acid molecule in pUC to generate pUC based PtxA-ZmUbi intron-SUS or SUS-like nucleic acid molecule construct followed by restriction enzyme digestion with AfeI and PmeI. PtxA-ZmUbi intron SUS or SUS-like gene cassette is cut out of a Seaplaque low melting temperature agarose gel (SeaPlaque® GTG® Agarose catalog No. 50110) after electrophoresis. A monocotyledonous base vector containing a selectable marker cassette (Monocot base vector) is digested with PmeI. The SUS or SUS-like nucleic acid molecule expression cassette containing ptxA promoter-ZmUbi intron is ligated into the Monocot base vector to generate PtxA-ZmUbi intron-SUS construct. Subsequently, the PtxA-ZmUbi intron-SUS or SUS-like nucleic acid molecule construct is transformed into a recombinant LBA4404 strain containing pSB1 (super vir plasmid) using electroporation following a general protocol in the art. *Agrobacterium*-mediated transformation

in maize is performed using immature embryo following a protocol described in US 5,591,616. An imidazolinoneherbicide selection is applied to obtain transgenic maize lines.

[Para 191] In general, a rice (or other monocot) SUS gene or SUS-like gene under a plant promoter like superpromoter could be transformed into corn, or another crop plant, to generate effects of monocot SUS genes in other monocots, or dicot SUS genes in other dicots, or monocot genes in dicots, or vice versa. The plasmids containing these SUS or SUS-like coding sequences, 5' of a promoter and 3' of a terminator would be constructed in a manner similar to those described for construction of other plasmids herein.

[Para 192] Example 12: In vivo Mutagenesis. In vivo mutagenesis of microorganisms can be performed by incorporation and passage of the plasmid (or other vector) DNA through *E. coli* or other microorganisms (e.g. *Bacillus* spp. or yeasts such as *Saccharomyces cerevisiae*) that are impaired in their capabilities to maintain the integrity of their genetic information. Typical mutator strains have mutations in the genes for the DNA repair system (e.g., *mutHLS*, *mutD*, *mutT*, etc.; for reference, see Rupp W.D. 1996, DNA repair mechanisms, in: *Escherichia coli* and *Salmonella*, p. 2277-2294, ASM: Washington). Such strains are well known to those skilled in the art. The use of such strains is illustrated, for example, in Greener and Callahan 1994, *Strategies* 7:32-34. Transfer of mutated DNA molecules into plants is preferably done after selection and testing in microorganisms. Transgenic plants are generated according to various examples within the exemplification of this document.

[Para 193] Example 13: Assessment of the mRNA Expression and Activity of a Recombinant Gene Product in the Transformed Organism. The activity of a recombinant gene product in the transformed host organism can be measured on the transcriptional or/and on the translational level. A useful method to ascertain the level of transcription of the gene (an indicator of the amount of mRNA available for translation to the gene product) is to perform a Northern blot (for reference see, for example, Ausubel et al. 1988, *Current Protocols in Molecular Biology*, Wiley: New York), in which a primer designed to bind to the gene of interest is labeled with a detectable tag (usually radioactive or chemiluminescent), such that when the total RNA of a culture of the organism is extracted, run on gel, transferred to a stable matrix and incubated with this probe, the binding and quantity of binding of the probe indicates the presence and also the quantity of mRNA for this gene. This information at least partially demonstrates the degree of transcription of the transformed gene. Total cellular RNA can be prepared from plant cells, tissues or organs by several methods, all well-known in the art, such as that described in Bormann et al. (1992, *Mol. Microbiol.* 6:317-326).

[Para 194] To assess the presence or relative quantity of protein translated from this mRNA, standard techniques, such as a Western blot, may be employed (see, for example, Ausubel et al. 1988, *Current Protocols in Molecular Biology*, Wiley: New York). In this process, total cellular proteins are extracted, separated by gel electrophoresis, transferred to a matrix such as nitrocellulose, and incubated with a probe, such as an antibody, which specifically binds to the desired protein. This probe is gen-

erally tagged with a chemiluminescent or colorimetric label, which may be readily detected. The presence and quantity of label observed indicates the presence and quantity of the desired mutant protein present in the cell.

[Para 195] The activity of LMPs that bind to DNA can be measured by several well-established methods, such as DNA band-shift assays (also called gel retardation assays). The effect of such LMP on the expression of other molecules can be measured using reporter gene assays (such as that described in Kolmar H. et al. 1995, EMBO J. 14:3895-3904 and references cited therein). Reporter gene test systems are well known and established for applications in both prokaryotic and eukaryotic cells, using enzymes such as beta-galactosidase, green fluorescent protein, and several others.

[Para 196] The determination of activity of lipid metabolism membrane-transport proteins can be performed according to techniques such as those described in Gennis R.B. (1989 Pores, Channels, and Transporters in Biomembranes, Molecular Structure and Function, Springer: Heidelberg, pp. 85-137, 199-234 and 270-322).

[Para 197] Example 14: In vitro Analysis of the Function of Arabidopsis thaliana, Brassica napus, Glycine max, Zea mays, Oryza sativa, Triticum aestivum or Physcomitrella patens SUS and SUS-like Genes in Transgenic Plants. The determination of activities and kinetic parameters of enzymes is well established in the art. Experiments to determine the activity of any given altered enzyme must be tailored to the specific activity of the wild-type enzyme, which is well within the ability of one skilled in the art. Overviews about enzymes in general, as well as specific details concerning structure, kinetics, principles, methods, applications and examples for the determination of many enzyme activities may be found, for example, in the following references: Dixon, M. & Webb, E.C. 1979, Enzymes. Longmans: London; Fersht, (1985) Enzyme Structure and Mechanism. Freeman: New York; Walsh (1979) Enzymatic Reaction Mechanisms. Freeman: San Francisco; Price, N.C., Stevens, L. (1982) Fundamentals of Enzymology. Oxford Univ. Press: Oxford; Boyer, P.D., ed. (1983) The Enzymes, 3rd ed. Academic Press: New York; Bisswanger, H., (1994) Enzymkinetik, 2nd ed. VCH: Weinheim (ISBN 3527300325); Bergmeyer, H.U., Bergmeyer, J., Graßl, M., eds. (1983-1986) Methods of Enzymatic Analysis, 3rd ed., vol. I-XII, Verlag Chemie: Weinheim; and Ullmann's Encyclopedia of Industrial Chemistry (1987) vol. A9, Enzymes. VCH: Weinheim, p. 352-363. For a more specific example for purification, preparation of membrane fractions, kinetic analyses and assay of sucrose synthase activities see, for example, Barratt, D.H.P. et al. (2001, Plant Physiol. 127: 655-664).

[Para 198] Example 15: Analysis of the Impact of Recombinant Proteins on the Production of a Desired Seed Storage Compound. Seeds from transformed Arabidopsis thaliana plants were analyzed by gas chromatography (GC) for total oil content and fatty acid profile. As an example for seed oil changes, GC analysis reveals that Arabidopsis plants transformed with pBPS-GB01 containing a USP promoter driving the Arabidopsis SUS gene (AtSUS-2b; named pk118 in the figure; SEQ ID:5 of the WIPO ST. 25 sequence listing) show an increase in total seed oil content by 10-18% compared with

the Columbia-2 control (the genetic background of the transformed lines) in both homozygous T3 and homozygous T4 seed generation (Figure 11). C24 represents a non-transformed high fatty acid seed control Columbia-24. Each circle represents the data obtained with 5 mg bulked seeds of one individual plant. The p values (as obtained by simple t-test) reveal significant increases in at least three independent transgenic events. The results suggest that SUS overexpression with a seed specific promoter allows the manipulation of total seed oil content. Similar results have been obtained when using a constitutive promoter as well (data not shown).

Arabidopsis was used to investigate the influence of SUS-like genes on seed storage compound accumulation. Total fatty acid content of seeds of control and transgenic plants were measured with bulked seeds (usually 5 mg seed weight) of a single plant. Three different types of controls have been used: Col-2 (Columbia-2, the Arabidopsis ecotypes LMP gene of interest have been transformed in), C-24 (an Arabidopsis ecotype found to accumulate high amounts of total fatty acids in seeds) and BPS empty (without LMP gene of interest) binary vector construct. The controls indicated in the tables below have been grown side by side with the transgenic lines. Differences in the total values of the controls are explained either by differences in the growth conditions, which were found to be very sensitive to small variations in the plant cultivation. Because of the seed bulking all values obtained with T2 seeds are the result of a mixture of homozygous (for the gene of interest) and heterozygous events, implying that these data underestimate the LMP gene effect. The binary vector pBPS-GB01 containing a USP promoter driving the gene of interest has been used. For PpSUS-2 the binary vector pSun2 containing a USP promoter was used as empty vector control and for delivery of the PpSUS-2 into the Arabidopsis ecotype Col-0.

Table 1. Determination of the T2 seed total fatty acid content of transgenic lines of BnSUS-2277 (containing SEQ ID NO:9 as used in the WIPO Standard ST. 25 Sequence listing). Shown are the means ($\pm$ standard deviation). Average mean values are shown $\pm$ standard deviation, number of individual measurements per plant line: 8-10; Col-2 is the Arabidopsis ecotype the LMP gene has been transformed in, C-24 is a high-oil Arabidopsis ecotype used as another control. Transgenic seeds of BnSUS-2277 show a significant increase relative to the empty vector control seeds ($p < 0.05$ as obtained by simple t-test)

Genotype	g total fatty acids/g seed weight
C-24 wild-type seeds	0.328 ± 0.024
Empty vector control seeds	0.312 ± 0.034
BnSUS-2277 transgenic seeds	0.345 ± 0.010

Table 2. Determination of the T2 seed total fatty acid content of transgenic lines of GmSUS-4819 (containing SEQ ID NO:17 of the WIPO Standard ST. 25 sequence listing). Shown are the means ($\pm$ standard deviation). Average mean values are shown $\pm$ standard deviation, number of individual measurements per plant line: 8-20; Col-2 is the Arabidopsis ecotype the LMP gene has been trans-

formed in, C-24 is a high-oil Arabidopsis ecotype used as another control. Transgenic seeds of GmSUS-4819 show a significant increase relative to the empty vector control seeds ($p < 0.05$ as obtained by simple t-test)

<u>Genotype</u>	<u>g total fatty acids/g seed weight</u>
C-24 wild-type seeds	0.312 ± 0.017
Empty vector control seeds	0.288 ± 0.022
GmSUS-4819 transgenic seeds	0.312 ± 0.025

Table 3. Determination of the T2 seed total fatty acid content of transgenic lines of ZmSUS-7691 (containing SEQ ID NO:21 of the WIPO Standard ST. 25 sequence listing). Shown are the means ($\pm$ standard deviation). Average mean values are shown $\pm$ standard deviation, number of individual measurements per plant line: 8-20; Col-2 is the Arabidopsis ecotype the LMP gene has been transformed in, C-24 is a high-oil Arabidopsis ecotype used as another control. Transgenic seeds of ZmSUS-7691 show a significant increase relative to the empty vector control seeds ($p < 0.05$ as obtained by simple t-test)

<u>Genotype</u>	<u>g total fatty acids/g seed weight</u>
C-24 wild-type seeds	0.331 ± 0.026
Empty vector control seeds	0.297 ± 0.041
ZmSUS-7691 transgenic seeds	0.312 ± 0.031

Table 4. Determination of the T2 seed total fatty acid content of transgenic lines of PpSUS-2 (containing SEQ ID NO:37). Shown are the means ($\pm$ standard deviation). Average mean values are shown $\pm$ standard deviation, number of individual measurements per plant line: 10; Col-0 is the Arabidopsis ecotype the LMP gene and the empty vector have been transformed in. Transgenic seeds of PpSUS-2 show a significant increase relative to the empty vector control seeds ($p < 0.05$ as obtained by simple t-test)

<u>Genotype</u>	<u>g total fatty acids/g seed weight</u>
Empty vector control seeds	0.304 ± 0.031
PpSUS-2 transgenic seeds	0.329 ± 0.011

[Para 199] The effect of the genetic modification in plants on a desired seed storage compound (such as a sugar, lipid or fatty acid) can be assessed by growing the modified plant under suitable conditions and analyzing the seeds or any other plant organ for increased production of the desired product (i.e., a lipid or a fatty acid). Such analysis techniques are well known to one skilled in the art, and include spectroscopy, thin layer chromatography, staining methods of various kinds, enzymatic and microbiological methods, and analytical chromatography such as high performance liquid chromatography (see, for example, Ullman 1985, Encyclopedia of Industrial Chemistry, vol. A2, pp. 89-90 and 443-613, VCH: Weinheim; Fallon, A. et al. 1987, Applications of HPLC in Biochemistry in: Laboratory Techniques in Biochemistry and Molecular Biology, vol. 17; Rehm et al., 1993 Product recovery and

purification, *Biotechnology*, vol. 3, Chapter III, pp. 469-714, VCH: Weinheim; Belter, P.A. et al., 1988 *Bioseparations: downstream processing for biotechnology*, John Wiley & Sons; Kennedy J.F. & Cabral J.M.S. 1992, *Recovery processes for biological materials*, John Wiley and Sons; Shaeiwitz J.A. & Henry J.D. 1988, *Biochemical separations in: Ulmann's Encyclopedia of Industrial Chemistry*, Separation and purification techniques in biotechnology, vol. B3, Chapter 11, pp. 1-27, VCH: Weinheim; and Dechow F.J. 1989).

[Para 200] Besides the above-mentioned methods, plant lipids are extracted from plant material as described by Cahoon et al. (1999, *Proc. Natl. Acad. Sci. USA* 96, 22:12935-12940) and Browse et al. (1986, *Anal. Biochemistry* 442:141-145). Qualitative and quantitative lipid or fatty acid analysis is described in Christie, William W., *Advances in Lipid Methodology*. Ayr/Scotland:Oily Press - (Oily Press Lipid Library; Christie, William W., *Gas Chromatography and Lipids. A Practical Guide* - Ayr, Scotland:Oily Press, 1989 Repr. 1992. - IX,307 S. - (Oily Press Lipid Library; and "Progress in Lipid Research," Oxford :Pergamon Press, 1 (1952) – 16 (1977) *Progress in the Chemistry of Fats and Other Lipids* CODEN.

[Para 201] Unequivocal proof of the presence of fatty acid products can be obtained by the analysis of transgenic plants following standard analytical procedures: GC, GC-MS or TLC as variously described by Christie and references therein (1997 in: *Advances on Lipid Methodology* 4th ed.: Christie, Oily Press, Dundee, pp. 119-169; 1998). Detailed methods are described for leaves by Lemieux et al. (1990, *Theor. Appl. Genet.* 80:234-240) and for seeds by Focks & Benning (1998, *Plant Physiol.* 118:91-101).

[Para 202] Positional analysis of the fatty acid composition at the sn-1, sn-2 or sn-3 positions of the glycerol backbone is determined by lipase digestion (see, e.g., Siebertz & Heinz 1977, *Z. Naturforsch.* 32c:193-205, and Christie 1987, *Lipid Analysis* 2nd Edition, Pergamon Press, Exeter, ISBN 0-08-023791-6).

[Para 203] Total seed oil levels can be measured by any appropriate method. Quantitation of seed oil contents is often performed with conventional methods, such as near infrared analysis (NIR) or nuclear magnetic resonance imaging (NMR). NIR spectroscopy has become a standard method for screening seed samples whenever the samples of interest have been amenable to this technique. Samples studied include canola, soybean, maize, wheat, rice, and others. NIR analysis of single seeds can be used (see e.g. Velasco et al., *Estimation of seed weight, oil content and fatty acid composition in intact single seeds of rapeseed (Brassica napus L.) by near-infrared reflectance spectroscopy*, *Euphytica*, Vol. 106, 1999, pp. 79-85). NMR has also been used to analyze oil content in seeds (see e.g. Robertson & Morrison, "Analysis of oil content of sunflower seed by wide-line NMR," *Journal of the American Oil Chemists Society*, 1979, Vol. 56, 1979, pp. 961-964, which is herein incorporated by reference in its entirety).

[Para 204] A typical way to gather information regarding the influence of increased or decreased protein activities on lipid and sugar biosynthetic pathways is for example via analyzing the carbon fluxes by labeling studies with leaves or seeds using ^{14}C -acetate or ^{14}C -pyruvate (see, e.g. Focks & Benning 1998, *Plant Physiol.* 118:91-101; Eccleston & Ohlrogge 1998, *Plant Cell* 10:613-621). The distribution of carbon-14 into lipids and aqueous soluble components can be determined by liquid scintillation counting after the respective separation (for example on TLC plates) including standards like ^{14}C -sucrose and ^{14}C -malate (Eccleston & Ohlrogge 1998, *Plant Cell* 10:613-621).

[Para 205] Material to be analyzed can be disintegrated via sonification, glass milling, liquid nitrogen, and grinding or via other applicable methods. The material has to be centrifuged after disintegration. The sediment is re-suspended in distilled water, heated for 10 minutes at 100°C , cooled on ice and centrifuged again followed by extraction in 0.5 M sulfuric acid in methanol containing 2% dimethoxypropane for 1 hour at 90°C leading to hydrolyzed oil and lipid compounds resulting in trans-methylated lipids. These fatty acid methyl esters are extracted in petrolether and finally subjected to GC analysis using a capillary column (Chrompack, WCOT Fused Silica, CP-Wax-52 CB, 25 m, 0.32mm) at a temperature gradient between 170°C and 240°C for 20 minutes, and 5 minutes at 240°C . The identity of resulting fatty acid methylesters is defined by the use of standards available from commercial sources (i.e., Sigma).

[Para 206] In case of fatty acids where standards are not available, molecule identity is shown via derivatization and subsequent GC-MS analysis. For example, the localization of triple bond fatty acids is shown via GC-MS after derivatization via 4,4-Dimethoxy-oxazolin-Derivaten (Christie, Oily Press, Dundee, 1998).

[Para 207] A common standard method for analyzing sugars, especially starch, is published by Stitt M., Lilley R.Mc.C., Gerhardt R. and Heldt M.W. (1989, "Determination of metabolite levels in specific cells and subcellular compartments of plant leaves," *Methods Enzymol.* 174:518-552; for other methods see also Härtel et al. 1998, *Plant Physiol. Biochem.* 36:407-417 and Focks & Benning 1998, *Plant Physiol.* 118:91-101).

[Para 208] For the extraction of soluble sugars and starch, 50 seeds are homogenized in 500 μl of 80% (v/v) ethanol in a 1.5-ml polypropylene test tube and incubated at 70°C for 90 minutes. Following centrifugation at 16,000g for 5 minutes, the supernatant is transferred to a new test tube. The pellet is extracted twice with 500 μl of 80% ethanol. The solvent of the combined supernatants is evaporated at room temperature under a vacuum. The residue is dissolved in 50 μl of water, representing the soluble carbohydrate fraction. The pellet left from the ethanol extraction, which contains the insoluble carbohydrates including starch, is homogenized in 200 μl of 0.2 N KOH, and the suspension is incubated at 95°C for 1 hour to dissolve the starch. Following the addition of 35 μl of 1 N acetic acid and centrifugation for 5 minutes at 16,000g, the supernatant is used for starch quantification.

[Para 209] To quantify soluble sugars, 10 μ l of the sugar extract is added to 990 μ l of reaction buffer containing 100mM imidazole, pH 6.9, 5mM MgCl₂, 2mM NADP, 1mM ATP, and 2 units 2 ml⁻¹ of Glucose-6-P-dehydrogenase. For enzymatic determination of glucose, fructose, and sucrose, 4.5 units of hexokinase, 1 unit of phosphoglucosomerase, and 2 μ l of a saturated fructosidase solution are added in succession. The production of NADPH is photometrically monitored at a wavelength of 340nm. Similarly, starch is assayed in 30 μ l of the insoluble carbohydrate fraction with a kit from Boehringer Mannheim.

[Para 210] An example for analyzing the protein content in leaves and seeds can be found by Bradford M.M. (1976, "A rapid and sensitive method for the quantification of microgram quantities of protein using the principle of protein dye binding," *Anal. Biochem.* 72:248-254). For quantification of total seed protein, 15-20 seeds are homogenized in 250 μ l of acetone in a 1.5-ml polypropylene test tube. Following centrifugation at 16,000g, the supernatant is discarded and the vacuum-dried pellet is resuspended in 250 μ l of extraction buffer containing 50mM Tris-HCl, pH 8.0, 250mM NaCl, 1 mM EDTA, and 1% (w/v) SDS. Following incubation for 2 hours at 25°C, the homogenate is centrifuged at 16,000g for 5 minutes and 200 μ l of the supernatant will be used for protein measurements. In the assay, γ -globulin is used for calibration. For protein measurements, Lowry DC protein assay (Bio-Rad) or Bradford-assay (Bio-Rad) is used.

[Para 211] Enzymatic assays of hexokinase and fructokinase are performed spectrophotometrically according to Renz et al. (1993, *Planta* 190:156-165), of phosphoglucosomerase, ATP-dependent 6-phosphofructokinase, pyrophosphate-dependent 6-phosphofructokinase, Fructose-1,6-bisphosphate aldolase, triose phosphate isomerase, glyceral-3-P dehydrogenase, phosphoglycerate kinase, phosphoglycerate mutase, enolase and pyruvate kinase are performed according to Burrell et al. (1994, *Planta* 194:95-101) and of UDP-Glucose-pyrophosphorylase according to Zrenner et al. (1995, *Plant J.* 7:97-107).

[Para 212] Intermediates of the carbohydrate metabolism, like Glucose-1-phosphate, Glucose-6-phosphate, Fructose-6-phosphate, Phosphoenolpyruvate, Pyruvate, and ATP are measured as described in Härtel et al. (1998, *Plant Physiol. Biochem.* 36:407-417) and metabolites are measured as described in Jelitto et al. (1992, *Planta* 188:238-244).

[Para 213] In addition to the measurement of the final seed storage compound (i.e., lipid, starch or storage protein) it is also possible to analyze other components of the metabolic pathways utilized for the production of a desired seed storage compound, such as intermediates and side-products, to determine the overall efficiency of production of the compound (Fiehn et al. 2000, *Nature Biotech.* 18:1447-1161).

[Para 214] For example, yeast expression vectors comprising the nucleic acids disclosed herein, or fragments thereof, can be constructed and transformed into *Saccharomyces cerevisiae* using standard

protocols. The resulting transgenic cells can then be assayed for alterations in sugar, oil, lipid, or fatty acid contents.

[Para 215] Similarly, plant expression vectors comprising the nucleic acids disclosed herein, or fragments thereof, can be constructed and transformed into an appropriate plant cell such as *Arabidopsis*, soybean, rapeseed, rice, maize, wheat, *Medicago truncatula*, etc., using standard protocols. The resulting transgenic cells and/or plants derived there from can then be assayed for alterations in sugar, oil, lipid, or fatty acid contents.

[Para 216] Additionally, the sequences disclosed herein, or fragments thereof, can be used to generate knockout mutations in the genomes of various organisms, such as bacteria, mammalian cells, yeast cells, and plant cells (Girke et al. 1998, *Plant J.* 15:39-48). The resultant knockout cells can then be evaluated for their composition and content in seed storage compounds, and the effect on the phenotype and/or genotype of the mutation. For other methods of gene inactivation include US 6004804 "Non-Chimeric Mutational Vectors" and Puttaraju et al. (1999, "Spliceosome-mediated RNA trans-splicing as a tool for gene therapy," *Nature Biotech.* 17:246-252). Example 16: Purification of the Desired Product from Transformed Organisms. An LMP can be recovered from plant material by various methods well known in the art. Organs of plants can be separated mechanically from other tissue or organs prior to isolation of the seed storage compound from the plant organ. Following homogenization of the tissue, cellular debris is removed by centrifugation and the supernatant fraction containing the soluble proteins is retained for further purification of the desired compound. If the product is secreted from cells grown in culture, then the cells are removed from the culture by low-speed centrifugation and the supernate fraction is retained for further purification.

[Para 218] The supernatant fraction from either purification method is subjected to chromatography with a suitable resin, in which the desired molecule is either retained on a chromatography resin while many of the impurities in the sample are not, or where the impurities are retained by the resin, while the sample is not. Such chromatography steps may be repeated as necessary, using the same or different chromatography resins. One skilled in the art would be well versed in the selection of appropriate chromatography resins and in their most efficacious application for a particular molecule to be purified. The purified product may be concentrated by filtration or ultrafiltration, and stored at a temperature at which the stability of the product is maximized.

[Para 219] There is a wide array of purification methods known to the art, and the preceding method of purification is not meant to be limiting. Such purification techniques are described, for example, in Bailey J.E. & Ollis D.F. 1986, *Biochemical Engineering Fundamentals*, McGraw-Hill:New York).

[Para 220] The identity and purity of the isolated compounds may be assessed by techniques standard in the art. These include high-performance liquid chromatography (HPLC), spectroscopic methods, staining methods, thin layer chromatography, analytical chromatography such as high performance liquid chromatography, NIRS, enzymatic assay, or microbiologically. Such analysis methods

are reviewed in: Patek et al. (1994, Appl. Environ. Microbiol. 60:133-140), Malakhova et al. (1996, Biotekhnologiya 11:27-32) and Schmidt et al. (1998, Bioprocess Engineer 19:67-70), Ulmann's Encyclopedia of Industrial Chemistry (1996, Vol. A27, VCH: Weinheim, p. 89-90, p. 521-540, p. 540-547, p. 559-566, 575-581 and p. 581-587) and Michal G. (1999, Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology, John Wiley and Sons; Fallon, A. et al. 1987, Applications of HPLC in Biochemistry in: Laboratory Techniques in Biochemistry and Molecular Biology, vol. 17).

[Para 221] Example 17: Screening for Increased Stress Tolerance and Plant Growth. The transgenic plants are screened for their improved stress tolerance demonstrating that transgene expression confers stress tolerance.

[Para 222] The transgenic plants are further screened for their growth rate demonstrating that transgene expression confers increased growth rates and/or increased seed yield.

[Para 223] Increased seed size might be reflected in an increased seed weight of gene overexpressors. Increased seed size leads to greater yield in many economically important crop plants. Therefore, increased seed size is one goal of genetically engineering and selection using LMPs as described in this application.

[Para 224] For in vitro root analysis square plates measuring 12 cm x 12 cm can be used. For each plate, 52 ml of MS media (0.5X MS salts, 0.5% sucrose, 0.5g/L MES buffer, 1% Phytagar) without selection will be used. Plates will be allowed to dry in the sterile hood for one hour to reduce future condensation.

[Para 225] Seed aliquots will be sterilized in glass vials with ethanol for 5 minutes, the ethanol was removed, and the seeds were allowed to dry in the sterile hood for one hour.

[Para 226] Seeds will be spotted in the plates using the Vacuseed Device (Lehle). After the seeds were spotted on the plates, the plates will be wrapped with Ventwrap and placed vertically in racks in the dark at 4°C for 4 days to stratify the seeds. The plates are transferred to a C5 Percival Growth Chamber and placed vertically. The growth chamber conditions will be 23°C day/21°C night, and 16-hour day/8-hour night.

[Para 227] For data collection a high resolution flat-bed scanner is used. Analysis of the roots is done using the WinRhizo software package.

[Para 228] For soil root analysis seeds may be imbibed at 4°C for 2 days in water and planted directly in soil with no selection. Deepots (Hummert D40) will be used with a saturated peat pellet (Jiffy 727) at the base and filled with water saturated Metromix. After planting, pots will be covered with plastic wrap to prevent drying. Plants may be grown using only water present at media preparation, as the water in the soil in these large pots is sufficient for 3 weeks of growth, and encourages rapid root growth. The plastic wrapping of the pots will be removed after 12 days and morphological data documented. At day 17 the aerial parts of the plant will be harvested, dried (65°C for 2 days) and dry weight measured. To examine the roots the peat pellet will be pushed towards the top of the pot

to remove the soil and roots as a unit. The soil will then be separated from the roots in a tray and the maximum root length will be measured. Root length of all plants for all transgenic lines will be averaged and compared against the average of the wild type plants.

Table 5.

Plant Lipid Classes

Neutral Lipids	Triacylglycerol (TAG)
	Diacylglycerol (DAG)
	Monoacylglycerol (MAG)
Polar Lipids	Monogalactosyldiacylglycerol (MGDG)
	Digalactosyldiacylglycerol (DGDG)
	Phosphatidylglycerol (PG)
	Phosphatidylcholine (PC)
	Phosphatidylethanolamine (PE)
	Phosphatidylinositol (PI)
	Phosphatidylserine (PS)
	Sulfoquinovosyldiacylglycerol

Table 6.

Common Plant Fatty Acids

16:0	Palmitic acid
16:1	Palmitoleic acid
16:3	Palmitolenic acid
18:0	Stearic acid
18:1	Oleic acid
18:2	Linoleic acid
18:3	Linolenic acid
Γ -18:3	Gamma-linolenic acid*
20:0	Arachidic acid
20:1	Eicosenoic acid
22:6	Docosahexanoic acid (DHA) *
20:2	Eicosadienoic acid
20:4	Arachidonic acid (AA) *
20:5	Eicosapentaenoic acid (EPA) *
22:1	Erucic acid

[Para 229] *These fatty acids do not normally occur in plant seed oils, but their production in transgenic plant seed oil is of importance in plant biotechnology.

Table 7.

A table of the putative functions of the SUS-like LMPs (the full length nucleic acid sequences can be found in Appendix A using the sequence codes ; column 2 shows the concordance of the sequence identifier used in Appendix A with or the sequence identifier of the WIPO Standard ST. 25 sequence listing)

Seq ID as used in Appendix A	SEQ ID as used in WIPO Standard ST. 25 sequence listing	Sequence name	Species	Function	ORF position
1	1	AtSUS-2 ^a	Arabidopsis thaliana	Sucrose synthase (Sucrose-UDP glucosyltransferase)	1-2415
4	3	AtSUS-2b	Arabidopsis thaliana	Sucrose synthase (Sucrose-UDP glucosyltransferase)	263-2653
7	7	BnSUS-2277	Brassica napus	Sucrose synthase (Sucrose-UDP glucosyltransferase)	93-2507
10	11	GmSUS-2025	Glycine max	Sucrose synthase (Sucrose-UDP glucosyltransferase)	244-2658
13	15	GmSUS-4819	Glycine max	Sucrose synthase (Sucrose-UDP glucosyltransferase)	84-2501
16	19	ZmSUS-7691	Zea mays	Sucrose synthase (Sucrose-UDP glucosyltransferase)	56-2461
19	23	ZmSUS-8057	Zea mays	Sucrose synthase (Sucrose-UDP glucosyltransferase)	156-2603
22	27	TaSUS-4775	Triticum aestivum	Sucrose synthase (Sucrose-UDP glucosyltransferase)	97-2544
25	31	OsSUS-9471	Oryza sativa	Sucrose synthase (Sucrose-UDP glucosyltransferase)	24-2447
28	35	PpSUS-2	Physcomitrella patens	Sucrose synthase (Sucrose-UDP glucosyltransferase)	227-2761

[Para 230] Those skilled in the art will recognize, or will be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the claims to the invention disclosed and claimed herein.

We claim:

1. An isolated polypeptide comprising an amino acid sequence as described by SEQ ID NO: 39, SEQ ID NO: 40, SEQ ID NO: 41 or SEQ ID NO: 42.
2. An isolated nucleic acid sequence encoding the polypeptide of claim 1.
3. An isolated polypeptide encoded by a nucleic acid sequence according to claim 2.
4. The isolated nucleic acid of claim 2 comprising a polynucleotide sequence selected from the group consisting of:
 - a. a polynucleotide sequence as described by SEQ ID NO: 7; SEQ ID NO: 9; SEQ ID NO: 5, SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35 or SEQ ID NO: 37;
 - b. a polynucleotide sequence encoding a polypeptide as described by SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 6, SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36 or SEQ ID NO: 38;
 - c. a polynucleotide sequence having at least 70% sequence identity with the nucleic acid of a) or b) above;
 - d. a polynucleotide sequence that is complementary to the nucleic acid of a) or b) above; and
 - e. a polynucleotide sequence that hybridizes under stringent conditions to nucleic acid of a) or b) above.
5. An isolated polypeptide encoded by a polynucleotide sequence as claimed in claim 4.
6. The isolated nucleic acid of claim 2, wherein the isolated nucleic acid encodes a polypeptide that functions as a modulator of a seed storage compound in microorganisms or plants.
7. The isolated polypeptide of claim 2, wherein the isolated polypeptide sequence functions as a modulator of a seed storage compound in microorganisms or plants.
8. An expression vector containing the nucleic acid of Claim 2, wherein the nucleic acid is operatively linked to a promoter selected from the group consisting of a seed-specific promoter, a root-specific promoter, and a non-tissue-specific promoter.
9. A method of producing a transgenic plant having a modified level of a seed storage compound weight percentage compared to the empty vector control comprising,
 - a. a first step of introduction into a plant cell of an expression vector containing a nucleic acid, and
 - b. a further step of generating from the plant cell the transgenic plant,wherein the nucleic acid encodes a polypeptide that functions as a modulator of a seed storage compound in the plant, and wherein the nucleic acid comprises a polynucleotide sequence of claim 2.

10. The method of Claim 6, wherein the nucleic acid comprises a polynucleotide sequence having at least 90% sequence identity with the polynucleotide sequence of a) or b) of Claim 4.
11. The method of Claim 9, wherein the total seed oil content weight percentage is increased in the transgenic plant as compared to an empty vector control.
12. A method of modulating the level of a seed storage compound weight percentage in a plant comprising, modifying the expression of a nucleic acid in the plant, comprising
 - a. a first step of introduction into a plant cell of an expression vector comprising a nucleic acid, and
 - b. a further step of generating from the plant cell the transgenic plant,wherein the nucleic acid encodes a polypeptide that functions as a modulator of a seed storage compound in the plant wherein the nucleic acid comprises a polynucleotide sequence of claim 2.
13. The method of Claim 12, wherein the total seed oil content weight percentage is increased in the transgenic plant as compared to an empty vector control.
14. A transgenic plant made by a method of claims 9 or 12.
15. The transgenic plant of Claim 14, wherein the total seed oil content weight percentage is increased in the transgenic plant as compared to an empty vector control.
16. The transgenic plant of Claim 14, wherein the plant is selected from the group consisting of rapeseed, canola, linseed, soybean, sunflower, maize, oat, rye, barley, wheat, pepper, tagetes, cotton, oil palm, coconut palm, flax, castor, sugarbeet, rice and peanut.
17. A seed produced by the transgenic plant of Claim 14, wherein the plant expresses the polypeptide that functions as a modulator of a seed storage compound and wherein the plant is true breeding for a modified level of seed storage compound weight percentage as compared to an empty vector control.

Figure 1A: Seq ID 1 - Nucleic acid sequence of AtSUS-2a

ATGGCGTCTTTCTTTGATCTCGTATACATAAGCTGGATTGGTTTCTTAAACAA
ATTCCTCTCCTTTTGGGTCTTCTGGGTTTGCCTTGTAAGATATGTAGCCCAG
GGAAAGGGGATATTGCAGTCCCACCAGCTGATTGATGAGTTCCTTAAGACT
GTGAAAGTTGATGGAACATTAGAAGATCTTAACAAAAGTCCATTCATGAAAG
TTCTCAGGAAGCCATAGTTTTGCCTCCATTTGTTGCTTTGGCTATACGTCC
CAGACCTGGTGTAGGGAATATGTCCGTGTGAATGTGTATGAGCTGAGCGT
AGATCATTTAACTGTTTCTGAATATCTTCGGTTTAAGGAAGAGCTCGTTAATG
GCCATGCCAATGGAGATTATCTCCTTGAACCTTGATTTTGAACCTTTCAATGCA
ACATTGCCTCGCCCAACTCGTTCATCATCCATTGGGAATGGGGTTCAGTTCC
TCAATCGTCACCTCTCTTCAATTATGTTCCGTAACAAAGAAAGCATGGAGCC
TTTGCTTGAGTTTCTCCGCACTCACAAACATGATGGCCGTCCTATGATGCTG
AATGATCGAATACAGAATATCCCCATACTTCAGGGAGCTTTGGCAAGAGCA
GAGGAGTTCCTTTCTAAACTTCCTCTGGCAACACCATACTCTGAATTCGAAT
TTGAACTACAAGGGATGGGATTTGAAAGGGGATGGGGTGACACAGCACAGA
AGGTTTCAGAAATGGTGCATCTTCTTCTGGACATACTCCAGGCACCTGATCC
TTCTGTCTTGGAGACGTTTCTAGGAAGGATTCTATGGTGTTCATGTTGTG
ATTTTGTCTCCGCATGGTTACTTTGGCCAAGCCAATGTCTTGGGTCTGCCTG
ATACTGGTGGACAGGTTGTCTACATTCTTGATCAAGTACGTGCATTGGAAAA
TGAGATGCTCCTTAGGATACAGAAGCAAGGACTGGAAGTTATTCCAAAGATT
CTCATTGTAACAAGACTGCTACCCGAAGCAAAGGGAACAACGTGCAACCAG
AGGTTAGAAAGAGTTAGTGGTACAGAACACGCACACATTCTGCGAATACCAT
TTAGGACTGAAAAGGGAATTCTTCGCAAGTGGATCTCAAGGTTTGATGTCTG
GCCATACCTGGAGACTTTTGCAGAGGATGCATCAAATGAAATTTCTGCGGA
GTTGCAGGGTGTACCAAATCTCATCATTGGCAACTACAGTGATGGAAATCTC
GTTGCTTCTTTGTTAGCTAGTAAGCTAGGTGTGATACAGTGTAATATTGCTCA
TGCTTTAGAGAAAACCAAGTACCCCGAGTCTGACATTTACTGGAGAAACCAT
GAAGATAAGTATCACTTTTCAAGTCAGTTCAGTGCAGATCTAATTGCCATGA
ATAATGCCGATTTTCATCATCACCAGCACATACCAAGAGATTGCGGGAAGCAA
GAACAATGTTGGGCAATACGAGAGCCACACAGCTTTCAGTATGCCTGGTCT
TTACCGAGTTGTTTCATGGAATTGATGTCTTTGATCCTAAGTTTAATATAGTCT
CTCCAGGAGCTGATATGACCATATACTTTCCATATTCTGACAAGGAAAGAAG
ACTCACTGCCCTTCATGAGTCAATTGAAGAACTCCTCTTTAGTGCCGAACAG
AATGATGAGCATGTTGGTTTACTGAGCGACCAATCGAAGCCAATCATCTTCT
CTATGGCAAGACTTGACAGGGTGAAAACTTGACTGGGCTAGTTGAATGCT
ATGCCAAGAATAGCAAGCTTAGAGAGCTTGCAAATCTTGTTATAGTCGGTGG
CTACATCGATGAGAATCAGTCCAGGGATAGAGAGGAAATGGCTGAGATACA
AAAGATGCACAGCCTGATTGAGCAGTATGATTTACACGGTGAGTTTAGGTG
GATAGCTGCTCAAATGAACCGTGCTCGAAATGGTGAGCTTTACCGTTATATC
GCAGACACAAAAGGTGTTTTTGTTCAGCCTGCTTTCTATGAAGCATTTGGGC
TTACGGTTGTGGAATCAATGACTTGTGCACTCCCAACGTTTGCTACCTGTCA
TGGTGGACCCGCAGAGATTATCGAAAACGGAGTTTCTGGGTTCCACATTGA
CCCATATCATCCAGACCAGGTTGCAGCTACCTTGGTCAGCTTCTTTGAGACC
TGTAACACCAATCCAAATCATTGGGTAAAATCTCTGAAGGAGGGCTCAAGC
GAATCTATGAAAGGTACACATGGAAGAAGTACTCAGAGAGACTGCTTACCCT

GGCTGGAGTCTATGCATTCTGGAAACATGTGTCTAAGCTCGAAAGGAGAGA
AACACGACGTTACCTAGAGATGTTTTACTCATTGAAATTTCTGTGATTTGGCCA
ATTCAATCCCGCTGGCAACAGATGAGAACTGA

Figure 1B: Seq ID 2 - Nucleic acid sequence of the open reading frame of AtSUS-2a

ATGGCGTCTTTCTTTGATCTCGTATACATAAGCTGGATTGGTTTCTTAAACAA
ATTCCTCTCCTTTTGGGTCTTCTGGGTTTGCCTTGTAAGATATGTAGCCCAG
GGAAAGGGGATATTGCAGTCCCACCAGCTGATTGATGAGTTCCTTAAGACT
GTGAAAGTTGATGGAACATTAGAAGATCTTAACAAAAGTCCATTCATGAAAG
TTCTGCAGGAAGCCATAGTTTTGCCTCCATTTGTTGCTTTGGCTATACGTCC
CAGACCTGGTGTAGGGAATATGTCCGTGTGAATGTGTATGAGCTGAGCGT
AGATCATTTAACTGTTTCTGAATATCTTCGGTTTAAGGAAGAGCTCGTTAATG
GCCATGCCAATGGAGATTATCTCCTTGAACCTTGATTTTGAACCTTTCAATGCA
ACATTGCCTCGCCCAACTCGTTCATCATCCATTGGGAATGGGGTTTCAGTTCC
TCAATCGTCACCTCTCTTCAATTATGTTCCGTAACAAAGAAAGCATGGAGCC
TTTGCTTGAGTTTCTCCGCACTCACAAACATGATGGCCGTCCTATGATGCTG
AATGATCGAATACAGAATATCCCCATACTTCAGGGAGCTTTGGCAAGAGCA
GAGGAGTTCCTTTCTAAACTTCCTCTGGCAACACCATACTCTGAATTCGAAT
TTGAACTACAAGGGATGGGATTTGAAAGGGGATGGGGTGACACAGCACAGA
AGGTTTCAGAAATGGTGCATCTTCTTCTGGACATACTCCAGGCACCTGATCC
TTCTGTCTTGGAGACGTTTCTAGGAAGGATTCCTATGGTGTTCATGTTGTG
ATTTTGTCTCCGCATGGTACTTTGGCCAAGCCAATGTCTTGGGTCTGCCTG
ATACTGGTGGACAGGTTGTCTACATTCTTGATCAAGTACGTGCATTGGAAAA
TGAGATGCTCCTTAGGATACAGAAGCAAGGACTGGAAGTTATTCCAAAGATT
CTCATTGTAACAAGACTGCTACCCGAAGCAAAGGGGAACAACGTGCAACCAG
AGGTTAGAAAGAGTTAGTGGTACAGAACACGCACACATTCTGCGAATACCAT
TTAGGACTGAAAAGGGAATTCTTCGCAAGTGGATCTCAAGGTTTGATGTCTG
GCCATACCTGGAGACTTTTGCAGAGGATGCATCAAATGAAATTTCTGCGGA
GTTGCAGGGTGTACCAAATCTCATCATTGGCAACTACAGTGATGGAAATCTC
GTTGCTTCTTTGTTAGCTAGTAAGCTAGGTGTGATACAGTGTAATATTGCTCA
TGCTTTAGAGAAAACCAAGTACCCCGAGTCTGACATTTACTGGAGAAACCAT
GAAGATAAGTATCACTTTTCAAGTCAGTTCAGTGCAGATCTAATTGCCATGA
ATAATGCCGATTTTCATCATCACCAGCACATACCAAGAGATTGCGGGAAGCAA
GAACAATGTTGGGCAATACGAGAGCCACACAGCTTTCACTATGCCTGGTCT
TTACCGAGTTGTTTCATGGAATTGATGTCTTTGATCCTAAGTTTAATATAGTCT
CTCCAGGAGCTGATATGACCATATACTTTCCATATTCTGACAAGGAAAGAAG
ACTCACTGCCCTTCATGAGTCAATTGAAGAACTCCTCTTTAGTGCCGAACAG
AATGATGAGCATGTTGGTTTACTGAGCGACCAATCGAAGCCAATCATCTTCT
CTATGGCAAGACTTGACAGGGTGAAAACTTGACTGGGCTAGTTGAATGCT
ATGCCAAGAATAGCAAGCTTAGAGAGCTTGCAAATCTTGTTATAGTCGGTGG
CTACATCGATGAGAATCAGTCCAGGGATAGAGAGGAAATGGCTGAGATACA
AAAGATGCACAGCCTGATTGAGCAGTATGATTTACACGGTGAGTTTAGGTG
GATAGCTGCTCAAATGAACCGTGCTCGAAATGGTGAGCTTTACCGTTATATC
GCAGACACAAAAGGTGTTTTTGTTCAGCCTGCTTTCTATGAAGCATTG

TTACGGTTGTGGAATCAATGACTTGTGCACTCCCAACGTTTGCTACCTGTCA
 TGGTGGACCCGCAGAGATTATCGAAAACGGAGTTTCTGGGTTCCACATTGA
 CCCATATCATCCAGACCAGGTTGCAGCTACCTTGGTCAGCTTCTTTGAGACC
 TGTAACACCAATCCAAATCATTGGGTTAAAATCTCTGAAGGAGGGCTCAAGC
 GAATCTATGAAAGGTACACATGGAAGAAGTACTCAGAGAGACTGCTTACCCT
 GGCTGGAGTCTATGCATTCTGGAAACATGTGTCTAAGCTCGAAAGGAGAGA
 AACACGACGTTACCTAGAGATGTTTTACTCATTGAAATTTCTGTGATTTGGCCA
 ATTC AATCCCGCTGGCAACAGATGAGAAC

Figure 1C: Seq ID 3 - Amino acid sequence of the open reading frame of AtSUS-2a

MASFFDLVYISWIGFLNKFLSFVFWVCLVRYVAQGKGILQSHQLIDFLKTVKV
 DGTLEDLNKSPFMKVLQEAIPLPFVALAIRPRPGVREYVRVNVYELSVDHLTVS
 EYLRKFKEELVNGHANGDYLLDFEPFNATLPRPTRSSSIGNGVQFLNRHLSSI
 MFRNKESMEPLLEFLRTHKHDGRPMMLNDRIQNIPILQGALARAEFLSKLPLA
 TPYSEFEFELQGMGFERGWGDTAQKVSEMVHLLLDILQAPDPSVLETFLGRIP
 MVFNVLSPHGYFGQANVLGLPDTGGQVYILDQVRALENEMLLRIQKQGLEV
 IPKILIVTRLLPEAKGTTCNQRLERVSGTEHAHILRIPRTEKGILRKWISRFDVWP
 YLETFAEDASNEISAELOQGVNLIIGNYSDGNLVLASLLASKLGVICQNIHALEKT
 KYPESDIYWRNHEDKYHFSSQFTADLIAMNNADFIITSTYQEIAGSKNNVGQYES
 HTAFTMPGLYRVVHGIDVDFPKFNIVSPGADMTIYFPYSDKERRLTALHESIEEL
 LFSAEQNDEHVGLLSDQSKPIIFSMARLDRVKNLTGLVECYAKNSKLRELANLVI
 VGGYIDENQSRDREEMAEIQKMHSLEQYDLHGFEFRWIAAQMNRRARNGELYRY
 IADTKGVFVQPAFYEAFLTVVESMTCALPTFATCHGGPAEIIENGVSFGFHIDPY
 HPDQVAATLVSSFFETCNTNPNHWVKISEGGLKRIYERYTWKKYSERLLTLAGVY
 AFWKHVSKLERRETRRYLEMFYSLKFRDLANSIPLATDEN

Figure 2A: Seq ID 4 – Nucleic acid sequence of AtSUS-2b

ATGGCGTCTTTCTTTGATCTCGTATACATAAGCTGGATTGGTTTCTTAAACAA
 ATTCCTCTCCTTTTGGGTCTTCTGGGTTTGCCTTGTAAGTGTTTGTGTTTTTG
 CCTCTGAGAAAAAATCATGCCGACTGGTAGGTTGAGACTATGCGTGAATG
 GGTTTATGACGCTATCTCTGCTCAACGCAATGAGCTCCTCTCTCTTTTCTCC
 AGGTTTCTTCTTAACCCTCTCGCTCTCTCTCCCTGAGCCTTGCTTATAAATGC
 ATGTGTGTTGTGATTGGTCTCAAGTCATGGGTAATGGTTTTGGTTGTTATCTT
 TATTAGATATGTAGCCCAGGGAAAGGGGATATTGCAGTCCCACCAGCTGAT
 TGATGAGTTCCTTAAGACTGTGAAAGTTGATGGAACATTAGAAGATCTTAAC
 AAAAGTCCATTCATGAAAGTTCTGCAGTCTGCAGAGGAAGCCATAGTTTTGC
 CTCCATTTGTTGCTTTGGCTATACGTCCCAGACCTGGTGTTAGGGAATATGT
 CCGTGTGAATGTGTATGAGCTGAGCGTAGATCATTTAACTGTTTCTGAATAT
 CTTCGGTTTAAGGAAGAGCTCGTTAATGGCCATGCCAATGGAGATTATCTCC
 TTGAACTTGATTTTGAACCTTTCAATGCAACATTGCCTCGCCCAACTCGTTCA
 TCATCCATTGGGAATGGGGTTCAGTTCCTCAATCGTCACCTCTCTTCAATTA
 TGTTCCGTAACAAAGAAAGCATGGAGCCTTTGCTTGAGTTTCTCCGCACTCA
 CAAACATGATGGCCGTCCTATGATGCTGAATGATCGAATACAGAATATCCCC

ATACTTCAGGGAGCTTTGGCAAGAGCAGAGGAGTTCCTTTCTAACTTCCTC
TGGCAACACCATACTCTGAATTCGAATTTGAACTACAAGGGATGGGATTTGA
AAGGGGATGGGGTGACACAGCACAGAAGGTTTCAGAAATGGTGCATCTTCT
TCTGGACATACTCCAGGCACCTGATCCTTCTGTCTTGGAGACGTTTCTAGGA
AGGATTCCTATGGTGTTCATGTTGTGATTTTGTCTCCGCATGGTTACTTTG
GCCAAGCCAATGTCTTGGGTCTGCCTGATACTGGTGGACAGGTTGTCTACA
TTCTTGATCAAGTACGTGCATTGGAAAATGAGATGCTCCTTAGGATACAGAA
GCAAGGACTGGAAGTTATTCCAAAGATTCTCATTGTAACAAGACTGCTACCC
GAAGCAAAGGGAACAACGTGCAACCAGAGGTTAGAAAGAGTTAGTGGTACA
GAACACGCACACATTCTGCGAATACCATTTAGGACTGAAAAGGGAATTCTTC
GCAAGTGGATCTCAAGGTTTGATGTCTGGCCATACCTGGAGACTTTTGCAG
AGGATGCATCAAATGAAATTTCTGCGGAGTTGCAGGGTGTACCAAATCTCAT
CATTGGCAACTACAGTGATGGAAATCTCGTTGCTTCTTTGTTAGCTAGTAAG
CTAGGTGTGATACAGTGTAATATTGCTCATGCTTTAGAGAAAACCAAGTACC
CCGAGTCTGACATTTACTGGAGAAACCATGAAGATAAGTATCACTTTTCAAG
TCAGTTCACTGCAGATCTAATTGCCATGAATAATGCCGATTTTCATCATCACC
AGCACATACCAAGAGATTGCGGGAAGCAAGAACAATGTTGGGCAATACGAG
AGCCACACAGCTTTCACTATGCCTGGTCTTTACCGAGTTGTTTCATGGAATTG
ATGTCTTTGATCCTAAGTTTAATATGGTCTCTCCAGGAGCTGATATGACCATA
TACTTTCCATATTCCGACAAGGAAAGAAGACTCACTGCCCTTCATGAGTCAA
TTGAAGAACTCCTCTTTAGTGCCGAACAGAATGATGAGCATGTTGGTTTACT
GAGCGACCAATCGAAGCCAATCATCTTCTCTATGGCAAGACTTGACAGGGT
GAAAACTTGACTGGGCTAGTTGAATGCTATGCCAAGAATAGCAAGCTTAGA
GAGCTTGCAAATCTTGTTATAGTTCGGTGGCTACATCGATGAGAATCAGTCCA
GGGATAGAGAGGAAATGGCTGAGATACAAAAGATGCACAGCCTGATTGAGC
AGTATGATTTACACGGTGAGTTTAGGTGGATAGCTGCTCAAATGAACCGTGC
TCGAAATGGTGAGCTTTACCGTTATATCGCAGACACAAAAGGTGTTTTTGT
CAGCCTGCTTTCTATGAAGCATTGTTGGGCTTACGGTTGTGGAATCAATGACTT
GTGCACTCCCAACGTTTGCTACCTGTCATGGTGGACCCGCAGAGATTATCG
AAAACGGAGTTTCTGGGTTCCACATTGACCCATATCATCCAGACCAGGTTGC
AGCTACCTTGGTCAGCTTCTTTGAGACCTGTAACACCAATCCAAATCATTGG
GTTAAAATCTCTGAAGGAGGGCTCAAGCGAATCTATGAAAGGTACACATGG
AAGAAGTACTCAGAGAGACTGCTTACCCTGGCTGGAGTCTATGCATTCTGG
AAACATGTGTCTAAGCTCGAAAGGAGAGAAACACGACGTTACCTAGAGATG
TTTTACTCATTGAAATTTCTGTGATTGGCCAATTCAATCCCGCTGGCAACAGA
TGAGAACTGATTAATTAA

Figure 2B: Seq ID 5 – Nucleic acid sequence of the open reading frame of AtSUS-2b

ATGTGTGTTGTGATTGGTCTCAAGTCATGGGTAATGGTTTTGGTTGTTATCTT
TATTAGATATGTAGCCCAGGGGAAAGGGGATATTGCAGTCCCACCAGCTGAT
TGATGAGTTCCTTAAGACTGTGAAAGTTGATGGAACATTAGAAGATCTTAAC
AAAAGTCCATTCATGAAAGTTCTGCAGTCTGCAGAGGAAGCCATAGTTTTGC
CTCCATTTGTTGCTTTGGCTATACGTCCAGACCTGGTGTAGGGAATATGT
CCGTGTGAATGTGTATGAGCTGAGCGTAGATCATTTAACTGTTTCTGAATAT

CTTCGGTTTAAGGAAGAGCTCGTTAATGGCCATGCCAATGGAGATTATCTCC
TTGAACTTGATTTTGAACCTTTCAATGCAACATTGCCTCGCCCAACTCGTTCA
TCATCCATTGGGAATGGGGTTTCAAGTTCTCAATCGTCCCTCTCTTCAATTA
TGTTCCGTAACAAAGAAAGCATGGAGCCTTTGCTTGAGTTTCTCCGCACTCA
CAAACATGATGGCCGTCCTATGATGCTGAATGATCGAATACAGAATATCCCC
ATACTTCAGGGAGCTTTGGCAAGAGCAGAGGAGTTCTTTCTAAACTTCCTC
TGGCAACACCATACTCTGAATTCGAATTTGAACTACAAGGGATGGGATTTGA
AAGGGGATGGGGTGACACAGCACAGAAGGTTTCAGAAATGGTGCATCTTCT
TCTGGACATACTCCAGGCACCTGATCCTTCTGTCTTGGAGACGTTTCTAGGA
AGGATTCCTATGGTGTTCAATGTTGTGATTTTGTCTCCGCATGGTTACTTTG
GCCAAGCCAATGTCTTGGGTCTGCCTGATACTGGTGGACAGGTTGTCTACA
TTCTTGATCAAGTACGTGCATTGGAAAATGAGATGCTCCTTAGGATACAGAA
GCAAGGACTGGAAGTTATTCCAAAGATTCTCATTGTAACAAGACTGCTACCC
GAAGCAAAGGGAACAACGTGCAACCAGAGGTTAGAAAGAGTTAGTGGTACA
GAACACGCACACATTCTGCGAATACCATTTAGGACTGAAAAGGGAATTCTTC
GCAAGTGGATCTCAAGGTTTGATGTCTGGCCATACCTGGAGACTTTTGCAG
AGGATGCATCAAATGAAATTTCTGCGGAGTTGCAGGGTGTACCAAATCTCAT
CATTGGCAACTACAGTGATGGAAATCTCGTTGCTTCTTTGTTAGCTAGTAAG
CTAGGTGTGATACAGTGTAATATTGCTCATGCTTTAGAGAAAACCAAGTACC
CCGAGTCTGACATTTACTGGAGAAACCATGAAGATAAGTATCACTTTTCAAG
TCAGTTCACTGCAGATCTAATTGCCATGAATAATGCCGATTTTCATCATCACC
AGCACATACCAAGAGATTGCGGGAAGCAAGAACAATGTTGGGCAATACGAG
AGCCACACAGCTTTCACTATGCCTGGTCTTTACCGAGTTGTTTCATGGAATTG
ATGTCTTTGATCCTAAGTTTAATATGGTCTCTCCAGGAGCTGATATGACCATA
TACTTTCCATATTCCGACAAGGAAAGAAGACTCACTGCCCTTCATGAGTCAA
TTGAAGAACTCCTCTTTAGTGCCGAACAGAATGATGAGCATGTTGGTTTACT
GAGCGACCAATCGAAGCCAATCATCTTCTCTATGGCAAGACTTGACAGGGT
GAAAACTTGACTGGGCTAGTTGAATGCTATGCCAAGAATAGCAAGCTTAGA
GAGCTTGCAAATCTTGTTATAGTCGGTGGCTACATCGATGAGAATCAGTCCA
GGGATAGAGAGGAAATGGCTGAGATACAAAAGATGCACAGCCTGATTGAGC
AGTATGATTTACACGGTGAGTTTAGGTGGATAGCTGCTCAAATGAACCGTGC
TCGAAATGGTGAGCTTTACCGTTATATCGCAGACACAAAAGGTGTTTTTGT
CAGCCTGCTTTCTATGAAGCATTGTTGGGCTTACGGTTGTGGAATCAATGACTT
GTGCACTCCCAACGTTTGCTACCTGTCATGGTGGACCCGCAGAGATTATCG
AAAACGGAGTTTCTGGGTTCCACATTGACCCATATCATCCAGACCAGGTTGC
AGCTACCTTGGTCAGCTTCTTTGAGACCTGTAACACCAATCCAAATCATTGG
GTAAAATCTCTGAAGGAGGGCTCAAGCGAATCTATGAAAGGTACACATGG
AAGAAGTACTCAGAGAGACTGCTTACCCTGGCTGGAGTCTATGCATTCTGG
AAACATGTGTCTAAGCTCGAAAGGAGAGAAACACGACGTTACCTAGAGATG
TTTTACTCATTGAAATTCGTGATTTGGCCAATTCAATCCCGCTGGCAACAGA
TGAGAAC

Figure 2C: Seq ID 6 - Amino acid sequence of the open reading frame of AtSUS-2b

MCVVIGLKSWVMVLVVIFIRYVAQGKGILQSHQLIDEFLKTVKVDGTLEDLNKSP
FMKVLQSAEEAIVLPPFVALAIRPRPGVREYVRVNVYELSDHDTVSEYLRFKKEE
LVNGHANGDYLLLELDFEPFNATLPRPTRSSSIGNGVQFLNRHLSSIMFRNKESM
EPLLEFLRTHKHHDGRPMMLNDRIQNIPILQGALARAEFLSKLPLATPYSEFEFE
LQGMGFERGWGDTAQKVSEMVHLLLDILQAPDPSVLETFLGRIPMVFNVVILSP
HGYFGQANVLGLPDTGGQVVYILDQVRALENEMLLRIQKQGLEVIPKILIVTRLLP
EAKGTTCNQRLERVSGTEHAHILRIPFRTEKGILRKWISRFDVWPYLETFAEDAS
NEISAEALQGVPNLIIGNYSDGNLVLASLLASKLGVICNIAHALEKTKYPESDIYWR
NHEDKYHFSSQFTADLIAMNNADFIITSTYQEIAGSKNNVGGQYESHTAFTMPGL
YRVVHGIDVFDPKFNMVSPGADMTIYFPYSDKERRLTALHESIEELLFSAEQND
EHVGLLSQSKPIIFSMARLDRVKNLTGLVECYAKNSKLRELANLVIVGGYIDEN
QSRDREEMAEIQKMHSLEQYDLHGEFRWIAAQMNRRARNGELYRYIADTKGVF
VQPAFYEAFLTVVESMTALPTFATCHGGPAEIIENGVS GFHIDPYHPDQVAA
TLVSFFETCNTNPNHWVKISEGGLKRIYERYTWKKYSERLLTAGVYAFWKHVS
KLERRETRRYLEM FYSLKFRDLANSIPLATDEN

Figure 3A: Seq ID 7 - Nucleic acid sequence of BnSUS-2277

CCCGGGTCGACCATTTCTGTTCTACTCCAGTTTCCAAGAAAAAAAAAAAAAAC
TCTGTTATCATTGTTTCCTGCGTCTCTCTGTTAAGGATCAATGGTGAACGGG
GTGTTAACGCGCGCCACAGCCAACGTGAGCGTTTGAACGAAACGCTCGCT
GCTCAAAGAAACGAAGTCCTTGCCCTTACTTTCCAGGGTTGAAGGCAAAGGT
AAAGGCATCTTGCAACAAAACCAGATCATTGCTGAATTCGAGGCTTTGCCTG
AAGAAACCCAAAAAAGATTGAAGGTGGTGGCTTTCTTCGACCTTCTCAAAC
CACTCAGGAAGCAATAGTGTTGCCACCATGGGTTGCTCTTGCTGTGAGGCC
AAGGCCTGGTGTGTTGGGAGTACATAAGAGTCAATCTCCATGCTCTTGTTGTT
GAGGAGCTTACACCTGCTGAGTTTCTTCATTTCAAGGAAGAGCTTGTTGATG
GAGTTAAGAACGGTGATTTACGCTTGAGCTTGACTTTGAGCCGTTTAATGC
CTCTGTCCCTCGCCCAACGCTCCCCAAGTACATTGGAGATGGTGTGAGTT
TCTCAACCGTCACCTCTCGGCTAAGCTCTTCCATGACAAGGAGAGTTTGCTT
CCTTTGCTCAAGTTCCTTCGACTACACAGCTACCAGGGAAAGACCCTGATG
CTGAACGAGAGAGTTCAAAACCTCAACAATCTGCAACACATCTTGAGGAAG
GCAGAGGAGTATCTGGCAGGTCTTTCACCTGAAACGCCTTATGAAGATTTTG
AGGCCAAGTTTGAGGAGATTGGTCTTGAGAGGGGATGGGGAAACAATGCC
GAGCGTGTCTTGACATGATCCGTCTTCTTTGGACCTCCTCGAGGCCCT
GACCCTTGCACTCTAGAGAACTTTCTTGGGAGAGTCCCTATGGTGTTCACG
TTGTCATCCTCTCTCCACATGGTTACTTTGCTCAGGACAATGTTCTTGGTTAC
CCTGACACTGGTGGTCAGGTTGTTTACATTCTGGATCAAGTCCGTGCCATG
GAGACAGAGATGCTCCAACGTATTCAGCAGCAAGGACTCAACATTACTCCA
AGAATTCTCATTCTCACTAGACTGCTCCCTGATGCAGTAGGAACCATGTG
GTGAGCGTCTTGAGAGAGTTGATGGATCTGAGTACTGTGACATCCTCCGTG
TGCCCTTCAGAACAGAGAAGGGTATCGTTCGCAAATGGATCTCAAGATTG
AAGTCTGGCCATATCTAGAGACTTACACCGAGGATGCTGCCGTTGAGCTTG

CTAAAGAGTTGAAGGGCAAGCCTGACCTTATCATTGGTAACTACAGTGATG
GCAACCTCGTTGCCTCTTTACTGGCACACAACTTGGTGTCACTCAGTGTAC
CATTGCTCACGCCCTGGAGAAGACAAAGTACCCTGACTCGGATATATACTG
GAAGAAGCTTGACGATAAGTACCATTTCTCATGCCAGTTCACTGCAGATCTA
TTTGCAATGAACCATACTGATTTTCATCATCACAAAGTACTCTCCAAGAAATTGC
TGGAAGCAAGGACACAGTTGGGCAGTATGAGAGCCACACGGCCTTTACTCT
TCCTGGACTGTACCGAGTTGTTTCATGGCATTGATGTGTTTGATCCCAAGTTC
AACATTGTCTCTCCTGGTGCTGACATGAGCATCTACTTCCCATACTGAGG
AGAAGCGTAGATTGACTAAGTTCCACCCAGAGATTGAGGAGCTCCTCTACA
GCAATGTTGAAAACGAAGAGCACTTATGTGTGCTCAAGGACAAGAAGAAGC
CAATCCTCTTCACAATGGCTAGGCTCGACCGTGTCAAGAACTTGTCAAGGTCT
TGTTGAGTGGTACGGGAAGAACAACGCCTGCGTGAGCTTGTTAACTTGGT
GGTTGTTGGAGGAGACAGGAGGAAAGAGTCAAAGGACAACGAAGAGAAGG
CCGAGATGAAGAAAATGTATGACCTCATTGAGGAATACAAGCTCAACGGCC
AGTTCCGTTGGATCTCCTCCCAGATGAACAGGGTTAGGAACGGTGAGCTCT
ACAGGTACATCTGTGACACCAAGGGTGCTTTTGTGCAACCAGCATTGTATGA
AGCCTTTGGTTTGACTGTTGTGGAGGCTATGACCTGTGGGTACCAACATTC
GCCACTTGCAAAGGTGGTCCAGCTGAGATCATTGTCCACGGCAAATCAGGT
TTCCACATTGACCCTTACCATGGTGATCAGGCGGCTGATACTCTTGCTGATT
TCTTCACCAAGTGTAAGGAAGATCCATCTCACTGGGACCAGATCTCAAAGG
GAGGGCTTCAGAGGATTGAGGAGAAATACACGTGGCAGATTTACTCACAGA
GGCTCTTGACATTGACTGGTGTGTATGGATTCTGGAAGCATGTCTCGAACCT
TGACCGTCTTGAGAGCCGTCGTTACCTTGAGATGTTCTATGCATTGAAGTAT
CGCCCACTGGCTCAGGCTGTTCTCTTGCTCAAGAGTAATGAAGAAAAGAG
AGAAAAAATGTGGAACCATTGAAGAGCTTGAAGATAATGGTTCCAGTGTTTG
AAGAATATCAAAAATGTCTTTTGATTCTATCATCATTTGGATTAGGAGCGC
TTTTGTGTTCTCTTGCTTACTTTGTTGTTTCAATCCTTTTGTCTTTGTGTATT
TTCTTTTCTTGCTTTTCCGGCAATGGTTTGTACTCTGATTGTTGAGGGATCT
CCCTTATGATTGGCAATGAGATATCGATTATTTTATCACGAGATCGTCGACC
CGGAAATT

Figure 3B: Seq ID 8 - Nucleic acid sequence of the open reading frame of BnSUS-2277

ATGGTGAACGGGGTGTTAACGCGCGCCACAGCCAACGTGAGCGTTTGAA
CGAAACGCTCGCTGCTCAAAGAAACGAAGTCCTTGCCTTACTTTCCAGGGT
TGAAGGCAAAGGTAAAGGCATCTTGCAACAAAACCAGATCATTGCTGAATTC
GAGGCTTTGCCTGAAGAAACCCAAAAAAGATTGAAGGTGGTGCTTTCTTC
GACCTTCTCAAACCACTCAGGAAGCAATAGTGTTGCCACCATGGGTGCT
CTTGCTGTGAGGCCAAGGCCTGGTGTGTTGGGAGTACATAAGAGTCAATCTC
CATGCTCTTGTTGTTGAGGAGCTTACACCTGCTGAGTTTCTTCATTTCAAGG
AAGAGCTTGTTGATGGAGTTAAGAACGGTGATTTACGCTTGAGCTTGACTT
TGAGCCGTTTAATGCCTCTGTCCCTCGCCCAACGCTCCCCAAGTACATTGG
AGATGGTGTTGAGTTTCTCAACCGTCACCTCTCGGCTAAGCTCTTCCATGAC
AAGGAGAGTTTGCTTCTTTGCTCAAGTTCCTTCGACTACACAGCTACCAGG
GAAAGACCCTGATGCTGAACGAGAGAGTTCAAACCTCAACAATCTGCAAC

ACATCTTGAGGAAGGCAGAGGAGTATCTGGCAGGTCTTTACCTGAAACGC
CTTATGAAGATTTTGAGGCCAAGTTTGAGGAGATTGGTCTTGAGAGGGGAT
GGGGAAACAATGCCGAGCGTGTCTTGACATGATCCGTCTTCTTTGGACC
TCCTCGAGGGCCCCTGACCCTTGCACTCTAGAGAACTTTCTTGGGAGAGTCC
CTATGGTGTTC AACGTTGTCATCCTCTCTCCACATGGTTACTTTGCTCAGGA
CAATGTTCTTGGTTACCCTGACACTGGTGGTCAGGTTGTTTACATTCTGGAT
CAAGTCCGTGCCATGGAGACAGAGATGCTCCAACGTATTCAGCAGCAAGGA
CTCAACATTACTCCAAGAATTCTCATTCTCACTAGACTGCTCCCTGATGCAG
TAGGAACCACATGTGGTGAGCGTCTTGAGAGAGTTGATGGATCTGAGTACT
GTGACATCCTCCGTGTGCCCTTCAGAACAGAGAAGGGTATCGTTCGCAAAT
GGATCTCAAGATTCGAAGTCTGGCCATATCTAGAGACTTACACCGAGGATG
CTGCCGTTGAGCTTGCTAAAGAGTTGAAGGGCAAGCCTGACCTTATCATTG
GTA ACTACAGTGATGGCAACCTCGTTGCCTCTTTACTGGCACACAACTTGG
TG TCACTCAGTGTACCATTGCTCACGCCCTGGAGAAGACAAAGTACCCTGA
CTCGGATATATACTGGAAGAAGCTTGACGATAAGTACCATTTCTCATGCCAG
T TCACTGCAGATCTATTTGCAATGAACCATACTGATTTTCATCATCACAAGTAC
TCTCCAAGAAATTGCTGGAAGCAAGGACACAGTTGGGCAGTATGAGAGCCA
CACGGCCTTTACTCTTCCTGGACTGTACCGAGTTGTTTCATGGCATTGATGTG
TTTGATCCCAAGTTCAACATTGTCTCTCCTGGTGCTGACATGAGCATCTACT
TCCCATACACTGAGGAGAAGCGTAGATTGACTAAGTTCCACCCAGAGATTG
AGGAGCTCCTCTACAGCAATGTTGAAAACGAAGAGCACTTATGTGTGCTCAA
GGACAAGAAGAAGCCAATCCTCTTCACAATGGCTAGGCTCGACCGTGTCAA
GAACTTGTCAGGTCTTGTTGAGTGGTACGGGAAGAACAACGCCTGCGTGA
GCTTGTTAACTTGGTGGTTGTTGGAGGAGACAGGAGGAAAGAGTCAAAGGA
CAACGAAGAGAAGGCCGAGATGAAGAAAATGTATGACCTCATTGAGGAATA
CAAGCTCAACGGCCAGTTCCGTTGGATCTCCTCCCAGATGAACAGGGT TAG
GAACGGTGAGCTCTACAGGTACATCTGTGACACCAAGGGTGCTTTTGTGCA
ACCAGCATTGTATGAAGCCTTTGGTTTGACTGTTGTGGAGGCTATGACCTGT
GGGTTACCAACATTGCGCACTTGCAAAGGTGGTCCAGCTGAGATCATTGTC
CACGGCAAATCAGGTTTCCACATTGACCCTTACCATGGTGATCAGGCGGCT
GATACTCTTGCTGATTTCTTCACCAAGTGTAAGGAAGATCCATCTCACTGGG
ACCAGATCTCAAAGGGAGGGCTTCAGAGGATTGAGGAGAAATACACGTGGC
AGATTTACTCACAGAGGCTCTTGACATTGACTGGTGTGTATGGATTCTGGAA
GCATGTCTCGAACCTTGACCGTCTTGAGAGCCGTCGTTACCTTGAGATGTTT
TATGCATTGAAGTATCGCCCACTGGCTCAGGCTGTTCTCTTGCTCAAGAG

Figure 3C: Seq ID 9 - Amino acid sequence of the open reading frame of BnSUS-2277

MVNGVLTRAHSQRERLNETLAAQRNEVLALLSRVEGKGKGILQQNQIIAEFEAL
PEETQKKIEGGAFFDLLKTTQEAI VLPWVALAVRPRPGVWEYIRVNLHALVVE
ELTPAEFLHFKEELVDGVKN GDFLELDFEPFNASVPRPTLPKYIGD GVEFLNR
HLSAKLFHDKESLLPLLKFLRLHSYQGKTLMLNERVQNLNNLQHILRKAEEYLAG
LSPETPYEDFEAKFEEIGLERGWGNNAERVLD MIRLLLDLLEAPDPCTLENFLG
RVPMVFN VVILSPHGYFAQDNVLGY PDTGGQVVYILDQVRAMETEM LQRIQQQ
GLNITPRILILTRLLPDAVGTT CGERLERVDGSEYCDILRV PFRTEKGIVRKWISR

FEVWPYLETYTEDAAVELAKELKGKPDLIIGNYSDGNLVASLLAHKLGVTQCTIA
HALEKTKYPDSDIYWKKLDDKYHFSCQFTADLFAMNHTDFIITSTLQEIAGSKDT
VGQYESHTAFTLPGLYRVVHGIDVFDPKFNIVSPGADMSIYFPYTEEKRRRLTKFH
PEIEELLYSNVENEEHLCVLKDKKKPILFTMARLDRVKNLSGLVEWYGKNKRLR
ELVNLVVVGDDRRESKDNEEKAEMKKMYDLIEEYKLNGQFRWISSQMNRVR
NGELYRYICDTKGAFVQPALYEAFLTVVEAMTCGLPTFATCKGGPAEIIIVHGK
SGFHIDPYHGDQAADTLADFFTKCKEDPSHWDQISKGGLQRIEEKYTWQIYSQ
RLLTLTGVYGFWKHVSNLDRLESRRYLEMFYALKYRPLAQAVPLAQE

Figure 4A: Seq ID 10 - Nucleic acid sequence of GmSUS-2025

GGAACCCTAGCATTGCGCCATTGCCGCCACCTTTCCCCACGCAGCTTTTGTT
GTAGCAAAAGATTGATATTTTCTTCTTCGTCTATAAAACCCACACGTCAGTC
ACTTCTTTTAACTCAATTTTGGGTCCCGTAGCTCTATATTTTCTCATTGCGA
GACAAGAAAAATCCAAGCCTCTCGATTACACCCCCCTCTCTTTTTTGC GTTC
ATTCTGTTTTCTGTTGAAGTCTTTCCTAGCCAAATGGCCACCGATCGTTTG
ACCCGGGTTCACAGTCTCCGTGAGAGGCTTGATGAAACCCTCACTGCCAAC
AGGAATGAAATTTTGGCCCTTCTGTCAAGGATCGAAGCCAAGGGCAAGGGC
ATCCTGCAACACCACCAGGTCATTGCTGAGTTTGAGGAAATCCCTGAGGAG
AACAGACAGAAGCTCACTGATGGTGCCTTTGGAGAAGTCTTGAGATCTACA
CAGGAAGCCATAGTTTTGCCACCATGGGTTGCTCTGGCTGTTCTGTCCAAGA
CCTGGTGTGTGGGAGTACCTGAGAGTGAATGTGCACGCTCTTGTTGTTGAG
GAGTTGCAACCTGCTGAGTACCTGCACTTCAAGGAAGAACTTGTTGATGGA
AGTTCTAATGGCAACTTTGTGCTTGAGTTGGACTTTGAACCATTCATGCAG
CCTTCCCTCGCCCAACTCTTAACAAGTCAATTGGAAATGGTGTGCAGTTCT
CAACCGCCACCTTTCTGCCAAACTCTTCCACGACAAGGAGAGCTTGCACCC
ACTTTTGGAATTCCTCAGGCTTCACAGCGTCAAGGGAAAGACTTTGATGTTG
AATGACAGAATTCAAAACCCAGATGCACTCCAACATGTTCTGAGGAAAGCTG
AGGAGTATCTGGGCACAGTGCCTCCTGAAACTCCCTACTCAGAATTTGAGC
ACAAGTTCCAGGAGATTGGTTTGGAGAGAGGGTGGGGTGACAACGCAGAG
CGTGTTCTTGAGTCAATTCAACTTCTCTTGATCTTCTTGAGGCCCTGACC
CTTGCAACCCTTGAGACTTTCCTTGGAAGAATTCCTATGGTCTTCAATGTTGT
CATTCTTTCTCCCCATGGTTACTTTGCCCAAGATAATGTCTTGGGATACCCT
GACACTGGTGGCCAGGTTGTTTACATCTTGATCAAGTTCTGTGCTTTGGAGA
ACGAGATGCTCCATCGCATTAAGCAACAAGGATTGGACATTGTACCTCGTAT
TCTCATTATCACCCGTCTTCTCCCCGATGCAATCGGAACTACTTGTGGCCAA
CGTCTTGAGAAGGTGTTCCGGAACCGAGCACTCCACATTCTTCGAGTTCCC
TTTAGAACTGAGAAGGGAATTGTTCTGTCAGTGGATCTCAAGATTGGAAGTCT
GGCCATACTTGGAACCTTACACTGAGGATGTTGCTCATGAGCTTGCCAAAG
AGTTGCAAGGCAAGCCAGATCTGATTGTGCGAAACTACAGTGATGGAAACA
TTGTCGCCTCTTTGTTGGCACATAAATTAGGAGTCACTCAGTGTACCATTGC
TCACGCACTTGAGAAGACCAATAACCCGAATCCGACATTTACTGGAAAAAA
TTGGAAGAGAGATACCACTTCTCTTGCCAATTCACAGCTGATCTATTTGCCA
TGAACCACACAGATTTTATTATCACCAGTACCTTCCAGGAGATTGCTGGAAG
CAAGGACACTGTTGGACAGTACGAGTCTCACACAGCCTTCACCCTTCTGG
ACTCTACCGTGTTGTGCACGGCATTGATGTCTTTGATCCAAAATTCAACATT

GTCTCCCCTGGAGCTGATCAAACCATTTACTTCCCCCCCCACCGAAACTAGC
CGTAGGTTGACCTCCTTCCACCCCGAAATCGAAGAACTTCTTTACAGCTCTG
TGGAGAATGAAGAACACATATGCGTGCTGAAGGACCGCAGCAAGCCGATTA
TCTTACCATGGCAAGGTTGGACCGTGTGAAGAACATCACAGGACTCGTGG
AGTGGTACGGTAAGAACGCGAAGCTGAGGGAGTTGGTGAACCTTGTGGTT
GTTGCCGGAGACAGGAGGAAGGAGTCGAAGGACTTGAAGAGAAGGCCGA
GATGAAGAAGATGTACGGCCTGATCGAGACCTACAAGTTGAACGGGCAATT
CAGATGGATTTTCATCTCAGATGAACCGTGTGAGGAACGGAGAGCTGTACCG
CGTGATCTGCGACACCAGGGGTGCTTTTCGTGCAGCCTGCTGTATACGAGGC
TTTTGGTTTGACAGTGGTTGAGGCCATGACTTGCGGCTTGCCAACATTCGC
CACATGCAATGGTGGTCCTGCTGAGATCATTGTGCACGGCAAGTCTGGCTT
CCACATTGACCCTTACCATGGTGACCGTGTGCTGCTGATCTCCTTGTGACTTC
TTTGAGAAGTGCAAGCTTGACCCAACTCACTGGGACAAGATCTCAAAGGCT
GGTCTCCAGCGTATTGAAGAGAAGTACACATGGCAAATTTACTCTCAGAGG
CTTCTCACTCTCACCGGTGTCTATGGCTTCTGGAAGCATGTGTCTAACCTTG
ACCGCCGTGAGAGCCGCCGCTATCTCGAGATGTTCTATGCTCTCAAGTACC
GCAAATTGGCTGAGTCTGTGCCCTTGCTGCTGAGTAACTGAGGATAAAG
AGTTGGATAAAGAAATGGAGGAACCGGCTTTTTCTTCTCATTGGAGTTTG
TCGCACTTGAGTTTTATAATAATGTCCGTGATTTTAGTTTTGTGATTAAGCT
TTCGATAAGAGGAGAGAAAGAGAAGGAAAAAAAAAAGTTGCTTTTTTTTTTGT
GTTGCATGATTTGGATCTTGATTGGAAAAGCTTCGAATTGGGGTAGTTTTAC
CCATCAATTCAATTTTAAGCCGTGCCTTCTTCAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAATCAAAAAAAAAAAAAAAAAAAAA

Figure 4B: Seq ID 11 - Nucleic acid sequence of the open reading frame of GmSUS-2025

ATGGCCACCGATCGTTTGACCCGGGTTACAGTCTCCGTGAGAGGCTTGAT
GAAACCCTCACTGCCAACAGGAATGAAATTTTGCCCTTCTGTCAAGGATC
GAAGCCAAGGGCAAGGGCATCCTGCAACACCACCAGGTCAATTGCTGAGTTT
GAGGAAATCCCTGAGGAGAACAGACAGAAGCTCACTGATGGTGCCTTTGGA
GAAGTCTTGAGATCTACACAGGAAGCCATAGTTTTGCCACCATGGGTGCT
CTGGCTGTTCTGTTCCAAGACCTGGTGTGTGGGAGTACCTGAGAGTGAATGTG
CACGCTCTTGTGTTGAGGAGTTGCAACCTGCTGAGTACCTGCACTTCAAG
GAAGAACTTGTTGATGGAAGTTCTAATGGCAACTTTGTGCTTGAGTTGGACT
TTGAACCATTCATGCAGCCTTCCCTCGCCCAACTCTTAACAAGTCAATTGG
AAATGGTGTGCAGTTCCTCAACCGCCACCTTTCTGCCAAACTCTTCCACGAC
AAGGAGAGCTTGACCCCACTTTTGGAATTCCTCAGGCTTCACAGCGTCAAG
GGAAAGACTTTGATGTTGAATGACAGAATTCAAACCCAGATGCACTCCAAC
ATGTTCTGAGGAAAGCTGAGGAGTATCTGGGCACAGTGCCTCCTGAAACTC
CCTACTCAGAATTTGAGCACAAGTTCCAGGAGATTGGTTTGGAGAGAGGGT
GGGGTGACAACGCAGAGCGTGTCTTGAGTCAATTCAACTTCTCTTGATCT
TCTTGAGGCCCTGACCCTTGACCCCTTGAGACTTTCCTTGGAAGAATTCCT
ATGGTCTTCAATGTTGTCATTCTTCTCCCATGGTTACTTTGCCCAAGATAA
TGTCTTGGGATACCCTGACACTGGTGGCCAGGTTGTTTACATCTTGGATCAA
GTTTCGTGCTTTGGAGAACGAGATGCTCCATCGCATTAAGCAACAAGGATTG

GACATTGTACCTCGTATTCTCATTATCACCCGTCTTCTCCCCGATGCAATCG
GAACTACTTGTGGCCAACGTCTTGAGAAGGTGTTGGAACCGAGCACTCCC
ACATTCTTCGAGTTCCCTTTAGAACTGAGAAGGGAATTGTTGTCAGTGGAT
CTCAAGATTGGAAGTCTGGCCATACTTGGAACTTACACTGAGGATGTTGCT
CATGAGCTTGCCAAAGAGTTGCAAGGCAAGCCAGATCTGATTGTCGGAAAC
TACAGTGATGGAAACATTGTCGCCTCTTTGTTGGCACATAAATTAGGAGTCA
CTCAGTGTACCATTGCTCACGCACTTGAGAAGACCAAATACCCCGAATCCG
ACATTTACTGGAAAAAATTGGAAGAGAGATACCACTTCTCTTGCCAATTCAC
AGCTGATCTATTTGCCATGAACCACACAGATTTTATTATCACCACTACCTTCC
AGGAGATTGCTGGAAGCAAGGACACTGTTGGACAGTACGAGTCTCACACAG
CCTTCACCTTCCCTGGACTCTACCGTGTTGTGCACGGCATTGATGTCTTTGA
TCCAAAATTCAACATTGTCTCCCCTGGAGCTGATCAAACCACTTACTTCCCC
CCCACCGAACTAGCCGTAGGTTGACCTCCTTCCACCCCGAAATCGAAGAA
CTTCTTTACAGCTCTGTGGAGAATGAAGAACACATATGCGTGCTGAAGGAC
CGCAGCAAGCCGATTATCTTCACCATGGCAAGGTTGGACCGTGTTGAAGAAC
ATCACAGGACTCGTGGAGTGGTACGGTAAGAACGCGAAGCTGAGGGAGTT
GGTGAACCTTGTGGTTGTTGCCGGAGACAGGAGGAAGGAGTGAAGGACT
TGGAAGAGAAGGCCGAGATGAAGAAGATGTACGGCCTGATCGAGACCTAC
AAGTTGAACGGGCAATTCAGATGGATTTTATCTCAGATGAACCGTGTTGAGG
AACGGAGAGCTGTACCGCGTGATCTGCGACACCAGGGGTGCTTTTCGTGCA
GCCTGCTGTATACGAGGCTTTTGGTTTGACAGTGGTTGAGGCCATGACTTG
CGGCTTGCCAACATTCGCCACATGCAATGGTGGTCCTGCTGAGATCATTGT
GCACGGCAAGTCTGGCTTCCACATTGACCCTTACCATGGTGACCGTGCTGC
TGATCTCCTTGTGACTTCTTTGAGAAGTGCAAGCTTGACCCAACTCACTGG
GACAAGATCTCAAAGGCTGGTCTCCAGCGTATTGAAGAGAAGTACACATGG
CAAATTTACTCTCAGAGGCTTCTCACTCTCACCGGTGTCTATGGCTTCTGGA
AGCATGTGTCTAACCTTGACCGCCGTGAGAGCCGCCGCTATCTCGAGATGT
TCTATGCTCTCAAGTACCGCAAATTGGCTGAGTCTGTGCCCTTGCTGCTGA
G

Figure 4C: Seq ID 12 - Amino acid sequence of the open reading frame of GmSUS-2025

MATDRLTRVHSLRERLDETLTANRNEILALLSRIEAKGKGILQHHQVIAEFEEIPE
ENRQKLTDGAFGEVLRSTQEAIVLPPWVALAVRPRPGWWEYLRVNVHALVVEE
LQPAEYLHFKEELVDGSSNGNFVLELDFEPFNAAFPRPTLNKSGNGVQFLNRH
LSAKLFHDKESLHPLLEFLRLHSVKGKTLMLNDRIQNPDALQHVLRKAEYYLGT
VPPETPYSEFEHKFQEIGLERGWGDNAERVLESIQLLLDLLEAPDPCTLETFLG
RIPMVFNVLSPHGYFAQDNVLGYPTDGGQVVYILDQVRALENEMLHRIKQQG
LDIVPRILIITRLLPDAIGTTGQRLEKVFGEHSHILRVFPRTEKGIVRQWISRFE
VWPYLETYTEDVAHELAKELQGKPDIVGNYS DGNIVASLLAHKLGVTQCTIAHA
LEKTKYPESDIYWKKLEERYHFSCQFTADLFAMNHTDFIITSTFQEIAGSKDTVG
QYESHTAFTLPGLYRVVHGIDVDFPKFNIVSPGADQTIYFPPTETSRRLTSFHPEI
EELLYSSVENEHICVLKDRSKPIIFTMARLDRVKNITGLVEWYGKNAKLRELVN
LVVVAGDRRKESKDLEEKAEKMYGLIETYKLNQFRWISSQMNRVRNGELY
RVICDTRGAFVQPAVYEAFGLTVVEAMTCGLPTFATCNGGPAEIVHGKSGFHI

DPYHGDRAADLLVDFFEKCKLDPTHWDKISKAGLQRIEEKYTWQIYSQRLTLT
GVYGFWKHVSNLDRRESRRYLEMFYALKYRKLAESVPLAAE

Figure 5A: Seq ID 13 – Nucleic acid sequence of GmSUS-4819

GCTGTCCGGTCTCTCTCTCTCTCTCTCCCCACCCCTTCCTTTTTTGGCGTTCATT
CTGTTTTTCATAGTGACGAACTTCTGAAGAAATGGCAAATCATCCTTTGACAC
ACTCTCACTCTTTCCGCGAGAGGTTTGATGAACTCTCACTGGTCACAGGAA
CGAAATTTTGGCCCTTCTGTCAAGGCTTGAAGCCAAGGGCAAGGGAATCCT
GCAACATCACCAAGTGGTTGCAGAGTTTGAAGAAATCCCTGAGGAGAGCAG
AAAGAACTCCAAGATGGTGTCTTTGGAGAAGTTTTGAGATCCACACAGGAA
GCCATAGTGCTGCCACCATTTGTAGCTCTTGCTGTTTCGACCAAGGCCTGGT
GTATGGGAATATCTGCGTGTGAATGTGCACATGCTTGTTGTTGATGAGCTTC
GTCCTGCTGAGTATCTGCGTTTCAAGGAGGAGCTTGTTGAGGGAAGTTCAA
ACGGCAACTTTGTGCTTGAGTTGGACTTTGAACCGTTTAATGCATCCTTCCC
TCGCCCAACTCTGAACAAGTCCATTGGAAATGGCGTCGAGTTCCTCAACCG
CCACCTTTTCGGCCAAGCTCTTCCATGACAAGGAGAGCATGCAGCCACTGCT
TGAGTTCCTCAGGCTTCACAGTTATAAGGGAAAGACCATGATGTTGAATGAC
AAAGTTCAAAGCCTGGATTCTCTCCAACATGTTTTGAGAAAAGCAGAAGAGT
ATCTGATTTTCAGTTGCTCCTGAAACACCCTACTCGGAATTCGAAAACAGATT
CCGGGAGATTGGTCTGGAGAGGGGGTGGGGTGACACTGCCGAGCGTGTC
CTCGAGATGATCCAGCTTCTCCTGGACCTTCTTGAAGCGCCTGACCCTTGC
ACCCTCGAGACATTCTTGGAAAGAGTCCCTATGGTCTTCAATGTTGTTATCC
TTTCTCCCCATGGTTACTTTGCCCAAGATAATGTCTTGGGATACCCTGACAC
TGGTGGACAGGTTGTTTACATCTTGGATCAAGTTCGTGCCTTGGAGAATGAG
ATGCTCAACCGCATCAAGAAACAAGGCCTTGATATCACCCCTCGTATTCTCA
TTATCACTCGTCTTCTCCCTGATGCAGTAGGAACTACCTGTGGCCAACGTCT
AGAGAGGGTATATGATACTGAATATTGTGACATTCTCCGAGTTCCTTTCAGA
ACCGAAAAGGGAATTGTTTCGCAAATGGATCTCAAGATTCTGAAGTCTGGCCA
TACCTAGAGACTTACACTGAGGATGTTGCCCTTGAAGTTGCCAAGGAGTTG
CAAGCCAAGCCAGATCTGATTGTTGGAACTACAGTGATGGAACATTGTTG
CCTCTTTGTTAGCACATAAATTAGGAGTAACTCAGTGTACCATTGCTCATGCT
CTAGAAAAGACCAAGTACCCTGAGTCTGACATTTACTGGAAAAAATTTGAAG
AGAAATATCATTTCTCATGCCAATTTACTGCTGATCTTTTTGCAATGAACCAC
ACAGACTTTATCATCACAGCACCTTCCAAGAGATTGCTGGAAGCAAGGAC
ACTGTTGGACAGTATGAGAGTCACACTGCCTTCACCCTTCCAGGACTTTACC
GTGTTGTTACGGTATTGATCCCTTTGATCCAAAGTTCAACATTGTCTCTCCC
GGTGCAGACATGGGCATATACTTCCCATACTGAACTGAGCGTAGGTTA
ACAGAATTCCTCTGACATTGAAGAGCTTCTTTACAGCTCAGTGGAGAATG
AGGAACACATATGCGTATTGAAGGACCGCAACAAACCAATAATCTTCACCAT
GGCAAGGCTTGACCGTGTGAAGAACATCACGGGGCTTGTGAGTGGTACG
GGAAGAACGCACGCCTCCGCGAGTTGGTGAACCTGGTGGTGGTGGCTGGA
GACAGGAGGAAGGAGTCGAAGGACTTGGAAGAGAAGGCCGAGATGAAGAA
GATGTATGGCCTCATCGAGACCTACAAGTTGAACGGCCAATTCAGATGGAT
ATCCTCTCAGATGAACCGTGTGAGGAACGGAGAGCTTACCGTGTCTCTG
TGACACAAGGGGTGCCTTTGTGCAGCCTGCAGTTTATGAGGCCTTTGGGTT

GA CTGTGGTTGAGGCCATGACTTGTGGGTTGCCAACGTTTGCCACATGCAA
TGGTGGTCCTGCTGAGATCATTGTGCATGGAAAATCTGGTTACCACATTGAT
CCTTACCATGGTGACCATGCTGCTGAGATCCTTGTTGAGTTCTTTGAGAAGA
GCAAGGCTGATCCATCTCACTGGGACAAAATCTCCCAGGGTGGACTCAAGC
GTATTCATGAGAAGTACACATGGCAAATTTACTCGGACAGGCTCTTGACACT
CACTGGTGTGTATGGCTTCTGGAAGCACGTGACCAATCTTGAACGCCGTGA
GAGCAAACGTTACCTCGAGATGTTCTATGCTCTCAAGTACCGCAAATTGGCT
GAGTCTGTGCCCTTGCTATTGAAGAGTAAATTCATGTTTGAAGAGAACATC
AATGGAGAAACCGGCTTTTGGTCGTTTGAAGTCTTATGGAGCTTTCATAAAT
AACGCCATTGATTTTGATTGTGATCAGCTTTTGGATTTAAAGAGTCTAAAAAA
AAGTGTGCTTTGTTTTTATATTTTGCTTAGGTTTTGGGTAAAAAAAGTAAAAA
AAGAAAGTCTTTTTTTTTTTTTTAAATTCTTTTGTTGTTGCAATATTGTGTATG
ATTGGAAATCTTTAAATTCGGGTTTTCCACCCCATCAATTCAATTAAAGATA
TGCCTTCTTCCATTCAATTCGGTAAATACTGGTCCAGAGAAGAGAAAAATG
GCAATCTCCATACCCAGTAGACAACCTTTTCATAGAAAACCTCGTGCCGCTTCT
CTTGTTCTTCAAGAATCCACACCAACCCCTCACAATTGGTACCTAACTT
CTGTTCCAATTTTATTTATTGGGTTGGAGTGTTCTTTGTGTTGATTTCAATT
GCCTGTTGCTTGTGCCTTCACAGGACCAAAGGAATTTGCAGCACTCTGGAN
TTGCCTTGGTCATTGGCGANGAATATTTGAGAGGTACGATAAANACAGGAG
TGGGAAAATTGATCCATTANAACTCANAGATGCTCTGTATGGTATTGGCTAT
GCTGTACCANGCTCNGTTCTNNNNNTACTGCTTTCAAAGTATGGTGATGGAA
NTGGTAGGANGGTTGAACTTGGTTTTGACAG

Figure 5B: Seq ID 14 – Nucleic acid sequence of the open reading frame of GmSUS-4819

ATGGCAAATCATCCTTTGACACACTCTCACTCTTTCCGCGAGAGGTTTGATG
AAACTCTCACTGGTCACAGGAACGAAATTTTGGCCCTTCTGTCAAGGCTTGA
AGCCAAGGGCAAGGGAATCCTGCAACATCACCAAGTGGTTGCAGAGTTTGA
AGAAATCCCTGAGGAGAGCAGAAAGAACTCCAAGATGGTGTCTTTGGAGA
AGTTTTGAGATCCACACAGGAAGCCATAGTGCTGCCACCATTGTAGCTCTT
GCTGTTGACCAAGGCCTGGTGTATGGGAATATCTGCGTGTGAATGTGCAC
ATGCTTGTTGTTGATGAGCTTCGTCCTGCTGAGTATCTGCGTTTCAAGGAGG
AGCTTGTTGAGGGAAGTTCAAACGGCAACTTTGTGCTTGAGTTGGACTTTGA
ACCGTTTAATGCATCCTTCCCTCGCCCAACTCTGAACAAGTCCATTGGAAAT
GGCGTCGAGTTCCTCAACCGCCACCTTTCGGCCAAGCTCTTCCATGACAAG
GAGAGCATGCAGCCACTGCTTGAGTTCCTCAGGCTTCACAGTTATAAGGGA
AAGACCATGATGTTGAATGACAAAGTTCAAAGCCTGGATTCTCTCCAACATG
TTTTGAGAAAAGCAGAAGAGTATCTGATTTCAAGTTGCTCCTGAAACACCCTA
CTCGGAATTCGAAAACAGATTCCGGGAGATTGGTCTGGAGAGGGGGTGGG
GTGACACTGCCGAGCGTGTCTCGAGATGATCCAGCTTCTCCTGGACCTTC
TTGAAGCGCCTGACCCTTGCAACCCTCGAGACATTCTTGGAAAGAGTCCCTA
TGGTCTTCAATGTTGTTATCCTTTCTCCCATGGTTACTTTGCCCAAGATAAT
GTCTTGGGATACCCTGACACTGGTGGACAGGTTGTTTACATCTTGGATCAAG
TTCGTGCCTTGGAGAATGAGATGCTCAACCGCATCAAGAAACAAGGCCTTG
ATATCACCCCTCGTATTCTCATTATCACTCGTCTTCTCCCTGATGCAGTAGG

AACTACCTGTGGCCAACGTCTAGAGAGGGTATATGATACTGAATATTGTGAC
ATTCTCCGAGTTCCTTTCAGAACCGAAAAGGGAATTGTTGCGCAAATGGATCT
CAAGATTGGAAGTCTGGCCATACCTAGAGACTTACACTGAGGATGTTGCCC
TTGAACTTGCCAAGGAGTTGCAAGCCAAGCCAGATCTGATTGTTGGAACTA
CAGTGATGGAAACATTGTTGCCTCTTTGTTAGCACATAAATTAGGAGTAACT
CAGTGTACCATTGCTCATGCTCTAGAAAAGACCAAGTACCCTGAGTCTGACA
TTTACTGGAAAAAATTTGAAGAGAAATATCATTCTCATGCCAATTTACTGCT
GATCTTTTTTGCAATGAACCACACAGACTTTATCATCACCAGCACCTTCCAAG
AGATTGCTGGAAGCAAGGACACTGTTGGACAGTATGAGAGTCACACTGCCT
TCACCCTTCCAGGACTTTACCGTGTTGTTACGGTATTGATCCCTTTGATCC
AAAGTTCAACATTGTCTCTCCCGGTGCAGACATGGGCATATACTTCCCATAC
ACTGAAACTGAGCGTAGGTTAACAGAATTCCACTCTGACATTGAAGAGCTTC
TTTACAGCTCAGTGGAGAATGAGGAACACATATGCGTATTGAAGGACCGCA
ACAAACCAATAATCTTCACCATGGCAAGGCTTGACCGTGTGAAGAACATCAC
GGGGCTTGTGAGTGGTACGGGAAGAACGCACGCCTCCGCGAGTTGGTGA
ACCTGGTGGTGGTGGCTGGAGACAGGAGGAAGGAGTCGAAGGACTTGGAA
GAGAAGGCCGAGATGAAGAAGATGTATGGCCTCATCGAGACCTACAAGTTG
AACGGCCAATTCAGATGGATATCCTCTCAGATGAACCGTGTGAGGAACGGA
GAGCTCTACCGTGTCTGTGACACAAGGGGTGCCTTTGTGCAGCCTGCA
GTTTATGAGGCCTTTGGGTTGACTGTGGTTGAGGCCATGACTTGTGGGTTG
CCAACGTTTGCCACATGCAATGGTGGTCCTGCTGAGATCATTGTGCATGGA
AAATCTGGTTACCACATTGATCCTTACCATGGTGACCATGCTGCTGAGATCC
TTGTTGAGTTCTTTGAGAAGAGCAAGGCTGATCCATCTCACTGGGACAAAAT
CTCCCAGGGTGGACTCAAGCGTATTCATGAGAAGTACACATGGCAAATTTA
CTCGGACAGGCTCTTGACACTCACTGGTGTGTATGGCTTCTGGAAGCACGT
GACCAATCTTGAACGCCGTGAGAGCAAACGTTACCTCGAGATGTTCTATGC
TCTCAAGTACCGCAAATTGGCTGAGTCTGTGCCCTTGCTATTGAAGAG

Figure 5C: Seq ID 15 - Amino acid sequence of the open reading frame of GmSUS-4818

MANHPLTHSHSFRERFDETLTGHRNEILALLSRLEAKGKGILQHHQVVAEFEEIP
EESRKKLQDGVFGEVLRSTQEAIVLPPFVALAVRPRPGVWEYLRVNVHMLVVD
ELRPAEYLRKFKEELVEGSSNGNFVLELDFEPFNASFPRPTLNKSIGNGVEFLNR
HLSAKLFHDKESMQPLLEFLRLHSYKGKTMMLNDKVQSLDSLQHVLRKAEELYI
SVAPETPYSEFENRFREIGLERGWGDTAERVLEMIQLLLDLLEAPDPCTLETFL
GRVPMVFNVLSPHGYFAQDNVLGY PDTGGQVVYILDQVRALENEMLNRIKK
QGLDITPRILIITRLLPDAVGTTGQRLERVYDTEYCDILRVFPRTEKGIVRKWISR
FEVWPYLETYTEDVALELAKELQAKPDLIVGNYS DGNIVASLLAHKLGVTQCTIA
HALEKTKYPESDIYWKKFEEKYHFSCQFTADLFAMNHTDFIITSTFQEIAGSKDT
VGQYESHTAFTLPGLYRVVHGIDPFDPKFNIVSPGADMGIYFPYTETERRLTEF
HSDIEELLYSSVENEHICVLKDRNKPIIFTMARLDRVKNITGLVEWYGKNARLR
ELVNLVVVAGDRRKESKDLEEKAEEMKKMYGLIETYKLNQQRWISSQMNRVRN
GELYRVICDTRGAFVQPAVYEAFGLTVVEAMTCGLPTFATCNGGPAEIVHGKS
GYHIDPYHGDHAAEILVEFFEKSKADPSHWDKISQGGGLKRIHEKYTWQIYSDRL
LTLTG VYGFWKHVTNLERRESKRYLEM FYALKYRKLAESVPLAIEE

Figure 6A: Seq ID 16 – Nucleic acid sequence of ZmSUS-7691

CTCCATTGGACTGCTTGCTCCCTGTTGACCATTGGGTATTCTGAACCATCGA
GCCATGGCTGCCAAGCTGACTCGCCTCCACAGTCTTCGCGAACGCCTTGGT
GCCACCTTCTCCTCCCATCCCAATGAACTGATAGCACTCTTTTCCAGGTATG
TTCACCAGGGCAAGGGAATGCTTCAGCGCCATCAGCTGCTTGCGGAGTTTG
ATGCCCTGTTTGATAGTGACAAGGAGAAGTATGCACCCTTTGAAGACATTCT
TCGTGCTGCTCAGGAAGCAATTGTGCTCCCCCATGGGTTGCACTTGCTAT
CAGGCCAAGGCCTGGTGTCTGGGATTACATTCGGGTGAATGTAAGTGAGCT
GGCTGTGGAGGAGCTGAGTGTTTCTGAGTACTTGGCATTCAAGGAACAGCT
GGTGGATGGACAATCCAACAGCAACTTTGTGCTTGAGCTTGATTTTGAGCCC
TTCAATGCCTCCTTTCCTCGTCCTTCCATGTGCGAAGTCCATCGGAAATGGAG
TGCAATTCCTTAACCGACACCTGTCGTCCAAGTTGTTCCAGGACAAGGAGA
GTTTGTACCCCTTGCTGAACTTCCTCAAGGCTCATAACTACAAGGGCACGAC
GATGATGTTGAATGACAGAATCCAAAGCCTTCGTGGTCTCCAATCATCCCTG
AGAAAGGCAGAGGAGTATCTACTGAGTGTTCTCAAGACACTCCCTACTCG
GAGTTCAACCATAGGTTCCAAGAGCTTGGTTTGAGAGAAGGGTTGGGGTGAC
ACTGCGAAGCGTGTACTCGACACACTCCACTTGCTTCTTGACCTTCTTGAGG
CCCCTGATCCTGCCAACTTGGAGAAGTTCCTTGGAATAACCAATGATGTT
CAATGTTGTTATCCTTTCTCCTCATGGCTACTTCGCTCAGTCCAATGTGCTTG
GATACCCTGACACTGGCGGTCAGGTTGTGTACATTCTGGATCAAGTCCGTG
CTTTGGAGAATGAGATGCTTCTGAGGATTAAGCAGCAAGGCCTTGATATCAC
TCCGAAGATCCTCATTGTTACCAGGCTGTTGCCTGATGCTGCTGGGACTAC
GTGCGGTCAGCGGCTGGAGAAGGTCATTGGTACTGAGCACACAGACATCAT
TCGCGTTCCGTTTCAGAAATGAGAATGGCATCCTCCGCAAGTGGATCTCTCG
TTTTGATGTCTGGCCATACCTGGAGACATACACTGAGGATGTTTCCAGTGAA
ATAATGAAAGAAATGCAGGCCAAGCCTGACCTTATCATTGGCAACTACAGC
GATGGCAACCTAGTCGCCACTCTGCTCGCGCACAAGTTGGGAGTCACTCAG
TGTACCATCGCTCATGCCTTGGAGAAAACCAAATACCCCAACTCGGACATCT
ACTTGGACAAGTTCGACAGCCAGTACCACTTCTCTTGCCAGTTCACAGCTGA
CCTTATTGCCATGAACCACACCGATTTCATCATCACCAGCACATTCCAAGAA
ATCGCGGGAAGCAAGGACACCGTGGGGCAGTACGAGTCCACATCGCGTT
CACTCTTCTGGGCTCTACCGTGTGTCGTCATGGCATCGATGTTTTCGATCCC
AAGTTCAACATTGTCTCCCCTGGAGCAGACATGAGTGTTTACTACCCGTATA
CGGAAACCGACAAGAGACTCACTGCCTTCCATCCTGAAATCGAGGAGCTCA
TCTACAGCGACGTGAGAACTCCGAGCACAAGTTCGTGCTGAAGGACAAGA
AGAAGCCGATCATCTTCTCGATGGCGCGTCTCGACCGCGTGAAGAACATGA
CAGGCCTGGTTCGAGATGTACGGCAAGAACGCGCGCCTGAGGGAGCTGGC
GAACCTCGTGATCGTTGCCGGCGACCAAGGAGTCCAAGGACAGGG
AGGAGCAGGCGGAGTTCAAGAAGATGTACAGCCTCATCGACGAGTACAAGT
TGAAGGGCCATATCCGGTGGATCTCGGCGCAGATGAACCGCGTCCGCAAC
GGGGAGCTGTACCGCTACATTTGCGATACGAAGGGCGCATTCTGTGAGCC
TGCGTTCTACGAAGCGTTCCGGCCTGACTGTGATCGAGTCCATGACGTGCGG
TCTGCCAACGATCGCGACCTGCCATGGTGGCCCTGCTGAGATCATCGTGGA
CGGGGTATCTGGCCTGCACATTGACCCTTACCACAGCGACAAGGCCGCGG
ATATCCTGGTCAACTTCTTTGACAAATGCAAGGCAGATCCGAGCTACTGGGA

CAAGATCTCACAGGGGCGGCCTGCAGAGAATTTATGAGAAGTACACCTGGAA
GCTCTACTCCGAGAGGCTGATGACCCTGACCGGCGTGTACGGGTTCTGGA
AGTACGTGAGCAACCTGGAGAGGCGCGAGACCCGCCGCTACATCGAGATG
TTCTACGCCCTGAAGTACCGTAGCCTGGCAAGCCAGGTTCCGCTGTCCTTC
GATTAGTACGGGGAAAGAAGGAGAAGAAGAAGCCCAGGCCGGAGAACCAT
CGCCTGCATTTGATCTGTTTCACCGCAATTCGCATTGTTAGTCGTGTATTG
GAGTTATGTGTACTTGGTTTCCAAGAACTTTGGTTCCTTCTCGTTTTTTTTCC
TTGTTTGAGCGTTTTTGGGCAGCGCTGGCCTGGTTCCTATTCTATGGTGG
GAATTGGCTGCACCTTTTGCTTCGAATAAAAATGCCTGCTCGTTCTCTGTG
AAAAAAAAAAAAAAAAAAAAAAAAAACTCAAGGGGGGGGCCCGGTACC

Figure 6B: Seq ID 17 – Nucleic acid sequence of the open reading frame of ZmSUS-7691

ATGGCTGCCAAGCTGACTCGCCTCCACAGTCTTCGCGAACGCCTTGGTGCC
ACCTTCTCCTCCCATCCCAATGAACTGATAGCACTCTTTTCCAGGTATGTTT
ACCAGGGCAAGGGAATGCTTCAGCGCCATCAGCTGCTTGCGGAGTTTGATG
CCCTGTTTGATAGTGACAAGGAGAAGTATGCACCCTTTGAAGACATTCTTCG
TGCTGCTCAGGAAGCAATTGTGCTCCCCCATGGGTTGCACTTGCTATCAG
GCCAAGGCCTGGTGTCTGGGATTACATTCGGGTGAATGTAAGTGAGCTGGC
TGTGGAGGAGCTGAGTGTTTCTGAGTACTTGGCATTCAAGGAACAGCTGGT
GGATGGACAATCCAACAGCAACTTTGTGCTTGAGCTTGATTTTGAGCCCTTC
AATGCCTCCTTTCTCGTCCTTCCATGTCGAAGTCCATCGGAAATGGAGTGC
AATTCCTTAACCGACACCTGTGCTCCAAGTTGTTCCAGGACAAGGAGAGTTT
GTACCCCTTGCTGAACTTCCTCAAGGCTCATAACTACAAGGGCACGACGAT
GATGTTGAATGACAGAATCCAAAGCCTTCGTGGTCTCCAATCATCCCTGAGA
AAGGCAGAGGAGTATCTACTGAGTGTTCTCAAGACACTCCCTACTCGGAG
TTCAACCATAGGTTCCAAGAGCTTGGTTTGGAGAAGGGTTGGGGTGACACT
GCGAAGCGTGTA CTGACACACTCCACTTGCTTCTTGACCTTCTTGAGGCC
CCTGATCCTGCCAACTTGGAGAAGTTCTTGGA ACTATACCAATGATGTTCA
ATGTTGTTATCCTTTCTCCTCATGGCTACTTCGCTCAGTCCAATGTGCTTGG
ATACCTGACACTGGCGGTCAGGTTGTGTACATTCTGGATCAAGTCCGTGC
TTTGGAGAATGAGATGCTTCTGAGGATTAAGCAGCAAGGCCTTGATATCACT
CCGAAGATCCTCATTGTTACCAGGCTGTTGCCTGATGCTGCTGGGACTACG
TGCGGTCAGCGGCTGGAGAAGGTCATTGGTACTGAGCACACAGACATCATT
CGCGTTCCGTT CAGAAATGAGAATGGCATCCTCCGCAAGTGGATCTCTCGT
TTTGATGTCTGGCCATACCTGGAGACATACTGAGGATGTTTCCAGTGAAA
TAATGAAAGAAATGCAGGCCAAGCCTGACCTTATCATTGGCAACTACAGCG
ATGGCAACCTAGTCGCCACTCTGCTCGCGCACAAGTTGGGAGTCACTCAGT
GTACCATCGCTCATGCCTTGGAGAAAACCAAATACCCCAACTCGGACATCTA
CTTGGACAAGTTGACAGCCAGTACCACTTCTCTTGCCAGTTACAGCTGAC
CTTATTGCCATGAACCACACCGATTTCATCATCACCAGCACATTCCAAGAAA
TCGCGGGAAGCAAGGACACCGTGGGGCAGTACGAGTCCCACATCGCGTTC
ACTCTTCCTGGGCTCTACCGTGTGCTCCATGGCATCGATGTTTTCGATCCCA
AGTTCAACATTGTCTCCCCTGGAGCAGACATGAGTGTTTACTACCCGTATAC
GGAAACCGACAAGAGACTCACTGCCTTCCATCCTGAAATCGAGGAGCTCAT

CTACAGCGACGTCGAGAACTCCGAGCACAAGTTCGTGCTGAAGGACAAGAA
GAAGCCGATCATCTTCTCGATGGCGCGTCTCGACCGCGTGAAGAACATGAC
AGGCCTGGTCGAGATGTACGGCAAGAACGCGCGCCTGAGGGAGCTGGCG
AACCTCGTGATCGTTGCCGGCGACACGGCAAGGAGTCCAAGGACAGGGA
GGAGCAGGCGGAGTTCAAGAAGATGTACAGCCTCATCGACGAGTACAAGTT
GAAGGGCCATATCCGGTGGATCTCGGCGCAGATGAACCGCGTCCGCAACG
GGGAGCTGTACCGCTACATTTGCGATACGAAGGGCGCATTCTGTGCAGCCT
GCGTTCTACGAAGCGTTCGGCCTGACTGTGATCGAGTCCATGACGTGCGGT
CTGCCAACGATCGCGACCTGCCATGGTGGCCCTGCTGAGATCATCGTGGA
CGGGGTATCTGGCCTGCACATTGACCCTTACCACAGCGACAAGGCCGCGG
ATATCCTGGTCAACTTCTTTGACAAATGCAAGGCAGATCCGAGCTACTGGGA
CAAGATCTCACAGGGCGGCCTGCAGAGAATTTATGAGAAGTACACCTGGAA
GCTCTACTCCGAGAGGCTGATGACCCTGACCGGCGTGTACGGGTTCTGGA
AGTACGTGAGCAACCTGGAGAGGCGCGAGACCCGCCGCTACATCGAGATG
TTCTACGCCCTGAAGTACCGTAGCCTGGCAAGCCAGGTTCCGCTGTCCTTC
GAT

Figure 6C: Seq ID 18 - Amino acid sequence of the open reading frame of ZmSUS-7691

MAAKLTRLHSLRERLGATFSSHPNELIALFSRYVHQKGMLQRHQLLAEFDALE
DSDKEYAPFEDILRAAQEAIVLPPWVALAIRPRPGVWDYIRVNVSELAVEELSV
SEYLAFKEQLVDGQSNFVLELDFEPFNASFPRPSMSKSIENGNGVQFLNRHLS
SKLFQDKESLYPLLNFLKAHNYKGTTMMLNDRIQSLRGLQSSLRKAEEYLLSVP
QDTPYSEFNHRFQELGLEKGWGD TAKRVLDTLHLLDLLEAPDPANLEKFLGTI
PMMFNVVILSPHGYFAQSNVLGYPD TGGQVVYILDQVRALENEMLLRIKQQGL
DITPKILIVTRLLPDAAGTTCGQRLEK VIGTEHTDIIRVPFRNENGILRKWISRFDV
WPLYETYTEDVSSEIMKEMQAKPD LIIGNYSDGNLVATLLAHKLGVTQCTIAHAL
EKT KYPNSDIYLDKFD SQYHFSCQFTADLIAMNHTDFIITSTFQEIAGSKDTVGQ
YESHIAFTLPGLYRVVHGIDVFD PKFNIVSPGADMSVYYPYTETDKRLTAFHPEI
EELIYSDVENSEHKFVLKDKKKPIIF SMARLDRVKNMTGLVEMYGKNARLRELA
NLVIVAGDHGKESKDREEQAEFKK MYSLIDEYKLKGHIRWISAQMNRVRNGELY
RYICDTKGAFVQPAFYEA FGLTVIESMTCGLPTIATCHGGPAEIIVDGV SGLHIDP
YHSDKAADILVNFFDKCKADPSYWDKISQ GGLQRIYEKYTWKLYSERLMTLTGV
YGFWKYVSNLERRETRRYIEMFYALKYRSLASQVPLSFD

Figure 7A: Seq ID 19 – Nucleic acid sequence of ZmSUS-8057

ACTCCACCTCCGTTACACCCGTCCATTTGATTTGCGTTCACTGCGTTGCGTT
TCCTTGAGGGGATTGTTCTCTCCTCTCCTTTGGATTGGAGGTCCCTCCTTC
TTCTCCTCTCTCTCAGAGGAAGGCCTGAGGATCCAGGAAGAGGACAGCA
ATGGGGGAAGGTGCAGGTGACCGTGTCTGAGCCGCCTCCACAGCGTCAG
GGAGCGCATTGGCGACTCACTCTCTGCCACCCCAATGAGCTTGTGCGCGT
CTTACCAGGCTGAAAAACCTTGAAAGGGTATGCTGCAGCCCCACCAGAT
CATTGCCGAGTACAACAATGCGATCCCTGAGGCTGAGCGCGAGAAGCTCAA
GGATGGTGCTTTTGAGGATGTCCTGAGGGCAGCTCAGGAGGCGATTGTCAT

CCCCCATGGGTTGCACTTGCCATCCGCCCTAGGCCTGGTGTCTGGGAGTA
TGTGAGGGTCAACGTCAGTGAGCTCGCTGTTGAGGAGCTGAGAGTTCCTGA
GTACCTGCAGTTCAAGGAACAGCTTGTGGAAGAAGGCCCAACAACAATT
TGTTCTTGAGCTGGACTTTGAGCCATTCAATGCCTCCTTCCCCCGTCCTTCT
CTGTCAAAGTCCATTGGCAATGGCGTGCAGTTCCTCAACAGGCACCTGTCA
TCAAAGCTCTTCCATGACAAGGAGAGCATGTACCCCTTGCTCAACTTCCTTC
GCGCCCACTACAAGGGGATGACCATGATGTTGAACGACAGAATCCGCA
GTCTCAGTGCTCTGCAAGGTGCGCTGAGGAAGGCTGAGGAGCACCTGTCC
ACCCTACAAGCTGATACCCCATACTCTGAATTTACCACAGGTTCCAGGAAC
TTGGTCTGGAGAAGGGTTGGGGTGATTGCGCTAAGCGTGCACAGGAGACT
ATCCACCTCCTCTTGACCTCCTGGAGGCCCCAGATCCGTCCACCCTGGAG
AAGTTCCTTGGAACGATCCCATGGTGTTCATGTCTGTTATCCTCTCCCTC
ATGGTTACTTCGCTCAAGCTAATGTCTTGGGTTACCCTGACACCGGAGGCC
AGGTTGTCTACATCTTGGATCAAGTGCGCGCTATGGAGAACGAAATGCTGC
TGAGGATCAAGCAGTGTGGTCTTGACATCACGCCGAAGATCCTTATTGTCA
CCAGGTTGCTCCCTGATGCAACTGGCACCACCTGTGGCCAGCGCCTTGAGA
AGGTCCTTGGCACCAGCACTGCCATATCCTTCGCGTGCCATTCAGAACAG
AAAACGGAATCGTTCGCAAGTGATCTCGCGATTTGAAGTCTGGCCGTACC
TGAGACTTAACTGATGACGTGGCGCATGAGATTGCTGGAGAGCTTCAGG
CCAATCCTGACCTGATCATCGGAACTACAGTGACGGAAACCTTGTTGCGT
GTTTGCTCGCCCAAGATGGGTGTTACTCACTGTACCATTGCCCATGCGC
TTGAGAAAATAAGTACCCTAACTCCGACCTCTACTGGAAGAAGTTTGAGGA
TCACTACCACTTCTCGTGCCAGTTCACCACTGACTTGATTGCAATGAACCAT
GCCGACTTCATCATCACCAGTACCTTCCAAGAGATCGCCGGAACAAGGAC
ACCGTCGGCCAGTACGAGTCACACATGGCGTTCACAATGCCTGGCCTGTAC
CGCGTTGTCCACGGCATTGATGTGTTGACCCCAAGTTCAACATCGTGTCT
CCTGGCGCGGACCTGTCCATCTACTTCCCGTACACCGAGTCGCACAAGAG
GCTGACCTCCCTTACCCGGAGATTGAGGAGCTCCTGTACAGCCAAACCGA
GAACACGGAGCACAAGTTCGTTCTGAACGACAGGAACAAGCCAATCATCTT
CTCCATGGCTCGTCTCGACCGTGTGAAGAACTTGACTGGGCTGGTGGAGCT
GTACGGCCGGAACAAGCGGCTGCAGGAGCTGGTGAACCTCGTGGTCTGTCT
GCGGCGACCATGGCAACCTTCCAAGGACAAGGAGGAGCAGGCCGAGTTC
AAGAAGATGTTTGACCTCATCGAGCAGTACAACCTGAACGGGCACATCCGC
TGATCTCCGCCAGATGAACCGCGTCCGCAACGGCGAGCTGTACCGCTA
CATCTGCGACACCAAGGGCGCCTTCGTGCAGCCTGCTTTCTACGAGGCTTT
CGGGCTGACGGTGGTTGAGGCCATGACCTGCGGCCTGCCACGTTTGCCA
CAGCCTACGGCGGTCCGGCCGAGATCATCGTGCACGGCGTGTCTGGCTAC
CACATCGACCCTTACCAGGGCGACAAGGCGTCGGCCCTGCTCGTGGACTT
CTTCGACAAGTGCCAGGCGGAGCCGAGCCACTGGAGCAAGATCTCCCAGG
GCGGGCTCCAGCGTATCGAGGAGAAGTACACCTGGAAGCTGTACTCGGAG
AGGCTGATGACCCTCACCGGCGTGTACGGGTTCTGGAAGTACGTGTCCAAC
CTGGAGAGGCGCGAGACCCGGCGGTACCTGGAGATGCTGTACGCGCTCAA
GTACCGCACCATGGCGAGCACCGTGCCGCTGGCCGTGGAGGGAGAGCCC
TCCAGCAAGTGATGCGCGACGGCGGCCACAGACCTGATCGATCGATGAGC
GAGAGGGAGCACTCGGAGTGTCTGTCTTTCCCTTGCCATTTCTTTCTTTC
TTTTTTTTCTTCCCGGAGGCGAAAAAAAGAGTCTGCTTTTGCTAGGCGGC

GGGCGTTCGTTGCTGCTCTTTGCTTCAAGAGTTAAATTTACCTACCTTGTC
AGGTCTTGTTCCATCATTGATCCGGGTGTCGCTTTTAGTAGTCTGATGGACT
GTTAGTAGTTTTCGTTGCGTTCGGTTGAGAGGGAACGGTGGTGGTGGTGGT
GTGTGTGCAGTCGGGTGTGGTGTCTCCCTTTGTTTCCTGGATGGGATGTTGC
TCCTTGAATAATAATCGTAGTGGCCTTGGAGCCCTTTTCCTGAAATAAGAGC
AGCATCCTAGTGCTTACCTTTGCAGCTGTAACCTGGTGTGTTTGCCTCAGGATTT
TTGCTCCTCCCTTCTATGGATTGTGCTTAAAATTCTCTATAAAAAAAAAAAAA
AAAAAACTCCAGGGGGGGGCCCGGGCCCCCATTCGCCCTATAGTGATTTTAA
TCACAGCNCCCGCGGGGGGGGCCCGGCCCTCCCGTCCCTTTTGGGGCGAG
AATTCGGCCCCGCCCTTTCCCGCCTCACTCCCCCGACCACCGTACNTCCGCG
CGCACGCCCCAGACCGCGCAGAGCCCCGCGCCCCCTGGTATTCCCGCGCACA
CCGCCCCCCCCCGCCCCCACCGTCCCCATAGTGTTATTCCACTCTCCTGT
ATCCCGCCCTTAGCCCCCTCCCCGCGCCCCCAATTTGACCCACCATCTCCT
CCCCCAAGATACGCCTTCACGACCAGATTTACCCGCGGGCCGCTCCCCC
GCACACACAAAAAAAAA

Figure 7B: Seq ID 20 – Nucleic acid sequence of the open reading frame of ZmSUS-8057

ATGGGGGAAGGTGCAGGTGACCGTGTCCTGAGCCGCCTCCACAGCGTCAG
GGAGCGCATTGGCGACTCACTCTCTGCCACCCCAATGAGCTTGTCGCCGT
CTTCACCAGGCTGAAAAACCTTGGAAGGGTATGCTGCAGCCCCACCAGAT
CATTGCCGAGTACAACAATGCGATCCCTGAGGCTGAGCGCGAGAAGCTCAA
GGATGGTGTCTTTGAGGATGTCCTGAGGGCAGCTCAGGAGGCGATTGTCAT
CCCCCATGGGTTGCACTTGCCATCCGCCCTAGGCCTGGTGTCTGGGAGTA
TGTGAGGGTCAACGTCAGTGAGCTCGCTGTTGAGGAGCTGAGAGTTCCTGA
GTACCTGCAGTTCAAGGAACAGCTTGTGGAAGAAGGCCCAACAACAATT
TGTTCTTGAGCTGGACTTTGAGCCATTCAATGCCTCCTTCCCCCGTCCTTCT
CTGTCAAAGTCCATTGGCAATGGCGTGCAAGTTCCTCAACAGGCACCTGTCA
TCAAAGCTCTTCCATGACAAGGAGAGCATGTACCCCTTGCTCAACTTCCTTC
GCGCCCACAACACTACAAGGGGATGACCATGATGTTGAACGACAGAATCCGCA
GTCTCAGTGCTCTGCAAGGTGCGCTGAGGAAGGCTGAGGAGCACCTGTCC
ACCCTACAAGCTGATACCCATACTCTGAATTTACCACAGGTTCCAGGAAC
TTGGTCTGGAGAAGGGTTGGGGTGATTGCGCTAAGCGTGACAGGAGACT
ATCCACCTCCTCTTGACCTCCTGGAGGCCCCAGATCCGTCCACCCTGGAG
AAGTTCCTTGGAACGATCCCCATGGTGTTCATGTGCTTATCCTCTCCCCTC
ATGGTTACTTCGCTCAAGCTAATGTCTTGGGTACCCTGACACCGGAGGCC
AGGTTGTCTACATCTTGGATCAAGTGCGCGCTATGGAGAACGAAATGCTGC
TGAGGATCAAGCAGTGTGGTCTTGACATCACGCCGAAGATCCTTATTGTCA
CCAGGTTGCTCCCTGATGCAACTGGCACCACCTGTGGCCAGCGCCTTGAGA
AGGTCCTTGGCACCGAGCACTGCCATATCCTTCGCGTGCCATTCAGAACAG
AAAACGGAATCGTTCGCAAGTGGATCTCGCGATTTGAAGTCTGGCCGTACC
TGGAGACTTAACTGATGACGTGGCGCATGAGATTGCTGGAGAGCTTCAGG
CCAATCCTGACCTGATCATCGGAACTACAGTGACGGAAACCTTGTTGCGT
GTTTGCTCGCCACAAGATGGGTGTTACTCACTGTACCATTGCCCATGCGC
TTGAGAAAATAAGTACCCTAACTCCGACCTCTACTGGAAGAAGTTTGAGGA

TCCTACTACCACTTCTCGTGCCAGTTCACCACTGACTTGATTGCAATGAACCAT
GCCGACTTCATCATCACCAGTACCTTCCAAGAGATCGCCGGAACAAGGAC
ACCGTCGGCCAGTACGAGTCACACATGGCGTTTACAATGCCTGGCCTGTAC
CGCGTTGTCCACGGCATTGATGTGTTTCGACCCCAAGTTCAACATCGTGTCT
CCTGGCGCGGACCTGTCCATCTACTTCCCGTACACCGAGTCGCACAAGAG
GCTGACCTCCCTTCACCCGGAGATTGAGGAGCTCCTGTACAGCCAAACCGA
GAACACGGAGCACAAGTTCGTTCTGAACGACAGGAACAAGCCAATCATCTT
CTCCATGGCTCGTCTCGACCGTGTGAAGAACTTGACTGGGCTGGTGGAGCT
GTACGGCCGGAACAAGCGGCTGCAGGAGCTGGTGAACCTCGTGGTCTGTCT
GCGGCGACCATGGCAACCCTTCCAAGGACAAGGAGGAGCAGGCCGAGTTC
AAGAAGATGTTTGACCTCATCGAGCAGTACAACCTGAACGGGCACATCCGC
TGGATCTCCGCCAGATGAACCGCGTCCGCAACGGCGAGCTGTACCGCTA
CATCTGCGACACCAAGGGCGCCTTCGTGCAGCCTGCTTTCTACGAGGCTTT
CGGGCTGACGGTGGTTGAGGCCATGACCTGCGGCCTGCCACGTTTGCCA
CAGCCTACGGCGGTCCGGCCGAGATCATCGTGCACGGCGTGTCTGGCTAC
CACATCGACCCTTACCAGGGCGACAAGGCGTCCGGCCCTGCTCGTGGACTT
CTTCGACAAGTGCCAGGCGGAGCCGAGCCACTGGAGCAAGATCTCCCAGG
GCGGGCTCCAGCGTATCGAGGAGAAGTACACCTGGAAGCTGTACTCGGAG
AGGCTGATGACCCTCACCGGCGTGTACGGGTTCTGGAAGTACGTGTCCAAC
CTGGAGAGGCGCGAGACCCGGCGGTACCTGGAGATGCTGTACGCGCTCAA
GTACCGCACCATGGCGAGCACCGTGCCGCTGGCCGTGGAGGGAGAGCCC
TCCAGCAAG

Figure 7C: Seq ID 21 - Amino acid sequence of the open reading frame of ZmSUS-8057

MGEGAGDRVLSRLHSVRERIGDSL SAHPNELVAVFTRLKNLKGMLQPHQIIAE
YNNAIPEAEREKLDGAFEDVLRAAQEAIVIPPWVALAIRPRPGVWEYVRNVNS
ELAVEELRVPEYLQFKEQLVEEGPNNNFVLELDFEPFNASFPRPSLSKSSIGNGV
QFLNRHLSSKLFHDKESMYPLLNFLRAHNYKGMTMMLNDRIRSLALQGALRK
AEEHLSTLQADTPYSEFHHRFQELGLEKGWGDCAKRAQETIHLLDLLEAPDPS
TLEKFLGTIPMVFNVLSPHGYFAQANVLGYPDTGGQVVYILDQVRAMENEML
LRIKQCGLDITPKILIVTRLLPDATGTTGQRLEKVLGTEHCHILRVFPRTENGIVR
KWISRFVWPYLETYTDDVAHEIAGELQANPDLIIGNYSDGNLVACLLAHKMGV
THCTIAHALEKTKYPNSDLYWKKFEDHYHFSCQFTTDLIAMNHADFIITSTFQEIA
GNKDTVGGQYESHMAFTMPGLYRVVHGIDVDFDPKFNIVSPGADLSIYFPYTESHK
RLTSLHPEIEELLYSQTEHTEHFKFVLNDRNKPIIFSMARLDRVKNLTGLVELYGR
NKRLQELVNLVVCGDHGNPSKDKEEQAEFKKMFDLIEQYNLNGHIRWISAQM
NRVRNGELYRYICDTKGAFVQPAFYEAFLTVVEAMTCGLPTFATAYGGPAEII
VHGVSGYHIDPYQGDKASALLVDFFDKCAEPSHWSKISQGGQLQRIEEKYTWK
LYSERLMTLTGVYGFWKYVSNLERRETRRYLEMLYALKYRTMASTVPLAVEGE
PSSK

Figure 8A: Seq ID 22 – Nucleic acid sequence of TaSUS-4775

TCGCTGTCGACGATTTCGTCGAGCCTGCTCCCCTGTGACGTTTCCTTGTTTC
TCGCCCTCCTCCTCCCTCCTCTGGAGCAAGGTTTGAGGACAGCAATGGGCG
AGACTGCCGGAGAGCGCGCCCTGAGCCGTATCCACAGCGTGAGGGAGCG
CATCGGCGATTCCCTCTCCGCGCACACCAATGAGCTTGTGCGCGTCTTCTC
AAGGCTTGTTAACCAAGGAAAGGGGATGCTGCAGCCCCATCAGATAACTGC
TGAGTACAATGCCGCGATCCCAGAGGCCGAGCGTGAGAAGCTCAAGGACA
CCGCCTTTGAGGATCTCCTAAGGGGCGCACAGGAGGCAATTGTCATCCCTC
CATGGGTTGCTCTCGCCATCCGGCCAAGGCCCGGCGTCTGGGAGTATGTG
AGGGTCAATGTGAGCGAGCTTGGTGTGAGGAGTTAAGCATCGCTGAGTAT
TTGCAGTTCAAGGAACAACCTGGCGAATGGAAGCATCGATAACAACCTTTGTG
CTTGAGCTGGACTTTGAGCCATTCAACGCCTCCTTCCCTCGCCCATCGCTG
TCGAAGTCCATTGGCAACGGTGTGCAGTTTCTGAACAGGCACTTGTGATCG
AAGCTGTTCCATGACAAGGAGAGCATGTACCCATTGCTCAACTTCCTTCGC
GCGCACAACCTACAAGGGGATGACCATGATGTTGAACGACAGAATTCGGAGT
CTCAGTACCCTCCAAGGTGCACTCAGGAAGGCAGAGACACATCTGTCAGGC
CTTCAGCTGACACCCCCTACTCGGAGTTCCACCACCGGTTCCAGGAACCT
GGTCTGGAGAAAGGTTGGGGCGACTGTGCTCAGCGTGCGAGTGAGACTAT
CCACCTTCTCTTGGACCTTCTCGAGGCCCTGATCCATCCTCCTTGGAGAA
GTTCTCGGAACAATCCCAATGGTGTTCATGTTGTTATCCTCTCTCCTCAT
GGTTACTTTGCTCAGGCCAATGTCTTGGGGTACCCTGATACTGGTGGACAG
ATTGTCTACATTTTGGACCAAGTCCGTGCTATGGAGAATGAGATGCTGTTGA
GAATCAAGCAGCAAGGTCTTGACATTACACCAAAGATTCTAATAGTCACCAG
GTTGCTCCCTGATGCACATGGCACCACCTGTGGCCAGCGCCTTGAGAAGGT
CCTTGGCACCGAGCACACCCACATCCTGCGTGTGCCATTCAAAACAGAAGA
TGGTATTGTTGCAAAATGGATCTCCCGCTTTGAAGTCTGGCCTTACCTGGAA
GCTTACACTGATGATGTGGCACACGAGATCGCCGGAGAGCTGCAGGCCAC
TCCTGACCTGATCATTGGAACTACAGTGATGGCAACCTAGTCGCGTGTTTG
TTGGCTCACAAGTTGGGAGTTACTCATTGTACCATTGCGCATGCACTCGAGA
AAACCAAGTATCCCAACTCCGACCTTTACTGGAAGAAATTTGAGGATCACTA
CCACTTCTCCTGCCAGTTCACAGCTGACCTGATTGCAATGAATCATGCTGAC
TTCATCATCACCAATACTTTCCAAGAGATTGCCGGAAACAAGGACACCGTAG
GGCAGTACGAGTCGCACATGGCATTACAATGCCAGGCCTCTATCGTGTTG
TCCATGGTATTGATGTCTTCGACCCCAAGTTCAACATCGTCTCCCCTGGTGC
TGACATGTCCATCTACTTCCCATACACTGAACAGCAGAAGAGGCTTACCTCC
CTCCATACTGAGATTGAGGAGCTACTCTTCAGTGATGTTGAGAATGCTGAGC
ACAAATTTGTGCTGAAGGACAAGAAGAAGCCGATCATCTTCTCGATGGCTA
GGCTGGACCGTGTCAAGAATATGACTGGCCTGGTAGAAATGTATGGGCGGA
ATCCTCGCCTACAGGAGCTGGTAAACCTGGTGGTTGTTTGTGGTGACCATG
GAAAGGTGTCCAAGGACAAGGAGGAGCAGGCAGAGTTCAAAAAGATGTTTG
ATCTTATCGAACAGTACAACCTGATTGGTCACATCCGCTGGATCTCTGCTCA
GATGAACCGTGTCCGCAATGGTGAGCTCTACCGCTACATCTGCGACATGAA
GGGAGCCTTTGTGCAGCCTGCTTTCTATGAGGCTTTCCGGTCTTACCGTGATA
GAGGCCATGACATGTGGCCTTCCAACATTGCGCCACTGCATATGGTGGTCCA
GCTGAGATCATTGTGCACGGTGTGTCCGGCTACCACATCGATCCTTACCAG

AATGACAAGGCCTCCGCACTGCTTGTGGACTTCTTTGGGAAGTGCCAGGAA
GACCCGAGCCACTGGAACAAGATCTCGCAGGGAGGACTCCAGCGCATCGA
GGAGAAGTACACCTGGAAGCTGTACTCTGAGAGGCTGATGACCCTTTCTGG
TGTCTATGGTTTCTGGAAGTATGTCTCCAACCTCGACAGGCGCGAGACTCG
TCGCTACCTTGAAATGCTCTACGCCCTCAAGTACCGCAAATGGCTGCAACT
GTCCCATTTGGCTGTTGAGGGCGAGACCTCGGGCAAATGATTTGTCCTTACC
AGAGAAATAAATGGCGGGCGAGCGCTCCGCTTTACCGTTCTTGATTCAGTG
ATGAAGCACAGATCGGAGAGTGTTATGCCTTTGATTGTCCTTTGTTACCGTT
CTTGATTCAGTATGAACTTTAGTTCTGTTGAGGCCCGGCCGTGTTGTGCG
GGTAGGGGAACTGGATGGATGCTGTGTTGTTGGTACTTTCTCATGGATACA
GTATTTGAATGAATGAATAAGATAGTCTTTGTTTTGAAAAAAAAAAAAAAAAA
GCGGCCGC

Figure 8B: Seq ID 23 – Nucleic acid sequence of the open reading frame of TaSUS-4775

ATGGGCGGAGACTGCCGGAGAGCGCGCCCTGAGCCGTATCCACAGCGTGAG
GGAGCGCATCGGCGATTCCCTCTCCGCGCACACCAATGAGCTTGTGCGCCG
TCTTCTCAAGGCTTGTTAACCAAGGAAAGGGGATGCTGCAGCCCCATCAGA
TAACTGCTGAGTACAATGCCGCGATCCCAGAGGCCGAGCGTGAGAAGCTC
AAGGACACCGCCTTTGAGGATCTCCTAAGGGGCGCACAGGAGGCAATTGT
CATCCCTCCATGGGTTGCTCTCGCCATCCGGCCAAGGCCCGGCGTCTGGG
AGTATGTGAGGGTCAATGTGAGCGAGCTTGGTGTGAGGAGTTAAGCATCG
CTGAGTATTTGCAGTTCAAGGAACAACCTGGCGAATGGAAGCATCGATAACA
ACTTTGTGCTTGAGCTGGACTTTGAGCCATTCAACGCCTCCTTCCCTCGCCC
ATCGCTGTCGAAGTCCATTGGCAACGGTGTGCAGTTTCTGAACAGGCACTT
GTCATCGAAGCTGTTCCATGACAAGGAGAGCATGTACCCATTGCTCAACTTC
CTTCGCGCGCACAACCTACAAGGGGATGACCATGATGTTGAACGACAGAATT
CGGAGTCTCAGTACCCTCCAAGGTGCACTCAGGAAGGCAGAGACACATCTG
TCAGGCCTTCCAGCTGACACCCCCTACTCGGAGTTCCACCACCGGTTCCAG
GAACTTGGTCTGGAGAAAGGTTGGGGCGACTGTGCTCAGCGTGCGAGTGA
GACTATCCACCTTCTCTTGACCTTCTCGAGGCCCTGATCCATCCTCCTTG
GAGAAGTTCCTCGGAACAATCCCAATGGTGTTCATGTTGTTATCCTCTCTC
CTCATGGTTACTTTGCTCAGGCCAATGTCTTGGGGTACCCTGATACTGGTG
GACAGATTGTCTACATTTTGGACCAAGTCCGTGCTATGGAGAATGAGATGCT
GTTGAGAATCAAGCAGCAAGGTCTTGACATTACACCAAGATTCTAATAGTC
ACCAGGTTGCTCCCTGATGCACATGGCACCACCTGTGGCCAGCGCCTTGAG
AAGGTCCTTGGCACCGAGCACACCCACATCCTGCGTGTGCCATTCAAAACA
GAAGATGGTATTGTTGCAATGGATCTCCCGCTTTGAAGTCTGGCCTTACC
TGGAAGCTTACACTGATGATGTGGCACACGAGATCGCCGGAGAGCTGCAG
GCCACTCCTGACCTGATCATTGGAACTACAGTGATGGCAACCTAGTCGCG
TGTTTGTGCTCACAAGTTGGGAGTTACTCATTGTACCATTGCGCATGCAC
TCGAGAAAACCAAGTATCCCAACTCCGACCTTTACTGGAAGAAATTTGAGGA
TCACTACCACTTCTCCTGCCAGTTCACAGCTGACCTGATTGCAATGAATCAT
GCTGACTTCATCATCAACCAATACTTTCCAAGAGATTGCCGGAAACAAGGACA
CCGTAGGGCAGTACGAGTCGCACATGGCATTACAATGCCAGGCCTCTATC

GTGTTGTCCATGGTATTGATGTCTTCGACCCCAAGTTCAACATCGTCTCCCC
TGGTGCTGACATGTCCATCTACTTCCCATACTGAACAGCAGAAGAGGCTT
ACCTCCCTCCATACTGAGATTGAGGAGCTACTCTTCAGTGATGTTGAGAATG
CTGAGCACAAATTTGTGCTGAAGGACAAGAAGAAGCCGATCATCTTCTCGAT
GGCTAGGCTGGACCGTGTCAAGAATATGACTGGCCTGGTAGAAATGTATGG
GCGGAATCCTCGCCTACAGGAGCTGGTAAACCTGGTGGTTGTTTGTGGTGA
CCATGGAAAGGTGTCCAAGGACAAGGAGGAGCAGGCAGAGTTCAAAAAGA
TGTTTGATCTTATCGAACAGTACAACCTGATTGGTCACATCCGCTGGATCTC
TGCTCAGATGAACCGTGTCCGCAATGGTGAGCTCTACCGCTACATCTGCGA
CATGAAGGGAGCCTTTGTGCAGCCTGCTTTCTATGAGGCTTTTCGGTCTTACC
GTGATAGAGGCCATGACATGTGGCCTTCCAACATTCGCCACTGCATATGGT
GGTCCAGCTGAGATCATTGTGCACGGTGTGTCCGGCTACCACATCGATCCT
TACCAGAATGACAAGGCCTCCGCACTGCTTGTGGACTTCTTTGGGAAGTGC
CAGGAAGACCCGAGCCACTGGAACAAGATCTCGCAGGGAGGACTCCAGCG
CATCGAGGAGAAGTACACCTGGAAGCTGTACTCTGAGAGGCTGATGACCCT
TTCTGGTGTCTATGGTTTCTGGAAGTATGTCTCCAACCTCGACAGGCGCGA
GACTCGTCGCTACCTTGAAATGCTCTACGCCCTCAAGTACCGCAAAATGGC
TGCAACTGTCCCATTGGCTGTTGAGGGCGAGACCTCGGGCAA

Figure 8C: Seq ID 24 - Amino acid sequence of the open reading frame of TaSUS-4775

MGETAGERALSRIHSVRERIGDSL SAHTNELVAVFSRLVNQGKGMLQPHQITAE
YNAAIPEAEREKLKDTAFEDLLRGAQEAIIPPWVALAIRPRPGVWEYVRNVSE
LGVEELSIAEYLQFKEQLANGSIDNNFVLELDFEPFNASFPRPSLSKSGINGVQF
LNRHLSSKLFHDKESMYPLLNFLRAHNYKGMTMMLNDRIRSLSTLQGALRKA
THLSGLPADTPYSEFHHRFQELGLEKGWGDCAQRASETIHLLLDLLEAPDPSSL
EKFLGTIPMVFNVLSPHGYFAQANVLGYPTDGGQIVYILDQVRAMENEMLLRI
KQQGLDITPKILIVTRLLPDAHGTTCGQRLEKVLGTEHTHILRVPFKTEDGIVRKW
ISRFEVWPYLEAYTDDVAHEIAGELQATPDLIIGNYSDGNLVACLLAHKLGVTHC
TIAHALEKTKYPNSDLYWKKFEDHYHFSCQFTADLIAMNHADFIITNTFQEIAGN
KDTVGGQYESHMAFTMPGLYRVVHGIDVDFPKFNIVSPGADMSIYFPYTEQQKR
LTSLHTEIEELLFSDVENAEHKFVLKDKKKPIIFSMARLDRVKNMTGLVEMYGRN
PRLQELVNLVVCGDHGKVS KDKEEQAEFKMFDLIEQYNLIGHIRWISAQMNR
VRNGELYRYICDMKGAFVQPAFYEAFLTVIEAMTCGLPTFATAYGGPAEIIVHG
VSGYHIDPYQNDKASALLVDFFGKCQEDPSHWNKISQGGGLQRIEEKYTWKLYS
ERLMTLSGVYGFWKYVSNLDRRETRRYLEMLYALKYRKMAATVPLAVEGETSG
K

Figure 9A: Seq ID 25 – Nucleic acid sequence of OsSUS-9471

CCATTGGTTCCTGAAATTGAGTCATGGCTGCCAAGCTAGCTCGCCTCCACA
GTCTCCGCGAACGCCTCGGTGCCACCTTCTCGTCTCATCCCAATGAGTTGA
TTGCACTCTTCTCTAGGTATGTTAACCAGGGAAAGGGAATGCTCCAGCGTC
ACCAGCTGCTTGCGGAGTTGATGCCTTGATCGAAGCTGACAAAGAGAAAT
ATGCTCCCTTTGAAGACATTCTCCGGGCTGCTCAGGAAGCCATTGTGCTGC

CGCCCTGGGTTGCACTGGCCATCAGGCCAAGGCCTGGTGTCTGGGACTAC
ATTCGGGTGAATGTAAGTGAGTTGGCAGTGGAAGAGCTGAGTGTCTGAG
TACTTGGCATTCAAGGAACAGCTTGTTGATGGACACACCAACAGCAACTTTG
TTCTTGAGCTTGATTTTGAGCCCTTCAATGCCTCCTTCCCGCGCCCGTCCAT
GTCCAAGTCCATCGGAAATGGGGTGCAGTTCCTTAACCGTCACCTGTCGTC
CAAGTTGTTCCAGGACAAGGAGAGCCTCTACCCCTTGCTGAACTTCCTGAA
AGCCCATAAACCACAAGGGCACGACAATGATGCTGAATGACAGAATTCAGAG
CCTTCGTGGGCTCCAATCATCCCTTAGAAAGGCAGAAGAATATCTGATGGG
CATTCTCAAGACACGCCCTACTCGGAGTTCAACCACAGGTTCCAAGAGCT
CGGTTTGGAGAAGGGTTGGGGTGAAGTGTGCAAAGCGTGTGCTTGACACCAT
CCACTTGCTTCTTGACCTTCTTGAGGCCCTGATCCGGCCAACCTTGAGAA
GTTCTTGGAACCTATTCCAATGATGTTCAATGTTGTTATCCTGTCTCCGCATG
GATACTTTGCCCAATCCAATGTGTTGGGATACCCTGATACTGGTGGTCAGGT
TGTGTACATTTTGGACCAAGTCCGCGCTTTGGAGAATGAGATGCTTTTGAGG
ATCAAGCAGCAAGGCCTTGATATCACACCTAAGATCCTCATTGTAACCAGGC
TGTTGCCTGATGCTGTTGGTACTACATGCGGCCAGCGTGTGGAGAAGGTTA
TTGGAAGTGAACACACTGACATTCTTCGTGTTCCATTGAGGAGTGAGAATGG
TATCTCCGCAAGTGGATCTCCCGTTTTGATGTCTGGCCATTCTTGGAACA
TAACTGAGGATGTTGCAAACGAAATTATGAGGGAAATGCAAGCCAAACCT
GATCTCATCATTGGCAATTACAGTGATGGAAACCTTGTTGCCACTCTGCTGG
CTCACAATTAGGAGTTACCCAGTGTACCATTGCTCATGCCTTGAGAGAAAC
CAAATACCCCAACTCAGACATATACTTGGACAAGTTTGACAGCCAGTACCAC
TTCTCATGCCAATTCAGTCTGATCTTATCGCCATGAATCACACTGATTTTCT
CATCACCAGTACATTCCAAGAAATTGCTGGAAGCAAGGACACTGTGGGGCA
GTATGAATCACACATTGCATTACCCCTTCTGGGCTTTACCGAGTTGTGCAT
GGCATAGATGTTTTTGATCCCAAGTTCAACATTGTCTCTCCTGGAGCTGACA
TGAGTGTCTACTTCCCGTACACCGAGGCTGACAAGAGGCTCACTGCTTTCC
ACCCTGAAATTGAGGAGCTTCTCTACAGTGAAGTCGAGAACGATGAACACA
AGTTTGTATTGAAGGACAAGAACAAGCCAATCATCTTCTCCATGGCTCGTCT
TGACCGAGTGAAGAACATGACAGGTCTGGTTGAGATGTATGGTAAGAATGC
ACATCTCAGGGATTTGGCAAACCTTGTTGATTGTTTGTGGTGACCAAGGCAAT
CAGTCCAAGGACAGGGAGGAGCAGGCTGAGTTCAAGAAGATGTACGGTCT
CATTGACCAGTACAAGTTGAAGGGGCATATCCGCTGGATCTCAGCTCAGAT
GAACCGTGTTGTAACGGGGAGTTGTACCGATACATTTGTGACACCAAGGG
AGTCTTTGTCCAGCCTGCATTCTATGAAGCGTTTGGTCTGACTGTCATCGAA
GCCATGACATGTGGTTTGCCAACAATCGCAACATGCCATGGTGGCCCTGCT
GAGATTATTGTTGATGGGGTGTCTGGTCTGCACATTGATCCTTACCACAGTG
ACAAGGCTGCTGATATCTTGGTCAACTTCTTTGAGAAGTGCAAGCAGGATTC
AACCTACTGGGACAATATTTTACAGGGAGGTCTGCAGAGGATTTACGAGAA
GTACACCTGGAAGCTGTACTCTGAGAGGCTGATGACCTTGACTGGTGTATA
CGGATTCTGGAAGTACGTAAGCAACCTTGAGAGGCGCGAGACTCGCCGTTA
CATTGAGATGTTCTATGCTCTGAAATACCGCAGCCTGGCCAGCGCCGTCCC
ATTGGCTGTGATGGAGAGAGCACATCCAAGTAATGGAGGGGAAAATATGC
ATCTTCAGCAGGAGAAGCCGTCAGCTGCATTGCAATTTGATAATTTTCTGTA
GTTGTCAATTTGGCATCCATGTTTGCCATGGATGTGTACTATGTCTAAGGTTT
CAGTACTTTTGCGAGATTTTGGGCAGTGCTTGCCTCAAATAAACGCCGGTTC

CTGGTGTTTATCGTTCAGAATTCAATGCAATGTTTTATTGCCGCGCTTCCTCA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAGTGCCCTACTGTAAAAAAGGTTGGTTCAAAGCCATGGATGA
ACCCAACATTTTTGTTACCTTGTTCTTTTCAGAATTTCTTAAGAGTTGTGATA
GTAGGCCCAGTGCGACCAAAATGTTGGGTTCTTGAGTTTTCAGTATAAAAT
TTCTGAT

Figure 9B: Seq ID 26 - Nucleic acid sequence of the open reading frame of
OsSUS-9471

ATGGCTGCCAAGCTAGCTCGCCTCCACAGTCTCCGCGAACGCCTCGGTGC
CACCTTCTCGTCTCATCCCAATGAGTTGATTGCACTCTTCTCTAGGTATGTTA
ACCAGGGAAAGGGAATGCTCCAGCGTCACCAGCTGCTTGCGGAGTTGAT
GCCTTGATCGAAGCTGACAAAGAGAAATATGCTCCCTTTGAAGACATTCTCC
GGGCTGCTCAGGAAGCCATTGTGCTGCCGCCCTGGGTTGCACTGGCCATC
AGGCCAAGGCCTGGTGTCTGGGACTACATTCGGGTGAATGTAAGTGAGTTG
GCAGTGGAAGAGCTGAGTGTTTCTGAGTACTTGGCATTCAAGGAACAGCTT
GTTGATGGACACACCAACAGCAACTTTGTTCTTGAGCTTGATTTTGAGCCCT
TCAATGCCTCCTTCCCGCGCCCGTCCATGTCCAAGTCCATCGGAAATGGGG
TGCAGTTCCTTAACCGTCACCTGTCTCCAAGTTGTTCCAGGACAAGGAGA
GCCTCTACCCCTTGCTGAACTTCCTGAAAGCCCATAACCACAAGGGGCACGA
CAATGATGCTGAATGACAGAATTCAGAGCCTTCGTGGGCTCCAATCATCCCT
TAGAAAGGCAGAAGAATATCTGATGGGCATTCTCAAGACACGCCCTACTC
GGAGTTCAACCACAGGTTCCAAGAGCTCGGTTTGGAGAAGGGTTGGGGTG
ACTGTGCAAAGCGTGTGCTTGACACCATCCACTTGCTTCTTGACCTTCTTGA
GGCCCCTGATCCGGCCAACTTGAGAGAAGTTCCTTGGAATATTCCAATGAT
GTTCAATGTTGTTATCCTGTCTCCGCATGGATACTTTGCCCAATCCAATGTG
TTGGGATACCCTGATACTGGTGGTCAGGTTGTGTACATTTTGACCAAGTCC
GCGCTTTGGAGAATGAGATGCTTTTGAGGATCAAGCAGCAAGGCCTTGATA
TCACACCTAAGATCCTCATTGTAACCAGGCTGTTGCCTGATGCTGTTGGTAC
TACATGCGGCCAGCGTGTGGAGAAGGTTATTGGAATGAGCACACTGACAT
TCTTCGTGTTCCATTGAGGAGTGAGAATGGTATCCTCCGCAAGTGGATCTCC
CGTTTTGATGTCTGGCCATTCTTGGAACATACTGAGGATGTTGCAAACG
AAATTATGAGGGAAATGCAAGCCAAACCTGATCTCATCATTGGCAATTACAG
TGATGGAAACCTTGTTGCCACTCTGCTGGCTCACAAATTAGGAGTTACCCAG
TGTACCATTGCTCATGCCTTGAGAGAAAACCAAATACCCCAACTCAGACATAT
ACTTGGAACAAGTTTGACAGCCAGTACCACTTCTCATGCCAATCACTGCTGA
TCTTATCGCCATGAATCACACTGATTTTCATCATCACCAAGTACATTCCAAGAAA
TTGCTGGAAGCAAGGACACTGTGGGGCAGTATGAATCACACATTGCATTCA
CCCTTCCTGGGCTTTACCGAGTTGTGCATGGCATAGATGTTTTTGATCCCAA
GTTCAACATTGTCTCTCCTGGAGCTGACATGAGTGTCTACTTCCCGTACACC
GAGGCTGACAAGAGGCTCACTGCTTTCCACCCTGAAATTGAGGAGCTTCTC
TACAGTGAAGTCGAGAACGATGAACACAAGTTTGTATTGAAGGACAAGAACA
AGCCAATCATCTTCTCCATGGCTCGTCTTGACCGAGTGAAGAACATGACAG
GTCTGGTTGAGATGTATGGTAAGAATGCACATCTCAGGGATTTGGCAAACCT
TGTGATTGTTTGTGGTGACCACGGCAATCAGTCCAAGGACAGGGAGGAGCA

GGCTGAGTTCAAGAAGATGTACGGTCTCATTGACCAGTACAAGTTGAAGGG
 GCATATCCGCTGGATCTCAGCTCAGATGAACCGTGTTTCGTAACGGGGAGTT
 GTACCGATACATTTGTGACACCAAGGGAGTCTTTGTCCAGCCTGCATTCTAT
 GAAGCGTTTGGTCTGACTGTCATCGAAGCCATGACATGTGGTTTGCCAACA
 ATCGCAACATGCCATGGTGGCCCTGCTGAGATTATTGTTGATGGGGTGTCT
 GGTCTGCACATTGATCCTTACCACAGTGACAAGGCTGCTGATATCTTGGTCA
 ACTTCTTTGAGAAGTGCAAGCAGGATTCAACCTACTGGGACAATATTTACA
 GGGAGGTCTGCAGAGGATTTACGAGAAGTACACCTGGAAGCTGTACTCTGA
 GAGGCTGATGACCTTGACTGGTGTATACGGATTCTGGAAGTACGTAAGCAA
 CCTTGAGAGGGCGCGAGACTCGCCGTTACATTGAGATGTTCTATGCTCTGAA
 ATACCGCAGCCTGGCCAGCGCCGTCCCATTGGCTGTGATGGAGAGAGCA
 CATCCAAG

Figure 9C: Seq ID 27 - Amino acid sequence of the open reading frame of OsSUS-9471

MAAKLARLHSLRERLGATFSSHPNELIALFSRYVNQKGMLQRHQLLAEFDALI
 EADKEYAPFEDILRAAQEAIVLPWVALAIRPRPGVWDYIRVNVSELAVEELSV
 SEYLAFFEQLVDGHTNSNFVLELDFEPFNASFPRPSMSKSIGNGVQFLNRHLSS
 KLFQDKESLYPLLNFLKAHNHKGTTMMLNDRIQSLRGLQSSLRKAEYLMGIPQ
 DTPYSEFNHRFQELGLEKGWGDCAKRVLDTIHLLDLLEAPDPANLEKFLGTIP
 MMFNVVILSPHGYFAQSNVLGYPDGTGGQVVYILDQVRALENEMLLRIKQQGLDI
 TPKILIVTRLLPDAVGTTGQQRVEKVGTEHTDILRVPFRSENGILRKWISRFDVW
 PFLETYTEDVANEIMREMQAAPDLIIGNYSDGNLVATLLAHKLGVTQCTIAHALE
 KTKYPNSDIYLDKFDSDQYHFSCQFTADLIAMNHTDFIITSTFQEIAGSKDTVQY
 ESHIAFTLPGLYRVVHGIDVFDPKFNIVSPGADMSVYFPYTEADKRLTAFHPEIE
 ELLYSEVENDEHKFVLKDKNKPIIFSMARLDRVKNMTGLVEMYGKNAHLRDLAN
 LVIVCGDHGNQSKDREEQAEFKKMYGLIDQYKLKGHIRWISAQMNRVRNGELY
 RYICDTKGVFVQPAFYEAFLTVIEAMTCGLPTIATCHGGPAEIIVDGVSGLHIDP
 YHSDKAADILVNFFEKCKQDSTYWDNISQGLQRIYEKYTWKLYSERLMTLTGV
 YGFWKYVSNLERRETRYIEMFYALKYRSLASAVPLAVDGESTSK

Figure 10A: Seq ID 28 - Nucleic acid sequence of PpSUS-2

GGCGCGCCAGGACAGGCAGAGGACGAGACAAGGGGGGCTTGTTTGTGGTATTGT
 GGATATTTTCGGAACGGTTTACGGGTTTTGTAGCGAGTTTTTGGACCTCATCCGGT
 TAGGATTTTGGGAATTAAGGCTGTTAGGCTGTTCAATTGAGTCGGCCCTGCAGGT
 GATTATTGTGGGAATCGGAGATCGCCTCTCGGGAGTGTTTTTAAGGTTTCGATCGC
 AACATGTGCGAGCCACGGCCGACGCTTCGGCGTCTTACGGGCCTCAAGGAGCGAG
 TGGAGAGCTCTTTGCAAGAGCATCGCAATGAGCTGCTCCACCTTCTGCAAGGCTAT
 GTGGTACAAGGTCGTTCAATTCTTCAACCGCATCATTTGCAAGATCAACTGGCTGCT
 GTGCACGATGCTGCCCATATTCAAGACACTGCTATTGGAAAGCTTCTTCAGAATTGC
 CAGGAAGCCATGGTGTGCGCCACCTTGGGTTGGATTTGCCGTGCGTCCAAGGCCTG
 GAATCTGGGAGTATGTGCGGATCAATGTGGAGGAGCTCATTGTGGAAGAGCTTAG
 CGTGTCCGAATACTTGGGTTTCAAAGAACAGCTTAGCCTTGGCTCTGATTCAATTGA
 TCTATATGTGCTGGAGTTGGATTTGAGCCCTTCAATGCGCACTTTCGCGCATGA
 CAAGACCATCATCCATTGGACATGGTGTACAGTTTTTGAATCGCCACTTATCTTCTA
 AACTCTTTCAAAATCCAGAGAGCATGGAGCCTTTGTTTCAATTCCTTCGCCTCCATA

CTTATCGAGGAGAGACGCTGATGCTCAATGAGCGGATTGCAACTTTTTCTAGGTTT
CGTCCTCAGTTAGTTAGAGCTGAAGAGGCCCTATCAAACTCCCAGAGGATACTCC
ATTTTCCAGCTTTGCACACAGGTTACAAGAGTTAGGTTTGAAAAAGGTTGGGGAA
ATACAGCAGGGCGTGTGCTGCAAACCTTGAAGTTGTTACTAGATCTGCTCCAGGCT
CCTGACCCAGACACACTAGAAAAGTTCTTAGCCAGAATTCCTATGATTTTACCGTT
TGCATTGTTTCTCCTCATGGTTACTTTGGTCAAGCAGGTGTTCTTGGCTTGCCTGAT
ACTGGTGGTCAGGTTGTTTACATATTGGATCAAGTGAGAGCTCTGGAAAATCAAATG
TTGGAAAATCTTCAGCTTCAAGGGTTGGATTTCAAGCCGCAAATTGTTATTCTCACA
CGATTGATTCCTAATGCTAATGGGACTACCGTCAACCAGCGTATCGAAAAGGTTTCA
GGCACTCAGCATTGAGGATTCTACGTGTTCTTTCCAACACGAAGGCAACATCCT
GAAGAACTGGATTTCTCGTTTTGATGTTTACCCCTACTTGGAAAATGCTCAGGA
CGCGGCCAGGGAAGTCCTTGGGGAGCTCCAAGGGCGGCCTGATCTGATTATTGGT
AATTACAGCGATGGTAACCTGGTCGCGACGCTTCTTTCTCACTACCTTGATGTAAT
CAGTGTATTATTGCACATGCGTTGGAGAAGACAAAGTATCCGGACTCGGACATTTAT
TGGAAGGACTTTGAAGAGAAGTATCACTTTTCATGCCAGTTCACTGCTGATCTCATT
GCGATGAATAGTGCGGACTTCATCATCACTAGCACCTATCAAGAGATTGCCGGAAG
TGCTGATACCGTGGGACAGTATGAAAGCCATCAGGCATTCACAATGCCAGGGTTGT
ATCGAGTCGTGAATGGCATAGACGCTTTTATGCCAAATTCAACATTGTGTCCCCAG
GAGCAGATATGAACATCTACTATCCATTGCGAGACAAAGAACGTCGCTTGACAAGT
CTACAGGAATCTATTGAGGAGCTCCTTTATAGCCCTGAACAGACGGATGAGCATAT
TGGGTTGATTGACAAGGAAAAGCCATTCTCTTTTCCATGGCTCGACTCGACAGAG
TAAAGAACCTCACAGGGTTAGTAGAGATGTACGGAAAGAACCAGAACTAAAAGAA
TTTGTACATTTGGTAATTGTTGGAGGCGAAATCAATCCTTCAAAGTCGAAGGACCGG
GAAGAAGTTAGAGAGATTGAGAAGATGCACAATCTCATCAAGCGGTACAAGCTTGA
AAATAACTTCAGATGGATCCGATCACAGACAAATAGAATTCGAAATGGAGAGTTGTA
CCGTTACATTGCAGATTCTCAAGGAGCTTTTGTACAGCCTGCACTTTATGAAGTTT
CGGTCTGACAGTGGTCGAAGCTATGACAAGTGGACTTCCACCTTTGCAACGAGTC
ATGGAGGCCCAGCTGAGATTATTGAGCATGGGATTTCTGGGTATCATATCGATCCT
TATTATCCTGATGAAGCTGCTGAACAGATCGTTGCTTCTTCGAGAAATGTAAAAAT
GAGCCTGGACTTTGGAATAAAGTTTCTGAAGCTGGATTACAACGCATATATTCCAGC
TACACTTGGAAGATATATGCCGAGCGGCTTATGACACTGTCTGCGGTGTACGGATT
CTGGAAGTATGTTTCAAATTTGCACAGGCAGGAAGCTCGGAGATATTTGGAGATGT
TTTACATCTTGAAGTTCCGTGAATTGGCAAGAACAGTACCGCTTCAAAGGATGATG
AGGATGTTCTAGAAAAAGTTGAGAAGAAAGCTCAGTTGGGTCCTGGCGTAGGGGC
AATTGTTGGTGAAGCAGCAACAGCAGTCGAAGCACGGAAAGCTGTTACTGGTCATA
CATAATCGCAGTGTAATTAACAGTTCGTCATAAAAGCTATCTGTATGCTTTGCACA
GAAGGATGTAATATATGTTTTGGGGTTTGGTTGAAGTGCATGGTTTGAACGCCAACA
TCACAGTTGATGACTTAGATCTTGTGTTGAAACATGTTATAGATTTTAAATCATTTAACT
TGTGAAACAATATGATACACCAGCAATTATTGTTGAAAGCAAAAAAAAAAAAAAAAAA
AAAAAGCGGCCGC

Figure 10B: Seq ID 29 - Nucleic acid sequence of the open reading frame of PpSUS-2

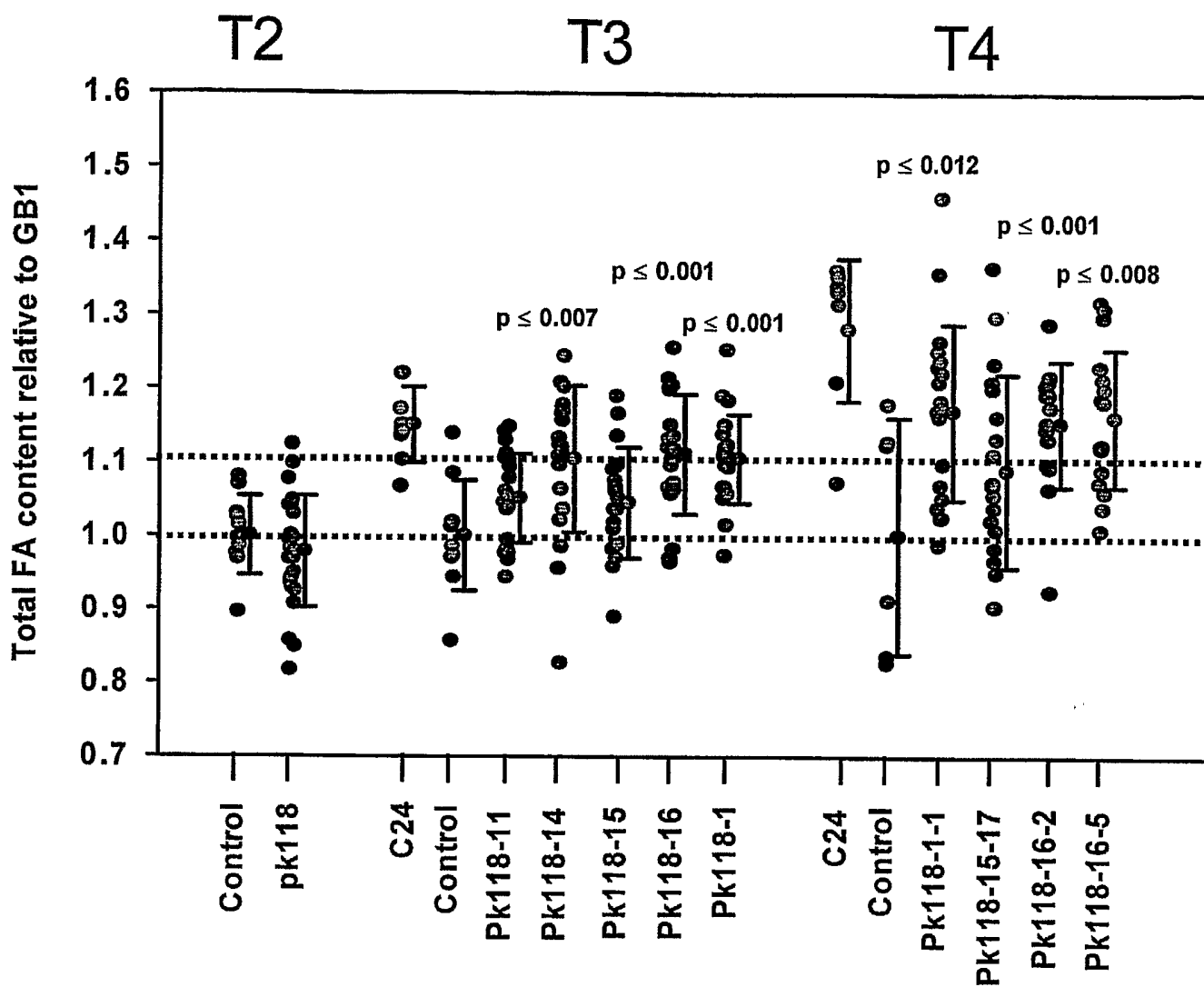
ATGGCTCGTGCTTCGACTAACTGGCTATCGTTCTCTCTCTCCCCATGGAAA
TGCTCCGAACCCCCGAACCTCAGTTCGTTCAATACGACGCCGCTTCCGACA
CTTCTCGCATCACTACTACCTCGACAACCTTGACACCAACGGGTGGGGGA
ACGGGAGCCTCAAGTTTGAGCAGAATCTGAACCACAGCGACGTGAGTTTCG
TTGAATCGTCGTCGACAGAGCGTCAGCCACGCGCCGCGGAAGCTGGAGGAT
TTTCTCGGCGACTCCTCCGCTGTTATGCGTTACTCCGACAGCCAGACGGAG

ACGCAGGACTCGTCGCTGACGCACATCTACGACCACCACCACCACCACCAC
CACCACCACCACCACGGTTCTTCTGCGTACTTCGGCGGTGACCACCAGGAT
CTCAAGGCCATTACTGGATTCCAAGCTTTTTCGACTAACTCTGGCTCCGAGG
TTGATGATTCTGCATCGATCGGAAAGGCGCAGGGCAGCGAGTTCGGGACT
CACTCTATTGAGTCCTCCGTCAACGAGTTCGCCGCGTTCTCCGGTGGCACC
AACACCGGTGGAACCTTGTGCTCGCCGTGCGGCAGAGCTCCGAGAAGGC
CGTCGCTGCTGCGGCGGAGTCCGATCGCTCGAAGAAGGTTGTGGATACCT
TCGGCCAGCGGACTTCTATATACAGAGGTGTCACTAGGCACCGATGGACAG
GAAGATATGAAGCGCATCTATGGGACAATAGTTGCAGAAGGGAGGGTCAAG
CCAGAAAAGGGCGTCAAGTTTATTTGGGTGGATATGATAAGGAAGAAAAGG
CCGCGAGAGCTTATGATTTGGCAGCTCTAAAGTACTGGGGTCCCCTGCTA
CCACCAACTTCCCTGTTTCCAATTATTCGAAGGAAGTGGAGGAGATGAAACA
TGTAACAAAGCAAGAATTTATTGCATCATTGCGGAGGAAAAGTAGTGGTTTC
TCCAGGGGAGCTTCCATATACAGAGGTGTTACAAGGCATCATCAACAGGGT
AGGTGGCAAGCAAGAATTGGCCGTGTAGCTGGAAACAAAGATTTATACTTG
GGAACATTCGCAACCGAGGAGGAAGCAGCAGAGGCATATGATATTGCAGC
CATAAAGTTCAGAGGTGCAAACGCGGTAACCAACTTTGAGATGAATAGATAT
GATGTGGAAGCTATAATGAAGAGTTCTCTTCCAGTGGGTGGGGCAGCAAAA
CGCTTGAGGCTTTCCCTTGAATCAGAGCAGAAAGCTCCTCCTGTGAACAGC
AGCAGTCAGCAGCAGAATCCACAGTGTGGTAACGTGAGTGGTAGCATCAAT
TTCTCAGCCATTCATCAGCCAATTGCTTCAATCCCTTGTGGAATTCCGTTTGA
TTCAACAACAGCATATTATCCTCACAACCTTTTCCAACATTTTCACCCTACCA
ACGCTGGTGCAGCAGCGTCTGCTGTTACTTCTGCCAATGCAACCGCACTAA
CTGCACTGCCAGCATCAGCAGCAACTGAGTTCCTTTATTTGGCCTCATCAGTC
TTATTGA

Figure 10C: Seq ID 30 - Amino acid sequence of the open reading frame of PpSUS-2

MSQPRPTLRRLTGLKERVESSLQEHRNELLHLLQGYVVQGRSILQPHHLQDQL
AAVHDAAHIQDTAIGKLLQNCQEAMVSPPWVGFVAVRPRPGIWEYVRINVEELIV
EELSVSEYLGFKQLSLGSDSIDLYVLELDFEPFNAHFPRMTRPSSIGHGVQFL
NRHLSSKLFQNPESMEPLFQFLRLHTYRGETLMLNERIATFSRFRPQLVRAEEA
LSKLPEDTPFSSFAHRLQELGLEKGWGNTAGRVLQTLKLLDLLQAPDPDTLEK
FLARIPMIFTVCIVSPHGYFGQAGVLGLPDTGGQVVYILDQVRALENQMLENLQL
QGLDFKPQIVILTRLIPNANGTTVNQRIEKVSGTQHSRILRVFPQHEGNILKNWIS
RFDVYPYLENYAQDAAREVLGELQGRPDLIIGNYS DGNLVATLLSHYLDVTQCII
AHAEKTKYPDSDIYWKDFEEKYHFSCQFTADLIAMNSADFIITSTYQEIAGSAD
TVGQYESHQAFTMPGLYRVVNGIDVDFPKFNIVSPGADMNIYYPFADKERRLTS
LQESIEELLYSPEQTDEHIGLIDKEKPILFSMARLDRVKNLTGLVEMYGKNQKLK
EFVHLVIVGGEINPSKSKDREEVREIEKMHNLIKRYKLENNFRWIRSQTNIRNG
ELYRYIADSQGA FVQPALYEGFGLTVVEAMTSGLP TFATSHGGPAEIIIEHGISGY
HIDPYYPDEAAEQIVAFFEKCKNEPGLWNKVSEAGLQRIYSSYTWKIYAERLMT
LSAVYGFWKYVSKLHRQEARRYLEM FYILKFRELARTVPLSKDDEDVLEKVEKK
AQLGPGVGAI VGEAATAVEARKAVTGHT

Figure 11



SEQUENCE LISTING

<110> BASF Plant Science GmbH
Haertel, Heiko A
Oswald, Oliver

<120> NUCLEIC ACID MOLECULES ENCODING SUCROSE SYNTHASE-LIKE
POLYPEPTIDES AND METHODS OF USE

<130> PF 2063015126

<150> US 60/595112

<151> 2005-06-07

<160> 42

<170> PatentIn version 3.3

<210> 1

<211> 2418

<212> DNA

<213> Arabidopsis thaliana

<220>

<221> CDS

<222> (1)..(2418)

<400> 1

atg gcg tct ttc ttt gat ctc gta tac ata agc tgg att ggt ttc tta	48
Met Ala Ser Phe Phe Asp Leu Val Tyr Ile Ser Trp Ile Gly Phe Leu	
1 5 10 15	
aac aaa ttc ctc tcc ttt tgg gtc ttc tgg gtt tgc ctt gta aga tat	96
Asn Lys Phe Leu Ser Phe Trp Val Phe Trp Val Cys Leu Val Arg Tyr	
20 25 30	
gta gcc cag gga aag ggg ata ttg cag tcc cac cag ctg att gat gag	144
Val Ala Gln Gly Lys Gly Ile Leu Gln Ser His Gln Leu Ile Asp Glu	
35 40 45	
ttc ctt aag act gtg aaa gtt gat gga aca tta gaa gat ctt aac aaa	192
Phe Leu Lys Thr Val Lys Val Asp Gly Thr Leu Glu Asp Leu Asn Lys	
50 55 60	
agt cca ttc atg aaa gtt ctg cag gaa gcc ata gtt ttg cct cca ttt	240
Ser Pro Phe Met Lys Val Leu Gln Glu Ala Ile Val Leu Pro Pro Phe	
65 70 75 80	
gtt gct ttg gct ata cgt ccc aga cct ggt gtt agg gaa tat gtc cgt	288
Val Ala Leu Ala Ile Arg Pro Arg Pro Gly Val Arg Glu Tyr Val Arg	
85 90 95	
gtg aat gtg tat gag ctg agc gta gat cat tta act gtt tct gaa tat	336
Val Asn Val Tyr Glu Leu Ser Val Asp His Leu Thr Val Ser Glu Tyr	
100 105 110	
ctt cgg ttt aag gaa gag ctc gtt aat ggc cat gcc aat gga gat tat	384
Leu Arg Phe Lys Glu Glu Leu Val Asn Gly His Ala Asn Gly Asp Tyr	
115 120 125	
ctc ctt gaa ctt gat ttt gaa cct ttc aat gca aca ttg cct cgc cca	432

Leu	Leu	Glu	Leu	Asp	Phe	Glu	Pro	Phe	Asn	Ala	Thr	Leu	Pro	Arg	Pro		
130						135					140						
act	cgt	tca	tca	tcc	att	ggg	aat	ggg	gtt	cag	ttc	ctc	aat	cgt	cac	480	
Thr	Arg	Ser	Ser	Ser	Ile	Gly	Asn	Gly	Val	Gln	Phe	Leu	Asn	Arg	His		
145					150					155					160		
ctc	tct	tca	att	atg	ttc	cgt	aac	aaa	gaa	agc	atg	gag	cct	ttg	ctt	528	
Leu	Ser	Ser	Ile	Met	Phe	Arg	Asn	Lys	Glu	Ser	Met	Glu	Pro	Leu	Leu		
				165					170					175			
gag	ttt	ctc	cgc	act	cac	aaa	cat	gat	ggc	cgt	cct	atg	atg	ctg	aat	576	
Glu	Phe	Leu	Arg	Thr	His	Lys	His	Asp	Gly	Arg	Pro	Met	Met	Leu	Asn		
			180					185					190				
gat	cga	ata	cag	aat	atc	ccc	ata	ctt	cag	gga	gct	ttg	gca	aga	gca	624	
Asp	Arg	Ile	Gln	Asn	Ile	Pro	Ile	Leu	Gln	Gly	Ala	Leu	Ala	Arg	Ala		
		195				200						205					
gag	gag	ttc	ctt	tct	aaa	ctt	cct	ctg	gca	aca	cca	tac	tct	gaa	ttc	672	
Glu	Glu	Phe	Leu	Ser	Lys	Leu	Pro	Leu	Ala	Thr	Pro	Tyr	Ser	Glu	Phe		
	210					215					220						
gaa	ttt	gaa	cta	caa	ggg	atg	gga	ttt	gaa	agg	gga	tgg	ggg	gac	aca	720	
Glu	Phe	Glu	Leu	Gln	Gly	Met	Gly	Phe	Glu	Arg	Gly	Trp	Gly	Asp	Thr		
225					230				235					240			
gca	cag	aag	gtt	tca	gaa	atg	gtg	cat	ctt	ctt	ctg	gac	ata	ctc	cag	768	
Ala	Gln	Lys	Val	Ser	Glu	Met	Val	His	Leu	Leu	Leu	Asp	Ile	Leu	Gln		
				245					250					255			
gca	cct	gat	cct	tct	gtc	ttg	gag	acg	ttt	cta	gga	agg	att	cct	atg	816	
Ala	Pro	Asp	Pro	Ser	Val	Leu	Glu	Thr	Phe	Leu	Gly	Arg	Ile	Pro	Met		
			260					265					270				
gtg	ttc	aat	gtt	gtg	att	ttg	tct	ccg	cat	ggg	tac	ttt	ggc	caa	gcc	864	
Val	Phe	Asn	Val	Val	Ile	Leu	Ser	Pro	His	Gly	Tyr	Phe	Gly	Gln	Ala		
		275				280						285					
aat	gtc	ttg	ggg	ctg	cct	gat	act	ggg	gga	cag	gtt	gtc	tac	att	ctt	912	
Asn	Val	Leu	Gly	Leu	Pro	Asp	Thr	Gly	Gly	Gln	Val	Val	Tyr	Ile	Leu		
		290				295					300						
gat	caa	gta	cgt	gca	ttg	gaa	aat	gag	atg	ctc	ctt	agg	ata	cag	aag	960	
Asp	Gln	Val	Arg	Ala	Leu	Glu	Asn	Glu	Met	Leu	Leu	Arg	Ile	Gln	Lys		
305					310					315				320			
caa	gga	ctg	gaa	gtt	att	cca	aag	att	ctc	att	gta	aca	aga	ctg	cta	1008	
Gln	Gly	Leu	Glu	Val	Ile	Pro	Lys	Ile	Leu	Ile	Val	Thr	Arg	Leu	Leu		
				325					330					335			
ccc	gaa	gca	aag	gga	aca	acg	tgc	aac	cag	agg	tta	gaa	aga	gtt	agt	1056	
Pro	Glu	Ala	Lys	Gly	Thr	Thr	Cys	Asn	Gln	Arg	Leu	Glu	Arg	Val	Ser		
			340					345					350				
ggg	aca	gaa	cac	gca	cac	att	ctg	cga	ata	cca	ttt	agg	act	gaa	aag	1104	
Gly	Thr	Glu	His	Ala	His	Ile	Leu	Arg	Ile	Pro	Phe	Arg	Thr	Glu	Lys		
		355				360						365					
gga	att	ctt	cgc	aag	tgg	atc	tca	agg	ttt	gat	gtc	tgg	cca	tac	ctg	1152	
Gly	Ile	Leu	Arg	Lys	Trp	Ile	Ser	Arg	Phe	Asp	Val	Trp	Pro	Tyr	Leu		

370	375	380	
gag act ttt gca gag gat gca tca aat gaa att tct gcg gag ttg cag			1200
Glu Thr Phe Ala Glu Asp Ala Ser Asn Glu Ile Ser Ala Glu Leu Gln			
385	390	395	400
ggt gta cca aat ctc atc att ggc aac tac agt gat gga aat ctc gtt			1248
Gly Val Pro Asn Leu Ile Ile Gly Asn Tyr Ser Asp Gly Asn Leu Val			
	405	410	415
gct tct ttg tta gct agt aag cta ggt gtg ata cag tgt aat att gct			1296
Ala Ser Leu Leu Ala Ser Lys Leu Gly Val Ile Gln Cys Asn Ile Ala			
	420	425	430
cat gct tta gag aaa acc aag tac ccc gag tct gac att tac tgg aga			1344
His Ala Leu Glu Lys Thr Lys Tyr Pro Glu Ser Asp Ile Tyr Trp Arg			
	435	440	445
aac cat gaa gat aag tat cac ttt tca agt cag ttc act gca gat cta			1392
Asn His Glu Asp Lys Tyr His Phe Ser Ser Gln Phe Thr Ala Asp Leu			
	450	455	460
att gcc atg aat aat gcc gat ttc atc atc acc agc aca tac caa gag			1440
Ile Ala Met Asn Asn Ala Asp Phe Ile Ile Thr Ser Thr Tyr Gln Glu			
	465	470	475
att gcg gga agc aag aac aat gtt ggg caa tac gag agc cac aca gct			1488
Ile Ala Gly Ser Lys Asn Asn Val Gly Gln Tyr Glu Ser His Thr Ala			
	485	490	495
ttc act atg cct ggt ctt tac cga gtt gtt cat gga att gat gtc ttt			1536
Phe Thr Met Pro Gly Leu Tyr Arg Val Val His Gly Ile Asp Val Phe			
	500	505	510
gat cct aag ttt aat ata gtc tct cca gga gct gat atg acc ata tac			1584
Asp Pro Lys Phe Asn Ile Val Ser Pro Gly Ala Asp Met Thr Ile Tyr			
	515	520	525
ttt cca tat tct gac aag gaa aga aga ctc act gcc ctt cat gag tca			1632
Phe Pro Tyr Ser Asp Lys Glu Arg Arg Leu Thr Ala Leu His Glu Ser			
	530	535	540
att gaa gaa ctc ctc ttt agt gcc gaa cag aat gat gag cat gtt ggt			1680
Ile Glu Glu Leu Leu Phe Ser Ala Glu Gln Asn Asp Glu His Val Gly			
	545	550	555
tta ctg agc gac caa tcg aag cca atc atc ttc tct atg gca aga ctt			1728
Leu Leu Ser Asp Gln Ser Lys Pro Ile Ile Phe Ser Met Ala Arg Leu			
	565	570	575
gac agg gtg aaa aac ttg act ggg cta gtt gaa tgc tat gcc aag aat			1776
Asp Arg Val Lys Asn Leu Thr Gly Leu Val Glu Cys Tyr Ala Lys Asn			
	580	585	590
agc aag ctt aga gag ctt gca aat ctt gtt ata gtc ggt ggc tac atc			1824
Ser Lys Leu Arg Glu Leu Ala Asn Leu Val Ile Val Gly Gly Tyr Ile			
	595	600	605
gat gag aat cag tcc agg gat aga gag gaa atg gct gag ata caa aag			1872
Asp Glu Asn Gln Ser Arg Asp Arg Glu Glu Met Ala Glu Ile Gln Lys			
	610	615	620

atg cac agc ctg att gag cag tat gat tta cac ggt gag ttt agg tgg 1920
 Met His Ser Leu Ile Glu Gln Tyr Asp Leu His Gly Glu Phe Arg Trp
 625 630 635 640
 ata gct gct caa atg aac cgt gct cga aat ggt gag ctt tac cgt tat 1968
 Ile Ala Ala Gln Met Asn Arg Ala Arg Asn Gly Glu Leu Tyr Arg Tyr
 645 650 655
 atc gca gac aca aaa ggt gtt ttt gtt cag cct gct ttc tat gaa gca 2016
 Ile Ala Asp Thr Lys Gly Val Phe Val Gln Pro Ala Phe Tyr Glu Ala
 660 665 670
 ttt ggg ctt acg gtt gtg gaa tca atg act tgt gca ctc cca acg ttt 2064
 Phe Gly Leu Thr Val Val Glu Ser Met Thr Cys Ala Leu Pro Thr Phe
 675 680 685
 gct acc tgt cat ggt gga ccc gca gag att atc gaa aac gga gtt tct 2112
 Ala Thr Cys His Gly Gly Pro Ala Glu Ile Ile Glu Asn Gly Val Ser
 690 695 700
 ggg ttc cac att gac cca tat cat cca gac cag gtt gca gct acc ttg 2160
 Gly Phe His Ile Asp Pro Tyr His Pro Asp Gln Val Ala Ala Thr Leu
 705 710 715 720
 gtc agc ttc ttt gag acc tgt aac acc aat cca aat cat tgg gtt aaa 2208
 Val Ser Phe Phe Glu Thr Cys Asn Thr Asn Pro Asn His Trp Val Lys
 725 730 735
 atc tct gaa gga ggg ctc aag cga atc tat gaa agg tac aca tgg aag 2256
 Ile Ser Glu Gly Gly Leu Lys Arg Ile Tyr Glu Arg Tyr Thr Trp Lys
 740 745 750
 aag tac tca gag aga ctg ctt acc ctg gct gga gtc tat gca ttc tgg 2304
 Lys Tyr Ser Glu Arg Leu Leu Thr Leu Ala Gly Val Tyr Ala Phe Trp
 755 760 765
 aaa cat gtg tct aag ctc gaa agg aga gaa aca cga cgt tac cta gag 2352
 Lys His Val Ser Lys Leu Glu Arg Arg Glu Thr Arg Arg Tyr Leu Glu
 770 775 780
 atg ttt tac tca ttg aaa ttt cgt gat ttg gcc aat tca atc ccg ctg 2400
 Met Phe Tyr Ser Leu Lys Phe Arg Asp Leu Ala Asn Ser Ile Pro Leu
 785 790 795 800
 gca aca gat gag aac tga 2418
 Ala Thr Asp Glu Asn
 805

<210> 2
 <211> 805
 <212> PRT
 <213> Arabidopsis thaliana

<400> 2
 Met Ala Ser Phe Phe Asp Leu Val Tyr Ile Ser Trp Ile Gly Phe Leu
 1 5 10 15

Asn Lys Phe Leu Ser Phe Trp Val Phe Trp Val Cys Leu Val Arg Tyr

20	25	30
Val Ala Gln Gly Lys Gly Ile Leu Gln Ser His Gln Leu Ile Asp Glu		
35	40	45
Phe Leu Lys Thr Val Lys Val Asp Gly Thr Leu Glu Asp Leu Asn Lys		
50	55	60
Ser Pro Phe Met Lys Val Leu Gln Glu Ala Ile Val Leu Pro Pro Phe		
65	70	75
Val Ala Leu Ala Ile Arg Pro Arg Pro Gly Val Arg Glu Tyr Val Arg		
85	90	95
Val Asn Val Tyr Glu Leu Ser Val Asp His Leu Thr Val Ser Glu Tyr		
100	105	110
Leu Arg Phe Lys Glu Glu Leu Val Asn Gly His Ala Asn Gly Asp Tyr		
115	120	125
Leu Leu Glu Leu Asp Phe Glu Pro Phe Asn Ala Thr Leu Pro Arg Pro		
130	135	140
Thr Arg Ser Ser Ser Ile Gly Asn Gly Val Gln Phe Leu Asn Arg His		
145	150	155
Leu Ser Ser Ile Met Phe Arg Asn Lys Glu Ser Met Glu Pro Leu Leu		
165	170	175
Glu Phe Leu Arg Thr His Lys His Asp Gly Arg Pro Met Met Leu Asn		
180	185	190
Asp Arg Ile Gln Asn Ile Pro Ile Leu Gln Gly Ala Leu Ala Arg Ala		
195	200	205
Glu Glu Phe Leu Ser Lys Leu Pro Leu Ala Thr Pro Tyr Ser Glu Phe		
210	215	220
Glu Phe Glu Leu Gln Gly Met Gly Phe Glu Arg Gly Trp Gly Asp Thr		
225	230	235
Ala Gln Lys Val Ser Glu Met Val His Leu Leu Leu Asp Ile Leu Gln		
245	250	255
Ala Pro Asp Pro Ser Val Leu Glu Thr Phe Leu Gly Arg Ile Pro Met		
260	265	270

Val Phe Asn Val Val Ile Leu Ser Pro His Gly Tyr Phe Gly Gln Ala
 275 280 285

Asn Val Leu Gly Leu Pro Asp Thr Gly Gly Gln Val Val Tyr Ile Leu
 290 295 300

Asp Gln Val Arg Ala Leu Glu Asn Glu Met Leu Leu Arg Ile Gln Lys
 305 310 315 320

Gln Gly Leu Glu Val Ile Pro Lys Ile Leu Ile Val Thr Arg Leu Leu
 325 330 335

Pro Glu Ala Lys Gly Thr Thr Cys Asn Gln Arg Leu Glu Arg Val Ser
 340 345 350

Gly Thr Glu His Ala His Ile Leu Arg Ile Pro Phe Arg Thr Glu Lys
 355 360 365

Gly Ile Leu Arg Lys Trp Ile Ser Arg Phe Asp Val Trp Pro Tyr Leu
 370 375 380

Glu Thr Phe Ala Glu Asp Ala Ser Asn Glu Ile Ser Ala Glu Leu Gln
 385 390 395 400

Gly Val Pro Asn Leu Ile Ile Gly Asn Tyr Ser Asp Gly Asn Leu Val
 405 410 415

Ala Ser Leu Leu Ala Ser Lys Leu Gly Val Ile Gln Cys Asn Ile Ala
 420 425 430

His Ala Leu Glu Lys Thr Lys Tyr Pro Glu Ser Asp Ile Tyr Trp Arg
 435 440 445

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

Ile Ala Met Asn Asn Ala Asp Phe Ile Ile Thr Ser Thr Tyr Gln Glu
 465 470 475 480

Ile Ala Gly Ser Lys Asn Asn Val Gly Gln Tyr Glu Ser His Thr Ala
 485 490 495

Phe Thr Met Pro Gly Leu Tyr Arg Val Val His Gly Ile Asp Val Phe
 500 505 510

Asp Pro Lys Phe Asn Ile Val Ser Pro Gly Ala Asp Met Thr Ile Tyr
 515 520 525

Phe Pro Tyr Ser Asp Lys Glu Arg Arg Leu Thr Ala Leu His Glu Ser
 530 535 540

Ile Glu Glu Leu Leu Phe Ser Ala Glu Gln Asn Asp Glu His Val Gly
 545 550 555 560

Leu Leu Ser Asp Gln Ser Lys Pro Ile Ile Phe Ser Met Ala Arg Leu
 565 570 575

Asp Arg Val Lys Asn Leu Thr Gly Leu Val Glu Cys Tyr Ala Lys Asn
 580 585 590

Ser Lys Leu Arg Glu Leu Ala Asn Leu Val Ile Val Gly Gly Tyr Ile
 595 600 605

Asp Glu Asn Gln Ser Arg Asp Arg Glu Glu Met Ala Glu Ile Gln Lys
 610 615 620

Met His Ser Leu Ile Glu Gln Tyr Asp Leu His Gly Glu Phe Arg Trp
 625 630 635 640

Ile Ala Ala Gln Met Asn Arg Ala Arg Asn Gly Glu Leu Tyr Arg Tyr
 645 650 655

Ile Ala Asp Thr Lys Gly Val Phe Val Gln Pro Ala Phe Tyr Glu Ala
 660 665 670

Phe Gly Leu Thr Val Val Glu Ser Met Thr Cys Ala Leu Pro Thr Phe
 675 680 685

Ala Thr Cys His Gly Gly Pro Ala Glu Ile Ile Glu Asn Gly Val Ser
 690 695 700

Gly Phe His Ile Asp Pro Tyr His Pro Asp Gln Val Ala Ala Thr Leu
 705 710 715 720

Val Ser Phe Phe Glu Thr Cys Asn Thr Asn Pro Asn His Trp Val Lys
 725 730 735

Ile Ser Glu Gly Gly Leu Lys Arg Ile Tyr Glu Arg Tyr Thr Trp Lys
 740 745 750

Lys Tyr Ser Glu Arg Leu Leu Thr Leu Ala Gly Val Tyr Ala Phe Trp
 755 760 765

Lys His Val Ser Lys Leu Glu Arg Arg Glu Thr Arg Arg Tyr Leu Glu
 770 775 780

Met Phe Tyr Ser Leu Lys Phe Arg Asp Leu Ala Asn Ser Ile Pro Leu
 785 790 795 800

Ala Thr Asp Glu Asn
 805

<210> 3
 <211> 2668
 <212> DNA
 <213> Arabidopsis thaliana

<220>
 <221> Intron
 <222> (97)..(325)

<220>
 <221> CDS
 <222> (5)..(96)

<220>
 <221> CDS
 <222> (326)..(2657)

<400> 3
 aaca atg gcg tct ttc ttt gat ctc gta tac ata agc tgg att ggt ttc 49
 Met Ala Ser Phe Phe Asp Leu Val Tyr Ile Ser Trp Ile Gly Phe
 1 5 10 15
 tta aac aaa ttc ctc tcc ttt tgg gtc ttc tgg gtt tgc ctt gta ag 96
 Leu Asn Lys Phe Leu Ser Phe Trp Val Phe Trp Val Cys Leu Val Arg
 20 25 30
 tgttttgtgtt tttgcctctg agaaaaaatc atgccgactg gtaggttcga gactatgcgt 156
 gaatggggttt atgacgctat ctctgctcaa cgcaatgagc tcctctctct tttctccagg 216
 tttcttctta accctctcgc tctctctccc tgagccttgc ttataaatgc atgtgtgttg 276
 tgattggtct caagtcattg gtaatgggtt tggttggttat ctttattag a tat gta 332
 Tyr Val
 gcc cag gga aag ggg ata ttg cag tcc cac cag ctg att gat gag ttc 380
 Ala Gln Gly Lys Gly Ile Leu Gln Ser His Gln Leu Ile Asp Glu Phe
 35 40 45
 ctt aag act gtg aaa gtt gat gga aca tta gaa gat ctt aac aaa agt 428
 Leu Lys Thr Val Lys Val Asp Gly Thr Leu Glu Asp Leu Asn Lys Ser
 50 55 60 65
 cca ttc atg aaa gtt ctg cag tct gca gag gaa gcc ata gtt ttg cct 476

Pro Phe Met Lys Val Leu Gln Ser Ala Glu Glu Ala Ile Val Leu Pro	
70 75 80	
cca ttt gtt gct ttg gct ata cgt ccc aga cct ggt gtt agg gaa tat	524
Pro Phe Val Ala Leu Ala Ile Arg Pro Arg Pro Gly Val Arg Glu Tyr	
85 90 95	
gtc cgt gtg aat gtg tat gag ctg agc gta gat cat tta act gtt tct	572
Val Arg Val Asn Val Tyr Glu Leu Ser Val Asp His Leu Thr Val Ser	
100 105 110	
gaa tat ctt cgg ttt aag gaa gag ctc gtt aat ggc cat gcc aat gga	620
Glu Tyr Leu Arg Phe Lys Glu Glu Leu Val Asn Gly His Ala Asn Gly	
115 120 125	
gat tat ctc ctt gaa ctt gat ttt gaa cct ttc aat gca aca ttg cct	668
Asp Tyr Leu Leu Glu Leu Asp Phe Glu Pro Phe Asn Ala Thr Leu Pro	
130 135 140 145	
cgc cca act cgt tca tca tcc att ggg aat ggg gtt cag ttc ctc aat	716
Arg Pro Thr Arg Ser Ser Ser Ile Gly Asn Gly Val Gln Phe Leu Asn	
150 155 160	
cgt cac ctc tct tca att atg ttc cgt aac aaa gaa agc atg gag cct	764
Arg His Leu Ser Ser Ile Met Phe Arg Asn Lys Glu Ser Met Glu Pro	
165 170 175	
ttg ctt gag ttt ctc cgc act cac aaa cat gat ggc cgt cct atg atg	812
Leu Leu Glu Phe Leu Arg Thr His Lys His Asp Gly Arg Pro Met Met	
180 185 190	
ctg aat gat cga ata cag aat atc ccc ata ctt cag gga gct ttg gca	860
Leu Asn Asp Arg Ile Gln Asn Ile Pro Ile Leu Gln Gly Ala Leu Ala	
195 200 205	
aga gca gag gag ttc ctt tct aaa ctt cct ctg gca aca cca tac tct	908
Arg Ala Glu Glu Phe Leu Ser Lys Leu Pro Leu Ala Thr Pro Tyr Ser	
210 215 220 225	
gaa ttc gaa ttt gaa cta caa ggg atg gga ttt gaa agg gga tgg ggt	956
Glu Phe Glu Phe Glu Leu Gln Gly Met Gly Phe Glu Arg Gly Trp Gly	
230 235 240	
gac aca gca cag aag gtt tca gaa atg gtg cat ctt ctt ctg gac ata	1004
Asp Thr Ala Gln Lys Val Ser Glu Met Val His Leu Leu Leu Asp Ile	
245 250 255	
ctc cag gca cct gat cct tct gtc ttg gag acg ttt cta gga agg att	1052
Leu Gln Ala Pro Asp Pro Ser Val Leu Glu Thr Phe Leu Gly Arg Ile	
260 265 270	
cct atg gtg ttc aat gtt gtg att ttg tct ccg cat ggt tac ttt ggc	1100
Pro Met Val Phe Asn Val Val Ile Leu Ser Pro His Gly Tyr Phe Gly	
275 280 285	
caa gcc aat gtc ttg ggt ctg cct gat act ggt gga cag gtt gtc tac	1148
Gln Ala Asn Val Leu Gly Leu Pro Asp Thr Gly Gly Gln Val Val Tyr	
290 295 300 305	
att ctt gat caa gta cgt gca ttg gaa aat gag atg ctc ctt agg ata	1196
Ile Leu Asp Gln Val Arg Ala Leu Glu Asn Glu Met Leu Leu Arg Ile	

310	315	320	
cag aag caa gga ctg gaa gtt att cca aag att ctc att gta aca aga Gln Lys Gln Gly Leu Glu Val Ile Pro Lys Ile Leu Ile Val Thr Arg 325 330 335			1244
ctg cta ccc gaa gca aag gga aca acg tgc aac cag agg tta gaa aga Leu Leu Pro Glu Ala Lys Gly Thr Thr Cys Asn Gln Arg Leu Glu Arg 340 345 350			1292
gtt agt ggt aca gaa cac gca cac att ctg cga ata cca ttt agg act Val Ser Gly Thr Glu His Ala His Ile Leu Arg Ile Pro Phe Arg Thr 355 360 365			1340
gaa aag gga att ctt cgc aag tgg atc tca agg ttt gat gtc tgg cca Glu Lys Gly Ile Leu Arg Lys Trp Ile Ser Arg Phe Asp Val Trp Pro 370 375 380 385			1388
tac ctg gag act ttt gca gag gat gca tca aat gaa att tct gcg gag Tyr Leu Glu Thr Phe Ala Glu Asp Ala Ser Asn Glu Ile Ser Ala Glu 390 395 400			1436
ttg cag ggt gta cca aat ctc atc att ggc aac tac agt gat gga aat Leu Gln Gly Val Pro Asn Leu Ile Ile Gly Asn Tyr Ser Asp Gly Asn 405 410 415			1484
ctc gtt gct tct ttg tta gct agt aag cta ggt gtg ata cag tgt aat Leu Val Ala Ser Leu Leu Ala Ser Lys Leu Gly Val Ile Gln Cys Asn 420 425 430			1532
att gct cat gct tta gag aaa acc aag tac ccc gag tct gac att tac Ile Ala His Ala Leu Glu Lys Thr Lys Tyr Pro Gln Ser Asp Ile Tyr 435 440 445			1580
tgg aga aac cat gaa gat aag tat cac ttt tca agt cag ttc act gca Trp Arg Asn His Glu Asp Lys Tyr His Phe Ser Ser Gln Phe Thr Ala 450 455 460 465			1628
gat cta att gcc atg aat aat gcc gat ttc atc atc acc agc aca tac Asp Leu Ile Ala Met Asn Asn Ala Asp Phe Ile Ile Thr Ser Thr Tyr 470 475 480			1676
caa gag att gcg gga agc aag aac aat gtt ggg caa tac gag agc cac Gln Glu Ile Ala Gly Ser Lys Asn Asn Val Gly Gln Tyr Glu Ser His 485 490 495			1724
aca gct ttc act atg cct ggt ctt tac cga gtt gtt cat gga att gat Thr Ala Phe Thr Met Pro Gly Leu Tyr Arg Val Val His Gly Ile Asp 500 505 510			1772
gtc ttt gat cct aag ttt aat atg gtc tct cca gga gct gat atg acc Val Phe Asp Pro Lys Phe Asn Met Val Ser Pro Gly Ala Asp Met Thr 515 520 525			1820
ata tac ttt cca tat tcc gac aag gaa aga aga ctc act gcc ctt cat Ile Tyr Phe Pro Tyr Ser Asp Lys Glu Arg Arg Leu Thr Ala Leu His 530 535 540 545			1868
gag tca att gaa gaa ctc ctc ttt agt gcc gaa cag aat gat gag cat Glu Ser Ile Glu Glu Leu Leu Phe Ser Ala Glu Gln Asn Asp Glu His 550 555 560			1916

gtt ggt tta ctg agc gac caa tcg aag cca atc atc ttc tct atg gca	1964
Val Gly Leu Leu Ser Asp Gln Ser Lys Pro Ile Ile Phe Ser Met Ala	
565 570 575	
aga ctt gac agg gtg aaa aac ttg act ggg cta gtt gaa tgc tat gcc	2012
Arg Leu Asp Arg Val Lys Asn Leu Thr Gly Leu Val Glu Cys Tyr Ala	
580 585 590	
aag aat agc aag ctt aga gag ctt gca aat ctt gtt ata gtc ggt ggc	2060
Lys Asn Ser Lys Leu Arg Glu Leu Ala Asn Leu Val Ile Val Gly Gly	
595 600 605	
tac atc gat gag aat cag tcc agg gat aga gag gaa atg gct gag ata	2108
Tyr Ile Asp Glu Asn Gln Ser Arg Asp Arg Glu Glu Met Ala Glu Ile	
610 615 620 625	
caa aag atg cac agc ctg att gag cag tat gat tta cac ggt gag ttt	2156
Gln Lys Met His Ser Leu Ile Glu Gln Tyr Asp Leu His Gly Glu Phe	
630 635 640	
agg tgg ata gct gct caa atg aac cgt gct cga aat ggt gag ctt tac	2204
Arg Trp Ile Ala Ala Gln Met Asn Arg Ala Arg Asn Gly Glu Leu Tyr	
645 650 655	
cgt tat atc gca gac aca aaa ggt gtt ttt gtt cag cct gct ttc tat	2252
Arg Tyr Ile Ala Asp Thr Lys Gly Val Phe Val Gln Pro Ala Phe Tyr	
660 665 670	
gaa gca ttt ggg ctt acg gtt gtg gaa tca atg act tgt gca ctc cca	2300
Glu Ala Phe Gly Leu Thr Val Glu Ser Met Thr Cys Ala Leu Pro	
675 680 685	
acg ttt gct acc tgt cat ggt gga ccc gca gag att atc gaa aac gga	2348
Thr Phe Ala Thr Cys His Gly Gly Pro Ala Glu Ile Ile Glu Asn Gly	
690 695 700 705	
gtt tct ggg ttc cac att gac cca tat cat cca gac cag gtt gca gct	2396
Val Ser Gly Phe His Ile Asp Pro Tyr His Pro Asp Gln Val Ala Ala	
710 715 720	
acc ttg gtc agc ttc ttt gag acc tgt aac acc aat cca aat cat tgg	2444
Thr Leu Val Ser Phe Phe Glu Thr Cys Asn Thr Asn Pro Asn His Trp	
725 730 735	
gtt aaa atc tct gaa gga ggg ctc aag cga atc tat gaa agg tac aca	2492
Val Lys Ile Ser Glu Gly Gly Leu Lys Arg Ile Tyr Glu Arg Tyr Thr	
740 745 750	
tgg aag aag tac tca gag aga ctg ctt acc ctg gct gga gtc tat gca	2540
Trp Lys Lys Tyr Ser Glu Arg Leu Leu Thr Leu Ala Gly Val Tyr Ala	
755 760 765	
ttc tgg aaa cat gtg tct aag ctc gaa agg aga gaa aca cga cgt tac	2588
Phe Trp Lys His Val Ser Lys Leu Glu Arg Arg Glu Thr Arg Arg Tyr	
770 775 780 785	
cta gag atg ttt tac tca ttg aaa ttt cgt gat ttg gcc aat tca atc	2636
Leu Glu Met Phe Tyr Ser Leu Lys Phe Arg Asp Leu Ala Asn Ser Ile	
790 795 800	

ccg ctg gca aca gat gag aac tgattaatta a
 Pro Leu Ala Thr Asp Glu Asn
 805

2668

<210> 4
 <211> 808
 <212> PRT
 <213> Arabidopsis thaliana

<400> 4
 Met Ala Ser Phe Phe Asp Leu Val Tyr Ile Ser Trp Ile Gly Phe Leu
 1 5 10 15

Asn Lys Phe Leu Ser Phe Trp Val Phe Trp Val Cys Leu Val Arg Tyr
 20 25 30

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

Phe Leu Lys Thr Val Lys Val Asp Gly Thr Leu Glu Asp Leu Asn Lys
 50 55 60

Ser Pro Phe Met Lys Val Leu Gln Ser Ala Glu Glu Ala Ile Val Leu
 65 70 75 80

Pro Pro Phe Val Ala Leu Ala Ile Arg Pro Arg Pro Gly Val Arg Glu
 85 90 95

Tyr Val Arg Val Asn Val Tyr Glu Leu Ser Val Asp His Leu Thr Val
 100 105 110

Ser Glu Tyr Leu Arg Phe Lys Glu Glu Leu Val Asn Gly His Ala Asn
 115 120 125

Gly Asp Tyr Leu Leu Glu Leu Asp Phe Glu Pro Phe Asn Ala Thr Leu
 130 135 140

Pro Arg Pro Thr Arg Ser Ser Ser Ile Gly Asn Gly Val Gln Phe Leu
 145 150 155 160

Asn Arg His Leu Ser Ser Ile Met Phe Arg Asn Lys Glu Ser Met Glu
 165 170 175

Pro Leu Leu Glu Phe Leu Arg Thr His Lys His Asp Gly Arg Pro Met
 180 185 190

Met Leu Asn Asp Arg Ile Gln Asn Ile Pro Ile Leu Gln Gly Ala Leu

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

Tyr Trp Arg Asn His Glu Asp Lys Tyr His Phe Ser Ser Gln Phe Thr
 450 455 460

Ala Asp Leu Ile Ala Met Asn Asn Ala Asp Phe Ile Ile Thr Ser Thr
 465 470 475 480

Tyr Gln Glu Ile Ala Gly Ser Lys Asn Asn Val Gly Gln Tyr Glu Ser
 485 490 495

His Thr Ala Phe Thr Met Pro Gly Leu Tyr Arg Val Val His Gly Ile
 500 505 510

Asp Val Phe Asp Pro Lys Phe Asn Met Val Ser Pro Gly Ala Asp Met
 515 520 525

Thr Ile Tyr Phe Pro Tyr Ser Asp Lys Glu Arg Arg Leu Thr Ala Leu
 530 535 540

His Glu Ser Ile Glu Glu Leu Leu Phe Ser Ala Glu Gln Asn Asp Glu
 545 550 555 560

His Val Gly Leu Leu Ser Asp Gln Ser Lys Pro Ile Ile Phe Ser Met
 565 570 575

Ala Arg Leu Asp Arg Val Lys Asn Leu Thr Gly Leu Val Glu Cys Tyr
 580 585 590

Ala Lys Asn Ser Lys Leu Arg Glu Leu Ala Asn Leu Val Ile Val Gly
 595 600 605

Gly Tyr Ile Asp Glu Asn Gln Ser Arg Asp Arg Glu Glu Met Ala Glu
 610 615 620

Ile Gln Lys Met His Ser Leu Ile Glu Gln Tyr Asp Leu His Gly Glu
 625 630 635 640

Phe Arg Trp Ile Ala Ala Gln Met Asn Arg Ala Arg Asn Gly Glu Leu
 645 650 655

Tyr Arg Tyr Ile Ala Asp Thr Lys Gly Val Phe Val Gln Pro Ala Phe
 660 665 670

Tyr Glu Ala Phe Gly Leu Thr Val Val Glu Ser Met Thr Cys Ala Leu
 675 680 685

Pro Thr Phe Ala Thr Cys His Gly Gly Pro Ala Glu Ile Ile Glu Asn
690 695 700

Gly Val Ser Gly Phe His Ile Asp Pro Tyr His Pro Asp Gln Val Ala
705 710 715 720

Ala Thr Leu Val Ser Phe Phe Glu Thr Cys Asn Thr Asn Pro Asn His
725 730 735

Trp Val Lys Ile Ser Glu Gly Gly Leu Lys Arg Ile Tyr Glu Arg Tyr
740 745 750

Thr Trp Lys Lys Tyr Ser Glu Arg Leu Leu Thr Leu Ala Gly Val Tyr
755 760 765

Ala Phe Trp Lys His Val Ser Lys Leu Glu Arg Arg Glu Thr Arg Arg
770 775 780

Tyr Leu Glu Met Phe Tyr Ser Leu Lys Phe Arg Asp Leu Ala Asn Ser
785 790 795 800

Ile Pro Leu Ala Thr Asp Glu Asn
805

<210> 5
<211> 2656
<212> DNA
<213> Arabidopsis thaliana

<220>
<221> Intron
<222> (93)..(321)

<220>
<221> CDS
<222> (1)..(92)

<220>
<221> CDS
<222> (322)..(2653)

<400> 5
atg gcg tct ttc ttt gat ctc gta tac ata agc tgg att ggt ttc tta 48
Met Ala Ser Phe Phe Asp Leu Val Tyr Ile Ser Trp Ile Gly Phe Leu
1 5 10 15

aac aaa ttc ctc tcc ttt tgg gtc ttc tgg gtt tgc ctt gta ag 92
Asn Lys Phe Leu Ser Phe Trp Val Phe Trp Val Cys Leu Val Arg
20 25 30

tgttttgtgtt tttgcctctg agaaaaaatc atgccgactg gtaggttcga gactatgcgt 152

gaatggggttt atgacgctat ctctgctcaa cgcaatgagc tcctctctct tttctccagg	212
tttcttctta accctctcgc tctctctccc tgagccttgc ttataaatgc atgtgtgttg	272
tgattgggtct caagtcatgg gtaatgggtt tggttggttat ctttattag a tat gta Tyr Val	328
gcc cag gga aag ggg ata ttg cag tcc cac cag ctg att gat gag ttc Ala Gln Gly Lys Gly Ile Leu Gln Ser His Gln Leu Ile Asp Glu Phe 35 40 45	376
ctt aag act gtg aaa gtt gat gga aca tta gaa gat ctt aac aaa agt Leu Lys Thr Val Lys Val Asp Gly Thr Leu Glu Asp Leu Asn Lys Ser 50 55 60 65	424
cca ttc atg aaa gtt ctg cag tct gca gag gaa gcc ata gtt ttg cct Pro Phe Met Lys Val Leu Gln Ser Ala Glu Glu Ala Ile Val Leu Pro 70 75 80	472
cca ttt gtt gct ttg gct ata cgt ccc aga cct ggt gtt agg gaa tat Pro Phe Val Ala Leu Ala Ile Arg Pro Arg Pro Gly Val Arg Glu Tyr 85 90 95	520
gtc cgt gtg aat gtg tat gag ctg agc gta gat cat tta act gtt tct Val Arg Val Asn Val Tyr Glu Leu Ser Val Asp His Leu Thr Val Ser 100 105 110	568
gaa tat ctt cgg ttt aag gaa gag ctc gtt aat ggc cat gcc aat gga Glu Tyr Leu Arg Phe Lys Glu Glu Leu Val Asn Gly His Ala Asn Gly 115 120 125	616
gat tat ctc ctt gaa ctt gat ttt gaa cct ttc aat gca aca ttg cct Asp Tyr Leu Leu Glu Leu Asp Phe Glu Pro Phe Asn Ala Thr Leu Pro 130 135 140 145	664
cgc cca act cgt tca tca tcc att ggg aat ggg gtt cag ttc ctc aat Arg Pro Thr Arg Ser Ser Ser Ile Gly Asn Gly Val Gln Phe Leu Asn 150 155 160	712
cgt cac ctc tct tca att atg ttc cgt aac aaa gaa agc atg gag cct Arg His Leu Ser Ser Ile Met Phe Arg Asn Lys Glu Ser Met Glu Pro 165 170 175	760
ttg ctt gag ttt ctc cgc act cac aaa cat gat ggc cgt cct atg atg Leu Leu Glu Phe Leu Arg Thr His Lys His Asp Gly Arg Pro Met Met 180 185 190	808
ctg aat gat cga ata cag aat atc ccc ata ctt cag gga gct ttg gca Leu Asn Asp Arg Ile Gln Asn Ile Pro Ile Leu Gln Gly Ala Leu Ala 195 200 205	856
aga gca gag gag ttc ctt tct aaa ctt cct ctg gca aca cca tac tct Arg Ala Glu Glu Phe Leu Ser Lys Leu Pro Leu Ala Thr Pro Tyr Ser 210 215 220 225	904
gaa ttc gaa ttt gaa cta caa ggg atg gga ttt gaa agg gga tgg ggt Glu Phe Glu Phe Glu Leu Gln Gly Met Gly Phe Glu Arg Gly Trp Gly 230 235 240	952

gac aca gca cag aag gtt tca gaa atg gtg cat ctt ctt ctg gac ata Asp Thr Ala Gln Lys Val Ser Glu Met Val His Leu Leu Leu Asp Ile 245 250 255	1000
ctc cag gca cct gat cct tct gtc ttg gag acg ttt cta gga agg att Leu Gln Ala Pro Asp Pro Ser Val Leu Glu Thr Phe Leu Gly Arg Ile 260 265 270	1048
cct atg gtg ttc aat gtt gtg att ttg tct ccg cat ggt tac ttt ggc Pro Met Val Phe Asn Val Val Ile Leu Ser Pro His Gly Tyr Phe Gly 275 280 285	1096
caa gcc aat gtc ttg ggt ctg cct gat act ggt gga cag gtt gtc tac Gln Ala Asn Val Leu Gly Leu Pro Asp Thr Gly Gly Gln Val Val Tyr 290 295 300 305	1144
att ctt gat caa gta cgt gca ttg gaa aat gag atg ctc ctt agg ata Ile Leu Asp Gln Val Arg Ala Leu Glu Asn Glu Met Leu Leu Arg Ile 310 315 320	1192
cag aag caa gga ctg gaa gtt att cca aag att ctc att gta aca aga Gln Lys Gln Gly Leu Glu Val Ile Pro Lys Ile Leu Ile Val Thr Arg 325 330 335	1240
ctg cta ccc gaa gca aag gga aca acg tgc aac cag agg tta gaa aga Leu Leu Pro Glu Ala Lys Gly Thr Thr Cys Asn Gln Arg Leu Glu Arg 340 345 350	1288
gtt agt ggt aca gaa cac gca cac att ctg cga ata cca ttt agg act Val Ser Gly Thr Glu His Ala His Ile Leu Arg Ile Pro Phe Arg Thr 355 360 365	1336
gaa aag gga att ctt cgc aag tgg atc tca agg ttt gat gtc tgg cca Glu Lys Gly Ile Leu Arg Lys Trp Ile Ser Arg Phe Asp Val Trp Pro 370 375 380 385	1384
tac ctg gag act ttt gca gag gat gca tca aat gaa att tct gcg gag Tyr Leu Glu Thr Phe Ala Glu Asp Ala Ser Asn Glu Ile Ser Ala Glu 390 395 400	1432
ttg cag ggt gta cca aat ctc atc att ggc aac tac agt gat gga aat Leu Gln Gly Val Pro Asn Leu Ile Ile Gly Asn Tyr Ser Asp Gly Asn 405 410 415	1480
ctc gtt gct tct ttg tta gct agt aag cta ggt gtg ata cag tgt aat Leu Val Ala Ser Leu Leu Ala Ser Lys Leu Gly Val Ile Gln Cys Asn 420 425 430	1528
att gct cat gct tta gag aaa acc aag tac ccc gag tct gac att tac Ile Ala His Ala Leu Glu Lys Thr Lys Tyr Pro Glu Ser Asp Ile Tyr 435 440 445	1576
tgg aga aac cat gaa gat aag tat cac ttt tca agt cag ttc act gca Trp Arg Asn His Glu Asp Lys Tyr His Phe Ser Ser Gln Phe Thr Ala 450 455 460 465	1624
gat cta att gcc atg aat aat gcc gat ttc atc atc acc agc aca tac Asp Leu Ile Ala Met Asn Asn Ala Asp Phe Ile Ile Thr Ser Thr Tyr 470 475 480	1672
caa gag att gcg gga agc aag aac aat gtt ggg caa tac gag agc cac	1720

Gln	Glu	Ile	Ala	Gly	Ser	Lys	Asn	Asn	Val	Gly	Gln	Tyr	Glu	Ser	His	
			485					490					495			
aca	gct	ttc	act	atg	cct	ggc	ctt	tac	cga	gtt	gtt	cat	gga	att	gat	1768
Thr	Ala	Phe	Thr	Met	Pro	Gly	Leu	Tyr	Arg	Val	Val	His	Gly	Ile	Asp	
		500					505					510				
gtc	ttt	gat	cct	aag	ttt	aat	atg	gtc	tct	cca	gga	gct	gat	atg	acc	1816
Val	Phe	Asp	Pro	Lys	Phe	Asn	Met	Val	Ser	Pro	Gly	Ala	Asp	Met	Thr	
	515					520					525					
ata	tac	ttt	cca	tat	tcc	gac	aag	gaa	aga	aga	ctc	act	gcc	ctt	cat	1864
Ile	Tyr	Phe	Pro	Tyr	Ser	Asp	Lys	Glu	Arg	Arg	Leu	Thr	Ala	Leu	His	
530					535					540					545	
gag	tca	att	gaa	gaa	ctc	ctc	ttt	agt	gcc	gaa	cag	aat	gat	gag	cat	1912
Glu	Ser	Ile	Glu	Glu	Leu	Leu	Phe	Ser	Ala	Glu	Gln	Asn	Asp	Glu	His	
			550						555					560		
gtt	ggc	tta	ctg	agc	gac	caa	tcg	aag	cca	atc	atc	ttc	tct	atg	gca	1960
Val	Gly	Leu	Leu	Ser	Asp	Gln	Ser	Lys	Pro	Ile	Ile	Phe	Ser	Met	Ala	
		565						570					575			
aga	ctt	gac	agg	gtg	aaa	aac	ttg	act	ggg	cta	gtt	gaa	tgc	tat	gcc	2008
Arg	Leu	Asp	Arg	Val	Lys	Asn	Leu	Thr	Gly	Leu	Val	Glu	Cys	Tyr	Ala	
		580					585					590				
aag	aat	agc	aag	ctt	aga	gag	ctt	gca	aat	ctt	gtt	ata	gtc	ggc	ggc	2056
Lys	Asn	Ser	Lys	Leu	Arg	Glu	Leu	Ala	Asn	Leu	Val	Ile	Val	Gly	Gly	
	595					600					605					
tac	atc	gat	gag	aat	cag	tcc	agg	gat	aga	gag	gaa	atg	gct	gag	ata	2104
Tyr	Ile	Asp	Glu	Asn	Gln	Ser	Arg	Asp	Arg	Glu	Glu	Met	Ala	Glu	Ile	
610				615						620					625	
caa	aag	atg	cac	agc	ctg	att	gag	cag	tat	gat	tta	cac	ggc	gag	ttt	2152
Gln	Lys	Met	His	Ser	Leu	Ile	Glu	Gln	Tyr	Asp	Leu	His	Gly	Glu	Phe	
			630					635					640			
agg	tgg	ata	gct	gct	caa	atg	aac	cgt	gct	cga	aat	ggc	gag	ctt	tac	2200
Arg	Trp	Ile	Ala	Ala	Gln	Met	Asn	Arg	Ala	Arg	Asn	Gly	Glu	Leu	Tyr	
		645						650					655			
cgt	tat	atc	gca	gac	aca	aaa	ggc	gtt	ttt	gtt	cag	cct	gct	ttc	tat	2248
Arg	Tyr	Ile	Ala	Asp	Thr	Lys	Gly	Val	Phe	Val	Gln	Pro	Ala	Phe	Tyr	
		660					665					670				
gaa	gca	ttt	ggc	ctt	acg	gtt	gtg	gaa	tca	atg	act	tgt	gca	ctc	cca	2296
Glu	Ala	Phe	Gly	Leu	Thr	Val	Val	Glu	Ser	Met	Thr	Cys	Ala	Leu	Pro	
	675					680					685					
acg	ttt	gct	acc	tgt	cat	ggc	gga	ccc	gca	gag	att	atc	gaa	aac	gga	2344
Thr	Phe	Ala	Thr	Cys	His	Gly	Gly	Pro	Ala	Glu	Ile	Ile	Glu	Asn	Gly	
690					695					700					705	
gtt	tct	ggc	ttc	cac	att	gac	cca	tat	cat	cca	gac	cag	gtt	gca	gct	2392
Val	Ser	Gly	Phe	His	Ile	Asp	Pro	Tyr	His	Pro	Asp	Gln	Val	Ala	Ala	
			710						715					720		
acc	ttg	gtc	agc	ttc	ttt	gag	acc	tgt	aac	acc	aat	cca	aat	cat	tggt	2440
Thr	Leu	Val	Ser	Phe	Phe	Glu	Thr	Cys	Asn	Thr	Asn	Pro	Asn	His	Trp	

725	730	735	
gtt aaa atc tct gaa gga ggg ctc aag cga atc tat gaa agg tac aca			2488
Val Lys Ile Ser Glu Gly Gly Leu Lys Arg Ile Tyr Glu Arg Tyr Thr			
740	745	750	
tgg aag aag tac tca gag aga ctg ctt acc ctg gct gga gtc tat gca			2536
Trp Lys Lys Tyr Ser Glu Arg Leu Leu Thr Leu Ala Gly Val Tyr Ala			
755	760	765	
ttc tgg aaa cat gtg tct aag ctc gaa agg aga gaa aca cga cgt tac			2584
Phe Trp Lys His Val Ser Lys Leu Glu Arg Arg Glu Thr Arg Arg Tyr			
770	775	780	785
cta gag atg ttt tac tca ttg aaa ttt cgt gat ttg gcc aat tca atc			2632
Leu Glu Met Phe Tyr Ser Leu Lys Phe Arg Asp Leu Ala Asn Ser Ile			
790	795	800	
ccg ctg gca aca gat gag aac tga			2656
Pro Leu Ala Thr Asp Glu Asn			
805			
<210> 6			
<211> 808			
<212> PRT			
<213> Arabidopsis thaliana			
<400> 6			
Met Ala Ser Phe Phe Asp Leu Val Tyr Ile Ser Trp Ile Gly Phe Leu			
1	5	10	15
Asn Lys Phe Leu Ser Phe Trp Val Phe Trp Val Cys Leu Val Arg Tyr			
20	25	30	
Val Ala Gln Gly Lys Gly Ile Leu Gln Ser His Gln Leu Ile Asp Glu			
35	40	45	
Phe Leu Lys Thr Val Lys Val Asp Gly Thr Leu Glu Asp Leu Asn Lys			
50	55	60	
Ser Pro Phe Met Lys Val Leu Gln Ser Ala Glu Glu Ala Ile Val Leu			
65	70	75	80
Pro Pro Phe Val Ala Leu Ala Ile Arg Pro Arg Pro Gly Val Arg Glu			
85	90	95	
Tyr Val Arg Val Asn Val Tyr Glu Leu Ser Val Asp His Leu Thr Val			
100	105	110	
Ser Glu Tyr Leu Arg Phe Lys Glu Glu Leu Val Asn Gly His Ala Asn			
115	120	125	

Gly Asp Tyr Leu Leu Glu Leu Asp Phe Glu Pro Phe Asn Ala Thr Leu
 130 135 140

Pro Arg Pro Thr Arg Ser Ser Ser Ile Gly Asn Gly Val Gln Phe Leu
 145 150 155 160

Asn Arg His Leu Ser Ser Ile Met Phe Arg Asn Lys Glu Ser Met Glu
 165 170 175

Pro Leu Leu Glu Phe Leu Arg Thr His Lys His Asp Gly Arg Pro Met
 180 185 190

Met Leu Asn Asp Arg Ile Gln Asn Ile Pro Ile Leu Gln Gly Ala Leu
 195 200 205

Ala Arg Ala Glu Glu Phe Leu Ser Lys Leu Pro Leu Ala Thr Pro Tyr
 210 215 220

Ser Glu Phe Glu Phe Glu Leu Gln Gly Met Gly Phe Glu Arg Gly Trp
 225 230 235 240

Gly Asp Thr Ala Gln Lys Val Ser Glu Met Val His Leu Leu Leu Asp
 245 250 255

Ile Leu Gln Ala Pro Asp Pro Ser Val Leu Glu Thr Phe Leu Gly Arg
 260 265 270

Ile Pro Met Val Phe Asn Val Val Ile Leu Ser Pro His Gly Tyr Phe
 275 280 285

Gly Gln Ala Asn Val Leu Gly Leu Pro Asp Thr Gly Gly Gln Val Val
 290 295 300

Tyr Ile Leu Asp Gln Val Arg Ala Leu Glu Asn Glu Met Leu Leu Arg
 305 310 315 320

Ile Gln Lys Gln Gly Leu Glu Val Ile Pro Lys Ile Leu Ile Val Thr
 325 330 335

Arg Leu Leu Pro Glu Ala Lys Gly Thr Thr Cys Asn Gln Arg Leu Glu
 340 345 350

Arg Val Ser Gly Thr Glu His Ala His Ile Leu Arg Ile Pro Phe Arg
 355 360 365

Thr Glu Lys Gly Ile Leu Arg Lys Trp Ile Ser Arg Phe Asp Val Trp

370 375 380
 Pro Tyr Leu Glu Thr Phe Ala Glu Asp Ala Ser Asn Glu Ile Ser Ala
 385 390 395 400
 Glu Leu Gln Gly Val Pro Asn Leu Ile Ile Gly Asn Tyr Ser Asp Gly
 405 410 415
 Asn Leu Val Ala Ser Leu Leu Ala Ser Lys Leu Gly Val Ile Gln Cys
 420 425 430
 Asn Ile Ala His Ala Leu Glu Lys Thr Lys Tyr Pro Glu Ser Asp Ile
 435 440 445
 Tyr Trp Arg Asn His Glu Asp Lys Tyr His Phe Ser Ser Gln Phe Thr
 450 455 460
 Ala Asp Leu Ile Ala Met Asn Asn Ala Asp Phe Ile Ile Thr Ser Thr
 465 470 475 480
 Tyr Gln Glu Ile Ala Gly Ser Lys Asn Asn Val Gly Gln Tyr Glu Ser
 485 490 495
 His Thr Ala Phe Thr Met Pro Gly Leu Tyr Arg Val Val His Gly Ile
 500 505 510
 Asp Val Phe Asp Pro Lys Phe Asn Met Val Ser Pro Gly Ala Asp Met
 515 520 525
 Thr Ile Tyr Phe Pro Tyr Ser Asp Lys Glu Arg Arg Leu Thr Ala Leu
 530 535 540
 His Glu Ser Ile Glu Glu Leu Leu Phe Ser Ala Glu Gln Asn Asp Glu
 545 550 555 560
 His Val Gly Leu Leu Ser Asp Gln Ser Lys Pro Ile Ile Phe Ser Met
 565 570 575
 Ala Arg Leu Asp Arg Val Lys Asn Leu Thr Gly Leu Val Glu Cys Tyr
 580 585 590
 Ala Lys Asn Ser Lys Leu Arg Glu Leu Ala Asn Leu Val Ile Val Gly
 595 600 605
 Gly Tyr Ile Asp Glu Asn Gln Ser Arg Asp Arg Glu Glu Met Ala Glu
 610 615 620

Ile Gln Lys Met His Ser Leu Ile Glu Gln Tyr Asp Leu His Gly Glu
625 630 635 640

Phe Arg Trp Ile Ala Ala Gln Met Asn Arg Ala Arg Asn Gly Glu Leu
645 650 655

Tyr Arg Tyr Ile Ala Asp Thr Lys Gly Val Phe Val Gln Pro Ala Phe
660 665 670

Tyr Glu Ala Phe Gly Leu Thr Val Val Glu Ser Met Thr Cys Ala Leu
675 680 685

Pro Thr Phe Ala Thr Cys His Gly Gly Pro Ala Glu Ile Ile Glu Asn
690 695 700

Gly Val Ser Gly Phe His Ile Asp Pro Tyr His Pro Asp Gln Val Ala
705 710 715 720

Ala Thr Leu Val Ser Phe Phe Glu Thr Cys Asn Thr Asn Pro Asn His
725 730 735

Trp Val Lys Ile Ser Glu Gly Gly Leu Lys Arg Ile Tyr Glu Arg Tyr
740 745 750

Thr Trp Lys Lys Tyr Ser Glu Arg Leu Leu Thr Leu Ala Gly Val Tyr
755 760 765

Ala Phe Trp Lys His Val Ser Lys Leu Glu Arg Arg Glu Thr Arg Arg
770 775 780

Tyr Leu Glu Met Phe Tyr Ser Leu Lys Phe Arg Asp Leu Ala Asn Ser
785 790 795 800

Ile Pro Leu Ala Thr Asp Glu Asn
805

<210> 7
<211> 2794
<212> DNA
<213> Brassica napus

<220>
<221> CDS
<222> (93)..(2510)

<400> 7
cccggtcga ccatttcgtt ctactccagt ttccaagaaa aaaaaaaaaa actctgttat 60

cattgtttcc	tgcgctctctc	tggttaaggat	ca	atg	gtg	aac	ggg	gtg	tta	acg	113
				Met	Val	Asn	Gly	Val	Leu	Thr	
				1				5			
cgc gcc cac agc caa cgt gag cgt ttg aac gaa acg ctc gct gct caa	161										
Arg Ala His Ser Gln Arg Glu Arg Leu Asn Glu Thr Leu Ala Ala Gln											
	10			15				20			
aga aac gaa gtc ctt gcc tta ctt tcc agg gtt gaa ggc aaa ggt aaa	209										
Arg Asn Glu Val Leu Ala Leu Leu Ser Arg Val Glu Gly Lys Gly Lys											
	25			30				35			
ggc atc ttg caa caa aac cag atc att gct gaa ttc gag gct ttg cct	257										
Gly Ile Leu Gln Gln Asn Gln Ile Ile Ala Glu Phe Glu Ala Leu Pro											
	40			45				50			55
gaa gaa acc caa aaa aag att gaa ggt ggt gct ttc ttc gac ctt ctc	305										
Glu Glu Thr Gln Lys Lys Ile Glu Gly Gly Ala Phe Phe Asp Leu Leu											
		60					65				70
aaa acc act cag gaa gca ata gtg ttg cca cca tgg gtt gct ctt gct	353										
Lys Thr Thr Gln Glu Ala Ile Val Leu Pro Pro Trp Val Ala Leu Ala											
		75				80				85	
gtg agg cca agg cct ggt gtt tgg gag tac ata aga gtc aat ctc cat	401										
Val Arg Pro Arg Pro Gly Val Trp Glu Tyr Ile Arg Val Asn Leu His											
	90					95				100	
gct ctt gtt gtt gag gag ctt aca cct gct gag ttt ctt cat ttc aag	449										
Ala Leu Val Val Glu Glu Leu Thr Pro Ala Glu Phe Leu His Phe Lys											
	105					110				115	
gaa gag ctt gtt gat gga gtt aag aac ggt gat ttc acg ctt gag ctt	497										
Glu Glu Leu Val Asp Gly Val Lys Asn Gly Asp Phe Thr Leu Glu Leu											
	120				125			130			135
gac ttt gag ccg ttt aat gcc tct gtc cct cgc cca acg ctc ccc aag	545										
Asp Phe Glu Pro Phe Asn Ala Ser Val Pro Arg Pro Thr Leu Pro Lys											
			140					145			150
tac att gga gat ggt gtt gag ttt ctc aac cgt cac ctc tcg gct aag	593										
Tyr Ile Gly Asp Gly Val Glu Phe Leu Asn Arg His Leu Ser Ala Lys											
		155				160				165	
ctc ttc cat gac aag gag agt ttg ctt cct ttg ctc aag ttc ctt cga	641										
Leu Phe His Asp Lys Glu Ser Leu Leu Pro Leu Leu Lys Phe Leu Arg											
	170					175				180	
cta cac agc tac cag gga aag acc ctg atg ctg aac gag aga gtt caa	689										
Leu His Ser Tyr Gln Gly Lys Thr Leu Met Leu Asn Glu Arg Val Gln											
	185					190				195	
aac ctc aac aat ctg caa cac atc ttg agg aag gca gag gag tat ctg	737										
Asn Leu Asn Asn Leu Gln His Ile Leu Arg Lys Ala Glu Glu Tyr Leu											
	200				205			210			215
gca ggt ctt tca cct gaa acg cct tat gaa gat ttt gag gcc aag ttt	785										
Ala Gly Leu Ser Pro Glu Thr Pro Tyr Glu Asp Phe Glu Ala Lys Phe											
			220				225				230

gag gag att ggt ctt gag agg gga tgg gga aac aat gcc gag cgt gtc Glu Glu Ile Gly Leu Glu Arg Gly Trp Gly Asn Asn Ala Glu Arg Val 235 240 245	833
ctt gac atg atc cgt ctt ctt ttg gac ctc ctc gag gcc cct gac cct Leu Asp Met Ile Arg Leu Leu Leu Asp Leu Leu Glu Ala Pro Asp Pro 250 255 260	881
tgc act cta gag aac ttt ctt ggg aga gtc cct atg gtg ttc aac gtt Cys Thr Leu Glu Asn Phe Leu Gly Arg Val Pro Met Val Phe Asn Val 265 270 275	929
gtc atc ctc tct cca cat ggt tac ttt gct cag gac aat gtt ctt ggt Val Ile Leu Ser Pro His Gly Tyr Phe Ala Gln Asp Asn Val Leu Gly 280 285 290 295	977
tac cct gac act ggt ggt cag gtt gtt tac att ctg gat caa gtc cgt Tyr Pro Asp Thr Gly Gly Gln Val Val Tyr Ile Leu Asp Gln Val Arg 300 305 310	1025
gcc atg gag aca gag atg ctc caa cgt att cag cag caa gga ctc aac Ala Met Glu Thr Glu Met Leu Gln Arg Ile Gln Gln Gln Gly Leu Asn 315 320 325	1073
att act cca aga att ctc att ctc act aga ctg ctc cct gat gca gta Ile Thr Pro Arg Ile Leu Ile Leu Thr Arg Leu Leu Pro Asp Ala Val 330 335 340	1121
gga acc aca tgt ggt gag cgt ctt gag aga gtt gat gga tct gag tac Gly Thr Thr Cys Gly Glu Arg Leu Glu Arg Val Asp Gly Ser Glu Tyr 345 350 355	1169
tgt gac atc ctc cgt gtg ccc ttc aga aca gag aag ggt atc gtt cgc Cys Asp Ile Leu Arg Val Pro Phe Arg Thr Glu Lys Gly Ile Val Arg 360 365 370 375	1217
aaa tgg atc tca aga ttc gaa gtc tgg cca tat cta gag act tac acc Lys Trp Ile Ser Arg Phe Glu Val Trp Pro Tyr Leu Glu Thr Tyr Thr 380 385 390	1265
gag gat gct gcc gtt gag ctt gct aaa gag ttg aag ggc aag cct gac Glu Asp Ala Ala Val Glu Leu Ala Lys Glu Leu Lys Gly Lys Pro Asp 395 400 405	1313
ctt atc att ggt aac tac agt gat ggc aac ctc gtt gcc tct tta ctg Leu Ile Ile Gly Asn Tyr Ser Asp Gly Asn Leu Val Ala Ser Leu Leu 410 415 420	1361
gca cac aaa ctt ggt gtc act cag tgt acc att gct cac gcc ctg gag Ala His Lys Leu Gly Val Thr Gln Cys Thr Ile Ala His Ala Leu Glu 425 430 435	1409
aag aca aag tac cct gac tcg gat ata tac tgg aag aag ctt gac gat Lys Thr Lys Tyr Pro Asp Ser Asp Ile Tyr Trp Lys Lys Leu Asp Asp 440 445 450 455	1457
aag tac cat ttc tca tgc cag ttc act gca gat cta ttt gca atg aac Lys Tyr His Phe Ser Cys Gln Phe Thr Ala Asp Leu Phe Ala Met Asn 460 465 470	1505
cat act gat ttc atc atc aca agt act ctc caa gaa att gct gga agc	1553

His	Thr	Asp	Phe	Ile	Ile	Thr	Ser	Thr	Leu	Gln	Glu	Ile	Ala	Gly	Ser	
			475					480					485			
aag	gac	aca	gtt	ggg	cag	tat	gag	agc	cac	acg	gcc	ttt	act	ctt	cct	1601
Lys	Asp	Thr	Val	Gly	Gln	Tyr	Glu	Ser	His	Thr	Ala	Phe	Thr	Leu	Pro	
			490				495					500				
gga	ctg	tac	cga	gtt	gtt	cat	ggc	att	gat	gtg	ttt	gat	ccc	aag	ttc	1649
Gly	Leu	Tyr	Arg	Val	Val	His	Gly	Ile	Asp	Val	Phe	Asp	Pro	Lys	Phe	
	505					510					515					
aac	att	gtc	tct	cct	ggg	gct	gac	atg	agc	atc	tac	ttc	cca	tac	act	1697
Asn	Ile	Val	Ser	Pro	Gly	Ala	Asp	Met	Ser	Ile	Tyr	Phe	Pro	Tyr	Thr	
520					525					530					535	
gag	gag	aag	cgt	aga	ttg	act	aag	ttc	cac	cca	gag	att	gag	gag	ctc	1745
Glu	Glu	Lys	Arg	Arg	Leu	Thr	Lys	Phe	His	Pro	Glu	Ile	Glu	Glu	Leu	
			540					545					550			
ctc	tac	agc	aat	gtt	gaa	aac	gaa	gag	cac	tta	tgt	gtg	ctc	aag	gac	1793
Leu	Tyr	Ser	Asn	Val	Glu	Asn	Glu	Glu	His	Leu	Cys	Val	Leu	Lys	Asp	
			555				560						565			
aag	aag	aag	cca	atc	ctc	ttc	aca	atg	gct	agg	ctc	gac	cgt	gtc	aag	1841
Lys	Lys	Lys	Pro	Ile	Leu	Phe	Thr	Met	Ala	Arg	Leu	Asp	Arg	Val	Lys	
			570				575					580				
aac	ttg	tca	ggg	ctt	gtt	gag	tgg	tac	ggg	aag	aac	aaa	cgc	ctg	cgt	1889
Asn	Leu	Ser	Gly	Leu	Val	Glu	Trp	Tyr	Gly	Lys	Asn	Lys	Arg	Leu	Arg	
	585					590					595					
gag	ctt	gtt	aac	ttg	gtg	gtt	gtt	gga	gga	gac	agg	agg	aaa	gag	tca	1937
Glu	Leu	Val	Asn	Leu	Val	Val	Val	Gly	Gly	Asp	Arg	Arg	Lys	Glu	Ser	
600					605					610					615	
aag	gac	aac	gaa	gag	aag	gcc	gag	atg	aag	aaa	atg	tat	gac	ctc	att	1985
Lys	Asp	Asn	Glu	Glu	Lys	Ala	Glu	Met	Lys	Lys	Met	Tyr	Asp	Leu	Ile	
			620					625					630			
gag	gaa	tac	aag	ctc	aac	ggc	cag	ttc	cgt	tgg	atc	tcc	tcc	cag	atg	2033
Glu	Glu	Tyr	Lys	Leu	Asn	Gly	Gln	Phe	Arg	Trp	Ile	Ser	Ser	Gln	Met	
			635				640						645			
aac	agg	gtt	agg	aac	ggg	gag	ctc	tac	agg	tac	atc	tgt	gac	acc	aag	2081
Asn	Arg	Val	Arg	Asn	Gly	Glu	Leu	Tyr	Arg	Tyr	Ile	Cys	Asp	Thr	Lys	
		650				655						660				
ggg	gct	ttt	gtg	caa	cca	gca	ttg	tat	gaa	gcc	ttt	ggg	ttg	act	gtt	2129
Gly	Ala	Phe	Val	Gln	Pro	Ala	Leu	Tyr	Glu	Ala	Phe	Gly	Leu	Thr	Val	
	665					670					675					
gtg	gag	gct	atg	acc	tgt	ggg	tta	cca	aca	ttc	gcc	act	tgc	aaa	ggg	2177
Val	Glu	Ala	Met	Thr	Cys	Gly	Leu	Pro	Thr	Phe	Ala	Thr	Cys	Lys	Gly	
680					685					690					695	
ggg	cca	gct	gag	atc	att	gtc	cac	ggc	aaa	tca	ggg	ttc	cac	att	gac	2225
Gly	Pro	Ala	Glu	Ile	Ile	Val	His	Gly	Lys	Ser	Gly	Phe	His	Ile	Asp	
			700					705					710			
cct	tac	cat	ggg	gat	cag	gcg	gct	gat	act	ctt	gct	gat	ttc	ttc	acc	2273
Pro	Tyr	His	Gly	Asp	Gln	Ala	Ala	Asp	Thr	Leu	Ala	Asp	Phe	Phe	Thr	

715	720	725	
aag tgt aag gaa gat cca tct cac tgg gac cag atc tca aag gga ggg			2321
Lys Cys Lys Glu Asp Pro Ser His Trp Asp Gln Ile Ser Lys Gly Gly			
730	735	740	
ctt cag agg att gag gag aaa tac acg tgg cag att tac tca cag agg			2369
Leu Gln Arg Ile Glu Glu Lys Tyr Thr Trp Gln Ile Tyr Ser Gln Arg			
745	750	755	
ctc ttg aca ttg act ggt gtg tat gga ttc tgg aag cat gtc tcg aac			2417
Leu Leu Thr Leu Thr Gly Val Tyr Gly Phe Trp Lys His Val Ser Asn			
760	765	770	775
ctt gac cgt ctt gag agc cgt cgt tac ctt gag atg ttc tat gca ttg			2465
Leu Asp Arg Leu Glu Ser Arg Arg Tyr Leu Glu Met Phe Tyr Ala Leu			
780	785	790	
aag tat cgc cca ctg gct cag gct gtt cct ctt gct caa gag taa			2510
Lys Tyr Arg Pro Leu Ala Gln Ala Val Pro Leu Ala Gln Glu			
795	800	805	
tgaagaaaag agagaaaaaa tgtggaacca ttgaagagct tgaagataat ggttccagtg			2570
tttgaagaat atcaaaaatg tcttttgatt cctatcatca tttggattta ggagcgcttt			2630
tgtgttccctc ttgcttactt tggtgtttca atccttttgt ctttgtgtat tttcttttcc			2690
ttgtcttttcc ggcaatgggt tgtactctga ttgttgaggg atctccctta tgattggcaa			2750
tgagatatcg attattttat cagcagatcg tcgacccgga aatt			2794

<210> 8
 <211> 805
 <212> PRT
 <213> Brassica napus

<400> 8

Met Val Asn Gly Val Leu Thr Arg Ala His Ser Gln Arg Glu Arg Leu
 1 5 10 15

Asn Glu Thr Leu Ala Ala Gln Arg Asn Glu Val Leu Ala Leu Leu Ser
 20 25 30

Arg Val Glu Gly Lys Gly Lys Gly Ile Leu Gln Gln Asn Gln Ile Ile
 35 40 45

Ala Glu Phe Glu Ala Leu Pro Glu Glu Thr Gln Lys Lys Ile Glu Gly
 50 55 60

Gly Ala Phe Phe Asp Leu Leu Lys Thr Thr Gln Glu Ala Ile Val Leu
 65 70 75 80

Pro Pro Trp Val Ala Leu Ala Val Arg Pro Arg Pro Gly Val Trp Glu

85	90	95
Tyr Ile Arg Val Asn Leu His Ala Leu Val Val Glu Glu Leu Thr Pro 100 105 110		
Ala Glu Phe Leu His Phe Lys Glu Glu Leu Val Asp Gly Val Lys Asn 115 120 125		
Gly Asp Phe Thr Leu Glu Leu Asp Phe Glu Pro Phe Asn Ala Ser Val 130 135 140		
Pro Arg Pro Thr Leu Pro Lys Tyr Ile Gly Asp Gly Val Glu Phe Leu 145 150 155 160		
Asn Arg His Leu Ser Ala Lys Leu Phe His Asp Lys Glu Ser Leu Leu 165 170 175		
Pro Leu Leu Lys Phe Leu Arg Leu His Ser Tyr Gln Gly Lys Thr Leu 180 185 190		
Met Leu Asn Glu Arg Val Gln Asn Leu Asn Asn Leu Gln His Ile Leu 195 200 205		
Arg Lys Ala Glu Glu Tyr Leu Ala Gly Leu Ser Pro Glu Thr Pro Tyr 210 215 220		
Glu Asp Phe Glu Ala Lys Phe Glu Glu Ile Gly Leu Glu Arg Gly Trp 225 230 235 240		
Gly Asn Asn Ala Glu Arg Val Leu Asp Met Ile Arg Leu Leu Leu Asp 245 250 255		
Leu Leu Glu Ala Pro Asp Pro Cys Thr Leu Glu Asn Phe Leu Gly Arg 260 265 270		
Val Pro Met Val Phe Asn Val Val Ile Leu Ser Pro His Gly Tyr Phe 275 280 285		
Ala Gln Asp Asn Val Leu Gly Tyr Pro Asp Thr Gly Gly Gln Val Val 290 295 300		
Tyr Ile Leu Asp Gln Val Arg Ala Met Glu Thr Glu Met Leu Gln Arg 305 310 315 320		
Ile Gln Gln Gln Gly Leu Asn Ile Thr Pro Arg Ile Leu Ile Leu Thr 325 330 335		

Arg Leu Leu Pro Asp Ala Val Gly Thr Thr Cys Gly Glu Arg Leu Glu
340 345 350

Arg Val Asp Gly Ser Glu Tyr Cys Asp Ile Leu Arg Val Pro Phe Arg
355 360 365

Thr Glu Lys Gly Ile Val Arg Lys Trp Ile Ser Arg Phe Glu Val Trp
370 375 380

Pro Tyr Leu Glu Thr Tyr Thr Glu Asp Ala Ala Val Glu Leu Ala Lys
385 390 395 400

Glu Leu Lys Gly Lys Pro Asp Leu Ile Ile Gly Asn Tyr Ser Asp Gly
405 410 415

Asn Leu Val Ala Ser Leu Leu Ala His Lys Leu Gly Val Thr Gln Cys
420 425 430

Thr Ile Ala His Ala Leu Glu Lys Thr Lys Tyr Pro Asp Ser Asp Ile
435 440 445

Tyr Trp Lys Lys Leu Asp Asp Lys Tyr His Phe Ser Cys Gln Phe Thr
450 455 460

Ala Asp Leu Phe Ala Met Asn His Thr Asp Phe Ile Ile Thr Ser Thr
465 470 475 480

Leu Gln Glu Ile Ala Gly Ser Lys Asp Thr Val Gly Gln Tyr Glu Ser
485 490 495

His Thr Ala Phe Thr Leu Pro Gly Leu Tyr Arg Val Val His Gly Ile
500 505 510

Asp Val Phe Asp Pro Lys Phe Asn Ile Val Ser Pro Gly Ala Asp Met
515 520 525

Ser Ile Tyr Phe Pro Tyr Thr Glu Glu Lys Arg Arg Leu Thr Lys Phe
530 535 540

His Pro Glu Ile Glu Glu Leu Leu Tyr Ser Asn Val Glu Asn Glu Glu
545 550 555 560

His Leu Cys Val Leu Lys Asp Lys Lys Lys Pro Ile Leu Phe Thr Met
565 570 575

Ala Arg Leu Asp Arg Val Lys Asn Leu Ser Gly Leu Val Glu Trp Tyr
 580 585 590

Gly Lys Asn Lys Arg Leu Arg Glu Leu Val Asn Leu Val Val Val Gly
 595 600 605

Gly Asp Arg Arg Lys Glu Ser Lys Asp Asn Glu Glu Lys Ala Glu Met
 610 615 620

Lys Lys Met Tyr Asp Leu Ile Glu Glu Tyr Lys Leu Asn Gly Gln Phe
 625 630 635 640

Arg Trp Ile Ser Ser Gln Met Asn Arg Val Arg Asn Gly Glu Leu Tyr
 645 650 655

Arg Tyr Ile Cys Asp Thr Lys Gly Ala Phe Val Gln Pro Ala Leu Tyr
 660 665 670

Glu Ala Phe Gly Leu Thr Val Val Glu Ala Met Thr Cys Gly Leu Pro
 675 680 685

Thr Phe Ala Thr Cys Lys Gly Gly Pro Ala Glu Ile Ile Val His Gly
 690 695 700

Lys Ser Gly Phe His Ile Asp Pro Tyr His Gly Asp Gln Ala Ala Asp
 705 710 715 720

Thr Leu Ala Asp Phe Phe Thr Lys Cys Lys Glu Asp Pro Ser His Trp
 725 730 735

Asp Gln Ile Ser Lys Gly Gly Leu Gln Arg Ile Glu Glu Lys Tyr Thr
 740 745 750

Trp Gln Ile Tyr Ser Gln Arg Leu Leu Thr Leu Thr Gly Val Tyr Gly
 755 760 765

Phe Trp Lys His Val Ser Asn Leu Asp Arg Leu Glu Ser Arg Arg Tyr
 770 775 780

Leu Glu Met Phe Tyr Ala Leu Lys Tyr Arg Pro Leu Ala Gln Ala Val
 785 790 795 800

Pro Leu Ala Gln Glu
 805

<210> 9
 <211> 2415
 <212> DNA
 <213> Brassica napus

 <220>
 <221> CDS
 <222> (1)..(2415)

 <400> 9
 atg gtg aac ggg gtg tta acg cgc gcc cac agc caa cgt gag cgt ttg 48
 Met Val Asn Gly Val Leu Thr Arg Ala His Ser Gln Arg Glu Arg Leu
 1 5 10 15

 aac gaa acg ctc gct gct caa aga aac gaa gtc ctt gcc tta ctt tcc 96
 Asn Glu Thr Leu Ala Ala Gln Arg Asn Glu Val Leu Ala Leu Leu Ser
 20 25 30

 agg gtt gaa ggc aaa ggt aaa ggc atc ttg caa caa aac cag atc att 144
 Arg Val Glu Gly Lys Gly Lys Gly Ile Leu Gln Gln Asn Gln Ile Ile
 35 40 45

 gct gaa ttc gag gct ttg cct gaa gaa acc caa aaa aag att gaa ggt 192
 Ala Glu Phe Glu Ala Leu Pro Glu Glu Thr Gln Lys Lys Ile Glu Gly
 50 55 60

 ggt gct ttc ttc gac ctt ctc aaa acc act cag gaa gca ata gtg ttg 240
 Gly Ala Phe Phe Asp Leu Leu Lys Thr Thr Gln Glu Ala Ile Val Leu
 65 70 75 80

 cca cca tgg gtt gct ctt gct gtg agg cca agg cct ggt gtt tgg gag 288
 Pro Pro Trp Val Ala Leu Ala Val Arg Pro Arg Pro Gly Val Trp Glu
 85 90 95

 tac ata aga gtc aat ctc cat gct ctt gtt gtt gag gag ctt aca cct 336
 Tyr Ile Arg Val Asn Leu His Ala Leu Val Val Glu Glu Leu Thr Pro
 100 105 110

 gct gag ttt ctt cat ttc aag gaa gag ctt gtt gat gga gtt aag aac 384
 Ala Glu Phe Leu His Phe Lys Glu Glu Leu Val Asp Gly Val Lys Asn
 115 120 125

 ggt gat ttc acg ctt gag ctt gac ttt gag ccg ttt aat gcc tct gtc 432
 Gly Asp Phe Thr Leu Glu Leu Asp Phe Glu Pro Phe Asn Ala Ser Val
 130 135 140

 cct cgc cca acg ctc ccc aag tac att gga gat ggt gtt gag ttt ctc 480
 Pro Arg Pro Thr Leu Pro Lys Tyr Ile Gly Asp Gly Val Glu Phe Leu
 145 150 155 160

 aac cgt cac ctc tcg gct aag ctc ttc cat gac aag gag agt ttg ctt 528
 Asn Arg His Leu Ser Ala Lys Leu Phe His Asp Lys Glu Ser Leu Leu
 165 170 175

 cct ttg ctc aag ttc ctt cga cta cac agc tac cag gga aag acc ctg 576
 Pro Leu Leu Lys Phe Leu Arg Leu His Ser Tyr Gln Gly Lys Thr Leu
 180 185 190

atg ctg aac gag aga gtt caa aac ctc aac aat ctg caa cac atc ttg	624
Met Leu Asn Glu Arg Val Gln Asn Leu Asn Asn Leu Gln His Ile Leu	
195 200 205	
agg aag gca gag gag tat ctg gca ggt ctt tca cct gaa acg cct tat	672
Arg Lys Ala Glu Glu Tyr Leu Ala Gly Leu Ser Pro Glu Thr Pro Tyr	
210 215 220	
gaa gat ttt gag gcc aag ttt gag gag att ggt ctt gag agg gga tgg	720
Glu Asp Phe Glu Ala Lys Phe Glu Glu Ile Gly Leu Glu Arg Gly Trp	
225 230 235 240	
gga aac aat gcc gag cgt gtc ctt gac atg atc cgt ctt ctt ttg gac	768
Gly Asn Asn Ala Glu Arg Val Leu Asp Met Ile Arg Leu Leu Leu Asp	
245 250 255	
ctc ctc gag gcc cct gac cct tgc act cta gag aac ttt ctt ggg aga	816
Leu Leu Glu Ala Pro Asp Pro Cys Thr Leu Glu Asn Phe Leu Gly Arg	
260 265 270	
gtc cct atg gtg ttc aac gtt gtc atc ctc tct cca cat ggt tac ttt	864
Val Pro Met Val Phe Asn Val Val Ile Leu Ser Pro His Gly Tyr Phe	
275 280 285	
gct cag gac aat gtt ctt ggt tac cct gac act ggt ggt cag gtt gtt	912
Ala Gln Asp Asn Val Leu Gly Tyr Pro Asp Thr Gly Gly Gln Val Val	
290 295 300	
tac att ctg gat caa gtc cgt gcc atg gag aca gag atg ctc caa cgt	960
Tyr Ile Leu Asp Gln Val Arg Ala Met Glu Thr Glu Met Leu Gln Arg	
305 310 315 320	
att cag cag caa gga ctc aac att act cca aga att ctc att ctc act	1008
Ile Gln Gln Gln Gly Leu Asn Ile Thr Pro Arg Ile Leu Ile Leu Thr	
325 330 335	
aga ctg ctc cct gat gca gta gga acc aca tgt ggt gag cgt ctt gag	1056
Arg Leu Leu Pro Asp Ala Val Gly Thr Thr Cys Gly Glu Arg Leu Glu	
340 345 350	
aga gtt gat gga tct gag tac tgt gac atc ctc cgt gtg ccc ttc aga	1104
Arg Val Asp Gly Ser Glu Tyr Cys Asp Ile Leu Arg Val Pro Phe Arg	
355 360 365	
aca gag aag ggt atc gtt cgc aaa tgg atc tca aga ttc gaa gtc tgg	1152
Thr Glu Lys Gly Ile Val Arg Lys Trp Ile Ser Arg Phe Glu Val Trp	
370 375 380	
cca tat cta gag act tac acc gag gat gct gcc gtt gag ctt gct aaa	1200
Pro Tyr Leu Glu Thr Tyr Thr Glu Asp Ala Ala Val Glu Leu Ala Lys	
385 390 395 400	
gag ttg aag ggc aag cct gac ctt atc att ggt aac tac agt gat ggc	1248
Glu Leu Lys Gly Lys Pro Asp Leu Ile Ile Gly Asn Tyr Ser Asp Gly	
405 410 415	
aac ctc gtt gcc tct tta ctg gca cac aaa ctt ggt gtc act cag tgt	1296
Asn Leu Val Ala Ser Leu Leu Ala His Lys Leu Gly Val Thr Gln Cys	
420 425 430	
acc att gct cac gcc ctg gag aag aca aag tac cct gac tcg gat ata	1344

Thr	Ile	Ala	His	Ala	Leu	Glu	Lys	Thr	Lys	Tyr	Pro	Asp	Ser	Asp	Ile	
		435					440					445				
tac	tgg	aag	aag	ctt	gac	gat	aag	tac	cat	ttc	tca	tgc	cag	ttc	act	1392
Tyr	Trp	Lys	Lys	Leu	Asp	Asp	Lys	Tyr	His	Phe	Ser	Cys	Gln	Phe	Thr	
	450						455				460					
gca	gat	cta	ttt	gca	atg	aac	cat	act	gat	ttc	atc	atc	aca	agt	act	1440
Ala	Asp	Leu	Phe	Ala	Met	Asn	His	Thr	Asp	Phe	Ile	Ile	Thr	Ser	Thr	
465					470					475					480	
ctc	caa	gaa	att	gct	gga	agc	aag	gac	aca	gtt	ggg	cag	tat	gag	agc	1488
Leu	Gln	Glu	Ile	Ala	Gly	Ser	Lys	Asp	Thr	Val	Gly	Gln	Tyr	Glu	Ser	
				485					490					495		
cac	acg	gcc	ttt	act	ctt	cct	gga	ctg	tac	cga	gtt	gtt	cat	ggc	att	1536
His	Thr	Ala	Phe	Thr	Leu	Pro	Gly	Leu	Tyr	Arg	Val	Val	His	Gly	Ile	
		500						505					510			
gat	gtg	ttt	gat	ccc	aag	ttc	aac	att	gtc	tct	cct	ggg	gct	gac	atg	1584
Asp	Val	Phe	Asp	Pro	Lys	Phe	Asn	Ile	Val	Ser	Pro	Gly	Ala	Asp	Met	
		515					520					525				
agc	atc	tac	ttc	cca	tac	act	gag	gag	aag	cgt	aga	ttg	act	aag	ttc	1632
Ser	Ile	Tyr	Phe	Pro	Tyr	Thr	Glu	Glu	Lys	Arg	Arg	Leu	Thr	Lys	Phe	
	530					535					540					
cac	cca	gag	att	gag	gag	ctc	ctc	tac	agc	aat	gtt	gaa	aac	gaa	gag	1680
His	Pro	Glu	Ile	Glu	Glu	Leu	Leu	Tyr	Ser	Asn	Val	Glu	Asn	Glu	Glu	
545					550					555					560	
cac	tta	tgt	gtg	ctc	aag	gac	aag	aag	aag	cca	atc	ctc	ttc	aca	atg	1728
His	Leu	Cys	Val	Leu	Lys	Asp	Lys	Lys	Lys	Pro	Ile	Leu	Phe	Thr	Met	
				565					570					575		
gct	agg	ctc	gac	cgt	gtc	aag	aac	ttg	tca	ggg	ctt	gtt	gag	tgg	tac	1776
Ala	Arg	Leu	Asp	Arg	Val	Lys	Asn	Leu	Ser	Gly	Leu	Val	Glu	Trp	Tyr	
			580					585					590			
ggg	aag	aac	aaa	cgc	ctg	cgt	gag	ctt	gtt	aac	ttg	gtg	gtt	gtt	gga	1824
Gly	Lys	Asn	Lys	Arg	Leu	Arg	Glu	Leu	Val	Asn	Leu	Val	Val	Val	Gly	
		595					600					605				
gga	gac	agg	agg	aaa	gag	tca	aag	gac	aac	gaa	gag	aag	gcc	gag	atg	1872
Gly	Asp	Arg	Arg	Lys	Glu	Ser	Lys	Asp	Asn	Glu	Glu	Lys	Ala	Glu	Met	
	610					615					620					
aag	aaa	atg	tat	gac	ctc	att	gag	gaa	tac	aag	ctc	aac	ggc	cag	ttc	1920
Lys	Lys	Met	Tyr	Asp	Leu	Ile	Glu	Glu	Tyr	Lys	Leu	Asn	Gly	Gln	Phe	
	625				630					635					640	
cgt	tgg	atc	tcc	tcc	cag	atg	aac	agg	gtt	agg	aac	ggg	gag	ctc	tac	1968
Arg	Trp	Ile	Ser	Ser	Gln	Met	Asn	Arg	Val	Arg	Asn	Gly	Glu	Leu	Tyr	
				645					650					655		
agg	tac	atc	tgt	gac	acc	aag	ggg	gct	ttt	gtg	caa	cca	gca	ttg	tat	2016
Arg	Tyr	Ile	Cys	Asp	Thr	Lys	Gly	Ala	Phe	Val	Gln	Pro	Ala	Leu	Tyr	
		660						665					670			
gaa	gcc	ttt	ggg	ttg	act	gtt	gtg	gag	gct	atg	acc	tgt	ggg	tta	cca	2064
Glu	Ala	Phe	Gly	Leu	Thr	Val	Val	Glu	Ala	Met	Thr	Cys	Gly	Leu	Pro	

675	680	685	
aca ttc gcc act tgc aaa ggt ggt cca gct gag atc att gtc cac ggc			2112
Thr Phe Ala Thr Cys Lys Gly Gly Pro Ala Glu Ile Ile Val His Gly			
690	695	700	
aaa tca ggt ttc cac att gac cct tac cat ggt gat cag gcg gct gat			2160
Lys Ser Gly Phe His Ile Asp Pro Tyr His Gly Asp Gln Ala Ala Asp			
705	710	715	720
act ctt gct gat ttc ttc acc aag tgt aag gaa gat cca tct cac tgg			2208
Thr Leu Ala Asp Phe Phe Thr Lys Cys Lys Glu Asp Pro Ser His Trp			
	725	730	735
gac cag atc tca aag gga ggg ctt cag agg att gag gag aaa tac acg			2256
Asp Gln Ile Ser Lys Gly Gly Leu Gln Arg Ile Glu Glu Lys Tyr Thr			
	740	745	750
tgg cag att tac tca cag agg ctc ttg aca ttg act ggt gtg tat gga			2304
Trp Gln Ile Tyr Ser Gln Arg Leu Leu Thr Leu Thr Gly Val Tyr Gly			
	755	760	765
ttc tgg aag cat gtc tcg aac ctt gac cgt ctt gag agc cgt cgt tac			2352
Phe Trp Lys His Val Ser Asn Leu Asp Arg Leu Glu Ser Arg Arg Tyr			
	770	775	780
ctt gag atg ttc tat gca ttg aag tat cgc cca ctg gct cag gct gtt			2400
Leu Glu Met Phe Tyr Ala Leu Lys Tyr Arg Pro Leu Ala Gln Ala Val			
785	790	795	800
cct ctt gct caa gag			2415
Pro Leu Ala Gln Glu			
	805		
<210> 10			
<211> 805			
<212> PRT			
<213> Brassica napus			
<400> 10			
Met Val Asn Gly Val Leu Thr Arg Ala His Ser Gln Arg Glu Arg Leu			
1	5	10	15
Asn Glu Thr Leu Ala Ala Gln Arg Asn Glu Val Leu Ala Leu Leu Ser			
	20	25	30
Arg Val Glu Gly Lys Gly Lys Gly Ile Leu Gln Gln Asn Gln Ile Ile			
	35	40	45
Ala Glu Phe Glu Ala Leu Pro Glu Glu Thr Gln Lys Lys Ile Glu Gly			
	50	55	60
Gly Ala Phe Phe Asp Leu Leu Lys Thr Thr Gln Glu Ala Ile Val Leu			
65	70	75	80

Pro Pro Trp Val Ala Leu Ala Val Arg Pro Arg Pro Gly Val Trp Glu
 85 90 95

Tyr Ile Arg Val Asn Leu His Ala Leu Val Val Glu Glu Leu Thr Pro
 100 105 110

Ala Glu Phe Leu His Phe Lys Glu Glu Leu Val Asp Gly Val Lys Asn
 115 120 125

Gly Asp Phe Thr Leu Glu Leu Asp Phe Glu Pro Phe Asn Ala Ser Val
 130 135 140

Pro Arg Pro Thr Leu Pro Lys Tyr Ile Gly Asp Gly Val Glu Phe Leu
 145 150 155 160

Asn Arg His Leu Ser Ala Lys Leu Phe His Asp Lys Glu Ser Leu Leu
 165 170 175

Pro Leu Leu Lys Phe Leu Arg Leu His Ser Tyr Gln Gly Lys Thr Leu
 180 185 190

Met Leu Asn Glu Arg Val Gln Asn Leu Asn Asn Leu Gln His Ile Leu
 195 200 205

Arg Lys Ala Glu Glu Tyr Leu Ala Gly Leu Ser Pro Glu Thr Pro Tyr
 210 215 220

Glu Asp Phe Glu Ala Lys Phe Glu Glu Ile Gly Leu Glu Arg Gly Trp
 225 230 235 240

Gly Asn Asn Ala Glu Arg Val Leu Asp Met Ile Arg Leu Leu Leu Asp
 245 250 255

Leu Leu Glu Ala Pro Asp Pro Cys Thr Leu Glu Asn Phe Leu Gly Arg
 260 265 270

Val Pro Met Val Phe Asn Val Val Ile Leu Ser Pro His Gly Tyr Phe
 275 280 285

Ala Gln Asp Asn Val Leu Gly Tyr Pro Asp Thr Gly Gly Gln Val Val
 290 295 300

Tyr Ile Leu Asp Gln Val Arg Ala Met Glu Thr Glu Met Leu Gln Arg
 305 310 315 320

Ile Gln Gln Gln Gly Leu Asn Ile Thr Pro Arg Ile Leu Ile Leu Thr
 325 330 335
 Arg Leu Leu Pro Asp Ala Val Gly Thr Thr Cys Gly Glu Arg Leu Glu
 340 345 350
 Arg Val Asp Gly Ser Glu Tyr Cys Asp Ile Leu Arg Val Pro Phe Arg
 355 360 365
 Thr Glu Lys Gly Ile Val Arg Lys Trp Ile Ser Arg Phe Glu Val Trp
 370 375 380
 Pro Tyr Leu Glu Thr Tyr Thr Glu Asp Ala Ala Val Glu Leu Ala Lys
 385 390 395 400
 Glu Leu Lys Gly Lys Pro Asp Leu Ile Ile Gly Asn Tyr Ser Asp Gly
 405 410 415
 Asn Leu Val Ala Ser Leu Leu Ala His Lys Leu Gly Val Thr Gln Cys
 420 425 430
 Thr Ile Ala His Ala Leu Glu Lys Thr Lys Tyr Pro Asp Ser Asp Ile
 435 440 445
 Tyr Trp Lys Lys Leu Asp Asp Lys Tyr His Phe Ser Cys Gln Phe Thr
 450 455 460
 Ala Asp Leu Phe Ala Met Asn His Thr Asp Phe Ile Ile Thr Ser Thr
 465 470 475 480
 Leu Gln Glu Ile Ala Gly Ser Lys Asp Thr Val Gly Gln Tyr Glu Ser
 485 490 495
 His Thr Ala Phe Thr Leu Pro Gly Leu Tyr Arg Val Val His Gly Ile
 500 505 510
 Asp Val Phe Asp Pro Lys Phe Asn Ile Val Ser Pro Gly Ala Asp Met
 515 520 525
 Ser Ile Tyr Phe Pro Tyr Thr Glu Glu Lys Arg Arg Leu Thr Lys Phe
 530 535 540
 His Pro Glu Ile Glu Glu Leu Leu Tyr Ser Asn Val Glu Asn Glu Glu
 545 550 555 560
 His Leu Cys Val Leu Lys Asp Lys Lys Lys Pro Ile Leu Phe Thr Met

565	570	575
Ala Arg Leu Asp Arg Val Lys Asn Leu Ser Gly Leu Val Glu Trp Tyr		
580	585	590
Gly Lys Asn Lys Arg Leu Arg Glu Leu Val Asn Leu Val Val Val Gly		
595	600	605
Gly Asp Arg Arg Lys Glu Ser Lys Asp Asn Glu Glu Lys Ala Glu Met		
610	615	620
Lys Lys Met Tyr Asp Leu Ile Glu Glu Tyr Lys Leu Asn Gly Gln Phe		
625	630	635
Arg Trp Ile Ser Ser Gln Met Asn Arg Val Arg Asn Gly Glu Leu Tyr		
645	650	655
Arg Tyr Ile Cys Asp Thr Lys Gly Ala Phe Val Gln Pro Ala Leu Tyr		
660	665	670
Glu Ala Phe Gly Leu Thr Val Val Glu Ala Met Thr Cys Gly Leu Pro		
675	680	685
Thr Phe Ala Thr Cys Lys Gly Gly Pro Ala Glu Ile Ile Val His Gly		
690	695	700
Lys Ser Gly Phe His Ile Asp Pro Tyr His Gly Asp Gln Ala Ala Asp		
705	710	715
Thr Leu Ala Asp Phe Phe Thr Lys Cys Lys Glu Asp Pro Ser His Trp		
725	730	735
Asp Gln Ile Ser Lys Gly Gly Leu Gln Arg Ile Glu Glu Lys Tyr Thr		
740	745	750
Trp Gln Ile Tyr Ser Gln Arg Leu Leu Thr Leu Thr Gly Val Tyr Gly		
755	760	765
Phe Trp Lys His Val Ser Asn Leu Asp Arg Leu Glu Ser Arg Arg Tyr		
770	775	780
Leu Glu Met Phe Tyr Ala Leu Lys Tyr Arg Pro Leu Ala Gln Ala Val		
785	790	795
		800

Pro Leu Ala Gln Glu

805

<210> 11

<211> 2968

<212> DNA

<213> Glycine max

<220>

<221> CDS

<222> (244) .. (2661)

<400> 11

ggaaccctag cattcgccat tgccgccacc tttccccacg cagcttttgt tgtagcaaaa 60

gattgatatt ttcttcttcg tctataaaac ccacacgtca agtcacttct ttttaactca 120

atatttgggtc ccgtagctct atattttctc attcgcagac aagaaaaatc caagcctctc 180

gattacaccc ccctctcttt tttgcgttca ttctgttttc ctgttgaagt ctttccttag 240

cca atg gcc acc gat cgt ttg acc cgg gtt cac agt ctc cgt gag agg 288

Met Ala Thr Asp Arg Leu Thr Arg Val His Ser Leu Arg Glu Arg

1 5 10 15

ctt gat gaa acc ctc act gcc aac agg aat gaa att ttg gcc ctt ctg 336

Leu Asp Glu Thr Leu Thr Ala Asn Arg Asn Glu Ile Leu Ala Leu Leu

20 25 30

tca agg atc gaa gcc aag ggc aag ggc atc ctg caa cac cac cag gtc 384

Ser Arg Ile Glu Ala Lys Gly Lys Gly Ile Leu Gln His His Gln Val

35 40 45

att gct gag ttt gag gaa atc cct gag gag aac aga cag aag ctc act 432

Ile Ala Glu Phe Glu Glu Ile Pro Glu Glu Asn Arg Gln Lys Leu Thr

50 55 60

gat ggt gcc ttt gga gaa gtc ttg aga tct aca cag gaa gcc ata gtt 480

Asp Gly Ala Phe Gly Glu Val Leu Arg Ser Thr Gln Glu Ala Ile Val

65 70 75

ttg cca cca tgg gtt gct ctg gct gtt cgt cca aga cct ggt gtg tgg 528

Leu Pro Pro Trp Val Ala Leu Ala Val Arg Pro Arg Pro Gly Val Trp

80 85 90 95

gag tac ctg aga gtg aat gtg cac gct ctt gtt gtt gag gag ttg caa 576

Glu Tyr Leu Arg Val Asn Val His Ala Leu Val Val Glu Glu Leu Gln

100 105 110

cct gct gag tac ctg cac ttc aag gaa gaa ctt gtt gat gga agt tct 624

Pro Ala Glu Tyr Leu His Phe Lys Glu Glu Leu Val Asp Gly Ser Ser

115 120 125

aat ggc aac ttt gtg ctt gag ttg gac ttt gaa cca ttc aat gca gcc 672

Asn Gly Asn Phe Val Leu Glu Leu Asp Phe Glu Pro Phe Asn Ala Ala

130 135 140

ttc cct cgc cca act ctt aac aag tca att gga aat ggt gtg cag ttc 720

Phe Pro Arg Pro Thr Leu Asn Lys Ser Ile Gly Asn Gly Val Gln Phe

145	150	155	
ctc aac cgc cac ctt tct gcc aaa ctc ttc cac gac aag gag agc ttg			768
Leu Asn Arg His Leu Ser Ala Lys Leu Phe His Asp Lys Glu Ser Leu			
160	165	170	175
cac cca ctt ttg gaa ttc ctc agg ctt cac agc gtc aag gga aag act			816
His Pro Leu Leu Glu Phe Leu Arg Leu His Ser Val Lys Gly Lys Thr			
	180	185	190
ttg atg ttg aat gac aga att caa aac cca gat gca ctc caa cat gtt			864
Leu Met Leu Asn Asp Arg Ile Gln Asn Pro Asp Ala Leu Gln His Val			
	195	200	205
ctg agg aaa gct gag gag tat ctg ggc aca gtg cct cct gaa act ccc			912
Leu Arg Lys Ala Glu Glu Tyr Leu Gly Thr Val Pro Pro Glu Thr Pro			
	210	215	220
tac tca gaa ttt gag cac aag ttc cag gag att ggt ttg gag aga ggg			960
Tyr Ser Glu Phe Glu His Lys Phe Gln Glu Ile Gly Leu Glu Arg Gly			
	225	230	235
tgg ggt gac aac gca gag cgt gtt ctt gag tca att caa ctt ctc ttg			1008
Trp Gly Asp Asn Ala Glu Arg Val Leu Glu Ser Ile Gln Leu Leu Leu			
	240	245	250
gat ctt ctt gag gcc cct gac cct tgc acc ctt gag act ttc ctt gga			1056
Asp Leu Leu Glu Ala Pro Asp Pro Cys Thr Leu Glu Thr Phe Leu Gly			
	260	265	270
aga att cct atg gtc ttc aat gtt gtc att ctt tct ccc cat ggt tac			1104
Arg Ile Pro Met Val Phe Asn Val Val Ile Leu Ser Pro His Gly Tyr			
	275	280	285
ttt gcc caa gat aat gtc ttg gga tac cct gac act ggt ggc cag gtt			1152
Phe Ala Gln Asp Asn Val Leu Gly Tyr Pro Asp Thr Gly Gly Gln Val			
	290	295	300
gtt tac atc ttg gat caa gtt cgt gct ttg gag aac gag atg ctc cat			1200
Val Tyr Ile Leu Asp Gln Val Arg Ala Leu Glu Asn Glu Met Leu His			
	305	310	315
cgc att aag caa caa gga ttg gac att gta cct cgt att ctc att atc			1248
Arg Ile Lys Gln Gln Gly Leu Asp Ile Val Pro Arg Ile Leu Ile Ile			
	320	325	330
acc cgt ctt ctc ccc gat gca atc gga act act tgt ggc caa cgt ctt			1296
Thr Arg Leu Leu Pro Asp Ala Ile Gly Thr Thr Cys Gly Gln Arg Leu			
	340	345	350
gag aag gtg ttc gga acc gag cac tcc cac att ctt cga gtt ccc ttt			1344
Glu Lys Val Phe Gly Thr Glu His Ser His Ile Leu Arg Val Pro Phe			
	355	360	365
aga act gag aag gga att gtt cgt cag tgg atc tca aga ttc gaa gtc			1392
Arg Thr Glu Lys Gly Ile Val Arg Gln Trp Ile Ser Arg Phe Glu Val			
	370	375	380
tgg cca tac ttg gaa act tac act gag gat gtt gct cat gag ctt gcc			1440
Trp Pro Tyr Leu Glu Thr Tyr Thr Glu Asp Val Ala His Glu Leu Ala			
	385	390	395

aaa gag ttg caa ggc aag cca gat ctg att gtc gga aac tac agt gat	1488
Lys Glu Leu Gln Gly Lys Pro Asp Leu Ile Val Gly Asn Tyr Ser Asp	
400 405 410 415	
gga aac att gtc gcc tct ttg ttg gca cat aaa tta gga gtc act cag	1536
Gly Asn Ile Val Ala Ser Leu Leu Ala His Lys Leu Gly Val Thr Gln	
420 425 430	
tgt acc att gct cac gca ctt gag aag acc aaa tac ccc gaa tcc gac	1584
Cys Thr Ile Ala His Ala Leu Glu Lys Thr Lys Tyr Pro Glu Ser Asp	
435 440 445	
att tac tgg aaa aaa ttg gaa gag aga tac cac ttc tct tgc caa ttc	1632
Ile Tyr Trp Lys Lys Leu Glu Glu Arg Tyr His Phe Ser Cys Gln Phe	
450 455 460	
aca gct gat cta ttt gcc atg aac cac aca gat ttc att atc acc agt	1680
Thr Ala Asp Leu Phe Ala Met Asn His Thr Asp Phe Ile Ile Thr Ser	
465 470 475	
acc ttc cag gag att gct gga agc aag gac act gtt gga cag tac gag	1728
Thr Phe Gln Glu Ile Ala Gly Ser Lys Asp Thr Val Gly Gln Tyr Glu	
480 485 490 495	
tct cac aca gcc ttc acc ctt cct gga ctc tac cgt gtt gtg cac ggc	1776
Ser His Thr Ala Phe Thr Leu Pro Gly Leu Tyr Arg Val Val His Gly	
500 505 510	
att gat gtc ttt gat cca aaa ttc aac att gtc tcc cct gga gct gat	1824
Ile Asp Val Phe Asp Pro Lys Phe Asn Ile Val Ser Pro Gly Ala Asp	
515 520 525	
caa acc att tac ttc ccc ccc acc gaa act agc cgt agg ttg acc tcc	1872
Gln Thr Ile Tyr Phe Pro Pro Thr Glu Thr Ser Arg Arg Leu Thr Ser	
530 535 540	
ttc cac ccc gaa atc gaa gaa ctt ctt tac agc tct gtg gag aat gaa	1920
Phe His Pro Glu Ile Glu Glu Leu Leu Tyr Ser Ser Val Glu Asn Glu	
545 550 555	
gaa cac ata tgc gtg ctg aag gac cgc agc aag ccg att atc ttc acc	1968
Glu His Ile Cys Val Leu Lys Asp Arg Ser Lys Pro Ile Ile Phe Thr	
560 565 570 575	
atg gca agg ttg gac cgt gtg aag aac atc aca gga ctc gtg gag tgg	2016
Met Ala Arg Leu Asp Arg Val Lys Asn Ile Thr Gly Leu Val Glu Trp	
580 585 590	
tac ggt aag aac gcg aag ctg agg gag ttg gtg aac ctt gtg gtt gtt	2064
Tyr Gly Lys Asn Ala Lys Leu Arg Glu Leu Val Asn Leu Val Val Val	
595 600 605	
gcc gga gac agg agg aag gag tcg aag gac ttg gaa gag aag gcc gag	2112
Ala Gly Asp Arg Arg Lys Glu Ser Lys Asp Leu Glu Glu Lys Ala Glu	
610 615 620	
atg aag aag atg tac ggc ctg atc gag acc tac aag ttg aac ggg caa	2160
Met Lys Lys Met Tyr Gly Leu Ile Glu Thr Tyr Lys Leu Asn Gly Gln	
625 630 635	

ttc aga tgg att tca tct cag atg aac cgt gtg agg aac gga gag ctg 2208
 Phe Arg Trp Ile Ser Ser Gln Met Asn Arg Val Arg Asn Gly Glu Leu
 640 645 650 655
 tac cgc gtg atc tgc gac acc agg ggt gct ttc gtg cag cct gct gta 2256
 Tyr Arg Val Ile Cys Asp Thr Arg Gly Ala Phe Val Gln Pro Ala Val
 660 665 670
 tac gag gct ttt ggt ttg aca gtg gtt gag gcc atg act tgc ggc ttg 2304
 Tyr Glu Ala Phe Gly Leu Thr Val Val Glu Ala Met Thr Cys Gly Leu
 675 680 685
 cca aca ttc gcc aca tgc aat ggt ggt cct gct gag atc att gtg cac 2352
 Pro Thr Phe Ala Thr Cys Asn Gly Gly Pro Ala Glu Ile Ile Val His
 690 695 700
 ggc aag tct ggc ttc cac att gac cct tac cat ggt gac cgt gct gct 2400
 Gly Lys Ser Gly Phe His Ile Asp Pro Tyr His Gly Asp Arg Ala Ala
 705 710 715
 gat ctc ctt gtt gac ttc ttt gag aag tgc aag ctt gac cca act cac 2448
 Asp Leu Leu Val Asp Phe Phe Glu Lys Cys Lys Leu Asp Pro Thr His
 720 725 730 735
 tgg gac aag atc tca aag gct ggt ctc cag cgt att gaa gag aag tac 2496
 Trp Asp Lys Ile Ser Lys Ala Gly Leu Gln Arg Ile Glu Glu Lys Tyr
 740 745 750
 aca tgg caa att tac tct cag agg ctt ctc act ctc acc ggt gtc tat 2544
 Thr Trp Gln Ile Tyr Ser Gln Arg Leu Leu Thr Leu Thr Gly Val Tyr
 755 760 765
 ggc ttc tgg aag cat gtg tct aac ctt gac cgc cgt gag agc cgc cgc 2592
 Gly Phe Trp Lys His Val Ser Asn Leu Asp Arg Arg Glu Ser Arg Arg
 770 775 780
 tat ctc gag atg ttc tat gct ctc aag tac cgc aaa ttg gct gag tct 2640
 Tyr Leu Glu Met Phe Tyr Ala Leu Lys Tyr Arg Lys Leu Ala Glu Ser
 785 790 795
 gtg ccc ctt gct gct gag taa actgaggata aagagttgga taaagaaatg 2691
 Val Pro Leu Ala Ala Glu
 800 805
 gaggaaccgg ctttttcttt ctcatattgga gtttgctcgca cttgagtttt ataaataatg 2751
 tccgtgatttt tagttttgtg attaagcttt cgataagagg agagaaagag aaggaaaaaa 2811
 aaagttgctt ttttttttgt tgttgcatga tttggatctt gatttgaaaa gcttcgaatt 2871
 ggggtagttt tacccatcaa ttcaatttta agccgtgcct tcttcaaaaa aaaaaaaaaa 2931
 aaaaaaaaaa aaaaaaaatc aaaaaaaaaa aaaaaaa 2968

<210> 12
 <211> 805
 <212> PRT
 <213> Glycine max

<400> 12

Met Ala Thr Asp Arg Leu Thr Arg Val His Ser Leu Arg Glu Arg Leu
 1 5 10 15
 Asp Glu Thr Leu Thr Ala Asn Arg Asn Glu Ile Leu Ala Leu Leu Ser
 20 25 30
 Arg Ile Glu Ala Lys Gly Lys Gly Ile Leu Gln His His Gln Val Ile
 35 40 45
 Ala Glu Phe Glu Glu Ile Pro Glu Glu Asn Arg Gln Lys Leu Thr Asp
 50 55 60
 Gly Ala Phe Gly Glu Val Leu Arg Ser Thr Gln Glu Ala Ile Val Leu
 65 70 75 80
 Pro Pro Trp Val Ala Leu Ala Val Arg Pro Arg Pro Gly Val Trp Glu
 85 90 95
 Tyr Leu Arg Val Asn Val His Ala Leu Val Val Glu Glu Leu Gln Pro
 100 105 110
 Ala Glu Tyr Leu His Phe Lys Glu Glu Leu Val Asp Gly Ser Ser Asn
 115 120 125
 Gly Asn Phe Val Leu Glu Leu Asp Phe Glu Pro Phe Asn Ala Ala Phe
 130 135 140
 Pro Arg Pro Thr Leu Asn Lys Ser Ile Gly Asn Gly Val Gln Phe Leu
 145 150 155 160
 Asn Arg His Leu Ser Ala Lys Leu Phe His Asp Lys Glu Ser Leu His
 165 170 175
 Pro Leu Leu Glu Phe Leu Arg Leu His Ser Val Lys Gly Lys Thr Leu
 180 185 190
 Met Leu Asn Asp Arg Ile Gln Asn Pro Asp Ala Leu Gln His Val Leu
 195 200 205
 Arg Lys Ala Glu Glu Tyr Leu Gly Thr Val Pro Pro Glu Thr Pro Tyr
 210 215 220
 Ser Glu Phe Glu His Lys Phe Gln Glu Ile Gly Leu Glu Arg Gly Trp
 225 230 235 240

Gly Asp Asn Ala Glu Arg Val Leu Glu Ser Ile Gln Leu Leu Leu Asp
 245 250 255
 Leu Leu Glu Ala Pro Asp Pro Cys Thr Leu Glu Thr Phe Leu Gly Arg
 260 265 270
 Ile Pro Met Val Phe Asn Val Val Ile Leu Ser Pro His Gly Tyr Phe
 275 280 285
 Ala Gln Asp Asn Val Leu Gly Tyr Pro Asp Thr Gly Gly Gln Val Val
 290 295 300
 Tyr Ile Leu Asp Gln Val Arg Ala Leu Glu Asn Glu Met Leu His Arg
 305 310 315 320
 Ile Lys Gln Gln Gly Leu Asp Ile Val Pro Arg Ile Leu Ile Ile Thr
 325 330 335
 Arg Leu Leu Pro Asp Ala Ile Gly Thr Thr Cys Gly Gln Arg Leu Glu
 340 345 350
 Lys Val Phe Gly Thr Glu His Ser His Ile Leu Arg Val Pro Phe Arg
 355 360 365
 Thr Glu Lys Gly Ile Val Arg Gln Trp Ile Ser Arg Phe Glu Val Trp
 370 375 380
 Pro Tyr Leu Glu Thr Tyr Thr Glu Asp Val Ala His Glu Leu Ala Lys
 385 390 395 400
 Glu Leu Gln Gly Lys Pro Asp Leu Ile Val Gly Asn Tyr Ser Asp Gly
 405 410 415
 Asn Ile Val Ala Ser Leu Leu Ala His Lys Leu Gly Val Thr Gln Cys
 420 425 430
 Thr Ile Ala His Ala Leu Glu Lys Thr Lys Tyr Pro Glu Ser Asp Ile
 435 440 445
 Tyr Trp Lys Lys Leu Glu Glu Arg Tyr His Phe Ser Cys Gln Phe Thr
 450 455 460
 Ala Asp Leu Phe Ala Met Asn His Thr Asp Phe Ile Ile Thr Ser Thr
 465 470 475 480
 Phe Gln Glu Ile Ala Gly Ser Lys Asp Thr Val Gly Gln Tyr Glu Ser

43

	725	730	735	
Asp Lys Ile Ser Lys Ala Gly Leu Gln Arg Ile Glu Glu Lys Tyr Thr	740	745	750	
Trp Gln Ile Tyr Ser Gln Arg Leu Leu Thr Leu Thr Gly Val Tyr Gly	755	760	765	
Phe Trp Lys His Val Ser Asn Leu Asp Arg Arg Glu Ser Arg Arg Tyr	770	775	780	
Leu Glu Met Phe Tyr Ala Leu Lys Tyr Arg Lys Leu Ala Glu Ser Val	785	790	795	800
Pro Leu Ala Ala Glu	805			
<210> 13				
<211> 2415				
<212> DNA				
<213> Glycine max				
<220>				
<221> CDS				
<222> (1)..(2415)				
<400> 13				
atg gcc acc gat cgt ttg acc cgg gtt cac agt ctc cgt gag agg ctt				48
Met Ala Thr Asp Arg Leu Thr Arg Val His Ser Leu Arg Glu Arg Leu				
1 5 10 15				
gat gaa acc ctc act gcc aac agg aat gaa att ttg gcc ctt ctg tca				96
Asp Glu Thr Leu Thr Ala Asn Arg Asn Glu Ile Leu Ala Leu Leu Ser				
20 25 30				
agg atc gaa gcc aag ggc aag ggc atc ctg caa cac cac cag gtc att				144
Arg Ile Glu Ala Lys Gly Lys Gly Ile Leu Gln His His Gln Val Ile				
35 40 45				
gct gag ttt gag gaa atc cct gag gag aac aga cag aag ctc act gat				192
Ala Glu Phe Glu Glu Ile Pro Glu Glu Asn Arg Gln Lys Leu Thr Asp				
50 55 60				
ggc gcc ttt gga gaa gtc ttg aga tct aca cag gaa gcc ata gtt ttg				240
Gly Ala Phe Gly Glu Val Leu Arg Ser Thr Gln Glu Ala Ile Val Leu				
65 70 75 80				
cca cca tgg gtt gct ctg gct gtt cgt cca aga cct ggt gtg tgg gag				288
Pro Pro Trp Val Ala Leu Ala Val Arg Pro Arg Pro Gly Val Trp Glu				
85 90 95				
tac ctg aga gtg aat gtg cac gct ctt gtt gtt gag gag ttg caa cct				336
Tyr Leu Arg Val Asn Val His Ala Leu Val Val Glu Glu Leu Gln Pro				
100 105 110				

gct gag tac ctg cac ttc aag gaa gaa ctt gtt gat gga agt tct aat Ala Glu Tyr Leu His Phe Lys Glu Glu Leu Val Asp Gly Ser Ser Asn 115 120 125	384
ggc aac ttt gtg ctt gag ttg gac ttt gaa cca ttc aat gca gcc ttc Gly Asn Phe Val Leu Glu Leu Asp Phe Glu Pro Phe Asn Ala Ala Phe 130 135 140	432
cct cgc cca act ctt aac aag tca att gga aat ggt gtg cag ttc ctc Pro Arg Pro Thr Leu Asn Lys Ser Ile Gly Asn Gly Val Gln Phe Leu 145 150 155 160	480
aac cgc cac ctt tct gcc aaa ctc ttc cac gac aag gag agc ttg cac Asn Arg His Leu Ser Ala Lys Leu Phe His Asp Lys Glu Ser Leu His 165 170 175	528
cca ctt ttg gaa ttc ctc agg ctt cac agc gtc aag gga aag act ttg Pro Leu Leu Glu Phe Leu Arg Leu His Ser Val Lys Gly Lys Thr Leu 180 185 190	576
atg ttg aat gac aga att caa aac cca gat gca ctc caa cat gtt ctg Met Leu Asn Asp Arg Ile Gln Asn Pro Asp Ala Leu Gln His Val Leu 195 200 205	624
agg aaa gct gag gag tat ctg ggc aca gtg cct cct gaa act ccc tac Arg Lys Ala Glu Glu Tyr Leu Gly Thr Val Pro Pro Glu Thr Pro Tyr 210 215 220	672
tca gaa ttt gag cac aag ttc cag gag att ggt ttg gag aga ggg tgg Ser Glu Phe Glu His Lys Phe Gln Glu Ile Gly Leu Glu Arg Gly Trp 225 230 235 240	720
ggg gac aac gca gag cgt gtt ctt gag tca att caa ctt ctc ttg gat Gly Asp Asn Ala Glu Arg Val Leu Glu Ser Ile Gln Leu Leu Leu Asp 245 250 255	768
ctt ctt gag gcc cct gac cct tgc acc ctt gag act ttc ctt gga aga Leu Leu Glu Ala Pro Asp Pro Cys Thr Leu Glu Thr Phe Leu Gly Arg 260 265 270	816
att cct atg gtc ttc aat gtt gtc att ctt tct ccc cat ggt tac ttt Ile Pro Met Val Phe Asn Val Val Ile Leu Ser Pro His Gly Tyr Phe 275 280 285	864
gcc caa gat aat gtc ttg gga tac cct gac act ggt ggc cag gtt gtt Ala Gln Asp Asn Val Leu Gly Tyr Pro Asp Thr Gly Gly Gln Val Val 290 295 300	912
tac atc ttg gat caa gtt cgt gct ttg gag aac gag atg ctc cat cgc Tyr Ile Leu Asp Gln Val Arg Ala Leu Glu Asn Glu Met Leu His Arg 305 310 315 320	960
att aag caa caa gga ttg gac att gta cct cgt att ctc att atc acc Ile Lys Gln Gln Gly Leu Asp Ile Val Pro Arg Ile Leu Ile Ile Thr 325 330 335	1008
cgt ctt ctc ccc gat gca atc gga act act tgt ggc caa cgt ctt gag Arg Leu Leu Pro Asp Ala Ile Gly Thr Thr Cys Gly Gln Arg Leu Glu 340 345 350	1056

aag gtg ttc gga acc gag cac tcc cac att ctt cga gtt ccc ttt aga Lys Val Phe Gly Thr Glu His Ser His Ile Leu Arg Val Pro Phe Arg 355 360 365	1104
act gag aag gga att gtt cgt cag tgg atc tca aga ttc gaa gtc tgg Thr Glu Lys Gly Ile Val Arg Gln Trp Ile Ser Arg Phe Glu Val Trp 370 375 380	1152
cca tac ttg gaa act tac act gag gat gtt gct cat gag ctt gcc aaa Pro Tyr Leu Glu Thr Tyr Thr Glu Asp Val Ala His Glu Leu Ala Lys 385 390 395 400	1200
gag ttg caa ggc aag cca gat ctg att gtc gga aac tac agt gat gga Glu Leu Gln Gly Lys Pro Asp Leu Ile Val Gly Asn Tyr Ser Asp Gly 405 410 415	1248
aac att gtc gcc tct ttg ttg gca cat aaa tta gga gtc act cag tgt Asn Ile Val Ala Ser Leu Leu Ala His Lys Leu Gly Val Thr Gln Cys 420 425 430	1296
acc att gct cac gca ctt gag aag acc aaa tac ccc gaa tcc gac att Thr Ile Ala His Ala Leu Glu Lys Thr Lys Tyr Pro Glu Ser Asp Ile 435 440 445	1344
tac tgg aaa aaa ttg gaa gag aga tac cac ttc tct tgc caa ttc aca Tyr Trp Lys Lys Leu Glu Glu Arg Tyr His Phe Ser Cys Gln Phe Thr 450 455 460	1392
gct gat cta ttt gcc atg aac cac aca gat ttc att atc acc agt acc Ala Asp Leu Phe Ala Met Asn His Thr Asp Phe Ile Ile Thr Ser Thr 465 470 475 480	1440
ttc cag gag att gct gga agc aag gac act gtt gga cag tac gag tct Phe Gln Glu Ile Ala Gly Ser Lys Asp Thr Val Gly Gln Tyr Glu Ser 485 490 495	1488
cac aca gcc ttc acc ctt cct gga ctc tac cgt gtt gtg cac ggc att His Thr Ala Phe Thr Leu Pro Gly Leu Tyr Arg Val Val His Gly Ile 500 505 510	1536
gat gtc ttt gat cca aaa ttc aac att gtc tcc cct gga gct gat caa Asp Val Phe Asp Pro Lys Phe Asn Ile Val Ser Pro Gly Ala Asp Gln 515 520 525	1584
acc att tac ttc ccc ccc acc gaa act agc cgt agg ttg acc tcc ttc Thr Ile Tyr Phe Pro Pro Thr Glu Thr Ser Arg Arg Leu Thr Ser Phe 530 535 540	1632
cac ccc gaa atc gaa gaa ctt ctt tac agc tct gtg gag aat gaa gaa His Pro Glu Ile Glu Glu Leu Leu Tyr Ser Ser Val Glu Asn Glu Glu 545 550 555 560	1680
cac ata tgc gtg ctg aag gac cgc agc aag ccg att atc ttc acc atg His Ile Cys Val Leu Lys Asp Arg Ser Lys Pro Ile Ile Phe Thr Met 565 570 575	1728
gca agg ttg gac cgt gtg aag aac atc aca gga ctc gtg gag tgg tac Ala Arg Leu Asp Arg Val Lys Asn Ile Thr Gly Leu Val Glu Trp Tyr 580 585 590	1776
ggt aag aac gcg aag ctg agg gag ttg gtg aac ctt gtg gtt gtt gcc	1824

Gly Lys Asn Ala Lys Leu Arg Glu Leu Val Asn Leu Val Val Val Ala	
595 600 605	
gga gac agg agg aag gag tcg aag gac ttg gaa gag aag gcc gag atg	1872
Gly Asp Arg Arg Lys Glu Ser Lys Asp Leu Glu Glu Lys Ala Glu Met	
610 615 620	
aag aag atg tac ggc ctg atc gag acc tac aag ttg aac ggg caa ttc	1920
Lys Lys Met Tyr Gly Leu Ile Glu Thr Tyr Lys Leu Asn Gly Gln Phe	
625 630 635 640	
aga tgg att tca tct cag atg aac cgt gtg agg aac gga gag ctg tac	1968
Arg Trp Ile Ser Ser Gln Met Asn Arg Val Arg Asn Gly Glu Leu Tyr	
645 650 655	
cgc gtg atc tgc gac acc agg ggt gct ttc gtg cag cct gct gta tac	2016
Arg Val Ile Cys Asp Thr Arg Gly Ala Phe Val Gln Pro Ala Val Tyr	
660 665 670	
gag gct ttt ggt ttg aca gtg gtt gag gcc atg act tgc ggc ttg cca	2064
Glu Ala Phe Gly Leu Thr Val Val Glu Ala Met Thr Cys Gly Leu Pro	
675 680 685	
aca ttc gcc aca tgc aat ggt ggt cct gct gag atc att gtg cac ggc	2112
Thr Phe Ala Thr Cys Asn Gly Gly Pro Ala Glu Ile Ile Val His Gly	
690 695 700	
aag tct ggc ttc cac att gac cct tac cat ggt gac cgt gct gct gat	2160
Lys Ser Gly Phe His Ile Asp Pro Tyr His Gly Asp Arg Ala Ala Asp	
705 710 715 720	
ctc ctt gtt gac ttc ttt gag aag tgc aag ctt gac cca act cac tgg	2208
Leu Leu Val Asp Phe Phe Glu Lys Cys Lys Leu Asp Pro Thr His Trp	
725 730 735	
gac aag atc tca aag gct ggt ctc cag cgt att gaa gag aag tac aca	2256
Asp Lys Ile Ser Lys Ala Gly Leu Gln Arg Ile Glu Glu Lys Tyr Thr	
740 745 750	
tgg caa att tac tct cag agg ctt ctc act ctc acc ggt gtc tat ggc	2304
Trp Gln Ile Tyr Ser Gln Arg Leu Leu Thr Leu Thr Gly Val Tyr Gly	
755 760 765	
ttc tgg aag cat gtg tct aac ctt gac cgc cgt gag agc cgc cgc tat	2352
Phe Trp Lys His Val Ser Asn Leu Asp Arg Arg Glu Ser Arg Arg Tyr	
770 775 780	
ctc gag atg ttc tat gct ctc aag tac cgc aaa ttg gct gag tct gtg	2400
Leu Glu Met Phe Tyr Ala Leu Lys Tyr Arg Lys Leu Ala Glu Ser Val	
785 790 795 800	
ccc ctt gct gct gag	2415
Pro Leu Ala Ala Glu	
805	
<210> 14	
<211> 805	
<212> PRT	
<213> Glycine max	

<400> 14

Met Ala Thr Asp Arg Leu Thr Arg Val His Ser Leu Arg Glu Arg Leu
 1 5 10 15

Asp Glu Thr Leu Thr Ala Asn Arg Asn Glu Ile Leu Ala Leu Leu Ser
 20 25 30

Arg Ile Glu Ala Lys Gly Lys Gly Ile Leu Gln His His Gln Val Ile
 35 40 45

Ala Glu Phe Glu Glu Ile Pro Glu Glu Asn Arg Gln Lys Leu Thr Asp
 50 55 60

Gly Ala Phe Gly Glu Val Leu Arg Ser Thr Gln Glu Ala Ile Val Leu
 65 70 75 80

Pro Pro Trp Val Ala Leu Ala Val Arg Pro Arg Pro Gly Val Trp Glu
 85 90 95

Tyr Leu Arg Val Asn Val His Ala Leu Val Val Glu Glu Leu Gln Pro
 100 105 110

Ala Glu Tyr Leu His Phe Lys Glu Glu Leu Val Asp Gly Ser Ser Asn
 115 120 125

Gly Asn Phe Val Leu Glu Leu Asp Phe Glu Pro Phe Asn Ala Ala Phe
 130 135 140

Pro Arg Pro Thr Leu Asn Lys Ser Ile Gly Asn Gly Val Gln Phe Leu
 145 150 155 160

Asn Arg His Leu Ser Ala Lys Leu Phe His Asp Lys Glu Ser Leu His
 165 170 175

Pro Leu Leu Glu Phe Leu Arg Leu His Ser Val Lys Gly Lys Thr Leu
 180 185 190

Met Leu Asn Asp Arg Ile Gln Asn Pro Asp Ala Leu Gln His Val Leu
 195 200 205

Arg Lys Ala Glu Glu Tyr Leu Gly Thr Val Pro Pro Glu Thr Pro Tyr
 210 215 220

Ser Glu Phe Glu His Lys Phe Gln Glu Ile Gly Leu Glu Arg Gly Trp
 225 230 235 240

49

Phe Gln Glu Ile Ala Gly Ser Lys Asp Thr Val Gly Gln Tyr Glu Ser
 485 490 495

His Thr Ala Phe Thr Leu Pro Gly Leu Tyr Arg Val Val His Gly Ile
 500 505 510

Asp Val Phe Asp Pro Lys Phe Asn Ile Val Ser Pro Gly Ala Asp Gln
 515 520 525

Thr Ile Tyr Phe Pro Pro Thr Glu Thr Ser Arg Arg Leu Thr Ser Phe
 530 535 540

His Pro Glu Ile Glu Glu Leu Leu Tyr Ser Ser Val Glu Asn Glu Glu
 545 550 555 560

His Ile Cys Val Leu Lys Asp Arg Ser Lys Pro Ile Ile Phe Thr Met
 565 570 575

Ala Arg Leu Asp Arg Val Lys Asn Ile Thr Gly Leu Val Glu Trp Tyr
 580 585 590

Gly Lys Asn Ala Lys Leu Arg Glu Leu Val Asn Leu Val Val Val Ala
 595 600 605

Gly Asp Arg Arg Lys Glu Ser Lys Asp Leu Glu Glu Lys Ala Glu Met
 610 615 620

Lys Lys Met Tyr Gly Leu Ile Glu Thr Tyr Lys Leu Asn Gly Gln Phe
 625 630 635 640

Arg Trp Ile Ser Ser Gln Met Asn Arg Val Arg Asn Gly Glu Leu Tyr
 645 650 655

Arg Val Ile Cys Asp Thr Arg Gly Ala Phe Val Gln Pro Ala Val Tyr
 660 665 670

Glu Ala Phe Gly Leu Thr Val Val Glu Ala Met Thr Cys Gly Leu Pro
 675 680 685

Thr Phe Ala Thr Cys Asn Gly Gly Pro Ala Glu Ile Ile Val His Gly
 690 695 700

Lys Ser Gly Phe His Ile Asp Pro Tyr His Gly Asp Arg Ala Ala Asp
 705 710 715 720

Leu Leu Val Asp Phe Phe Glu Lys Cys Lys Leu Asp Pro Thr His Trp
 725 730 735

Asp Lys Ile Ser Lys Ala Gly Leu Gln Arg Ile Glu Glu Lys Tyr Thr
 740 745 750

Trp Gln Ile Tyr Ser Gln Arg Leu Leu Thr Leu Thr Gly Val Tyr Gly
 755 760 765

Phe Trp Lys His Val Ser Asn Leu Asp Arg Arg Glu Ser Arg Arg Tyr
 770 775 780

Leu Glu Met Phe Tyr Ala Leu Lys Tyr Arg Lys Leu Ala Glu Ser Val
 785 790 795 800

Pro Leu Ala Ala Glu
 805

<210> 15
 <211> 3238
 <212> DNA
 <213> Glycine max

<220>
 <221> misc_feature
 <222> (3052)..(3217)
 <223> n located at 3052, 3130, 3072, 3123, 3096, 3165, 3170, 3176-3180,
 3208, 3217, each n is a or c or g or t/u, unknown, or other

<220>
 <221> CDS
 <222> (84)..(2504)

<400> 15
 gctgtcgggtc tctctctctc tctctcccca ccccttcctt ttttgcgttc attctgtttt 60
 catagtgcgc aacttctgaa gaa atg gca aat cat cct ttg aca cac tct cac 113
 Met Ala Asn His Pro Leu Thr His Ser His
 1 5 10
 tct ttc cgc gag agg ttt gat gaa act ctc act ggt cac agg aac gaa 161
 Ser Phe Arg Glu Arg Phe Asp Glu Thr Leu Thr Gly His Arg Asn Glu
 15 20 25
 att ttg gcc ctt ctg tca agg ctt gaa gcc aag ggc aag gga atc ctg 209
 Ile Leu Ala Leu Leu Ser Arg Leu Glu Ala Lys Gly Lys Gly Ile Leu
 30 35 40
 caa cat cac caa gtg gtt gca gag ttt gaa gaa atc cct gag gag agc 257
 Gln His His Gln Val Val Ala Glu Phe Glu Glu Ile Pro Glu Glu Ser
 45 50 55
 aga aag aaa ctc caa gat ggt gtc ttt gga gaa gtt ttg aga tcc aca 305
 Arg Lys Lys Leu Gln Asp Gly Val Phe Gly Glu Val Leu Arg Ser Thr
 60 65 70

cag gaa gcc ata gtg ctg cca cca ttt gta gct ctt gct gtt cga cca	353
Gln Glu Ala Ile Val Leu Pro Pro Phe Val Ala Leu Ala Val Arg Pro	
75 80 85 90	
agg cct ggt gta tgg gaa tat ctg cgt gtg aat gtg cac atg ctt gtt	401
Arg Pro Gly Val Trp Glu Tyr Leu Arg Val Asn Val His Met Leu Val	
95 100 105	
gtt gat gag ctt cgt cct gct gag tat ctg cgt ttc aag gag gag ctt	449
Val Asp Glu Leu Arg Pro Ala Glu Tyr Leu Arg Phe Lys Glu Glu Leu	
110 115 120	
gtt gag gga agt tca aac ggc aac ttt gtg ctt gag ttg gac ttt gaa	497
Val Glu Gly Ser Ser Asn Gly Asn Phe Val Leu Glu Leu Asp Phe Glu	
125 130 135	
ccg ttt aat gca tcc ttc cct cgc cca act ctg aac aag tcc att gga	545
Pro Phe Asn Ala Ser Phe Pro Arg Pro Thr Leu Asn Lys Ser Ile Gly	
140 145 150	
aat ggc gtc gag ttc ctc aac cgc cac ctt tcg gcc aag ctc ttc cat	593
Asn Gly Val Glu Phe Leu Asn Arg His Leu Ser Ala Lys Leu Phe His	
155 160 165 170	
gac aag gag agc atg cag cca ctg ctt gag ttc ctc agg ctt cac agt	641
Asp Lys Glu Ser Met Gln Pro Leu Leu Glu Phe Leu Arg Leu His Ser	
175 180 185	
tat aag gga aag acc atg atg ttg aat gac aaa gtt caa agc ctg gat	689
Tyr Lys Gly Lys Thr Met Met Leu Asn Asp Lys Val Gln Ser Leu Asp	
190 195 200	
tct ctc caa cat gtt ttg aga aaa gca gaa gag tat ctg att tca gtt	737
Ser Leu Gln His Val Leu Arg Lys Ala Glu Glu Tyr Leu Ile Ser Val	
205 210 215	
gct cct gaa aca ccc tac tcg gaa ttc gaa aac aga ttc cgg gag att	785
Ala Pro Glu Thr Pro Tyr Ser Glu Phe Glu Asn Arg Phe Arg Glu Ile	
220 225 230	
ggt ctg gag agg ggg tgg ggt gac act gcc gag cgt gtc ctc gag atg	833
Gly Leu Glu Arg Gly Trp Gly Asp Thr Ala Glu Arg Val Leu Glu Met	
235 240 245 250	
atc cag ctt ctc ctg gac ctt ctt gaa gcg cct gac cct tgc acc ctc	881
Ile Gln Leu Leu Leu Asp Leu Leu Glu Ala Pro Asp Pro Cys Thr Leu	
255 260 265	
gag aca ttc ctt gga aga gtc cct atg gtc ttc aat gtt gtt atc ctt	929
Glu Thr Phe Leu Gly Arg Val Pro Met Val Phe Asn Val Val Ile Leu	
270 275 280	
tct ccc cat ggt tac ttt gcc caa gat aat gtc ttg gga tac cct gac	977
Ser Pro His Gly Tyr Phe Ala Gln Asp Asn Val Leu Gly Tyr Pro Asp	
285 290 295	
act ggt gga cag gtt gtt tac atc ttg gat caa gtt cgt gcc ttg gag	1025
Thr Gly Gly Gln Val Val Tyr Ile Leu Asp Gln Val Arg Ala Leu Glu	
300 305 310	

aat gag atg ctc aac cgc atc aag aaa caa ggc ctt gat atc acc cct	1073
Asn Glu Met Leu Asn Arg Ile Lys Lys Gln Gly Leu Asp Ile Thr Pro	
315 320 325 330	
cgt att ctc att atc act cgt ctt ctc cct gat gca gta gga act acc	1121
Arg Ile Leu Ile Ile Thr Arg Leu Leu Pro Asp Ala Val Gly Thr Thr	
335 340 345	
tgt ggc caa cgt cta gag agg gta tat gat act gaa tat tgt gac att	1169
Cys Gly Gln Arg Leu Glu Arg Val Tyr Asp Thr Glu Tyr Cys Asp Ile	
350 355 360	
ctc cga gtt cct ttc aga acc gaa aag gga att gtt cgc aaa tgg atc	1217
Leu Arg Val Pro Phe Arg Thr Glu Lys Gly Ile Val Arg Lys Trp Ile	
365 370 375	
tca aga ttc gaa gtc tgg cca tac cta gag act tac act gag gat gtt	1265
Ser Arg Phe Glu Val Trp Pro Tyr Leu Glu Thr Tyr Thr Glu Asp Val	
380 385 390	
gcc ctt gaa ctt gcc aag gag ttg caa gcc aag cca gat ctg att gtt	1313
Ala Leu Glu Leu Ala Lys Glu Leu Gln Ala Lys Pro Asp Leu Ile Val	
395 400 405 410	
gga aac tac agt gat gga aac att gtt gcc tct ttg tta gca cat aaa	1361
Gly Asn Tyr Ser Asp Gly Asn Ile Val Ala Ser Leu Leu Ala His Lys	
415 420 425	
tta gga gta act cag tgt acc att gct cat gct cta gaa aag acc aag	1409
Leu Gly Val Thr Gln Cys Thr Ile Ala His Ala Leu Glu Lys Thr Lys	
430 435 440	
tac cct gag tct gac att tac tgg aaa aaa ttt gaa gag aaa tat cat	1457
Tyr Pro Glu Ser Asp Ile Tyr Trp Lys Lys Phe Glu Glu Lys Tyr His	
445 450 455	
ttc tca tgc caa ttt act gct gat ctt ttt gca atg aac cac aca gac	1505
Phe Ser Cys Gln Phe Thr Ala Asp Leu Phe Ala Met Asn His Thr Asp	
460 465 470	
ttt atc atc acc agc acc ttc caa gag att gct gga agc aag gac act	1553
Phe Ile Ile Thr Ser Thr Phe Gln Glu Ile Ala Gly Ser Lys Asp Thr	
475 480 485 490	
gtt gga cag tat gag agt cac act gcc ttc acc ctt cca gga ctt tac	1601
Val Gly Gln Tyr Glu Ser His Thr Ala Phe Thr Leu Pro Gly Leu Tyr	
495 500 505	
cgt gtt gtt cac ggt att gat ccc ttt gat cca aag ttc aac att gtc	1649
Arg Val Val His Gly Ile Asp Pro Phe Asp Pro Lys Phe Asn Ile Val	
510 515 520	
tct ccc ggt gca gac atg ggc ata tac ttc cca tac act gaa act gag	1697
Ser Pro Gly Ala Asp Met Gly Ile Tyr Phe Pro Tyr Thr Glu Thr Glu	
525 530 535	
cgt agg tta aca gaa ttc cac tct gac att gaa gag ctt ctt tac agc	1745
Arg Arg Leu Thr Glu Phe His Ser Asp Ile Glu Glu Leu Leu Tyr Ser	
540 545 550	
tca gtg gag aat gag gaa cac ata tgc gta ttg aag gac cgc aac aaa	1793

Ser	Val	Glu	Asn	Glu	Glu	His	Ile	Cys	Val	Leu	Lys	Asp	Arg	Asn	Lys	
555					560					565					570	
cca	ata	atc	ttc	acc	atg	gca	agg	ctt	gac	cgt	gtg	aag	aac	atc	acg	1841
Pro	Ile	Ile	Phe	Thr	Met	Ala	Arg	Leu	Asp	Arg	Val	Lys	Asn	Ile	Thr	
			575						580					585		
ggg	ctt	gtc	gag	tgg	tac	ggg	aag	aac	gca	cgc	ctc	cgc	gag	ttg	gtg	1889
Gly	Leu	Val	Glu	Trp	Tyr	Gly	Lys	Asn	Ala	Arg	Leu	Arg	Glu	Leu	Val	
		590					595						600			
aac	ctg	gtg	gtg	gtg	gct	gga	gac	agg	agg	aag	gag	tcg	aag	gac	ttg	1937
Asn	Leu	Val	Val	Val	Ala	Gly	Asp	Arg	Arg	Lys	Glu	Ser	Lys	Asp	Leu	
		605					610						615			
gaa	gag	aag	gcc	gag	atg	aag	aag	atg	tat	ggc	ctc	atc	gag	acc	tac	1985
Glu	Glu	Lys	Ala	Glu	Met	Lys	Lys	Met	Tyr	Gly	Leu	Ile	Glu	Thr	Tyr	
		620				625					630					
aag	ttg	aac	ggc	caa	ttc	aga	tgg	ata	tcc	tct	cag	atg	aac	cgt	gtg	2033
Lys	Leu	Asn	Gly	Gln	Phe	Arg	Trp	Ile	Ser	Ser	Gln	Met	Asn	Arg	Val	
635					640					645					650	
agg	aac	gga	gag	ctc	tac	cgt	gtc	atc	tgt	gac	aca	agg	ggt	gcc	ttt	2081
Arg	Asn	Gly	Glu	Leu	Tyr	Arg	Val	Ile	Cys	Asp	Thr	Arg	Gly	Ala	Phe	
				655					660					665		
gtg	cag	cct	gca	gtt	tat	gag	gcc	ttt	ggg	ttg	act	gtg	gtt	gag	gcc	2129
Val	Gln	Pro	Ala	Val	Tyr	Glu	Ala	Phe	Gly	Leu	Thr	Val	Val	Glu	Ala	
			670					675					680			
atg	act	tgt	ggg	ttg	cca	acg	ttt	gcc	aca	tgc	aat	ggt	ggt	cct	gct	2177
Met	Thr	Cys	Gly	Leu	Pro	Thr	Phe	Ala	Thr	Cys	Asn	Gly	Gly	Pro	Ala	
		685					690					695				
gag	atc	att	gtg	cat	gga	aaa	tct	ggt	tac	cac	att	gat	cct	tac	cat	2225
Glu	Ile	Ile	Val	His	Gly	Lys	Ser	Gly	Tyr	His	Ile	Asp	Pro	Tyr	His	
	700					705					710					
ggt	gac	cat	gct	gct	gag	atc	ctt	gtt	gag	ttc	ttt	gag	aag	agc	aag	2273
Gly	Asp	His	Ala	Ala	Glu	Ile	Leu	Val	Glu	Phe	Phe	Glu	Lys	Ser	Lys	
715					720				725						730	
gct	gat	cca	tct	cac	tgg	gac	aaa	atc	tcc	cag	ggt	gga	ctc	aag	cgt	2321
Ala	Asp	Pro	Ser	His	Trp	Asp	Lys	Ile	Ser	Gln	Gly	Gly	Leu	Lys	Arg	
				735				740						745		
att	cat	gag	aag	tac	aca	tgg	caa	att	tac	tcg	gac	agg	ctc	ttg	aca	2369
Ile	His	Glu	Lys	Tyr	Thr	Trp	Gln	Ile	Tyr	Ser	Asp	Arg	Leu	Leu	Thr	
			750					755					760			
ctc	act	ggt	gtg	tat	ggc	ttc	tgg	aag	cac	gtg	acc	aat	ctt	gaa	cgc	2417
Leu	Thr	Gly	Val	Tyr	Gly	Phe	Trp	Lys	His	Val	Thr	Asn	Leu	Glu	Arg	
		765					770					775				
cgt	gag	agc	aaa	cgt	tac	ctc	gag	atg	ttc	tat	gct	ctc	aag	tac	cgc	2465
Arg	Glu	Ser	Lys	Arg	Tyr	Leu	Glu	Met	Phe	Tyr	Ala	Leu	Lys	Tyr	Arg	
		780				785					790					
aaa	ttg	gct	gag	tct	gtg	ccc	ctt	gct	att	gaa	gag	taa	attcatgttt			2514
Lys	Leu	Ala	Glu	Ser	Val	Pro	Leu	Ala	Ile	Glu	Glu					

795 800 805

gaagagaaca tcaatggaga aaccggcgttt tggtcgtttg aagtcttatg gagctttcat 2574

aaataacgcc attgattttg attgtgatca gcttttggat ttaaagagtc taaaaaaaaag 2634

tgtcgtttgt ttttatattt tgcttaggtt ttgggtaaaa aaagtaaaaa aagaaagtct 2694

tttttttttt tttaaattct ttttgttgtt gcaatattgt gtatgattgg aaatctttaa 2754

attcggggttt tccaccccat caattcaatt taaagatatg ctttcttcca ttcaatttcg 2814

gtaaatactg gtccagagaa gagaaaaatg gcaatctcca taccagtag acaacttttc 2874

atagaaaact cgtgccgctt ctcttggttc tcttcaagaa tccacaccaa cccctcacia 2934

ttggtacctt acttctgttc caattttatt tattgggttg gagtgttcct ttgtgttgat 2994

ttcaattgcc tgttgcttgt gccttcacag gaccaaagga atttgcagca ctctggantt 3054

gccttggtca ttggcganga atatttgaga ggtacgataa anacaggagt gggaaaattg 3114

atccattana actcanagat gctctgtatg gtattggcta tgctgtacca ngctcngttc 3174

tnnnntact gctttcaaag tatgggtgat gaantggtag ganggttgaa cttgggtttg 3234

acag 3238

<210> 16

<211> 806

<212> PRT

<213> Glycine max

<220>

<221> misc_feature

<222> (3052)..(3217)

<223> n located at 3052, 3130, 3072, 3123, 3096, 3165, 3170, 3176-3180, 3208, 3217, each n is a or c or g or t/u, unknown, or other

<400> 16

Met Ala Asn His Pro Leu Thr His Ser His Ser Phe Arg Glu Arg Phe
1 5 10 15

Asp Glu Thr Leu Thr Gly His Arg Asn Glu Ile Leu Ala Leu Leu Ser
20 25 30

Arg Leu Glu Ala Lys Gly Lys Gly Ile Leu Gln His His Gln Val Val
35 40 45

Ala Glu Phe Glu Glu Ile Pro Glu Glu Ser Arg Lys Lys Leu Gln Asp
50 55 60

Gly Val Phe Gly Glu Val Leu Arg Ser Thr Gln Glu Ala Ile Val Leu
65 70 75 80

Pro Pro Phe Val Ala Leu Ala Val Arg Pro Arg Pro Gly Val Trp Glu
 85 90 95

Tyr Leu Arg Val Asn Val His Met Leu Val Val Asp Glu Leu Arg Pro
 100 105 110

Ala Glu Tyr Leu Arg Phe Lys Glu Glu Leu Val Glu Gly Ser Ser Asn
 115 120 125

Gly Asn Phe Val Leu Glu Leu Asp Phe Glu Pro Phe Asn Ala Ser Phe
 130 135 140

Pro Arg Pro Thr Leu Asn Lys Ser Ile Gly Asn Gly Val Glu Phe Leu
 145 150 155 160

Asn Arg His Leu Ser Ala Lys Leu Phe His Asp Lys Glu Ser Met Gln
 165 170 175

Pro Leu Leu Glu Phe Leu Arg Leu His Ser Tyr Lys Gly Lys Thr Met
 180 185 190

Met Leu Asn Asp Lys Val Gln Ser Leu Asp Ser Leu Gln His Val Leu
 195 200 205

Arg Lys Ala Glu Glu Tyr Leu Ile Ser Val Ala Pro Glu Thr Pro Tyr
 210 215 220

Ser Glu Phe Glu Asn Arg Phe Arg Glu Ile Gly Leu Glu Arg Gly Trp
 225 230 235 240

Gly Asp Thr Ala Glu Arg Val Leu Glu Met Ile Gln Leu Leu Leu Asp
 245 250 255

Leu Leu Glu Ala Pro Asp Pro Cys Thr Leu Glu Thr Phe Leu Gly Arg
 260 265 270

Val Pro Met Val Phe Asn Val Val Ile Leu Ser Pro His Gly Tyr Phe
 275 280 285

Ala Gln Asp Asn Val Leu Gly Tyr Pro Asp Thr Gly Gly Gln Val Val
 290 295 300

Tyr Ile Leu Asp Gln Val Arg Ala Leu Glu Asn Glu Met Leu Asn Arg
 305 310 315 320

Ile Lys Lys Gln Gly Leu Asp Ile Thr Pro Arg Ile Leu Ile Ile Thr
 325 330 335
 Arg Leu Leu Pro Asp Ala Val Gly Thr Thr Cys Gly Gln Arg Leu Glu
 340 345 350
 Arg Val Tyr Asp Thr Glu Tyr Cys Asp Ile Leu Arg Val Pro Phe Arg
 355 360 365
 Thr Glu Lys Gly Ile Val Arg Lys Trp Ile Ser Arg Phe Glu Val Trp
 370 375 380
 Pro Tyr Leu Glu Thr Tyr Thr Glu Asp Val Ala Leu Glu Leu Ala Lys
 385 390 395 400
 Glu Leu Gln Ala Lys Pro Asp Leu Ile Val Gly Asn Tyr Ser Asp Gly
 405 410 415
 Asn Ile Val Ala Ser Leu Leu Ala His Lys Leu Gly Val Thr Gln Cys
 420 425 430
 Thr Ile Ala His Ala Leu Glu Lys Thr Lys Tyr Pro Glu Ser Asp Ile
 435 440 445
 Tyr Trp Lys Lys Phe Glu Glu Lys Tyr His Phe Ser Cys Gln Phe Thr
 450 455 460
 Ala Asp Leu Phe Ala Met Asn His Thr Asp Phe Ile Ile Thr Ser Thr
 465 470 475 480
 Phe Gln Glu Ile Ala Gly Ser Lys Asp Thr Val Gly Gln Tyr Glu Ser
 485 490 495
 His Thr Ala Phe Thr Leu Pro Gly Leu Tyr Arg Val Val His Gly Ile
 500 505 510
 Asp Pro Phe Asp Pro Lys Phe Asn Ile Val Ser Pro Gly Ala Asp Met
 515 520 525
 Gly Ile Tyr Phe Pro Tyr Thr Glu Thr Glu Arg Arg Leu Thr Glu Phe
 530 535 540
 His Ser Asp Ile Glu Glu Leu Leu Tyr Ser Ser Val Glu Asn Glu Glu
 545 550 555 560
 His Ile Cys Val Leu Lys Asp Arg Asn Lys Pro Ile Ile Phe Thr Met

58

<210> 17
 <211> 2418
 <212> DNA
 <213> Glycine max

<220>
 <221> CDS
 <222> (1)..(2418)

<400> 17
 atg gca aat cat cct ttg aca cac tct cac tct ttc cgc gag agg ttt 48
 Met Ala Asn His Pro Leu Thr His Ser His Ser Phe Arg Glu Arg Phe
 1 5 10 15
 gat gaa act ctc act ggt cac agg aac gaa att ttg gcc ctt ctg tca 96
 Asp Glu Thr Leu Thr Gly His Arg Asn Glu Ile Leu Ala Leu Leu Ser
 20 25 30
 agg ctt gaa gcc aag ggc aag gga atc ctg caa cat cac caa gtg gtt 144
 Arg Leu Glu Ala Lys Gly Lys Gly Ile Leu Gln His His Gln Val Val
 35 40 45
 gca gag ttt gaa gaa atc cct gag gag agc aga aag aaa ctc caa gat 192
 Ala Glu Phe Glu Glu Ile Pro Glu Glu Ser Arg Lys Lys Leu Gln Asp
 50 55 60
 ggt gtc ttt gga gaa gtt ttg aga tcc aca cag gaa gcc ata gtg ctg 240
 Gly Val Phe Gly Glu Val Leu Arg Ser Thr Gln Glu Ala Ile Val Leu
 65 70 75 80
 cca cca ttt gta gct ctt gct gtt cga cca agg cct ggt gta tgg gaa 288
 Pro Pro Phe Val Ala Leu Ala Val Arg Pro Arg Pro Gly Val Trp Glu
 85 90 95
 tat ctg cgt gtg aat gtg cac atg ctt gtt gtt gat gag ctt cgt cct 336
 Tyr Leu Arg Val Asn Val His Met Leu Val Val Asp Glu Leu Arg Pro
 100 105 110
 gct gag tat ctg cgt ttc aag gag gag ctt gtt gag gga agt tca aac 384
 Ala Glu Tyr Leu Arg Phe Lys Glu Glu Leu Val Glu Gly Ser Ser Asn
 115 120 125
 ggc aac ttt gtg ctt gag ttg gac ttt gaa ccg ttt aat gca tcc ttc 432
 Gly Asn Phe Val Leu Glu Leu Asp Phe Glu Pro Phe Asn Ala Ser Phe
 130 135 140
 cct cgc cca act ctg aac aag tcc att gga aat ggc gtc gag ttc ctc 480
 Pro Arg Pro Thr Leu Asn Lys Ser Ile Gly Asn Gly Val Glu Phe Leu
 145 150 155 160
 aac cgc cac ctt tcg gcc aag ctc ttc cat gac aag gag agc atg cag 528
 Asn Arg His Leu Ser Ala Lys Leu Phe His Asp Lys Glu Ser Met Gln
 165 170 175
 cca ctg ctt gag ttc ctc agg ctt cac agt tat aag gga aag acc atg 576
 Pro Leu Leu Glu Phe Leu Arg Leu His Ser Tyr Lys Gly Lys Thr Met
 180 185 190
 atg ttg aat gac aaa gtt caa agc ctg gat tct ctc caa cat gtt ttg 624

Met	Leu	Asn	Asp	Lys	Val	Gln	Ser	Leu	Asp	Ser	Leu	Gln	His	Val	Leu		
		195					200					205					
aga	aaa	gca	gaa	gag	tat	ctg	att	tca	gtt	gct	cct	gaa	aca	ccc	tac	672	
Arg	Lys	Ala	Glu	Glu	Tyr	Leu	Ile	Ser	Val	Ala	Pro	Glu	Thr	Pro	Tyr		
	210					215					220						
tcg	gaa	ttc	gaa	aac	aga	ttc	cgg	gag	att	ggt	ctg	gag	agg	ggg	tgg	720	
Ser	Glu	Phe	Glu	Asn	Arg	Phe	Arg	Glu	Ile	Gly	Leu	Glu	Arg	Gly	Trp		
225				230						235					240		
ggt	gac	act	gcc	gag	cgt	gtc	ctc	gag	atg	atc	cag	ctt	ctc	ctg	gac	768	
Gly	Asp	Thr	Ala	Glu	Arg	Val	Leu	Glu	Met	Ile	Gln	Leu	Leu	Leu	Asp		
			245						250					255			
ctt	ctt	gaa	gcg	cct	gac	cct	tgc	acc	ctc	gag	aca	ttc	ctt	gga	aga	816	
Leu	Leu	Glu	Ala	Pro	Asp	Pro	Cys	Thr	Leu	Glu	Thr	Phe	Leu	Gly	Arg		
		260					265						270				
gtc	cct	atg	gtc	ttc	aat	gtt	gtt	atc	ctt	tct	ccc	cat	ggt	tac	ttt	864	
Val	Pro	Met	Val	Phe	Asn	Val	Val	Ile	Leu	Ser	Pro	His	Gly	Tyr	Phe		
	275					280						285					
gcc	caa	gat	aat	gtc	ttg	gga	tac	cct	gac	act	ggt	gga	cag	gtt	gtt	912	
Ala	Gln	Asp	Asn	Val	Leu	Gly	Tyr	Pro	Asp	Thr	Gly	Gly	Gln	Val	Val		
	290					295					300						
tac	atc	ttg	gat	caa	gtt	cgt	gcc	ttg	gag	aat	gag	atg	ctc	aac	cgc	960	
Tyr	Ile	Leu	Asp	Gln	Val	Arg	Ala	Leu	Glu	Asn	Glu	Met	Leu	Asn	Arg		
305				310						315					320		
atc	aag	aaa	caa	ggc	ctt	gat	atc	acc	cct	cgt	att	ctc	att	atc	act	1008	
Ile	Lys	Lys	Gln	Gly	Leu	Asp	Ile	Thr	Pro	Arg	Ile	Leu	Ile	Ile	Thr		
			325						330					335			
cgt	ctt	ctc	cct	gat	gca	gta	gga	act	acc	tgt	ggc	caa	cgt	cta	gag	1056	
Arg	Leu	Leu	Pro	Asp	Ala	Val	Gly	Thr	Thr	Cys	Gly	Gln	Arg	Leu	Glu		
			340					345					350				
agg	gta	tat	gat	act	gaa	tat	tgt	gac	att	ctc	cga	gtt	cct	ttc	aga	1104	
Arg	Val	Tyr	Asp	Thr	Glu	Tyr	Cys	Asp	Ile	Leu	Arg	Val	Pro	Phe	Arg		
	355						360					365					
acc	gaa	aag	gga	att	gtt	cgc	aaa	tgg	atc	tca	aga	ttc	gaa	gtc	tgg	1152	
Thr	Glu	Lys	Gly	Ile	Val	Arg	Lys	Trp	Ile	Ser	Arg	Phe	Glu	Val	Trp		
	370					375					380						
cca	tac	cta	gag	act	tac	act	gag	gat	gtt	gcc	ctt	gaa	ctt	gcc	aag	1200	
Pro	Tyr	Leu	Glu	Thr	Tyr	Thr	Glu	Asp	Val	Ala	Leu	Glu	Leu	Ala	Lys		
385				390						395					400		
gag	ttg	caa	gcc	aag	cca	gat	ctg	att	gtt	gga	aac	tac	agt	gat	gga	1248	
Glu	Leu	Gln	Ala	Lys	Pro	Asp	Leu	Ile	Val	Gly	Asn	Tyr	Ser	Asp	Gly		
			405						410					415			
aac	att	gtt	gcc	tct	ttg	tta	gca	cat	aaa	tta	gga	gta	act	cag	tgt	1296	
Asn	Ile	Val	Ala	Ser	Leu	Leu	Ala	His	Lys	Leu	Gly	Val	Thr	Gln	Cys		
		420						425					430				
acc	att	gct	cat	gct	cta	gaa	aag	acc	aag	tac	cct	gag	tct	gac	att	1344	
Thr	Ile	Ala	His	Ala	Leu	Glu	Lys	Thr	Lys	Tyr	Pro	Glu	Ser	Asp	Ile		

435	440	445	
tac tgg aaa aaa ttt gaa gag aaa tat cat ttc tca tgc caa ttt act			1392
Tyr Trp Lys Lys Phe Glu Glu Lys Tyr His Phe Ser Cys Gln Phe Thr			
450	455	460	
gct gat ctt ttt gca atg aac cac aca gac ttt atc atc acc agc acc			1440
Ala Asp Leu Phe Ala Met Asn His Thr Asp Phe Ile Ile Thr Ser Thr			
465	470	475	480
ttc caa gag att gct gga agc aag gac act gtt gga cag tat gag agt			1488
Phe Gln Glu Ile Ala Gly Ser Lys Asp Thr Val Gly Gln Tyr Glu Ser			
485	490	495	
cac act gcc ttc acc ctt cca gga ctt tac cgt gtt gtt cac ggt att			1536
His Thr Ala Phe Thr Leu Pro Gly Leu Tyr Arg Val Val His Gly Ile			
500	505	510	
gat ccc ttt gat cca aag ttc aac att gtc tct ccc ggt gca gac atg			1584
Asp Pro Phe Asp Pro Lys Phe Asn Ile Val Ser Pro Gly Ala Asp Met			
515	520	525	
ggc ata tac ttc cca tac act gaa act gag cgt agg tta aca gaa ttc			1632
Gly Ile Tyr Phe Pro Tyr Thr Glu Thr Glu Arg Arg Leu Thr Glu Phe			
530	535	540	
cac tct gac att gaa gag ctt ctt tac agc tca gtg gag aat gag gaa			1680
His Ser Asp Ile Glu Glu Leu Leu Tyr Ser Ser Val Glu Asn Glu Glu			
545	550	555	560
cac ata tgc gta ttg aag gac cgc aac aaa cca ata atc ttc acc atg			1728
His Ile Cys Val Leu Lys Asp Arg Asn Lys Pro Ile Ile Phe Thr Met			
565	570	575	
gca agg ctt gac cgt gtg aag aac atc acg ggg ctt gtc gag tgg tac			1776
Ala Arg Leu Asp Arg Val Lys Asn Ile Thr Gly Leu Val Glu Trp Tyr			
580	585	590	
ggg aag aac gca cgc ctc cgc gag ttg gtg aac ctg gtg gtg gtg gct			1824
Gly Lys Asn Ala Arg Leu Arg Glu Leu Val Asn Leu Val Val Val Ala			
595	600	605	
gga gac agg agg aag gag tcg aag gac ttg gaa gag aag gcc gag atg			1872
Gly Asp Arg Arg Lys Glu Ser Lys Asp Leu Glu Glu Lys Ala Glu Met			
610	615	620	
aag aag atg tat ggc ctc atc gag acc tac aag ttg aac ggc caa ttc			1920
Lys Lys Met Tyr Gly Leu Ile Glu Thr Tyr Lys Leu Asn Gly Gln Phe			
625	630	635	640
aga tgg ata tcc tct cag atg aac cgt gtg agg aac gga gag ctc tac			1968
Arg Trp Ile Ser Ser Gln Met Asn Arg Val Arg Asn Gly Glu Leu Tyr			
645	650	655	
cgt gtc atc tgt gac aca agg ggt gcc ttt gtg cag cct gca gtt tat			2016
Arg Val Ile Cys Asp Thr Arg Gly Ala Phe Val Gln Pro Ala Val Tyr			
660	665	670	
gag gcc ttt ggg ttg act gtg gtt gag gcc atg act tgt ggg ttg cca			2064
Glu Ala Phe Gly Leu Thr Val Val Glu Ala Met Thr Cys Gly Leu Pro			
675	680	685	

acg ttt gcc aca tgc aat ggt ggt cct gct gag atc att gtg cat gga 2112
 Thr Phe Ala Thr Cys Asn Gly Gly Pro Ala Glu Ile Ile Val His Gly
 690 695 700
 aaa tct ggt tac cac att gat cct tac cat ggt gac cat gct gct gag 2160
 Lys Ser Gly Tyr His Ile Asp Pro Tyr His Gly Asp His Ala Ala Glu
 705 710 715 720
 atc ctt gtt gag ttc ttt gag aag agc aag gct gat cca tct cac tgg 2208
 Ile Leu Val Glu Phe Phe Glu Lys Ser Lys Ala Asp Pro Ser His Trp
 725 730 735
 gac aaa atc tcc cag ggt gga ctc aag cgt att cat gag aag tac aca 2256
 Asp Lys Ile Ser Gln Gly Gly Leu Lys Arg Ile His Glu Lys Tyr Thr
 740 745 750
 tgg caa att tac tcg gac agg ctc ttg aca ctc act ggt gtg tat ggc 2304
 Trp Gln Ile Tyr Ser Asp Arg Leu Leu Thr Leu Thr Gly Val Tyr Gly
 755 760 765
 ttc tgg aag cac gtg acc aat ctt gaa cgc cgt gag agc aaa cgt tac 2352
 Phe Trp Lys His Val Thr Asn Leu Glu Arg Arg Glu Ser Lys Arg Tyr
 770 775 780
 ctc gag atg ttc tat gct ctc aag tac cgc aaa ttg gct gag tct gtg 2400
 Leu Glu Met Phe Tyr Ala Leu Lys Tyr Arg Lys Leu Ala Glu Ser Val
 785 790 795 800
 ccc ctt gct att gaa gag 2418
 Pro Leu Ala Ile Glu Glu
 805

<210> 18
 <211> 806
 <212> PRT
 <213> Glycine max

<400> 18
 Met Ala Asn His Pro Leu Thr His Ser His Ser Phe Arg Glu Arg Phe
 1 5 10 15

Asp Glu Thr Leu Thr Gly His Arg Asn Glu Ile Leu Ala Leu Leu Ser
 20 25 30

Arg Leu Glu Ala Lys Gly Lys Gly Ile Leu Gln His His Gln Val Val
 35 40 45

Ala Glu Phe Glu Glu Ile Pro Glu Glu Ser Arg Lys Lys Leu Gln Asp
 50 55 60

Gly Val Phe Gly Glu Val Leu Arg Ser Thr Gln Glu Ala Ile Val Leu
 65 70 75 80

Pro Pro Phe Val Ala Leu Ala Val Arg Pro Arg Pro Gly Val Trp Glu

85	90	95
Tyr Leu Arg Val Asn Val His Met Leu Val Val Asp Glu Leu Arg Pro 100 105 110		
Ala Glu Tyr Leu Arg Phe Lys Glu Glu Leu Val Glu Gly Ser Ser Asn 115 120 125		
Gly Asn Phe Val Leu Glu Leu Asp Phe Glu Pro Phe Asn Ala Ser Phe 130 135 140		
Pro Arg Pro Thr Leu Asn Lys Ser Ile Gly Asn Gly Val Glu Phe Leu 145 150 155 160		
Asn Arg His Leu Ser Ala Lys Leu Phe His Asp Lys Glu Ser Met Gln 165 170 175		
Pro Leu Leu Glu Phe Leu Arg Leu His Ser Tyr Lys Gly Lys Thr Met 180 185 190		
Met Leu Asn Asp Lys Val Gln Ser Leu Asp Ser Leu Gln His Val Leu 195 200 205		
Arg Lys Ala Glu Glu Tyr Leu Ile Ser Val Ala Pro Glu Thr Pro Tyr 210 215 220		
Ser Glu Phe Glu Asn Arg Phe Arg Glu Ile Gly Leu Glu Arg Gly Trp 225 230 235 240		
Gly Asp Thr Ala Glu Arg Val Leu Glu Met Ile Gln Leu Leu Leu Asp 245 250 255		
Leu Leu Glu Ala Pro Asp Pro Cys Thr Leu Glu Thr Phe Leu Gly Arg 260 265 270		
Val Pro Met Val Phe Asn Val Val Ile Leu Ser Pro His Gly Tyr Phe 275 280 285		
Ala Gln Asp Asn Val Leu Gly Tyr Pro Asp Thr Gly Gly Gln Val Val 290 295 300		
Tyr Ile Leu Asp Gln Val Arg Ala Leu Glu Asn Glu Met Leu Asn Arg 305 310 315 320		
Ile Lys Lys Gln Gly Leu Asp Ile Thr Pro Arg Ile Leu Ile Ile Thr 325 330 335		

Arg Leu Leu Pro Asp Ala Val Gly Thr Thr Cys Gly Gln Arg Leu Glu
340 345 350

Arg Val Tyr Asp Thr Glu Tyr Cys Asp Ile Leu Arg Val Pro Phe Arg
355 360 365

Thr Glu Lys Gly Ile Val Arg Lys Trp Ile Ser Arg Phe Glu Val Trp
370 375 380

Pro Tyr Leu Glu Thr Tyr Thr Glu Asp Val Ala Leu Glu Leu Ala Lys
385 390 395 400

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

Asn Ile Val Ala Ser Leu Leu Ala His Lys Leu Gly Val Thr Gln Cys
420 425 430

Thr Ile Ala His Ala Leu Glu Lys Thr Lys Tyr Pro Glu Ser Asp Ile
435 440 445

Tyr Trp Lys Lys Phe Glu Glu Lys Tyr His Phe Ser Cys Gln Phe Thr
450 455 460

Ala Asp Leu Phe Ala Met Asn His Thr Asp Phe Ile Ile Thr Ser Thr
465 470 475 480

Phe Gln Glu Ile Ala Gly Ser Lys Asp Thr Val Gly Gln Tyr Glu Ser
485 490 495

His Thr Ala Phe Thr Leu Pro Gly Leu Tyr Arg Val Val His Gly Ile
500 505 510

Asp Pro Phe Asp Pro Lys Phe Asn Ile Val Ser Pro Gly Ala Asp Met
515 520 525

Gly Ile Tyr Phe Pro Tyr Thr Glu Thr Glu Arg Arg Leu Thr Glu Phe
530 535 540

His Ser Asp Ile Glu Glu Leu Leu Tyr Ser Ser Val Glu Asn Glu Glu
545 550 555 560

His Ile Cys Val Leu Lys Asp Arg Asn Lys Pro Ile Ile Phe Thr Met
565 570 575

Ala Arg Leu Asp Arg Val Lys Asn Ile Thr Gly Leu Val Glu Trp Tyr
580 585 590

Gly Lys Asn Ala Arg Leu Arg Glu Leu Val Asn Leu Val Val Val Ala
595 600 605

Gly Asp Arg Arg Lys Glu Ser Lys Asp Leu Glu Glu Lys Ala Glu Met
610 615 620

Lys Lys Met Tyr Gly Leu Ile Glu Thr Tyr Lys Leu Asn Gly Gln Phe
625 630 635 640

Arg Trp Ile Ser Ser Gln Met Asn Arg Val Arg Asn Gly Glu Leu Tyr
645 650 655

Arg Val Ile Cys Asp Thr Arg Gly Ala Phe Val Gln Pro Ala Val Tyr
660 665 670

Glu Ala Phe Gly Leu Thr Val Val Glu Ala Met Thr Cys Gly Leu Pro
675 680 685

Thr Phe Ala Thr Cys Asn Gly Gly Pro Ala Glu Ile Ile Val His Gly
690 695 700

Lys Ser Gly Tyr His Ile Asp Pro Tyr His Gly Asp His Ala Ala Glu
705 710 715 720

Ile Leu Val Glu Phe Phe Glu Lys Ser Lys Ala Asp Pro Ser His Trp
725 730 735

Asp Lys Ile Ser Gln Gly Gly Leu Lys Arg Ile His Glu Lys Tyr Thr
740 745 750

Trp Gln Ile Tyr Ser Asp Arg Leu Leu Thr Leu Thr Gly Val Tyr Gly
755 760 765

Phe Trp Lys His Val Thr Asn Leu Glu Arg Arg Glu Ser Lys Arg Tyr
770 775 780

Leu Glu Met Phe Tyr Ala Leu Lys Tyr Arg Lys Leu Ala Glu Ser Val
785 790 795 800

Pro Leu Ala Ile Glu Glu
805

<210> 19
 <211> 2765
 <212> DNA
 <213> Zea mays

<220>
 <221> CDS
 <222> (56)..(2464)

<400> 19
 ctccattgga ctgcttgctc cctgttgacc attgggtatt ctgaaccatc gagcc atg 58
 Met
 1

gct gcc aag ctg act cgc ctc cac agt ctt cgc gaa cgc ctt ggt gcc 106
 Ala Ala Lys Leu Thr Arg Leu His Ser Leu Arg Glu Arg Leu Gly Ala
 5 10 15

acc ttc tcc tcc cat ccc aat gaa ctg ata gca ctc ttt tcc agg tat 154
 Thr Phe Ser Ser His Pro Asn Glu Leu Ile Ala Leu Phe Ser Arg Tyr
 20 25 30

gtt cac cag ggc aag gga atg ctt cag cgc cat cag ctg ctt gcg gag 202
 Val His Gln Gly Lys Gly Met Leu Gln Arg His Gln Leu Leu Ala Glu
 35 40 45

ttt gat gcc ctg ttt gat agt gac aag gag aag tat gca ccc ttt gaa 250
 Phe Asp Ala Leu Phe Asp Ser Asp Lys Glu Lys Tyr Ala Pro Phe Glu
 50 55 60 65

gac att ctt cgt gct gct cag gaa gca att gtg ctc ccc cca tgg gtt 298
 Asp Ile Leu Arg Ala Ala Gln Glu Ala Ile Val Leu Pro Pro Trp Val
 70 75 80

gca ctt gct atc agg cca agg cct ggt gtc tgg gat tac att cgg gtg 346
 Ala Leu Ala Ile Arg Pro Arg Pro Gly Val Trp Asp Tyr Ile Arg Val
 85 90 95

aat gta agt gag ctg gct gtg gag gag ctg agt gtt tct gag tac ttg 394
 Asn Val Ser Glu Leu Ala Val Glu Glu Leu Ser Val Ser Glu Tyr Leu
 100 105 110

gca ttc aag gaa cag ctg gtg gat gga caa tcc aac agc aac ttt gtg 442
 Ala Phe Lys Glu Gln Leu Val Asp Gly Gln Ser Asn Ser Asn Phe Val
 115 120 125

ctt gag ctt gat ttt gag ccc ttc aat gcc tcc ttt cct cgt cct tcc 490
 Leu Glu Leu Asp Phe Glu Pro Phe Asn Ala Ser Phe Pro Arg Pro Ser
 130 135 140 145

atg tcg aag tcc atc gga aat gga gtg caa ttc ctt aac cga cac ctg 538
 Met Ser Lys Ser Ile Gly Asn Gly Val Gln Phe Leu Asn Arg His Leu
 150 155 160

tcg tcc aag ttg ttc cag gac aag gag agt ttg tac ccc ttg ctg aac 586
 Ser Ser Lys Leu Phe Gln Asp Lys Glu Ser Leu Tyr Pro Leu Leu Asn
 165 170 175

ttc ctc aag gct cat aac tac aag ggc acg acg atg atg ttg aat gac 634
 Phe Leu Lys Ala His Asn Tyr Lys Gly Thr Thr Met Met Leu Asn Asp

180	185	190	
aga atc caa agc ctt cgt ggt ctc caa tca tcc ctg aga aag gca gag Arg Ile Gln Ser Leu Arg Gly Leu Gln Ser Ser Leu Arg Lys Ala Glu 195 200 205			682
gag tat cta ctg agt gtt cct caa gac act ccc tac tcg gag ttc aac Glu Tyr Leu Leu Ser Val Pro Gln Asp Thr Pro Tyr Ser Glu Phe Asn 210 215 220 225			730
cat agg ttc caa gag ctt ggt ttg gag aag ggt tgg ggt gac act gcg His Arg Phe Gln Glu Leu Gly Leu Glu Lys Gly Trp Gly Asp Thr Ala 230 235 240			778
aag cgt gta ctc gac aca ctc cac ttg ctt ctt gac ctt ctt gag gcc Lys Arg Val Leu Asp Thr Leu His Leu Leu Leu Asp Leu Leu Glu Ala 245 250 255			826
cct gat cct gcc aac ttg gag aag ttc ctt gga act ata cca atg atg Pro Asp Pro Ala Asn Leu Glu Lys Phe Leu Gly Thr Ile Pro Met Met 260 265 270			874
ttc aat gtt gtt atc ctt tct cct cat ggc tac ttc gct cag tcc aat Phe Asn Val Val Ile Leu Ser Pro His Gly Tyr Phe Ala Gln Ser Asn 275 280 285			922
gtg ctt gga tac cct gac act ggc ggt cag gtt gtg tac att ctg gat Val Leu Gly Tyr Pro Asp Thr Gly Gly Gln Val Val Tyr Ile Leu Asp 290 295 300 305			970
caa gtc cgt gct ttg gag aat gag atg ctt ctg agg att aag cag caa Gln Val Arg Ala Leu Glu Asn Glu Met Leu Leu Arg Ile Lys Gln Gln 310 315 320			1018
ggc ctt gat atc act ccg aag atc ctc att gtt acc agg ctg ttg cct Gly Leu Asp Ile Thr Pro Lys Ile Leu Ile Val Thr Arg Leu Leu Pro 325 330 335			1066
gat gct gct ggg act acg tgc ggt cag cgg ctg gag aag gtc att ggt Asp Ala Ala Gly Thr Thr Cys Gly Gln Arg Leu Glu Lys Val Ile Gly 340 345 350			1114
act gag cac aca gac atc att cgc gtt ccg ttc aga aat gag aat ggc Thr Glu His Thr Asp Ile Ile Arg Val Pro Phe Arg Asn Glu Asn Gly 355 360 365			1162
atc ctc cgc aag tgg atc tct cgt ttt gat gtc tgg cca tac ctg gag Ile Leu Arg Lys Trp Ile Ser Arg Phe Asp Val Trp Pro Tyr Leu Glu 370 375 380 385			1210
aca tac act gag gat gtt tcc agt gaa ata atg aaa gaa atg cag gcc Thr Tyr Thr Glu Asp Val Ser Ser Glu Ile Met Lys Glu Met Gln Ala 390 395 400			1258
aag cct gac ctt atc att ggc aac tac agc gat ggc aac cta gtc gcc Lys Pro Asp Leu Ile Ile Gly Asn Tyr Ser Asp Gly Asn Leu Val Ala 405 410 415			1306
act ctg ctc gcg cac aag ttg gga gtc act cag tgt acc atc gct cat Thr Leu Leu Ala His Lys Leu Gly Val Thr Gln Cys Thr Ile Ala His 420 425 430			1354

gcc ttg gag aaa acc aaa tac ccc aac tcg gac atc tac ttg gac aag Ala Leu Glu Lys Thr Lys Tyr Pro Asn Ser Asp Ile Tyr Leu Asp Lys 435 440 445	1402
ttc gac agc cag tac cac ttc tct tgc cag ttc aca gct gac ctt att Phe Asp Ser Gln Tyr His Phe Ser Cys Gln Phe Thr Ala Asp Leu Ile 450 455 460 465	1450
gcc atg aac cac acc gat ttc atc atc acc agc aca ttc caa gaa atc Ala Met Asn His Thr Asp Phe Ile Ile Thr Ser Thr Phe Gln Glu Ile 470 475 480	1498
gcg gga agc aag gac acc gtg ggg cag tac gag tcc cac atc gcg ttc Ala Gly Ser Lys Asp Thr Val Gly Gln Tyr Glu Ser His Ile Ala Phe 485 490 495	1546
act ctt cct ggg ctc tac cgt gtc gtc cat ggc atc gat gtt ttc gat Thr Leu Pro Gly Leu Tyr Arg Val Val His Gly Ile Asp Val Phe Asp 500 505 510	1594
ccc aag ttc aac att gtc tcc cct gga gca gac atg agt gtt tac tac Pro Lys Phe Asn Ile Val Ser Pro Gly Ala Asp Met Ser Val Tyr Tyr 515 520 525	1642
ccg tat acg gaa acc gac aag aga ctc act gcc ttc cat cct gaa atc Pro Tyr Thr Glu Thr Asp Lys Arg Leu Thr Ala Phe His Pro Glu Ile 530 535 540 545	1690
gag gag ctc atc tac agc gac gtc gag aac tcc gag cac aag ttc gtg Glu Glu Leu Ile Tyr Ser Asp Val Glu Asn Ser Glu His Lys Phe Val 550 555 560	1738
ctg aag gac aag aag aag ccg atc atc ttc tcg atg gcg cgt ctc gac Leu Lys Asp Lys Lys Lys Pro Ile Ile Phe Ser Met Ala Arg Leu Asp 565 570 575	1786
cgc gtg aag aac atg aca ggc ctg gtc gag atg tac ggc aag aac gcg Arg Val Lys Asn Met Thr Gly Leu Val Glu Met Tyr Gly Lys Asn Ala 580 585 590	1834
cgc ctg agg gag ctg gcg aac ctc gtg atc gtt gcc ggc gac cac ggc Arg Leu Arg Glu Leu Ala Asn Leu Val Ile Val Ala Gly Asp His Gly 595 600 605	1882
aag gag tcc aag gac agg gag gag cag gcg gag ttc aag aag atg tac Lys Glu Ser Lys Asp Arg Glu Glu Gln Ala Glu Phe Lys Lys Met Tyr 610 615 620 625	1930
agc ctc atc gac gag tac aag ttg aag ggc cat atc cgg tgg atc tcg Ser Leu Ile Asp Glu Tyr Lys Leu Lys Gly His Ile Arg Trp Ile Ser 630 635 640	1978
gcg cag atg aac cgc gtc cgc aac ggg gag ctg tac cgc tac att tgc Ala Gln Met Asn Arg Val Arg Asn Gly Glu Leu Tyr Arg Tyr Ile Cys 645 650 655	2026
gat acg aag ggc gca ttc gtg cag cct gcg ttc tac gaa gcg ttc ggc Asp Thr Lys Gly Ala Phe Val Gln Pro Ala Phe Tyr Glu Ala Phe Gly 660 665 670	2074

ctg act gtg atc gag tcc atg acg tgc ggt ctg cca acg atc gcg acc 2122
 Leu Thr Val Ile Glu Ser Met Thr Cys Gly Leu Pro Thr Ile Ala Thr
 675 680 685

tgc cat ggt ggc cct gct gag atc atc gtg gac ggg gta tct ggc ctg 2170
 Cys His Gly Gly Pro Ala Glu Ile Ile Val Asp Gly Val Ser Gly Leu
 690 695 700 705

cac att gac cct tac cac agc gac aag gcc gcg gat atc ctg gtc aac 2218
 His Ile Asp Pro Tyr His Ser Asp Lys Ala Ala Asp Ile Leu Val Asn
 710 715 720

ttc ttt gac aaa tgc aag gca gat ccg agc tac tgg gac aag atc tca 2266
 Phe Phe Asp Lys Cys Lys Ala Asp Pro Ser Tyr Trp Asp Lys Ile Ser
 725 730 735

cag ggc ggc ctg cag aga att tat gag aag tac acc tgg aag ctc tac 2314
 Gln Gly Gly Leu Gln Arg Ile Tyr Glu Lys Tyr Thr Trp Lys Leu Tyr
 740 745 750

tcc gag agg ctg atg acc ctg acc ggc gtg tac ggg ttc tgg aag tac 2362
 Ser Glu Arg Leu Met Thr Leu Thr Gly Val Tyr Gly Phe Trp Lys Tyr
 755 760 765

gtg agc aac ctg gag agg cgc gag acc cgc cgc tac atc gag atg ttc 2410
 Val Ser Asn Leu Glu Arg Arg Glu Thr Arg Arg Tyr Ile Glu Met Phe
 770 775 780 785

tac gcc ctg aag tac cgt agc ctg gca agc cag gtt ccg ctg tcc ttc 2458
 Tyr Ala Leu Lys Tyr Arg Ser Leu Ala Ser Gln Val Pro Leu Ser Phe
 790 795 800

gat tag tacggggaaa gaaggagaag aagaagccca ggccggagaa ccatcgctg 2514
 Asp

catttcgatac tgtttcaccg caattcgcat tgtagtcgt gtattggagt tatgtgtact 2574

tggtttccaa gaactttggt tccttctcgt tttttttcct tgtttgagcg tttttgggca 2634

gcgctggcct ggttcttatt cctatggtgg gaattggctg caccttttgc ttcgaataaaa 2694

aatgcctgct cggttctcctg tcaaaaaaaaa aaaaaaaaaa aaaaaaaaaa ctcaaggggg 2754

ggcccgtac c 2765

<210> 20
 <211> 802
 <212> PRT
 <213> Zea mays

<400> 20
 Met Ala Ala Lys Leu Thr Arg Leu His Ser Leu Arg Glu Arg Leu Gly
 1 5 10 15
 Ala Thr Phe Ser Ser His Pro Asn Glu Leu Ile Ala Leu Phe Ser Arg
 20 25 30

Tyr Val His Gln Gly Lys Gly Met Leu Gln Arg His Gln Leu Leu Ala
 35 40 45

Glu Phe Asp Ala Leu Phe Asp Ser Asp Lys Glu Lys Tyr Ala Pro Phe
 50 55 60

Glu Asp Ile Leu Arg Ala Ala Gln Glu Ala Ile Val Leu Pro Pro Trp
 65 70 75 80

Val Ala Leu Ala Ile Arg Pro Arg Pro Gly Val Trp Asp Tyr Ile Arg
 85 90 95

Val Asn Val Ser Glu Leu Ala Val Glu Glu Leu Ser Val Ser Glu Tyr
 100 105 110

Leu Ala Phe Lys Glu Gln Leu Val Asp Gly Gln Ser Asn Ser Asn Phe
 115 120 125

Val Leu Glu Leu Asp Phe Glu Pro Phe Asn Ala Ser Phe Pro Arg Pro
 130 135 140

Ser Met Ser Lys Ser Ile Gly Asn Gly Val Gln Phe Leu Asn Arg His
 145 150 155 160

Leu Ser Ser Lys Leu Phe Gln Asp Lys Glu Ser Leu Tyr Pro Leu Leu
 165 170 175

Asn Phe Leu Lys Ala His Asn Tyr Lys Gly Thr Thr Met Met Leu Asn
 180 185 190

Asp Arg Ile Gln Ser Leu Arg Gly Leu Gln Ser Ser Leu Arg Lys Ala
 195 200 205

Glu Glu Tyr Leu Leu Ser Val Pro Gln Asp Thr Pro Tyr Ser Glu Phe
 210 215 220

Asn His Arg Phe Gln Glu Leu Gly Leu Glu Lys Gly Trp Gly Asp Thr
 225 230 235 240

Ala Lys Arg Val Leu Asp Thr Leu His Leu Leu Leu Asp Leu Leu Glu
 245 250 255

Ala Pro Asp Pro Ala Asn Leu Glu Lys Phe Leu Gly Thr Ile Pro Met
 260 265 270

Met Phe Asn Val Val Ile Leu Ser Pro His Gly Tyr Phe Ala Gln Ser

275	280	285
Asn Val Leu Gly Tyr Pro Asp Thr Gly Gly Gln Val Val Tyr Ile Leu		
290	295	300
Asp Gln Val Arg Ala Leu Glu Asn Glu Met Leu Leu Arg Ile Lys Gln		
305	310	315
Gln Gly Leu Asp Ile Thr Pro Lys Ile Leu Ile Val Thr Arg Leu Leu		
325	330	335
Pro Asp Ala Ala Gly Thr Thr Cys Gly Gln Arg Leu Glu Lys Val Ile		
340	345	350
Gly Thr Glu His Thr Asp Ile Ile Arg Val Pro Phe Arg Asn Glu Asn		
355	360	365
Gly Ile Leu Arg Lys Trp Ile Ser Arg Phe Asp Val Trp Pro Tyr Leu		
370	375	380
Glu Thr Tyr Thr Glu Asp Val Ser Ser Glu Ile Met Lys Glu Met Gln		
385	390	395
Ala Lys Pro Asp Leu Ile Ile Gly Asn Tyr Ser Asp Gly Asn Leu Val		
405	410	415
Ala Thr Leu Leu Ala His Lys Leu Gly Val Thr Gln Cys Thr Ile Ala		
420	425	430
His Ala Leu Glu Lys Thr Lys Tyr Pro Asn Ser Asp Ile Tyr Leu Asp		
435	440	445
Lys Phe Asp Ser Gln Tyr His Phe Ser Cys Gln Phe Thr Ala Asp Leu		
450	455	460
Ile Ala Met Asn His Thr Asp Phe Ile Ile Thr Ser Thr Phe Gln Glu		
465	470	475
Ile Ala Gly Ser Lys Asp Thr Val Gly Gln Tyr Glu Ser His Ile Ala		
485	490	495
Phe Thr Leu Pro Gly Leu Tyr Arg Val Val His Gly Ile Asp Val Phe		
500	505	510
Asp Pro Lys Phe Asn Ile Val Ser Pro Gly Ala Asp Met Ser Val Tyr		

515	520	525
Tyr Pro Tyr Thr Glu Thr Asp Lys Arg Leu Thr Ala Phe His Pro Glu 530 535 540		
Ile Glu Glu Leu Ile Tyr Ser Asp Val Glu Asn Ser Glu His Lys Phe 545 550 555 560		
Val Leu Lys Asp Lys Lys Lys Pro Ile Ile Phe Ser Met Ala Arg Leu 565 570 575		
Asp Arg Val Lys Asn Met Thr Gly Leu Val Glu Met Tyr Gly Lys Asn 580 585 590		
Ala Arg Leu Arg Glu Leu Ala Asn Leu Val Ile Val Ala Gly Asp His 595 600 605		
Gly Lys Glu Ser Lys Asp Arg Glu Glu Gln Ala Glu Phe Lys Lys Met 610 615 620		
Tyr Ser Leu Ile Asp Glu Tyr Lys Leu Lys Gly His Ile Arg Trp Ile 625 630 635 640		
Ser Ala Gln Met Asn Arg Val Arg Asn Gly Glu Leu Tyr Arg Tyr Ile 645 650 655		
Cys Asp Thr Lys Gly Ala Phe Val Gln Pro Ala Phe Tyr Glu Ala Phe 660 665 670		
Gly Leu Thr Val Ile Glu Ser Met Thr Cys Gly Leu Pro Thr Ile Ala 675 680 685		
Thr Cys His Gly Gly Pro Ala Glu Ile Ile Val Asp Gly Val Ser Gly 690 695 700		
Leu His Ile Asp Pro Tyr His Ser Asp Lys Ala Ala Asp Ile Leu Val 705 710 715 720		
Asn Phe Phe Asp Lys Cys Lys Ala Asp Pro Ser Tyr Trp Asp Lys Ile 725 730 735		
Ser Gln Gly Gly Leu Gln Arg Ile Tyr Glu Lys Tyr Thr Trp Lys Leu 740 745 750		
Tyr Ser Glu Arg Leu Met Thr Leu Thr Gly Val Tyr Gly Phe Trp Lys 755 760 765		

Tyr Val Ser Asn Leu Glu Arg Arg Glu Thr Arg Arg Tyr Ile Glu Met
 770 775 780

Phe Tyr Ala Leu Lys Tyr Arg Ser Leu Ala Ser Gln Val Pro Leu Ser
 785 790 795 800

Phe Asp

<210> 21
 <211> 2406
 <212> DNA
 <213> Zea mays

<220>
 <221> CDS
 <222> (1)..(2406)

<400> 21
 atg gct gcc aag ctg act cgc ctc cac agt ctt cgc gaa cgc ctt ggt 48
 Met Ala Ala Lys Leu Thr Arg Leu His Ser Leu Arg Glu Arg Leu Gly
 1 5 10 15
 gcc acc ttc tcc tcc cat ccc aat gaa ctg ata gca ctc ttt tcc agg 96
 Ala Thr Phe Ser Ser His Pro Asn Glu Leu Ile Ala Leu Phe Ser Arg
 20 25 30
 tat gtt cac cag ggc aag gga atg ctt cag cgc cat cag ctg ctt gcg 144
 Tyr Val His Gln Gly Lys Gly Met Leu Gln Arg His Gln Leu Leu Ala
 35 40 45
 gag ttt gat gcc ctg ttt gat agt gac aag gag aag tat gca ccc ttt 192
 Glu Phe Asp Ala Leu Phe Asp Ser Asp Lys Glu Lys Tyr Ala Pro Phe
 50 55 60
 gaa gac att ctt cgt gct gct cag gaa gca att gtg ctc ccc cca tgg 240
 Glu Asp Ile Leu Arg Ala Ala Gln Glu Ala Ile Val Leu Pro Pro Trp
 65 70 75 80
 gtt gca ctt gct atc agg cca agg cct ggt gtc tgg gat tac att cgg 288
 Val Ala Leu Ala Ile Arg Pro Arg Pro Gly Val Trp Asp Tyr Ile Arg
 85 90 95
 gtg aat gta agt gag ctg gct gtg gag gag ctg agt gtt tct gag tac 336
 Val Asn Val Ser Glu Leu Ala Val Glu Glu Leu Ser Val Ser Glu Tyr
 100 105 110
 ttg gca ttc aag gaa cag ctg gtg gat gga caa tcc aac agc aac ttt 384
 Leu Ala Phe Lys Glu Gln Leu Val Asp Gly Gln Ser Asn Ser Asn Phe
 115 120 125
 gtg ctt gag ctt gat ttt gag ccc ttc aat gcc tcc ttt cct cgt cct 432
 Val Leu Glu Leu Asp Phe Glu Pro Phe Asn Ala Ser Phe Pro Arg Pro
 130 135 140

tcc atg tcg aag tcc atc gga aat gga gtg caa ttc ctt aac cga cac Ser Met Ser Lys Ser Ile Gly Asn Gly Val Gln Phe Leu Asn Arg His 145 150 155 160	480
ctg tcg tcc aag ttg ttc cag gac aag gag agt ttg tac ccc ttg ctg Leu Ser Ser Lys Leu Phe Gln Asp Lys Glu Ser Leu Tyr Pro Leu Leu 165 170 175	528
aac ttc ctc aag gct cat aac tac aag ggc acg acg atg atg ttg aat Asn Phe Leu Lys Ala His Asn Tyr Lys Gly Thr Thr Met Met Leu Asn 180 185 190	576
gac aga atc caa agc ctt cgt ggt ctc caa tca tcc ctg aga aag gca Asp Arg Ile Gln Ser Leu Arg Gly Leu Gln Ser Ser Leu Arg Lys Ala 195 200 205	624
gag gag tat cta ctg agt gtt cct caa gac act ccc tac tcg gag ttc Glu Glu Tyr Leu Leu Ser Val Pro Gln Asp Thr Pro Tyr Ser Glu Phe 210 215 220	672
aac cat agg ttc caa gag ctt ggt ttg gag aag ggt tgg ggt gac act Asn His Arg Phe Gln Glu Leu Gly Leu Glu Lys Gly Trp Gly Asp Thr 225 230 235 240	720
gcg aag cgt gta ctc gac aca ctc cac ttg ctt ctt gac ctt ctt gag Ala Lys Arg Val Leu Asp Thr Leu His Leu Leu Leu Asp Leu Leu Glu 245 250 255	768
gcc cct gat cct gcc aac ttg gag aag ttc ctt gga act ata cca atg Ala Pro Asp Pro Ala Asn Leu Glu Lys Phe Leu Gly Thr Ile Pro Met 260 265 270	816
atg ttc aat gtt gtt atc ctt tct cct cat ggc tac ttc gct cag tcc Met Phe Asn Val Val Ile Leu Ser Pro His Gly Tyr Phe Ala Gln Ser 275 280 285	864
aat gtg ctt gga tac cct gac act ggc ggt cag gtt gtg tac att ctg Asn Val Leu Gly Tyr Pro Asp Thr Gly Gly Gln Val Val Tyr Ile Leu 290 295 300	912
gat caa gtc cgt gct ttg gag aat gag atg ctt ctg agg att aag cag Asp Gln Val Arg Ala Leu Glu Asn Glu Met Leu Leu Arg Ile Lys Gln 305 310 315 320	960
caa ggc ctt gat atc act ccg aag atc ctc att gtt acc agg ctg ttg Gln Gly Leu Asp Ile Thr Pro Lys Ile Leu Ile Val Thr Arg Leu Leu 325 330 335	1008
cct gat gct gct ggg act acg tgc ggt cag cgg ctg gag aag gtc att Pro Asp Ala Ala Gly Thr Thr Cys Gly Gln Arg Leu Glu Lys Val Ile 340 345 350	1056
ggt act gag cac aca gac atc att cgc gtt ccg ttc aga aat gag aat Gly Thr Glu His Thr Asp Ile Ile Arg Val Pro Phe Arg Asn Glu Asn 355 360 365	1104
ggc atc ctc cgc aag tgg atc tct cgt ttt gat gtc tgg cca tac ctg Gly Ile Leu Arg Lys Trp Ile Ser Arg Phe Asp Val Trp Pro Tyr Leu 370 375 380	1152
gag aca tac act gag gat gtt tcc agt gaa ata atg aaa gaa atg cag	1200

Glu Thr Tyr Thr Glu Asp Val Ser Ser Glu Ile Met Lys Glu Met Gln	
385 390 395 400	
gcc aag cct gac ctt atc att ggc aac tac agc gat ggc aac cta gtc	1248
Ala Lys Pro Asp Leu Ile Ile Gly Asn Tyr Ser Asp Gly Asn Leu Val	
405 410 415	
gcc act ctg ctc gcg cac aag ttg gga gtc act cag tgt acc atc gct	1296
Ala Thr Leu Leu Ala His Lys Leu Gly Val Thr Gln Cys Thr Ile Ala	
420 425 430	
cat gcc ttg gag aaa acc aaa tac ccc aac tcg gac atc tac ttg gac	1344
His Ala Leu Glu Lys Thr Lys Tyr Pro Asn Ser Asp Ile Tyr Leu Asp	
435 440 445	
aag ttc gac agc cag tac cac ttc tct tgc cag ttc aca gct gac ctt	1392
Lys Phe Asp Ser Gln Tyr His Phe Ser Cys Gln Phe Thr Ala Asp Leu	
450 455 460	
att gcc atg aac cac acc gat ttc atc atc acc agc aca ttc caa gaa	1440
Ile Ala Met Asn His Thr Asp Phe Ile Ile Thr Ser Thr Phe Gln Glu	
465 470 475 480	
atc gcg gga agc aag gac acc gtg ggg cag tac gag tcc cac atc gcg	1488
Ile Ala Gly Ser Lys Asp Thr Val Gly Gln Tyr Glu Ser His Ile Ala	
485 490 495	
ttc act ctt cct ggg ctc tac cgt gtc gtc cat ggc atc gat gtt ttc	1536
Phe Thr Leu Pro Gly Leu Tyr Arg Val Val His Gly Ile Asp Val Phe	
500 505 510	
gat ccc aag ttc aac att gtc tcc cct gga gca gac atg agt gtt tac	1584
Asp Pro Lys Phe Asn Ile Val Ser Pro Gly Ala Asp Met Ser Val Tyr	
515 520 525	
tac ccg tat acg gaa acc gac aag aga ctc act gcc ttc cat cct gaa	1632
Tyr Pro Tyr Thr Glu Thr Asp Lys Arg Leu Thr Ala Phe His Pro Glu	
530 535 540	
atc gag gag ctc atc tac agc gac gtc gag aac tcc gag cac aag ttc	1680
Ile Glu Glu Leu Ile Tyr Ser Asp Val Glu Asn Ser Glu His Lys Phe	
545 550 555 560	
gtg ctg aag gac aag aag aag ccg atc atc ttc tcg atg gcg cgt ctc	1728
Val Leu Lys Asp Lys Lys Lys Pro Ile Ile Phe Ser Met Ala Arg Leu	
565 570 575	
gac cgc gtg aag aac atg aca ggc ctg gtc gag atg tac ggc aag aac	1776
Asp Arg Val Lys Asn Met Thr Gly Leu Val Glu Met Tyr Gly Lys Asn	
580 585 590	
gcg cgc ctg agg gag ctg gcg aac ctc gtg atc gtt gcc ggc gac cac	1824
Ala Arg Leu Arg Glu Leu Ala Asn Leu Val Ile Val Ala Gly Asp His	
595 600 605	
ggc aag gag tcc aag gac agg gag gag cag gcg gag ttc aag aag atg	1872
Gly Lys Glu Ser Lys Asp Arg Glu Glu Gln Ala Glu Phe Lys Lys Met	
610 615 620	
tac agc ctc atc gac gag tac aag ttg aag ggc cat atc cgg tgg atc	1920
Tyr Ser Leu Ile Asp Glu Tyr Lys Leu Lys Gly His Ile Arg Trp Ile	

625	630	635	640	
tcg gcg cag atg aac cgc gtc cgc aac ggg gag ctg tac cgc tac att				1968
Ser Ala Gln Met Asn Arg Val Arg Asn Gly Glu Leu Tyr Arg Tyr Ile	645	650	655	
tgc gat acg aag ggc gca ttc gtg cag cct gcg ttc tac gaa gcg ttc				2016
Cys Asp Thr Lys Gly Ala Phe Val Gln Pro Ala Phe Tyr Glu Ala Phe	660	665	670	
ggc ctg act gtg atc gag tcc atg acg tgc ggt ctg cca acg atc gcg				2064
Gly Leu Thr Val Ile Glu Ser Met Thr Cys Gly Leu Pro Thr Ile Ala	675	680	685	
acc tgc cat ggt ggc cct gct gag atc atc gtg gac ggg gta tct ggc				2112
Thr Cys His Gly Gly Pro Ala Glu Ile Ile Val Asp Gly Val Ser Gly	690	695	700	
ctg cac att gac cct tac cac agc gac aag gcc gcg gat atc ctg gtc				2160
Leu His Ile Asp Pro Tyr His Ser Asp Lys Ala Ala Asp Ile Leu Val	705	710	715	720
aac ttc ttt gac aaa tgc aag gca gat ccg agc tac tgg gac aag atc				2208
Asn Phe Phe Asp Lys Cys Lys Ala Asp Pro Ser Tyr Trp Asp Lys Ile	725	730	735	
tca cag ggc ggc ctg cag aga att tat gag aag tac acc tgg aag ctc				2256
Ser Gln Gly Gly Leu Gln Arg Ile Tyr Glu Lys Tyr Thr Trp Lys Leu	740	745	750	
tac tcc gag agg ctg atg acc ctg acc ggc gtg tac ggg ttc tgg aag				2304
Tyr Ser Glu Arg Leu Met Thr Leu Thr Gly Val Tyr Gly Phe Trp Lys	755	760	765	
tac gtg agc aac ctg gag agg cgc gag acc cgc cgc tac atc gag atg				2352
Tyr Val Ser Asn Leu Glu Arg Arg Glu Thr Arg Arg Tyr Ile Glu Met	770	775	780	
ttc tac gcc ctg aag tac cgt agc ctg gca agc cag gtt ccg ctg tcc				2400
Phe Tyr Ala Leu Lys Tyr Arg Ser Leu Ala Ser Gln Val Pro Leu Ser	785	790	795	800
ttc gat				2406
Phe Asp				
<210> 22				
<211> 802				
<212> PRT				
<213> Zea mays				
<400> 22				
Met Ala Ala Lys Leu Thr Arg Leu His Ser Leu Arg Glu Arg Leu Gly	1	5	10	15
Ala Thr Phe Ser Ser His Pro Asn Glu Leu Ile Ala Leu Phe Ser Arg	20	25	30	

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

Met Phe Asn Val Val Ile Leu Ser Pro His Gly Tyr Phe Ala Gln Ser
275 280 285

Asn Val Leu Gly Tyr Pro Asp Thr Gly Gly Gln Val Val Tyr Ile Leu
290 295 300

Asp Gln Val Arg Ala Leu Glu Asn Glu Met Leu Leu Arg Ile Lys Gln
305 310 315 320

Gln Gly Leu Asp Ile Thr Pro Lys Ile Leu Ile Val Thr Arg Leu Leu
325 330 335

Pro Asp Ala Ala Gly Thr Thr Cys Gly Gln Arg Leu Glu Lys Val Ile
340 345 350

Gly Thr Glu His Thr Asp Ile Ile Arg Val Pro Phe Arg Asn Glu Asn
355 360 365

Gly Ile Leu Arg Lys Trp Ile Ser Arg Phe Asp Val Trp Pro Tyr Leu
370 375 380

Glu Thr Tyr Thr Glu Asp Val Ser Ser Glu Ile Met Lys Glu Met Gln
385 390 395 400

Ala Lys Pro Asp Leu Ile Ile Gly Asn Tyr Ser Asp Gly Asn Leu Val
405 410 415

Ala Thr Leu Leu Ala His Lys Leu Gly Val Thr Gln Cys Thr Ile Ala
420 425 430

His Ala Leu Glu Lys Thr Lys Tyr Pro Asn Ser Asp Ile Tyr Leu Asp
435 440 445

Lys Phe Asp Ser Gln Tyr His Phe Ser Cys Gln Phe Thr Ala Asp Leu
450 455 460

Ile Ala Met Asn His Thr Asp Phe Ile Ile Thr Ser Thr Phe Gln Glu
465 470 475 480

Ile Ala Gly Ser Lys Asp Thr Val Gly Gln Tyr Glu Ser His Ile Ala
485 490 495

Phe Thr Leu Pro Gly Leu Tyr Arg Val Val His Gly Ile Asp Val Phe
500 505 510

Asp Pro Lys Phe Asn Ile Val Ser Pro Gly Ala Asp Met Ser Val Tyr

515	520	525
Tyr Pro Tyr Thr Glu Thr Asp Lys Arg Leu Thr Ala Phe His Pro Glu		
530	535	540
Ile Glu Glu Leu Ile Tyr Ser Asp Val Glu Asn Ser Glu His Lys Phe		
545	550	555
		560
Val Leu Lys Asp Lys Lys Lys Pro Ile Ile Phe Ser Met Ala Arg Leu		
	565	570
		575
Asp Arg Val Lys Asn Met Thr Gly Leu Val Glu Met Tyr Gly Lys Asn		
	580	585
		590
Ala Arg Leu Arg Glu Leu Ala Asn Leu Val Ile Val Ala Gly Asp His		
	595	600
		605
Gly Lys Glu Ser Lys Asp Arg Glu Glu Gln Ala Glu Phe Lys Lys Met		
610	615	620
Tyr Ser Leu Ile Asp Glu Tyr Lys Leu Lys Gly His Ile Arg Trp Ile		
625	630	635
		640
Ser Ala Gln Met Asn Arg Val Arg Asn Gly Glu Leu Tyr Arg Tyr Ile		
	645	650
		655
Cys Asp Thr Lys Gly Ala Phe Val Gln Pro Ala Phe Tyr Glu Ala Phe		
	660	665
		670
Gly Leu Thr Val Ile Glu Ser Met Thr Cys Gly Leu Pro Thr Ile Ala		
	675	680
		685
Thr Cys His Gly Gly Pro Ala Glu Ile Ile Val Asp Gly Val Ser Gly		
	690	695
		700
Leu His Ile Asp Pro Tyr His Ser Asp Lys Ala Ala Asp Ile Leu Val		
705	710	715
		720
Asn Phe Phe Asp Lys Cys Lys Ala Asp Pro Ser Tyr Trp Asp Lys Ile		
	725	730
		735
Ser Gln Gly Gly Leu Gln Arg Ile Tyr Glu Lys Tyr Thr Trp Lys Leu		
	740	745
		750
Tyr Ser Glu Arg Leu Met Thr Leu Thr Gly Val Tyr Gly Phe Trp Lys		

755 760 765

Tyr Val Ser Asn Leu Glu Arg Arg Glu Thr Arg Arg Tyr Ile Glu Met
770 775 780

Phe Tyr Ala Leu Lys Tyr Arg Ser Leu Ala Ser Gln Val Pro Leu Ser
785 790 795 800

Phe Asp

<210>	23
<211>	3481
<212>	DNA
<213>	Zea mays

```
<220>
<221> CDS
<222> (156)..(2606)
```

```
<220>
<221> misc_feature
<222> (3169)..(3255)
<223> n located at 3169, 3255, each n is a or c or g or t/u, unknown, or
other
```

```

<400> 23
actccacctc cgttcacccc gtccatttga tttgcgttca ctgcggttgcg tttccttggga      60
ggggattggt ctctcctctc ctttggattg gaggtccctc cttcttctctc tctctctctctc      120
agaggaaggc ctgaggatcc aggaagagga cagca atg ggg gaa ggt gca ggt      173
          Met Gly Glu Gly Ala Gly
                1                5

```

gac cgt gtc ctg agc cgc ctc cac agc gtc agg gag cgc att ggc gac 221
Asp Arg Val Leu Ser Arg Leu His Ser Val Arg Glu Arg Ile Gly Asp
10 15 20

tca ctc tct gcc cac ccc aat gag ctt gtc gcc gtc ttc acc agg ctg 269
Ser Leu Ser Ala His Pro Asn Glu Leu Val Ala Val Phe Thr Arg Leu
25 30 35

aaa aac ctt gga aag ggt atg ctg cag ccc cac cag atc att gcc gag 317
Lys Asn Leu Gly Lys Gly Met Leu Gln Pro His Gln Ile Ile Ala Glu
40 45 50

tac aac aat gcg atc cct gag gct gag cgc gag aag ctc aag gat ggt. 365
Tyr Asn Asn Ala Ile Pro Glu Ala Glu Arg Glu Lys Leu Lys Asp Gly
55 60 65 70

gct ttt gag gat gtc ctg agg gca gct cag gag gcg att gtc atc ccc 413
Ala Phe Glu Asp Val Leu Arg Ala Ala Gln Glu Ala Ile Val Ile Pro
75 80 85

cca tgg gtt gca ctt gcc atc cgc cct agg cct ggt gtc tgg gag tat 461
Pro Trp Val Ala Leu Ala Ile Arg Pro Arg Pro Gly Val Trp Glu Tyr

90						95						100						
gtg	agg	gtc	aac	gtc	agt	gag	ctc	gct	gtt	gag	gag	ctg	aga	gtt	cct	509		
Val	Arg	Val	Asn	Val	Ser	Glu	Leu	Ala	Val	Glu	Glu	Leu	Arg	Val	Pro			
		105					110					115						
gag	tac	ctg	cag	ttc	aag	gaa	cag	ctt	gtg	gaa	gaa	ggc	ccc	aac	aac	557		
Glu	Tyr	Leu	Gln	Phe	Lys	Glu	Gln	Leu	Val	Glu	Glu	Gly	Pro	Asn	Asn			
	120					125						130						
aac	ttt	gtt	ctt	gag	ctg	gac	ttt	gag	cca	ttc	aat	gcc	tcc	ttc	ccc	605		
Asn	Phe	Val	Leu	Glu	Leu	Asp	Phe	Glu	Pro	Phe	Asn	Ala	Ser	Phe	Pro			
135					140					145					150			
cgt	cct	tct	ctg	tca	aag	tcc	att	ggc	aat	ggc	gtg	cag	ttc	ctc	aac	653		
Arg	Pro	Ser	Leu	Ser	Lys	Ser	Ile	Gly	Asn	Gly	Val	Gln	Phe	Leu	Asn			
				155				160						165				
agg	cac	ctg	tca	tca	aag	ctc	ttc	cat	gac	aag	gag	agc	atg	tac	ccc	701		
Arg	His	Leu	Ser	Ser	Lys	Leu	Phe	His	Asp	Lys	Glu	Ser	Met	Tyr	Pro			
		170					175						180					
ttg	ctc	aac	ttc	ctt	cgc	gcc	cac	aac	tac	aag	ggg	atg	acc	atg	atg	749		
Leu	Leu	Asn	Phe	Leu	Arg	Ala	His	Asn	Tyr	Lys	Gly	Met	Thr	Met	Met			
		185					190					195						
ttg	aac	gac	aga	atc	cgc	agt	ctc	agt	gct	ctg	caa	ggc	gcg	ctg	agg	797		
Leu	Asn	Asp	Arg	Ile	Arg	Ser	Leu	Ser	Ala	Leu	Gln	Gly	Ala	Leu	Arg			
	200					205					210							
aag	gct	gag	gag	cac	ctg	tcc	acc	cta	caa	gct	gat	acc	cca	tac	tct	845		
Lys	Ala	Glu	Glu	His	Leu	Ser	Thr	Leu	Gln	Ala	Asp	Thr	Pro	Tyr	Ser			
215					220					225					230			
gaa	ttt	cac	cac	agg	ttc	cag	gaa	ctt	ggc	ctg	gag	aag	ggc	tgg	ggc	893		
Glu	Phe	His	His	Arg	Phe	Gln	Glu	Leu	Gly	Leu	Glu	Lys	Gly	Trp	Gly			
				235					240					245				
gat	tgc	gct	aag	cgt	gca	cag	gag	act	atc	cac	ctc	ctc	ttg	gac	ctc	941		
Asp	Cys	Ala	Lys	Arg	Ala	Gln	Glu	Thr	Ile	His	Leu	Leu	Leu	Asp	Leu			
			250					255					260					
ctg	gag	gcc	cca	gat	ccg	tcc	acc	ctg	gag	aag	ttc	ctt	gga	acg	atc	989		
Leu	Glu	Ala	Pro	Asp	Pro	Ser	Thr	Leu	Glu	Lys	Phe	Leu	Gly	Thr	Ile			
		265					270					275						
ccc	atg	gtg	ttc	aat	gtc	gtt	atc	ctc	tcc	cct	cat	ggc	tac	ttc	gct	1037		
Pro	Met	Val	Phe	Asn	Val	Val	Ile	Leu	Ser	Pro	His	Gly	Tyr	Phe	Ala			
	280						285					290						
caa	gct	aat	gtc	ttg	ggc	tac	cct	gac	acc	gga	ggc	cag	gtt	gtc	tac	1085		
Gln	Ala	Asn	Val	Leu	Gly	Tyr	Pro	Asp	Thr	Gly	Gly	Gln	Val	Val	Tyr			
295					300					305					310			
atc	ttg	gat	caa	gtg	cgc	gct	atg	gag	aac	gaa	atg	ctg	ctg	agg	atc	1133		
Ile	Leu	Asp	Gln	Val	Arg	Ala	Met	Glu	Asn	Glu	Met	Leu	Leu	Arg	Ile			
			315					320						325				
aag	cag	tgt	ggc	ctt	gac	atc	acg	ccg	aag	atc	ctt	att	gtc	acc	agg	1181		
Lys	Gln	Cys	Gly	Leu	Asp	Ile	Thr	Pro	Lys	Ile	Leu	Ile	Val	Thr	Arg			
		330						335					340					

ttg ctc cct gat gca act ggc acc acc tgt ggc cag cgc ctt gag aag Leu Leu Pro Asp Ala Thr Gly Thr Thr Cys Gly Gln Arg Leu Glu Lys 345 350 355	1229
gtc ctt ggc acc gag cac tgc cat atc ctt cgc gtg cca ttc aga aca Val Leu Gly Thr Glu His Cys His Ile Leu Arg Val Pro Phe Arg Thr 360 365 370	1277
gaa aac gga atc gtt cgc aag tgg atc tcg cga ttt gaa gtc tgg ccg Glu Asn Gly Ile Val Arg Lys Trp Ile Ser Arg Phe Glu Val Trp Pro 375 380 385 390	1325
tac ctg gag act tac act gat gac gtg gcg cat gag att gct gga gag Tyr Leu Glu Thr Tyr Thr Asp Asp Val Ala His Glu Ile Ala Gly Glu 395 400 405	1373
ctt cag gcc aat cct gac ctg atc atc gga aac tac agt gac gga aac Leu Gln Ala Asn Pro Asp Leu Ile Ile Gly Asn Tyr Ser Asp Gly Asn 410 415 420	1421
ctt gtt gcg tgt ttg ctc gcc cac aag atg ggt gtt act cac tgt acc Leu Val Ala Cys Leu Leu Ala His Lys Met Gly Val Thr His Cys Thr 425 430 435	1469
att gcc cat gcg ctt gag aaa act aag tac cct aac tcc gac ctc tac Ile Ala His Ala Leu Glu Lys Thr Lys Tyr Pro Asn Ser Asp Leu Tyr 440 445 450	1517
tgg aag aag ttt gag gat cac tac cac ttc tcg tgc cag ttc acc act Trp Lys Lys Phe Glu Asp His Tyr His Phe Ser Cys Gln Phe Thr Thr 455 460 465 470	1565
gac ttg att gca atg aac cat gcc gac ttc atc atc acc agt acc ttc Asp Leu Ile Ala Met Asn His Ala Asp Phe Ile Ile Thr Ser Thr Phe 475 480 485	1613
caa gag atc gcc gga aac aag gac acc gtc ggc cag tac gag tca cac Gln Glu Ile Ala Gly Asn Lys Asp Thr Val Gly Gln Tyr Glu Ser His 490 495 500	1661
atg gcg ttc aca atg cct ggc ctg tac cgc gtt gtc cac ggc att gat Met Ala Phe Thr Met Pro Gly Leu Tyr Arg Val Val His Gly Ile Asp 505 510 515	1709
gtg ttc gac ccc aag ttc aac atc gtg tct cct ggc gcg gac ctg tcc Val Phe Asp Pro Lys Phe Asn Ile Val Ser Pro Gly Ala Asp Leu Ser 520 525 530	1757
atc tac ttc ccg tac acc gag tcg cac aag agg ctg acc tcc ctt cac Ile Tyr Phe Pro Tyr Thr Glu Ser His Lys Arg Leu Thr Ser Leu His 535 540 545 550	1805
ccg gag att gag gag ctc ctg tac agc caa acc gag aac acg gag cac Pro Glu Ile Glu Glu Leu Leu Tyr Ser Gln Thr Glu Asn Thr Glu His 555 560 565	1853
aag ttc gtt ctg aac gac agg aac aag cca atc atc ttc tcc atg gct Lys Phe Val Leu Asn Asp Arg Asn Lys Pro Ile Ile Phe Ser Met Ala 570 575 580	1901

cgt ctc gac cgt gtg aag aac ttg act ggg ctg gtg gag ctg tac ggc Arg Leu Asp Arg Val Lys Asn Leu Thr Gly Leu Val Glu Leu Tyr Gly 585 590 595	1949
cgg aac aag cgg ctg cag gag ctg gtg aac ctc gtg gtc gtc tgc ggc Arg Asn Lys Arg Leu Gln Glu Leu Val Asn Leu Val Val Val Cys Gly 600 605 610	1997
gac cat ggc aac cct tcc aag gac aag gag gag cag gcc gag ttc aag Asp His Gly Asn Pro Ser Lys Asp Lys Glu Glu Gln Ala Glu Phe Lys 615 620 625 630	2045
aag atg ttt gac ctc atc gag cag tac aac ctg aac ggg cac atc cgc Lys Met Phe Asp Leu Ile Glu Gln Tyr Asn Leu Asn Gly His Ile Arg 635 640 645	2093
tgg atc tcc gcc cag atg aac cgc gtc cgc aac ggc gag ctg tac cgc Trp Ile Ser Ala Gln Met Asn Arg Val Arg Asn Gly Glu Leu Tyr Arg 650 655 660	2141
tac atc tgc gac acc aag ggc gcc ttc gtg cag cct gct ttc tac gag Tyr Ile Cys Asp Thr Lys Gly Ala Phe Val Gln Pro Ala Phe Tyr Glu 665 670 675	2189
gct ttc ggg ctg acg gtg gtt gag gcc atg acc tgc ggc ctg ccc acg Ala Phe Gly Leu Thr Val Val Glu Ala Met Thr Cys Gly Leu Pro Thr 680 685 690	2237
ttt gcc aca gcc tac ggc ggt ccg gcc gag atc atc gtg cac ggc gtg Phe Ala Thr Ala Tyr Gly Gly Pro Ala Glu Ile Ile Val His Gly Val 695 700 705 710	2285
tct ggc tac cac atc gac cct tac cag ggc gac aag gcg tcg gcc ctg Ser Gly Tyr His Ile Asp Pro Tyr Gln Gly Asp Lys Ala Ser Ala Leu 715 720 725	2333
ctc gtg gac ttc ttc gac aag tgc cag gcg gag ccg agc cac tgg agc Leu Val Asp Phe Phe Asp Lys Cys Gln Ala Glu Pro Ser His Trp Ser 730 735 740	2381
aag atc tcc cag ggc ggg ctc cag cgt atc gag gag aag tac acc tgg Lys Ile Ser Gln Gly Gly Leu Gln Arg Ile Glu Glu Lys Tyr Thr Trp 745 750 755	2429
aag ctg tac tcg gag agg ctg atg acc ctc acc ggc gtg tac ggg ttc Lys Leu Tyr Ser Glu Arg Leu Met Thr Leu Thr Gly Val Tyr Gly Phe 760 765 770	2477
tgg aag tac gtg tcc aac ctg gag agg cgc gag acc cgg cgg tac ctg Trp Lys Tyr Val Ser Asn Leu Glu Arg Arg Glu Thr Arg Arg Tyr Leu 775 780 785 790	2525
gag atg ctg tac gcg ctc aag tac cgc acc atg gcg agc acc gtg ccg Glu Met Leu Tyr Ala Leu Lys Tyr Arg Thr Met Ala Ser Thr Val Pro 795 800 805	2573
ctg gcc gtg gag gga gag ccc tcc agc aag tga tgcgcgacgg cggccacaga Leu Ala Val Glu Gly Glu Pro Ser Ser Lys 810 815	2626
cctgatcgat cgatgagcga gagggagcac tcggagtgtc gtgtcttttc ccttgccatt	2686

```

tctttctttc ttttttttcc ttcccggagg cgaaaaaaaa gagtctgctt ttgctaggcg 2746
gcgggcggttc gttgctgctc tttgcttcaa gagttaaaatt tacctacctt gtcaaggctc 2806
tgttccatca ttgatccggg tgtcgctttt agtagtctga tggactgtta gtagtttgcg 2866
ttgcgctcggg tgagagggaa cgggtgggtgt ggtgggtgtgt gtgcagtcgg gtgtgggtgct 2926
ccctttgttt cctggatggg atgttgctcc ttgaataata atcgtagtgg ccttgagacc 2986
cttttctga aataagagca gcacccatgt gcttaccttt gcagctgtaa ctgggtgtttg 3046
cctcaggatt tttgctcctc ccttctatgg attgtgctta aaattctcta taaaaaaaaa 3106
aaaaaaaaa ctccaggggg ggcccggggc ccatttcgcc ctatagtgat ttaatacaca 3166
gcncgcgcgg ggggggcccgc cctcccgtc ctttttgggg cgagaattcg gcccgccctt 3226
tcccgcctca ctcccccgac caccgtacnt ccgcgcgcac gcccagacc ggcagagacc 3286
cgcgccctg gtattcccgc gcacaccgcc cccccccgc cccaccgtc cccatagtgt 3346
tattccactc tctgtatcc cgcccttagc cctccccgc cgccccaat ttgaccacc 3406
atctcctccc cccaagatac gccttcacga ccagatttca cccgcggccg ctccccccgc 3466
acacacaaaa aaaaaa 3481

```

<210> 24
 <211> 816
 <212> PRT
 <213> Zea mays

<400> 24
 Met Gly Glu Gly Ala Gly Asp Arg Val Leu Ser Arg Leu His Ser Val
 1 5 10 15
 Arg Glu Arg Ile Gly Asp Ser Leu Ser Ala His Pro Asn Glu Leu Val
 20 25 30
 Ala Val Phe Thr Arg Leu Lys Asn Leu Gly Lys Gly Met Leu Gln Pro
 35 40 45
 His Gln Ile Ile Ala Glu Tyr Asn Asn Ala Ile Pro Glu Ala Glu Arg
 50 55 60
 Glu Lys Leu Lys Asp Gly Ala Phe Glu Asp Val Leu Arg Ala Ala Gln
 65 70 75 80
 Glu Ala Ile Val Ile Pro Pro Trp Val Ala Leu Ala Ile Arg Pro Arg
 85 90 95

Pro Gly Val Trp Glu Tyr Val Arg Val Asn Val Ser Glu Leu Ala Val
 100 105 110

Glu Glu Leu Arg Val Pro Glu Tyr Leu Gln Phe Lys Glu Gln Leu Val
 115 120 125

Glu Glu Gly Pro Asn Asn Asn Phe Val Leu Glu Leu Asp Phe Glu Pro
 130 135 140

Phe Asn Ala Ser Phe Pro Arg Pro Ser Leu Ser Lys Ser Ile Gly Asn
 145 150 155 160

Gly Val Gln Phe Leu Asn Arg His Leu Ser Ser Lys Leu Phe His Asp
 165 170 175

Lys Glu Ser Met Tyr Pro Leu Leu Asn Phe Leu Arg Ala His Asn Tyr
 180 185 190

Lys Gly Met Thr Met Met Leu Asn Asp Arg Ile Arg Ser Leu Ser Ala
 195 200 205

Leu Gln Gly Ala Leu Arg Lys Ala Glu Glu His Leu Ser Thr Leu Gln
 210 215 220

Ala Asp Thr Pro Tyr Ser Glu Phe His His Arg Phe Gln Glu Leu Gly
 225 230 235 240

Leu Glu Lys Gly Trp Gly Asp Cys Ala Lys Arg Ala Gln Glu Thr Ile
 245 250 255

His Leu Leu Leu Asp Leu Leu Glu Ala Pro Asp Pro Ser Thr Leu Glu
 260 265 270

Lys Phe Leu Gly Thr Ile Pro Met Val Phe Asn Val Val Ile Leu Ser
 275 280 285

Pro His Gly Tyr Phe Ala Gln Ala Asn Val Leu Gly Tyr Pro Asp Thr
 290 295 300

Gly Gly Gln Val Val Tyr Ile Leu Asp Gln Val Arg Ala Met Glu Asn
 305 310 315 320

Glu Met Leu Leu Arg Ile Lys Gln Cys Gly Leu Asp Ile Thr Pro Lys
 325 330 335

Ile Leu Ile Val Thr Arg Leu Leu Pro Asp Ala Thr Gly Thr Thr Cys

340

Gly Gln Arg Leu Glu Lys Val Leu Gly Thr Glu His Cys His Ile Leu
355 360 365

Arg Val Pro Phe Arg Thr Glu Asn Gly Ile Val Arg Lys Trp Ile Ser
370 375 380

Arg Phe Glu Val Trp Pro Tyr Leu Glu Thr Tyr Thr Asp Asp Val Ala
385 390 395 400

His Glu Ile Ala Gly Glu Leu Gln Ala Asn Pro Asp Leu Ile Ile Gly
405 410 415

Asn Tyr Ser Asp Gly Asn Leu Val Ala Cys Leu Leu Ala His Lys Met
420 425 430

Gly Val Thr His Cys Thr Ile Ala His Ala Leu Glu Lys Thr Lys Tyr
435 440 445

Pro Asn Ser Asp Leu Tyr Trp Lys Lys Phe Glu Asp His Tyr His Phe
450 455 460

Ser Cys Gln Phe Thr Thr Asp Leu Ile Ala Met Asn His Ala Asp Phe
465 470 475 480

Ile Ile Thr Ser Thr Phe Gln Glu Ile Ala Gly Asn Lys Asp Thr Val
485 490 495

Gly Gln Tyr Glu Ser His Met Ala Phe Thr Met Pro Gly Leu Tyr Arg
500 505 510

Val Val His Gly Ile Asp Val Phe Asp Pro Lys Phe Asn Ile Val Ser
515 520 525

Pro Gly Ala Asp Leu Ser Ile Tyr Phe Pro Tyr Thr Glu Ser His Lys
530 535 540

Arg Leu Thr Ser Leu His Pro Glu Ile Glu Glu Leu Leu Tyr Ser Gln
545 550 555 560

Thr Glu Asn Thr Glu His Lys Phe Val Leu Asn Asp Arg Asn Lys Pro
565 570 575

Ile Ile Phe Ser Met Ala Arg Leu Asp Arg Val Lys Asn Leu Thr Gly
580 585 590

Leu Val Glu Leu Tyr Gly Arg Asn Lys Arg Leu Gln Glu Leu Val Asn
 595 600 605

Leu Val Val Val Cys Gly Asp His Gly Asn Pro Ser Lys Asp Lys Glu
 610 615 620

Glu Gln Ala Glu Phe Lys Lys Met Phe Asp Leu Ile Glu Gln Tyr Asn
 625 630 635 640

Leu Asn Gly His Ile Arg Trp Ile Ser Ala Gln Met Asn Arg Val Arg
 645 650 655

Asn Gly Glu Leu Tyr Arg Tyr Ile Cys Asp Thr Lys Gly Ala Phe Val
 660 665 670

Gln Pro Ala Phe Tyr Glu Ala Phe Gly Leu Thr Val Val Glu Ala Met
 675 680 685

Thr Cys Gly Leu Pro Thr Phe Ala Thr Ala Tyr Gly Gly Pro Ala Glu
 690 695 700

Ile Ile Val His Gly Val Ser Gly Tyr His Ile Asp Pro Tyr Gln Gly
 705 710 715 720

Asp Lys Ala Ser Ala Leu Leu Val Asp Phe Phe Asp Lys Cys Gln Ala
 725 730 735

Glu Pro Ser His Trp Ser Lys Ile Ser Gln Gly Gly Leu Gln Arg Ile
 740 745 750

Glu Glu Lys Tyr Thr Trp Lys Leu Tyr Ser Glu Arg Leu Met Thr Leu
 755 760 765

Thr Gly Val Tyr Gly Phe Trp Lys Tyr Val Ser Asn Leu Glu Arg Arg
 770 775 780

Glu Thr Arg Arg Tyr Leu Glu Met Leu Tyr Ala Leu Lys Tyr Arg Thr
 785 790 795 800

Met Ala Ser Thr Val Pro Leu Ala Val Glu Gly Glu Pro Ser Ser Lys
 805 810 815

<210> 25
 <211> 2448

<212> DNA

<213> Zea mays

<220>

<221> CDS

<222> (1)..(2448)

<400> 25

atg ggg gaa ggt gca ggt gac cgt gtc ctg agc cgc ctc cac agc gtc	48
Met Gly Glu Gly Ala Gly Asp Arg Val Leu Ser Arg Leu His Ser Val	
1 5 10 15	
agg gag cgc att ggc gac tca ctc tct gcc cac ccc aat gag ctt gtc	96
Arg Glu Arg Ile Gly Asp Ser Leu Ser Ala His Pro Asn Glu Leu Val	
20 25 30	
gcc gtc ttc acc agg ctg aaa aac ctt gga aag ggt atg ctg cag ccc	144
Ala Val Phe Thr Arg Leu Lys Asn Leu Gly Lys Gly Met Leu Gln Pro	
35 40 45	
cac cag atc att gcc gag tac aac aat gcg atc cct gag gct gag cgc	192
His Gln Ile Ile Ala Glu Tyr Asn Asn Ala Ile Pro Glu Ala Glu Arg	
50 55 60	
gag aag ctc aag gat ggt gct ttt gag gat gtc ctg agg gca gct cag	240
Glu Lys Leu Lys Asp Gly Ala Phe Glu Asp Val Leu Arg Ala Ala Gln	
65 70 75 80	
gag gcg att gtc atc ccc cca tgg gtt gca ctt gcc atc cgc cct agg	288
Glu Ala Ile Val Ile Pro Pro Trp Val Ala Leu Ala Ile Arg Pro Arg	
85 90 95	
cct ggt gtc tgg gag tat gtg agg gtc aac gtc agt gag ctc gct gtt	336
Pro Gly Val Trp Glu Tyr Val Arg Val Asn Val Ser Glu Leu Ala Val	
100 105 110	
gag gag ctg aga gtt cct gag tac ctg cag ttc aag gaa cag ctt gtg	384
Glu Glu Leu Arg Val Pro Glu Tyr Leu Gln Phe Lys Glu Gln Leu Val	
115 120 125	
gaa gaa ggc ccc aac aac aac ttt gtt ctt gag ctg gac ttt gag cca	432
Glu Glu Gly Pro Asn Asn Asn Phe Val Leu Glu Leu Asp Phe Glu Pro	
130 135 140	
ttc aat gcc tcc ttc ccc cgt cct tct ctg tca aag tcc att ggc aat	480
Phe Asn Ala Ser Phe Pro Arg Pro Ser Leu Ser Lys Ser Ile Gly Asn	
145 150 155 160	
ggc gtg cag ttc ctc aac agg cac ctg tca tca aag ctc ttc cat gac	528
Gly Val Gln Phe Leu Asn Arg His Leu Ser Ser Lys Leu Phe His Asp	
165 170 175	
aag gag agc atg tac ccc ttg ctc aac ttc ctt cgc gcc cac aac tac	576
Lys Glu Ser Met Tyr Pro Leu Leu Asn Phe Leu Arg Ala His Asn Tyr	
180 185 190	
aag ggg atg acc atg atg ttg aac gac aga atc cgc agt ctc agt gct	624
Lys Gly Met Thr Met Met Leu Asn Asp Arg Ile Arg Ser Leu Ser Ala	
195 200 205	
ctg caa ggt gcg ctg agg aag gct gag gag cac ctg tcc acc cta caa	672

Leu Gln Gly Ala Leu Arg Lys Ala Glu Glu His Leu Ser Thr Leu Gln	
210 215 220	
gct gat acc cca tac tct gaa ttt cac cac agg ttc cag gaa ctt ggt	720
Ala Asp Thr Pro Tyr Ser Glu Phe His His Arg Phe Gln Glu Leu Gly	
225 230 235 240	
ctg gag aag ggt tgg ggt gat tgc gct aag cgt gca cag gag act atc	768
Leu Glu Lys Gly Trp Gly Asp Cys Ala Lys Arg Ala Gln Glu Thr Ile	
245 250 255	
cac ctc ctc ttg gac ctc ctg gag gcc cca gat ccg tcc acc ctg gag	816
His Leu Leu Leu Asp Leu Leu Glu Ala Pro Asp Pro Ser Thr Leu Glu	
260 265 270	
aag ttc ctt gga acg atc ccc atg gtg ttc aat gtc gtt atc ctc tcc	864
Lys Phe Leu Gly Thr Ile Pro Met Val Phe Asn Val Val Ile Leu Ser	
275 280 285	
cct cat ggt tac ttc gct caa gct aat gtc ttg ggt tac cct gac acc	912
Pro His Gly Tyr Phe Ala Gln Ala Asn Val Leu Gly Tyr Pro Asp Thr	
290 295 300	
gga ggc cag gtt gtc tac atc ttg gat caa gtg cgc gct atg gag aac	960
Gly Gly Gln Val Val Tyr Ile Leu Asp Gln Val Arg Ala Met Glu Asn	
305 310 315 320	
gaa atg ctg ctg agg atc aag cag tgt ggt ctt gac atc acg ccg aag	1008
Glu Met Leu Leu Arg Ile Lys Gln Cys Gly Leu Asp Ile Thr Pro Lys	
325 330 335	
atc ctt att gtc acc agg ttg ctc cct gat gca act ggc acc acc tgt	1056
Ile Leu Ile Val Thr Arg Leu Leu Pro Asp Ala Thr Gly Thr Thr Cys	
340 345 350	
ggc cag cgc ctt gag aag gtc ctt ggc acc gag cac tgc cat atc ctt	1104
Gly Gln Arg Leu Glu Lys Val Leu Gly Thr Glu His Cys His Ile Leu	
355 360 365	
cgc gtg cca ttc aga aca gaa aac gga atc gtt cgc aag tgg atc tcg	1152
Arg Val Pro Phe Arg Thr Glu Asn Gly Ile Val Arg Lys Trp Ile Ser	
370 375 380	
cga ttt gaa gtc tgg ccg tac ctg gag act tac act gat gac gtg gcg	1200
Arg Phe Glu Val Trp Pro Tyr Leu Glu Thr Tyr Thr Asp Asp Val Ala	
385 390 395 400	
cat gag att gct gga gag ctt cag gcc aat cct gac ctg atc atc gga	1248
His Glu Ile Ala Gly Glu Leu Gln Ala Asn Pro Asp Leu Ile Ile Gly	
405 410 415	
aac tac agt gac gga aac ctt gtt gcg tgt ttg ctc gcc cac aag atg	1296
Asn Tyr Ser Asp Gly Asn Leu Val Ala Cys Leu Leu Ala His Lys Met	
420 425 430	
ggt gtt act cac tgt acc att gcc cat gcg ctt gag aaa act aag tac	1344
Gly Val Thr His Cys Thr Ile Ala His Ala Leu Glu Lys Thr Lys Tyr	
435 440 445	
cct aac tcc gac ctc tac tgg aag aag ttt gag gat cac tac cac ttc	1392
Pro Asn Ser Asp Leu Tyr Trp Lys Lys Phe Glu Asp His Tyr His Phe	

450	455	460	
tcg tgc cag ttc acc act gac ttg att gca atg aac cat gcc gac ttc			1440
Ser Cys Gln Phe Thr Thr Asp Leu Ile Ala Met Asn His Ala Asp Phe			
465	470	475 480	
atc atc acc agt acc ttc caa gag atc gcc gga aac aag gac acc gtc			1488
Ile Ile Thr Ser Thr Phe Gln Glu Ile Ala Gly Asn Lys Asp Thr Val			
	485	490 495	
ggc cag tac gag tca cac atg gcg ttc aca atg cct ggc ctg tac cgc			1536
Gly Gln Tyr Glu Ser His Met Ala Phe Thr Met Pro Gly Leu Tyr Arg			
	500	505 510	
gtt gtc cac ggc att gat gtg ttc gac ccc aag ttc aac atc gtg tct			1584
Val Val His Gly Ile Asp Val Phe Asp Pro Lys Phe Asn Ile Val Ser			
	515	520 525	
cct ggc gcg gac ctg tcc atc tac ttc ccg tac acc gag tcg cac aag			1632
Pro Gly Ala Asp Leu Ser Ile Tyr Phe Pro Tyr Thr Glu Ser His Lys			
	530	535 540	
agg ctg acc tcc ctt cac ccg gag att gag gag ctc ctg tac agc caa			1680
Arg Leu Thr Ser Leu His Pro Glu Ile Glu Glu Leu Leu Tyr Ser Gln			
	545	550 555 560	
acc gag aac acg gag cac aag ttc gtt ctg aac gac agg aac aag cca			1728
Thr Glu Asn Thr Glu His Lys Phe Val Leu Asn Asp Arg Asn Lys Pro			
	565	570 575	
atc atc ttc tcc atg gct cgt ctc gac cgt gtg aag aac ttg act ggg			1776
Ile Ile Phe Ser Met Ala Arg Leu Asp Arg Val Lys Asn Leu Thr Gly			
	580	585 590	
ctg gtg gag ctg tac ggc cgg aac aag cgg ctg cag gag ctg gtg aac			1824
Leu Val Glu Leu Tyr Gly Arg Asn Lys Arg Leu Gln Glu Leu Val Asn			
	595	600 605	
ctc gtg gtc gtc tgc ggc gac cat ggc aac cct tcc aag gac aag gag			1872
Leu Val Val Val Cys Gly Asp His Gly Asn Pro Ser Lys Asp Lys Glu			
	610	615 620	
gag cag gcc gag ttc aag aag atg ttt gac ctc atc gag cag tac aac			1920
Glu Gln Ala Glu Phe Lys Lys Met Phe Asp Leu Ile Glu Gln Tyr Asn			
	625	630 635 640	
ctg aac ggg cac atc cgc tgg atc tcc gcc cag atg aac cgc gtc cgc			1968
Leu Asn Gly His Ile Arg Trp Ile Ser Ala Gln Met Asn Arg Val Arg			
	645	650 655	
aac ggc gag ctg tac cgc tac atc tgc gac acc aag ggc gcc ttc gtg			2016
Asn Gly Glu Leu Tyr Arg Tyr Ile Cys Asp Thr Lys Gly Ala Phe Val			
	660	665 670	
cag cct gct ttc tac gag gct ttc ggg ctg acg gtg gtt gag gcc atg			2064
Gln Pro Ala Phe Tyr Glu Ala Phe Gly Leu Thr Val Val Glu Ala Met			
	675	680 685	
acc tgc ggc ctg ccc acg ttt gcc aca gcc tac ggc ggt ccg gcc gag			2112
Thr Cys Gly Leu Pro Thr Phe Ala Thr Ala Tyr Gly Gly Pro Ala Glu			
	690	695 700	

atc atc gtg cac ggc gtg tct ggc tac cac atc gac cct tac cag ggc 2160
 Ile Ile Val His Gly Val Ser Gly Tyr His Ile Asp Pro Tyr Gln Gly
 705 710 715 720

gac aag gcg tcg gcc ctg ctc gtg gac ttc ttc gac aag tgc cag gcg 2208
 Asp Lys Ala Ser Ala Leu Leu Val Asp Phe Phe Asp Lys Cys Gln Ala
 725 730 735

gag ccg agc cac tgg agc aag atc tcc cag ggc ggg ctc cag cgt atc 2256
 Glu Pro Ser His Trp Ser Lys Ile Ser Gln Gly Gly Leu Gln Arg Ile
 740 745 750

gag gag aag tac acc tgg aag ctg tac tcg gag agg ctg atg acc ctc 2304
 Glu Glu Lys Tyr Thr Trp Lys Leu Tyr Ser Glu Arg Leu Met Thr Leu
 755 760 765

acc ggc gtg tac ggg ttc tgg aag tac gtg tcc aac ctg gag agg cgc 2352
 Thr Gly Val Tyr Gly Phe Trp Lys Tyr Val Ser Asn Leu Glu Arg Arg
 770 775 780

gag acc cgg cgg tac ctg gag atg ctg tac gcg ctc aag tac cgc acc 2400
 Glu Thr Arg Arg Tyr Leu Glu Met Leu Tyr Ala Leu Lys Tyr Arg Thr
 785 790 795 800

atg gcg agc acc gtg ccg ctg gcc gtg gag gga gag ccc tcc agc aag 2448
 Met Ala Ser Thr Val Pro Leu Ala Val Glu Gly Glu Pro Ser Ser Lys
 805 810 815

<210> 26
 <211> 816
 <212> PRT
 <213> Zea mays

<400> 26
 Met Gly Glu Gly Ala Gly Asp Arg Val Leu Ser Arg Leu His Ser Val
 1 5 10 15

Arg Glu Arg Ile Gly Asp Ser Leu Ser Ala His Pro Asn Glu Leu Val
 20 25 30

Ala Val Phe Thr Arg Leu Lys Asn Leu Gly Lys Gly Met Leu Gln Pro
 35 40 45

His Gln Ile Ile Ala Glu Tyr Asn Asn Ala Ile Pro Glu Ala Glu Arg
 50 55 60

Glu Lys Leu Lys Asp Gly Ala Phe Glu Asp Val Leu Arg Ala Ala Gln
 65 70 75 80

Glu Ala Ile Val Ile Pro Pro Trp Val Ala Leu Ala Ile Arg Pro Arg
 85 90 95

Pro Gly Val Trp Glu Tyr Val Arg Val Asn Val Ser Glu Leu Ala Val

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

340 345 350
 Gly Gln Arg Leu Glu Lys Val Leu Gly Thr Glu His Cys His Ile Leu
 355 360 365
 Arg Val Pro Phe Arg Thr Glu Asn Gly Ile Val Arg Lys Trp Ile Ser
 370 375 380
 Arg Phe Glu Val Trp Pro Tyr Leu Glu Thr Tyr Thr Asp Asp Val Ala
 385 390 395 400
 His Glu Ile Ala Gly Glu Leu Gln Ala Asn Pro Asp Leu Ile Ile Gly
 405 410 415
 Asn Tyr Ser Asp Gly Asn Leu Val Ala Cys Leu Leu Ala His Lys Met
 420 425 430
 Gly Val Thr His Cys Thr Ile Ala His Ala Leu Glu Lys Thr Lys Tyr
 435 440 445
 Pro Asn Ser Asp Leu Tyr Trp Lys Lys Phe Glu Asp His Tyr His Phe
 450 455 460
 Ser Cys Gln Phe Thr Thr Asp Leu Ile Ala Met Asn His Ala Asp Phe
 465 470 475 480
 Ile Ile Thr Ser Thr Phe Gln Glu Ile Ala Gly Asn Lys Asp Thr Val
 485 490 495
 Gly Gln Tyr Glu Ser His Met Ala Phe Thr Met Pro Gly Leu Tyr Arg
 500 505 510
 Val Val His Gly Ile Asp Val Phe Asp Pro Lys Phe Asn Ile Val Ser
 515 520 525
 Pro Gly Ala Asp Leu Ser Ile Tyr Phe Pro Tyr Thr Glu Ser His Lys
 530 535 540
 Arg Leu Thr Ser Leu His Pro Glu Ile Glu Glu Leu Leu Tyr Ser Gln
 545 550 555 560
 Thr Glu Asn Thr Glu His Lys Phe Val Leu Asn Asp Arg Asn Lys Pro
 565 570 575
 Ile Ile Phe Ser Met Ala Arg Leu Asp Arg Val Lys Asn Leu Thr Gly
 580 585 590

Leu Val Glu Leu Tyr Gly Arg Asn Lys Arg Leu Gln Glu Leu Val Asn
 595 600 605

Leu Val Val Val Cys Gly Asp His Gly Asn Pro Ser Lys Asp Lys Glu
 610 615 620

Glu Gln Ala Glu Phe Lys Lys Met Phe Asp Leu Ile Glu Gln Tyr Asn
 625 630 635 640

Leu Asn Gly His Ile Arg Trp Ile Ser Ala Gln Met Asn Arg Val Arg
 645 650 655

Asn Gly Glu Leu Tyr Arg Tyr Ile Cys Asp Thr Lys Gly Ala Phe Val
 660 665 670

Gln Pro Ala Phe Tyr Glu Ala Phe Gly Leu Thr Val Val Glu Ala Met
 675 680 685

Thr Cys Gly Leu Pro Thr Phe Ala Thr Ala Tyr Gly Gly Pro Ala Glu
 690 695 700

Ile Ile Val His Gly Val Ser Gly Tyr His Ile Asp Pro Tyr Gln Gly
 705 710 715 720

Asp Lys Ala Ser Ala Leu Leu Val Asp Phe Phe Asp Lys Cys Gln Ala
 725 730 735

Glu Pro Ser His Trp Ser Lys Ile Ser Gln Gly Gly Leu Gln Arg Ile
 740 745 750

Glu Glu Lys Tyr Thr Trp Lys Leu Tyr Ser Glu Arg Leu Met Thr Leu
 755 760 765

Thr Gly Val Tyr Gly Phe Trp Lys Tyr Val Ser Asn Leu Glu Arg Arg
 770 775 780

Glu Thr Arg Arg Tyr Leu Glu Met Leu Tyr Ala Leu Lys Tyr Arg Thr
 785 790 795 800

Met Ala Ser Thr Val Pro Leu Ala Val Glu Gly Glu Pro Ser Ser Lys
 805 810 815

<210> 27
 <211> 2826

<212> DNA
 <213> *Triticum aestivum*
 <220>
 <221> CDS
 <222> (97)..(2547)
 <400> 27
 tcgctgtcga cgatttcgtc gagcctgctc ccctgtgacg tttccttggt tctcgcctc 60
 ctctccctc ctctggagca aggtttgagg acagca atg ggc gag act gcc gga 114
 Met Gly Glu Thr Ala Gly
 1 5
 gag cgc gcc ctg agc cgt atc cac agc gtg agg gag cgc atc ggc gat 162
 Glu Arg Ala Leu Ser Arg Ile His Ser Val Arg Glu Arg Ile Gly Asp
 10 15 20
 tcc ctc tcc gcg cac acc aat gag ctt gtc gcc gtc ttc tca agg ctt 210
 Ser Leu Ser Ala His Thr Asn Glu Leu Val Ala Val Phe Ser Arg Leu
 25 30 35
 gtt aac caa gga aag ggg atg ctg cag ccc cat cag ata act gct gag 258
 Val Asn Gln Gly Lys Gly Met Leu Gln Pro His Gln Ile Thr Ala Glu
 40 45 50
 tac aat gcc gcg atc cca gag gcc gag cgt gag aag ctc aag gac acc 306
 Tyr Asn Ala Ala Ile Pro Glu Ala Glu Arg Glu Lys Leu Lys Asp Thr
 55 60 65 70
 gcc ttt gag gat ctc cta agg ggc gca cag gag gca att gtc atc cct 354
 Ala Phe Glu Asp Leu Leu Arg Gly Ala Gln Glu Ala Ile Val Ile Pro
 75 80 85
 cca tgg gtt gct ctc gcc atc cgg cca agg ccc ggc gtc tgg gag tat 402
 Pro Trp Val Ala Leu Ala Ile Arg Pro Arg Pro Gly Val Trp Glu Tyr
 90 95 100
 gtg agg gtc aat gtg agc gag ctt ggt gtt gag gag tta agc atc gct 450
 Val Arg Val Asn Val Ser Glu Leu Gly Val Glu Glu Leu Ser Ile Ala
 105 110 115
 gag tat ttg cag ttc aag gaa caa ctg gcg aat gga agc atc gat aac 498
 Glu Tyr Leu Gln Phe Lys Glu Gln Leu Ala Asn Gly Ser Ile Asp Asn
 120 125 130
 aac ttt gtg ctt gag ctg gac ttt gag cca ttc aac gcc tcc ttc cct 546
 Asn Phe Val Leu Glu Leu Asp Phe Glu Pro Phe Asn Ala Ser Phe Pro
 135 140 145 150
 cgc cca tcg ctg tcg aag tcc att ggc aac ggt gtg cag ttt ctg aac 594
 Arg Pro Ser Leu Ser Lys Ser Ile Gly Asn Gly Val Gln Phe Leu Asn
 155 160 165
 agg cac ttg tca tcg aag ctg ttc cat gac aag gag agc atg tac cca 642
 Arg His Leu Ser Ser Lys Leu Phe His Asp Lys Glu Ser Met Tyr Pro
 170 175 180
 ttg ctc aac ttc ctt cgc gcg cac aac tac aag ggg atg acc atg atg 690
 Leu Leu Asn Phe Leu Arg Ala His Asn Tyr Lys Gly Met Thr Met Met
 185 190 195

ttg aac gac aga att cgg agt ctc agt acc ctc caa ggt gca ctc agg Leu Asn Asp Arg Ile Arg Ser Leu Ser Thr Leu Gln Gly Ala Leu Arg 200 205 210	738
aag gca gag aca cat ctg tca ggc ctt cca gct gac acc ccc tac tcg Lys Ala Glu Thr His Leu Ser Gly Leu Pro Ala Asp Thr Pro Tyr Ser 215 220 225 230	786
gag ttc cac cac cgg ttc cag gaa ctt ggt ctg gag aaa ggt tgg ggc Glu Phe His His Arg Phe Gln Glu Leu Gly Leu Glu Lys Gly Trp Gly 235 240 245	834
gac tgt gct cag cgt gcg agt gag act atc cac ctt ctc ttg gac ctt Asp Cys Ala Gln Arg Ala Ser Glu Thr Ile His Leu Leu Leu Asp Leu 250 255 260	882
ctc gag gcc cct gat cca tcc tcc ttg gag aag ttc ctc gga aca atc Leu Glu Ala Pro Asp Pro Ser Ser Leu Glu Lys Phe Leu Gly Thr Ile 265 270 275	930
cca atg gtg ttc aat gtt gtt atc ctc tct cct cat ggt tac ttt gct Pro Met Val Phe Asn Val Val Ile Leu Ser Pro His Gly Tyr Phe Ala 280 285 290	978
cag gcc aat gtc ttg ggg tac cct gat act ggt gga cag att gtc tac Gln Ala Asn Val Leu Gly Tyr Pro Asp Thr Gly Gly Gln Ile Val Tyr 295 300 305 310	1026
att ttg gac caa gtc cgt gct atg gag aat gag atg ctg ttg aga atc Ile Leu Asp Gln Val Arg Ala Met Glu Asn Glu Met Leu Leu Arg Ile 315 320 325	1074
aag cag caa ggt ctt gac att aca cca aag att cta ata gtc acc agg Lys Gln Gln Gly Leu Asp Ile Thr Pro Lys Ile Leu Ile Val Thr Arg 330 335 340	1122
ttg ctc cct gat gca cat ggc acc acc tgt ggc cag cgc ctt gag aag Leu Leu Pro Asp Ala His Gly Thr Thr Cys Gly Gln Arg Leu Glu Lys 345 350 355	1170
gtc ctt ggc acc gag cac acc cac atc ctg cgt gtg cca ttc aaa aca Val Leu Gly Thr Glu His Thr His Ile Leu Arg Val Pro Phe Lys Thr 360 365 370	1218
gaa gat ggt att gtt cgc aaa tgg atc tcc cgc ttt gaa gtc tgg cct Glu Asp Gly Ile Val Arg Lys Trp Ile Ser Arg Phe Glu Val Trp Pro 375 380 385 390	1266
tac ctg gaa gct tac act gat gat gtg gca cac gag atc gcc gga gag Tyr Leu Glu Ala Tyr Thr Asp Asp Val Ala His Glu Ile Ala Gly Glu 395 400 405	1314
ctg cag gcc act cct gac ctg atc att gga aac tac agt gat ggc aac Leu Gln Ala Thr Pro Asp Leu Ile Ile Gly Asn Tyr Ser Asp Gly Asn 410 415 420	1362
cta gtc gcg tgt ttg ttg gct cac aag ttg gga gtt act cat tgt acc Leu Val Ala Cys Leu Leu Ala His Lys Leu Gly Val Thr His Cys Thr 425 430 435	1410

att gcg cat gca ctc gag aaa acc aag tat ccc aac tcc gac ctt tac Ile Ala His Ala Leu Glu Lys Thr Lys Tyr Pro Asn Ser Asp Leu Tyr 440 445 450	1458
tgg aag aaa ttt gag gat cac tac cac ttc tcc tgc cag ttc aca gct Trp Lys Lys Phe Glu Asp His Tyr His Phe Ser Cys Gln Phe Thr Ala 455 460 465 470	1506
gac ctg att gca atg aat cat gct gac ttc atc atc acc aat act ttc Asp Leu Ile Ala Met Asn His Ala Asp Phe Ile Ile Thr Asn Thr Phe 475 480 485	1554
caa gag att gcc gga aac aag gac acc gta ggg cag tac gag tcg cac Gln Glu Ile Ala Gly Asn Lys Asp Thr Val Gly Gln Tyr Glu Ser His 490 495 500	1602
atg gca ttc aca atg cca ggc ctc tat cgt gtt gtc cat ggt att gat Met Ala Phe Thr Met Pro Gly Leu Tyr Arg Val Val His Gly Ile Asp 505 510 515	1650
gtc ttc gac ccc aag ttc aac atc gtc tcc cct ggt gct gac atg tcc Val Phe Asp Pro Lys Phe Asn Ile Val Ser Pro Gly Ala Asp Met Ser 520 525 530	1698
atc tac ttc cca tac act gaa cag cag aag agg ctt acc tcc ctc cat Ile Tyr Phe Pro Tyr Thr Glu Gln Gln Lys Arg Leu Thr Ser Leu His 535 540 545 550	1746
act gag att gag gag cta ctc ttc agt gat gtt gag aat gct gag cac Thr Glu Ile Glu Glu Leu Leu Phe Ser Asp Val Glu Asn Ala Glu His 555 560 565	1794
aaa ttt gtg ctg aag gac aag aag aag ccg atc atc ttc tcg atg gct Lys Phe Val Leu Lys Asp Lys Lys Lys Pro Ile Ile Phe Ser Met Ala 570 575 580	1842
agg ctg gac cgt gtc aag aat atg act ggc ctg gta gaa atg tat ggg Arg Leu Asp Arg Val Lys Asn Met Thr Gly Leu Val Glu Met Tyr Gly 585 590 595	1890
cgg aat cct cgc cta cag gag ctg gta aac ctg gtg gtt gtt tgt ggt Arg Asn Pro Arg Leu Gln Glu Leu Val Asn Leu Val Val Val Cys Gly 600 605 610	1938
gac cat gga aag gtg tcc aag gac aag gag gag cag gca gag ttc aaa Asp His Gly Lys Val Ser Lys Asp Lys Glu Glu Gln Ala Glu Phe Lys 615 620 625 630	1986
aag atg ttt gat ctt atc gaa cag tac aac ctg att ggt cac atc cgc Lys Met Phe Asp Leu Ile Glu Gln Tyr Asn Leu Ile Gly His Ile Arg 635 640 645	2034
tgg atc tct gct cag atg aac cgt gtc cgc aat ggt gag ctc tac cgc Trp Ile Ser Ala Gln Met Asn Arg Val Arg Asn Gly Glu Leu Tyr Arg 650 655 660	2082
tac atc tgc gac atg aag gga gcc ttt gtg cag cct gct ttc tat gag Tyr Ile Cys Asp Met Lys Gly Ala Phe Val Gln Pro Ala Phe Tyr Glu 665 670 675	2130
gct ttc ggt ctt acc gtg ata gag gcc atg aca tgt ggc ctt cca aca	2178

Ala Phe Gly Leu Thr Val Ile Glu Ala Met Thr Cys Gly Leu Pro Thr
680 685 690

ttc gcc act gca tat ggt ggt cca gct gag atc att gtg cac ggt gtg 2226
Phe Ala Thr Ala Tyr Gly Gly Pro Ala Glu Ile Ile Val His Gly Val
695 700 705 710

tcc ggc tac cac atc gat cct tac cag aat gac aag gcc tcc gca ctg 2274
Ser Gly Tyr His Ile Asp Pro Tyr Gln Asn Asp Lys Ala Ser Ala Leu
715 720 725

ctt gtg gac ttc ttt ggg aag tgc cag gaa gac ccg agc cac tgg aac 2322
Leu Val Asp Phe Phe Gly Lys Cys Gln Glu Asp Pro Ser His Trp Asn
730 735 740

aag atc tcg cag gga gga ctc cag cgc atc gag gag aag tac acc tgg 2370
Lys Ile Ser Gln Gly Gly Leu Gln Arg Ile Glu Glu Lys Tyr Thr Trp
745 750 755

aag ctg tac tct gag agg ctg atg acc ctt tct ggt gtc tat ggt ttc 2418
Lys Leu Tyr Ser Glu Arg Leu Met Thr Leu Ser Gly Val Tyr Gly Phe
760 765 770

tgg aag tat gtc tcc aac ctc gac agg cgc gag act cgt cgc tac ctt 2466
Trp Lys Tyr Val Ser Asn Leu Asp Arg Arg Glu Thr Arg Arg Tyr Leu
775 780 785 790

gaa atg ctc tac gcc ctc aag tac cgc aaa atg gct gca act gtc cca 2514
Glu Met Leu Tyr Ala Leu Lys Tyr Arg Lys Met Ala Ala Thr Val Pro
795 800 805

ttg gct gtt gag ggc gag acc tcg ggc aaa tga tttgtcctta ccagagaaat 2567
Leu Ala Val Glu Gly Glu Thr Ser Gly Lys
810 815

aaatggcggg cgagcgctcc gctttaccgt tcttgattca gtgatgaagc acagatcgga 2627

gagtgttatg cctttgattg tcctttgtta ccgttcttga ttcagtatga acttttagttc 2687

ctgttgaggc cccggccgtg ttgtcgcggt aggggaactg gatggatgct gtgttggttg 2747

tactttctca tggatacagt atttgaatga atgaataaga tagtctttgt tttgaaaaaa 2807

aaaaaaaaa agcggccgc 2826

<210> 28

<211> 816

<212> PRT

<213> Triticum aestivum

<400> 28

Met Gly Glu Thr Ala Gly Glu Arg Ala Leu Ser Arg Ile His Ser Val
1 5 10 15

Arg Glu Arg Ile Gly Asp Ser Leu Ser Ala His Thr Asn Glu Leu Val
20 25 30

Ala Val Phe Ser Arg Leu Val Asn Gln Gly Lys Gly Met Leu Gln Pro

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

Pro His Gly Tyr Phe Ala Gln Ala Asn Val Leu Gly Tyr Pro Asp Thr
 290 295 300

Gly Gly Gln Ile Val Tyr Ile Leu Asp Gln Val Arg Ala Met Glu Asn
 305 310 315 320

Glu Met Leu Leu Arg Ile Lys Gln Gln Gly Leu Asp Ile Thr Pro Lys
 325 330 335

Ile Leu Ile Val Thr Arg Leu Leu Pro Asp Ala His Gly Thr Thr Cys
 340 345 350

Gly Gln Arg Leu Glu Lys Val Leu Gly Thr Glu His Thr His Ile Leu
 355 360 365

Arg Val Pro Phe Lys Thr Glu Asp Gly Ile Val Arg Lys Trp Ile Ser
 370 375 380

Arg Phe Glu Val Trp Pro Tyr Leu Glu Ala Tyr Thr Asp Asp Val Ala
 385 390 395 400

His Glu Ile Ala Gly Glu Leu Gln Ala Thr Pro Asp Leu Ile Ile Gly
 405 410 415

Asn Tyr Ser Asp Gly Asn Leu Val Ala Cys Leu Leu Ala His Lys Leu
 420 425 430

Gly Val Thr His Cys Thr Ile Ala His Ala Leu Glu Lys Thr Lys Tyr
 435 440 445

Pro Asn Ser Asp Leu Tyr Trp Lys Lys Phe Glu Asp His Tyr His Phe
 450 455 460

Ser Cys Gln Phe Thr Ala Asp Leu Ile Ala Met Asn His Ala Asp Phe
 465 470 475 480

Ile Ile Thr Asn Thr Phe Gln Glu Ile Ala Gly Asn Lys Asp Thr Val
 485 490 495

Gly Gln Tyr Glu Ser His Met Ala Phe Thr Met Pro Gly Leu Tyr Arg
 500 505 510

Val Val His Gly Ile Asp Val Phe Asp Pro Lys Phe Asn Ile Val Ser
 515 520 525

Pro Gly Ala Asp Met Ser Ile Tyr Phe Pro Tyr Thr Glu Gln Gln Lys
 530 535 540

Arg Leu Thr Ser Leu His Thr Glu Ile Glu Glu Leu Leu Phe Ser Asp
 545 550 555 560

Val Glu Asn Ala Glu His Lys Phe Val Leu Lys Asp Lys Lys Lys Pro
 565 570 575

Ile Ile Phe Ser Met Ala Arg Leu Asp Arg Val Lys Asn Met Thr Gly
 580 585 590

Leu Val Glu Met Tyr Gly Arg Asn Pro Arg Leu Gln Glu Leu Val Asn
 595 600 605

Leu Val Val Val Cys Gly Asp His Gly Lys Val Ser Lys Asp Lys Glu
 610 615 620

Glu Gln Ala Glu Phe Lys Lys Met Phe Asp Leu Ile Glu Gln Tyr Asn
 625 630 635 640

Leu Ile Gly His Ile Arg Trp Ile Ser Ala Gln Met Asn Arg Val Arg
 645 650 655

Asn Gly Glu Leu Tyr Arg Tyr Ile Cys Asp Met Lys Gly Ala Phe Val
 660 665 670

Gln Pro Ala Phe Tyr Glu Ala Phe Gly Leu Thr Val Ile Glu Ala Met
 675 680 685

Thr Cys Gly Leu Pro Thr Phe Ala Thr Ala Tyr Gly Gly Pro Ala Glu
 690 695 700

Ile Ile Val His Gly Val Ser Gly Tyr His Ile Asp Pro Tyr Gln Asn
 705 710 715 720

Asp Lys Ala Ser Ala Leu Leu Val Asp Phe Phe Gly Lys Cys Gln Glu
 725 730 735

Asp Pro Ser His Trp Asn Lys Ile Ser Gln Gly Gly Leu Gln Arg Ile
 740 745 750

Glu Glu Lys Tyr Thr Trp Lys Leu Tyr Ser Glu Arg Leu Met Thr Leu
 755 760 765

Ser Gly Val Tyr Gly Phe Trp Lys Tyr Val Ser Asn Leu Asp Arg Arg
770 775 780

Glu Thr Arg Arg Tyr Leu Glu Met Leu Tyr Ala Leu Lys Tyr Arg Lys
785 790 795 800

Met Ala Ala Thr Val Pro Leu Ala Val Glu Gly Glu Thr Ser Gly Lys
805 810 815

<210> 29
<211> 2448
<212> DNA
<213> Triticum aestivum

<220>
<221> CDS
<222> (1)..(2448)

<400> 29
atg ggc gag act gcc gga gag cgc gcc ctg agc cgt atc cac agc gtg 48
Met Gly Glu Thr Ala Gly Glu Arg Ala Leu Ser Arg Ile His Ser Val
1 5 10 15
agg gag cgc atc ggc gat tcc ctc tcc gcg cac acc aat gag ctt gtc 96
Arg Glu Arg Ile Gly Asp Ser Leu Ser Ala His Thr Asn Glu Leu Val
20 25 30
gcc gtc ttc tca agg ctt gtt aac caa gga aag ggg atg ctg cag ccc 144
Ala Val Phe Ser Arg Leu Val Asn Gln Gly Lys Gly Met Leu Gln Pro
35 40 45
cat cag ata act gct gag tac aat gcc gcg atc cca gag gcc gag cgt 192
His Gln Ile Thr Ala Glu Tyr Asn Ala Ala Ile Pro Glu Ala Glu Arg
50 55 60
gag aag ctc aag gac acc gcc ttt gag gat ctc cta agg ggc gca cag 240
Glu Lys Leu Lys Asp Thr Ala Phe Glu Asp Leu Leu Arg Gly Ala Gln
65 70 75 80
gag gca att gtc atc cct cca tgg gtt gct ctc gcc atc cgg cca agg 288
Glu Ala Ile Val Ile Pro Pro Trp Val Ala Leu Ala Ile Arg Pro Arg
85 90 95
ccc ggc gtc tgg gag tat gtg agg gtc aat gtg agc gag ctt ggt gtt 336
Pro Gly Val Trp Glu Tyr Val Arg Val Asn Val Ser Glu Leu Gly Val
100 105 110
gag gag tta agc atc gct gag tat ttg cag ttc aag gaa caa ctg gcg 384
Glu Glu Leu Ser Ile Ala Glu Tyr Leu Gln Phe Lys Glu Gln Leu Ala
115 120 125
aat gga agc atc gat aac aac ttt gtg ctt gag ctg gac ttt gag cca 432
Asn Gly Ser Ile Asp Asn Asn Phe Val Leu Glu Leu Asp Phe Glu Pro
130 135 140
ttc aac gcc tcc ttc cct cgc cca tcg ctg tcg aag tcc att ggc aac 480
Phe Asn Ala Ser Phe Pro Arg Pro Ser Leu Ser Lys Ser Ile Gly Asn
145 150 155 160

ggt gtg cag ttt ctg aac agg cac ttg tca tgc aag ctg ttc cat gac Gly Val Gln Phe Leu Asn Arg His Leu Ser Ser Lys Leu Phe His Asp 165 170 175	528
aag gag agc atg tac cca ttg ctc aac ttc ctt cgc gcg cac aac tac Lys Glu Ser Met Tyr Pro Leu Leu Asn Phe Leu Arg Ala His Asn Tyr 180 185 190	576
aag ggg atg acc atg atg ttg aac gac aga att cgg agt ctc agt acc Lys Gly Met Thr Met Met Leu Asn Asp Arg Ile Arg Ser Leu Ser Thr 195 200 205	624
ctc caa ggt gca ctc agg aag gca gag aca cat ctg tca ggc ctt cca Leu Gln Gly Ala Leu Arg Lys Ala Glu Thr His Leu Ser Gly Leu Pro 210 215 220	672
gct gac acc ccc tac tgc gag ttc cac cac cgg ttc cag gaa ctt ggt Ala Asp Thr Pro Tyr Ser Glu Phe His His Arg Phe Gln Glu Leu Gly 225 230 235 240	720
ctg gag aaa ggt tgg ggc gac tgt gct cag cgt gcg agt gag act atc Leu Glu Lys Gly Trp Gly Asp Cys Ala Gln Arg Ala Ser Glu Thr Ile 245 250 255	768
cac ctt ctc ttg gac ctt ctc gag gcc cct gat cca tcc tcc ttg gag His Leu Leu Leu Asp Leu Leu Glu Ala Pro Asp Pro Ser Ser Leu Glu 260 265 270	816
aag ttc ctc gga aca atc cca atg gtg ttc aat gtt gtt atc ctc tct Lys Phe Leu Gly Thr Ile Pro Met Val Phe Asn Val Val Ile Leu Ser 275 280 285	864
cct cat ggt tac ttt gct cag gcc aat gtc ttg ggg tac cct gat act Pro His Gly Tyr Phe Ala Gln Ala Asn Val Leu Gly Tyr Pro Asp Thr 290 295 300	912
ggt gga cag att gtc tac att ttg gac caa gtc cgt gct atg gag aat Gly Gly Gln Ile Val Tyr Ile Leu Asp Gln Val Arg Ala Met Glu Asn 305 310 315 320	960
gag atg ctg ttg aga atc aag cag caa ggt ctt gac att aca cca aag Glu Met Leu Leu Arg Ile Lys Gln Gln Gly Leu Asp Ile Thr Pro Lys 325 330 335	1008
att cta ata gtc acc agg ttg ctc cct gat gca cat ggc acc acc tgt Ile Leu Ile Val Thr Arg Leu Leu Pro Asp Ala His Gly Thr Thr Cys 340 345 350	1056
ggc cag cgc ctt gag aag gtc ctt ggc acc gag cac acc cac atc ctg Gly Gln Arg Leu Glu Lys Val Leu Gly Thr Glu His Thr His Ile Leu 355 360 365	1104
cgt gtg cca ttc aaa aca gaa gat ggt att gtt cgc aaa tgg atc tcc Arg Val Pro Phe Lys Thr Glu Asp Gly Ile Val Arg Lys Trp Ile Ser 370 375 380	1152
cgc ttt gaa gtc tgg cct tac ctg gaa gct tac act gat gat gtg gca Arg Phe Glu Val Trp Pro Tyr Leu Glu Ala Tyr Thr Asp Asp Val Ala 385 390 395 400	1200

cac gag atc gcc gga gag ctg cag gcc act cct gac ctg atc att gga His Glu Ile Ala Gly Glu Leu Gln Ala Thr Pro Asp Leu Ile Ile Gly 405 410 415	1248
aac tac agt gat ggc aac cta gtc gcg tgt ttg ttg gct cac aag ttg Asn Tyr Ser Asp Gly Asn Leu Val Ala Cys Leu Leu Ala His Lys Leu 420 425 430	1296
gga gtt act cat tgt acc att gcg cat gca ctc gag aaa acc aag tat Gly Val Thr His Cys Thr Ile Ala His Ala Leu Glu Lys Thr Lys Tyr 435 440 445	1344
ccc aac tcc gac ctt tac tgg aag aaa ttt gag gat cac tac cac ttc Pro Asn Ser Asp Leu Tyr Trp Lys Lys Phe Glu Asp His Tyr His Phe 450 455 460	1392
tcc tgc cag ttc aca gct gac ctg att gca atg aat cat gct gac ttc Ser Cys Gln Phe Thr Ala Asp Leu Ile Ala Met Asn His Ala Asp Phe 465 470 475 480	1440
atc atc acc aat act ttc caa gag att gcc gga aac aag gac acc gta Ile Ile Thr Asn Thr Phe Gln Glu Ile Ala Gly Asn Lys Asp Thr Val 485 490 495	1488
ggg cag tac gag tcg cac atg gca ttc aca atg cca ggc ctc tat cgt Gly Gln Tyr Glu Ser His Met Ala Phe Thr Met Pro Gly Leu Tyr Arg 500 505 510	1536
gtt gtc cat ggt att gat gtc ttc gac ccc aag ttc aac atc gtc tcc Val Val His Gly Ile Asp Val Phe Asp Pro Lys Phe Asn Ile Val Ser 515 520 525	1584
cct ggt gct gac atg tcc atc tac ttc cca tac act gaa cag cag aag Pro Gly Ala Asp Met Ser Ile Tyr Phe Pro Tyr Thr Glu Gln Gln Lys 530 535 540	1632
agg ctt acc tcc ctc cat act gag att gag gag cta ctc ttc agt gat Arg Leu Thr Ser Leu His Thr Glu Ile Glu Glu Leu Leu Phe Ser Asp 545 550 555 560	1680
gtt gag aat gct gag cac aaa ttt gtg ctg aag gac aag aag aag ccg Val Glu Asn Ala Glu His Lys Phe Val Leu Lys Asp Lys Lys Lys Pro 565 570 575	1728
atc atc ttc tcg atg gct agg ctg gac cgt gtc aag aat atg act ggc Ile Ile Phe Ser Met Ala Arg Leu Asp Arg Val Lys Asn Met Thr Gly 580 585 590	1776
ctg gta gaa atg tat ggg cgg aat cct cgc cta cag gag ctg gta aac Leu Val Glu Met Tyr Gly Arg Asn Pro Arg Leu Gln Glu Leu Val Asn 595 600 605	1824
ctg gtg gtt gtt tgt ggt gac cat gga aag gtg tcc aag gac aag gag Leu Val Val Val Cys Gly Asp His Gly Lys Val Ser Lys Asp Lys Glu 610 615 620	1872
gag cag gca gag ttc aaa aag atg ttt gat ctt atc gaa cag tac aac Glu Gln Ala Glu Phe Lys Lys Met Phe Asp Leu Ile Glu Gln Tyr Asn 625 630 635 640	1920
ctg att ggt cac atc cgc tgg atc tct gct cag atg aac cgt gtc cgc	1968

Leu Ile Gly His Ile Arg Trp Ile Ser Ala Gln Met Asn Arg Val Arg
 645 650 655
 aat ggt gag ctc tac cgc tac atc tgc gac atg aag gga gcc ttt gtg 2016
 Asn Gly Glu Leu Tyr Arg Tyr Ile Cys Asp Met Lys Gly Ala Phe Val
 660 665 670
 cag cct gct ttc tat gag gct ttc ggt ctt acc gtg ata gag gcc atg 2064
 Gln Pro Ala Phe Tyr Glu Ala Phe Gly Leu Thr Val Ile Glu Ala Met
 675 680 685
 aca tgt ggc ctt cca aca ttc gcc act gca tat ggt ggt cca gct gag 2112
 Thr Cys Gly Leu Pro Thr Phe Ala Thr Ala Tyr Gly Gly Pro Ala Glu
 690 695 700
 atc att gtg cac ggt gtg tcc ggc tac cac atc gat cct tac cag aat 2160
 Ile Ile Val His Gly Val Ser Gly Tyr His Ile Asp Pro Tyr Gln Asn
 705 710 715 720
 gac aag gcc tcc gca ctg ctt gtg gac ttc ttt ggg aag tgc cag gaa 2208
 Asp Lys Ala Ser Ala Leu Leu Val Asp Phe Phe Gly Lys Cys Gln Glu
 725 730 735
 gac ccg agc cac tgg aac aag atc tcg cag gga gga ctc cag cgc atc 2256
 Asp Pro Ser His Trp Asn Lys Ile Ser Gln Gly Gly Leu Gln Arg Ile
 740 745 750
 gag gag aag tac acc tgg aag ctg tac tct gag agg ctg atg acc ctt 2304
 Glu Glu Lys Tyr Thr Trp Lys Leu Tyr Ser Glu Arg Leu Met Thr Leu
 755 760 765
 tct ggt gtc tat ggt ttc tgg aag tat gtc tcc aac ctc gac agg cgc 2352
 Ser Gly Val Tyr Gly Phe Trp Lys Tyr Val Ser Asn Leu Asp Arg Arg
 770 775 780
 gag act cgt cgc tac ctt gaa atg ctc tac gcc ctc aag tac cgc aaa 2400
 Glu Thr Arg Arg Tyr Leu Glu Met Leu Tyr Ala Leu Lys Tyr Arg Lys
 785 790 795 800
 atg gct gca act gtc cca ttg gct gtt gag ggc gag acc tcg ggc aaa 2448
 Met Ala Ala Thr Val Pro Leu Ala Val Glu Gly Glu Thr Ser Gly Lys
 805 810 815
 <210> 30
 <211> 816
 <212> PRT
 <213> Triticum aestivum
 <400> 30
 Met Gly Glu Thr Ala Gly Glu Arg Ala Leu Ser Arg Ile His Ser Val
 1 5 10 15
 Arg Glu Arg Ile Gly Asp Ser Leu Ser Ala His Thr Asn Glu Leu Val
 20 25 30
 Ala Val Phe Ser Arg Leu Val Asn Gln Gly Lys Gly Met Leu Gln Pro
 35 40 45

His Gln Ile Thr Ala Glu Tyr Asn Ala Ala Ile Pro Glu Ala Glu Arg
 50 55 60

Glu Lys Leu Lys Asp Thr Ala Phe Glu Asp Leu Leu Arg Gly Ala Gln
 65 70 75 80

Glu Ala Ile Val Ile Pro Pro Trp Val Ala Leu Ala Ile Arg Pro Arg
 85 90 95

Pro Gly Val Trp Glu Tyr Val Arg Val Asn Val Ser Glu Leu Gly Val
 100 105 110

Glu Glu Leu Ser Ile Ala Glu Tyr Leu Gln Phe Lys Glu Gln Leu Ala
 115 120 125

Asn Gly Ser Ile Asp Asn Asn Phe Val Leu Glu Leu Asp Phe Glu Pro
 130 135 140

Phe Asn Ala Ser Phe Pro Arg Pro Ser Leu Ser Lys Ser Ile Gly Asn
 145 150 155 160

Gly Val Gln Phe Leu Asn Arg His Leu Ser Ser Lys Leu Phe His Asp
 165 170 175

Lys Glu Ser Met Tyr Pro Leu Leu Asn Phe Leu Arg Ala His Asn Tyr
 180 185 190

Lys Gly Met Thr Met Met Leu Asn Asp Arg Ile Arg Ser Leu Ser Thr
 195 200 205

Leu Gln Gly Ala Leu Arg Lys Ala Glu Thr His Leu Ser Gly Leu Pro
 210 215 220

Ala Asp Thr Pro Tyr Ser Glu Phe His His Arg Phe Gln Glu Leu Gly
 225 230 235 240

Leu Glu Lys Gly Trp Gly Asp Cys Ala Gln Arg Ala Ser Glu Thr Ile
 245 250 255

His Leu Leu Leu Asp Leu Leu Glu Ala Pro Asp Pro Ser Ser Leu Glu
 260 265 270

Lys Phe Leu Gly Thr Ile Pro Met Val Phe Asn Val Val Ile Leu Ser
 275 280 285

Pro His Gly Tyr Phe Ala Gln Ala Asn Val Leu Gly Tyr Pro Asp Thr
 290 295 300

Gly Gly Gln Ile Val Tyr Ile Leu Asp Gln Val Arg Ala Met Glu Asn
 305 310 315 320

Glu Met Leu Leu Arg Ile Lys Gln Gln Gly Leu Asp Ile Thr Pro Lys
 325 330 335

Ile Leu Ile Val Thr Arg Leu Leu Pro Asp Ala His Gly Thr Thr Cys
 340 345 350

Gly Gln Arg Leu Glu Lys Val Leu Gly Thr Glu His Thr His Ile Leu
 355 360 365

Arg Val Pro Phe Lys Thr Glu Asp Gly Ile Val Arg Lys Trp Ile Ser
 370 375 380

Arg Phe Glu Val Trp Pro Tyr Leu Glu Ala Tyr Thr Asp Asp Val Ala
 385 390 395 400

His Glu Ile Ala Gly Glu Leu Gln Ala Thr Pro Asp Leu Ile Ile Gly
 405 410 415

Asn Tyr Ser Asp Gly Asn Leu Val Ala Cys Leu Leu Ala His Lys Leu
 420 425 430

Gly Val Thr His Cys Thr Ile Ala His Ala Leu Glu Lys Thr Lys Tyr
 435 440 445

Pro Asn Ser Asp Leu Tyr Trp Lys Lys Phe Glu Asp His Tyr His Phe
 450 455 460

Ser Cys Gln Phe Thr Ala Asp Leu Ile Ala Met Asn His Ala Asp Phe
 465 470 475 480

Ile Ile Thr Asn Thr Phe Gln Glu Ile Ala Gly Asn Lys Asp Thr Val
 485 490 495

Gly Gln Tyr Glu Ser His Met Ala Phe Thr Met Pro Gly Leu Tyr Arg
 500 505 510

Val Val His Gly Ile Asp Val Phe Asp Pro Lys Phe Asn Ile Val Ser
 515 520 525

Pro Gly Ala Asp Met Ser Ile Tyr Phe Pro Tyr Thr Glu Gln Gln Lys

530	535	540
Arg Leu Thr Ser Leu His Thr Glu Ile Glu Glu Leu Leu Phe Ser Asp		
545	550	555 560
Val Glu Asn Ala Glu His Lys Phe Val Leu Lys Asp Lys Lys Lys Pro		
	565	570 575
Ile Ile Phe Ser Met Ala Arg Leu Asp Arg Val Lys Asn Met Thr Gly		
	580	585 590
Leu Val Glu Met Tyr Gly Arg Asn Pro Arg Leu Gln Glu Leu Val Asn		
	595	600 605
Leu Val Val Val Cys Gly Asp His Gly Lys Val Ser Lys Asp Lys Glu		
	610	615 620
Glu Gln Ala Glu Phe Lys Lys Met Phe Asp Leu Ile Glu Gln Tyr Asn		
	625	630 635 640
Leu Ile Gly His Ile Arg Trp Ile Ser Ala Gln Met Asn Arg Val Arg		
	645	650 655
Asn Gly Glu Leu Tyr Arg Tyr Ile Cys Asp Met Lys Gly Ala Phe Val		
	660	665 670
Gln Pro Ala Phe Tyr Glu Ala Phe Gly Leu Thr Val Ile Glu Ala Met		
	675	680 685
Thr Cys Gly Leu Pro Thr Phe Ala Thr Ala Tyr Gly Gly Pro Ala Glu		
	690	695 700
Ile Ile Val His Gly Val Ser Gly Tyr His Ile Asp Pro Tyr Gln Asn		
	705	710 715 720
Asp Lys Ala Ser Ala Leu Leu Val Asp Phe Phe Gly Lys Cys Gln Glu		
	725	730 735
Asp Pro Ser His Trp Asn Lys Ile Ser Gln Gly Gly Leu Gln Arg Ile		
	740	745 750
Glu Glu Lys Tyr Thr Trp Lys Leu Tyr Ser Glu Arg Leu Met Thr Leu		
	755	760 765
Ser Gly Val Tyr Gly Phe Trp Lys Tyr Val Ser Asn Leu Asp Arg Arg		
	770	775 780

Glu Thr Arg Arg Tyr Leu Glu Met Leu Tyr Ala Leu Lys Tyr Arg Lys
785 790 795 800

Met Ala Ala Thr Val Pro Leu Ala Val Glu Gly Glu Thr Ser Gly Lys
805 810 815

<210> 31
<211> 2894
<212> DNA
<213> Oryza sativa

<220>
<221> CDS
<222> (24)..(2450)

<400> 31
ccattgggttc ctgaaattga gtc atg gct gcc aag cta gct cgc ctc cac agt 53
Met Ala Ala Lys Leu Ala Arg Leu His Ser
1 5 10

ctc cgc gaa cgc ctc ggt gcc acc ttc tcg tct cat ccc aat gag ttg 101
Leu Arg Glu Arg Leu Gly Ala Thr Phe Ser Ser His Pro Asn Glu Leu
15 20 25

att gca ctc ttc tct agg tat gtt aac cag gga aag gga atg ctc cag 149
Ile Ala Leu Phe Ser Arg Tyr Val Asn Gln Gly Lys Gly Met Leu Gln
30 35 40

cgt cac cag ctg ctt gcg gag ttc gat gcc ttg atc gaa gct gac aaa 197
Arg His Gln Leu Leu Ala Glu Phe Asp Ala Leu Ile Glu Ala Asp Lys
45 50 55

gag aaa tat gct ccc ttt gaa gac att ctc cgg gct gct cag gaa gcc 245
Glu Lys Tyr Ala Pro Phe Glu Asp Ile Leu Arg Ala Ala Gln Glu Ala
60 65 70

att gtg ctg ccg ccc tgg gtt gca ctg gcc atc agg cca agg cct ggt 293
Ile Val Leu Pro Pro Trp Val Ala Leu Ala Ile Arg Pro Arg Pro Gly
75 80 85 90

gtc tgg gac tac att cgg gtg aat gta agt gag ttg gca gtg gaa gag 341
Val Trp Asp Tyr Ile Arg Val Asn Val Ser Glu Leu Ala Val Glu Glu
95 100 105

ctg agt gtt tct gag tac ttg gca ttc aag gaa cag ctt gtt gat gga 389
Leu Ser Val Ser Glu Tyr Leu Ala Phe Lys Glu Gln Leu Val Asp Gly
110 115 120

cac acc aac agc aac ttt gtt ctt gag ctt gat ttt gag ccc ttc aat 437
His Thr Asn Ser Asn Phe Val Leu Glu Leu Asp Phe Glu Pro Phe Asn
125 130 135

gcc tcc ttc ccg cgc ccg tcc atg tcc aag tcc atc gga aat ggg gtg 485
Ala Ser Phe Pro Arg Pro Ser Met Ser Lys Ser Ile Gly Asn Gly Val
140 145 150

cag ttc ctt aac cgt cac ctg tcg tcc aag ttg ttc cag gac aag gag 533

Gln	Phe	Leu	Asn	Arg	His	Leu	Ser	Ser	Lys	Leu	Phe	Gln	Asp	Lys	Glu	
155					160					165					170	
agc	ctc	tac	ccc	ttg	ctg	aac	ttc	ctg	aaa	gcc	cat	aac	cac	aag	ggc	581
Ser	Leu	Tyr	Pro	Leu	Leu	Asn	Phe	Leu	Lys	Ala	His	Asn	His	Lys	Gly	
				175					180					185		
acg	aca	atg	atg	ctg	aat	gac	aga	att	cag	agc	ctt	cgt	ggg	ctc	caa	629
Thr	Thr	Met	Met	Leu	Asn	Asp	Arg	Ile	Gln	Ser	Leu	Arg	Gly	Leu	Gln	
				190				195					200			
tca	tcc	ctt	aga	aag	gca	gaa	gaa	tat	ctg	atg	ggc	att	cct	caa	gac	677
Ser	Ser	Leu	Arg	Lys	Ala	Glu	Glu	Tyr	Leu	Met	Gly	Ile	Pro	Gln	Asp	
		205					210					215				
acg	ccc	tac	tcg	gag	ttc	aac	cac	agg	ttc	caa	gag	ctc	ggg	ttg	gag	725
Thr	Pro	Tyr	Ser	Glu	Phe	Asn	His	Arg	Phe	Gln	Glu	Leu	Gly	Leu	Glu	
	220					225					230					
aag	ggg	tgg	ggg	gac	tgt	gca	aag	cgt	gtg	ctt	gac	acc	atc	cac	ttg	773
Lys	Gly	Trp	Gly	Asp	Cys	Ala	Lys	Arg	Val	Leu	Asp	Thr	Ile	His	Leu	
235					240				245						250	
ctt	ctt	gac	ctt	ctt	gag	gcc	cct	gat	ccg	gcc	aac	ttg	gag	aag	ttc	821
Leu	Leu	Asp	Leu	Leu	Glu	Ala	Pro	Asp	Pro	Ala	Asn	Leu	Glu	Lys	Phe	
				255					260					265		
ctt	gga	act	att	cca	atg	atg	ttc	aat	gtt	gtt	atc	ctg	tct	ccg	cat	869
Leu	Gly	Thr	Ile	Pro	Met	Met	Phe	Asn	Val	Val	Ile	Leu	Ser	Pro	His	
			270					275					280			
gga	tac	ttt	gcc	caa	tcc	aat	gtg	ttg	gga	tac	cct	gat	act	ggg	ggg	917
Gly	Tyr	Phe	Ala	Gln	Ser	Asn	Val	Leu	Gly	Tyr	Pro	Asp	Thr	Gly	Gly	
		285					290					295				
cag	gtt	gtg	tac	att	ttg	gac	caa	gtc	cgc	gct	ttg	gag	aat	gag	atg	965
Gln	Val	Val	Tyr	Ile	Leu	Asp	Gln	Val	Arg	Ala	Leu	Glu	Asn	Glu	Met	
	300					305					310					
ctt	ttg	agg	atc	aag	cag	caa	ggc	ctt	gat	atc	aca	cct	aag	atc	ctc	1013
Leu	Leu	Arg	Ile	Lys	Gln	Gln	Gly	Leu	Asp	Ile	Thr	Pro	Lys	Ile	Leu	
315					320					325					330	
att	gta	acc	agg	ctg	ttg	cct	gat	gct	gtt	ggg	act	aca	tgc	ggc	cag	1061
Ile	Val	Thr	Arg	Leu	Leu	Pro	Asp	Ala	Val	Gly	Thr	Thr	Cys	Gly	Gln	
				335					340					345		
cgt	gtg	gag	aag	gtt	att	gga	act	gag	cac	act	gac	att	ctt	cgt	gtt	1109
Arg	Val	Glu	Lys	Val	Ile	Gly	Thr	Glu	His	Thr	Asp	Ile	Leu	Arg	Val	
			350					355					360			
cca	ttc	agg	agt	gag	aat	ggg	atc	ctc	cgc	aag	tgg	atc	tcc	cgt	ttt	1157
Pro	Phe	Arg	Ser	Glu	Asn	Gly	Ile	Leu	Arg	Lys	Trp	Ile	Ser	Arg	Phe	
		365					370					375				
gat	gtc	tgg	cca	ttc	ctg	gaa	aca	tac	act	gag	gat	gtt	gca	aac	gaa	1205
Asp	Val	Trp	Pro	Phe	Leu	Glu	Thr	Tyr	Thr	Glu	Asp	Val	Ala	Asn	Glu	
	380					385					390					
att	atg	agg	gaa	atg	caa	gcc	aaa	cct	gat	ctc	atc	att	ggc	aat	tac	1253
Ile	Met	Arg	Glu	Met	Gln	Ala	Lys	Pro	Asp	Leu	Ile	Ile	Gly	Asn	Tyr	

395	400	405	410	
agt gat gga aac ctt gtt gcc act ctg ctg gct cac aaa tta gga gtt				1301
Ser Asp Gly Asn Leu Val Ala Thr Leu Leu Ala His Lys Leu Gly Val				
	415	420	425	
acc cag tgt acc att gct cat gcc ttg gag aaa acc aaa tac ccc aac				1349
Thr Gln Cys Thr Ile Ala His Ala Leu Glu Lys Thr Lys Tyr Pro Asn				
	430	435	440	
tca gac ata tac ttg gac aag ttt gac agc cag tac cac ttc tca tgc				1397
Ser Asp Ile Tyr Leu Asp Lys Phe Asp Ser Gln Tyr His Phe Ser Cys				
	445	450	455	
caa ttc act gct gat ctt atc gcc atg aat cac act gat ttc atc atc				1445
Gln Phe Thr Ala Asp Leu Ile Ala Met Asn His Thr Asp Phe Ile Ile				
	460	465	470	
acc agt aca ttc caa gaa att gct gga agc aag gac act gtg ggg cag				1493
Thr Ser Thr Phe Gln Glu Ile Ala Gly Ser Lys Asp Thr Val Gly Gln				
	475	480	485	490
tat gaa tca cac att gca ttc acc ctt cct ggg ctt tac cga gtt gtg				1541
Tyr Glu Ser His Ile Ala Phe Thr Leu Pro Gly Leu Tyr Arg Val Val				
	495	500	505	
cat ggc ata gat gtt ttt gat ccc aag ttc aac att gtc tct cct gga				1589
His Gly Ile Asp Val Phe Asp Pro Lys Phe Asn Ile Val Ser Pro Gly				
	510	515	520	
gct gac atg agt gtc tac ttc ccg tac acc gag gct gac aag agg ctc				1637
Ala Asp Met Ser Val Tyr Phe Pro Tyr Thr Glu Ala Asp Lys Arg Leu				
	525	530	535	
act gct ttc cac cct gaa att gag gag ctt ctc tac agt gaa gtc gag				1685
Thr Ala Phe His Pro Glu Ile Glu Glu Leu Leu Tyr Ser Glu Val Glu				
	540	545	550	
aac gat gaa cac aag ttt gta ttg aag gac aag aac aag cca atc atc				1733
Asn Asp Glu His Lys Phe Val Leu Lys Asp Lys Asn Lys Pro Ile Ile				
	555	560	565	570
ttc tcc atg gct cgt ctt gac cga gtg aag aac atg aca ggt ctg gtt				1781
Phe Ser Met Ala Arg Leu Asp Arg Val Lys Asn Met Thr Gly Leu Val				
	575	580	585	
gag atg tat ggt aag aat gca cat ctc agg gat ttg gca aac ctt gtg				1829
Glu Met Tyr Gly Lys Asn Ala His Leu Arg Asp Leu Ala Asn Leu Val				
	590	595	600	
att gtt tgt ggt gac cac ggc aat cag tcc aag gac agg gag gag cag				1877
Ile Val Cys Gly Asp His Gly Asn Gln Ser Lys Asp Arg Glu Glu Gln				
	605	610	615	
gct gag ttc aag aag atg tac ggt ctc att gac cag tac aag ttg aag				1925
Ala Glu Phe Lys Lys Met Tyr Gly Leu Ile Asp Gln Tyr Lys Leu Lys				
	620	625	630	
ggg cat atc cgc tgg atc tca gct cag atg aac cgt gtt cgt aac ggg				1973
Gly His Ile Arg Trp Ile Ser Ala Gln Met Asn Arg Val Arg Asn Gly				
	635	640	645	650

gag ttg tac cga tac att tgt gac acc aag gga gtc ttt gtc cag cct 2021
 Glu Leu Tyr Arg Tyr Ile Cys Asp Thr Lys Gly Val Phe Val Gln Pro
 655 660 665

gca ttc tat gaa gcg ttt ggt ctg act gtc atc gaa gcc atg aca tgt 2069
 Ala Phe Tyr Glu Ala Phe Gly Leu Thr Val Ile Glu Ala Met Thr Cys
 670 675 680

ggt ttg cca aca atc gca aca tgc cat ggt ggc cct gct gag att att 2117
 Gly Leu Pro Thr Ile Ala Thr Cys His Gly Gly Pro Ala Glu Ile Ile
 685 690 695

gtt gat ggg gtg tct ggt ctg cac att gat cct tac cac agt gac aag 2165
 Val Asp Gly Val Ser Gly Leu His Ile Asp Pro Tyr His Ser Asp Lys
 700 705 710

gct gct gat atc ttg gtc aac ttc ttt gag aag tgc aag cag gat tca 2213
 Ala Ala Asp Ile Leu Val Asn Phe Phe Glu Lys Cys Lys Gln Asp Ser
 715 720 725 730

acc tac tgg gac aat att tca cag gga ggt ctg cag agg att tac gag 2261
 Thr Tyr Trp Asp Asn Ile Ser Gln Gly Gly Leu Gln Arg Ile Tyr Glu
 735 740 745

aag tac acc tgg aag ctg tac tct gag agg ctg atg acc ttg act ggt 2309
 Lys Tyr Thr Trp Lys Leu Tyr Ser Glu Arg Leu Met Thr Leu Thr Gly
 750 755 760

gta tac gga ttc tgg aag tac gta agc aac ctt gag agg cgc gag act 2357
 Val Tyr Gly Phe Trp Lys Tyr Val Ser Asn Leu Glu Arg Arg Glu Thr
 765 770 775

cgc cgt tac att gag atg ttc tat gct ctg aaa tac cgc agc ctg gcc 2405
 Arg Arg Tyr Ile Glu Met Phe Tyr Ala Leu Lys Tyr Arg Ser Leu Ala
 780 785 790

agc gcc gtc cca ttg gct gtc gat gga gag agc aca tcc aag taa 2450
 Ser Ala Val Pro Leu Ala Val Asp Gly Glu Ser Thr Ser Lys
 795 800 805

tggaggggaa aatatgcac ttcagcagga gaagccgtca gctgcattcg aatttgataa 2510

ttttctgtag ttgtcatttg gcatccatgt ttgccatgga tgtgtactat gtctaagggt 2570

tcagtacttt tgcgagattt tgggcagtgc ttgcctcaaa taaacgccgg ttcttggtgt 2630

ttatcgttca gaattcaatg caatgtttta ttgccgcgct tcctcaaaaa aaaaaaaaaa 2690

aaaaaaaca aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaag tgcctactg 2750

ttaaaaaagg ttggttcaaa gccatggatg aaccaacat ttttgttacc ttgttctttt 2810

cagaatttcc ttaagagttg tgatagtagg cccagtgcga ccaaatggtt gggttccttg 2870

agttttcagt ataaaatttc tgat 2894

<210> 32
 <211> 808
 <212> PRT

<213> Oryza sativa

<400> 32

Met Ala Ala Lys Leu Ala Arg Leu His Ser Leu Arg Glu Arg Leu Gly
 1 5 10 15

Ala Thr Phe Ser Ser His Pro Asn Glu Leu Ile Ala Leu Phe Ser Arg
 20 25 30

Tyr Val Asn Gln Gly Lys Gly Met Leu Gln Arg His Gln Leu Leu Ala
 35 40 45

Glu Phe Asp Ala Leu Ile Glu Ala Asp Lys Glu Lys Tyr Ala Pro Phe
 50 55 60

Glu Asp Ile Leu Arg Ala Ala Gln Glu Ala Ile Val Leu Pro Pro Trp
 65 70 75 80

Val Ala Leu Ala Ile Arg Pro Arg Pro Gly Val Trp Asp Tyr Ile Arg
 85 90 95

Val Asn Val Ser Glu Leu Ala Val Glu Glu Leu Ser Val Ser Glu Tyr
 100 105 110

Leu Ala Phe Lys Glu Gln Leu Val Asp Gly His Thr Asn Ser Asn Phe
 115 120 125

Val Leu Glu Leu Asp Phe Glu Pro Phe Asn Ala Ser Phe Pro Arg Pro
 130 135 140

Ser Met Ser Lys Ser Ile Gly Asn Gly Val Gln Phe Leu Asn Arg His
 145 150 155 160

Leu Ser Ser Lys Leu Phe Gln Asp Lys Glu Ser Leu Tyr Pro Leu Leu
 165 170 175

Asn Phe Leu Lys Ala His Asn His Lys Gly Thr Thr Met Met Leu Asn
 180 185 190

Asp Arg Ile Gln Ser Leu Arg Gly Leu Gln Ser Ser Leu Arg Lys Ala
 195 200 205

Glu Glu Tyr Leu Met Gly Ile Pro Gln Asp Thr Pro Tyr Ser Glu Phe
 210 215 220

Asn His Arg Phe Gln Glu Leu Gly Leu Glu Lys Gly Trp Gly Asp Cys

225	230	235	240
Ala Lys Arg Val	Leu Asp Thr Ile His	Leu Leu Leu Asp	Leu Leu Glu
	245	250	255
Ala Pro Asp	Pro Ala Asn Leu Glu	Lys Phe Leu Gly Thr	Ile Pro Met
	260	265	270
Met Phe Asn Val	Val Ile Leu Ser Pro His	Gly Tyr Phe Ala Gln	Ser
	275	280	285
Asn Val Leu Gly Tyr	Pro Asp Thr Gly Gly Gln	Val Val Tyr Ile Leu	
	290	295	300
Asp Gln Val Arg Ala	Leu Glu Asn Glu Met	Leu Leu Arg Ile Lys	Gln
305	310	315	320
Gln Gly Leu Asp	Ile Thr Pro Lys Ile Leu Ile	Val Thr Arg Leu Leu	
	325	330	335
Pro Asp Ala Val Gly Thr	Thr Cys Gly Gln Arg Val Glu	Lys Val Ile	
	340	345	350
Gly Thr Glu His Thr Asp	Ile Leu Arg Val Pro Phe Arg	Ser Glu Asn	
	355	360	365
Gly Ile Leu Arg Lys Trp	Ile Ser Arg Phe Asp	Val Trp Pro Phe Leu	
	370	375	380
Glu Thr Tyr Thr Glu Asp	Val Ala Asn Glu Ile Met Arg	Glu Met Gln	
385	390	395	400
Ala Lys Pro Asp Leu Ile	Ile Gly Asn Tyr Ser Asp Gly	Asn Leu Val	
	405	410	415
Ala Thr Leu Leu Ala His	Lys Leu Gly Val Thr Gln Cys	Thr Ile Ala	
	420	425	430
His Ala Leu Glu Lys Thr	Lys Tyr Pro Asn Ser Asp	Ile Tyr Leu Asp	
	435	440	445
Lys Phe Asp Ser Gln Tyr	His Phe Ser Cys Gln Phe Thr	Ala Asp Leu	
	450	455	460
Ile Ala Met Asn His Thr	Asp Phe Ile Ile Thr Ser Thr	Phe Gln Glu	
465	470	475	480

Ile Ala Gly Ser Lys Asp Thr Val Gly Gln Tyr Glu Ser His Ile Ala
 485 490 495

Phe Thr Leu Pro Gly Leu Tyr Arg Val Val His Gly Ile Asp Val Phe
 500 505 510

Asp Pro Lys Phe Asn Ile Val Ser Pro Gly Ala Asp Met Ser Val Tyr
 515 520 525

Phe Pro Tyr Thr Glu Ala Asp Lys Arg Leu Thr Ala Phe His Pro Glu
 530 535 540

Ile Glu Glu Leu Leu Tyr Ser Glu Val Glu Asn Asp Glu His Lys Phe
 545 550 555 560

Val Leu Lys Asp Lys Asn Lys Pro Ile Ile Phe Ser Met Ala Arg Leu
 565 570 575

Asp Arg Val Lys Asn Met Thr Gly Leu Val Glu Met Tyr Gly Lys Asn
 580 585 590

Ala His Leu Arg Asp Leu Ala Asn Leu Val Ile Val Cys Gly Asp His
 595 600 605

Gly Asn Gln Ser Lys Asp Arg Glu Glu Gln Ala Glu Phe Lys Lys Met
 610 615 620

Tyr Gly Leu Ile Asp Gln Tyr Lys Leu Lys Gly His Ile Arg Trp Ile
 625 630 635 640

Ser Ala Gln Met Asn Arg Val Arg Asn Gly Glu Leu Tyr Arg Tyr Ile
 645 650 655

Cys Asp Thr Lys Gly Val Phe Val Gln Pro Ala Phe Tyr Glu Ala Phe
 660 665 670

Gly Leu Thr Val Ile Glu Ala Met Thr Cys Gly Leu Pro Thr Ile Ala
 675 680 685

Thr Cys His Gly Gly Pro Ala Glu Ile Ile Val Asp Gly Val Ser Gly
 690 695 700

Leu His Ile Asp Pro Tyr His Ser Asp Lys Ala Ala Asp Ile Leu Val
 705 710 715 720

Asn Phe Phe Glu Lys Cys Lys Gln Asp Ser Thr Tyr Trp Asp Asn Ile
 725 730 735

Ser Gln Gly Gly Leu Gln Arg Ile Tyr Glu Lys Tyr Thr Trp Lys Leu
 740 745 750

Tyr Ser Glu Arg Leu Met Thr Leu Thr Gly Val Tyr Gly Phe Trp Lys
 755 760 765

Tyr Val Ser Asn Leu Glu Arg Arg Glu Thr Arg Arg Tyr Ile Glu Met
 770 775 780

Phe Tyr Ala Leu Lys Tyr Arg Ser Leu Ala Ser Ala Val Pro Leu Ala
 785 790 795 800

Val Asp Gly Glu Ser Thr Ser Lys
 805

<210> 33
 <211> 2424
 <212> DNA
 <213> Oryza sativa

<220>
 <221> CDS
 <222> (1)..(2424)

<400> 33
 atg gct gcc aag cta gct cgc ctc cac agt ctc cgc gaa cgc ctc ggt 48
 Met Ala Ala Lys Leu Ala Arg Leu His Ser Leu Arg Glu Arg Leu Gly
 1 5 10 15
 gcc acc ttc tgc tct cat ccc aat gag ttg att gca ctc ttc tct agg 96
 Ala Thr Phe Ser Ser His Pro Asn Glu Leu Ile Ala Leu Phe Ser Arg
 20 25 30
 tat gtt aac cag gga aag gga atg ctc cag cgt cac cag ctg ctt gcg 144
 Tyr Val Asn Gln Gly Lys Gly Met Leu Gln Arg His Gln Leu Leu Ala
 35 40 45
 gag ttc gat gcc ttg atc gaa gct gac aaa gag aaa tat gct ccc ttt 192
 Glu Phe Asp Ala Leu Ile Glu Ala Asp Lys Glu Lys Tyr Ala Pro Phe
 50 55 60
 gaa gac att ctc cgg gct gct cag gaa gcc att gtg ctg ccg ccc tgg 240
 Glu Asp Ile Leu Arg Ala Ala Gln Glu Ala Ile Val Leu Pro Pro Trp
 65 70 75 80
 gtt gca ctg gcc atc agg cca agg cct ggt gtc tgg gac tac att cgg 288
 Val Ala Leu Ala Ile Arg Pro Arg Pro Gly Val Trp Asp Tyr Ile Arg
 85 90 95
 gtg aat gta agt gag ttg gca gtg gaa gag ctg agt gtt tct gag tac 336

Val Asn Val Ser Glu Leu Ala Val Glu Glu Leu Ser Val Ser Glu Tyr	
100 105 110	
ttg gca ttc aag gaa cag ctt gtt gat gga cac acc aac agc aac ttt	384
Leu Ala Phe Lys Glu Gln Leu Val Asp Gly His Thr Asn Ser Asn Phe	
115 120 125	
gtt ctt gag ctt gat ttt gag ccc ttc aat gcc tcc ttc ccg cgc cgg	432
Val Leu Glu Leu Asp Phe Glu Pro Phe Asn Ala Ser Phe Pro Arg Pro	
130 135 140	
tcc atg tcc aag tcc atc gga aat ggg gtg cag ttc ctt aac cgt cac	480
Ser Met Ser Lys Ser Ile Gly Asn Gly Val Gln Phe Leu Asn Arg His	
145 150 155 160	
ctg tcg tcc aag ttg ttc cag gac aag gag agc ctc tac ccc ttg ctg	528
Leu Ser Ser Lys Leu Phe Gln Asp Lys Glu Ser Leu Tyr Pro Leu Leu	
165 170 175	
aac ttc ctg aaa gcc cat aac cac aag ggc acg aca atg atg ctg aat	576
Asn Phe Leu Lys Ala His Asn His Lys Gly Thr Thr Met Met Leu Asn	
180 185 190	
gac aga att cag agc ctt cgt ggg ctc caa tca tcc ctt aga aag gca	624
Asp Arg Ile Gln Ser Leu Arg Gly Leu Gln Ser Ser Leu Arg Lys Ala	
195 200 205	
gaa gaa tat ctg atg ggc att cct caa gac acg ccc tac tcg gag ttc	672
Glu Glu Tyr Leu Met Gly Ile Pro Gln Asp Thr Pro Tyr Ser Glu Phe	
210 215 220	
aac cac agg ttc caa gag ctc ggt ttg gag aag ggt tgg ggt gac tgt	720
Asn His Arg Phe Gln Glu Leu Gly Leu Glu Lys Gly Trp Gly Asp Cys	
225 230 235 240	
gca aag cgt gtg ctt gac acc atc cac ttg ctt ctt gac ctt ctt gag	768
Ala Lys Arg Val Leu Asp Thr Ile His Leu Leu Leu Asp Leu Leu Glu	
245 250 255	
gcc cct gat ccg gcc aac ttg gag aag ttc ctt gga act att cca atg	816
Ala Pro Asp Pro Ala Asn Leu Glu Lys Phe Leu Gly Thr Ile Pro Met	
260 265 270	
atg ttc aat gtt gtt atc ctg tct ccg cat gga tac ttt gcc caa tcc	864
Met Phe Asn Val Val Ile Leu Ser Pro His Gly Tyr Phe Ala Gln Ser	
275 280 285	
aat gtg ttg gga tac cct gat act ggt ggt cag gtt gtg tac att ttg	912
Asn Val Leu Gly Tyr Pro Asp Thr Gly Gly Gln Val Val Tyr Ile Leu	
290 295 300	
gac caa gtc cgc gct ttg gag aat gag atg ctt ttg agg atc aag cag	960
Asp Gln Val Arg Ala Leu Glu Asn Glu Met Leu Leu Arg Ile Lys Gln	
305 310 315 320	
caa ggc ctt gat atc aca cct aag atc ctc att gta acc agg ctg ttg	1008
Gln Gly Leu Asp Ile Thr Pro Lys Ile Leu Ile Val Thr Arg Leu Leu	
325 330 335	
cct gat gct gtt ggt act aca tgc ggc cag cgt gtg gag aag gtt att	1056
Pro Asp Ala Val Gly Thr Thr Cys Gly Gln Arg Val Glu Lys Val Ile	

340	345	350	
gga act gag cac act gac att ctt cgt gtt cca ttc agg agt gag aat			1104
Gly Thr Glu His Thr Asp Ile Leu Arg Val Pro Phe Arg Ser Glu Asn			
355	360	365	
ggt atc ctc cgc aag tgg atc tcc cgt ttt gat gtc tgg cca ttc ctg			1152
Gly Ile Leu Arg Lys Trp Ile Ser Arg Phe Asp Val Trp Pro Phe Leu			
370	375	380	
gaa aca tac act gag gat gtt gca aac gaa att atg agg gaa atg caa			1200
Glu Thr Tyr Thr Glu Asp Val Ala Asn Glu Ile Met Arg Glu Met Gln			
385	390	395	400
gcc aaa cct gat ctc atc att ggc aat tac agt gat gga aac ctt gtt			1248
Ala Lys Pro Asp Leu Ile Ile Gly Asn Tyr Ser Asp Gly Asn Leu Val			
405	410	415	
gcc act ctg ctg gct cac aaa tta gga gtt acc cag tgt acc att gct			1296
Ala Thr Leu Leu Ala His Lys Leu Gly Val Thr Gln Cys Thr Ile Ala			
420	425	430	
cat gcc ttg gag aaa acc aaa tac ccc aac tca gac ata tac ttg gac			1344
His Ala Leu Glu Lys Thr Lys Tyr Pro Asn Ser Asp Ile Tyr Leu Asp			
435	440	445	
aag ttt gac agc cag tac cac ttc tca tgc caa ttc act gct gat ctt			1392
Lys Phe Asp Ser Gln Tyr His Phe Ser Cys Gln Phe Thr Ala Asp Leu			
450	455	460	
atc gcc atg aat cac act gat ttc atc atc acc agt aca ttc caa gaa			1440
Ile Ala Met Asn His Thr Asp Phe Ile Ile Thr Ser Thr Phe Gln Glu			
465	470	475	480
att gct gga agc aag gac act gtg ggg cag tat gaa tca cac att gca			1488
Ile Ala Gly Ser Lys Asp Thr Val Gly Gln Tyr Glu Ser His Ile Ala			
485	490	495	
ttc acc ctt cct ggg ctt tac cga gtt gtg cat ggc ata gat gtt ttt			1536
Phe Thr Leu Pro Gly Leu Tyr Arg Val Val His Gly Ile Asp Val Phe			
500	505	510	
gat ccc aag ttc aac att gtc tct cct gga gct gac atg agt gtc tac			1584
Asp Pro Lys Phe Asn Ile Val Ser Pro Gly Ala Asp Met Ser Val Tyr			
515	520	525	
ttc ccg tac acc gag gct gac aag agg ctc act gct ttc cac cct gaa			1632
Phe Pro Tyr Thr Glu Ala Asp Lys Arg Leu Thr Ala Phe His Pro Glu			
530	535	540	
att gag gag ctt ctc tac agt gaa gtc gag aac gat gaa cac aag ttt			1680
Ile Glu Glu Leu Leu Tyr Ser Glu Val Glu Asn Asp Glu His Lys Phe			
545	550	555	560
gta ttg aag gac aag aac aag cca atc atc ttc tcc atg gct cgt ctt			1728
Val Leu Lys Asp Lys Asn Lys Pro Ile Ile Phe Ser Met Ala Arg Leu			
565	570	575	
gac cga gtg aag aac atg aca ggt ctg gtt gag atg tat ggt aag aat			1776
Asp Arg Val Lys Asn Met Thr Gly Leu Val Glu Met Tyr Gly Lys Asn			
580	585	590	

gca cat ctc agg gat ttg gca aac ctt gtg att gtt tgt ggt gac cac 1824
Ala His Leu Arg Asp Leu Ala Asn Leu Val Ile Val Cys Gly Asp His
595 600 605

ggc aat cag tcc aag gac agg gag gag cag gct gag ttc aag aag atg 1872
Gly Asn Gln Ser Lys Asp Arg Glu Glu Gln Ala Glu Phe Lys Lys Met
610 615 620

tac ggt ctc att gac cag tac aag ttg aag ggg cat atc cgc tgg atc 1920
Tyr Gly Leu Ile Asp Gln Tyr Lys Leu Lys Gly His Ile Arg Trp Ile
625 630 635 640

tca gct cag atg aac cgt gtt cgt aac ggg gag ttg tac cga tac att 1968
Ser Ala Gln Met Asn Arg Val Arg Asn Gly Glu Leu Tyr Arg Tyr Ile
645 650 655

tgt gac acc aag gga gtc ttt gtc cag cct gca ttc tat gaa gcg ttt 2016
Cys Asp Thr Lys Gly Val Phe Val Gln Pro Ala Phe Tyr Glu Ala Phe
660 665 670

ggc ctg act gtc atc gaa gcc atg aca tgt ggt ttg cca aca atc gca 2064
Gly Leu Thr Val Ile Glu Ala Met Thr Cys Gly Leu Pro Thr Ile Ala
675 680 685

aca tgc cat ggt ggc cct gct gag att att gtt gat ggg gtg tct ggt 2112
Thr Cys His Gly Gly Pro Ala Glu Ile Ile Val Asp Gly Val Ser Gly
690 695 700

ctg cac att gat cct tac cac agt gac aag gct gct gat atc ttg gtc 2160
Leu His Ile Asp Pro Tyr His Ser Asp Lys Ala Ala Asp Ile Leu Val
705 710 715 720

aac ttc ttt gag aag tgc aag cag gat tca acc tac tgg gac aat att 2208
Asn Phe Phe Glu Lys Cys Lys Gln Asp Ser Thr Tyr Trp Asp Asn Ile
725 730 735

tca cag gga ggt ctg cag agg att tac gag aag tac acc tgg aag ctg 2256
Ser Gln Gly Gly Leu Gln Arg Ile Tyr Glu Lys Tyr Thr Trp Lys Leu
740 745 750

tac tct gag agg ctg atg acc ttg act ggt gta tac gga ttc tgg aag 2304
Tyr Ser Glu Arg Leu Met Thr Leu Thr Gly Val Tyr Gly Phe Trp Lys
755 760 765

tac gta agc aac ctt gag agg cgc gag act cgc cgt tac att gag atg 2352
Tyr Val Ser Asn Leu Glu Arg Arg Glu Thr Arg Arg Tyr Ile Glu Met
770 775 780

ttc tat gct ctg aaa tac cgc agc ctg gcc agc gcc gtc cca ttg gct 2400
Phe Tyr Ala Leu Lys Tyr Arg Ser Leu Ala Ser Ala Val Pro Leu Ala
785 790 795 800

gtc gat gga gag agc aca tcc aag 2424
Val Asp Gly Glu Ser Thr Ser Lys
805

<210> 34
<211> 808
<212> PRT

<213> Oryza sativa

<400> 34

Met Ala Ala Lys Leu Ala Arg Leu His Ser Leu Arg Glu Arg Leu Gly
 1 5 10 15

Ala Thr Phe Ser Ser His Pro Asn Glu Leu Ile Ala Leu Phe Ser Arg
 20 25 30

Tyr Val Asn Gln Gly Lys Gly Met Leu Gln Arg His Gln Leu Leu Ala
 35 40 45

Glu Phe Asp Ala Leu Ile Glu Ala Asp Lys Glu Lys Tyr Ala Pro Phe
 50 55 60

Glu Asp Ile Leu Arg Ala Ala Gln Glu Ala Ile Val Leu Pro Pro Trp
 65 70 75 80

Val Ala Leu Ala Ile Arg Pro Arg Pro Gly Val Trp Asp Tyr Ile Arg
 85 90 95

Val Asn Val Ser Glu Leu Ala Val Glu Glu Leu Ser Val Ser Glu Tyr
 100 105 110

Leu Ala Phe Lys Glu Gln Leu Val Asp Gly His Thr Asn Ser Asn Phe
 115 120 125

Val Leu Glu Leu Asp Phe Glu Pro Phe Asn Ala Ser Phe Pro Arg Pro
 130 135 140

Ser Met Ser Lys Ser Ile Gly Asn Gly Val Gln Phe Leu Asn Arg His
 145 150 155 160

Leu Ser Ser Lys Leu Phe Gln Asp Lys Glu Ser Leu Tyr Pro Leu Leu
 165 170 175

Asn Phe Leu Lys Ala His Asn His Lys Gly Thr Thr Met Met Leu Asn
 180 185 190

Asp Arg Ile Gln Ser Leu Arg Gly Leu Gln Ser Ser Leu Arg Lys Ala
 195 200 205

Glu Glu Tyr Leu Met Gly Ile Pro Gln Asp Thr Pro Tyr Ser Glu Phe
 210 215 220

Asn His Arg Phe Gln Glu Leu Gly Leu Glu Lys Gly Trp Gly Asp Cys

225	230	235	240
Ala Lys Arg Val	Leu Asp Thr Ile His	Leu Leu Leu Asp	Leu Leu Glu
	245	250	255
Ala Pro Asp Pro	Ala Asn Leu Glu Lys	Phe Leu Gly Thr	Ile Pro Met
	260	265	270
Met Phe Asn Val	Val Ile Leu Ser Pro	His Gly Tyr Phe	Ala Gln Ser
	275	280	285
Asn Val Leu Gly	Tyr Pro Asp Thr	Gly Gly Gln Val	Val Tyr Ile Leu
	290	295	300
Asp Gln Val Arg	Ala Leu Glu Asn	Glu Met Leu Leu	Arg Ile Lys Gln
	305	310	315
Gln Gly Leu Asp	Ile Thr Pro Lys	Ile Leu Ile Val	Thr Arg Leu Leu
	325	330	335
Pro Asp Ala Val	Gly Thr Thr Cys	Gly Gln Arg Val	Glu Lys Val Ile
	340	345	350
Gly Thr Glu His	Thr Asp Ile Leu	Arg Val Pro Phe	Arg Ser Glu Asn
	355	360	365
Gly Ile Leu Arg	Lys Trp Ile Ser	Arg Phe Asp Val	Trp Pro Phe Leu
	370	375	380
Glu Thr Tyr Thr	Glu Asp Val Ala	Asn Glu Ile Met	Arg Glu Met Gln
	385	390	395
Ala Lys Pro Asp	Leu Ile Ile Gly	Asn Tyr Ser Asp	Gly Asn Leu Val
	405	410	415
Ala Thr Leu Leu	Ala His Lys Leu	Gly Val Thr Gln	Cys Thr Ile Ala
	420	425	430
His Ala Leu Glu	Lys Thr Lys Tyr	Pro Asn Ser Asp	Ile Tyr Leu Asp
	435	440	445
Lys Phe Asp Ser	Gln Tyr His Phe	Ser Cys Gln Phe	Thr Ala Asp Leu
	450	455	460
Ile Ala Met Asn	His Thr Asp Phe	Ile Ile Thr Ser	Thr Phe Gln Glu
	465	470	475
			480

Ile Ala Gly Ser Lys Asp Thr Val Gly Gln Tyr Glu Ser His Ile Ala
 485 490 495

Phe Thr Leu Pro Gly Leu Tyr Arg Val Val His Gly Ile Asp Val Phe
 500 505 510

Asp Pro Lys Phe Asn Ile Val Ser Pro Gly Ala Asp Met Ser Val Tyr
 515 520 525

Phe Pro Tyr Thr Glu Ala Asp Lys Arg Leu Thr Ala Phe His Pro Glu
 530 535 540

Ile Glu Glu Leu Leu Tyr Ser Glu Val Glu Asn Asp Glu His Lys Phe
 545 550 555 560

Val Leu Lys Asp Lys Asn Lys Pro Ile Ile Phe Ser Met Ala Arg Leu
 565 570 575

Asp Arg Val Lys Asn Met Thr Gly Leu Val Glu Met Tyr Gly Lys Asn
 580 585 590

Ala His Leu Arg Asp Leu Ala Asn Leu Val Ile Val Cys Gly Asp His
 595 600 605

Gly Asn Gln Ser Lys Asp Arg Glu Glu Gln Ala Glu Phe Lys Lys Met
 610 615 620

Tyr Gly Leu Ile Asp Gln Tyr Lys Leu Lys Gly His Ile Arg Trp Ile
 625 630 635 640

Ser Ala Gln Met Asn Arg Val Arg Asn Gly Glu Leu Tyr Arg Tyr Ile
 645 650 655

Cys Asp Thr Lys Gly Val Phe Val Gln Pro Ala Phe Tyr Glu Ala Phe
 660 665 670

Gly Leu Thr Val Ile Glu Ala Met Thr Cys Gly Leu Pro Thr Ile Ala
 675 680 685

Thr Cys His Gly Gly Pro Ala Glu Ile Ile Val Asp Gly Val Ser Gly
 690 695 700

Leu His Ile Asp Pro Tyr His Ser Asp Lys Ala Ala Asp Ile Leu Val
 705 710 715 720

123

gga aag ctt ctt cag aat tgc cag gaa gcc atg gtg tgc cca cct tgg	475
Gly Lys Leu Leu Gln Asn Cys Gln Glu Ala Met Val Ser Pro Pro Trp	
70 75 80	
gtt gga ttt gcc gtg cgt cca agg cct gga atc tgg gag tat gtg cgg	523
Val Gly Phe Ala Val Arg Pro Arg Pro Gly Ile Trp Glu Tyr Val Arg	
85 90 95	
atc aat gtg gag gag ctc att gtg gaa gag ctt agc gtg tcc gaa tac	571
Ile Asn Val Glu Glu Leu Ile Val Glu Glu Leu Ser Val Ser Glu Tyr	
100 105 110 115	
ttg ggt ttc aaa gaa cag ctt agc ctt ggc tct gat tca att gat cta	619
Leu Gly Phe Lys Glu Gln Leu Ser Leu Gly Ser Asp Ser Ile Asp Leu	
120 125 130	
tat gtg ctg gag ttg gat ttt gag ccc ttc aat gcg cac ttt ccg cgc	667
Tyr Val Leu Glu Leu Asp Phe Glu Pro Phe Asn Ala His Phe Pro Arg	
135 140 145	
atg aca aga cca tca tcc att gga cat ggt gta cag ttt ttg aat cgc	715
Met Thr Arg Pro Ser Ser Ile Gly His Gly Val Gln Phe Leu Asn Arg	
150 155 160	
cac tta tct tct aaa ctc ttt caa aat cca gag agc atg gag cct ttg	763
His Leu Ser Ser Lys Leu Phe Gln Asn Pro Glu Ser Met Glu Pro Leu	
165 170 175	
ttt caa ttc ctt cgc ctc cat act tat cga gga gag acg ctg atg ctc	811
Phe Gln Phe Leu Arg Leu His Thr Tyr Arg Gly Glu Thr Leu Met Leu	
180 185 190 195	
aat gag cgg att gca act ttt tct agg ttt cgt cct cag tta gtt aga	859
Asn Glu Arg Ile Ala Thr Phe Ser Arg Phe Arg Pro Gln Leu Val Arg	
200 205 210	
gct gaa gag gcc cta tca aaa ctc cca gag gat act cca ttt tcc agc	907
Ala Glu Glu Ala Leu Ser Lys Leu Pro Glu Asp Thr Pro Phe Ser Ser	
215 220 225	
ttt gca cac agg tta caa gag tta ggt ttg gaa aaa ggt tgg gga aat	955
Phe Ala His Arg Leu Gln Glu Leu Gly Leu Glu Lys Gly Trp Gly Asn	
230 235 240	
aca gca ggg cgt gtg ctg caa acc ttg aag ttg tta cta gat ctg ctc	1003
Thr Ala Gly Arg Val Leu Gln Thr Leu Lys Leu Leu Leu Asp Leu Leu	
245 250 255	
cag gct cct gac cca gac aca cta gaa aag ttc tta gcc aga att cct	1051
Gln Ala Pro Asp Pro Asp Thr Leu Glu Lys Phe Leu Ala Arg Ile Pro	
260 265 270 275	
atg att ttc acc gtt tgc att gtt tct cct cat ggt tac ttt ggt caa	1099
Met Ile Phe Thr Val Cys Ile Val Ser Pro His Gly Tyr Phe Gly Gln	
280 285 290	
gca ggt gtt ctt ggc ttg cct gat act ggt ggt cag gtt gtt tac ata	1147
Ala Gly Val Leu Gly Leu Pro Asp Thr Gly Gly Gln Val Val Tyr Ile	
295 300 305	

ttg gat caa gtg aga gct ctg gaa aat caa atg ttg gaa aat ctt cag Leu Asp Gln Val Arg Ala Leu Glu Asn Gln Met Leu Glu Asn Leu Gln 310 315 320	1195
ctt caa ggg ttg gat ttc aag ccg caa att gtt att ctc aca cga ttg Leu Gln Gly Leu Asp Phe Lys Pro Gln Ile Val Ile Leu Thr Arg Leu 325 330 335	1243
att cct aat gct aat ggg act acc gtc aac cag cgt atc gaa aag gtt Ile Pro Asn Ala Asn Gly Thr Thr Val Asn Gln Arg Ile Glu Lys Val 340 345 350 355	1291
tca ggc act cag cat tcg agg att cta cgt gtt cct ttc caa cac gaa Ser Gly Thr Gln His Ser Arg Ile Leu Arg Val Pro Phe Gln His Glu 360 365 370	1339
ggc aac atc ctg aag aac tgg att tct cgt ttt gat gtt tac ccc tac Gly Asn Ile Leu Lys Asn Trp Ile Ser Arg Phe Asp Val Tyr Pro Tyr 375 380 385	1387
ttg gaa aac tat gct cag gac gcg gcc agg gaa gtc ctt ggg gag ctc Leu Glu Asn Tyr Ala Gln Asp Ala Ala Arg Glu Val Leu Gly Glu Leu 390 395 400	1435
caa ggg cgg cct gat ctg att att ggt aat tac agc gat ggt aac ctg Gln Gly Arg Pro Asp Leu Ile Ile Gly Asn Tyr Ser Asp Gly Asn Leu 405 410 415	1483
gtc gcg acg ctt ctt tct cac tac ctt gat gta act cag tgt att att Val Ala Thr Leu Leu Ser His Tyr Leu Asp Val Thr Gln Cys Ile Ile 420 425 430 435	1531
gca cat gcg ttg gag aag aca aag tat ccg gac tcg gac att tat tgg Ala His Ala Leu Glu Lys Thr Lys Tyr Pro Asp Ser Asp Ile Tyr Trp 440 445 450	1579
aag gac ttt gaa gag aag tat cac ttt tca tgc cag ttc act gct gat Lys Asp Phe Glu Glu Lys Tyr His Phe Ser Cys Gln Phe Thr Ala Asp 455 460 465	1627
ctc att gcg atg aat agt gcg gac ttc atc atc act agc acc tat caa Leu Ile Ala Met Asn Ser Ala Asp Phe Ile Ile Thr Ser Thr Tyr Gln 470 475 480	1675
gag att gcc gga agt gct gat acc gtg gga cag tat gaa agc cat cag Glu Ile Ala Gly Ser Ala Asp Thr Val Gly Gln Tyr Glu Ser His Gln 485 490 495	1723
gca ttc aca atg cca ggg ttg tat cga gtc gtg aat ggc ata gac gtc Ala Phe Thr Met Pro Gly Leu Tyr Arg Val Val Asn Gly Ile Asp Val 500 505 510 515	1771
ttt gat ccc aaa ttc aac att gtg tcc cca gga gca gat atg aac atc Phe Asp Pro Lys Phe Asn Ile Val Ser Pro Gly Ala Asp Met Asn Ile 520 525 530	1819
tac tat cca ttc gca gac aaa gaa cgt cgc ttg aca agt cta cag gaa Tyr Tyr Pro Phe Ala Asp Lys Glu Arg Arg Leu Thr Ser Leu Gln Glu 535 540 545	1867
tct att gag gag ctc ctt tat agc cct gaa cag acg gat gag cat att	1915

Ser Ile Glu Glu Leu Leu Tyr Ser Pro Glu Gln Thr Asp Glu His Ile
550 555 560

ggg ttg att gac aag gaa aag ccc att ctc ttt tcc atg gct cga ctc 1963
Gly Leu Ile Asp Lys Glu Lys Pro Ile Leu Phe Ser Met Ala Arg Leu
565 570 575

gac aga gta aag aac ctc aca ggg tta gta gag atg tac gga aag aac 2011
Asp Arg Val Lys Asn Leu Thr Gly Leu Val Glu Met Tyr Gly Lys Asn
580 585 590 595

cag aaa cta aaa gaa ttt gta cat ttg gta att gtt gga ggc gaa atc 2059
Gln Lys Leu Lys Glu Phe Val His Leu Val Ile Val Gly Gly Glu Ile
600 605 610

aat cct tca aag tcg aag gac cgg gaa gaa gtt aga gag att gag aag 2107
Asn Pro Ser Lys Ser Lys Asp Arg Glu Glu Val Arg Glu Ile Glu Lys
615 620 625

atg cac aat ctc atc aag cgg tac aag ctt gaa aat aac ttc aga tgg 2155
Met His Asn Leu Ile Lys Arg Tyr Lys Leu Glu Asn Asn Phe Arg Trp
630 635 640

atc cga tca cag aca aat aga att cga aat gga gag ttg tac cgt tac 2203
Ile Arg Ser Gln Thr Asn Arg Ile Arg Asn Gly Glu Leu Tyr Arg Tyr
645 650 655

att gca gat tct caa gga gct ttt gta cag cct gca ctt tat gaa ggt 2251
Ile Ala Asp Ser Gln Gly Ala Phe Val Gln Pro Ala Leu Tyr Glu Gly
660 665 670 675

ttc ggt ctg aca gtg gtc gaa gct atg aca agt gga ctt ccc acc ttt 2299
Phe Gly Leu Thr Val Val Glu Ala Met Thr Ser Gly Leu Pro Thr Phe
680 685 690

gca acg agt cat gga ggc cca gct gag att att gag cat ggg att tct 2347
Ala Thr Ser His Gly Gly Pro Ala Glu Ile Ile Glu His Gly Ile Ser
695 700 705

ggg tat cat atc gat cct tat tat cct gat gaa gct gct gaa cag atc 2395
Gly Tyr His Ile Asp Pro Tyr Tyr Pro Asp Glu Ala Ala Glu Gln Ile
710 715 720

gtt gct ttc ttc gag aaa tgt aaa aat gag cct gga ctt tgg aat aaa 2443
Val Ala Phe Phe Glu Lys Cys Lys Asn Glu Pro Gly Leu Trp Asn Lys
725 730 735

gtt tct gaa gct gga tta caa cgc ata tat tcc agc tac act tgg aag 2491
Val Ser Glu Ala Gly Leu Gln Arg Ile Tyr Ser Ser Tyr Thr Trp Lys
740 745 750 755

ata tat gcc gag cgg ctt atg aca ctg tct gcg gtg tac gga ttc tgg 2539
Ile Tyr Ala Glu Arg Leu Met Thr Leu Ser Ala Val Tyr Gly Phe Trp
760 765 770

aag tat gtt tca aaa ttg cac agg cag gaa gct cgg aga tat ttg gag 2587
Lys Tyr Val Ser Lys Leu His Arg Gln Glu Ala Arg Arg Tyr Leu Glu
775 780 785

atg ttt tac atc ttg aag ttc cgt gaa ttg gca aga aca gta ccg ctt 2635
Met Phe Tyr Ile Leu Lys Phe Arg Glu Leu Ala Arg Thr Val Pro Leu

790	795	800	
tca aag gat gat gag gat gtt cta gaa aaa gtt gag aag aaa gct cag			2683
Ser Lys Asp Asp Glu Asp Val Leu Glu Lys Val Glu Lys Lys Ala Gln			
805	810	815	
ttg ggt cct ggc gta ggg gca att gtt ggt gaa gca gca aca gca gtc			2731
Leu Gly Pro Gly Val Gly Ala Ile Val Gly Glu Ala Ala Thr Ala Val			
820	825	830	835
gaa gca cgg aaa gct gtt act ggt cat aca taa tcgcagtgtg attaaacagt			2784
Glu Ala Arg Lys Ala Val Thr Gly His Thr			
840	845		
tcgtcataaa agctatctgt atgcttttgca cagaaggatg taatatatgt tttgggggttt			2844
ggttgaagtg catgggtttga acgccaacat cacagttgat gacttagatc ttgtttgaaa			2904
catgttatag atttttaaadc atttaacttg tgaaacaata tgatacacca gcaattattg			2964
ttgaaagcaa aaaaaaaaaa aaaaaaaaaa agcggccgc			3003
<210> 36			
<211> 845			
<212> PRT			
<213> Physcomitrella patens			
<400> 36			
Met Ser Gln Pro Arg Pro Thr Leu Arg Arg Leu Thr Gly Leu Lys Glu			
1	5	10	15
Arg Val Glu Ser Ser Leu Gln Glu His Arg Asn Glu Leu Leu His Leu			
20	25	30	
Leu Gln Gly Tyr Val Val Gln Gly Arg Ser Ile Leu Gln Pro His His			
35	40	45	
Leu Gln Asp Gln Leu Ala Ala Val His Asp Ala Ala His Ile Gln Asp			
50	55	60	
Thr Ala Ile Gly Lys Leu Leu Gln Asn Cys Gln Glu Ala Met Val Ser			
65	70	75	80
Pro Pro Trp Val Gly Phe Ala Val Arg Pro Arg Pro Gly Ile Trp Glu			
85	90	95	
Tyr Val Arg Ile Asn Val Glu Glu Leu Ile Val Glu Glu Leu Ser Val			
100	105	110	
Ser Glu Tyr Leu Gly Phe Lys Glu Gln Leu Ser Leu Gly Ser Asp Ser			
115	120	125	

Ile Asp Leu Tyr Val Leu Glu Leu Asp Phe Glu Pro Phe Asn Ala His
 130 135 140

Phe Pro Arg Met Thr Arg Pro Ser Ser Ile Gly His Gly Val Gln Phe
 145 150 155 160

Leu Asn Arg His Leu Ser Ser Lys Leu Phe Gln Asn Pro Glu Ser Met
 165 170 175

Glu Pro Leu Phe Gln Phe Leu Arg Leu His Thr Tyr Arg Gly Glu Thr
 180 185 190

Leu Met Leu Asn Glu Arg Ile Ala Thr Phe Ser Arg Phe Arg Pro Gln
 195 200 205

Leu Val Arg Ala Glu Glu Ala Leu Ser Lys Leu Pro Glu Asp Thr Pro
 210 215 220

Phe Ser Ser Phe Ala His Arg Leu Gln Glu Leu Gly Leu Glu Lys Gly
 225 230 235 240

Trp Gly Asn Thr Ala Gly Arg Val Leu Gln Thr Leu Lys Leu Leu Leu
 245 250 255

Asp Leu Leu Gln Ala Pro Asp Pro Asp Thr Leu Glu Lys Phe Leu Ala
 260 265 270

Arg Ile Pro Met Ile Phe Thr Val Cys Ile Val Ser Pro His Gly Tyr
 275 280 285

Phe Gly Gln Ala Gly Val Leu Gly Leu Pro Asp Thr Gly Gly Gln Val
 290 295 300

Val Tyr Ile Leu Asp Gln Val Arg Ala Leu Glu Asn Gln Met Leu Glu
 305 310 315 320

Asn Leu Gln Leu Gln Gly Leu Asp Phe Lys Pro Gln Ile Val Ile Leu
 325 330 335

Thr Arg Leu Ile Pro Asn Ala Asn Gly Thr Thr Val Asn Gln Arg Ile
 340 345 350

Glu Lys Val Ser Gly Thr Gln His Ser Arg Ile Leu Arg Val Pro Phe
 355 360 365

Gln His Glu Gly Asn Ile Leu Lys Asn Trp Ile Ser Arg Phe Asp Val
 370 375 380

Tyr Pro Tyr Leu Glu Asn Tyr Ala Gln Asp Ala Ala Arg Glu Val Leu
 385 390 395 400

Gly Glu Leu Gln Gly Arg Pro Asp Leu Ile Ile Gly Asn Tyr Ser Asp
 405 410 415

Gly Asn Leu Val Ala Thr Leu Leu Ser His Tyr Leu Asp Val Thr Gln
 420 425 430

Cys Ile Ile Ala His Ala Leu Glu Lys Thr Lys Tyr Pro Asp Ser Asp
 435 440 445

Ile Tyr Trp Lys Asp Phe Glu Glu Lys Tyr His Phe Ser Cys Gln Phe
 450 455 460

Thr Ala Asp Leu Ile Ala Met Asn Ser Ala Asp Phe Ile Ile Thr Ser
 465 470 475 480

Thr Tyr Gln Glu Ile Ala Gly Ser Ala Asp Thr Val Gly Gln Tyr Glu
 485 490 495

Ser His Gln Ala Phe Thr Met Pro Gly Leu Tyr Arg Val Val Asn Gly
 500 505 510

Ile Asp Val Phe Asp Pro Lys Phe Asn Ile Val Ser Pro Gly Ala Asp
 515 520 525

Met Asn Ile Tyr Tyr Pro Phe Ala Asp Lys Glu Arg Arg Leu Thr Ser
 530 535 540

Leu Gln Glu Ser Ile Glu Glu Leu Leu Tyr Ser Pro Glu Gln Thr Asp
 545 550 555 560

Glu His Ile Gly Leu Ile Asp Lys Glu Lys Pro Ile Leu Phe Ser Met
 565 570 575

Ala Arg Leu Asp Arg Val Lys Asn Leu Thr Gly Leu Val Glu Met Tyr
 580 585 590

Gly Lys Asn Gln Lys Leu Lys Glu Phe Val His Leu Val Ile Val Gly
 595 600 605

Gly Glu Ile Asn Pro Ser Lys Ser Lys Asp Arg Glu Glu Val Arg Glu

610					615					620					
Ile	Glu	Lys	Met	His	Asn	Leu	Ile	Lys	Arg	Tyr	Lys	Leu	Glu	Asn	Asn
625					630					635					640
Phe	Arg	Trp	Ile	Arg	Ser	Gln	Thr	Asn	Arg	Ile	Arg	Asn	Gly	Glu	Leu
				645					650					655	
Tyr	Arg	Tyr	Ile	Ala	Asp	Ser	Gln	Gly	Ala	Phe	Val	Gln	Pro	Ala	Leu
			660					665					670		
Tyr	Glu	Gly	Phe	Gly	Leu	Thr	Val	Val	Glu	Ala	Met	Thr	Ser	Gly	Leu
		675					680					685			
Pro	Thr	Phe	Ala	Thr	Ser	His	Gly	Gly	Pro	Ala	Glu	Ile	Ile	Glu	His
	690					695					700				
Gly	Ile	Ser	Gly	Tyr	His	Ile	Asp	Pro	Tyr	Tyr	Pro	Asp	Glu	Ala	Ala
705					710					715					720
Glu	Gln	Ile	Val	Ala	Phe	Phe	Glu	Lys	Cys	Lys	Asn	Glu	Pro	Gly	Leu
				725					730					735	
Trp	Asn	Lys	Val	Ser	Glu	Ala	Gly	Leu	Gln	Arg	Ile	Tyr	Ser	Ser	Tyr
			740					745					750		
Thr	Trp	Lys	Ile	Tyr	Ala	Glu	Arg	Leu	Met	Thr	Leu	Ser	Ala	Val	Tyr
			755					760					765		
Gly	Phe	Trp	Lys	Tyr	Val	Ser	Lys	Leu	His	Arg	Gln	Glu	Ala	Arg	Arg
	770					775					780				
Tyr	Leu	Glu	Met	Phe	Tyr	Ile	Leu	Lys	Phe	Arg	Glu	Leu	Ala	Arg	Thr
785					790					795					800
Val	Pro	Leu	Ser	Lys	Asp	Asp	Glu	Asp	Val	Leu	Glu	Lys	Val	Glu	Lys
				805					810					815	
Lys	Ala	Gln	Leu	Gly	Pro	Gly	Val	Gly	Ala	Ile	Val	Gly	Glu	Ala	Ala
			820					825					830		
Thr	Ala	Val	Glu	Ala	Arg	Lys	Ala	Val	Thr	Gly	His	Thr			
			835				840					845			

<210> 37

<211> 2538
 <212> DNA
 <213> *Physcomitrella patens*

<220>
 <221> CDS
 <222> (1)..(2538)

<400> 37
 atg tcg cag cca cgg ccg acg ctt cgg cgt ctt acg ggc ctc aag gag 48
 Met Ser Gln Pro Arg Pro Thr Leu Arg Arg Leu Thr Gly Leu Lys Glu
 1 5 10 15
 cga gtg gag agc tct ttg caa gag cat cgc aat gag ctg ctc cac ctt 96
 Arg Val Glu Ser Ser Leu Gln Glu His Arg Asn Glu Leu Leu His Leu
 20 25 30
 ctg caa ggc tat gtg gta caa ggt cgt tca att ctt caa ccg cat cat 144
 Leu Gln Gly Tyr Val Val Gln Gly Arg Ser Ile Leu Gln Pro His His
 35 40 45
 ttg caa gat caa ctg gct gct gtg cac gat gct gcc cat att caa gac 192
 Leu Gln Asp Gln Leu Ala Ala Val His Asp Ala Ala His Ile Gln Asp
 50 55 60
 act gct att gga aag ctt ctt cag aat tgc cag gaa gcc atg gtg tcg 240
 Thr Ala Ile Gly Lys Leu Leu Gln Asn Cys Gln Glu Ala Met Val Ser
 65 70 75 80
 cca cct tgg gtt gga ttt gcc gtg cgt cca agg cct gga atc tgg gag 288
 Pro Pro Trp Val Gly Phe Ala Val Arg Pro Arg Pro Gly Ile Trp Glu
 85 90 95
 tat gtg cgg atc aat gtg gag gag ctc att gtg gaa gag ctt agc gtg 336
 Tyr Val Arg Ile Asn Val Glu Glu Leu Ile Val Glu Glu Leu Ser Val
 100 105 110
 tcc gaa tac ttg ggt ttc aaa gaa cag ctt agc ctt ggc tct gat tca 384
 Ser Glu Tyr Leu Gly Phe Lys Glu Gln Leu Ser Leu Gly Ser Asp Ser
 115 120 125
 att gat cta tat gtg ctg gag ttg gat ttt gag ccc ttc aat gcg cac 432
 Ile Asp Leu Tyr Val Leu Glu Leu Asp Phe Glu Pro Phe Asn Ala His
 130 135 140
 ttt ccg cgc atg aca aga cca tca tcc att gga cat ggt gta cag ttt 480
 Phe Pro Arg Met Thr Arg Pro Ser Ser Ile Gly His Gly Val Gln Phe
 145 150 155 160
 ttg aat cgc cac tta tct tct aaa ctc ttt caa aat cca gag agc atg 528
 Leu Asn Arg His Leu Ser Ser Lys Leu Phe Gln Asn Pro Glu Ser Met
 165 170 175
 gag cct ttg ttt caa ttc ctt cgc ctc cat act tat cga gga gag acg 576
 Glu Pro Leu Phe Gln Phe Leu Arg Leu His Thr Tyr Arg Gly Glu Thr
 180 185 190
 ctg atg ctc aat gag cgg att gca act ttt tct agg ttt cgt cct cag 624
 Leu Met Leu Asn Glu Arg Ile Ala Thr Phe Ser Arg Phe Arg Pro Gln
 195 200 205

tta gtt aga gct gaa gag gcc cta tca aaa ctc cca gag gat act cca Leu Val Arg Ala Glu Glu Ala Leu Ser Lys Leu Pro Glu Asp Thr Pro 210 215 220	672
ttt tcc agc ttt gca cac agg tta caa gag tta ggt ttg gaa aaa ggt Phe Ser Ser Phe Ala His Arg Leu Gln Glu Leu Gly Leu Glu Lys Gly 225 230 235 240	720
tgg gga aat aca gca ggg cgt gtg ctg caa acc ttg aag ttg tta cta Trp Gly Asn Thr Ala Gly Arg Val Leu Gln Thr Leu Lys Leu Leu Leu 245 250 255	768
gat ctg ctc cag gct cct gac cca gac aca cta gaa aag ttc tta gcc Asp Leu Leu Gln Ala Pro Asp Pro Asp Thr Leu Glu Lys Phe Leu Ala 260 265 270	816
aga att cct atg att ttc acc gtt tgc att gtt tct cct cat ggt tac Arg Ile Pro Met Ile Phe Thr Val Cys Ile Val Ser Pro His Gly Tyr 275 280 285	864
ttt ggt caa gca ggt gtt ctt ggc ttg cct gat act ggt ggt cag gtt Phe Gly Gln Ala Gly Val Leu Gly Leu Pro Asp Thr Gly Gly Gln Val 290 295 300	912
gtt tac ata ttg gat caa gtg aga gct ctg gaa aat caa atg ttg gaa Val Tyr Ile Leu Asp Gln Val Arg Ala Leu Glu Asn Gln Met Leu Glu 305 310 315 320	960
aat ctt cag ctt caa ggg ttg gat ttc aag ccg caa att gtt att ctc Asn Leu Gln Leu Gln Gly Leu Asp Phe Lys Pro Gln Ile Val Ile Leu 325 330 335	1008
aca cga ttg att cct aat gct aat ggg act acc gtc aac cag cgt atc Thr Arg Leu Ile Pro Asn Ala Asn Gly Thr Thr Val Asn Gln Arg Ile 340 345 350	1056
gaa aag gtt tca ggc act cag cat tcg agg att cta cgt gtt cct ttc Glu Lys Val Ser Gly Thr Gln His Ser Arg Ile Leu Arg Val Pro Phe 355 360 365	1104
caa cac gaa ggc aac atc ctg aag aac tgg att tct cgt ttt gat gtt Gln His Glu Gly Asn Ile Leu Lys Asn Trp Ile Ser Arg Phe Asp Val 370 375 380	1152
tac ccc tac ttg gaa aac tat gct cag gac gcg gcc agg gaa gtc ctt Tyr Pro Tyr Leu Glu Asn Tyr Ala Gln Asp Ala Ala Arg Glu Val Leu 385 390 395 400	1200
ggg gag ctc caa ggg cgg cct gat ctg att att ggt aat tac agc gat Gly Glu Leu Gln Gly Arg Pro Asp Leu Ile Ile Gly Asn Tyr Ser Asp 405 410 415	1248
ggt aac ctg gtc gcg acg ctt ctt tct cac tac ctt gat gta act cag Gly Asn Leu Val Ala Thr Leu Leu Ser His Tyr Leu Asp Val Thr Gln 420 425 430	1296
tgt att att gca cat gcg ttg gag aag aca aag tat ccg gac tcg gac Cys Ile Ile Ala His Ala Leu Glu Lys Thr Lys Tyr Pro Asp Ser Asp 435 440 445	1344
att tat tgg aag gac ttt gaa gag aag tat cac ttt tca tgc cag ttc	1392

Ile Tyr Trp Lys Asp Phe Glu Glu Lys Tyr His Phe Ser Cys Gln Phe	
450 455 460	
act gct gat ctc att gcg atg aat agt gcg gac ttc atc atc act agc	1440
Thr Ala Asp Leu Ile Ala Met Asn Ser Ala Asp Phe Ile Ile Thr Ser	
465 470 475 480	
acc tat caa gag att gcc gga agt gct gat acc gtg gga cag tat gaa	1488
Thr Tyr Gln Glu Ile Ala Gly Ser Ala Asp Thr Val Gly Gln Tyr Glu	
485 490 495	
agc cat cag gca ttc aca atg cca ggg ttg tat cga gtc gtg aat ggc	1536
Ser His Gln Ala Phe Thr Met Pro Gly Leu Tyr Arg Val Val Asn Gly	
500 505 510	
ata gac gtc ttt gat ccc aaa ttc aac att gtg tcc cca gga gca gat	1584
Ile Asp Val Phe Asp Pro Lys Phe Asn Ile Val Ser Pro Gly Ala Asp	
515 520 525	
atg aac atc tac tat cca ttc gca gac aaa gaa cgt cgc ttg aca agt	1632
Met Asn Ile Tyr Tyr Pro Phe Ala Asp Lys Glu Arg Arg Leu Thr Ser	
530 535 540	
cta cag gaa tct att gag gag ctc ctt tat agc cct gaa cag acg gat	1680
Leu Gln Glu Ser Ile Glu Glu Leu Leu Tyr Ser Pro Glu Gln Thr Asp	
545 550 555 560	
gag cat att ggg ttg att gac aag gaa aag ccc att ctc ttt tcc atg	1728
Glu His Ile Gly Leu Ile Asp Lys Glu Lys Pro Ile Leu Phe Ser Met	
565 570 575	
gct cga ctc gac aga gta aag aac ctc aca ggg tta gta gag atg tac	1776
Ala Arg Leu Asp Arg Val Lys Asn Leu Thr Gly Leu Val Glu Met Tyr	
580 585 590	
gga aag aac cag aaa cta aaa gaa ttt gta cat ttg gta att gtt gga	1824
Gly Lys Asn Gln Lys Leu Lys Glu Phe Val His Leu Val Ile Val Gly	
595 600 605	
ggc gaa atc aat cct tca aag tcg aag gac cgg gaa gaa gtt aga gag	1872
Gly Glu Ile Asn Pro Ser Lys Ser Lys Asp Arg Glu Glu Val Arg Glu	
610 615 620	
att gag aag atg cac aat ctc atc aag cgg tac aag ctt gaa aat aac	1920
Ile Glu Lys Met His Asn Leu Ile Lys Arg Tyr Lys Leu Glu Asn Asn	
625 630 635 640	
ttc aga tgg atc cga tca cag aca aat aga att cga aat gga gag ttg	1968
Phe Arg Trp Ile Arg Ser Gln Thr Asn Arg Ile Arg Asn Gly Glu Leu	
645 650 655	
tac cgt tac att gca gat tct caa gga gct ttt gta cag cct gca ctt	2016
Tyr Arg Tyr Ile Ala Asp Ser Gln Gly Ala Phe Val Gln Pro Ala Leu	
660 665 670	
tat gaa ggt ttc ggt ctg aca gtg gtc gaa gct atg aca agt gga ctt	2064
Tyr Glu Gly Phe Gly Leu Thr Val Val Glu Ala Met Thr Ser Gly Leu	
675 680 685	
ccc acc ttt gca acg agt cat gga ggc cca gct gag att att gag cat	2112
Pro Thr Phe Ala Thr Ser His Gly Gly Pro Ala Glu Ile Ile Glu His	

690	695	700	
ggg att tct ggg tat cat atc gat cct tat tat cct gat gaa gct gct			2160
Gly Ile Ser Gly Tyr His Ile Asp Pro Tyr Tyr Pro Asp Glu Ala Ala			
705	710	715	720
gaa cag atc gtt gct ttc ttc gag aaa tgt aaa aat gag cct gga ctt			2208
Glu Gln Ile Val Ala Phe Phe Glu Lys Cys Lys Asn Glu Pro Gly Leu			
725	730		735
tgg aat aaa gtt tct gaa gct gga tta caa cgc ata tat tcc agc tac			2256
Trp Asn Lys Val Ser Glu Ala Gly Leu Gln Arg Ile Tyr Ser Ser Tyr			
740	745		750
act tgg aag ata tat gcc gag cgg ctt atg aca ctg tct gcg gtg tac			2304
Thr Trp Lys Ile Tyr Ala Glu Arg Leu Met Thr Leu Ser Ala Val Tyr			
755	760		765
gga ttc tgg aag tat gtt tca aaa ttg cac agg cag gaa gct cgg aga			2352
Gly Phe Trp Lys Tyr Val Ser Lys Leu His Arg Gln Glu Ala Arg Arg			
770	775		780
tat ttg gag atg ttt tac atc ttg aag ttc cgt gaa ttg gca aga aca			2400
Tyr Leu Glu Met Phe Tyr Ile Leu Lys Phe Arg Glu Leu Ala Arg Thr			
785	790		800
gta ccg ctt tca aag gat gat gag gat gtt cta gaa aaa gtt gag aag			2448
Val Pro Leu Ser Lys Asp Asp Glu Asp Val Leu Glu Lys Val Glu Lys			
805	810		815
aaa gct cag ttg ggt cct ggc gta ggg gca att gtt ggt gaa gca gca			2496
Lys Ala Gln Leu Gly Pro Gly Val Gly Ala Ile Val Gly Glu Ala Ala			
820	825		830
aca gca gtc gaa gca cgg aaa gct gtt act ggt cat aca taa			2538
Thr Ala Val Glu Ala Arg Lys Ala Val Thr Gly His Thr			
835	840		845
<210> 38			
<211> 845			
<212> PRT			
<213> Physcomitrella patens			
<400> 38			
Met Ser Gln Pro Arg Pro Thr Leu Arg Arg Leu Thr Gly Leu Lys Glu			
1	5	10	15
Arg Val Glu Ser Ser Leu Gln Glu His Arg Asn Glu Leu Leu His Leu			
20	25		30
Leu Gln Gly Tyr Val Val Gln Gly Arg Ser Ile Leu Gln Pro His His			
35	40		45
Leu Gln Asp Gln Leu Ala Ala Val His Asp Ala Ala His Ile Gln Asp			
50	55		60

Thr Ala Ile Gly Lys Leu Leu Gln Asn Cys Gln Glu Ala Met Val Ser
 65 70 75 80

Pro Pro Trp Val Gly Phe Ala Val Arg Pro Arg Pro Gly Ile Trp Glu
 85 90 95

Tyr Val Arg Ile Asn Val Glu Glu Leu Ile Val Glu Glu Leu Ser Val
 100 105 110

Ser Glu Tyr Leu Gly Phe Lys Glu Gln Leu Ser Leu Gly Ser Asp Ser
 115 120 125

Ile Asp Leu Tyr Val Leu Glu Leu Asp Phe Glu Pro Phe Asn Ala His
 130 135 140

Phe Pro Arg Met Thr Arg Pro Ser Ser Ile Gly His Gly Val Gln Phe
 145 150 155 160

Leu Asn Arg His Leu Ser Ser Lys Leu Phe Gln Asn Pro Glu Ser Met
 165 170 175

Glu Pro Leu Phe Gln Phe Leu Arg Leu His Thr Tyr Arg Gly Glu Thr
 180 185 190

Leu Met Leu Asn Glu Arg Ile Ala Thr Phe Ser Arg Phe Arg Pro Gln
 195 200 205

Leu Val Arg Ala Glu Glu Ala Leu Ser Lys Leu Pro Glu Asp Thr Pro
 210 215 220

Phe Ser Ser Phe Ala His Arg Leu Gln Glu Leu Gly Leu Glu Lys Gly
 225 230 235 240

Trp Gly Asn Thr Ala Gly Arg Val Leu Gln Thr Leu Lys Leu Leu Leu
 245 250 255

Asp Leu Leu Gln Ala Pro Asp Pro Asp Thr Leu Glu Lys Phe Leu Ala
 260 265 270

Arg Ile Pro Met Ile Phe Thr Val Cys Ile Val Ser Pro His Gly Tyr
 275 280 285

Phe Gly Gln Ala Gly Val Leu Gly Leu Pro Asp Thr Gly Gly Gln Val
 290 295 300

Val Tyr Ile Leu Asp Gln Val Arg Ala Leu Glu Asn Gln Met Leu Glu

305 310 315 320
 Asn Leu Gln Leu Gln Gly Leu Asp Phe Lys Pro Gln Ile Val Ile Leu
 325 330 335
 Thr Arg Leu Ile Pro Asn Ala Asn Gly Thr Thr Val Asn Gln Arg Ile
 340 345 350
 Glu Lys Val Ser Gly Thr Gln His Ser Arg Ile Leu Arg Val Pro Phe
 355 360 365
 Gln His Glu Gly Asn Ile Leu Lys Asn Trp Ile Ser Arg Phe Asp Val
 370 375 380
 Tyr Pro Tyr Leu Glu Asn Tyr Ala Gln Asp Ala Ala Arg Glu Val Leu
 385 390 395 400
 Gly Glu Leu Gln Gly Arg Pro Asp Leu Ile Ile Gly Asn Tyr Ser Asp
 405 410 415
 Gly Asn Leu Val Ala Thr Leu Leu Ser His Tyr Leu Asp Val Thr Gln
 420 425 430
 Cys Ile Ile Ala His Ala Leu Glu Lys Thr Lys Tyr Pro Asp Ser Asp
 435 440 445
 Ile Tyr Trp Lys Asp Phe Glu Glu Lys Tyr His Phe Ser Cys Gln Phe
 450 455 460
 Thr Ala Asp Leu Ile Ala Met Asn Ser Ala Asp Phe Ile Ile Thr Ser
 465 470 475 480
 Thr Tyr Gln Glu Ile Ala Gly Ser Ala Asp Thr Val Gly Gln Tyr Glu
 485 490 495
 Ser His Gln Ala Phe Thr Met Pro Gly Leu Tyr Arg Val Val Asn Gly
 500 505 510
 Ile Asp Val Phe Asp Pro Lys Phe Asn Ile Val Ser Pro Gly Ala Asp
 515 520 525
 Met Asn Ile Tyr Tyr Pro Phe Ala Asp Lys Glu Arg Arg Leu Thr Ser
 530 535 540
 Leu Gln Glu Ser Ile Glu Glu Leu Leu Tyr Ser Pro Glu Gln Thr Asp

545	550	555	560
Glu His Ile Gly Leu Ile Asp Lys Glu Lys Pro Ile Leu Phe Ser Met			
	565	570	575
Ala Arg Leu Asp Arg Val Lys Asn Leu Thr Gly Leu Val Glu Met Tyr			
	580	585	590
Gly Lys Asn Gln Lys Leu Lys Glu Phe Val His Leu Val Ile Val Gly			
	595	600	605
Gly Glu Ile Asn Pro Ser Lys Ser Lys Asp Arg Glu Glu Val Arg Glu			
	610	615	620
Ile Glu Lys Met His Asn Leu Ile Lys Arg Tyr Lys Leu Glu Asn Asn			
	625	630	635
Phe Arg Trp Ile Arg Ser Gln Thr Asn Arg Ile Arg Asn Gly Glu Leu			
	645	650	655
Tyr Arg Tyr Ile Ala Asp Ser Gln Gly Ala Phe Val Gln Pro Ala Leu			
	660	665	670
Tyr Glu Gly Phe Gly Leu Thr Val Val Glu Ala Met Thr Ser Gly Leu			
	675	680	685
Pro Thr Phe Ala Thr Ser His Gly Gly Pro Ala Glu Ile Ile Glu His			
	690	695	700
Gly Ile Ser Gly Tyr His Ile Asp Pro Tyr Tyr Pro Asp Glu Ala Ala			
	705	710	715
Glu Gln Ile Val Ala Phe Phe Glu Lys Cys Lys Asn Glu Pro Gly Leu			
	725	730	735
Trp Asn Lys Val Ser Glu Ala Gly Leu Gln Arg Ile Tyr Ser Ser Tyr			
	740	745	750
Thr Trp Lys Ile Tyr Ala Glu Arg Leu Met Thr Leu Ser Ala Val Tyr			
	755	760	765
Gly Phe Trp Lys Tyr Val Ser Lys Leu His Arg Gln Glu Ala Arg Arg			
	770	775	780
Tyr Leu Glu Met Phe Tyr Ile Leu Lys Phe Arg Glu Leu Ala Arg Thr			
	785	790	795
			800

Val Pro Leu Ser Lys Asp Asp Glu Asp Val Leu Glu Lys Val Glu Lys
 805 810 815

Lys Ala Gln Leu Gly Pro Gly Val Gly Ala Ile Val Gly Glu Ala Ala
 820 825 830

Thr Ala Val Glu Ala Arg Lys Ala Val Thr Gly His Thr
 835 840 845

<210> 39
 <211> 11
 <212> PRT
 <213> Artificial sequence

<220>
 <221> Variant
 <222> (5)..(5)
 <223> Xaa in position 5 is Met, Gly, Pro, Ala, Val, Leu, Ile, Cys, Tyr,
 Trp, His, Lys, Arg, Gln, Asn, Glu, Asp, Thr or Phe

<220>
 <221> Variant
 <222> (9)..(9)
 <223> Xaa in position 9 is Met, Gly, Pro, Ala, Val, Leu, Ile, Cys, Tyr,
 Trp, His, Lys, Arg, Gln, Asn, Glu, Asp, Ser, Thr or Phe

<400> 39
 Tyr His Phe Ser Xaa Gln Phe Thr Xaa Asp Leu
 1 5 10

<210> 40
 <211> 10
 <212> PRT
 <213> Artificial sequence

<220>
 <221> Variant
 <222> (1)..(2)
 <223> Xaa in position 1 to 2 is Met, Gly, Pro, Ala, Val, Leu, Ile, Cys,
 Tyr, Trp, His, Lys, Arg, Gln, Glu, Asp, Ser, Thr or Phe

<400> 40
 Xaa Xaa Thr Val Gly Gln Tyr Glu Ser His
 1 5 10

<210> 41
 <211> 7
 <212> PRT
 <213> Artificial sequence

<220>
 <221> Variant
 <222> (1)..(1)

<223> Xaa in position 1 is Met, Gly, Pro, Val, Leu, Ile, Cys, Tyr, Trp, His, Lys, Arg, Gln, Asn, Glu, Asp, Ser, Thr or Phe

<220>

<221> Variant

<222> (5)..(5)

<223> Xaa in position 5 is Met, Gly, Pro, Ala, Val, Leu, Ile, Cys, Tyr, Trp, His, Lys, Arg, Gln, Asn, Glu, Asp, Ser, Thr or Phe

<400> 41

Xaa Leu Pro Thr Xaa Ala Thr
1 5

<210> 42

<211> 8

<212> PRT

<213> Artificial sequence

<220>

<221> Variant

<222> (3)..(3)

<223> Xaa in position 3 is Met, Gly, Pro, Val, Leu, Ile, Cys, Tyr, Trp, His, Lys, Arg, Gln, Asn, Glu, Asp, Ser, Thr or Phe

<400> 42

Val Tyr Xaa Phe Trp Lys His Val
1 5