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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
30 May 2013 (30.05.2013)

WIPO | PCT

(10) International Publication Number
WO 2013/076650 A1

(51) **International Patent Classification**:
C12N 15/29 (2006.01) *A01H 5/00* (2006.01)
C12N 15/82 (2006.01) *C12N 15/14* (2006.01)

(21) **International Application Number**:
PCT/IB2012/056570

(22) **International Filing Date**:
20 November 2012 (20.11.2012)

(25) **Filing Language**: English

(26) **Publication Language**: English

(30) **Priority Data**:
61/563,625 25 November 2011 (25.11.2011) US
11190757.2 25 November 2011 (25.11.2011) EP

(71) **Applicant** *(for all designated States except US)*: **BASF PLANT SCIENCE COMPANY GMBH** [DE/DE]; 67056 Ludwigshafen (DE).

(71) **Applicant** *(for MN only)*: **BASF (CHINA) COMPANY LIMITED** [CN/CN]; 300 Jiangxinsha Road, Shanghai, 200137 (CN).

(72) **Inventor; and**
(71) **Applicant** *(for US only)*: **Vandenabeele, Steven** [BE/BE]; Meersport 23, B-9700 Oudenaarde (BE).

(74) **Agent**: **Popp, Andreas**; Basf Se, Global Intellectual Property, GVX/B - C006, 67056 Ludwigshafen (DE).

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

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

Published:

— *with international search report (Art. 21(3))*

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

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

WO 2013/076650 A1

(54) **Title**: PLANTS HAVING ENHANCED YIELD-RELATED TRAITS AND METHOD FOR MAKING THE SAME

(57) **Abstract**: A method for enhancing yield-related traits in plants by modulating expression in a plant of a nucleic acid encoding a PRR-like (Pseudo Response Regulator-like) polypeptide is provided. Plants having modulated expression of a nucleic acid encoding a PRR-like polypeptide, which plants have enhanced yield-related traits relative to control plants, hitherto unknown PRR-like-encoding nucleic acids, and constructs comprising the same, which are useful in performing the methods are also provided.

PLANTS HAVING ENHANCED YIELD-RELATED TRAITS AND METHOD FOR MAKING THE SAME

Background

The present invention relates generally to the field of molecular biology and concerns a method for enhancing yield-related traits in plants by modulating expression in a plant of a nucleic acid encoding a PRR-like (Pseudo Response Regulator-like) polypeptide. The present invention also concerns plants having modulated expression of a nucleic acid encoding a PRR-like polypeptide, which plants have enhanced yield-related traits relative to corresponding wild type plants or other control plants. The invention also provides constructs useful in the methods of the invention.

The ever-increasing world population and the dwindling supply of arable land available for agriculture fuels research towards increasing the efficiency of agriculture. Conventional means for crop and horticultural improvements utilise selective breeding techniques to identify plants having desirable characteristics. However, such selective breeding techniques have several drawbacks, namely that these techniques are typically labour intensive and result in plants that often contain heterogeneous genetic components that may not always result in the desirable trait being passed on from parent plants. Advances in molecular biology have allowed mankind to modify the germplasm of animals and plants. Genetic engineering of plants entails the isolation and manipulation of genetic material (typically in the form of DNA or RNA) and the subsequent introduction of that genetic material into a plant. Such technology has the capacity to deliver crops or plants having various improved economic, agronomic or horticultural traits.

A trait of particular economic interest is increased yield. Yield is normally defined as the measurable produce of economic value from a crop. This may be defined in terms of quantity and/or quality. Yield is directly dependent on several factors, for example, the number and size of the organs, plant architecture (for example, the number of branches), seed production, leaf senescence and more. Root development, nutrient uptake, stress tolerance and early vigour may also be important factors in determining yield. Optimizing the abovementioned factors may therefore contribute to increasing crop yield.

Seed yield is a particularly important trait, since the seeds of many plants are important for human and animal nutrition. Crops such as corn, rice, wheat, canola and soybean account for over half the total human caloric intake, whether through direct consumption of the seeds themselves or through consumption of meat products raised on processed seeds. They are also a source of sugars, oils and many kinds of metabolites used in industrial processes. Seeds contain an embryo (the source of new shoots and roots) and an endosperm (the source of nutrients for embryo growth during germination and during early growth of seedlings). The development of a seed involves many genes, and requires the transfer of metabolites from the roots, leaves and stems into the growing seed. The

endosperm, in particular, assimilates the metabolic precursors of carbohydrates, oils and proteins and synthesizes them into storage macromolecules to fill out the grain.

Another important trait for many crops is early vigour. Improving early vigour is an important objective of modern rice breeding programs in both temperate and tropical rice cultivars. Long roots are important for proper soil anchorage in water-seeded rice. Where rice is sown directly into flooded fields, and where plants must emerge rapidly through water, longer shoots are associated with vigour. Where drill-seeding is practiced, longer mesocotyls and coleoptiles are important for good seedling emergence. The ability to engineer early vigour into plants would be of great importance in agriculture. For example, poor early vigour has been a limitation to the introduction of maize (Zea mays L.) hybrids based on Corn Belt germplasm in the European Atlantic.

A further important trait is that of improved abiotic stress tolerance. Abiotic stress is a primary cause of crop loss worldwide, reducing average yields for most major crop plants by more than 50% (Wang et al., Planta 218, 1-14, 2003). Abiotic stresses may be caused by drought, salinity, extremes of temperature, chemical toxicity and oxidative stress. The ability to improve plant tolerance to abiotic stress would be of great economic advantage to farmers worldwide and would allow for the cultivation of crops during adverse conditions and in territories where cultivation of crops may not otherwise be possible.

Crop yield may therefore be increased by optimising one of the above-mentioned factors.

Pseudo Response Regulators (PRRs) share a conserved domain, the receiver –like domain (RLD), along with another domain the CONSTANS/CONSTANS-like/TOC1 (CCT) domain. The RLD is similar to the receiver domain of the Response Regulators in the histidine to aspartic acid (His-Asp) phosphorelay, a versatile signal transduction system in organisms from bacteria to eukaryotes other than animals.

In the model dicot Arabidopsis thaliana, a representative set of such component genes is the pseudo-response regulator (PRR) gene family, which comprises five member genes, TOC1 (also called PRR1)/PRR3/PRR5/PRR7/PRR9, which play regulatory roles at multiple nodes in the interlocked loops of the A. thaliana circadian network.

Depending on the end use, the modification of certain yield traits may be favoured over others. For example for applications such as forage or wood production, or bio-fuel resource, an increase in the vegetative parts of a plant may be desirable, and for applications such as flour, starch or oil production, an increase in seed parameters may be particularly desirable. Even amongst the seed parameters, some may be favoured over others, depending on the application. Various mechanisms may contribute to increasing seed yield, whether that is in the form of increased seed size or increased seed number.

It has now been found that various yield-related traits may be improved in plants by modulating expression in a plant of a nucleic acid encoding a PRR-like (Pseudo Response Regulator-like) polypeptide in a plant.

Detailed description of the invention

The present invention shows that modulating expression in a plant of a nucleic acid encoding a PRR-like polypeptide gives plants having enhanced yield-related traits relative to control plants.

According to a first embodiment, the present invention provides a method for enhancing yield-related traits in plants relative to control plants, comprising modulating expression in a plant of a nucleic acid encoding a PRR-like polypeptide and optionally selecting for plants having enhanced yield-related traits. According to another embodiment, the present invention provides a method for producing plants having enhanced yield-related traits relative to control plants, wherein said method comprises the steps of modulating expression in said plant of a nucleic acid encoding a PRR-like polypeptide as described herein and optionally selecting for plants having enhanced yield-related traits.

A preferred method for modulating, preferably increasing, expression of a nucleic acid encoding a PRR-like polypeptide is by introducing and expressing in a plant a nucleic acid encoding a PRR-like polypeptide.

Any reference hereinafter to a "protein useful in the methods of the invention" is taken to mean a PRR-like polypeptide as defined herein. Any reference hereinafter to a "nucleic acid useful in the methods of the invention" is taken to mean a nucleic acid capable of encoding such a PRR-like polypeptide. The nucleic acid to be introduced into a plant (and therefore useful in performing the methods of the invention) is any nucleic acid encoding the type of protein which will now be described, hereafter also named "*PRR-like* nucleic acid" or "*PRR-like* gene".

A "PRR-like polypeptide" as defined herein refers to any polypeptide comprising an InterPro accession IPR001789 signal transduction response regulator, receiver domain corresponding to PFAM accession number PF00072 and/or an InterPro accession IPR010402 CCT domain corresponding to PFAM accession number PF06203.

According one embodiment, there is provided a method for improving yield-related traits as provided herein in plants relative to control plants, comprising modulating expression in a plant of a nucleic acid encoding a PRR-like polypeptide as defined herein.

Motifs 1 to 3 were derived using the MEME algorithm (Bailey and Elkan, Proceedings of the Second International Conference on Intelligent Systems for Molecular Biology, pp. 28-36, AAAI Press, Menlo Park, California, 1994). At each position within a MEME motif, the

residues are shown that are present in the query set of sequences with a frequency higher than 0.2. Residues within square brackets represent alternatives.

In one embodiment, the PRR-like polypeptide as used herein comprises at least one of the motifs 1, 2 or 3.

Motif 1 (SEQ ID NO: 43): MS[ST]NDSMSMVFKCLSKGAVDFLVKP[LI]RKNELKNLWQH[VI]WRRCHSSSGS[GE]S

Motif 2 (SEQ ID NO: 44): L[LV][VD][EP][NP]D[DS][SCT][TC][RA][QH]V[VI][SH][AP]L[LS][RE][KI]C[CS][YN][EK][VW][IL]P[AT][EA]N[GK][SLR][HN]A[WK][RKQ]Y[LK]E[ND][LK][QD][NE][NS][IM][DG][LR][VY]LT[EI]

Motif 3 (SEQ ID NO: 45): HDDEENDD[DG]DDDDFSVGLNARDGSDNGSGTQSSWTKRAVEIDSPQP[MI]SPD

In still another embodiment, the PRR-like polypeptide comprises in increasing order of preference, at least 2, or all 3 motifs as defined above.

Additionally or alternatively, the PRR-like protein has in increasing order of preference at least 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% overall sequence identity to the amino acid sequence represented by SEQ ID NO: 2, provided that the homologous protein comprises any one or more of the conserved motifs as outlined above. The overall sequence identity is determined using a global alignment algorithm, such as the Needleman Wunsch algorithm in the program GAP (GCG Wisconsin Package, Accelrys), preferably with default parameters and preferably with sequences of mature proteins (i.e. without taking into account secretion signals or transit peptides). In one embodiment the sequence identity level is determined by comparison of the polypeptide sequences over the entire length of the sequence of SEQ ID NO: 2.

In another embodiment, the sequence identity level is determined by comparison of one or more conserved domains or motifs in SEQ ID NO: 2 with corresponding conserved domains or motifs in other PRR-like polypeptides. Compared to overall sequence identity, the sequence identity will generally be higher when only conserved domains or motifs are considered. Preferably the motifs in a PRR-like polypeptide have, in increasing order of preference, at least 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to any one or more of the motifs represented by SEQ ID NO: 43 to SEQ ID NO: 45 (Motifs 1 to 3). In other words, in another embodiment a method for enhancing yield-related traits in plants is provided wherein said PRR-like polypeptide comprises a conserved domain (or motif) with at least 70%, 71%, 72%, 73%, 74%, 75%,

76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the conserved domain starting with amino acid 161 up to amino acid 211 and/or the conserved domain starting with amino acid 84 up to amino acid 133 and/or the conserved domain starting with amino acid 237 up to amino acid 286 in SEQ ID NO:2.

The terms “domain”, “signature” and “motif” are defined in the “definitions” section herein.

Nucleic acids encoding PRR-like polypeptides, when expressed in rice according to the methods of the present invention as outlined in Examples 7 and 9, give plants having increased yield related traits, in particular aboveground biomass, root biomass and seed yield, which included total weight of seeds, number of seeds, fillrate and number of flowers per panicle. Another function of the nucleic acid sequences encoding PRR-like polypeptides is to confer information for synthesis of the PRR-like protein that increases yield or yield related traits as described herein, when such a nucleic acid sequence of the invention is transcribed and translated in a living plant cell.

The present invention is illustrated by transforming plants with the nucleic acid sequence represented by SEQ ID NO: 1, encoding the polypeptide sequence of SEQ ID NO: 2. However, performance of the invention is not restricted to these sequences; the methods of the invention may advantageously be performed using any PRR-like-encoding nucleic acid or PRR-like polypeptide as defined herein. The term “PRR-like” or “PRR-like polypeptide” as used herein also intends to include homologues as defined hereunder of SEQ ID NO: 2.

Examples of nucleic acids encoding PRR-like polypeptides are given in Table A of the Examples section herein. Such nucleic acids are useful in performing the methods of the invention. The amino acid sequences given in Table A of the Examples section are example sequences of orthologues and paralogues of the PRR-like polypeptide represented by SEQ ID NO: 2, the terms “orthologues” and “paralogues” being as defined herein. Further orthologues and paralogues may readily be identified by performing a so-called reciprocal blast search as described in the definitions section; where the query sequence is SEQ ID NO: 1 or SEQ ID NO: 2, the second BLAST (back-BLAST) would be against *Saccharum officinarum* sequences.

The invention also provides hitherto unknown PRR-like-encoding nucleic acids and PRR-like polypeptides useful for conferring enhanced yield-related traits in plants relative to control plants.

According to a further embodiment of the present invention, there is therefore provided an isolated nucleic acid molecule selected from:

(i) a nucleic acid represented by SEQ ID NO: 39;

(ii) the complement of a nucleic acid represented by SEQ ID NO: 39;

(iii) a nucleic acid encoding a PRR-like polypeptide having in increasing order of preference at least 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence represented by SEQ ID NO: 40, and additionally or alternatively comprising one or more motifs having in increasing order of preference at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more sequence identity to any one or more of the motifs given in SEQ ID NO: 43 to SEQ ID NO: 45, and further preferably conferring enhanced yield-related traits relative to control plants.

(iv) a nucleic acid molecule which hybridizes with a nucleic acid molecule of (i) to (iii) under high stringency hybridization conditions and preferably confers enhanced yield-related traits relative to control plants.

According to a further embodiment of the present invention, there is also provided an isolated polypeptide selected from:

(i) an amino acid sequence represented by SEQ ID NO: 40;

(ii) an amino acid sequence having, in increasing order of preference, at least 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence represented by SEQ ID NO: 40, and additionally or alternatively comprising one or more motifs having in increasing order of preference at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more sequence identity to any one or more of the motifs given in SEQ ID NO: 43 to SEQ ID NO: 45, and further preferably conferring enhanced yield-related traits relative to control plants;

(iii) derivatives of any of the amino acid sequences given in (i) or (ii) above.

Nucleic acid variants may also be useful in practising the methods of the invention. Examples of such variants include nucleic acids encoding homologues and derivatives of any one of the amino acid sequences given in Table A of the Examples section, the terms “homologue” and “derivative” being as defined herein. Also useful in the methods of the invention are nucleic acids encoding homologues and derivatives of orthologues or

paralogues of any one of the amino acid sequences given in Table A of the Examples section. Homologues and derivatives useful in the methods of the present invention have substantially the same biological and functional activity as the unmodified protein from which they are derived. Further variants useful in practising the methods of the invention are variants in which codon usage is optimised or in which miRNA target sites are removed.

Further nucleic acid variants useful in practising the methods of the invention include portions of nucleic acids encoding PRR-like polypeptides, nucleic acids hybridising to nucleic acids encoding PRR-like polypeptides, splice variants of nucleic acids encoding PRR-like polypeptides, allelic variants of nucleic acids encoding PRR-like polypeptides and variants of nucleic acids encoding PRR-like polypeptides obtained by gene shuffling. The terms hybridising sequence, splice variant, allelic variant and gene shuffling are as described herein.

Nucleic acids encoding PRR-like polypeptides need not be full-length nucleic acids, since performance of the methods of the invention does not only rely on the use of full-length nucleic acid sequences. According to the present invention, there is provided a method for enhancing yield-related traits in plants, comprising introducing and expressing in a plant a portion of any one of the nucleic acid sequences given in Table A of the Examples section, or a portion of a nucleic acid encoding an orthologue, paralogue or homologue of any of the amino acid sequences given in Table A of the Examples section.

A portion of a nucleic acid may be prepared, for example, by making one or more deletions to the nucleic acid. The portions may be used in isolated form or they may be fused to other coding (or non-coding) sequences in order to, for example, produce a protein that combines several activities. When fused to other coding sequences, the resultant polypeptide produced upon translation may be bigger than that predicted for the protein portion.

Portions useful in the methods of the invention, encode a PRR-like polypeptide as defined herein or at least part thereof, and have substantially the same biological activity as the amino acid sequences given in Table A of the Examples section. Preferably, the portion is a portion of any one of the nucleic acids given in Table A of the Examples section, or is a portion of a nucleic acid encoding an orthologue or paralogue of any one of the amino acid sequences given in Table A of the Examples section. Preferably the portion is at least 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, 1450, 1500, 1550, 1600, 1650, 1700, 1750, 1800, 1850, 1900, 1950, 2000, 2050, 2100, 2150, 2200, 2250, 2300 consecutive nucleotides in length, the consecutive nucleotides being of any one of the nucleic acid sequences given in Table A of the Examples section, or of a nucleic acid encoding an orthologue or paralogue of any one of the amino acid sequences given in Table A of the Examples section. Most preferably the portion is a portion of the nucleic acid of SEQ ID NO: 1. Preferably, the portion encodes a fragment of an amino acid sequence which comprises any of the motifs 1 to 3 (SEQ ID NO:

43 to 45), and/or has at least 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to SEQ ID NO: 2.

Another nucleic acid variant useful in the methods of the invention is a nucleic acid capable of hybridising, under reduced stringency conditions, preferably under stringent conditions, with a nucleic acid encoding a PRR-like polypeptide as defined herein, or with a portion as defined herein. According to the present invention, there is provided a method for enhancing yield-related traits in plants, comprising introducing and expressing in a plant a nucleic acid capable of hybridizing to the complement of a nucleic acid encoding any one of the proteins given in Table A of the Examples section, or to the complement of a nucleic acid encoding an orthologue, paralogue or homologue of any one of the proteins given in Table A.

Hybridising sequences useful in the methods of the invention encode a PRR-like polypeptide as defined herein, having substantially the same biological activity as the amino acid sequences given in Table A of the Examples section. Preferably, the hybridising sequence is capable of hybridising to the complement of a nucleic acid encoding any one of the proteins given in Table A of the Examples section, or to a portion of any of these sequences, a portion being as defined herein, or the hybridising sequence is capable of hybridising to the complement of a nucleic acid encoding an orthologue or paralogue of any one of the amino acid sequences given in Table A of the Examples section. Most preferably, the hybridising sequence is capable of hybridising to the complement of a nucleic acid encoding the polypeptide as represented by SEQ ID NO: 2 or to a portion thereof. In one embodiment, the hybridization conditions are of medium stringency, preferably of high stringency, as defined herein.

Preferably, the hybridising sequence encodes a polypeptide with an amino acid sequence which comprises any of the motifs 1 to 3 (SEQ ID NO: 43 to 45), and/or has at least 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to SEQ ID NO: 2.

In another embodiment, there is provided a method for enhancing yield-related traits in plants, comprising introducing and expressing in a plant a splice variant of a nucleic acid

encoding any one of the proteins given in Table A of the Examples section, or a splice variant of a nucleic acid encoding an orthologue, paralogue or homologue of any of the amino acid sequences given in Table A of the Examples section.

Preferred splice variants are splice variants of a nucleic acid represented by SEQ ID NO: 1, or a splice variant of a nucleic acid encoding an orthologue or paralogue of SEQ ID NO: 2. Preferably, the amino acid sequence encoded by the splice variant comprises any of the motifs 1 to 3 (SEQ ID NO: 43 to 45), and/or has at least 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to SEQ ID NO: 2.

In yet another embodiment, there is provided a method for enhancing yield-related traits in plants, comprising introducing and expressing in a plant an allelic variant of a nucleic acid encoding any one of the proteins given in Table A of the Examples section, or comprising introducing and expressing in a plant an allelic variant of a nucleic acid encoding an orthologue, paralogue or homologue of any of the amino acid sequences given in Table A of the Examples section.

The polypeptides encoded by allelic variants useful in the methods of the present invention have substantially the same biological activity as the PRR-like polypeptide of SEQ ID NO: 2 and any of the amino acid sequences depicted in Table A of the Examples section. Allelic variants exist in nature, and encompassed within the methods of the present invention is the use of these natural alleles. Preferably, the allelic variant is an allelic variant of SEQ ID NO: 1 or an allelic variant of a nucleic acid encoding an orthologue or paralogue of SEQ ID NO: 2. Preferably, the amino acid sequence encoded by the allelic variant comprises any of the motifs 1 to 3 (SEQ ID NO: 43 to 45), and/or has at least 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to SEQ ID NO: 2.

In yet another embodiment, there is provided a method for enhancing yield-related traits in plants, comprising introducing and expressing in a plant a variant of a nucleic acid encoding any one of the proteins given in Table A of the Examples section, or comprising introducing and expressing in a plant a variant of a nucleic acid encoding an orthologue, paralogue or homologue of any of the amino acid sequences given in Table A of the Examples section, which variant nucleic acid is obtained by gene shuffling.

Preferably, the amino acid sequence encoded by the variant nucleic acid obtained by gene shuffling comprises any of the motifs 1 to 3 (SEQ ID NO: 43 to 45), and/or has at least 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to SEQ ID NO: 2.

Furthermore, nucleic acid variants may also be obtained by site-directed mutagenesis. Several methods are available to achieve site-directed mutagenesis, the most common being PCR based methods (Current Protocols in Molecular Biology. Wiley Eds.). PRR-like polypeptides differing from the sequence of SEQ ID NO: 2 by one or several amino acids (substitution(s), insertion(s) and/or deletion(s) as defined herein) may equally be useful to increase the yield of plants in the methods and constructs and plants of the invention.

Nucleic acids encoding PRR-like polypeptides may be derived from any natural or artificial source. The nucleic acid may be modified from its native form in composition and/or genomic environment through deliberate human manipulation. Preferably the PRR-like polypeptide-encoding nucleic acid is from a plant, further preferably from a monocotyledonous plant, more preferably from the family Poaceae, most preferably the nucleic acid is from *Saccharum officinarum*.

In another embodiment the present invention extends to recombinant chromosomal DNA comprising a nucleic acid sequence useful in the methods of the invention, wherein said nucleic acid is present in the chromosomal DNA as a result of recombinant methods, but is not in its natural genetic environment. In a further embodiment the recombinant chromosomal DNA of the invention is comprised in a plant cell.

Performance of the methods of the invention gives plants having enhanced yield-related traits. In particular performance of the methods of the invention gives plants having increased yield, especially increased aboveground biomass, increased root biomass and/or increased seed yield relative to control plants. The terms “yield” and “seed yield” are described in more detail in the “definitions” section herein.

The present invention thus provides a method for increasing yield-related traits, more in particular yield, especially aboveground biomass, root biomass and seed yield of plants, relative to control plants, which method comprises modulating expression in a plant of a nucleic acid encoding a PRR-like polypeptide as defined herein.

According to a preferred feature of the present invention, performance of the methods of the invention gives plants having an increased growth rate relative to control plants. Therefore, according to the present invention, there is provided a method for increasing the growth rate of plants, which method comprises modulating expression in a plant of a nucleic acid encoding a PRR-like polypeptide as defined herein.

Performance of the methods of the invention gives plants grown under non-stress conditions or under mild drought conditions increased yield-related traits relative to control plants grown under comparable conditions. Therefore, according to the present invention, there is provided a method for increasing yield-related traits in plants grown under non-stress conditions or under mild drought conditions, which method comprises modulating expression in a plant of a nucleic acid encoding a PRR-like polypeptide.

Performance of the methods of the invention gives plants grown under conditions of drought, increased yield-related traits relative to control plants grown under comparable conditions. Therefore, according to the present invention, there is provided a method for increasing yield-related traits in plants grown under conditions of drought which method comprises modulating expression in a plant of a nucleic acid encoding a PRR-like polypeptide.

Performance of the methods of the invention gives plants grown under conditions of nutrient deficiency, particularly under conditions of nitrogen deficiency, increased yield-related traits relative to control plants grown under comparable conditions. Therefore, according to the present invention, there is provided a method for increasing yield-related traits in plants grown under conditions of nutrient deficiency, which method comprises modulating expression in a plant of a nucleic acid encoding a PRR-like polypeptide.

Performance of the methods of the invention gives plants grown under conditions of salt stress, increased yield-related traits relative to control plants grown under comparable conditions. Therefore, according to the present invention, there is provided a method for increasing yield-related traits in plants grown under conditions of salt stress, which method comprises modulating expression in a plant of a nucleic acid encoding a PRR-like polypeptide.

The invention also provides genetic constructs and vectors to facilitate introduction and/or expression in plants of nucleic acids encoding PRR-like polypeptides. The gene constructs may be inserted into vectors, which may be commercially available, suitable for transforming into plants or host cells and suitable for expression of the gene of interest in the transformed cells. The invention also provides use of a gene construct as defined herein in the methods of the invention.

More specifically, the present invention provides a construct comprising:

(a) a nucleic acid encoding a PRR-like polypeptide as defined above;
(b) one or more control sequences capable of driving expression of the nucleic acid sequence of (a); and optionally
(c) a transcription termination sequence.

Preferably, the nucleic acid encoding a PRR-like polypeptide is as defined above. The term "control sequence" and "termination sequence" are as defined herein.

The genetic construct of the invention may be comprised in a host cell, plant cell, seed, agricultural product or plant. Plants or host cells are transformed with a genetic construct such as a vector or an expression cassette comprising any of the nucleic acids described above. Thus the invention furthermore provides plants or host cells transformed with a construct as described above. In particular, the invention provides plants transformed with a construct as described above, which plants have increased yield-related traits as described herein.

In one embodiment the genetic construct of the invention confers increased yield or yield related traits(s) to a plant when it has been introduced into said plant, which plant expresses the nucleic acid encoding the PRR-like comprised in the genetic construct. In another embodiment the genetic construct of the invention confers increased yield or yield related traits(s) to a plant comprising plant cells in which the construct has been introduced, which plant cells express the nucleic acid encoding the PRR-like comprised in the genetic construct.

The skilled artisan is well aware of the genetic elements that must be present on the genetic construct in order to successfully transform, select and propagate host cells containing the sequence of interest. The sequence of interest is operably linked to one or more control sequences, at least to a promoter.

Advantageously, any type of promoter, whether natural or synthetic, may be used to drive expression of the nucleic acid sequence, but preferably the promoter is of plant origin. A constitutive promoter is particularly useful in the methods. See the "Definitions" section herein for definitions of the various promoter types.

The constitutive promoter is preferably a ubiquitous constitutive promoter of medium strength. More preferably it is a plant derived promoter, e.g. a promoter of plant chromosomal origin, such as a GOS2 promoter or a promoter of substantially the same strength and having substantially the same expression pattern (a functionally equivalent promoter), more preferably the promoter is the promoter GOS2 promoter from rice. Further preferably the constitutive promoter is represented by a nucleic acid sequence substantially similar to SEQ ID NO: 46, most preferably the constitutive promoter is as represented by

SEQ ID NO: 46. See the "Definitions" section herein for further examples of constitutive promoters.

It should be clear that the applicability of the present invention is not restricted to the PRR-like polypeptide-encoding nucleic acid represented by SEQ ID NO: 1, nor is the applicability of the invention restricted to the rice GOS2 promoter when expression of a PRR-like polypeptide-encoding nucleic acid is driven by a constitutive promoter.

Optionally, one or more terminator sequences may be used in the construct introduced into a plant. Those skilled in the art will be aware of terminator sequences that may be suitable for use in performing the invention. Preferably, the construct comprises an expression cassette comprising a GOS2 promoter, substantially similar to SEQ ID NO: 46, operably linked to the nucleic acid encoding the PRR-like polypeptide. More preferably, the construct furthermore comprises a zein terminator (t-zein) linked to the 3' end of the PRR-like coding sequence. Furthermore, one or more sequences encoding selectable markers may be present on the construct introduced into a plant.

According to a preferred feature of the invention, the modulated expression is increased expression. Methods for increasing expression of nucleic acids or genes, or gene products, are well documented in the art and examples are provided in the definitions section.

As mentioned above, a preferred method for modulating expression of a nucleic acid encoding a PRR-like polypeptide is by introducing and expressing in a plant a nucleic acid encoding a PRR-like polypeptide; however the effects of performing the method, i.e. enhancing yield-related traits may also be achieved using other well known techniques, including but not limited to T-DNA activation tagging, TILLING, homologous recombination. A description of these techniques is provided in the definitions section.

The invention also provides a method for the production of transgenic plants having enhanced yield-related traits relative to control plants, comprising introduction and expression in a plant of any nucleic acid encoding a PRR-like polypeptide as defined herein.

More specifically, the present invention provides a method for the production of transgenic plants having enhanced yield-related traits, particularly increased biomass and seed yield, which method comprises:

(i) introducing and expressing in a plant or plant cell a PRR-like polypeptide-encoding nucleic acid or a genetic construct comprising a PRR-like polypeptide-encoding nucleic acid; and

(ii) cultivating the plant cell under conditions promoting plant growth and development.

The nucleic acid of (i) may be any of the nucleic acids capable of encoding a PRR-like polypeptide as defined herein.

Cultivating the plant cell under conditions promoting plant growth and development, may or may not include regeneration and/or growth to maturity. Accordingly, in a particular embodiment of the invention, the plant cell transformed by the method according to the invention is regenerable into a transformed plant. In another particular embodiment, the plant cell transformed by the method according to the invention is not regenerable into a transformed plant, i.e. cells that are not capable to regenerate into a plant using cell culture techniques known in the art. While plants cells generally have the characteristic of totipotency, some plant cells can not be used to regenerate or propagate intact plants from said cells. In one embodiment of the invention the plant cells of the invention are such cells. In another embodiment the plant cells of the invention are plant cells that do not sustain themselves in an autotrophic way.

The nucleic acid may be introduced directly into a plant cell or into the plant itself (including introduction into a tissue, organ or any other part of a plant). According to a preferred feature of the present invention, the nucleic acid is preferably introduced into a plant or plant cell by transformation. The term “transformation” is described in more detail in the “definitions” section herein.

In one embodiment the present invention extends to any plant cell or plant produced by any of the methods described herein, and to all plant parts and propagules thereof.

The present invention encompasses plants or parts thereof, including seeds, obtainable by the methods according to the present invention. The plants or plant parts or plant cells comprise a nucleic acid transgene encoding a PRR-like polypeptide as defined above, preferably in a genetic construct such as an expression cassette. The present invention extends further to encompass the progeny of a primary transformed or transfected cell, tissue, organ or whole plant that has been produced by any of the aforementioned methods, the only requirement being that progeny exhibit the same genotypic and/or phenotypic characteristic(s) as those produced by the parent in the methods according to the invention.

In a further embodiment the invention extends to seeds comprising the expression cassettes of the invention, the genetic constructs of the invention, or the nucleic acids encoding the PRR-like and/or the PRR-like polypeptides as described above.

The invention also includes host cells containing an isolated nucleic acid encoding a PRR-like polypeptide as defined above. In one embodiment host cells according to the invention are plant cells, yeasts, bacteria or fungi. Host plants for the nucleic acids, construct, expression cassette or the vector used in the method according to the invention are, in principle, advantageously all plants which are capable of synthesizing the polypeptides

used in the inventive method. In a particular embodiment the plant cells of the invention overexpress the nucleic acid molecule of the invention.

The methods of the invention are advantageously applicable to any plant, in particular to any plant as defined herein. Plants that are particularly useful in the methods of the invention include all plants which belong to the superfamily Viridiplantae, in particular monocotyledonous and dicotyledonous plants including fodder or forage legumes, ornamental plants, food crops, trees or shrubs. According to an embodiment of the present invention, the plant is a crop plant. Examples of crop plants include but are not limited to chicory, carrot, cassava, trefoil, soybean, beet, sugar beet, sunflower, canola, alfalfa, rapeseed, linseed, cotton, tomato, potato and tobacco. According to another embodiment of the present invention, the plant is a monocotyledonous plant. Examples of monocotyledonous plants include sugarcane. According to another embodiment of the present invention, the plant is a cereal. Examples of cereals include rice, maize, wheat, barley, millet, rye, triticale, sorghum, emmer, spelt, einkorn, teff, milo and oats. In a particular embodiment the plants used in the methods of the invention are selected from the group consisting of maize, wheat, rice, soybean, cotton, oilseed rape including canola, sugarcane, sugar beet and alfalfa. Advantageously the methods of the invention are more efficient than the known methods, because the plants of the invention have increased yield and/or tolerance to an environmental stress compared to control plants used in comparable methods.

The invention also extends to harvestable parts of a plant such as, but not limited to seeds, leaves, fruits, flowers, stems, roots, rhizomes, tubers and bulbs, which harvestable parts comprise a recombinant nucleic acid encoding a PRR-like polypeptide. The invention furthermore relates to products derived or produced, preferably directly derived or produced, from a harvestable part of such a plant, such as dry pellets, meal or powders, oil, fat and fatty acids, starch or proteins.

The invention also includes methods for manufacturing a product comprising a) growing the plants of the invention and b) producing said product from or by the plants of the invention or parts thereof, including seeds. In a further embodiment the methods comprise the steps of a) growing the plants of the invention, b) removing the harvestable parts as described herein from the plants and c) producing said product from, or with the harvestable parts of plants according to the invention.

In one embodiment the products produced by the methods of the invention are plant products such as, but not limited to, a foodstuff, feedstuff, a food supplement, feed supplement, fiber, cosmetic or pharmaceutical. In another embodiment the methods for production are used to make agricultural products such as, but not limited to, plant extracts, proteins, amino acids, carbohydrates, fats, oils, polymers, vitamins, and the like.

In yet another embodiment the polynucleotides or the polypeptides of the invention are comprised in an agricultural product. In a particular embodiment the nucleic acid sequences and protein sequences of the invention may be used as product markers, for example where an agricultural product was produced by the methods of the invention. Such a marker can be used to identify a product to have been produced by an advantageous process resulting not only in a greater efficiency of the process but also improved quality of the product due to increased quality of the plant material and harvestable parts used in the process. Such markers can be detected by a variety of methods known in the art, for example but not limited to PCR based methods for nucleic acid detection or antibody based methods for protein detection.

The present invention also encompasses use of nucleic acids encoding PRR-like polypeptides as described herein and use of these PRR-like polypeptides in enhancing any of the aforementioned yield-related traits in plants. For example, nucleic acids encoding PRR-like polypeptide described herein, or the PRR-like polypeptides themselves, may find use in breeding programmes in which a DNA marker is identified which may be genetically linked to a PRR-like polypeptide-encoding gene. The nucleic acids/genes, or the PRR-like polypeptides themselves may be used to define a molecular marker. This DNA or protein marker may then be used in breeding programmes to select plants having enhanced yield-related traits as defined herein in the methods of the invention. Furthermore, allelic variants of a PRR-like polypeptide-encoding nucleic acid/gene may find use in marker-assisted breeding programmes. Nucleic acids encoding PRR-like polypeptides may also be used as probes for genetically and physically mapping the genes that they are a part of, and as markers for traits linked to those genes. Such information may be useful in plant breeding in order to develop lines with desired phenotypes.

Moreover, the present invention relates to the following specific embodiments:

1. A method for enhancing yield-related traits in plants relative to control plants, comprising modulating expression in a plant of a nucleic acid encoding a PRR-like polypeptide, wherein said PRR-like polypeptide comprises an InterPro accession IPR001789 signal transduction response regulator, receiver domain corresponding to PFAM accession number PF00072 and/or an InterPro accession IPR010402 CCT domain corresponding to PFAM accession number PF06203.
2. Method according to embodiment 1, wherein said modulated expression is effected by introducing and expressing in a plant said nucleic acid encoding said PRR-like polypeptide.
3. Method according to embodiment 1 or 2, wherein said enhanced yield-related traits comprise increased yield relative to control plants, and preferably comprise increased biomass and/or increased seed yield relative to control plants.
4. Method according to any one of embodiments 1 to 3, wherein said enhanced yield-related traits are obtained under non-stress conditions.

5. Method according to any one of embodiments 1 to 3, wherein said enhanced yield-related traits are obtained under conditions of drought stress, salt stress or nitrogen deficiency.

6. Method according to any of embodiments 1 to 5, wherein said PRR-like polypeptide comprises one or more of the following motifs:
   (i) Motif 1 represented by SEQ ID NO: 43,
   (ii) Motif 2 represented by SEQ ID NO: 44,
   (iii) Motif 3 represented by SEQ ID NO: 45.

7. Method according to any one of embodiments 1 to 6, wherein said nucleic acid encoding a PRR-like polypeptide is of plant origin, preferably from a dicotyledonous plant, further preferably from the family Poaceae, more preferably from the genus Saccharum, most preferably from Saccharum officinarum.

8. Method according to any one of embodiments 1 to 7, wherein said nucleic acid encoding a PRR-like polypeptide encodes any one of the polypeptides listed in Table A or is a portion of such a nucleic acid, or a nucleic acid capable of hybridising with such a nucleic acid.

9. Method according to any one of embodiments 1 to 8, wherein said nucleic acid sequence encodes an orthologue or paralogue of any of the polypeptides given in Table A.

10. Method according to any one of embodiments 1 to 9, wherein said nucleic acid encodes the polypeptide represented by SEQ ID NO: 2.

11. Method according to any one of embodiments 1 to 10, wherein said nucleic acid is operably linked to a constitutive promoter of plant origin, preferably to a medium strength constitutive promoter of plant origin, more preferably to a GOS2 promoter, most preferably to a GOS2 promoter from rice.

12. Plant, or part thereof, or plant cell, obtainable by a method according to any one of embodiments 1 to 11, wherein said plant, plant part or plant cell comprises a recombinant nucleic acid encoding a PRR-like polypeptide as defined in any of embodiments 1 and 6 to 10.

13. Construct comprising:
   (i) nucleic acid encoding an PRR-like as defined in any of embodiments 1 and 6 to 10;
   (ii) one or more control sequences capable of driving expression of the nucleic acid sequence of (i); and optionally
   (iii) a transcription termination sequence.

14. Construct according to embodiment 13, wherein one of said control sequences is a constitutive promoter of plant origin, preferably to a medium strength constitutive promoter of plant origin, more preferably to a GOS2 promoter, most preferably to a GOS2 promoter from rice.

15. Use of a construct according to embodiment 13 or 14 in a method for making plants having enhanced yield-related traits, preferably increased yield relative to control

plants, and more preferably increased seed yield and/or increased biomass relative to control plants.

16. Plant, plant part or plant cell transformed with a construct according to embodiment 13 or 14.

17. Method for the production of a transgenic plant having enhanced yield-related traits relative to control plants, preferably increased yield relative to control plants, and more preferably increased seed yield and/or increased biomass relative to control plants, comprising:
    (i) introducing and expressing in a plant cell or plant a nucleic acid encoding an PRR-like polypeptide as defined in any of embodiments 1 and 6 to 10; and
    (ii) cultivating said plant cell or plant under conditions promoting plant growth and development.

18. Transgenic plant having enhanced yield-related traits relative to control plants, preferably increased yield relative to control plants, and more preferably increased seed yield and/or increased biomass, resulting from modulated expression of a nucleic acid encoding an PRR-like polypeptide as defined in any of embodiments 1 and 6 to 10 or a transgenic plant cell derived from said transgenic plant.

19. Transgenic plant according to embodiment 12, 16 or 18, or a transgenic plant cell derived therefrom, wherein said plant is a crop plant, such as beet, sugarbeet or alfalfa; or a monocotyledonous plant such as sugarcane; or a cereal, such as rice, maize, wheat, barley, millet, rye, triticale, sorghum, emmer, spelt, einkorn, teff, milo or oats.

20. Harvestable parts of a plant according to embodiment 19, wherein said harvestable parts are preferably root biomass and/or seeds.

21. Products derived from a plant according to embodiment 19 and/or from harvestable parts of a plant according to embodiment 20.

22. Use of a nucleic acid encoding an PRR-like polypeptide as defined in any of embodiments 1 and 6 to 10 for enhancing yield-related traits in plants relative to control plants, preferably for increasing yield, and more preferably for increasing seed yield and/or for increasing biomass in plants relative to control plants.

23. A method for manufacturing a product comprising the steps of growing the plants according to embodiment 12, 16, 19 or 20 and producing said product from or by said plants; or parts thereof, including seeds.

24. Products produced from a plant according to embodiment 19 and/or from harvestable parts of a plant according to embodiment 20.

25. Construct according to embodiment 13 or 14 comprised in a plant cell.

26. Recombinant chromosomal DNA comprising the construct according to embodiment 13 or 14.

27. Method according to any one of embodiments 1 to 11, wherein said polypeptide is encoded by a nucleic acid molecule comprising a nucleic acid molecule selected from the group consisting of:
    (i) a nucleic acid represented by SEQ ID NO: 1;

(ii) the complement of a nucleic acid represented by SEQ ID NO: 1;

(iii) a nucleic acid encoding the polypeptide as represented by SEQ ID NO: 1, and further preferably confers enhanced yield-related traits relative to control plants;

(iv) a nucleic acid having, in increasing order of preference at least 30 %, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity with the nucleic acid sequences of SEQ ID NO: 1, and further preferably conferring enhanced yield-related traits relative to control plants,

(v) a nucleic acid molecule which hybridizes to the complement of a nucleic acid molecule of (i) to (iv) under stringent hybridization conditions and preferably confers enhanced yield-related traits relative to control plants;

(vi) a nucleic acid encoding said polypeptide having, in increasing order of preference, at least 30 %, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence represented by SEQ ID NO: 2 and preferably conferring enhanced yield-related traits relative to control plants; or

(vii) a nucleic acid comprising any combination(s) of features of (i) to (vi) above.

Definitions

The following definitions will be used throughout the present application. The section captions and headings in this application are for convenience and reference purpose only and should not affect in any way the meaning or interpretation of this application. The technical terms and expressions used within the scope of this application are generally to be given the meaning commonly applied to them in the pertinent art of plant biology, molecular biology, bioinformatics and plant breeding. All of the following term definitions apply to the complete content of this application. The term “essentially”, “about”, “approximately” and the like in connection with an attribute or a value, particularly also define exactly the attribute or exactly the value, respectively. The term “about” in the context of a given numeric value or range relates in particular to a value or range that is within 20%, within 10%, or within 5% of the value or range given.

Peptide(s)/Protein(s)

The terms “peptides”, “oligopeptides”, “polypeptide” and “protein” are used interchangeably herein and refer to amino acids in a polymeric form of any length, linked together by peptide bonds, unless mentioned herein otherwise.

Polynucleotide(s)/Nucleic acid(s)/Nucleic acid sequence(s)/nucleotide sequence(s)

The terms "polynucleotide(s)", "nucleic acid sequence(s)", "nucleotide sequence(s)", “nucleic acid(s)”, “nucleic acid molecule” are used interchangeably herein and refer to nucleotides, either ribonucleotides or deoxyribonucleotides or a combination of both, in a polymeric unbranched form of any length.

Homologue(s)

“Homologues” of a protein encompass peptides, oligopeptides, polypeptides, proteins and enzymes having amino acid substitutions, deletions and/or insertions relative to the unmodified protein in question and having similar biological and functional activity as the unmodified protein from which they are derived.

Orthologues and paralogues are two different forms of homologues and encompass evolutionary concepts used to describe the ancestral relationships of genes. Paralogues are genes within the same species that have originated through duplication of an ancestral gene; orthologues are genes from different organisms that have originated through speciation, and are also derived from a common ancestral gene.

A “deletion” refers to removal of one or more amino acids from a protein.

An “insertion” refers to one or more amino acid residues being introduced into a predetermined site in a protein. Insertions may comprise N-terminal and/or C-terminal fusions as well as intra-sequence insertions of single or multiple amino acids. Generally, insertions within the amino acid sequence will be smaller than N- or C-terminal fusions, of the order of about 1 to 10 residues. Examples of N- or C-terminal fusion proteins or peptides include the binding domain or activation domain of a transcriptional activator as used in the yeast two-hybrid system, phage coat proteins, (histidine)-6-tag, glutathione S-transferase-tag, protein A, maltose-binding protein, dihydrofolate reductase, Tag•100 epitope, c-myc epitope, FLAG®-epitope, lacZ, CMP (calmodulin-binding peptide), HA epitope, protein C epitope and VSV epitope.

A “substitution” refers to replacement of amino acids of the protein with other amino acids having similar properties (such as similar hydrophobicity, hydrophilicity, antigenicity, propensity to form or break α-helical structures or β-sheet structures). Amino acid substitutions are typically of single residues, but may be clustered depending upon functional constraints placed upon the polypeptide and may range from 1 to 10 amino acids. The amino acid substitutions are preferably conservative amino acid substitutions. Conservative substitution tables are well known in the art (see for example Creighton (1984) Proteins. W.H. Freeman and Company (Eds) and Table 1 below).

Table 1: Examples of conserved amino acid substitutions

Residue	Conservative Substitutions	Residue	Conservative Substitutions
Ala	Ser	Leu	Ile; Val
Arg	Lys	Lys	Arg; Gln
Asn	Gln; His	Met	Leu; Ile
Asp	Glu	Phe	Met; Leu; Tyr
Gln	Asn	Ser	Thr; Gly
Cys	Ser	Thr	Ser; Val
Glu	Asp	Trp	Tyr
Gly	Pro	Tyr	Trp; Phe
His	Asn; Gln	Val	Ile; Leu
Ile	Leu, Val		

Amino acid substitutions, deletions and/or insertions may readily be made using peptide synthetic techniques known in the art, such as solid phase peptide synthesis and the like, or by recombinant DNA manipulation. Methods for the manipulation of DNA sequences to produce substitution, insertion or deletion variants of a protein are well known in the art. For example, techniques for making substitution mutations at predetermined sites in DNA are well known to those skilled in the art and include M13 mutagenesis, T7-Gen in vitro mutagenesis (USB, Cleveland, OH), QuickChange Site Directed mutagenesis (Stratagene, San Diego, CA), PCR-mediated site-directed mutagenesis or other site-directed mutagenesis protocols (see Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989 and yearly updates)).

Derivatives

"Derivatives" include peptides, oligopeptides, polypeptides which may, compared to the amino acid sequence of the naturally-occurring form of the protein, such as the protein of interest, comprise substitutions of amino acids with non-naturally occurring amino acid residues, or additions of non-naturally occurring amino acid residues. "Derivatives" of a protein also encompass peptides, oligopeptides, polypeptides which comprise naturally occurring altered (glycosylated, acylated, prenylated, phosphorylated, myristoylated, sulphated etc.) or non-naturally altered amino acid residues compared to the amino acid sequence of a naturally-occurring form of the polypeptide. A derivative may also comprise one or more non-amino acid substituents or additions compared to the amino acid sequence from which it is derived, for example a reporter molecule or other ligand, covalently or non-covalently bound to the amino acid sequence, such as a reporter molecule which is bound to facilitate its detection, and non-naturally occurring amino acid residues relative to the amino acid sequence of a naturally-occurring protein. Furthermore, "derivatives" also include fusions of the naturally-occurring form of the protein with tagging peptides such as FLAG, HIS6 or thioredoxin (for a review of tagging peptides, see Terpe, Appl. Microbiol. Biotechnol. 60, 523-533, 2003).

Domain, Motif/Consensus sequence/Signature

The term "domain" refers to a set of amino acids conserved at specific positions along an alignment of sequences of evolutionarily related proteins. While amino acids at other positions can vary between homologues, amino acids that are highly conserved at specific positions indicate amino acids that are likely essential in the structure, stability or function of a protein. Identified by their high degree of conservation in aligned sequences of a family of protein homologues, they can be used as identifiers to determine if any polypeptide in question belongs to a previously identified polypeptide family.

The term “motif” or “consensus sequence” or “signature” refers to a short conserved region in the sequence of evolutionarily related proteins. Motifs are frequently highly conserved parts of domains, but may also include only part of the domain, or be located outside of conserved domain (if all of the amino acids of the motif fall outside of a defined domain).

Specialist databases exist for the identification of domains, for example, SMART (Schultz et al. (1998) Proc. Natl. Acad. Sci. USA 95, 5857-5864; Letunic et al. (2002) Nucleic Acids Res 30, 242-244), InterPro (Mulder et al., (2003) Nucl. Acids. Res. 31, 315-318), Prosite (Bucher and Bairoch (1994), A generalized profile syntax for biomolecular sequences motifs and its function in automatic sequence interpretation. (In) ISMB-94; Proceedings 2nd International Conference on Intelligent Systems for Molecular Biology. Altman R., Brutlag D., Karp P., Lathrop R., Searls D., Eds., pp53-61, AAAI Press, Menlo Park; Hulo et al., Nucl. Acids. Res. 32:D134-D137, (2004)), or Pfam (Bateman et al., Nucleic Acids Research 30(1): 276-280 (2002)). A set of tools for *in silico* analysis of protein sequences is available on the ExPASy proteomics server (Swiss Institute of Bioinformatics (Gasteiger et al., ExPASy: the proteomics server for in-depth protein knowledge and analysis, Nucleic Acids Res. 31:3784-3788(2003)). Domains or motifs may also be identified using routine techniques, such as by sequence alignment.

Methods for the alignment of sequences for comparison are well known in the art, such methods include GAP, BESTFIT, BLAST, FASTA and TFASTA. GAP uses the algorithm of Needleman and Wunsch ((1970) J Mol Biol 48: 443-453) to find the global (i.e. spanning the complete sequences) alignment of two sequences that maximizes the number of matches and minimizes the number of gaps. The BLAST algorithm (Altschul et al. (1990) J Mol Biol 215: 403-10) calculates percent sequence identity and performs a statistical analysis of the similarity between the two sequences. The software for performing BLAST analysis is publicly available through the National Centre for Biotechnology Information (NCBI). Homologues may readily be identified using, for example, the ClustalW multiple sequence alignment algorithm (version 1.83), with the default pairwise alignment parameters, and a scoring method in percentage. Global percentages of similarity and identity may also be determined using one of the methods available in the MatGAT software package (Campanella et al., BMC Bioinformatics. 2003 Jul 10;4:29. MatGAT: an application that generates similarity/identity matrices using protein or DNA sequences.). Minor manual

editing may be performed to optimise alignment between conserved motifs, as would be apparent to a person skilled in the art. Furthermore, instead of using full-length sequences for the identification of homologues, specific domains may also be used. The sequence identity values may be determined over the entire nucleic acid or amino acid sequence or over selected domains or conserved motif(s), using the programs mentioned above using the default parameters. For local alignments, the Smith-Waterman algorithm is particularly useful (Smith TF, Waterman MS (1981) J. Mol. Biol 147(1);195-7).

Reciprocal BLAST

Typically, this involves a first BLAST involving BLASTing a query sequence (for example using any of the sequences listed in Table A of the Examples section) against any sequence database, such as the publicly available NCBI database. BLASTN or TBLASTX (using standard default values) are generally used when starting from a nucleotide sequence, and BLASTP or TBLASTN (using standard default values) when starting from a protein sequence. The BLAST results may optionally be filtered. The full-length sequences of either the filtered results or non-filtered results are then BLASTed back (second BLAST) against sequences from the organism from which the query sequence is derived. The results of the first and second BLASTs are then compared. A paralogue is identified if a high-ranking hit from the first blast is from the same species as from which the query sequence is derived, a BLAST back then ideally results in the query sequence amongst the highest hits; an orthologue is identified if a high-ranking hit in the first BLAST is not from the same species as from which the query sequence is derived, and preferably results upon BLAST back in the query sequence being among the highest hits.

High-ranking hits are those having a low E-value. The lower the E-value, the more significant the score (or in other words the lower the chance that the hit was found by chance). Computation of the E-value is well known in the art. In addition to E-values, comparisons are also scored by percentage identity. Percentage identity refers to the number of identical nucleotides (or amino acids) between the two compared nucleic acid (or polypeptide) sequences over a particular length. In the case of large families, ClustalW may be used, followed by a neighbour joining tree, to help visualize clustering of related genes and to identify orthologues and paralogues.

Hybridisation

The term "hybridisation" as defined herein is a process wherein substantially homologous complementary nucleotide sequences anneal to each other. The hybridisation process can occur entirely in solution, i.e. both complementary nucleic acids are in solution. The hybridisation process can also occur with one of the complementary nucleic acids immobilised to a matrix such as magnetic beads, Sepharose beads or any other resin. The hybridisation process can furthermore occur with one of the complementary nucleic acids immobilised to a solid support such as a nitro-cellulose or nylon membrane or immobilised by e.g. photolithography to, for example, a siliceous glass support (the latter known as

nucleic acid arrays or microarrays or as nucleic acid chips). In order to allow hybridisation to occur, the nucleic acid molecules are generally thermally or chemically denatured to melt a double strand into two single strands and/or to remove hairpins or other secondary structures from single stranded nucleic acids.

The term “stringency” refers to the conditions under which a hybridisation takes place. The stringency of hybridisation is influenced by conditions such as temperature, salt concentration, ionic strength and hybridisation buffer composition. Generally, low stringency conditions are selected to be about 30°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. Medium stringency conditions are when the temperature is 20°C below T_m, and high stringency conditions are when the temperature is 10°C below T_m. High stringency hybridisation conditions are typically used for isolating hybridising sequences that have high sequence similarity to the target nucleic acid sequence. However, nucleic acids may deviate in sequence and still encode a substantially identical polypeptide, due to the degeneracy of the genetic code. Therefore medium stringency hybridisation conditions may sometimes be needed to identify such nucleic acid molecules.

The T_m is the temperature under defined ionic strength and pH, at which 50% of the target sequence hybridises to a perfectly matched probe. The T_m is dependent upon the solution conditions and the base composition and length of the probe. For example, longer sequences hybridise specifically at higher temperatures. The maximum rate of hybridisation is obtained from about 16°C up to 32°C below T_m. The presence of monovalent cations in the hybridisation solution reduce the electrostatic repulsion between the two nucleic acid strands thereby promoting hybrid formation; this effect is visible for sodium concentrations of up to 0.4M (for higher concentrations, this effect may be ignored). Formamide reduces the melting temperature of DNA-DNA and DNA-RNA duplexes with 0.6 to 0.7°C for each percent formamide, and addition of 50% formamide allows hybridisation to be performed at 30 to 45°C, though the rate of hybridisation will be lowered. Base pair mismatches reduce the hybridisation rate and the thermal stability of the duplexes. On average and for large probes, the Tm decreases about 1°C per % base mismatch. The T_m may be calculated using the following equations, depending on the types of hybrids:

1) DNA-DNA hybrids (Meinkoth and Wahl, Anal. Biochem., 138: 267-284, 1984):

$$T_m = 81.5°C + 16.6x\log_{10}[Na^+]^a + 0.41x\%[G/C^b] - 500x[L^c]^{-1} - 0.61x\% \text{ formamide}$$

2) DNA-RNA or RNA-RNA hybrids:

$$T_m = 79.8°C + 18.5\ (\log_{10}[Na^+]^a) + 0.58\ (\%G/C^b) + 11.8\ (\%G/C^b)^2 - 820/L^c$$

3) oligo-DNA or oligo-RNA^d hybrids:

For <20 nucleotides: $T_m = 2\ (l_n)$

For 20–35 nucleotides: $T_m = 22 + 1.46\ (l_n)$

[a] or for other monovalent cation, but only accurate in the 0.01–0.4 M range.

[b] only accurate for %GC in the 30% to 75% range.

[c] L = length of duplex in base pairs.

[d] oligo, oligonucleotide; l_n, = effective length of primer = 2×(no. of G/C)+(no. of A/T).

Non-specific binding may be controlled using any one of a number of known techniques such as, for example, blocking the membrane with protein containing solutions, additions of heterologous RNA, DNA, and SDS to the hybridisation buffer, and treatment with Rnase. For non-homologous probes, a series of hybridizations may be performed by varying one of (i) progressively lowering the annealing temperature (for example from 68°C to 42°C) or (ii) progressively lowering the formamide concentration (for example from 50% to 0%). The skilled artisan is aware of various parameters which may be altered during hybridisation and which will either maintain or change the stringency conditions.

Besides the hybridisation conditions, specificity of hybridisation typically also depends on the function of post-hybridisation washes. To remove background resulting from non-specific hybridisation, samples are washed with dilute salt solutions. Critical factors of such washes include the ionic strength and temperature of the final wash solution: the lower the salt concentration and the higher the wash temperature, the higher the stringency of the wash. Wash conditions are typically performed at or below hybridisation stringency. A positive hybridisation gives a signal that is at least twice of that of the background. Generally, suitable stringent conditions for nucleic acid hybridisation assays or gene amplification detection procedures are as set forth above. More or less stringent conditions may also be selected. The skilled artisan is aware of various parameters which may be altered during washing and which will either maintain or change the stringency conditions.

For example, typical high stringency hybridisation conditions for DNA hybrids longer than 50 nucleotides encompass hybridisation at 65°C in 1x SSC or at 42°C in 1x SSC and 50% formamide, followed by washing at 65°C in 0.3x SSC. Examples of medium stringency hybridisation conditions for DNA hybrids longer than 50 nucleotides encompass hybridisation at 50°C in 4x SSC or at 40°C in 6x SSC and 50% formamide, followed by washing at 50°C in 2x SSC. The length of the hybrid is the anticipated length for the hybridising nucleic acid. When nucleic acids of known sequence are hybridised, the hybrid length may be determined by aligning the sequences and identifying the conserved regions described herein. 1×SSC is 0.15M NaCl and 15mM sodium citrate; the hybridisation solution and wash solutions may additionally include 5x Denhardt's reagent, 0.5-1.0% SDS, 100 µg/ml denatured, fragmented salmon sperm DNA, 0.5% sodium pyrophosphate.

For the purposes of defining the level of stringency, reference can be made to Sambrook et al. (2001) Molecular Cloning: a laboratory manual, 3rd Edition, Cold Spring Harbor Laboratory Press, CSH, New York or to Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989 and yearly updates).

Splice variant

The term "splice variant" as used herein encompasses variants of a nucleic acid sequence in which selected introns and/or exons have been excised, replaced, displaced or added, or in which introns have been shortened or lengthened. Such variants will be ones in which the biological activity of the protein is substantially retained; this may be achieved by selectively retaining functional segments of the protein. Such splice variants may be found in nature or may be manmade. Methods for predicting and isolating such splice variants are well known in the art (see for example Foissac and Schiex (2005) BMC Bioinformatics 6: 25).

Allelic variant

"Alleles" or "allelic variants" are alternative forms of a given gene, located at the same chromosomal position. Allelic variants encompass Single Nucleotide Polymorphisms (SNPs), as well as Small Insertion/Deletion Polymorphisms (INDELs). The size of INDELs is usually less than 100 bp. SNPs and INDELs form the largest set of sequence variants in naturally occurring polymorphic strains of most organisms.

Endogenous gene

Reference herein to an "endogenous" gene not only refers to the gene in question as found in a plant in its natural form (i.e., without there being any human intervention), but also refers to that same gene (or a substantially homologous nucleic acid/gene) in an isolated form subsequently (re)introduced into a plant (a transgene). For example, a transgenic plant containing such a transgene may encounter a substantial reduction of the transgene expression and/or substantial reduction of expression of the endogenous gene. The isolated gene may be isolated from an organism or may be manmade, for example by chemical synthesis.

Gene shuffling/Directed evolution

"Gene shuffling" or "directed evolution" consists of iterations of DNA shuffling followed by appropriate screening and/or selection to generate variants of nucleic acids or portions thereof encoding proteins having a modified biological activity (Castle et al., (2004) Science 304(5674): 1151-4; US patents 5,811,238 and 6,395,547).

Construct

Artificial DNA (such as but, not limited to plasmids or viral DNA) capable of replication in a host cell and used for introduction of a DNA sequence of interest into a host cell or host organism. Host cells of the invention may be any cell selected from bacterial cells, such as Escherichia coli or Agrobacterium species cells, yeast cells, fungal, algal or cyanobacterial cells or plant cells. The skilled artisan is well aware of the genetic elements that must be present on the genetic construct in order to successfully transform, select and propagate host cells containing the sequence of interest. The sequence of interest is operably linked to one or more control sequences (at least to a promoter) as described herein. Additional regulatory elements may include transcriptional as well as translational enhancers. Those

skilled in the art will be aware of terminator and enhancer sequences that may be suitable for use in performing the invention. An intron sequence may also be added to the 5' untranslated region (UTR) or in the coding sequence to increase the amount of the mature message that accumulates in the cytosol, as described in the definitions section. Other control sequences (besides promoter, enhancer, silencer, intron sequences, 3'UTR and/or 5'UTR regions) may be protein and/or RNA stabilizing elements. Such sequences would be known or may readily be obtained by a person skilled in the art.

The genetic constructs of the invention may further include an origin of replication sequence that is required for maintenance and/or replication in a specific cell type. One example is when a genetic construct is required to be maintained in a bacterial cell as an episomal genetic element (e.g. plasmid or cosmid molecule). Preferred origins of replication include, but are not limited to, the f1-ori and colE1.

For the detection of the successful transfer of the nucleic acid sequences as used in the methods of the invention and/or selection of transgenic plants comprising these nucleic acids, it is advantageous to use marker genes (or reporter genes). Therefore, the genetic construct may optionally comprise a selectable marker gene. Selectable markers are described in more detail in the “definitions” section herein. The marker genes may be removed or excised from the transgenic cell once they are no longer needed. Techniques for marker removal are known in the art, useful techniques are described above in the definitions section.

Regulatory element/Control sequence/Promoter

The terms “regulatory element”, “control sequence” and “promoter” are all used interchangeably herein and are to be taken in a broad context to refer to regulatory nucleic acid sequences capable of effecting expression of the sequences to which they are ligated. The term “promoter” typically refers to a nucleic acid control sequence located upstream from the transcriptional start of a gene and which is involved in recognising and binding of RNA polymerase and other proteins, thereby directing transcription of an operably linked nucleic acid. Encompassed by the aforementioned terms are transcriptional regulatory sequences derived from a classical eukaryotic genomic gene (including the TATA box which is required for accurate transcription initiation, with or without a CCAAT box sequence) and additional regulatory elements (i.e. upstream activating sequences, enhancers and silencers) which alter gene expression in response to developmental and/or external stimuli, or in a tissue-specific manner. Also included within the term is a transcriptional regulatory sequence of a classical prokaryotic gene, in which case it may include a –35 box sequence and/or –10 box transcriptional regulatory sequences. The term “regulatory element” also encompasses a synthetic fusion molecule or derivative that confers, activates or enhances expression of a nucleic acid molecule in a cell, tissue or organ.

A “plant promoter” comprises regulatory elements, which mediate the expression of a coding sequence segment in plant cells. Accordingly, a plant promoter need not be of plant origin, but may originate from viruses or micro-organisms, for example from viruses which attack plant cells. The “plant promoter” can also originate from a plant cell, e.g. from the plant which is transformed with the nucleic acid sequence to be expressed in the inventive process and described herein. This also applies to other “plant” regulatory signals, such as “plant” terminators. The promoters upstream of the nucleotide sequences useful in the methods of the present invention can be modified by one or more nucleotide substitution(s), insertion(s) and/or deletion(s) without interfering with the functionality or activity of either the promoters, the open reading frame (ORF) or the 3'-regulatory region such as terminators or other 3' regulatory regions which are located away from the ORF. It is furthermore possible that the activity of the promoters is increased by modification of their sequence, or that they are replaced completely by more active promoters, even promoters from heterologous organisms. For expression in plants, the nucleic acid molecule must, as described above, be linked operably to or comprise a suitable promoter which expresses the gene at the right point in time and with the required spatial expression pattern.

For the identification of functionally equivalent promoters, the promoter strength and/or expression pattern of a candidate promoter may be analysed for example by operably linking the promoter to a reporter gene and assaying the expression level and pattern of the reporter gene in various tissues of the plant. Suitable well-known reporter genes include for example beta-glucuronidase or beta-galactosidase. The promoter activity is assayed by measuring the enzymatic activity of the beta-glucuronidase or beta-galactosidase. The promoter strength and/or expression pattern may then be compared to that of a reference promoter (such as the one used in the methods of the present invention). Alternatively, promoter strength may be assayed by quantifying mRNA levels or by comparing mRNA levels of the nucleic acid used in the methods of the present invention, with mRNA levels of housekeeping genes such as 18S rRNA, using methods known in the art, such as Northern blotting with densitometric analysis of autoradiograms, quantitative real-time PCR or RT-PCR (Heid et al., 1996 Genome Methods 6: 986-994). Generally by “weak promoter” is intended a promoter that drives expression of a coding sequence at a low level. By “low level” is intended at levels of about 1/10,000 transcripts to about 1/100,000 transcripts, to about 1/500,0000 transcripts per cell. Conversely, a “strong promoter” drives expression of a coding sequence at high level, or at about 1/10 transcripts to about 1/100 transcripts to about 1/1000 transcripts per cell. Generally, by “medium strength promoter” is intended a promoter that drives expression of a coding sequence at a lower level than a strong promoter, in particular at a level that is in all instances below that obtained when under the control of a 35S CaMV promoter.

Operably linked

The term "operably linked" as used herein refers to a functional linkage between the promoter sequence and the gene of interest, such that the promoter sequence is able to initiate transcription of the gene of interest.

Constitutive promoter

A "constitutive promoter" refers to a promoter that is transcriptionally active during most, but not necessarily all, phases of growth and development and under most environmental conditions, in at least one cell, tissue or organ. Table 2a below gives examples of constitutive promoters.

Table 2a: Examples of constitutive promoters

Gene Source	Reference
Actin	McElroy et al, Plant Cell, 2: 163-171, 1990
HMGP	WO 2004/070039
CAMV 35S	Odell et al, Nature, 313: 810-812, 1985
CaMV 19S	Nilsson et al., Physiol. Plant. 100:456-462, 1997
GOS2	de Pater et al, Plant J Nov;2(6):837-44, 1992, WO 2004/065596
Ubiquitin	Christensen et al, Plant Mol. Biol. 18: 675-689, 1992
Rice cyclophilin	Buchholz et al, Plant Mol Biol. 25(5): 837-43, 1994
Maize H3 histone	Lepetit et al, Mol. Gen. Genet. 231:276-285, 1992
Alfalfa H3 histone	Wu et al. Plant Mol. Biol. 11:641-649, 1988
Actin 2	An et al, Plant J. 10(1); 107-121, 1996
34S FMV	Sanger et al., Plant. Mol. Biol., 14, 1990: 433-443
Rubisco small subunit	US 4,962,028
OCS	Leisner (1988) Proc Natl Acad Sci USA 85(5): 2553
SAD1	Jain et al., Crop Science, 39 (6), 1999: 1696
SAD2	Jain et al., Crop Science, 39 (6), 1999: 1696
nos	Shaw et al. (1984) Nucleic Acids Res. 12(20):7831-7846
V-ATPase	WO 01/14572
Super promoter	WO 95/14098
G-box proteins	WO 94/12015

Ubiquitous promoter

A "ubiquitous promoter" is active in substantially all tissues or cells of an organism.

Developmentally-regulated promoter

A "developmentally-regulated promoter" is active during certain developmental stages or in parts of the plant that undergo developmental changes.

Inducible promoter

An “inducible promoter” has induced or increased transcription initiation in response to a chemical (for a review see Gatz 1997, Annu. Rev. Plant Physiol. Plant Mol. Biol., 48:89-108), environmental or physical stimulus, or may be “stress-inducible”, i.e. activated when a plant is exposed to various stress conditions, or a “pathogen-inducible" i.e. activated when a plant is exposed to exposure to various pathogens.

Organ-specific/Tissue-specific promoter

An “organ-specific” or “tissue-specific promoter” is one that is capable of preferentially initiating transcription in certain organs or tissues, such as the leaves, roots, seed tissue etc. For example, a “root-specific promoter” is a promoter that is transcriptionally active predominantly in plant roots, substantially to the exclusion of any other parts of a plant, whilst still allowing for any leaky expression in these other plant parts. Promoters able to initiate transcription in certain cells only are referred to herein as “cell-specific”.

Examples of root-specific promoters are listed in Table 2b below:

Table 2b: Examples of root-specific promoters

Gene Source	Reference
RCc3	Plant Mol Biol. 1995 Jan;27(2):237-48
Arabidopsis PHT1	Koyama et al. J Biosci Bioeng. 2005 Jan;99(1):38-42.; Mudge et al. (2002, Plant J. 31:341)
Medicago phosphate transporter	Xiao et al., 2006, Plant Biol (Stuttg). 2006 Jul;8(4):439-49
Arabidopsis Pyk10	Nitz et al. (2001) Plant Sci 161(2): 337-346
root-expressible genes	Tingey et al., EMBO J. 6: 1, 1987.
tobacco auxin-inducible gene	Van der Zaal et al., Plant Mol. Biol. 16, 983, 1991.
β-tubulin	Oppenheimer, et al., Gene 63: 87, 1988.
tobacco root-specific genes	Conkling, et al., Plant Physiol. 93: 1203, 1990.
B. napus G1-3b gene	United States Patent No. 5, 401, 836
SbPRP1	Suzuki et al., Plant Mol. Biol. 21: 109-119, 1993.
LRX1	Baumberger et al. 2001, Genes & Dev. 15:1128
BTG-26 Brassica napus	US 20050044585
LeAMT1 (tomato)	Lauter et al. (1996, PNAS 3:8139)
The LeNRT1-1 (tomato)	Lauter et al. (1996, PNAS 3:8139)
class I patatin gene (potato)	Liu et al., Plant Mol. Biol. 17 (6): 1139-1154
KDC1 (Daucus carota)	Downey et al. (2000, J. Biol. Chem. 275:39420)
TobRB7 gene	W Song (1997) PhD Thesis, North Carolina State University, Raleigh, NC USA
OsRAB5a (rice)	Wang et al. 2002, Plant Sci. 163:273
ALF5 (Arabidopsis)	Diener et al. (2001, Plant Cell 13:1625)
NRT2;1Np (N. plumbaginifolia)	Quesada et al. (1997, Plant Mol. Biol. 34:265)

A “seed-specific promoter” is transcriptionally active predominantly in seed tissue, but not necessarily exclusively in seed tissue (in cases of leaky expression). The seed-specific promoter may be active during seed development and/or during germination. The seed specific promoter may be endosperm/aleurone/embryo specific. Examples of seed-specific promoters (endosperm/aleurone/embryo specific) are shown in Table 2c to Table 2f below. Further examples of seed-specific promoters are given in Qing Qu and Takaiwa (Plant Biotechnol. J. 2, 113-125, 2004), which disclosure is incorporated by reference herein as if fully set forth.

Table 2c: Examples of seed-specific promoters

Gene source	Reference
seed-specific genes	Simon et al., Plant Mol. Biol. 5: 191, 1985;
	Scofield et al., J. Biol. Chem. 262: 12202, 1987.;
	Baszczynski et al., Plant Mol. Biol. 14: 633, 1990.
Brazil Nut albumin	Pearson et al., Plant Mol. Biol. 18: 235-245, 1992.
legumin	Ellis et al., Plant Mol. Biol. 10: 203-214, 1988.
glutelin (rice)	Takaiwa et al., Mol. Gen. Genet. 208: 15-22, 1986;
	Takaiwa et al., FEBS Letts. 221: 43-47, 1987.
zein	Matzke et al Plant Mol Biol, 14(3):323-32 1990
napA	Stalberg et al, Planta 199: 515-519, 1996.
wheat LMW and HMW glutenin-1	Mol Gen Genet 216:81-90, 1989; NAR 17:461-2, 1989
wheat SPA	Albani et al, Plant Cell, 9: 171-184, 1997
wheat α, β, γ-gliadins	EMBO J. 3:1409-15, 1984
barley ltr1 promoter	Diaz et al. (1995) Mol Gen Genet 248(5):592-8
barley B1, C, D, hordein	Theor Appl Gen 98:1253-62, 1999; Plant J 4:343-55, 1993; Mol Gen Genet 250:750-60, 1996
barley DOF	Mena et al, The Plant Journal, 116(1): 53-62, 1998
blz2	EP99106056.7
synthetic promoter	Vicente-Carbajosa et al., Plant J. 13: 629-640, 1998.
rice prolamin NRP33	Wu et al, Plant Cell Physiology 39(8) 885-889, 1998
rice a-globulin Glb-1	Wu et al, Plant Cell Physiology 39(8) 885-889, 1998
rice OSH1	Sato et al, Proc. Natl. Acad. Sci. USA, 93: 8117-8122, 1996
rice α-globulin REB/OHP-1	Nakase et al. Plant Mol. Biol. 33: 513-522, 1997
rice ADP-glucose pyrophos-phorylase	Trans Res 6:157-68, 1997
maize ESR gene family	Plant J 12:235-46, 1997
sorghum α-kafirin	DeRose et al., Plant Mol. Biol 32:1029-35, 1996
KNOX	Postma-Haarsma et al, Plant Mol. Biol. 39:257-71, 1999
rice oleosin	Wu et al, J. Biochem. 123:386, 1998

sunflower oleosin	Cummins et al., Plant Mol. Biol. 19: 873-876, 1992
PRO0117, putative rice 40S ribosomal protein	WO 2004/070039
PRO0136, rice alanine aminotransferase	unpublished
PRO0147, trypsin inhibitor ITR1 (barley)	unpublished
PRO0151, rice WSI18	WO 2004/070039
PRO0175, rice RAB21	WO 2004/070039
PRO005	WO 2004/070039
PRO0095	WO 2004/070039
α-amylase (Amy32b)	Lanahan et al, Plant Cell 4:203-211, 1992; Skriver et al, Proc Natl Acad Sci USA 88:7266-7270, 1991
cathepsin β-like gene	Cejudo et al, Plant Mol Biol 20:849-856, 1992
Barley Ltp2	Kalla et al., Plant J. 6:849-60, 1994
Chi26	Leah et al., Plant J. 4:579-89, 1994
Maize B-Peru	Selinger et al., Genetics 149;1125-38,1998

Table 2d: examples of endosperm-specific promoters

Gene source	Reference
glutelin (rice)	Takaiwa et al. (1986) Mol Gen Genet 208:15-22; Takaiwa et al. (1987) FEBS Letts. 221:43-47
zein	Matzke et al., (1990) Plant Mol Biol 14(3): 323-32
wheat LMW and HMW glutenin-1	Colot et al. (1989) Mol Gen Genet 216:81-90, Anderson et al. (1989) NAR 17:461-2
wheat SPA	Albani et al. (1997) Plant Cell 9:171-184
wheat gliadins	Rafalski et al. (1984) EMBO 3:1409-15
barley Itr1 promoter	Diaz et al. (1995) Mol Gen Genet 248(5):592-8
barley B1, C, D, hordein	Cho et al. (1999) Theor Appl Genet 98:1253-62; Muller et al. (1993) Plant J 4:343-55; Sorenson et al. (1996) Mol Gen Genet 250:750-60
barley DOF	Mena et al, (1998) Plant J 116(1): 53-62
blz2	Onate et al. (1999) J Biol Chem 274(14):9175-82
synthetic promoter	Vicente-Carbajosa et al. (1998) Plant J 13:629-640
rice prolamin NRP33	Wu et al, (1998) Plant Cell Physiol 39(8) 885-889
rice globulin Glb-1	Wu et al. (1998) Plant Cell Physiol 39(8) 885-889
rice globulin REB/OHP-1	Nakase et al. (1997) Plant Molec Biol 33: 513-522
rice ADP-glucose pyrophosphorylase	Russell et al. (1997) Trans Res 6:157-68
maize ESR gene family	Opsahl-Ferstad et al. (1997) Plant J 12:235-46
sorghum kafirin	DeRose et al. (1996) Plant Mol Biol 32:1029-35

Table 2e: Examples of embryo specific promoters:

Gene source	Reference
rice OSH1	Sato et al, Proc. Natl. Acad. Sci. USA, 93: 8117-8122, 1996
KNOX	Postma-Haarsma et al, Plant Mol. Biol. 39:257-71, 1999
PRO0151	WO 2004/070039
PRO0175	WO 2004/070039
PRO005	WO 2004/070039
PRO0095	WO 2004/070039

Table 2f: Examples of aleurone-specific promoters:

Gene source	Reference
α-amylase (Amy32b)	Lanahan et al, Plant Cell 4:203-211, 1992; Skriver et al, Proc Natl Acad Sci USA 88:7266-7270, 1991
cathepsin β-like gene	Cejudo et al, Plant Mol Biol 20:849-856, 1992
Barley Ltp2	Kalla et al., Plant J. 6:849-60, 1994
Chi26	Leah et al., Plant J. 4:579-89, 1994
Maize B-Peru	Selinger et al., Genetics 149;1125-38,1998

A "green tissue-specific promoter" as defined herein is a promoter that is transcriptionally active predominantly in green tissue, substantially to the exclusion of any other parts of a plant, whilst still allowing for any leaky expression in these other plant parts.

Examples of green tissue-specific promoters which may be used to perform the methods of the invention are shown in Table 2g below.

Table 2g: Examples of green tissue-specific promoters

Gene	Expression	Reference
Maize Orthophosphate dikinase	Leaf specific	Fukavama et al., Plant Physiol. 2001 Nov;127(3):1136-46
Maize Phosphoenolpyruvate carboxylase	Leaf specific	Kausch et al., Plant Mol Biol. 2001 Jan;45(1):1-15
Rice Phosphoenolpyruvate carboxylase	Leaf specific	Lin et al., 2004 DNA Seq. 2004 Aug;15(4):269-76
Rice small subunit Rubisco	Leaf specific	Nomura et al., Plant Mol Biol. 2000 Sep;44(1):99-106
rice beta expansin EXBP9	Shoot specific	WO 2004/070039
Pigeonpea small subunit Rubisco	Leaf specific	Panguluri et al., Indian J Exp Biol. 2005 Apr;43(4):369-72
Pea RBCS3A	Leaf specific	

Another example of a tissue-specific promoter is a meristem-specific promoter, which is transcriptionally active predominantly in meristematic tissue, substantially to the exclusion

of any other parts of a plant, whilst still allowing for any leaky expression in these other plant parts. Examples of green meristem-specific promoters which may be used to perform the methods of the invention are shown in Table 2h below.

Table 2h: Examples of meristem-specific promoters

Gene source	Expression pattern	Reference
rice OSH1	Shoot apical meristem, from embryo globular stage to seedling stage	Sato *et al.* (1996) Proc. Natl. Acad. Sci. USA, 93: 8117-8122
Rice metallothionein	Meristem specific	BAD87835.1
WAK1 & WAK 2	Shoot and root apical meristems, and in expanding leaves and sepals	Wagner & Kohorn (2001) Plant Cell 13(2): 303–318

Terminator

The term “terminator” encompasses a control sequence which is a DNA sequence at the end of a transcriptional unit which signals 3’ processing and polyadenylation of a primary transcript and termination of transcription. The terminator can be derived from the natural gene, from a variety of other plant genes, or from T-DNA. The terminator to be added may be derived from, for example, the nopaline synthase or octopine synthase genes, or alternatively from another plant gene, or less preferably from any other eukaryotic gene.

Selectable marker (gene)/Reporter gene

"Selectable marker", "selectable marker gene" or “reporter gene” includes any gene that confers a phenotype on a cell in which it is expressed to facilitate the identification and/or selection of cells that are transfected or transformed with a nucleic acid construct of the invention. These marker genes enable the identification of a successful transfer of the nucleic acid molecules via a series of different principles. Suitable markers may be selected from markers that confer antibiotic or herbicide resistance, that introduce a new metabolic trait or that allow visual selection. Examples of selectable marker genes include genes conferring resistance to antibiotics (such as nptII that phosphorylates neomycin and kanamycin, or hpt, phosphorylating hygromycin, or genes conferring resistance to, for example, bleomycin, streptomycin, tetracyclin, chloramphenicol, ampicillin, gentamycin, geneticin (G418), spectinomycin or blasticidin), to herbicides (for example bar which provides resistance to Basta®; aroA or gox providing resistance against glyphosate, or the genes conferring resistance to, for example, imidazolinone, phosphinothricin or sulfonylurea), or genes that provide a metabolic trait (such as manA that allows plants to use mannose as sole carbon source or xylose isomerase for the utilisation of xylose, or antinutritive markers such as the resistance to 2-deoxyglucose). Expression of visual marker genes results in the formation of colour (for example β-glucuronidase, GUS or β-galactosidase with its coloured substrates, for example X-Gal), luminescence (such as the

luciferin/luceferase system) or fluorescence (Green Fluorescent Protein, GFP, and derivatives thereof). This list represents only a small number of possible markers. The skilled worker is familiar with such markers. Different markers are preferred, depending on the organism and the selection method.

It is known that upon stable or transient integration of nucleic acids into plant cells, only a minority of the cells takes up the foreign DNA and, if desired, integrates it into its genome, depending on the expression vector used and the transfection technique used. To identify and select these integrants, a gene coding for a selectable marker (such as the ones described above) is usually introduced into the host cells together with the gene of interest. These markers can for example be used in mutants in which these genes are not functional by, for example, deletion by conventional methods. Furthermore, nucleic acid molecules encoding a selectable marker can be introduced into a host cell on the same vector that comprises the sequence encoding the polypeptides of the invention or used in the methods of the invention, or else in a separate vector. Cells which have been stably transfected with the introduced nucleic acid can be identified for example by selection (for example, cells which have integrated the selectable marker survive whereas the other cells die).

Since the marker genes, particularly genes for resistance to antibiotics and herbicides, are no longer required or are undesired in the transgenic host cell once the nucleic acids have been introduced successfully, the process according to the invention for introducing the nucleic acids advantageously employs techniques which enable the removal or excision of these marker genes. One such a method is what is known as co-transformation. The co-transformation method employs two vectors simultaneously for the transformation, one vector bearing the nucleic acid according to the invention and a second bearing the marker gene(s). A large proportion of transformants receives or, in the case of plants, comprises (up to 40% or more of the transformants), both vectors. In case of transformation with Agrobacteria, the transformants usually receive only a part of the vector, i.e. the sequence flanked by the T-DNA, which usually represents the expression cassette. The marker genes can subsequently be removed from the transformed plant by performing crosses. In another method, marker genes integrated into a transposon are used for the transformation together with desired nucleic acid (known as the Ac/Ds technology). The transformants can be crossed with a transposase source or the transformants are transformed with a nucleic acid construct conferring expression of a transposase, transiently or stable. In some cases (approx. 10%), the transposon jumps out of the genome of the host cell once transformation has taken place successfully and is lost. In a further number of cases, the transposon jumps to a different location. In these cases the marker gene must be eliminated by performing crosses. In microbiology, techniques were developed which make possible, or facilitate, the detection of such events. A further advantageous method relies on what is known as recombination systems; whose advantage is that elimination by crossing can be dispensed with. The best-known system of this type is what is known as the Cre/lox system. Cre1 is a recombinase that removes the sequences located between the loxP sequences. If the

marker gene is integrated between the loxP sequences, it is removed once transformation has taken place successfully, by expression of the recombinase. Further recombination systems are the HIN/HIX, FLP/FRT and REP/STB system (Tribble et al., J. Biol. Chem., 275, 2000: 22255-22267; Velmurugan et al., J. Cell Biol., 149, 2000: 553-566). A site-specific integration into the plant genome of the nucleic acid sequences according to the invention is possible. Naturally, these methods can also be applied to microorganisms such as yeast, fungi or bacteria.

Transgenic/Transgene/Recombinant

For the purposes of the invention, "transgenic", “transgene” or "recombinant" means with regard to, for example, a nucleic acid sequence, an expression cassette, gene construct or a vector comprising the nucleic acid sequence or an organism transformed with the nucleic acid sequences, expression cassettes or vectors according to the invention, all those constructions brought about by recombinant methods in which either

(a) the nucleic acid sequences encoding proteins useful in the methods of the invention, or

(b) genetic control sequence(s) which is operably linked with the nucleic acid sequence according to the invention, for example a promoter, or

(c) a) and b)

are not located in their natural genetic environment or have been modified by recombinant methods, it being possible for the modification to take the form of, for example, a substitution, addition, deletion, inversion or insertion of one or more nucleotide residues. The natural genetic environment is understood as meaning the natural genomic or chromosomal locus in the original plant or the presence in a genomic library. In the case of a genomic library, the natural genetic environment of the nucleic acid sequence is preferably retained, at least in part. The environment flanks the nucleic acid sequence at least on one side and has a sequence length of at least 50 bp, preferably at least 500 bp, especially preferably at least 1000 bp, most preferably at least 5000 bp. A naturally occurring expression cassette – for example the naturally occurring combination of the natural promoter of the nucleic acid sequences with the corresponding nucleic acid sequence encoding a polypeptide useful in the methods of the present invention, as defined above – becomes a transgenic expression cassette when this expression cassette is modified by non-natural, synthetic ("artificial") methods such as, for example, mutagenic treatment. Suitable methods are described, for example, in US 5,565,350 or WO 00/15815.

A transgenic plant for the purposes of the invention is thus understood as meaning, as above, that the nucleic acids used in the method of the invention are not present in, or originating from, the genome of said plant, or are present in the genome of said plant but not at their natural locus in the genome of said plant, it being possible for the nucleic acids to be expressed homologously or heterologously. However, as mentioned, transgenic also means that, while the nucleic acids according to the invention or used in the inventive method are at their natural position in the genome of a plant, the sequence has been

modified with regard to the natural sequence, and/or that the regulatory sequences of the natural sequences have been modified. Transgenic is preferably understood as meaning the expression of the nucleic acids according to the invention at an unnatural locus in the genome, i.e. homologous or, preferably, heterologous expression of the nucleic acids takes place. Preferred transgenic plants are mentioned herein.

It shall further be noted that in the context of the present invention, the term "isolated nucleic acid" or "isolated polypeptide" may in some instances be considered as a synonym for a "recombinant nucleic acid" or a "recombinant polypeptide", respectively and refers to a nucleic acid or polypeptide that is not located in its natural genetic environment and/or that has been modified by recombinant methods.

Modulation

The term "modulation" means in relation to expression or gene expression, a process in which the expression level is changed by said gene expression in comparison to the control plant, the expression level may be increased or decreased. The original, unmodulated expression may be of any kind of expression of a structural RNA (rRNA, tRNA) or mRNA with subsequent translation. For the purposes of this invention, the original unmodulated expression may also be absence of any expression. The term "modulating the activity" shall mean any change of the expression of the inventive nucleic acid sequences or encoded proteins, which leads to increased yield and/or increased growth of the plants. The expression can increase from zero (absence of, or immeasurable expression) to a certain amount, or can decrease from a certain amount to immeasurable small amounts or zero.

Expression

The term "expression" or "gene expression" means the transcription of a specific gene or specific genes or specific genetic construct. The term "expression" or "gene expression" in particular means the transcription of a gene or genes or genetic construct into structural RNA (rRNA, tRNA) or mRNA with or without subsequent translation of the latter into a protein. The process includes transcription of DNA and processing of the resulting mRNA product.

Increased expression/overexpression

The term "increased expression" or "overexpression" as used herein means any form of expression that is additional to the original wild-type expression level. For the purposes of this invention, the original wild-type expression level might also be zero, i.e. absence of expression or immeasurable expression.

Methods for increasing expression of genes or gene products are well documented in the art and include, for example, overexpression driven by appropriate promoters, the use of transcription enhancers or translation enhancers. Isolated nucleic acids which serve as promoter or enhancer elements may be introduced in an appropriate position (typically

upstream) of a non-heterologous form of a polynucleotide so as to upregulate expression of a nucleic acid encoding the polypeptide of interest. For example, endogenous promoters may be altered in vivo by mutation, deletion, and/or substitution (see, Kmiec, US 5,565,350; Zarling et al., WO9322443), or isolated promoters may be introduced into a plant cell in the proper orientation and distance from a gene of the present invention so as to control the expression of the gene.

If polypeptide expression is desired, it is generally desirable to include a polyadenylation region at the 3'-end of a polynucleotide coding region. The polyadenylation region can be derived from the natural gene, from a variety of other plant genes, or from T-DNA. The 3' end sequence to be added may be derived from, for example, the nopaline synthase or octopine synthase genes, or alternatively from another plant gene, or less preferably from any other eukaryotic gene.

An intron sequence may also be added to the 5' untranslated region (UTR) or the coding sequence of the partial coding sequence to increase the amount of the mature message that accumulates in the cytosol. Inclusion of a spliceable intron in the transcription unit in both plant and animal expression constructs has been shown to increase gene expression at both the mRNA and protein levels up to 1000-fold (Buchman and Berg (1988) Mol. Cell biol. 8: 4395-4405; Callis et al. (1987) Genes Dev 1:1183-1200). Such intron enhancement of gene expression is typically greatest when placed near the 5' end of the transcription unit. Use of the maize introns Adh1-S intron 1, 2, and 6, the Bronze-1 intron are known in the art. For general information see: The Maize Handbook, Chapter 116, Freeling and Walbot, Eds., Springer, N.Y. (1994).

Decreased expression

Reference herein to “decreased expression” or “reduction or substantial elimination” of expression is taken to mean a decrease in endogenous gene expression and/or polypeptide levels and/or polypeptide activity relative to control plants. The reduction or substantial elimination is in increasing order of preference at least 10%, 20%, 30%, 40% or 50%, 60%, 70%, 80%, 85%, 90%, or 95%, 96%, 97%, 98%, 99% or more reduced compared to that of control plants.

For the reduction or substantial elimination of expression an endogenous gene in a plant, a sufficient length of substantially contiguous nucleotides of a nucleic acid sequence is required. In order to perform gene silencing, this may be as little as 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10 or fewer nucleotides, alternatively this may be as much as the entire gene (including the 5’ and/or 3’ UTR, either in part or in whole). The stretch of substantially contiguous nucleotides may be derived from the nucleic acid encoding the protein of interest (target gene), or from any nucleic acid capable of encoding an orthologue, paralogue or homologue of the protein of interest. Preferably, the stretch of substantially contiguous nucleotides is capable of forming hydrogen bonds with the target gene (either

sense or antisense strand), more preferably, the stretch of substantially contiguous nucleotides has, in increasing order of preference, 50%, 60%, 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 100% sequence identity to the target gene (either sense or antisense strand). A nucleic acid sequence encoding a (functional) polypeptide is not a requirement for the various methods discussed herein for the reduction or substantial elimination of expression of an endogenous gene.

This reduction or substantial elimination of expression may be achieved using routine tools and techniques. A preferred method for the reduction or substantial elimination of endogenous gene expression is by introducing and expressing in a plant a genetic construct into which the nucleic acid (in this case a stretch of substantially contiguous nucleotides derived from the gene of interest, or from any nucleic acid capable of encoding an orthologue, paralogue or homologue of any one of the protein of interest) is cloned as an inverted repeat (in part or completely), separated by a spacer (non-coding DNA).

In such a preferred method, expression of the endogenous gene is reduced or substantially eliminated through RNA-mediated silencing using an inverted repeat of a nucleic acid or a part thereof (in this case a stretch of substantially contiguous nucleotides derived from the gene of interest, or from any nucleic acid capable of encoding an orthologue, paralogue or homologue of the protein of interest), preferably capable of forming a hairpin structure. The inverted repeat is cloned in an expression vector comprising control sequences. A non-coding DNA nucleic acid sequence (a spacer, for example a matrix attachment region fragment (MAR), an intron, a polylinker, etc.) is located between the two inverted nucleic acids forming the inverted repeat. After transcription of the inverted repeat, a chimeric RNA with a self-complementary structure is formed (partial or complete). This double-stranded RNA structure is referred to as the hairpin RNA (hpRNA). The hpRNA is processed by the plant into siRNAs that are incorporated into an RNA-induced silencing complex (RISC). The RISC further cleaves the mRNA transcripts, thereby substantially reducing the number of mRNA transcripts to be translated into polypeptides. For further general details see for example, Grierson et al. (1998) WO 98/53083; Waterhouse et al. (1999) WO 99/53050).

Performance of the methods of the invention does not rely on introducing and expressing in a plant a genetic construct into which the nucleic acid is cloned as an inverted repeat, but any one or more of several well-known “gene silencing” methods may be used to achieve the same effects.

One such method for the reduction of endogenous gene expression is RNA-mediated silencing of gene expression (downregulation). Silencing in this case is triggered in a plant by a double stranded RNA sequence (dsRNA) that is substantially similar to the target endogenous gene. This dsRNA is further processed by the plant into about 20 to about 26 nucleotides called short interfering RNAs (siRNAs). The siRNAs are incorporated into an RNA-induced silencing complex (RISC) that cleaves the mRNA transcript of the

endogenous target gene, thereby substantially reducing the number of mRNA transcripts to be translated into a polypeptide. Preferably, the double stranded RNA sequence corresponds to a target gene.

Another example of an RNA silencing method involves the introduction of nucleic acid sequences or parts thereof (in this case a stretch of substantially contiguous nucleotides derived from the gene of interest, or from any nucleic acid capable of encoding an orthologue, paralogue or homologue of the protein of interest) in a sense orientation into a plant. “Sense orientation” refers to a DNA sequence that is homologous to an mRNA transcript thereof. Introduced into a plant would therefore be at least one copy of the nucleic acid sequence. The additional nucleic acid sequence will reduce expression of the endogenous gene, giving rise to a phenomenon known as co-suppression. The reduction of gene expression will be more pronounced if several additional copies of a nucleic acid sequence are introduced into the plant, as there is a positive correlation between high transcript levels and the triggering of co-suppression.

Another example of an RNA silencing method involves the use of antisense nucleic acid sequences. An "antisense" nucleic acid sequence comprises a nucleotide sequence that is complementary to a "sense" nucleic acid sequence encoding a protein, i.e. complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA transcript sequence. The antisense nucleic acid sequence is preferably complementary to the endogenous gene to be silenced. The complementarity may be located in the “coding region” and/or in the “non-coding region” of a gene. The term “coding region" refers to a region of the nucleotide sequence comprising codons that are translated into amino acid residues. The term “non-coding region" refers to 5' and 3' sequences that flank the coding region that are transcribed but not translated into amino acids (also referred to as 5' and 3' untranslated regions).

Antisense nucleic acid sequences can be designed according to the rules of Watson and Crick base pairing. The antisense nucleic acid sequence may be complementary to the entire nucleic acid sequence (in this case a stretch of substantially contiguous nucleotides derived from the gene of interest, or from any nucleic acid capable of encoding an orthologue, paralogue or homologue of the protein of interest), but may also be an oligonucleotide that is antisense to only a part of the nucleic acid sequence (including the mRNA 5’ and 3’ UTR). For example, the antisense oligonucleotide sequence may be complementary to the region surrounding the translation start site of an mRNA transcript encoding a polypeptide. The length of a suitable antisense oligonucleotide sequence is known in the art and may start from about 50, 45, 40, 35, 30, 25, 20, 15 or 10 nucleotides in length or less. An antisense nucleic acid sequence according to the invention may be constructed using chemical synthesis and enzymatic ligation reactions using methods known in the art. For example, an antisense nucleic acid sequence (e.g., an antisense oligonucleotide sequence) may 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 acid sequences, e.g., phosphorothioate derivatives and acridine substituted nucleotides may be used. Examples of modified nucleotides that may be used to generate the antisense nucleic acid sequences are well known in the art. Known nucleotide modifications include methylation, cyclization and 'caps' and substitution of one or more of the naturally occurring nucleotides with an analogue such as inosine. Other modifications of nucleotides are well known in the art.

The antisense nucleic acid sequence can be produced biologically using an expression vector into which a nucleic acid sequence 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). Preferably, production of antisense nucleic acid sequences in plants occurs by means of a stably integrated nucleic acid construct comprising a promoter, an operably linked antisense oligonucleotide, and a terminator.

The nucleic acid molecules used for silencing in the methods of the invention (whether introduced into a plant or generated in situ) hybridize with or bind to mRNA transcripts and/or genomic DNA encoding a polypeptide 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 sequence which binds to DNA duplexes, through specific interactions in the major groove of the double helix. Antisense nucleic acid sequences may be introduced into a plant by transformation or direct injection at a specific tissue site. Alternatively, antisense nucleic acid sequences can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense nucleic acid sequences can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, e.g., by linking the antisense nucleic acid sequence to peptides or antibodies which bind to cell surface receptors or antigens. The antisense nucleic acid sequences can also be delivered to cells using the vectors described herein.

According to a further aspect, the antisense nucleic acid sequence is an a-anomeric nucleic acid sequence. An a-anomeric nucleic acid sequence forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual b-units, the strands run parallel to each other (Gaultier et al. (1987) Nucl Ac Res 15: 6625-6641). The antisense nucleic acid sequence may also comprise a 2'-o-methylribonucleotide (Inoue et al. (1987) Nucl Ac Res 15, 6131-6148) or a chimeric RNA-DNA analogue (Inoue et al. (1987) FEBS Lett. 215, 327-330).

The reduction or substantial elimination of endogenous gene expression may also be performed using ribozymes. Ribozymes are catalytic RNA molecules with ribonuclease activity that are capable of cleaving a single-stranded nucleic acid sequence, such as an

mRNA, to which they have a complementary region. Thus, ribozymes (e.g., hammerhead ribozymes (described in Haselhoff and Gerlach (1988) Nature 334, 585-591) can be used to catalytically cleave mRNA transcripts encoding a polypeptide, thereby substantially reducing the number of mRNA transcripts to be translated into a polypeptide. A ribozyme having specificity for a nucleic acid sequence can be designed (see for example: Cech et al. U.S. Patent No. 4,987,071; and Cech et al. U.S. Patent No. 5,116,742). Alternatively, mRNA transcripts corresponding to a nucleic acid sequence can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules (Bartel and Szostak (1993) Science 261, 1411-1418). The use of ribozymes for gene silencing in plants is known in the art (e.g., Atkins et al. (1994) WO 94/00012; Lenne et al. (1995) WO 95/03404; Lutziger et al. (2000) WO 00/00619; Prinsen et al. (1997) WO 97/13865 and Scott et al. (1997) WO 97/38116).

Gene silencing may also be achieved by insertion mutagenesis (for example, T-DNA insertion or transposon insertion) or by strategies as described by, among others, Angell and Baulcombe ((1999) Plant J 20(3): 357-62), (Amplicon VIGS WO 98/36083), or Baulcombe (WO 99/15682).

Gene silencing may also occur if there is a mutation on an endogenous gene and/or a mutation on an isolated gene/nucleic acid subsequently introduced into a plant. The reduction or substantial elimination may be caused by a non-functional polypeptide. For example, the polypeptide may bind to various interacting proteins; one or more mutation(s) and/or truncation(s) may therefore provide for a polypeptide that is still able to bind interacting proteins (such as receptor proteins) but that cannot exhibit its normal function (such as signalling ligand).

A further approach to gene silencing is by targeting nucleic acid sequences complementary to the regulatory region of the gene (e.g., the promoter and/or enhancers) to form triple helical structures that prevent transcription of the gene in target cells. See Helene, C., Anticancer Drug Res. 6, 569-84, 1991; Helene et al., Ann. N.Y. Acad. Sci. 660, 27-36 1992; and Maher, L.J. Bioassays 14, 807-15, 1992.

Other methods, such as the use of antibodies directed to an endogenous polypeptide for inhibiting its function in planta, or interference in the signalling pathway in which a polypeptide is involved, will be well known to the skilled man. In particular, it can be envisaged that manmade molecules may be useful for inhibiting the biological function of a target polypeptide, or for interfering with the signalling pathway in which the target polypeptide is involved.

Alternatively, a screening program may be set up to identify in a plant population natural variants of a gene, which variants encode polypeptides with reduced activity. Such natural variants may also be used for example, to perform homologous recombination.

Artificial and/or natural microRNAs (miRNAs) may be used to knock out gene expression and/or mRNA translation. Endogenous miRNAs are single stranded small RNAs of typically 19-24 nucleotides long. They function primarily to regulate gene expression and/ or mRNA translation. Most plant microRNAs (miRNAs) have perfect or near-perfect complementarity with their target sequences. However, there are natural targets with up to five mismatches. They are processed from longer non-coding RNAs with characteristic fold-back structures by double-strand specific RNases of the Dicer family. Upon processing, they are incorporated in the RNA-induced silencing complex (RISC) by binding to its main component, an Argonaute protein. MiRNAs serve as the specificity components of RISC, since they base-pair to target nucleic acids, mostly mRNAs, in the cytoplasm. Subsequent regulatory events include target mRNA cleavage and destruction and/or translational inhibition. Effects of miRNA overexpression are thus often reflected in decreased mRNA levels of target genes.

Artificial microRNAs (amiRNAs), which are typically 21 nucleotides in length, can be genetically engineered specifically to negatively regulate gene expression of single or multiple genes of interest. Determinants of plant microRNA target selection are well known in the art. Empirical parameters for target recognition have been defined and can be used to aid in the design of specific amiRNAs, (Schwab et al., Dev. Cell 8, 517–527, 2005). Convenient tools for design and generation of amiRNAs and their precursors are also available to the public (Schwab et al., Plant Cell 18, 1121-1133, 2006).

For optimal performance, the gene silencing techniques used for reducing expression in a plant of an endogenous gene requires the use of nucleic acid sequences from monocotyledonous plants for transformation of monocotyledonous plants, and from dicotyledonous plants for transformation of dicotyledonous plants. Preferably, a nucleic acid sequence from any given plant species is introduced into that same species. For example, a nucleic acid sequence from rice is transformed into a rice plant. However, it is not an absolute requirement that the nucleic acid sequence to be introduced originates from the same plant species as the plant in which it will be introduced. It is sufficient that there is substantial homology between the endogenous target gene and the nucleic acid to be introduced.

Described above are examples of various methods for the reduction or substantial elimination of expression in a plant of an endogenous gene. A person skilled in the art would readily be able to adapt the aforementioned methods for silencing so as to achieve reduction of expression of an endogenous gene in a whole plant or in parts thereof through the use of an appropriate promoter, for example.

Transformation

The term "introduction" or "transformation" as referred to herein encompasses the transfer of an exogenous polynucleotide into a host cell, irrespective of the method used for transfer. Plant tissue capable of subsequent clonal propagation, whether by organogenesis or embryogenesis, may be transformed with a genetic construct of the present invention and a whole plant regenerated there from. The particular tissue chosen will vary depending on the clonal propagation systems available for, and best suited to, the particular species being transformed. Exemplary tissue targets include leaf disks, pollen, embryos, cotyledons, hypocotyls, megagametophytes, callus tissue, existing meristematic tissue (e.g., apical meristem, axillary buds, and root meristems), and induced meristem tissue (e.g., cotyledon meristem and hypocotyl meristem). The polynucleotide may be transiently or stably introduced into a host cell and may be maintained non-integrated, for example, as a plasmid. Alternatively, it may be integrated into the host genome. The resulting transformed plant cell may then be used to regenerate a transformed plant in a manner known to persons skilled in the art. Alternatively, a plant cell that cannot be regenerated into a plant may be chosen as host cell, i.e. the resulting transformed plant cell does not have the capacity to regenerate into a (whole) plant.

The transfer of foreign genes into the genome of a plant is called transformation. Transformation of plant species is now a fairly routine technique. Advantageously, any of several transformation methods may be used to introduce the gene of interest into a suitable ancestor cell. The methods described for the transformation and regeneration of plants from plant tissues or plant cells may be utilized for transient or for stable transformation. Transformation methods include the use of liposomes, electroporation, chemicals that increase free DNA uptake, injection of the DNA directly into the plant, particle gun bombardment, transformation using viruses or pollen and microprojection. Methods may be selected from the calcium/polyethylene glycol method for protoplasts (Krens, F.A. et al., (1982) Nature 296, 72-74; Negrutiu I et al. (1987) Plant Mol Biol 8: 363-373); electroporation of protoplasts (Shillito R.D. et al. (1985) Bio/Technol 3, 1099-1102); microinjection into plant material (Crossway A et al., (1986) Mol. Gen Genet 202: 179-185); DNA or RNA-coated particle bombardment (Klein TM et al., (1987) Nature 327: 70) infection with (non-integrative) viruses and the like. Transgenic plants, including transgenic crop plants, are preferably produced via *Agrobacterium*-mediated transformation. An advantageous transformation method is the transformation *in planta*. To this end, it is possible, for example, to allow the agrobacteria to act on plant seeds or to inoculate the plant meristem with agrobacteria. It has proved particularly expedient in accordance with the invention to allow a suspension of transformed agrobacteria to act on the intact plant or at least on the flower primordia. The plant is subsequently grown on until the seeds of the treated plant are obtained (Clough and Bent, Plant J. (1998) 16, 735–743). Methods for *Agrobacterium*-mediated transformation of rice include well known methods for rice transformation, such as those described in any of the following: European patent application EP 1198985 A1, Aldemita and Hodges (Planta 199: 612-617, 1996); Chan et al. (Plant Mol

Biol 22 (3): 491-506, 1993), Hiei et al. (Plant J 6 (2): 271-282, 1994), which disclosures are incorporated by reference herein as if fully set forth. In the case of corn transformation, the preferred method is as described in either Ishida et al. (Nat. Biotechnol 14(6): 745-50, 1996) or Frame et al. (Plant Physiol 129(1): 13-22, 2002), which disclosures are incorporated by reference herein as if fully set forth. Said methods are further described by way of example in B. Jenes et al., Techniques for Gene Transfer, in: Transgenic Plants, Vol. 1, Engineering and Utilization, eds. S.D. Kung and R. Wu, Academic Press (1993) 128-143 and in Potrykus Annu. Rev. Plant Physiol. Plant Molec. Biol. 42 (1991) 205-225). The nucleic acids or the construct to be expressed is preferably cloned into a vector, which is suitable for transforming *Agrobacterium tumefaciens*, for example pBin19 (Bevan et al., Nucl. Acids Res. 12 (1984) 8711). Agrobacteria transformed by such a vector can then be used in known manner for the transformation of plants, such as plants used as a model, like *Arabidopsis* (*Arabidopsis thaliana* is within the scope of the present invention not considered as a crop plant), or crop plants such as, by way of example, tobacco plants, for example by immersing bruised leaves or chopped leaves in an agrobacterial solution and then culturing them in suitable media. The transformation of plants by means of *Agrobacterium tumefaciens* is described, for example, by Höfgen and Willmitzer in Nucl. Acid Res. (1988) 16, 9877 or is known inter alia from F.F. White, Vectors for Gene Transfer in Higher Plants; in Transgenic Plants, Vol. 1, Engineering and Utilization, eds. S.D. Kung and R. Wu, Academic Press, 1993, pp. 15-38.

In addition to the transformation of somatic cells, which then have to be regenerated into intact plants, it is also possible to transform the cells of plant meristems and in particular those cells which develop into gametes. In this case, the transformed gametes follow the natural plant development, giving rise to transgenic plants. Thus, for example, seeds of *Arabidopsis* are treated with agrobacteria and seeds are obtained from the developing plants of which a certain proportion is transformed and thus transgenic [Feldman, KA and Marks MD (1987). Mol Gen Genet 208:1-9; Feldmann K (1992). In: C Koncz, N-H Chua and J Shell, eds, Methods in Arabidopsis Research. Word Scientific, Singapore, pp. 274-289]. Alternative methods are based on the repeated removal of the inflorescences and incubation of the excision site in the center of the rosette with transformed agrobacteria, whereby transformed seeds can likewise be obtained at a later point in time (Chang (1994). Plant J. 5: 551-558; Katavic (1994). Mol Gen Genet, 245: 363-370). However, an especially effective method is the vacuum infiltration method with its modifications such as the “floral dip” method. In the case of vacuum infiltration of *Arabidopsis*, intact plants under reduced pressure are treated with an agrobacterial suspension [Bechthold, N (1993). C R Acad Sci Paris Life Sci, 316: 1194-1199], while in the case of the ”floral dip” method the developing floral tissue is incubated briefly with a surfactant-treated agrobacterial suspension [Clough, SJ and Bent AF (1998) The Plant J. 16, 735-743]. A certain proportion of transgenic seeds are harvested in both cases, and these seeds can be distinguished from non-transgenic seeds by growing under the above-described selective conditions. In addition the stable transformation of plastids is of advantages because plastids are inherited maternally is most

crops reducing or eliminating the risk of transgene flow through pollen. The transformation of the chloroplast genome is generally achieved by a process which has been schematically displayed in Klaus et al., 2004 [Nature Biotechnology 22 (2), 225-229]. Briefly the sequences to be transformed are cloned together with a selectable marker gene between flanking sequences homologous to the chloroplast genome. These homologous flanking sequences direct site specific integration into the plastome. Plastidal transformation has been described for many different plant species and an overview is given in Bock (2001) Transgenic plastids in basic research and plant biotechnology. J Mol Biol. 2001 Sep 21; 312 (3):425-38 or Maliga, P (2003) Progress towards commercialization of plastid transformation technology. Trends Biotechnol. 21, 20-28. Further biotechnological progress has recently been reported in form of marker free plastid transformants, which can be produced by a transient co-integrated maker gene (Klaus et al., 2004, Nature Biotechnology 22(2), 225-229).

The genetically modified plant cells can be regenerated via all methods with which the skilled worker is familiar. Suitable methods can be found in the abovementioned publications by S.D. Kung and R. Wu, Potrykus or Höfgen and Willmitzer. Alternatively, the genetically modified plant cells are non-regenerable into a whole plant.

Generally after transformation, plant cells or cell groupings are selected for the presence of one or more markers which are encoded by plant-expressible genes co-transferred with the gene of interest, following which the transformed material is regenerated into a whole plant. To select transformed plants, the plant material obtained in the transformation is, as a rule, subjected to selective conditions so that transformed plants can be distinguished from untransformed plants. For example, the seeds obtained in the above-described manner can be planted and, after an initial growing period, subjected to a suitable selection by spraying. A further possibility consists in growing the seeds, if appropriate after sterilization, on agar plates using a suitable selection agent so that only the transformed seeds can grow into plants. Alternatively, the transformed plants are screened for the presence of a selectable marker such as the ones described above.

Following DNA transfer and regeneration, putatively transformed plants may also be evaluated, for instance using Southern analysis, for the presence of the gene of interest, copy number and/or genomic organisation. Alternatively or additionally, expression levels of the newly introduced DNA may be monitored using Northern and/or Western analysis, both techniques being well known to persons having ordinary skill in the art.

The generated transformed plants may be propagated by a variety of means, such as by clonal propagation or classical breeding techniques. For example, a first generation (or T1) transformed plant may be selfed and homozygous second-generation (or T2) transformants selected, and the T2 plants may then further be propagated through classical breeding techniques. The generated transformed organisms may take a variety of forms. For example, they may be chimeras of transformed cells and non-transformed cells; clonal

transformants (e.g., all cells transformed to contain the expression cassette); grafts of transformed and untransformed tissues (e.g., in plants, a transformed rootstock grafted to an untransformed scion).

T-DNA activation tagging

"T-DNA activation" tagging (Hayashi et al. Science (1992) 1350-1353), involves insertion of T-DNA, usually containing a promoter (may also be a translation enhancer or an intron), in the genomic region of the gene of interest or 10 kb up- or downstream of the coding region of a gene in a configuration such that the promoter directs expression of the targeted gene. Typically, regulation of expression of the targeted gene by its natural promoter is disrupted and the gene falls under the control of the newly introduced promoter. The promoter is typically embedded in a T-DNA. This T-DNA is randomly inserted into the plant genome, for example, through *Agrobacterium* infection and leads to modified expression of genes near the inserted T-DNA. The resulting transgenic plants show dominant phenotypes due to modified expression of genes close to the introduced promoter.

TILLING

The term "TILLING" is an abbreviation of "Targeted Induced Local Lesions In Genomes" and refers to a mutagenesis technology useful to generate and/or identify nucleic acids encoding proteins with modified expression and/or activity. TILLING also allows selection of plants carrying such mutant variants. These mutant variants may exhibit modified expression, either in strength or in location or in timing (if the mutations affect the promoter for example). These mutant variants may exhibit higher activity than that exhibited by the gene in its natural form. TILLING combines high-density mutagenesis with high-throughput screening methods. The steps typically followed in TILLING are: (a) EMS mutagenesis (Redei GP and Koncz C (1992) In Methods in Arabidopsis Research, Koncz C, Chua NH, Schell J, eds. Singapore, World Scientific Publishing Co, pp. 16–82; Feldmann et al., (1994) In Meyerowitz EM, Somerville CR, eds, Arabidopsis. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp 137-172; Lightner J and Caspar T (1998) In J Martinez-Zapater, J Salinas, eds, Methods on Molecular Biology, Vol. 82. Humana Press, Totowa, NJ, pp 91-104); (b) DNA preparation and pooling of individuals; (c) PCR amplification of a region of interest; (d) denaturation and annealing to allow formation of heteroduplexes; (e) DHPLC, where the presence of a heteroduplex in a pool is detected as an extra peak in the chromatogram; (f) identification of the mutant individual; and (g) sequencing of the mutant PCR product. Methods for TILLING are well known in the art (McCallum et al., (2000) Nat Biotechnol 18: 455-457; reviewed by Stemple (2004) Nat Rev Genet 5(2): 145-50).

Homologous recombination

"Homologous recombination" allows introduction in a genome of a selected nucleic acid at a defined selected position. Homologous recombination is a standard technology used routinely in biological sciences for lower organisms such as yeast or the moss *Physcomitrella*. Methods for performing homologous recombination in plants have been

described not only for model plants (Offringa et al. (1990) EMBO J 9(10): 3077-84) but also for crop plants, for example rice (Terada et al. (2002) Nat Biotech 20(10): 1030-4; Iida and Terada (2004) Curr Opin Biotech 15(2): 132-8), and approaches exist that are generally applicable regardless of the target organism (Miller et al, Nature Biotechnol. 25, 778-785, 2007).

Yield related Trait(s)

A "Yield related trait" is a trait or feature which is related to plant yield. Yield-related traits may comprise one or more of the following non-limitative list of features: early flowering time, yield, biomass, seed yield, early vigour, greenness index, growth rate, agronomic traits, such as e.g. tolerance to submergence (which leads to yield in rice), Water Use Efficiency (WUE), Nitrogen Use Efficiency (NUE), etc.

Reference herein to enhanced yield-related traits, relative to of control plants is taken to mean one or more of an increase in early vigour and/or in biomass (weight) of one or more parts of a plant, which may include (i) aboveground parts and preferably aboveground harvestable parts and/or (ii) parts below ground and preferably harvestable below ground. In particular, such harvestable parts are seeds.

Yield

The term "yield" in general means a measurable produce of economic value, typically related to a specified crop, to an area, and to a period of time. Individual plant parts directly contribute to yield based on their number, size and/or weight, or the actual yield is the yield per square meter for a crop and year, which is determined by dividing total production (includes both harvested and appraised production) by planted square meters.

The terms "yield" of a plant and "plant yield" are used interchangeably herein and are meant to refer to vegetative biomass such as root and/or shoot biomass, to reproductive organs, and/or to propagules such as seeds of that plant.

Flowers in maize are unisexual; male inflorescences (tassels) originate from the apical stem and female inflorescences (ears) arise from axillary bud apices. The female inflorescence produces pairs of spikelets on the surface of a central axis (cob). Each of the female spikelets encloses two fertile florets, one of them will usually mature into a maize kernel once fertilized. Hence a yield increase in maize may be manifested as one or more of the following: increase in the number of plants established per square meter, an increase in the number of ears per plant, an increase in the number of rows, number of kernels per row, kernel weight, thousand kernel weight, ear length/diameter, increase in the seed filling rate, which is the number of filled florets (i.e. florets containing seed) divided by the total number of florets and multiplied by 100), among others.

Inflorescences in rice plants are named panicles. The panicle bears spikelets, which are the basic units of the panicles, and which consist of a pedicel and a floret. The floret is borne on the pedicel and includes a flower that is covered by two protective glumes: a larger glume (the lemma) and a shorter glume (the palea). Hence, taking rice as an example, a yield increase may manifest itself as an increase in one or more of the following: number of plants per square meter, number of panicles per plant, panicle length, number of spikelets per panicle, number of flowers (or florets) per panicle; an increase in the seed filling rate which is the number of filled florets (i.e. florets containing seeds) divided by the total number of florets and multiplied by 100; an increase in thousand kernel weight, among others.

Early flowering time

Plants having an "early flowering time" as used herein are plants which start to flower earlier than control plants. Hence this term refers to plants that show an earlier start of flowering. Flowering time of plants can be assessed by counting the number of days ("time to flower") between sowing and the emergence of a first inflorescence. The "flowering time" of a plant can for instance be determined using the method as described in WO 2007/093444.

Early vigour

"Early vigour" refers to active healthy well-balanced growth especially during early stages of plant growth, and may result from increased plant fitness due to, for example, the plants being better adapted to their environment (i.e. optimizing the use of energy resources and partitioning between shoot and root). Plants having early vigour also show increased seedling survival and a better establishment of the crop, which often results in highly uniform fields (with the crop growing in uniform manner, i.e. with the majority of plants reaching the various stages of development at substantially the same time), and often better and higher yield. Therefore, early vigour may be determined by measuring various factors, such as thousand kernel weight, percentage germination, percentage emergence, seedling growth, seedling height, root length, root and shoot biomass and many more.

Increased growth rate

The increased growth rate may be specific to one or more parts of a plant (including seeds), or may be throughout substantially the whole plant. Plants having an increased growth rate may have a shorter life cycle. The life cycle of a plant may be taken to mean the time needed to grow from a mature seed up to the stage where the plant has produced mature seeds, similar to the starting material. This life cycle may be influenced by factors such as speed of germination, early vigour, growth rate, greenness index, flowering time and speed of seed maturation. The increase in growth rate may take place at one or more stages in the life cycle of a plant or during substantially the whole plant life cycle. Increased growth rate during the early stages in the life cycle of a plant may reflect enhanced vigour. The increase in growth rate may alter the harvest cycle of a plant allowing plants to be sown later and/or harvested sooner than would otherwise be possible (a similar effect may be

obtained with earlier flowering time). If the growth rate is sufficiently increased, it may allow for the further sowing of seeds of the same plant species (for example sowing and harvesting of rice plants followed by sowing and harvesting of further rice plants all within one conventional growing period). Similarly, if the growth rate is sufficiently increased, it may allow for the further sowing of seeds of different plants species (for example the sowing and harvesting of corn plants followed by, for example, the sowing and optional harvesting of soybean, potato or any other suitable plant). Harvesting additional times from the same rootstock in the case of some crop plants may also be possible. Altering the harvest cycle of a plant may lead to an increase in annual biomass production per square meter (due to an increase in the number of times (say in a year) that any particular plant may be grown and harvested). An increase in growth rate may also allow for the cultivation of transgenic plants in a wider geographical area than their wild-type counterparts, since the territorial limitations for growing a crop are often determined by adverse environmental conditions either at the time of planting (early season) or at the time of harvesting (late season). Such adverse conditions may be avoided if the harvest cycle is shortened. The growth rate may be determined by deriving various parameters from growth curves, such parameters may be: T-Mid (the time taken for plants to reach 50% of their maximal size) and T-90 (time taken for plants to reach 90% of their maximal size), amongst others.

Stress resistance

An increase in yield and/or growth rate occurs whether the plant is under non-stress conditions or whether the plant is exposed to various stresses compared to control plants. Plants typically respond to exposure to stress by growing more slowly. In conditions of severe stress, the plant may even stop growing altogether. Mild stress on the other hand is defined herein as being any stress to which a plant is exposed which does not result in the plant ceasing to grow altogether without the capacity to resume growth. Mild stress in the sense of the invention leads to a reduction in the growth of the stressed plants of less than 40%, 35%, 30% or 25%, more preferably less than 20% or 15% in comparison to the control plant under non-stress conditions. Due to advances in agricultural practices (irrigation, fertilization, pesticide treatments) severe stresses are not often encountered in cultivated crop plants. As a consequence, the compromised growth induced by mild stress is often an undesirable feature for agriculture. Abiotic stresses may be due to drought or excess water, anaerobic stress, salt stress, chemical toxicity, oxidative stress and hot, cold or freezing temperatures.

"Biotic stresses" are typically those stresses caused by pathogens, such as bacteria, viruses, fungi, nematodes and insects.

The "abiotic stress" may be an osmotic stress caused by a water stress, e.g. due to drought, salt stress, or freezing stress. Abiotic stress may also be an oxidative stress or a cold stress. "Freezing stress" is intended to refer to stress due to freezing temperatures, i.e. temperatures at which available water molecules freeze and turn into ice. "Cold stress", also

called “chilling stress”, is intended to refer to cold temperatures, e.g. temperatures below 10°, or preferably below 5°C, but at which water molecules do not freeze. As reported in Wang et al. (Planta (2003) 218: 1-14), abiotic stress leads to a series of morphological, physiological, biochemical and molecular changes that adversely affect plant growth and productivity. Drought, salinity, extreme temperatures and oxidative stress are known to be interconnected and may induce growth and cellular damage through similar mechanisms. Rabbani et al. (Plant Physiol (2003) 133: 1755-1767) describes a particularly high degree of “cross talk” between drought stress and high-salinity stress. For example, drought and/or salinisation are manifested primarily as osmotic stress, resulting in the disruption of homeostasis and ion distribution in the cell. Oxidative stress, which frequently accompanies high or low temperature, salinity or drought stress, may cause denaturing of functional and structural proteins. As a consequence, these diverse environmental stresses often activate similar cell signalling pathways and cellular responses, such as the production of stress proteins, up-regulation of anti-oxidants, accumulation of compatible solutes and growth arrest. The term “non-stress” conditions as used herein are those environmental conditions that allow optimal growth of plants. Persons skilled in the art are aware of normal soil conditions and climatic conditions for a given location. Plants with optimal growth conditions, (grown under non-stress conditions) typically yield in increasing order of preference at least 97%, 95%, 92%, 90%, 87%, 85%, 83%, 80%, 77% or 75% of the average production of such plant in a given environment. Average production may be calculated on harvest and/or season basis. Persons skilled in the art are aware of average yield productions of a crop.

In particular, the methods of the present invention may be performed under non-stress conditions. In an example, the methods of the present invention may be performed under non-stress conditions such as mild drought to give plants having increased yield relative to control plants.

In another embodiment, the methods of the present invention may be performed under stress conditions.

In an example, the methods of the present invention may be performed under stress conditions such as drought to give plants having increased yield relative to control plants.

In another example, the methods of the present invention may be performed under stress conditions such as nutrient deficiency to give plants having increased yield relative to control plants.

Nutrient deficiency may result from a lack of nutrients such as nitrogen, phosphates and other phosphorous-containing compounds, potassium, calcium, magnesium, manganese, iron and boron, amongst others.

In yet another example, the methods of the present invention may be performed under stress conditions such as salt stress to give plants having increased yield relative to control plants. The term salt stress is not restricted to common salt (NaCl), but may be any one or more of: NaCl, KCl, LiCl, $MgCl_2$, $CaCl_2$, amongst others.

In yet another example, the methods of the present invention may be performed under stress conditions such as cold stress or freezing stress to give plants having increased yield relative to control plants.

Increase/Improve/Enhance

The terms "increase", "improve" or "enhance" are interchangeable and shall mean in the sense of the application at least a 3%, 4%, 5%, 6%, 7%, 8%, 9% or 10%, preferably at least 15% or 20%, more preferably 25%, 30%, 35% or 40% more yield and/or growth in comparison to control plants as defined herein.

Seed yield

Increased seed yield may manifest itself as one or more of the following:

a) an increase in seed biomass (total seed weight) which may be on an individual seed basis and/or per plant and/or per square meter;
b) increased number of flowers per plant;
c) increased number of seeds;
d) increased seed filling rate (which is expressed as the ratio between the number of filled florets divided by the total number of florets);
e) increased harvest index, which is expressed as a ratio of the yield of harvestable parts, such as seeds, divided by the biomass of aboveground plant parts; and
f) increased thousand kernel weight (TKW), which is extrapolated from the number of seeds counted and their total weight. An increased TKW may result from an increased seed size and/or seed weight, and may also result from an increase in embryo and/or endosperm size.

The terms "filled florets" and "filled seeds" may be considered synonyms.

An increase in seed yield may also be manifested as an increase in seed size and/or seed volume. Furthermore, an increase in seed yield may also manifest itself as an increase in seed area and/or seed length and/or seed width and/or seed perimeter.

Greenness Index

The "greenness index" as used herein is calculated from digital images of plants. For each pixel belonging to the plant object on the image, the ratio of the green value versus the red value (in the RGB model for encoding color) is calculated. The greenness index is expressed as the percentage of pixels for which the green-to-red ratio exceeds a given threshold. Under normal growth conditions, under salt stress growth conditions, and under reduced nutrient availability growth conditions, the greenness index of plants is measured in the last imaging before flowering. In contrast, under drought stress growth conditions, the greenness index of plants is measured in the first imaging after drought.

Biomass

The term "biomass" as used herein is intended to refer to the total weight of a plant. Within the definition of biomass, a distinction may be made between the biomass of one or more parts of a plant, which may include any one or more of the following:

- aboveground parts such as but not limited to shoot biomass, seed biomass, leaf biomass, etc.;
- aboveground harvestable parts such as but not limited to shoot biomass, seed biomass, leaf biomass, etc.;
- parts below ground, such as but not limited to root biomass, tubers, bulbs, etc.;
- harvestable parts below ground, such as but not limited to root biomass, tubers, bulbs, etc.;
- harvestable parts partially below ground such as but not limited to beets and other hypocotyl areas of a plant, rhizomes, stolons or creeping rootstalks;
- vegetative biomass such as root biomass, shoot biomass, etc.;
- reproductive organs; and
- propagules such as seed.

Marker assisted breeding

Such breeding programmes sometimes require introduction of allelic variation by mutagenic treatment of the plants, using for example EMS mutagenesis; alternatively, the programme may start with a collection of allelic variants of so called "natural" origin caused unintentionally. Identification of allelic variants then takes place, for example, by PCR. This is followed by a step for selection of superior allelic variants of the sequence in question and which give increased yield. Selection is typically carried out by monitoring growth performance of plants containing different allelic variants of the sequence in question. Growth performance may be monitored in a greenhouse or in the field. Further optional steps include crossing plants in which the superior allelic variant was identified with another plant. This could be used, for example, to make a combination of interesting phenotypic features.

Use as probes in (gene mapping)

Use of nucleic acids encoding the protein of interest for genetically and physically mapping the genes requires only a nucleic acid sequence of at least 15 nucleotides in length. These nucleic acids may be used as restriction fragment length polymorphism (RFLP) markers. Southern blots (Sambrook J, Fritsch EF and Maniatis T (1989) Molecular Cloning, A Laboratory Manual) of restriction-digested plant genomic DNA may be probed with the nucleic acids encoding the protein of interest. The resulting banding patterns may then be subjected to genetic analyses using computer programs such as MapMaker (Lander et al. (1987) Genomics 1: 174-181) in order to construct a genetic map. In addition, the nucleic acids may be used to probe Southern blots containing restriction endonuclease-treated genomic DNAs of a set of individuals representing parent and progeny of a defined genetic cross. Segregation of the DNA polymorphisms is noted and used to calculate the position of

the nucleic acid encoding the protein of interest in the genetic map previously obtained using this population (Botstein et al. (1980) Am. J. Hum. Genet. 32:314-331).

The production and use of plant gene-derived probes for use in genetic mapping is described in Bernatzky and Tanksley (1986) Plant Mol. Biol. Reporter 4: 37-41. Numerous publications describe genetic mapping of specific cDNA clones using the methodology outlined above or variations thereof. For example, F2 intercross populations, backcross populations, randomly mated populations, near isogenic lines, and other sets of individuals may be used for mapping. Such methodologies are well known to those skilled in the art.

The nucleic acid probes may also be used for physical mapping (i.e., placement of sequences on physical maps; see Hoheisel et al. In: Non-mammalian Genomic Analysis: A Practical Guide, Academic press 1996, pp. 319-346, and references cited therein).

In another embodiment, the nucleic acid probes may be used in direct fluorescence in situ hybridisation (FISH) mapping (Trask (1991) Trends Genet. 7:149-154). Although current methods of FISH mapping favour use of large clones (several kb to several hundred kb; see Laan et al. (1995) Genome Res. 5:13-20), improvements in sensitivity may allow performance of FISH mapping using shorter probes.

A variety of nucleic acid amplification-based methods for genetic and physical mapping may be carried out using the nucleic acids. Examples include allele-specific amplification (Kazazian (1989) J. Lab. Clin. Med 11:95-96), polymorphism of PCR-amplified fragments (CAPS; Sheffield et al. (1993) Genomics 16:325-332), allele-specific ligation (Landegren et al. (1988) Science 241:1077-1080), nucleotide extension reactions (Sokolov (1990) Nucleic Acid Res. 18:3671), Radiation Hybrid Mapping (Walter et al. (1997) Nat. Genet. 7:22-28) and Happy Mapping (Dear and Cook (1989) Nucleic Acid Res. 17:6795-6807). For these methods, the sequence of a nucleic acid is used to design and produce primer pairs for use in the amplification reaction or in primer extension reactions. The design of such primers is well known to those skilled in the art. In methods employing PCR-based genetic mapping, it may be necessary to identify DNA sequence differences between the parents of the mapping cross in the region corresponding to the instant nucleic acid sequence. This, however, is generally not necessary for mapping methods.

Plant

The term “plant” as used herein encompasses whole plants, ancestors and progeny of the plants and plant parts, including seeds, shoots, stems, leaves, roots (including tubers), flowers, and tissues and organs, wherein each of the aforementioned comprise the gene/nucleic acid of interest. The term “plant” also encompasses plant cells, suspension cultures, callus tissue, embryos, meristematic regions, gametophytes, sporophytes, pollen and microspores, again wherein each of the aforementioned comprises the gene/nucleic acid of interest.

Plants that are particularly useful in the methods of the invention include all plants which belong to the superfamily Viridiplantae, in particular monocotyledonous and dicotyledonous plants including fodder or forage legumes, ornamental plants, food crops, trees or shrubs selected from the list comprising *Acer* spp., *Actinidia* spp., *Abelmoschus* spp., *Agave sisalana, Agropyron* spp., *Agrostis stolonifera, Allium* spp., *Amaranthus* spp., *Ammophila arenaria, Ananas comosus, Annona* spp., *Apium graveolens, Arachis spp, Artocarpus* spp., *Asparagus officinalis, Avena* spp. *(e.g. Avena sativa, Avena fatua, Avena byzantina, Avena fatua var. sativa, Avena hybrida), Averrhoa carambola, Bambusa sp., Benincasa hispida, Bertholletia excelsea, Beta vulgaris, Brassica* spp. *(e.g. Brassica napus, Brassica rapa* ssp. [canola, oilseed rape, turnip rape]), *Cadaba farinosa, Camellia sinensis, Canna indica, Cannabis sativa, Capsicum* spp., *Carex elata, Carica papaya, Carissa macrocarpa, Carya* spp., *Carthamus tinctorius, Castanea* spp., *Ceiba pentandra, Cichorium endivia, Cinnamomum* spp., *Citrullus lanatus, Citrus* spp., *Cocos* spp., *Coffea* spp., *Colocasia esculenta, Cola* spp., *Corchorus sp., Coriandrum sativum, Corylus* spp., *Crataegus* spp., *Crocus sativus, Cucurbita* spp., *Cucumis* spp., *Cynara* spp., *Daucus carota, Desmodium* spp., *Dimocarpus longan, Dioscorea* spp., *Diospyros* spp., *Echinochloa* spp., *Elaeis (e.g. Elaeis guineensis, Elaeis oleifera), Eleusine coracana, Eragrostis tef, Erianthus sp., Eriobotrya japonica, Eucalyptus sp., Eugenia uniflora, Fagopyrum* spp., *Fagus* spp., *Festuca arundinacea, Ficus carica, Fortunella* spp., *Fragaria* spp., *Ginkgo biloba, Glycine* spp. *(e.g. Glycine max, Soja hispida or Soja max), Gossypium hirsutum, Helianthus* spp. *(e.g. Helianthus annuus), Hemerocallis fulva, Hibiscus* spp., *Hordeum* spp. *(e.g. Hordeum vulgare), Ipomoea batatas, Juglans* spp., *Lactuca sativa, Lathyrus* spp., *Lens culinaris, Linum usitatissimum, Litchi chinensis, Lotus* spp., *Luffa acutangula, Lupinus* spp., *Luzula sylvatica, Lycopersicon* spp. *(e.g. Lycopersicon esculentum, Lycopersicon lycopersicum, Lycopersicon pyriforme), Macrotyloma* spp., *Malus* spp., *Malpighia emarginata, Mammea americana, Mangifera indica, Manihot* spp., *Manilkara zapota, Medicago sativa, Melilotus* spp., *Mentha* spp., *Miscanthus sinensis, Momordica* spp., *Morus nigra, Musa* spp., *Nicotiana* spp., *Olea* spp., *Opuntia* spp., *Ornithopus* spp., *Oryza* spp. *(e.g. Oryza sativa, Oryza latifolia), Panicum miliaceum, Panicum virgatum, Passiflora edulis, Pastinaca sativa, Pennisetum sp., Persea* spp., *Petroselinum crispum, Phalaris arundinacea, Phaseolus* spp., *Phleum pratense, Phoenix* spp., *Phragmites australis, Physalis* spp., *Pinus* spp., *Pistacia vera, Pisum* spp., *Poa* spp., *Populus* spp., *Prosopis* spp., *Prunus* spp., *Psidium* spp., *Punica granatum, Pyrus communis, Quercus* spp., *Raphanus sativus, Rheum rhabarbarum, Ribes* spp., *Ricinus communis, Rubus* spp., *Saccharum* spp., *Salix sp., Sambucus* spp., *Secale cereale, Sesamum* spp., *Sinapis sp., Solanum* spp. *(e.g. Solanum tuberosum, Solanum integrifolium or Solanum lycopersicum), Sorghum bicolor, Spinacia* spp., *Syzygium* spp., *Tagetes* spp., *Tamarindus indica, Theobroma cacao, Trifolium* spp., *Tripsacum dactyloides, Triticosecale rimpaui, Triticum* spp. *(e.g. Triticum aestivum, Triticum durum, Triticum turgidum, Triticum hybernum, Triticum macha, Triticum sativum, Triticum monococcum or Triticum vulgare), Tropaeolum minus, Tropaeolum majus, Vaccinium* spp.,

Vicia spp., *Vigna* spp., *Viola odorata, Vitis* spp., *Zea mays, Zizania palustris, Ziziphus* spp., amongst others.

Control plant(s)

The choice of suitable control plants is a routine part of an experimental setup and may include corresponding wild type plants or corresponding plants without the gene of interest. The control plant is typically of the same plant species or even of the same variety as the plant to be assessed. The control plant may also be a nullizygote of the plant to be assessed. Nullizygotes (or null control plants) are individuals missing the transgene by segregation. Further, control plants are grown under equal growing conditions to the growing conditions of the plants of the invention, i.e. in the vicinity of, and simultaneously with, the plants of the invention. A "control plant" as used herein refers not only to whole plants, but also to plant parts, including seeds and seed parts.

Description of figures

The present invention will now be described with reference to the following figures in which:

Fig. 1 represents the domain structure of SEQ ID NO: 2 with conserved motifs 1 to 3.

Fig. 2 represents a multiple alignment of various PRR-like polypeptides. These alignments can be used for defining further motifs or signature sequences, when using conserved amino acids.

The corresponding SEQ ID NO's for the aligned polypeptide sequences shown in Fig. 2 are:

SEQ ID NO: 4 for C.maculosa_TA1222_215693
SEQ ID NO: 6 for L.sativa_TC22097
SEQ ID NO: 8 for C.endivia_TA960_114280
SEQ ID NO: 10 for S.officinarum_TC81972
SEQ ID NO: 2 for S. officinarum _PRR-like
SEQ ID NO: 12 for S.bicolor_Sb01g038820.1
SEQ ID NO: 14 for Z.mays_GRMZM2G095727_T03
SEQ ID NO: 16 for Z.mays_GRMZM2G095727_T02
SEQ ID NO: 18 for Z.mays_GRMZM2G095727_T04
SEQ ID NO: 20 for Z.mays_GRMZM2G095727_T05
SEQ ID NO: 22 for Z.mays_GRMZM2G095727_T01
SEQ ID NO: 24 for S.officinarum_TC80591
SEQ ID NO: 26 for P.virgatum_TC27226
SEQ ID NO: 28 for S.officinarum_TC111093
SEQ ID NO: 30 for S.officinarum_TC84822
SEQ ID NO: 32 for P.virgatum_TC42896
SEQ ID NO: 34 for P.virgatum_TC22980
SEQ ID NO: 36 for T.aestivum_TC284871
SEQ ID NO: 38 for H.vulgare_TC165635
SEQ ID NO: 40 for T.aestivum_c59844892@6464
SEQ ID NO: 42 for B.distachyon_DV469398

Fig. 3 shows the MATGAT table of Example 3.

Fig. 4 represents the binary vector used for increased expression in Oryza sativa of a PRR-like-encoding nucleic acid under the control of a rice GOS2 promoter (pGOS2).

Examples

The present invention will now be described with reference to the following examples, which are by way of illustration only. The following examples are not intended to limit the scope of the invention. Unless otherwise indicated, the present invention employs conventional techniques and methods of plant biology, molecular biology, bioinformatics and plant breedings.

DNA manipulation: unless otherwise stated, recombinant DNA techniques are performed according to standard protocols described in Sambrook (2001) Molecular Cloning: a laboratory manual, 3rd Edition Cold Spring Harbor Laboratory Press, CSH, New York or in Volumes 1 and 2 of Ausubel et al. (1994), Current Protocols in Molecular Biology, Current Protocols. Standard materials and methods for plant molecular work are described in Plant Molecular Biology Labfax (1993) by R.D.D. Croy, published by BIOS Scientific Publications Ltd (UK) and Blackwell Scientific Publications (UK).

Example 1: Identification of sequences related to SEQ ID NO: 1 and SEQ ID NO: 2

Sequences (full length cDNA, ESTs or genomic) related to SEQ ID NO: 1 and SEQ ID NO: 2 were identified amongst those maintained in the Entrez Nucleotides database at the National Center for Biotechnology Information (NCBI) using database sequence search tools, such as the Basic Local Alignment Tool (BLAST) (Altschul et al. (1990) J. Mol. Biol. 215:403-410; and Altschul et al. (1997) Nucleic Acids Res. 25:3389-3402). The program is used to find regions of local similarity between sequences by comparing nucleic acid or polypeptide sequences to sequence databases and by calculating the statistical significance of matches. For example, the polypeptide encoded by the nucleic acid of SEQ ID NO: 1 was used for the TBLASTN algorithm, with default settings and the filter to ignore low complexity sequences set off. The output of the analysis was viewed by pairwise comparison, and ranked according to the probability score (E-value), where the score reflect the probability that a particular alignment occurs by chance (the lower the E-value, the more significant the hit). In addition to E-values, comparisons were also scored by percentage identity. Percentage identity refers to the number of identical nucleotides (or amino acids) between the two compared nucleic acid (or polypeptide) sequences over a particular length. In some instances, the default parameters may be adjusted to modify the stringency of the search. For example the E-value may be increased to show less stringent matches. This way, short nearly exact matches may be identified.

Table A provides a list of nucleic acid sequences related to SEQ ID NO: 1 and SEQ ID NO: 2.

Table A: Examples of PRR-like nucleic acids and polypeptides:

Plant Source	Nucleic acid SEQ ID NO:	Protein SEQ ID NO:
Saccharum_officinarum	1	2
Centaurea_maculosa	3	4
Lactuca_sativa	5	6
Cichorium_endivia	7	8
Saccharum_officinarum	9	10
Sorghum_bicolor	11	12
Zea_mays	13	14
Zea_mays	15	16
Zea_mays	17	18
Zea_mays	19	20
Zea_mays	21	22
Saccharum_officinarum	23	24
Panicum_virgatum	25	26
Saccharum_officinarum	27	28
Saccharum_officinarum	29	30
Panicum_virgatum	31	32
Panicum_virgatum	33	34
Triticum_aestivum	35	36
Hordeum_vulgare	37	38
Triticum_aestivum	39	40
Brachypodium_distachyon	41	42

Sequences have been tentatively assembled and publicly disclosed by research institutions, such as The Institute for Genomic Research (TIGR; beginning with TA). For instance, the Eukaryotic Gene Orthologs (EGO) database may be used to identify such related sequences, either by keyword search or by using the BLAST algorithm with the nucleic acid sequence or polypeptide sequence of interest. Special nucleic acid sequence databases have been created for particular organisms, e.g. for certain prokaryotic organisms, such as by the Joint Genome Institute. Furthermore, access to proprietary databases, has allowed the identification of novel nucleic acid and polypeptide sequences.

Example 2: Alignment of PRR-like polypeptide sequences

Alignment of the polypeptide sequences was performed using the ClustalW 2.0.11 algorithm of progressive alignment (Thompson et al. (1997) Nucleic Acids Res 25:4876-4882; Chenna et al. (2003). Nucleic Acids Res 31:3497-3500) with standard setting (slow alignment, similarity matrix: Gonnet, gap opening penalty 10, gap extension penalty: 0.2). Minor manual editing was done to further optimise the alignment. The PRR-like polypeptides are aligned in Figure 2.

Example 3: Calculation of global percentage identity between polypeptide sequences

Global percentages of similarity and identity between full length polypeptide sequences useful in performing the methods of the invention were determined using MatGAT (Matrix Global Alignment Tool) software (BMC Bioinformatics. 2003 4:29. MatGAT: an application that generates similarity/identity matrices using protein or DNA sequences. Campanella JJ, Bitincka L, Smalley J; software hosted by Ledion Bitincka). MatGAT generates similarity/identity matrices for DNA or protein sequences without needing pre-alignment of the data. The program performs a series of pair-wise alignments using the Myers and Miller global alignment algorithm, calculates similarity and identity, and then places the results in a distance matrix.

Results of the MatGAT analysis are shown in Figure 3 with global similarity and identity percentages over the full length of the polypeptide sequences. Sequence similarity is shown in the bottom half of the dividing line and sequence identity is shown in the top half of the diagonal dividing line. Parameters used in the analysis were: Scoring matrix: Blosum62, First Gap: 12, Extending Gap: 2. The sequence identity (in %) between the PRR-like polypeptide sequences useful in performing the methods of the invention can be as low as 7,8 %, but is generally higher than 67% compared to SEQ ID NO: 2.

Like for full length sequences, a MATGAT table based on subsequences of a specific domain, may be generated. Based on a multiple alignment of PRR-like polypeptides, such as for example the one of Example 2, a skilled person may select conserved sequences and submit as input for a MaTGAT analysis.

Example 4: Identification of domains comprised in polypeptide sequences useful in performing the methods of the invention

The Integrated Resource of Protein Families, Domains and Sites (InterPro) database is an integrated interface for the commonly used signature databases for text- and sequence-based searches. The InterPro database combines these databases, which use different methodologies and varying degrees of biological information about well-characterized proteins to derive protein signatures. Collaborating databases include SWISS-PROT, PROSITE, TrEMBL, PRINTS, ProDom and Pfam, Smart and TIGRFAMs. Pfam is a large collection of multiple sequence alignments and hidden Markov models covering many common protein domains and families. Pfam is hosted at the Sanger Institute server in the United Kingdom. Interpro is hosted at the European Bioinformatics Institute in the United Kingdom.

The results of the InterPro scan (InterPro database, version 4.8) of the polypeptide sequence as represented by SEQ ID NO: 2 are presented in Table B.

Table B: InterPro scan results (major accession numbers) of the polypeptide sequence as represented by SEQ ID NO: 2.

Database	Accession number	Accession name	Amino acid coordinates on SEQ ID NO : 2
PFAM	PF00072	Signal transduction response regulator, receiver domain	84-163
PFAM	PF06203	CCT domain	711-754

In one embodiment a PRR-like polypeptide comprises a conserved domain or motif with at least 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to a conserved domain from amino acid 84 to 163 in SEQ ID NO:2 or from amino acid 711 to 754 in SEQ ID NO: 2.

Example 5: Topology prediction of the PRR-like polypeptide sequences

TargetP 1.1 predicts the subcellular location of eukaryotic proteins. The location assignment is based on the predicted presence of any of the N-terminal pre-sequences: chloroplast transit peptide (cTP), mitochondrial targeting peptide (mTP) or secretory pathway signal peptide (SP). Scores on which the final prediction is based are not really probabilities, and they do not necessarily add to one. However, the location with the highest score is the most likely according to TargetP, and the relationship between the scores (the reliability class) may be an indication of how certain the prediction is. The reliability class (RC) ranges from 1 to 5, where 1 indicates the strongest prediction. For the sequences predicted to contain an N-terminal presequence a potential cleavage site can also be predicted. TargetP is maintained at the server of the Technical University of Denmark.

A number of parameters must be selected before analysing a sequence, such as organism group (non-plant or plant), cutoff sets (none, predefined set of cutoffs, or user-specified set of cutoffs), and the calculation of prediction of cleavage sites (yes or no).

The results of TargetP 1.1 analysis of the polypeptide sequence as represented by SEQ ID NO: 2 are presented Table C. The “plant” organism group has been selected, no cutoffs defined, and the predicted length of the transit peptide requested. The subcellular localization of the polypeptide sequence as represented by SEQ ID NO: 2 may be the cytoplasm or nucleus, no transit peptide is predicted.

Table C: TargetP 1.1 analysis of the polypeptide sequence as represented by SEQ ID NO: 2

Length (AA)	766
Chloroplastic transit peptide	0.061
Mitochondrial transit peptide	0.157
Secretory pathway signal peptide	0.104

Other subcellular targeting	0.908
Predicted Location	/
Reliability class	2
Predicted transit peptide length	/

Many other algorithms can be used to perform such analyses, including:

- ChloroP 1.1 hosted on the server of the Technical University of Denmark;
- Protein Prowler Subcellular Localisation Predictor version 1.2 hosted on the server of the Institute for Molecular Bioscience, University of Queensland, Brisbane, Australia;
- PENCE Proteome Analyst PA-GOSUB 2.5 hosted on the server of the University of Alberta, Edmonton, Alberta, Canada;
- TMHMM, hosted on the server of the Technical University of Denmark
- PSORT (URL: psort.org)
- PLOC (Park and Kanehisa, Bioinformatics, 19, 1656-1663, 2003).

Example 6: Cloning of the PRR-like encoding nucleic acid sequence

The nucleic acid sequence was amplified by PCR using as template a custom-made *Saccharum officinarum* seedlings cDNA library. PCR was performed using a commercially available proofreading Taq DNA polymerase in standard conditions, using 200 ng of template in a 50 µl PCR mix. The primers used were prm15559 (SEQ ID NO: 47; sense): 5'-ggggacaagtttgtacaaaaaagcaggcttaaacaatgggtagcgcttgcc-3' and prm15560 (SEQ ID NO:48; reverse, complementary): 5'-ggggaccactttgtacaagaaagctgggtacgaagatgccatgttgtatt-3', which include the AttB sites for Gateway recombination. The amplified PCR fragment was purified also using standard methods. The first step of the Gateway procedure, the BP reaction, was then performed, during which the PCR fragment recombined in vivo with the pDONR201 plasmid to produce, according to the Gateway terminology, an "entry clone", pPRR-like. Plasmid pDONR201 was purchased from Invitrogen, as part of the Gateway® technology.

The entry clone comprising SEQ ID NO: 1 was then used in an LR reaction with a destination vector used for *Oryza sativa* transformation. This vector contained as functional elements within the T-DNA borders: a plant selectable marker; a screenable marker expression cassette; and a Gateway cassette intended for LR *in vivo* recombination with the nucleic acid sequence of interest already cloned in the entry clone. A rice GOS2 promoter (SEQ ID NO: 46) for constitutive expression was located upstream of this Gateway cassette.

After the LR recombination step, the resulting expression vector pGOS2::PRR-like (Figure 4) was transformed into *Agrobacterium* strain LBA4044 according to methods well known in the art.

Example 7: Plant transformation

Rice transformation

The *Agrobacterium* containing the expression vector was used to transform *Oryza sativa* plants. Mature dry seeds of the rice japonica cultivar Nipponbare were dehusked. Sterilization was carried out by incubating for one minute in 70% ethanol, followed by 30 to 60 minutes, preferably 30 minutes in sodium hypochlorite solution (depending on the grade of contamination), followed by a 3 to 6 times, preferably 4 time wash with sterile distilled water. The sterile seeds were then germinated on a medium containing 2,4-D (callus induction medium). After incubation in light for 6 days scutellum-derived calli is transformed with Agrobacterium as described herein below.

Agrobacterium strain LBA4404 containing the expression vector was used for co-cultivation. *Agrobacterium* was inoculated on AB medium with the appropriate antibiotics and cultured for 3 days at 28°C. The bacteria were then collected and suspended in liquid co-cultivation medium to a density (OD_{600}) of about 1. The calli were immersed in the suspension for 1 to 15 minutes. The callus tissues were then blotted dry on a filter paper and transferred to solidified, co-cultivation medium and incubated for 3 days in the dark at 25°C. After washing away the *Agrobacterium,* the calli were grown on 2,4-D-containing medium for 10 to 14 days (growth time for indica: 3 weeks) under light at 28°C - 32°C in the presence of a selection agent. During this period, rapidly growing resistant callus developed. After transfer of this material to regeneration media, the embryogenic potential was released and shoots developed in the next four to six weeks. Shoots were excised from the calli and incubated for 2 to 3 weeks on an auxin-containing medium from which they were transferred to soil. Hardened shoots were grown under high humidity and short days in a greenhouse.

Transformation of rice cultivar indica can also be done in a similar way as give above according to techniques well known to a skilled person.

35 to 90 independent T0 rice transformants were generated for one construct. The primary transformants were transferred from a tissue culture chamber to a greenhouse. After a quantitative PCR analysis to verify copy number of the T-DNA insert, only single copy transgenic plants that exhibit tolerance to the selection agent were kept for harvest of T1 seed. Seeds were then harvested three to five months after transplanting. The method yielded single locus transformants at a rate of over 50 % (Aldemita and Hodges1996, Chan et al. 1993, Hiei et al. 1994).

Example 8: Transformation of other crops

Corn transformation

Transformation of maize (*Zea mays*) is performed with a modification of the method described by Ishida et al. (1996) Nature Biotech 14(6): 745-50. Transformation is genotype-dependent in corn and only specific genotypes are amenable to transformation and regeneration. The inbred line A188 (University of Minnesota) or hybrids with A188 as a

parent are good sources of donor material for transformation, but other genotypes can be used successfully as well. Ears are harvested from corn plant approximately 11 days after pollination (DAP) when the length of the immature embryo is about 1 to 1.2 mm. Immature embryos are cocultivated with *Agrobacterium tumefaciens* containing the expression vector, and transgenic plants are recovered through organogenesis. Excised embryos are grown on callus induction medium, then maize regeneration medium, containing the selection agent (for example imidazolinone but various selection markers can be used). The Petri plates are incubated in the light at 25 °C for 2-3 weeks, or until shoots develop. The green shoots are transferred from each embryo to maize rooting medium and incubated at 25 °C for 2-3 weeks, until roots develop. The rooted shoots are transplanted to soil in the greenhouse. T1 seeds are produced from plants that exhibit tolerance to the selection agent and that contain a single copy of the T-DNA insert.

Wheat transformation

Transformation of wheat is performed with the method described by Ishida et al. (1996) Nature Biotech 14(6): 745-50. The cultivar Bobwhite (available from CIMMYT, Mexico) is commonly used in transformation. Immature embryos are co-cultivated with *Agrobacterium tumefaciens* containing the expression vector, and transgenic plants are recovered through organogenesis. After incubation with *Agrobacterium*, the embryos are grown in vitro on callus induction medium, then regeneration medium, containing the selection agent (for example imidazolinone but various selection markers can be used). The Petri plates are incubated in the light at 25 °C for 2-3 weeks, or until shoots develop. The green shoots are transferred from each embryo to rooting medium and incubated at 25 °C for 2-3 weeks, until roots develop. The rooted shoots are transplanted to soil in the greenhouse. T1 seeds are produced from plants that exhibit tolerance to the selection agent and that contain a single copy of the T-DNA insert.

Soybean transformation

Soybean is transformed according to a modification of the method described in the Texas A&M patent US 5,164,310. Several commercial soybean varieties are amenable to transformation by this method. The cultivar Jack (available from the Illinois Seed foundation) is commonly used for transformation. Soybean seeds are sterilised for *in vitro* sowing. The hypocotyl, the radicle and one cotyledon are excised from seven-day old young seedlings. The epicotyl and the remaining cotyledon are further grown to develop axillary nodes. These axillary nodes are excised and incubated with *Agrobacterium tumefaciens* containing the expression vector. After the cocultivation treatment, the explants are washed and transferred to selection media. Regenerated shoots are excised and placed on a shoot elongation medium. Shoots no longer than 1 cm are placed on rooting medium until roots develop. The rooted shoots are transplanted to soil in the greenhouse. T1 seeds are produced from plants that exhibit tolerance to the selection agent and that contain a single copy of the T-DNA insert.

Rapeseed/canola transformation

Cotyledonary petioles and hypocotyls of 5-6 day old young seedling are used as explants for tissue culture and transformed according to Babic et al. (1998, Plant Cell Rep 17: 183-188). The commercial cultivar Westar (Agriculture Canada) is the standard variety used for transformation, but other varieties can also be used. Canola seeds are surface-sterilized for in vitro sowing. The cotyledon petiole explants with the cotyledon attached are excised from the in vitro seedlings, and inoculated with *Agrobacterium* (containing the expression vector) by dipping the cut end of the petiole explant into the bacterial suspension. The explants are then cultured for 2 days on MSBAP-3 medium containing 3 mg/l BAP, 3 % sucrose, 0.7 % Phytagar at 23 °C, 16 hr light. After two days of co-cultivation with *Agrobacterium*, the petiole explants are transferred to MSBAP-3 medium containing 3 mg/l BAP, cefotaxime, carbenicillin, or timentin (300 mg/l) for 7 days, and then cultured on MSBAP-3 medium with cefotaxime, carbenicillin, or timentin and selection agent until shoot regeneration. When the shoots are 5 – 10 mm in length, they are cut and transferred to shoot elongation medium (MSBAP-0.5, containing 0.5 mg/l BAP). Shoots of about 2 cm in length are transferred to the rooting medium (MS0) for root induction. The rooted shoots are transplanted to soil in the greenhouse. T1 seeds are produced from plants that exhibit tolerance to the selection agent and that contain a single copy of the T-DNA insert.

Alfalfa transformation

A regenerating clone of alfalfa (*Medicago sativa*) is transformed using the method of (McKersie et al., 1999 Plant Physiol 119: 839–847). Regeneration and transformation of alfalfa is genotype dependent and therefore a regenerating plant is required. Methods to obtain regenerating plants have been described. For example, these can be selected from the cultivar Rangelander (Agriculture Canada) or any other commercial alfalfa variety as described by Brown DCW and A Atanassov (1985. Plant Cell Tissue Organ Culture 4: 111-112). Alternatively, the RA3 variety (University of Wisconsin) has been selected for use in tissue culture (Walker et al., 1978 Am J Bot 65:654-659). Petiole explants are cocultivated with an overnight culture of *Agrobacterium tumefaciens* C58C1 pMP90 (McKersie et al., 1999 Plant Physiol 119: 839–847) or LBA4404 containing the expression vector. The explants are cocultivated for 3 d in the dark on SH induction medium containing 288 mg/ L Pro, 53 mg/ L thioproline, 4.35 g/ L K2SO4, and 100 µm acetosyringinone. The explants are washed in half-strength Murashige-Skoog medium (Murashige and Skoog, 1962) and plated on the same SH induction medium without acetosyringinone but with a suitable selection agent and suitable antibiotic to inhibit *Agrobacterium* growth. After several weeks, somatic embryos are transferred to BOi2Y development medium containing no growth regulators, no antibiotics, and 50 g/ L sucrose. Somatic embryos are subsequently germinated on half-strength Murashige-Skoog medium. Rooted seedlings were transplanted into pots and grown in a greenhouse. T1 seeds are produced from plants that exhibit tolerance to the selection agent and that contain a single copy of the T-DNA insert.

Cotton transformation

Cotton is transformed using *Agrobacterium tumefaciens* according to the method described in US 5,159,135. Cotton seeds are surface sterilised in 3% sodium hypochlorite solution during 20 minutes and washed in distilled water with 500 µg/ml cefotaxime. The seeds are then transferred to SH-medium with 50µg/ml benomyl for germination. Hypocotyls of 4 to 6 days old seedlings are removed, cut into 0.5 cm pieces and are placed on 0.8% agar. An *Agrobacterium* suspension (approx. 108 cells per ml, diluted from an overnight culture transformed with the gene of interest and suitable selection markers) is used for inoculation of the hypocotyl explants. After 3 days at room temperature and lighting, the tissues are transferred to a solid medium (1.6 g/l Gelrite) with Murashige and Skoog salts with B5 vitamins (Gamborg et al., Exp. Cell Res. 50:151-158 (1968)), 0.1 mg/l 2,4-D, 0.1 mg/l 6-furfurylaminopurine and 750 µg/ml MgCL2, and with 50 to 100 µg/ml cefotaxime and 400-500 µg/ml carbenicillin to kill residual bacteria. Individual cell lines are isolated after two to three months (with subcultures every four to six weeks) and are further cultivated on selective medium for tissue amplification (30°C, 16 hr photoperiod). Transformed tissues are subsequently further cultivated on non-selective medium during 2 to 3 months to give rise to somatic embryos. Healthy looking embryos of at least 4 mm length are transferred to tubes with SH medium in fine vermiculite, supplemented with 0.1 mg/l indole acetic acid, 6 furfurylaminopurine and gibberellic acid. The embryos are cultivated at 30°C with a photoperiod of 16 hrs, and plantlets at the 2 to 3 leaf stage are transferred to pots with vermiculite and nutrients. The plants are hardened and subsequently moved to the greenhouse for further cultivation.

Sugarbeet transformation

Seeds of sugarbeet (Beta vulgaris L.) are sterilized in 70% ethanol for one minute followed by 20 min. shaking in 20% Hypochlorite bleach e.g. Clorox® regular bleach (commercially available from Clorox, 1221 Broadway, Oakland, CA 94612, USA). Seeds are rinsed with sterile water and air dried followed by plating onto germinating medium (Murashige and Skoog (MS) based medium (Murashige, T., and Skoog, ., 1962. Physiol. Plant, vol. 15, 473-497) including B5 vitamins (Gamborg et al.; Exp. Cell Res., vol. 50, 151-8.) supplemented with 10 g/l sucrose and 0,8% agar). Hypocotyl tissue is used essentially for the initiation of shoot cultures according to Hussey and Hepher (Hussey, G., and Hepher, A., 1978. Annals of Botany, 42, 477-9) and are maintained on MS based medium supplemented with 30g/l sucrose plus 0,25mg/l benzylamino purine and 0,75% agar, pH 5,8 at 23-25°C with a 16-hour photoperiod. Agrobacterium tumefaciens strain carrying a binary plasmid harbouring a selectable marker gene, for example nptII, is used in transformation experiments. One day before transformation, a liquid LB culture including antibiotics is grown on a shaker (28°C, 150rpm) until an optical density (O.D.) at 600 nm of ~1 is reached. Overnight-grown bacterial cultures are centrifuged and resuspended in inoculation medium (O.D. ~1) including Acetosyringone, pH 5,5. Shoot base tissue is cut into slices (1.0 cm x 1.0 cm x 2.0 mm approximately). Tissue is immersed for 30s in liquid bacterial inoculation medium. Excess liquid is removed by filter paper blotting. Co-cultivation occurred for 24-72 hours on MS based medium incl. 30g/l sucrose followed by a non-selective period including MS

based medium, 30g/l sucrose with 1 mg/l BAP to induce shoot development and cefotaxim for eliminating the Agrobacterium. After 3-10 days explants are transferred to similar selective medium harbouring for example kanamycin or G418 (50-100 mg/l genotype dependent). Tissues are transferred to fresh medium every 2-3 weeks to maintain selection pressure. The very rapid initiation of shoots (after 3-4 days) indicates regeneration of existing meristems rather than organogenesis of newly developed transgenic meristems. Small shoots are transferred after several rounds of subculture to root induction medium containing 5 mg/l NAA and kanamycin or G418. Additional steps are taken to reduce the potential of generating transformed plants that are chimeric (partially transgenic). Tissue samples from regenerated shoots are used for DNA analysis. Other transformation methods for sugarbeet are known in the art, for example those by Linsey & Gallois (Linsey, K., and Gallois, P., 1990. Journal of Experimental Botany; vol. 41, No. 226; 529-36) or the methods published in the international application published as WO9623891A.

Sugarcane transformation

Spindles are isolated from 6-month-old field grown sugarcane plants (Arencibia et al., 1998. Transgenic Research, vol. 7, 213-22; Enriquez-Obregon et al., 1998. Planta, vol. 206, 20-27). Material is sterilized by immersion in a 20% Hypochlorite bleach e.g. Clorox® regular bleach (commercially available from Clorox, 1221 Broadway, Oakland, CA 94612, USA) for 20 minutes. Transverse sections around 0,5cm are placed on the medium in the top-up direction. Plant material is cultivated for 4 weeks on MS (Murashige, T., and Skoog, ., 1962. Physiol. Plant, vol. 15, 473-497) based medium incl. B5 vitamins (Gamborg, O., et al., 1968. Exp. Cell Res., vol. 50, 151-8) supplemented with 20g/l sucrose, 500 mg/l casein hydrolysate, 0,8% agar and 5mg/l 2,4-D at 23°C in the dark. Cultures are transferred after 4 weeks onto identical fresh medium. Agrobacterium tumefaciens strain carrying a binary plasmid harbouring a selectable marker gene, for example hpt, is used in transformation experiments. One day before transformation, a liquid LB culture including antibiotics is grown on a shaker (28°C, 150rpm) until an optical density (O.D.) at 600 nm of ~0,6 is reached. Overnight-grown bacterial cultures are centrifuged and resuspended in MS based inoculation medium (O.D. ~0,4) including acetosyringone, pH 5,5. Sugarcane embryogenic callus pieces (2-4 mm) are isolated based on morphological characteristics as compact structure and yellow colour and dried for 20 min. in the flow hood followed by immersion in a liquid bacterial inoculation medium for 10-20 minutes. Excess liquid is removed by filter paper blotting. Co-cultivation occurred for 3-5 days in the dark on filter paper which is placed on top of MS based medium incl. B5 vitamins containing 1 mg/l 2,4-D. After co-cultivation calli are washed with sterile water followed by a non-selective cultivation period on similar medium containing 500 mg/l cefotaxime for eliminating remaining Agrobacterium cells. After 3-10 days explants are transferred to MS based selective medium incl. B5 vitamins containing 1 mg/l 2,4-D for another 3 weeks harbouring 25 mg/l of hygromycin (genotype dependent). All treatments are made at 23°C under dark conditions. Resistant calli are further cultivated on medium lacking 2,4-D including 1 mg/l BA and 25 mg/l hygromycin under 16 h light photoperiod resulting in the development of shoot structures.

Shoots are isolated and cultivated on selective rooting medium (MS based including, 20g/l sucrose, 20 mg/l hygromycin and 500 mg/l cefotaxime). Tissue samples from regenerated shoots are used for DNA analysis. Other transformation methods for sugarcane are known in the art, for example from the in-ternational application published as WO2010/151634A and the granted European patent EP1831378.

Example 9: Phenotypic evaluation procedure

9.1 Evaluation setup

35 to 90 independent T0 rice transformants were generated. The primary transformants were transferred from a tissue culture chamber to a greenhouse for growing and harvest of T1 seed. Six events, of which the T1 progeny segregated 3:1 for presence/absence of the transgene, were retained. For each of these events, approximately 10 T1 seedlings containing the transgene (hetero- and homo-zygotes) and approximately 10 T1 seedlings lacking the transgene (nullizygotes) were selected by monitoring visual marker expression. The transgenic plants and the corresponding nullizygotes were grown side-by-side at random positions. Greenhouse conditions were of short days (12 hours light), 28°C in the light and 22°C in the dark, and a relative humidity of 70%. Plants grown under non-stress conditions were watered at regular intervals to ensure that water and nutrients were not limiting and to satisfy plant needs to complete growth and development, unless they were used in a stress screen.

From the stage of sowing until the stage of maturity the plants were passed several times through a digital imaging cabinet. At each time point digital images (2048x1536 pixels, 16 million colours) were taken of each plant from at least 6 different angles.

T1 events can be further evaluated in the T2 generation following the same evaluation procedure as for the T1 generation, e.g. with less events and/or with more individuals per event.

Drought screen

T1 or T2 plants are grown in potting soil under normal conditions until they approached the heading stage. They are then transferred to a "dry" section where irrigation is withheld. Soil moisture probes are inserted in randomly chosen pots to monitor the soil water content (SWC). When SWC goes below certain thresholds, the plants are automatically re-watered continuously until a normal level is reached again. The plants are then re-transferred again to normal conditions. The rest of the cultivation (plant maturation, seed harvest) is the same as for plants not grown under abiotic stress conditions. Growth and yield parameters are recorded as detailed for growth under normal conditions.

Nitrogen use efficiency screen

T1 or T2 plants are grown in potting soil under normal conditions except for the nutrient solution. The pots are watered from transplantation to maturation with a specific nutrient

solution containing reduced N nitrogen (N) content, usually between 7 to 8 times less. The rest of the cultivation (plant maturation, seed harvest) is the same as for plants not grown under abiotic stress. Growth and yield parameters are recorded as detailed for growth under normal conditions.

Salt stress screen

T1 or T2 plants are grown on a substrate made of coco fibers and particles of baked clay (Argex) (3 to 1 ratio). A normal nutrient solution is used during the first two weeks after transplanting the plantlets in the greenhouse. After the first two weeks, 25 mM of salt (NaCl) is added to the nutrient solution, until the plants are harvested. Growth and yield parameters are recorded as detailed for growth under normal conditions.

9.2 Statistical analysis: F test

A two factor ANOVA (analysis of variants) was used as a statistical model for the overall evaluation of plant phenotypic characteristics. An F test was carried out on all the parameters measured of all the plants of all the events transformed with the gene of the present invention. The F test was carried out to check for an effect of the gene over all the transformation events and to verify for an overall effect of the gene, also known as a global gene effect. The threshold for significance for a true global gene effect was set at a 5% probability level for the F test. A significant F test value points to a gene effect, meaning that it is not only the mere presence or position of the gene that is causing the differences in phenotype.

9.3 Parameters measured

From the stage of sowing until the stage of maturity the plants were passed several times through a digital imaging cabinet. At each time point digital images (2048x1536 pixels, 16 million colours) were taken of each plant from at least 6 different angles as described in WO2010/031780. These measurements were used to determine different parameters.

Biomass-related parameter measurement

The plant aboveground area, or leafy biomass, was determined by counting the total number of pixels on the digital images from aboveground plant parts discriminated from the background. This value was averaged for the pictures taken on the same time point from the different angles and was converted to a physical surface value expressed in square mm by calibration. Experiments show that the aboveground plant area measured this way correlates with the biomass of plant parts above ground. The above ground area is the area measured at the time point at which the plant had reached its maximal leafy biomass.

Increase in root biomass is expressed as an increase in total root biomass, measured as maximum biomass of roots observed during the lifespan of a plant; or as an increase in the root/shoot index, measured as the ratio between root mass and shoot mass in the period of active growth of root and shoot. In other words, the root/shoot index is defined as the ratio

of the rapidity of root growth to the rapidity of shoot growth in the period of active growth of root and shoot. Root biomass can be determined using a method as described in WO 2006/029987.

Parameters related to development time

The early vigour is the plant aboveground area three weeks post-germination. Early vigour was determined by counting the total number of pixels from aboveground plant parts discriminated from the background. This value was averaged for the pictures taken on the same time point from different angles and was converted to a physical surface value expressed in square mm by calibration.

AreaEmer is an indication of quick early development when this value is decreased compared to control plants. It is the ratio, expressed in %, between the time a plant needs to make 30 % of the final biomass and the time needs to make 90 % of its final biomass.

The "time to flower" or "flowering time" of the plant can be determined using the method as described in WO 2007/093444.

Seed-related parameter measurements

The mature primary panicles were harvested, counted, bagged, barcode-labelled and then dried for three days in an oven at 37°C. The panicles were then threshed and all the seeds were collected and counted. The seeds are usually covered by a dry outer covering, the husk. The filled husks (herein also named filled florets) were separated from the empty ones using an air-blowing device. The empty husks were discarded and the remaining fraction was counted again. The filled husks were weighed on an analytical balance.

The total number of seeds was determined by counting the number of filled husks that remained after the separation step. The total seed weight was measured by weighing all filled husks harvested from a plant.

The total number of seeds (or florets) per plant was determined by counting the number of husks (whether filled or not) harvested from a plant.

Thousand Kernel Weight (TKW) is extrapolated from the number of seeds counted and their total weight.

The Harvest Index (HI) in the present invention is defined as the ratio between the total seed weight and the above ground area (mm^2), multiplied by a factor 10^6.

The number of flowers per panicle as defined in the present invention is the ratio between the total number of seeds over the number of mature primary panicles.

The "seed fill rate" or "seed filling rate" as defined in the present invention is the proportion, expressed as a %, of the number of filled seeds, i.e. florets containing seeds, over the total number of seeds, i.e. total number of florets. In other words, the seed filling rate is the percentage of florets that are filled with seed.

Example 10: Results of the phenotypic evaluation of the transgenic plants

The results of the evaluation of transgenic rice plants in the T1 generation and expressing a nucleic acid encoding the PRR-like polypeptide of SEQ ID NO: 2 under non-stress conditions are presented below in Table D. When grown under non-stress conditions, an increase of at least 5 % was observed for aboveground biomass (AreaMax), root biomass (RootMax), and for seed yield, including total weight of seeds (totalwgseeds), number of seeds (nrfilledseeds), fill rate (fillrate) and the number of flowers per panicle (flowerperpan).

Table D: Data summary for transgenic rice plants; for each parameter, the overall percent increase is shown for the confirmation (T1 generation), for each parameter the p-value is <0.05.

Parameter	Overall increase
AreaMax	10.9
RootMax	9.7
totalwgseeds	26.7
nrfilledseed	22.9
fillrate	17.3
flowerperpan	17.1

Claims

1. A method for the production of a transgenic plant having increased biomass and/or seed yield relative to a control plant, comprising the steps of:
   - introducing and expressing in a plant cell or plant a nucleic acid encoding a PRR-like polypeptide, wherein said nucleic acid is operably linked to a constitutive plant promoter, and wherein said PRR-like polypeptide comprises the polypeptide represented by one of: SEQ ID NO: 2, or a homologue thereof which has at least 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% overall sequence identity to SEQ ID NO : 2, and
   - cultivating said plant cell or plant under conditions promoting plant growth and development.

2. Method according to claim 1, wherein said increased seed yield comprises at least one parameter selected from the group comprising increased total weight of seeds, increased number of seeds, increased fill rate and increased number of flowers per panicle.

3. Method according to claim 1 or 2, wherein said increase in seed yield comprises an increase of at least 5 % in said plant when compared to control plants for at least one of said parameters.

4. Method according to claim 1, wherein said increase in biomass comprises at least one parameter selected from the group comprising increased aboveground biomass and increased root biomass.

5. Method according to any of claims 1 to 4, wherein said increased biomass and/or seed yield is obtained under non-stress conditions.

6. Method according to any one of claims 1 to 5, wherein said nucleic acid is operably linked to a GOS2 promoter.

7. Method according to claim 6, wherein said GOS2 promoter is the GOS2 promoter from rice.

8. Method according to any one for claims 1 to 7, wherein said plant is a monocotyledonous plant.

9. Method according to claim 8, wherein said plant is a cereal.

10. Construct comprising:
    (i) nucleic acid encoding a PRR-like polypeptide as defined in claim 1,

(ii) one or more control sequences capable of driving expression of the nucleic acid sequence of (i); and optionally

(iii) a transcription termination sequence.

11. Construct of claim 10, wherein one of said one or more control sequences is a GOS2 promoter.

12. Transgenic plant having enhanced seed yield as defined in claim 2 or 3 relative to control plants, resulting from introduction and expression of a nucleic acid encoding a PRR-like polypeptide as defined in claim 1 in said plant, or a transgenic plant cell derived from said transgenic plant.

13. Transgenic plant according to claim 12, or a transgenic plant cell derived therefrom, wherein said plant is a crop plant, such as beet, sugarbeet or alfalfa; or a monocotyledonous plant such as sugarcane; or a cereal, such as rice, maize, wheat, barley, millet, rye, triticale, sorghum, emmer, spelt, einkorn, teff, milo or oats.

14. Harvestable parts of a plant according to claim 13, wherein said harvestable parts are preferably root biomass and/or seeds.

15. Products derived from a plant according to claim 13 and/or from harvestable parts of a plant according to claim 14.

16. Use of a nucleic acid encoding an PRR-like polypeptide as defined in any of claims 1, 6 or 7 for increasing biomass and/or seed yield in plants relative to control plants.

17. A method for manufacturing a product comprising the steps of growing the plants according to claims 12 or 13 and producing said product from or by said plants; or parts thereof, including seeds.

18. Products produced from a plant according to claim 13 and/or from harvestable parts of a plant according to claim 14.

19. Construct according to claim 10 or 11 comprised in a plant cell.

20. Recombinant chromosomal DNA comprising the construct according to claim 10 or 11.

21. Method according to any one of claims 1 to 9, wherein said polypeptide is encoded by a nucleic acid molecule comprising a nucleic acid molecule selected from the group consisting of:

(i) a nucleic acid represented by SEQ ID NO: 1;

(ii) the complement of a nucleic acid represented by SEQ ID NO: 1;

(iii) a nucleic acid encoding the polypeptide as represented by SEQ ID NO: 1, and further preferably confers increased biomass and/or seed yield relative to control plants;

(iv) a nucleic acid having, in increasing order of preference at least 30 %, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity with the nucleic acid sequences of SEQ ID NO: 1, and further preferably conferring increased biomass and/or seed yield relative to control plants,

(v) a nucleic acid molecule which hybridizes to the complement of a nucleic acid molecule of (i) to (iv) under stringent hybridization conditions and preferably confers increased biomass and/or seed yield relative to control plants;

(vi) a nucleic acid encoding said polypeptide having, in increasing order of preference, at least 30 %, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence represented by SEQ ID NO: 2 and preferably conferring increased biomass and/or seed yield relative to control plants; or

(vii) a nucleic acid comprising any combination(s) of features of (i) to (vi) above.

MGSACQAGTDGSSRKDVLGIGNVTLDNGHHEAEADADEWREKEDDLANGRSAPPGMQQ

VDEQKEQQGQSIHWERFLPVKTLRV**LLVENDDSTRQVVSALLRKCCYEVIPAENGSHA**

MEME motif 2

WRYLENLQNNIDLVLTEVFMPCLSGIGLLSKITSHKICKDIPVIM**MSSNDSMSMVFKC**

MEME motif 2 MEME motif 1

LSKGAVDFLVKPLRKNELKNLWQHVWRRCHSSSGSGSESGIQTQKYAKPNTGDEYEND

MEME motif 1

SDSNHDDEENDDDDDDDF**SVGLNARDGSDNGSGTQSSWTKRAVEIDSPQPMSPD**QLAD

MEME motif 3

PPDSTCAQVIHPKSEICSNKWLPTANKRNGKKHKENKDESMGRYLEIGAPRNSSAEYQ

SSLNDVSVNPTEKRHETHMPQCKSKKKMMAEDDCTDIPSETNTETADLISSIARNTEG

QQAVRAVDAPDGPSKMPDGNGKNHDSHIEVTPHELGLKRLRTDGATAEIHDERNILKR

SDQSAFTRYHTSVASNQGGARCGESSSPQDNSSEAVKTDSTCKMKSNSDAAPIKQGSN

GSSNNDVGSSTKNVVAKPSANRERVTSPSTIKSTQHASAFHTIHNQTSPANLVGKDKA

DEGISNAVKMSHPTEVPQSCVQHHHHVHYYLHVLTQKQLSIDRGSSDVQCGSSNVFDP

PVEGHAANYSVNGGVSVGHNGCNGQNGTSAVPNIATPNIESVNGTMSQNIAGGGIVSG

SGSGNDVYQNRFPQREAALNKFRLKRKDRNFGKKVRYQSRKRLAEQRPRVRGQFVRQS

GQEDQAAQGSER

FIGURE 1

```
CLUSTAL 2.0.11 multiple sequence alignment

Z.mays_GRMZM2G095727_T03        --------------------------------------------------
Z.mays_GRMZM2G095727_T02        --------------------------------------------------
Z.mays_GRMZM2G095727_T01        --------------------------------------------------
S.officinarum_TC81972           --------------------------------------------------
P.virgatum_TC27226              --------------------------------------------------
S.bicolor_Sb01g038820.1         ---MGSACQAGMDGPSRKDVLGIGNVALENGHHEVGADADEWREKEEDLA
S.OFFICINARUM_PRR-LIKE          --MGSACQAGTDGSSRKDVLGIGNVTLDNGHHEAEADADEWREKEDDLAN
Z.mays_GRMZM2G095727_T05        MGSACQAGTDGPSRKDVLGIGNAALENGHHQAEADADEWREKEEDLANNG
Z.mays_GRMZM2G095727_T04        MGSACQAGTDGPSRKDVLGIGNAALENGHHQAEADADEWREKEEDLANNG
S.officinarum_TC80591           --------------------------------------------------
S.officinarum_TC111093          --------------------------------------------------
P.virgatum_TC42896              --------------------------------------------------
S.officinarum_TC84822           --------------------------------------------------
P.virgatum_TC22980              --------------------------------------------------
L.sativa_TC22097                --------------------------------------------------
C.endivia_TA960_114280          --------------------------------------------------
C.maculosa_TA1222_215693        --------------------------------------------------
B.distachyon_DV469398           --------------------------------------------------
T.aestivum_TC284871             MVSAGQAGADGPSTSDIRGTGNGAVEN--GHALKANEDKEWRGGSKEEDW
H.vulgare_TC165635              MVSAGQAGADGPSSSDIRGIGNGAVENQNGHALKPNEDKEWRGGSKEEDW
T.aestivum_c59844892@6464       MVMP---------SSQTR-TRYGAAAS-----------RKRTG-------


Z.mays_GRMZM2G095727_T03        --------------------------------------------------
Z.mays_GRMZM2G095727_T02        --------------------------------------------------
Z.mays_GRMZM2G095727_T01        --------------------------------------------------
S.officinarum_TC81972           --------------------------------------------------
P.virgatum_TC27226              --------------------------------------------------
S.bicolor_Sb01g038820.1         NGHSAPPGMQQVDEQEQQGQSIHWERFLPVKTLRVMLVENDDSTRQVVSA
S.OFFICINARUM_PRR-LIKE          GRSAPPGMQQVDEQKEQQGQSIHWERFLPVKTLRVLLVENDDSTRQVVSA
Z.mays_GRMZM2G095727_T05        HSAPPPGMQQVDEHKEEQRQSIHWERFLPVKTLRVLLVENDDSTRQVVSA
Z.mays_GRMZM2G095727_T04        HSAPPPGMQQVDEHKEEQRQSIHWERFLPVKTLRVLLVENDDSTRQVVSA
S.officinarum_TC80591           --------------------------------------------------
S.officinarum_TC111093          --------------------------------------------------
P.virgatum_TC42896              --------------------------------------------------
S.officinarum_TC84822           --------------------------------------------------
P.virgatum_TC22980              --------------------------------------------------
L.sativa_TC22097                ----------------------------------MLVEDDDCTRHIVTA
C.endivia_TA960_114280          ----------------------------------MLVENDDCTRHIVTA
C.maculosa_TA1222_215693        --------------------------------------------------
B.distachyon_DV469398           --------------------------------------------------
T.aestivum_TC284871             PSTHSAPPGLDEHKQQQ-GRVIRWEKFLPVETLRVLLVENDDCTRHVVRA
H.vulgare_TC165635              PSTHSAPPGLDEHKQQQQDRVIRWEKFLPVKTLRVLLVENDDCTRHVVRA
T.aestivum_c59844892@6464       PTRTVRQTGLDEQKQQQ-DRVIRWEKFLPVKTLRVLLVENDDCTRHVVRA
```

FIGURE 2

```
Z.mays_GRMZM2G095727_T03        --------------------------------------------------
Z.mays_GRMZM2G095727_T02        --------------------------------------------------
Z.mays_GRMZM2G095727_T01        --------------------------------------------------
S.officinarum_TC81972           --------------------------------------------------
P.virgatum_TC27226              --------------------------------------------------
S.bicolor_Sb01g038820.1         LLRKCCYEVIPAENGSHAWRYLEDLQNNIDLVLTEVFMPCLSGIGLLSKI
S.OFFICINARUM_PRR-LIKE          LLRKCCYEVIPAENGSHAWRYLENLQNNIDLVLTEVFMPCLSGIGLLSKI
Z.mays_GRMZM2G095727_T05        LLRKCCYEVIPAENGLHAWRYLEDLQNNIDLVLTEVFMPCLSGIGLLSKI
Z.mays_GRMZM2G095727_T04        LLRKCCYEVIPAENGLHAWRYLEDLQNNIDLVLTEVFMPCLSGIGLLSKI
S.officinarum_TC80591           --------------------------------------------------
S.officinarum_TC111093          --------------------------------------------------
P.virgatum_TC42896              --------------------------------------------------
S.officinarum_TC84822           --------------------------------------------------
P.virgatum_TC22980              --------------------------------------------------
L.sativa_TC22097                LLRNCNYEVIQAANGLQAWKILENLSNHIDIVLTEVVMPCLSGIGLLCKI
C.endivia_TA960_114280          LLRNCNYEVIEASNGFQAWKILEDLSNHIDIVLTEVVMPSFSGVGLLCKI
C.maculosa_TA1222_215693        ----------------------------------MPCLSGIGLLCKI
B.distachyon_DV469398           ----------------------------------MPCLSGISLLSKI
T.aestivum_TC284871             LLRKCGYEVIAAENGLHAWHYLEDVQNRIDLVLTEVAMPCLSGIGLLSKI
H.vulgare_TC165635              LLRKCGYEVISAENGLDAWQYLEDVQNRIDLVLTEVAMPCLSGIGLLSKI
T.aestivum_c59844892@6464       LLRKCGYEVISAENGLHAWQYLEDVQNRIDLVLTEVAMPCLSGIGLLSKI

Z.mays_GRMZM2G095727_T03        ------------MSTNDSMSMVFKCLSKGAVDFLVKPLRKNELKNLWQH
Z.mays_GRMZM2G095727_T02        ------------MSTNDSMSMVFKCLSKGAVDFLVKPLRKNELKNLWQH
Z.mays_GRMZM2G095727_T01        ------------MSTNDSMSMVFKCLSKGAVDFLVKPLRKNELKNLWQH
S.officinarum_TC81972           ------------------MSMVFKCLSKGAVDFLVKPLRKNELKNLWQH
P.virgatum_TC27226              ------------MSSNDSMSMVFKCLSKGAVDFLVKPLRKNELKNLWQH
S.bicolor_Sb01g038820.1         TSHKICKDIPVIMMSSNDSMSMVFKCLSKGAVDFLVKPLRKNELKNLWQH
S.OFFICINARUM_PRR-LIKE          TSHKICKDIPVIMMSSNDSMSMVFKCLSKGAVDFLVKPLRKNELKNLWQH
Z.mays_GRMZM2G095727_T05        TSHKICKDIPVIMMSTNDSMSMVFKCLSKGAVDFLVKPLRKNELKNLWQH
Z.mays_GRMZM2G095727_T04        TSHKICKDIPVIMMSTNDSMSMVFKCLSKGAVDFLVKPLRKNELKNLWQH
S.officinarum_TC80591           --------------------------------------------------
S.officinarum_TC111093          --------------------------------------------------
P.virgatum_TC42896              --------------------------------------------------
S.officinarum_TC84822           --------------------------------------------------
P.virgatum_TC22980              --------------------------------------------------
L.sativa_TC22097                MSHKTRKNIPVIMMSSHDSMGLVFKCLSKGAVDFLVKPIRKNELKNLWQH
C.endivia_TA960_114280          MSHKTRKNIPVIMMSSHDSMGLVFKCLSKGAVDFLLKPIRKNELKNLWQH
C.maculosa_TA1222_215693        MSHKTRKNIPVIMMSSHDSMGLVFKCLSKGAVDFLVKPVRKNELKNLWQH
B.distachyon_DV469398           MSHKICKDIPVIMMSKNDSMGTVFKCLSKGAVDFLVKPIRKNELKNLWQH
T.aestivum_TC284871             TSHSICKGIPVIMMSKNDSMSTVFRCLSKGAVDFLVKPIRKNELKTLWQH
H.vulgare_TC165635              TSHSICKGIPVIMMSKNDSMSTVFKCLSKGAVDFLVKPIRKNELKTLWQH
T.aestivum_c59844892@6464       TSHSICKGIPVIMMSKNDSMSTVFKCLSKGAVDFLVKPIRKNELKTLWQH
```

FIGURE 2 (continued)

```
Z.mays_GRMZM2G095727_T03        VWRRCHS------SSGS--ESGIQTQKCAKLNTGDEYENGSDSNHDDEEN
Z.mays_GRMZM2G095727_T02        VWRRCHS------SSGS--ESGIQTQKCAKLNTGDEYENGSDSNHDDEEN
Z.mays_GRMZM2G095727_T01        VWRRCHS------SSGS--ESGIQTQKCAKLNTGDEYENGSDSNHDDEEN
S.officinarum_TC81972           VWRRCHS------SSGSGSESGIQTQKCAKPNTGDEYENDSDSNHDDEEN
P.virgatum_TC27226              VWRRCHS------SSGSRSESGIQTQKCAIPNTGDKYENNSNSNHDDDE-
S.bicolor_Sb01g038820.1         VWRRCHS------SSGSGSESGIQTQKCAKPNTGDEYENDSDSNHDDEEN
S.OFFICINARUM_PRR-LIKE          VWRRCHS------SSGSGSESGIQTQKYAKPNTGDEYENDSDSNHDDEEN
Z.mays_GRMZM2G095727_T05        VWRRCHS------SSGS--ESGIQTQKCAKLNTGDEYENGSDSNHDDEEN
Z.mays_GRMZM2G095727_T04        VWRRCHSVSCLFLSSGS--ESGIQTQKCAKLNTGDEYENGSDSNHDDEEN
S.officinarum_TC80591           MGSACQAG-----TDGSSRKDVLGIGNVTLDNGHHEAEADADEWREKED-
S.officinarum_TC111093          MGSACQAG-----TDGSSRKDVLGIGNVTLDNGHHEAEADADEWREKED-
P.virgatum_TC42896              MGSACQAG-----LDG-PHKDVRGIANVATENGHHGAEADADEWREKEE-
S.officinarum_TC84822           MGSACQAG-----TDGSSRKDVLGIGNVTLDNGHHEAEADADEWREKED-
P.virgatum_TC22980              MGSTCQAG-----SDGP-HKDVRGIANGATENGHHGAEADADEWREKEE-
L.sativa_TC22097                VWRRCHS------SSGSGSESGTQAQKSVNSKSNVRYDNSSK--------
C.endivia_TA960_114280          IWRRCHS------SSGSGSESGTNAQKSVNSKRSD--DNVSI--------
C.maculosa_TA1222_215693        VWRRCHS------SSGSGSESGTQAQKSVNSKSNLRYNNGSDK-------
B.distachyon_DV469398           IWRR-----------------------------------------------
T.aestivum_TC284871             IWRRCHS------SSGS--ESGIHTQKCSKPKAGDEYENNSHD-------
H.vulgare_TC165635              IWRRCHS------SSGS--ESGIHIQKCSKPKTGDEYAKNSG--------
T.aestivum_c59844892@6464       IWRRCHS------SSGS--ESGIHTQKCSKPKAGDEYENNSG--------
                                :

Z.mays_GRMZM2G095727_T03        DDGDDDDFSVGLNARDGSDNGSGTQSSWTKRAVEIDSPQPISPD-QLVDP
Z.mays_GRMZM2G095727_T02        DDGDDDDFSVGLNARDGSDNGSGTQSSWTKRAVEIDSPQPISPD-QLVDP
Z.mays_GRMZM2G095727_T01        DDGDDDDFSVGLNARDGSDNGSGTQSSWTKRAVEIDSPQPISPD-QLVDP
S.officinarum_TC81972           DDDDDDDFSVGLNARDGSDNGSGTQSSWTKRAVEIDSPQPMSPD-QLADP
P.virgatum_TC27226              ----DDDLSVRLNARDGSDNGSGTQSSWTKRAVEIDSPQPMSPD-QLVDP
S.bicolor_Sb01g038820.1         DEDDDDDFSVGLNARDGSDNGSGTQSSWTKRAVEIDSPEPMSPD-QLADP
S.OFFICINARUM_PRR-LIKE          DDDDDDDFSVGLNARDGSDNGSGTQSSWTKRAVEIDSPQPMSPD-QLADP
Z.mays_GRMZM2G095727_T05        DDGDDDDFSVGLNARDGSDNGSGTQSSWTKRAVEIDSPQPISPD-QLVDP
Z.mays_GRMZM2G095727_T04        DDGDDDDFSVGLNARDGSDNGSGTQSSWTKRAVEIDSPQPISPD-QLVDP
S.officinarum_TC80591           ------DLANGRSAPPGMQQVD-EQKEQQGQSIHWERFLPVKTLRVLLVE
S.officinarum_TC111093          ------DLTNGRSAPPGMQQVD-EQKEQQGQSIHWERFLPVKTLRVLLVE
P.virgatum_TC42896              ------DLPNGHSAPPGAQQLD-EQKEQ--QTILWERFLPVKTLRVLLVE
S.officinarum_TC84822           ------DLANGHSAPPGMQQVD-EQKEQQGQSIHWERFLPVKTLRVLLVE
P.virgatum_TC22980              ------DLPNGHSAPLGAQQLD-EQKEQ--QTIQWERFLPVKTLRVLLVE
L.sativa_TC22097                ------DGDDNENTSGGSDDGSG-------------------------
C.endivia_TA960_114280          ------AGDDDGSTRDGSDDGSGTQSSWTKQAEASPC-----------
C.maculosa_TA1222_215693        ------DGNDNGSTSGGSDDGSGTQSSWTKQAVEPESPEAASPCDQITEH
B.distachyon_DV469398           -----------------------------------------------
T.aestivum_TC284871             ------DDDDCGSHDDDDDDDDD-------------------------
H.vulgare_TC165635              -----------GSHDDDDDDD--------------------------
T.aestivum_c59844892@6464       -----------GSHDDDDDDDDD-----------------------
```

FIGURE 2 (continued)

```
Z.mays_GRMZM2G095727_T03        PDSTCAQVIHPRSEICSNKWLPTANKRNVKKQKENKDESMGRYLGIGAPR
Z.mays_GRMZM2G095727_T02        PDSTCAQVIHPRSEICSNKWLPTANKRNVKKQKENKDESMGRYLGIGAPR
Z.mays_GRMZM2G095727_T01        PDSTCAQVIHPRSEICSNKWLPTANKRNVKKQKENKDESMGRYLGIGAPR
S.officinarum_TC81972           PDSTCAQVIHPKSEICSNKWLPTANKRNGKKHKENKDESMGRYLEIGAPR
P.virgatum_TC27226              PDSTCAQVIHPKSEICSNKWLPGANKRNSKKKKENKDESMGKYLEIGAPR
S.bicolor_Sb01g038820.1         PDSTCAQVIHPKSEICSNKWLPTANKRNGKKHKENKDESMGRYLEIGAPR
S.OFFICINARUM_PRR-LIKE          PDSTCAQVIHPKSEICSNKWLPTANKRNGKKHKENKDESMGRYLEIGAPR
Z.mays_GRMZM2G095727_T05        PDSTCAQVIHPRSEICSNKWLPTANKRNVKKQKENKDESMGRYLGIGAPR
Z.mays_GRMZM2G095727_T04        PDSTCAQVIHPRSEICSNKWLPTANKRNVKKQKENKDESMGRYLGIGAPR
S.officinarum_TC80591           NDDSTRQVVSALLRKCCYEVIPAENGSHAWRYLENLQNNIDLVL------
S.officinarum_TC111093          NDDSTRQVVSALLRKCCYEVIPAENGSHAWRYLENLQNNIDLVL------
P.virgatum_TC42896              NDDSTRQVVSALLRKCCYEVIPAENGLHAWQHLEDLQNNIDLVL------
S.officinarum_TC84822           NDDSTRQVVSALLRKCCYEVIPAENGSHAWRYLENLQNNIDLVL------
P.virgatum_TC22980              NDESTRQVVSALLRKCCYEVIPAENGLHAWQHLEDLQNNIDLVL------
L.sativa_TC22097                ---------HS-----------------------------------------
C.endivia_TA960_114280          -DSTCGLVIHSAETPTT----KDFQVQ----VDEG---------------
C.maculosa_TA1222_215693        PDSTCGLVIRSVHAQAT----RDSNDQEGRPYDEGKAKEIAECSFRNSEM
B.distachyon_DV469398           ------------------------------------------------
T.aestivum_TC284871             ------------------------------------------------
H.vulgare_TC165635              ------------------------------------------------
T.aestivum_c59844892@6464       ------------------------------------------------

Z.mays_GRMZM2G095727_T03        NSSAEYQSSLNDVSVNPIEKGHENHMSKCKSKKETMAEDDCTNMPSATNA
Z.mays_GRMZM2G095727_T02        NSSAEYQSSLNDVSVNPIEKGHENHMSKCKSKKETMAEDDCTNMPSATNA
Z.mays_GRMZM2G095727_T01        NSSAEYQSSLNDVSVNPIEKGHENHMSKCKSKKETMAEDDCTNMPSATNA
S.officinarum_TC81972           NSSAEYQSSLNDVSVNPTEKRHETHMPQCKSKKKMMAEDDCTDIPSETNT
P.virgatum_TC27226              NSSVEYQSSLNDTSVNPTGKRHEIHIPQFKSKKKVMADDDCTNMMSEPNT
S.bicolor_Sb01g038820.1         NSSAEYQSSLNDVSVNPTEKRHETHMPQCKSKKKMMAEDDCTDIPSEINT
S.OFFICINARUM_PRR-LIKE          NSSAEYQSSLNDVSVNPTEKRHETHMPQCKSKKKMMAEDDCTDIPSETNT
Z.mays_GRMZM2G095727_T05        NSSAEYQSSLNDVSVNPIEKGHENHMSKCKSKKETMAEDDCTNMPSATNA
Z.mays_GRMZM2G095727_T04        NSSAEYQSSLNDVSVNPIEKGHENHMSKCKSKKETMAEDDCTNMPSATNA
S.officinarum_TC80591           --TEVFMPCLSGIGL----------LSKITSHK------ICKDIPVISK-
S.officinarum_TC111093          --TEVFMPCLSGIGL----------LSKITSHK------ICKDIPVISK-
P.virgatum_TC42896              --TEVFMPRLSGIGL----------LSKITSHK------VCKDIPVIMMS
S.officinarum_TC84822           --TEVFMPCLSGIGL----------LSKITSHK------ICKDIPVIMMS
P.virgatum_TC22980              --TEVFMPCLSGIGL----------LSKITSHK------VCKDIPVIMMS
L.sativa_TC22097                ------------------------------------------------
C.endivia_TA960_114280          ------------------------------------DVEMDKELETGGS
C.maculosa_TA1222_215693        QIEFPIQAPVKHNGIEQSTHQAFDPTLNFKSKEMGISDIRREHSLNKQKA
B.distachyon_DV469398           ------------------------------------------------
T.aestivum_TC284871             ------------------------------------------------
H.vulgare_TC165635              ------------------------------------------------
T.aestivum_c59844892@6464       ------------------------------------------------
```

FIGURE 2 (continued)

```
Z.mays_GRMZM2G095727_T03        ETADLISSIARNTEGQQ--AVQAVDAPDGPSKMANGNDKNHDSHIEVTPH
Z.mays_GRMZM2G095727_T02        ETADLISSIARNTEGQQ--AVQAVDAPDGPSKMANGNDKNHDSHIEVTPH
Z.mays_GRMZM2G095727_T01        ETADLISSIARNTEGQQ--AVQAVDAPDGPSKMANGNDKNHDSHIEVTPH
S.officinarum_TC81972           ETADLISSIARNTEGQQ--AVRAVDAPDGPSKMPDGNDKNHDSHIEVTPH
P.virgatum_TC27226              ETADLISSIARNTEGQQ--AVQVADAPGCPSRIPDGNDKNHDSHIQVTPH
S.bicolor_Sb01g038820.1         ETADLISSIARNTEGQQ--AVRAVDAPDGPSKMPDGNDKNHDSHIVVTPH
S.OFFICINARUM_PRR-LIKE          ETADLISSIARNTEGQQ--AVRAVDAPDGPSKMPDGNGKNHDSHIEVTPH
Z.mays_GRMZM2G095727_T05        ETADLISSIARNTEGQQ--AVQAVDAPDGPSKMANGNDKNHDSHIEVTPH
Z.mays_GRMZM2G095727_T04        ETADLISSIARNTEGQQ--AVQAVDAPDGPSKMANGNDKNHDSHIEVTPH
S.officinarum_TC80591           --------------------------------------------------
S.officinarum_TC111093          --------------------------------------------------
P.virgatum_TC42896              SNDSMSMVFKCLSKGAVDFLVKPLRKNELKNLWQHVWRRCHSGFLFLVSL
S.officinarum_TC84822           SNDSMSMVFKCLSKGAVDFLVKPLRKNELKNLWQHVWRRCHS--------
P.virgatum_TC22980              SNDSMSMVFKCLSKGAVDFLVKPLRKNELKNLWQHVWRRCHSSSGSGSES
L.sativa_TC22097                --------------------------------------------------
C.endivia_TA960_114280          KDSKK------IGESEV--------VKDTNNSPHTDTKVMNESKETLA-E
C.maculosa_TA1222_215693        KDSKMPEPYMENEELEG--------QGEPENIMDANTKVLDDSNGAIVGE
B.distachyon_DV469398           --------------------------------------------------
T.aestivum_TC284871             ----------------------------------------ADDDFS-----
H.vulgare_TC165635              ----------------------------------------ADDDFS-----
T.aestivum_c59844892@6464       ----------------------------------------ADDDFS-----


Z.mays_GRMZM2G095727_T03        ELGLKRSRTNGATAEIHDERNILKRSDQSAFTRYHTSVASNQGGARYGES
Z.mays_GRMZM2G095727_T02        ELGLKRSRTNGATAEIHDERNILKRSDQSAFTRYHTSVASNQGGARYGES
Z.mays_GRMZM2G095727_T01        ELGLKRSRTNGATAEIHDERNILKRSDQSAFTRYHTSVASNQGGARYGES
S.officinarum_TC81972           ELGLKRLRTDGATAENP---------------------------------
P.virgatum_TC27226              ELGLKRLKTNGATTEIHDERNILRRSDLSAFTRYHTSVASNQGGARFGES
S.bicolor_Sb01g038820.1         ELGLKRLRTDGAADEIHDERNILKRSDQSAFTRYHTSVASNQGGARCGES
S.OFFICINARUM_PRR-LIKE          ELGLKRLRTDGATAEIHDERNILKRSDQSAFTRYHTSVASNQGGARCGES
Z.mays_GRMZM2G095727_T05        ELGLKRSRTNGATAEIHDERNILKRSDQSAFTRYHTSVASNQGGARYGES
Z.mays_GRMZM2G095727_T04        ELGLKRSRTNGATAEIHDERNILKRSDQSAFTRYHTSVASNQGGARYGES
S.officinarum_TC80591           --------------------------------------------------
S.officinarum_TC111093          --------------------------------------------------
P.virgatum_TC42896              FCWILILQNHDALMHSFFYSTSCSGFSLPLLLVGFASLVLSPVVVEVKVA
S.officinarum_TC84822           --------------------SSGS---------GSESGIQTQKCAKPNTG
P.virgatum_TC22980              AIQTQKCAKPNTSDEYENNSNSNHDDDEND---DDLSVGLNARDGSDNGS
L.sativa_TC22097                --------------------------------------------------
C.endivia_TA960_114280          FSLKRPRE----VTQVQNGPNILTHSELSAFTRYNT--------------
C.maculosa_TA1222_215693        LGLXRPRADKHSGTEVQTGCNILRHSELSAFTRYKTTLNAVKGTPGITTS
B.distachyon_DV469398           --------------------------------------------------
T.aestivum_TC284871             --------------------------------------------------
H.vulgare_TC165635              --------------------------------------------------
T.aestivum_c59844892@6464       --------------------------------------------------
```

FIGURE 2 (continued)

```
Z.mays_GRMZM2G095727_T03        SSPQDNSSEAMKTDSTCKMKSNSDAAPIKQGSNGSSNN-DVGSSTKNVAA
Z.mays_GRMZM2G095727_T02        SSPQDNSSEAMKTDSTCKMKSNSDAAPIKQGSNGSSNN-DVGSSTKNVAA
Z.mays_GRMZM2G095727_T01        SSPQDNSSEAMKTDSTCKMKSNSDAAPIKQGSNGSSNN-DVGSSTKNVAA
S.officinarum_TC81972           ------------------------------------------------
P.virgatum_TC27226              SSPQDNSSEAVKTDSTCKMKSNSDAPPIKQGSNGSSNNNDMGSSTKNVVV
S.bicolor_Sb01g038820.1         SSPQDNSSEAVKTDSTCKMKSNSDAAPIKQGSNGSSNN-DVGSSTKNVIA
S.OFFICINARUM_PRR-LIKE          SSPQDNSSEAVKTDSTCKMKSNSDAAPIKQGSNGSSNN-DVGSSTKNVVA
Z.mays_GRMZM2G095727_T05        SSPQDNSSEAMKTDSTCKMKSNSDAAPIKQGSNGSSNN-DVGSSTKNVAA
Z.mays_GRMZM2G095727_T04        SSPQDNSSEAMKTDSTCKMKSNSDAAPIKQGSNGSSNN-DVGSSTKNVAA
S.officinarum_TC80591           ------------------------------------------------
S.officinarum_TC111093          ------------------------------------------------
P.virgatum_TC42896              SRHKSVPNQILVTSMRTTVTAIMMMTKMMT--------------------
S.officinarum_TC84822           DEYENDSDSNHDDEENDEHNDDNFSVRLQC--------------------
P.virgatum_TC22980              GTQSSWTKRAVEIDSPQPMSPDQLVDPPDSTCAQVI--------------
L.sativa_TC22097                ------------------------------------------------
C.endivia_TA960_114280          ----------------------------TSKQD--------VPGKSSDI
C.maculosa_TA1222_215693        RSQPDYRSNDVKKESKRDALSDGYLIYQGSSEQV--------IPSKAEGM
B.distachyon_DV469398           ------------------------------------------------
T.aestivum_TC284871             ------------------------------------------------
H.vulgare_TC165635              ------------------------------------------------
T.aestivum_c59844892@6464       ------------------------------------------------

Z.mays_GRMZM2G095727_T03        RPSGDRERVASPLAIKSTQHASAFHTIQNQTSPANLIGEDKADEGISNTV
Z.mays_GRMZM2G095727_T02        RPSGDRERVASPLAIKSTQHASAFHTIQNQTSPANLIGEDKADEGISNTV
Z.mays_GRMZM2G095727_T01        RPSGDRERVASPLAIKSTQHASAFHTIQNQTSPANLIGEDKADEGISNTV
S.officinarum_TC81972           ------------------------------------------------
P.virgatum_TC27226              KPSGNRERVTSPSAAKSIQHTSAFHPMPHQTSPANVVGKDKTDEGISNAV
S.bicolor_Sb01g038820.1         KPSANRERVTSPSAIKSTQHASAFHTIQNQTSPANLVGKDKADEGISNAV
S.OFFICINARUM_PRR-LIKE          KPSANRERVTSPSTIKSTQHASAFHTIHNQTSPANLVGKDKADEGISNAV
Z.mays_GRMZM2G095727_T05        RPSGDRERVASPLAIKSTQHASAFHTIQNQTSPANLIGEDKADEGISNTV
Z.mays_GRMZM2G095727_T04        RPSGDRERVASPLAIKSTQHASAFHTIQNQTSPANLIGEDKADEGISNTV
S.officinarum_TC80591           ------------------------------------------------
S.officinarum_TC111093          ------------------------------------------------
P.virgatum_TC42896              ------------------------------------------------
S.officinarum_TC84822           ------------------------------------------------
P.virgatum_TC22980              ------------------------------------------------
L.sativa_TC22097                ------------------------------------------------
C.endivia_TA960_114280          P-GADLVGPPQVHHIHHHHHVHHYHNIESNQPPD----------------
C.maculosa_TA1222_215693        PPGGVLHQEHRIQHIHHHHHVHHYHNLEEEQPPS----------------
B.distachyon_DV469398           ------------------------------------------------
T.aestivum_TC284871             ------------------------------------------------
H.vulgare_TC165635              ------------------------------------------------
T.aestivum_c59844892@6464       ------------------------------------------------
```

FIGURE 2 (continued)

```
Z.mays_GRMZM2G095727_T03        KMSHPTEVPQGCVQHHHHVHYYLHVMTQKQPSTDRGSSDVHCGSSNVFDP
Z.mays_GRMZM2G095727_T02        KMSHPTEVPQGCVQHHHHVHYYLHVMTQKQPSTDRGSSDVHCGSSNVFDP
Z.mays_GRMZM2G095727_T01        KMSHPTEVPQGCVQHHHHVHYYLHVMTQKQPSTDRGSSDVHCGSSNVFDP
S.officinarum_TC81972           --------------------------------------------------
P.virgatum_TC27226              KVG-QTEVPQSCVQHHHHVHYYLHVMTQKQPSIDHGSSDAQCGSSNVFDP
S.bicolor_Sb01g038820.1         KMSHPTEVPQSCVQHHHHVHYYLHVMTQKQSSIDRGSSDVQCGSSNVFDP
S.OFFICINARUM_PRR-LIKE          KMSHPTEVPQSCVQHHHHVHYYLHVLTQKQLSIDRGSSDVQCGSSNVFDP
Z.mays_GRMZM2G095727_T05        KMSHPTEVPQGCVQHHHHVHYYLHVMTQKQPSTDRGSSDVHCGSSNVFDP
Z.mays_GRMZM2G095727_T04        KMSHPTEVPQGCVQHHHHVHYYLHVMTQKQPSTDRGSSDVHCGSSNVFDP
S.officinarum_TC80591           --------------------------------------------------
S.officinarum_TC111093          --------------------------------------------------
P.virgatum_TC42896              --------------------------------------------------
S.officinarum_TC84822           --------------------------------------------------
P.virgatum_TC22980              --------------------------------------------------
L.sativa_TC22097                --------------------------------------------------
C.endivia_TA960_114280          ------------------NHEESGLKNLAAIAPHCGSSNVLGGGG-VEG
C.maculosa_TA1222_215693        ------------------NHDDFGLNRLGADAPYCGSSNIVGGPGPVEG
B.distachyon_DV469398           --------------------------------------------------
T.aestivum_TC284871             --------------------------------------------------
H.vulgare_TC165635              --------------------------------------------------
T.aestivum_c59844892@6464       --------------------------------------------------

Z.mays_GRMZM2G095727_T03        PVEGHAANYSVNGGVSVGHNGCNGQNGSSAVPNIARPNIESINGTMSQNI
Z.mays_GRMZM2G095727_T02        PVEGHAANYSVNGGVSVGHNGCNGQNGSSAVPNIARPNIESINGTMSQNI
Z.mays_GRMZM2G095727_T01        PVEGHAANYSVNGGVSVGHNGCNGQNGSSAVPNIARPNIESINGTMSQNI
S.officinarum_TC81972           --------------------------------------------------
P.virgatum_TC27226              PVEGHAANYSVNGGVSVGHNGCNGQNGSSAAPNIARPNMECVNGTMSKNV
S.bicolor_Sb01g038820.1         PVEGHAANYSVNGGVSVGHNGCNGQNGTSTVPNIARPNIESVNGTVSQNI
S.OFFICINARUM_PRR-LIKE          PVEGHAANYSVNGGVSVGHNGCNGQNGTSAVPNIATPNIESVNGTMSQNI
Z.mays_GRMZM2G095727_T05        PVEGHAANYSVNGGVSVGHNGCNGQNGSSAVPNIARPNIESINGTMSQNI
Z.mays_GRMZM2G095727_T04        PVEGHAANYSVNGGVSVGHNGCNGQNGSSAVPNIARPNIESINGTMSQNI
S.officinarum_TC80591           --------------------------------------------------
S.officinarum_TC111093          --------------------------------------------------
P.virgatum_TC42896              --------------------------------------------------
S.officinarum_TC84822           --------------------------------------------------
P.virgatum_TC22980              --------------------------------------------------
L.sativa_TC22097                --------------------------------------------------
C.endivia_TA960_114280          IIGNCSLN--GSGSGSK--HGSNGQNGSSAAVNLEGKNVESGGGGGLAGK
C.maculosa_TA1222_215693        NIENYSLNRSASGTGSK--HGSNLPNGSNAAVNFEATNVESDVGGLVISG
B.distachyon_DV469398           --------------------------------------------------
T.aestivum_TC284871             -----------VGPNAR--DGSD--NGSGT--------------------
H.vulgare_TC165635              -----------VGPNAR--DGSD--NGSGTQSSWTKRAVEIDSPQLVSSD
T.aestivum_c59844892@6464       -----------VGPNAR--DGSD--NGSGTQSSWTSVLWRLIV-------
```

FIGURE 2 (continued)

```
Z.mays_GRMZM2G095727_T03        AGGGIVSGSGSGNDMYQNRFLQREAALNKFRLKRKDRNFGKKVACFQLHA
Z.mays_GRMZM2G095727_T02        AGGGIVSGSGSGNDMYQNRFLQREAALNKFRLKRKDRNFGSLPK--QEEA
Z.mays_GRMZM2G095727_T01        AGGGIVSGSGSGNDMYQNRFLQREAALNKFRLKRKDRNFGKKVRYQSRKR
S.officinarum_TC81972           --------------------------------------------------
P.virgatum_TC27226              AGGG--SGSGSGNDMYQNRFPQREAALNKFRLKRKDRNFDKKVRYQSRKR
S.bicolor_Sb01g038820.1         AGGGIVSGSGSGNDVYQNRFPQREAALNKFRLKRKDRNFGKKVRYQSRKR
S.OFFICINARUM_PRR-LIKE          AGGGIVSGSGSGNDVYQNRFPQREAALNKFRLKRKDRNFGKKVRYQSRKR
Z.mays_GRMZM2G095727_T05        AGGGIVSGSGSGNDMYQNRFLQREAALNKFRLKRKDRNFGKKVRYQSRKR
Z.mays_GRMZM2G095727_T04        AGGGIVSGSGSGNDMYQNRFLQREAALNKFRLKRKDRNFGSLPKQEEAC-
S.officinarum_TC80591           --------------------------------------------------
S.officinarum_TC111093          --------------------------------------------------
P.virgatum_TC42896              --------------------------------------------------
S.officinarum_TC84822           --------------------------------------------------
P.virgatum_TC22980              --------------------------------------------------
L.sativa_TC22097                --------------------------------------------------
C.endivia_TA960_114280          SGGGGGGEN----RIDEEKSSQREAALMKFREKRKNRCFQKKVRYQNRKR
C.maculosa_TA1222_215693        SGDASESASGTGIIMDRRNSSQREAALNKFRQKREVRCFQKNVRYQNRKK
B.distachyon_DV469398           --------------------------------------------------
T.aestivum_TC284871             --------------------------------------------------
H.vulgare_TC165635              HLSDSPDSTCAQVIHPRSEIGSNRGADCK--------------------
T.aestivum_c59844892@6464       --------------------------------------------------


Z.mays_GRMZM2G095727_T03        CCSYDLKSVIVI----------------
Z.mays_GRMZM2G095727_T02        C---------------------------
Z.mays_GRMZM2G095727_T01        LAEQRPRVRGQFVRQSEQEDQTAQGSER
S.officinarum_TC81972           ----------------------------
P.virgatum_TC27226              LAEQRPRVRGQFVRQSGQEDKAGQDSER
S.bicolor_Sb01g038820.1         LAEQRPRVRGQFVRQSGQEDQAAQGSER
S.OFFICINARUM_PRR-LIKE          LAEQRPRVRGQFVRQSGQEDQAAQGSER
Z.mays_GRMZM2G095727_T05        LAEQRPRVRGQFVRQSEQEDQTAQGSER
Z.mays_GRMZM2G095727_T04        ----------------------------
S.officinarum_TC80591           ----------------------------
S.officinarum_TC111093          ----------------------------
P.virgatum_TC42896              ----------------------------
S.officinarum_TC84822           ----------------------------
P.virgatum_TC22980              ----------------------------
L.sativa_TC22097                ----------------------------
C.endivia_TA960_114280          LAEERPRVRGQFVKQTGQESSSNGEDR-
C.maculosa_TA1222_215693        LAEQRPRVR-------------------
B.distachyon_DV469398           ----------------------------
T.aestivum_TC284871             ----------------------------
H.vulgare_TC165635              ----------------------------
T.aestivum_c59844892@6464       ----------------------------
```

FIGURE 2 (continued)

	1	2	3	4	5	6	7	8	9	10	11
1. C.maculosa_TA1222_215693		18,6	44,3	25,4	31,8	32,0	33,1	33,0	29,6	30,9	34,2
2. L.sativa_TC22097	21,2		30,2	15,9	15,9	15,8	11,0	11,3	16,2	15,6	10,7
3. C.endivia_TA960_114280	58,8	32,8		22,9	33,4	33,9	26,9	26,8	30,7	33,6	30,0
4. S.officinarum_TC81972	33,9	28,0	32,8		37,3	36,9	44,3	45,3	35,1	34,0	43,1
5. S.officinarum_PRR	42,2	18,3	43,9	37,5		96,5	67,8	67,5	86,3	91,0	71,8
6. S.bicolor_Sb01g038820.1	42,2	18,3	44,2	37,3	97,8		67,6	67,3	86,1	90,9	71,6
7. Z.mays_GRMZM2G095727_T03	49,2	16,7	43,1	46,3	71,1	71,1		96,8	75,1	74,4	94,5
8. Z.mays_GRMZM2G095727_T02	49,8	17,1	43,2	47,4	70,6	70,6	96,8		77,3	74,2	94,2
9. Z.mays_GRMZM2G095727_T04	41,5	18,8	42,1	36,6	90,9	90,7	76,4	77,3		94,7	73,6
10. Z.mays_GRMZM2G095727_T05	42,2	18,3	44,5	35,5	94,9	94,6	74,9	74,4	95,7		78,7
11. Z.mays_GRMZM2G095727_T01	49,9	16,3	45,8	45,1	74,7	74,6	95,2	94,5	76,7	78,7	
12. S.officinarum_TC80591	16,3	38,8	18,5	27,0	21,0	20,1	16,4	15,7	21,0	20,2	16,1
13. P.virgatum_TC27226	51,0	16,6	45,8	44,3	72,3	72,5	86,5	85,6	68,8	70,8	89,9
14. S.officinarum_TC111093	16,3	38,8	18,1	27,0	20,9	20,0	15,5	15,7	20,9	20,1	16,1
15. S.officinarum_TC84822	22,7	51,6	29,3	34,6	32,9	32,4	22,3	22,6	33,1	32,0	21,7
16. P.virgatum_TC42896	25,1	41,4	32,8	35,9	31,5	31,9	24,4	23,2	33,0	31,9	23,2
17. P.virgatum_TC22980	26,2	47,8	34,5	43,0	35,9	36,1	25,4	25,3	37,3	36,2	24,4
18. T.aestivum_TC284871	23,6	48,9	30,4	31,5	28,5	28,8	23,3	23,9	29,5	28,6	22,7
19. H.vulgare_TC165635	26,6	41,4	33,7	41,4	33,7	34,2	23,0	23,5	35,0	34,1	23,7
20. T.aestivum_c59844892@6464	23,4	53,0	31,0	34,6	26,9	26,9	20,3	20,2	27,7	27,0	19,2
21. B.distachyon_DV469398	11,2	35,9	12,5	12,1	8,1	8,1	6,5	6,6	8,3	8,1	6,3

FIGURE 3

	12	13	14	15	16	17	18	19	20	21
1. C.maculosa_TA1222_215693	10,0	35,9	10,1	13,6	14,4	14,1	14,2	15,9	13,7	10,4
2. L.sativa_TC22097	22,5	11,2	22,5	44,9	35,2	40,6	42,5	37,0	46,2	34,1
3. C.endivia_TA960_114280	10,8	30,6	10,6	21,8	21,9	23,9	22,2	23,9	23,5	11,4
4. S.officinarum_TC81972	16,6	41,0	16,6	18,9	17,1	26,9	14,0	25,5	20,5	10,3
5. S.officinarum_PRR	20,9	68,2	20,8	32,5	28,3	33,9	24,9	30,0	24,0	7,8
6. S.bicolor_Sb01g038820.1	19,5	68,3	19,3	31,6	28,4	34,1	25,2	30,2	24,3	7,8
7. Z.mays_GRMZM2G095727_T03	9,7	78,7	9,5	13,4	13,1	12,9	13,9	13,5	10,8	5,9
8. Z.mays_GRMZM2G095727_T02	9,1	78,4	9,1	13,9	13,9	12,5	14,3	13,8	10,9	6,0
9. Z.mays_GRMZM2G095727_T04	20,2	61,3	20,1	31,6	29,7	35,3	26,3	31,7	24,7	8,1
10. Z.mays_GRMZM2G095727_T05	19,3	65,7	19,2	30,5	28,7	34,1	25,5	30,7	23,9	7,8
11. Z.mays_GRMZM2G095727_T01	8,8	83,3	8,5	13,4	13,9	12,8	13,6	12,5	10,4	5,7
12. S.officinarum_TC80591		8,9	99,4	62,9	46,8	45,3	40,3	34,7	32,0	9,3
13. P.virgatum_TC27226	15,1		9,2	13,4	13,0	16,9	12,3	14,5	12,7	5,8
14. S.officinarum_TC111093	99,4	15,6		62,5	46,8	45,3	40,3	34,7	32,0	9,3
15. S.officinarum_TC84822	62,9	20,6	62,5		63,8	73,0	64,7	56,3	56,4	23,4
16. P.virgatum_TC42896	50,0	22,6	50,0	70,3		69,4	52,9	52,2	47,1	20,0
17. P.virgatum_TC22980	49,5	21,6	49,5	79,2	76,1		62,1	69,3	59,3	20,1
18. T.aestivum_TC284871	46,3	21,6	46,3	76,3	61,4	74,1		77,0	74,9	22,2
19. H.vulgare_TC165635	39,2	24,2	39,2	65,3	60,5	77,7	80,6		68,4	19,4
20. T.aestivum_c59844892@6464	40,6	19,7	40,6	71,9	57,6	68,9	81,5	70,4		24,3
21. B.distachyon_DV469398	21,6	6,4	21,6	24,2	21,0	21,2	22,6	19,4	24,3	

FIGURE 3 (continued)

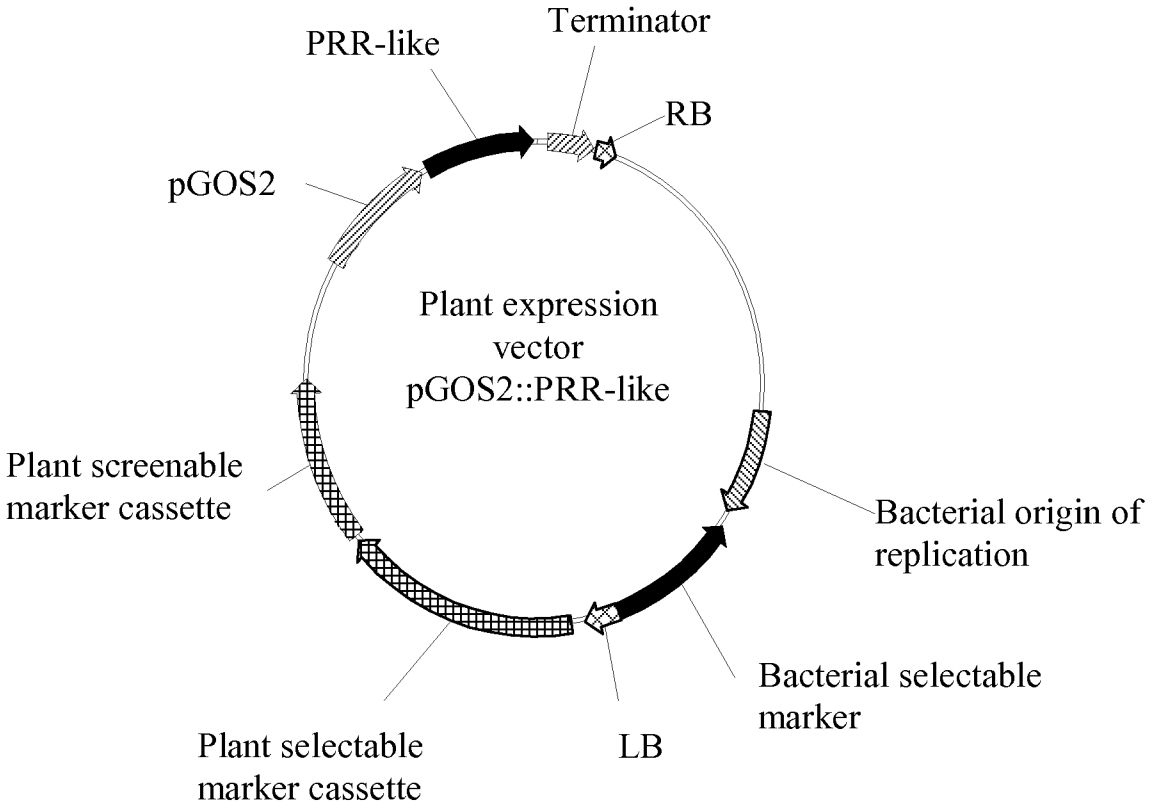
Terminator
PRR-like
RB
pGOS2
Plant expression
vector
pGOS2::PRR-like
Plant screenable
marker cassette
Bacterial origin of
replication
Bacterial selectable
marker
Plant selectable
marker cassette
LB

FIGURE 4

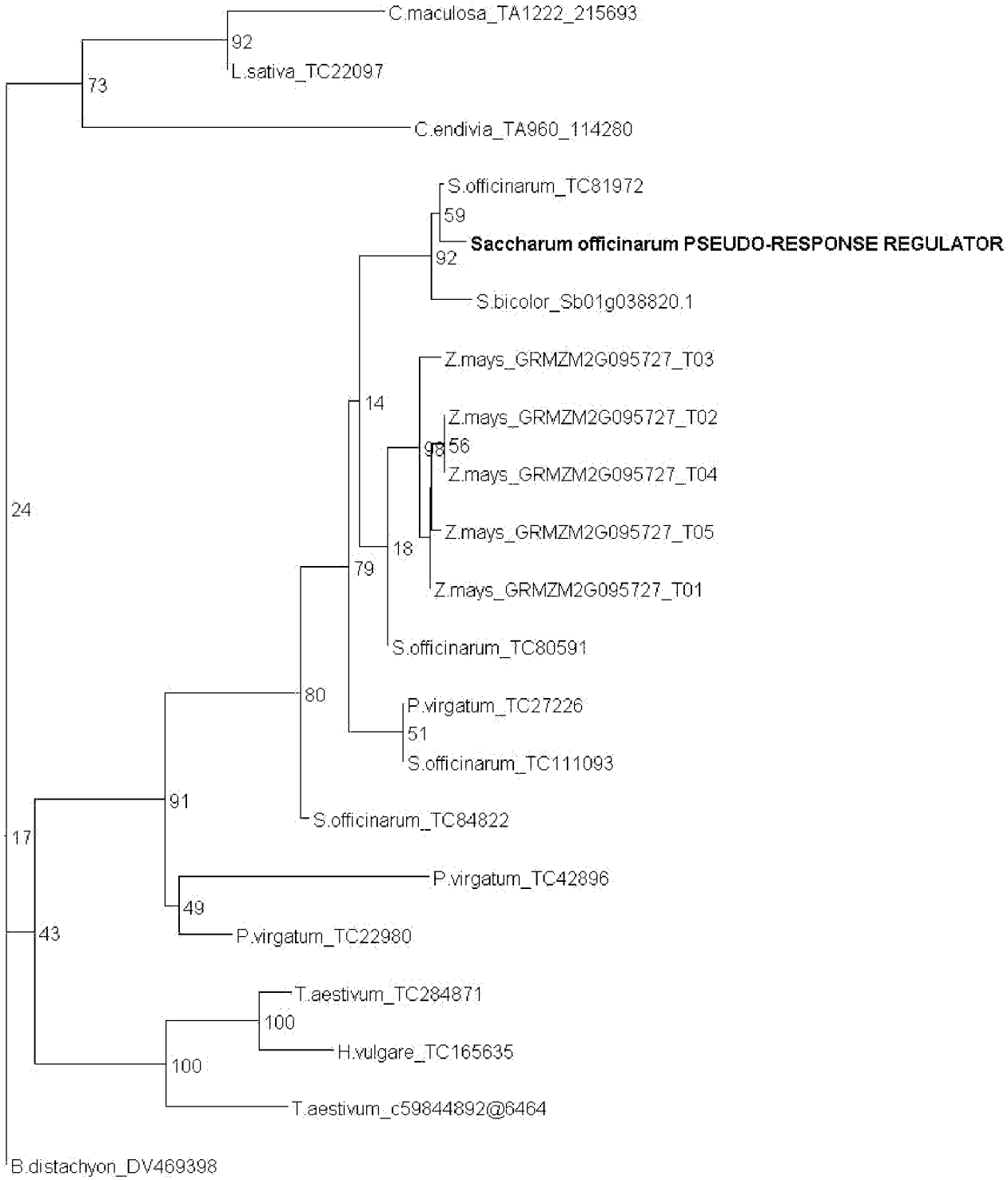
C.maculosa_TA1222_215693
92
L.sativa_TC22097
73
C.endivia_TA960_114280
S.officinarum_TC81972
59
Saccharum officinarum PSEUDO-RESPONSE REGULATOR
92
S.bicolor_Sb01g038820.1
Z.mays_GRMZM2G095727_T03
14
Z.mays_GRMZM2G095727_T02
56
Z.mays_GRMZM2G095727_T04
Z.mays_GRMZM2G095727_T05
18
79
Z.mays_GRMZM2G095727_T01
S.officinarum_TC80591
24
80
P.virgatum_TC27226
51
S.officinarum_TC111093
91
S.officinarum_TC84822
17
P.virgatum_TC42896
49
43
P.virgatum_TC22980
T.aestivum_TC284871
100
H.vulgare_TC165635
100
T.aestivum_c59844892@6464
B.distachyon_DV469398

FIGURE 5

SEQUENCE LISTING

<110> BASF Plant Science Company GmbH

<120> Plants having enhanced yield-related traits and a method for making the same

<130> PF73120_PCT

<150> EP 11190757.2
<151> 2011-11-25

<150> US 61/563,625
<151> 2011-11-25

<160> 48

<170> PatentIn version 3.5

<210> 1
<211> 2301
<212> DNA
<213> Saccharum officinarum

<400> 1
atgggtagcg cttgccaagc tggcacggac gggtcttccc gcaaggatgt gttggggata 60

gggaatgtca ccttagataa tggccaccat gaggctgaag ctgatgcaga tgaatggagg 120

gaaaaggaag atgacttagc caatgggcgc agtgcgccac cgggcatgca gcaggtggat 180

gaacagaagg agcaacaagg acaaagcatt cactgggaga ggttcctacc tgtgaagaca 240

ctgagagtct tgctggtgga gaatgatgac tctactcgtc aggtggtcag tgccctgctc 300

cgtaagtgct gctatgaagt tatccctgct gaaaatggtt cacatgcatg gcgatatctt 360

gaaaatctgc agaacaacat tgaccttgta ttgactgagg ttttcatgcc ttgtctatct 420

ggcatcggtc tgcttagcaa aatcactagt cacaaaattt gcaaggacat tcctgtgatt 480

atgatgtctt caaatgactc tatgagtatg gtgttttaagt gttttgtcgaa gggagcagtt 540

gacttcttgg taaagccact acgtaagaat gagcttaaga acctttggca gcacgttttgg 600

aggcgatgcc acagttccag tggcagtgga agtgaaagtg gcatccagac acagaagtat 660

gccaaaccaa atactggtga cgagtatgag aacgacagtg acagcaatca tgatgatgaa 720

gaaaatgacg acgacgacga tgacgacttc agtgtcggac tcaatgctag ggatggaagt 780

gataatggca gtggtactca aagctcatgg acaaagcgtg ctgtggagat tgacagtcca 840

caacctatgt ctcctgatca actagcagat ccacctgata gtacatgtgc acaagtgatt 900

caccccaaat cagagatatg cagtaacaag tggctaccga cagcaaacaa aaggaatggc 960

aagaaacata aggagaataa agatgaatct atgggaagat acctagaaat aggtgctcct 1020

aggaactcaa gtgcagaata tcaatcatct ctcaatgacg tatctgttaa tccaacagaa 1080

aaacgacatg agactcacat gccccaatgc aaatccaaaa agaaaatgat ggcagaagat 1140

gattgtacag acatacctag tgaaacaaat actgaaactg ctgatttaat tagctcaata 1200

gccagaaaca cagaaggcca acaagcagta cgagccgttg atgcacctga tggccttcc 1260

aagatgcccg atggaaatgg taagaatcat gattctcata tcgaggtgac accccatgag 1320

ttgggtttga agagattgag aacagatgga gctacagccg aaatccatga tgagcgaaat 1380

attctgaaaa gatcagatca gtcagccttc accaggtacc atacatctgt ggcttccaat 1440

caaggtggag caagatgtgg ggaaagctct tcaccacaag ataacagttc tgaggctgtg 1500

aaaacggact ctacatgcaa gatgaagtca aattcagatg ctgctccaat aaagcaggga 1560

tccaatggca gtagcaacaa cgatgtgggc tccagtacaa agaatgttgt tgcaaagcct 1620

tcggctaaca gggagagagt aacgtcacca tcaaccatca aatctaccca gcatgcctca 1680

gcattttcata ctatacataa tcaaacatca ccagctaatc tggttgggaa agacaaagct 1740

gatgaaggaa tttccaatgc agtgaaaatg agccacccaa cagaggttcc acaaagctgc 1800

gtccagcatc atcatcacgt gcattattac ctccatgttt tgacacagaa acagctatcc 1860

atcgaccgtg gatcatcaga tgttcagtgt ggttcatcaa atgtgtttga tcctcctgtt 1920

gaaggacatg ctgctaacta cagtgtgaat gggggtgtct cagttggtca taatgggtgc 1980

aatgggcaga atggaacgag cgctgtcccc aatattgcaa caccaaacat agagagtgtt 2040

aatggtacca tgagccaaaa tatcgctgga ggtggcattg taagtgggag tgggagtggc 2100

aatgatgttt atcagaatcg gttccccccaa cgagaagctg cattgaacaa attcagactg 2160

aagcggaaag atcggaactt cggtaaaaag gttcgctacc aaagcaggaa gaggcttgct 2220

gagcaacggc cacgggtccg tggacagttt gtgcgacaat ctgggcaaga agatcaagca 2280

gcgcaaggtt cagaaagatg a 2301

<210> 2
<211> 766
<212> PRT
<213> Saccharum officinarum

<400> 2

Met Gly Ser Ala Cys Gln Ala Gly Thr Asp Gly Ser Ser Arg Lys Asp
1 5 10 15

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

Glu Ala Asp Ala Asp Glu Trp Arg Glu Lys Glu Asp Asp Leu Ala Asn
35 40 45

Gly Arg Ser Ala Pro Pro Gly Met Gln Gln Val Asp Glu Gln Lys Glu
50 55 60

Gln Gln Gly Gln Ser Ile His Trp Glu Arg Phe Leu Pro Val Lys Thr
65 70 75 80

Leu Arg Val Leu Leu Val Glu Asn Asp Asp Ser Thr Arg Gln Val Val
85 90 95

Ser Ala Leu Leu Arg Lys Cys Cys Tyr Glu Val Ile Pro Ala Glu Asn

```
                100                 105                 110

Gly Ser His Ala Trp Arg Tyr Leu Glu Asn Leu Gln Asn Asn Ile Asp
        115                 120                 125

Leu Val Leu Thr Glu Val Phe Met Pro Cys Leu Ser Gly Ile Gly Leu
    130                 135                 140

Leu Ser Lys Ile Thr Ser His Lys Ile Cys Lys Asp Ile Pro Val Ile
145                 150                 155                 160

Met Met Ser Ser Asn Asp Ser Met Ser Met Val Phe Lys Cys Leu Ser
                165                 170                 175

Lys Gly Ala Val Asp Phe Leu Val Lys Pro Leu Arg Lys Asn Glu Leu
            180                 185                 190

Lys Asn Leu Trp Gln His Val Trp Arg Arg Cys His Ser Ser Ser Gly
        195                 200                 205

Ser Gly Ser Glu Ser Gly Ile Gln Thr Gln Lys Tyr Ala Lys Pro Asn
    210                 215                 220

Thr Gly Asp Glu Tyr Glu Asn Asp Ser Asp Ser Asn His Asp Asp Glu
225                 230                 235                 240

Glu Asn Asp Asp Asp Asp Asp Asp Asp Phe Ser Val Gly Leu Asn Ala
                245                 250                 255

Arg Asp Gly Ser Asp Asn Gly Ser Gly Thr Gln Ser Ser Trp Thr Lys
            260                 265                 270

Arg Ala Val Glu Ile Asp Ser Pro Gln Pro Met Ser Pro Asp Gln Leu
        275                 280                 285

Ala Asp Pro Pro Asp Ser Thr Cys Ala Gln Val Ile His Pro Lys Ser
    290                 295                 300

Glu Ile Cys Ser Asn Lys Trp Leu Pro Thr Ala Asn Lys Arg Asn Gly
305                 310                 315                 320

Lys Lys His Lys Glu Asn Lys Asp Glu Ser Met Gly Arg Tyr Leu Glu
                325                 330                 335

Ile Gly Ala Pro Arg Asn Ser Ser Ala Glu Tyr Gln Ser Ser Leu Asn
            340                 345                 350

Asp Val Ser Val Asn Pro Thr Glu Lys Arg His Glu Thr His Met Pro
        355                 360                 365

Gln Cys Lys Ser Lys Lys Lys Met Met Ala Glu Asp Asp Cys Thr Asp
```

370 375 380

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

Ala Arg Asn Thr Glu Gly Gln Gln Ala Val Arg Ala Val Asp Ala Pro
405 410 415

Asp Gly Pro Ser Lys Met Pro Asp Gly Asn Gly Lys Asn His Asp Ser
420 425 430

His Ile Glu Val Thr Pro His Glu Leu Gly Leu Lys Arg Leu Arg Thr
435 440 445

Asp Gly Ala Thr Ala Glu Ile His Asp Glu Arg Asn Ile Leu Lys Arg
450 455 460

Ser Asp Gln Ser Ala Phe Thr Arg Tyr His Thr Ser Val Ala Ser Asn
465 470 475 480

Gln Gly Gly Ala Arg Cys Gly Glu Ser Ser Ser Pro Gln Asp Asn Ser
485 490 495

Ser Glu Ala Val Lys Thr Asp Ser Thr Cys Lys Met Lys Ser Asn Ser
500 505 510

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

Val Gly Ser Ser Thr Lys Asn Val Val Ala Lys Pro Ser Ala Asn Arg
530 535 540

Glu Arg Val Thr Ser Pro Ser Thr Ile Lys Ser Thr Gln His Ala Ser
545 550 555 560

Ala Phe His Thr Ile His Asn Gln Thr Ser Pro Ala Asn Leu Val Gly
565 570 575

Lys Asp Lys Ala Asp Glu Gly Ile Ser Asn Ala Val Lys Met Ser His
580 585 590

Pro Thr Glu Val Pro Gln Ser Cys Val Gln His His His His Val His
595 600 605

Tyr Tyr Leu His Val Leu Thr Gln Lys Gln Leu Ser Ile Asp Arg Gly
610 615 620

Ser Ser Asp Val Gln Cys Gly Ser Ser Asn Val Phe Asp Pro Pro Val
625 630 635 640

Glu Gly His Ala Ala Asn Tyr Ser Val Asn Gly Gly Val Ser Val Gly

645 650 655

His Asn Gly Cys Asn Gly Gln Asn Gly Thr Ser Ala Val Pro Asn Ile
660 665 670

Ala Thr Pro Asn Ile Glu Ser Val Asn Gly Thr Met Ser Gln Asn Ile
675 680 685

Ala Gly Gly Gly Ile Val Ser Gly Ser Gly Ser Gly Asn Asp Val Tyr
690 695 700

Gln Asn Arg Phe Pro Gln Arg Glu Ala Ala Leu Asn Lys Phe Arg Leu
705 710 715 720

Lys Arg Lys Asp Arg Asn Phe Gly Lys Lys Val Arg Tyr Gln Ser Arg
725 730 735

Lys Arg Leu Ala Glu Gln Arg Pro Arg Val Arg Gly Gln Phe Val Arg
740 745 750

Gln Ser Gly Gln Glu Asp Gln Ala Ala Gln Gly Ser Glu Arg
755 760 765

<210> 3
<211> 1638
<212> DNA
<213> Centaurea maculosa

<220>
<221> misc_feature
<222> (856)..(856)
<223> n is a, c, g, or t

<400> 3
atgccttgtc tttcaggaat tggtctatta tgcaagatta tgagccacaa gacacgcaag 60

aatatccctg tgattatgat gtcttctcat gattcaatgg gttttggttt taagtgttta 120

tcaaaaggtg cagtagattt tctagtgaag ccagttcgga aaaatgagct taaaaacctt 180

tggcagcatg tgtggaggag gtgtcacagc tctagtggta gcgggagcga aagtggcaca 240

caggcccaaa aatctgtaaa ctcaaaaagc aatttaaggt acaataatgg cagcgacaaa 300

gatgggaatg acaatgggag caccagtggt ggcagcgatg atggtagtgg cactcagagt 360

tcttggacca aacaagctgt tgaacctgag agcccagaag cagcatctcc atgtgaccag 420

ataactgagc atccagacag cacttgcggc ctcgttatcc gttctgtaca tgctcaagca 480

acaagagact ccaatgatca agaaggacga ccatatgatg aaggaaaggc caaagagatt 540

gcagaatgca gctttagaaa ctcagaaatg caaattgagt ttccaattca ggctcctgta 600

aaacataatg gtatagaaca gagcactcat caagcatttg accccacatt gaacttaaa 660

agcaaggaga tggggatttc agatatcaga agggagcact cgttgaacaa acaaaaagct 720

aaagatagca aaatgcctga accttacatg gaaaatgaag aacttgaggg tcaaggagaa 780

cctgaaaaca ttatggatgc aaataccaag gttcttgatg attctaatgg agcaatagtg 840

ggtgagcttg gcttanagag gccgcgggca gataaacata gtgggacaga agttcagact 900

ggctgcaata tattaagaca ttcagagctt tcagccttca cgaggtacaa aacaacctta 960

aacgctgtta aaggtacccc tggaatcacc actagccgtt ctcaacctga ttatagatca 1020

aatgatgtga agaaagaatc caagcgtgat gcacttttcag atggatatct tatttatcaa 1080

ggctcaagtg agcaagtcat accaagcaag gccgaaggca tgccacctgg cggtgtgctg 1140

catcaagagc accgtattca acacatccat caccatcacc atgttcatca ttaccataac 1200

ttagaagaag agcagccacc gtccaatcat gatgattttg gcttaaacag gttgggtgca 1260

gatgctccat actgtgggtc atcaaatatc gtgggcgggc ccggacccgt tgaaggtaac 1320

attgaaaatt atagttttgaa cagaagtgcc tcgggcacgg gcagcaagca tggaagcaat 1380

ttgccgaacg gaagtaacgc tgctgttaat tttgaagcca caaatgtaga aagcgatgtt 1440

ggtggttttag tcataagtgg aagtggtgat gctagtgaga gtgccagtgg caccggaatt 1500

ataatggatc gacgcaactc ttcacagaga gaagcagcct tgaataagtt ccgccaaaag 1560

agagaagttc gatgcttcca aaagaacgtg cggtatcaaa atcgaaagaa actggctgaa 1620

caaaggccac gtgtgaga 1638

<210> 4
<211> 546
<212> PRT
<213> Centaurea maculosa

<220>
<221> misc_feature
<222> (286)..(286)
<223> Xaa can be any naturally occurring amino acid

<400> 4

```
Met Pro Cys Leu Ser Gly Ile Gly Leu Leu Cys Lys Ile Met Ser His
1               5                  10                  15

Lys Thr Arg Lys Asn Ile Pro Val Ile Met Met Ser Ser His Asp Ser
            20                  25                  30

Met Gly Leu Val Phe Lys Cys Leu Ser Lys Gly Ala Val Asp Phe Leu
        35                  40                  45

Val Lys Pro Val Arg Lys Asn Glu Leu Lys Asn Leu Trp Gln His Val
    50                  55                  60

Trp Arg Arg Cys His Ser Ser Ser Gly Ser Gly Ser Glu Ser Gly Thr
65                  70                  75                  80

Gln Ala Gln Lys Ser Val Asn Ser Lys Ser Asn Leu Arg Tyr Asn Asn
                85                  90                  95
```

```
Gly Ser Asp Lys Asp Gly Asn Asp Asn Gly Ser Thr Ser Gly Gly Ser
            100                 105                 110

Asp Asp Gly Ser Gly Thr Gln Ser Ser Trp Thr Lys Gln Ala Val Glu
        115                 120                 125

Pro Glu Ser Pro Glu Ala Ala Ser Pro Cys Asp Gln Ile Thr Glu His
    130                 135                 140

Pro Asp Ser Thr Cys Gly Leu Val Ile Arg Ser Val His Ala Gln Ala
145                 150                 155                 160

Thr Arg Asp Ser Asn Asp Gln Glu Gly Arg Pro Tyr Asp Glu Gly Lys
                165                 170                 175

Ala Lys Glu Ile Ala Glu Cys Ser Phe Arg Asn Ser Glu Met Gln Ile
            180                 185                 190

Glu Phe Pro Ile Gln Ala Pro Val Lys His Asn Gly Ile Glu Gln Ser
        195                 200                 205

Thr His Gln Ala Phe Asp Pro Thr Leu Asn Phe Lys Ser Lys Glu Met
    210                 215                 220

Gly Ile Ser Asp Ile Arg Arg Glu His Ser Leu Asn Lys Gln Lys Ala
225                 230                 235                 240

Lys Asp Ser Lys Met Pro Glu Pro Tyr Met Glu Asn Glu Glu Leu Glu
                245                 250                 255

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

Asp Asp Ser Asn Gly Ala Ile Val Gly Glu Leu Gly Leu Xaa Arg Pro
        275                 280                 285

Arg Ala Asp Lys His Ser Gly Thr Glu Val Gln Thr Gly Cys Asn Ile
    290                 295                 300

Leu Arg His Ser Glu Leu Ser Ala Phe Thr Arg Tyr Lys Thr Thr Leu
305                 310                 315                 320

Asn Ala Val Lys Gly Thr Pro Gly Ile Thr Thr Ser Arg Ser Gln Pro
                325                 330                 335

Asp Tyr Arg Ser Asn Asp Val Lys Lys Glu Ser Lys Arg Asp Ala Leu
            340                 345                 350

Ser Asp Gly Tyr Leu Ile Tyr Gln Gly Ser Ser Glu Gln Val Ile Pro
        355                 360                 365
```

Ser Lys Ala Glu Gly Met Pro Pro Gly Gly Val Leu His Gln Glu His
370 375 380

Arg Ile Gln His Ile His His His His His Val His His Tyr His Asn
385 390 395 400

Leu Glu Glu Glu Gln Pro Pro Ser Asn His Asp Asp Phe Gly Leu Asn
405 410 415

Arg Leu Gly Ala Asp Ala Pro Tyr Cys Gly Ser Ser Asn Ile Val Gly
420 425 430

Gly Pro Gly Pro Val Glu Gly Asn Ile Glu Asn Tyr Ser Leu Asn Arg
435 440 445

Ser Ala Ser Gly Thr Gly Ser Lys His Gly Ser Asn Leu Pro Asn Gly
450 455 460

Ser Asn Ala Ala Val Asn Phe Glu Ala Thr Asn Val Glu Ser Asp Val
465 470 475 480

Gly Gly Leu Val Ile Ser Gly Ser Gly Asp Ala Ser Glu Ser Ala Ser
485 490 495

Gly Thr Gly Ile Ile Met Asp Arg Arg Asn Ser Ser Gln Arg Glu Ala
500 505 510

Ala Leu Asn Lys Phe Arg Gln Lys Arg Glu Val Arg Cys Phe Gln Lys
515 520 525

Asn Val Arg Tyr Gln Asn Arg Lys Lys Leu Ala Glu Gln Arg Pro Arg
530 535 540

Val Arg
545

<210> 5
<211> 507
<212> DNA
<213> Lactuca sativa

<400> 5
atgttggttg aagatgatga ttgtacacgt cacattgtaa ctgcattgct tcgcaactgt 60

aattatgaag ttattcaagc tgccaatgga cttcaagcct ggaagatctt ggaaaatcta 120

tccaatcaca ttgacattgt tttaactgaa gtagtcatgc cttgtctttc aggaatcggt 180

cttttatgca agattatgag ccacaagaca cgcaagaata tccctgtgat catgatgtct 240

tctcatgatt caatgggttt ggttttaag tgtttatcaa aaggtgcagt agacttttta 300

gtgaagccta ttcggaaaaa tgagcttaaa aacctttggc aacatgtatg gaggaggtgt 360

cacagttcaa gtggtagtgg aagtgaaagt ggtacacaag ctcaaaaatc tgtaaactca 420

aaaagcaatg taaggtacga taacagcagc aaagatgggg atgacaatga gaacaccagt 480

ggtggtagtg atgatggtag tgggcac 507

<210> 6
<211> 170
<212> PRT
<213> Lactuca sativa

<400> 6

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

Leu Arg Asn Cys Asn Tyr Glu Val Ile Gln Ala Ala Asn Gly Leu Gln
20 25 30

Ala Trp Lys Ile Leu Glu Asn Leu Ser Asn His Ile Asp Ile Val Leu
35 40 45

Thr Glu Val Val Met Pro Cys Leu Ser Gly Ile Gly Leu Leu Cys Lys
50 55 60

Ile Met Ser His Lys Thr Arg Lys Asn Ile Pro Val Ile Met Met Ser
65 70 75 80

Ser His Asp Ser Met Gly Leu Val Phe Lys Cys Leu Ser Lys Gly Ala
85 90 95

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

Trp Gln His Val Trp Arg Arg Cys His Ser Ser Ser Gly Ser Gly Ser
115 120 125

Glu Ser Gly Thr Gln Ala Gln Lys Ser Val Asn Ser Lys Ser Asn Val
130 135 140

Arg Tyr Asp Asn Ser Ser Lys Asp Gly Asp Asp Asn Glu Asn Thr Ser
145 150 155 160

Gly Gly Ser Asp Asp Gly Ser Gly His Ser
165 170

<210> 7
<211> 1446
<212> DNA
<213> Cichorium endivia

<400> 7
atgttggttg aaaatgatga ttgtacacgt cacattgtca ctgcattgct tcgcaactgt 60

aattatgaag ttattgaagc atctaatgga tttcaagcgt ggaagattct agaagatcta 120

tccaatcaca tagacattgt tttaaccgaa gtagttatgc cttctttttc tggtgttggt 180

ctt ctatgca agatt atgag ccacaagaca cgcaagaata t acccgt aat t atgat gt ca 240

t cgcat gatt caat gggt ct agt t t t t aag t gt t t gt caa aaggcgcagt agat t t t t t a 300

t t aaagccca t t cggaaaaa cgagct t aaa aat ct t t ggc agcat at t t g gaggagat gt 360

cacagt t ct a gt ggt agt gg gagt gaaagt ggt acaaat g ct caaaaat c t gt aaact ca 420

aaaagat ct g at gacaat gt t agcat t gct ggagat gat g at gggagt ac ccgt gat gga 480

agt gat gat g gt agt ggcac ccagagt t ct t ggacaaaac aagcagaagc at ct cct t gt 540

gat agcact t gt ggcct agt cat ccact ct gcagaaact c ct acaacaaa agat t t t caa 600

gt ccaggt gg at gaaggt ga t gt ggaaat g gacaaagagt t ggaaacagg aggat ct aaa 660

gact ccaaaa agat t ggaga at ct gaggt a gt aaaagat a ccaacaacag t ccacacaca 720

gacact aagg t aat gaat ga at ct aaagaa acact cgcgg aat t cagct t aaagaggcct 780

agggaagt t a cacaagt t ca gaat ggccca aacat t t t ga cacat t caga gct t t cagcc 840

t t t acaagat at aacacaac ct caaagcaa gat gt acccg gaaaat ccag t gat at ccca 900

ggagccgat t t ggt gggccc accacaagt t cat cacat t c at cat cat ca ccat gt t cat 960

cact accaca acat agaat c gaat cagcca ccggat aat c at gaagaat c gggt ct aaaa 1020

aact t ggct g caat t gct cc acat t gcggg t cat caaat g t ct t aggt gg t ggaggt gt t 1080

gagggt at t a t t ggaaat t g t agt t t gaat ggaagt gggt cgggt agt aa acat ggaagc 1140

aat ggacaga at gggagcag t gct gct gt g aat t t agaag gt aaaaat gt cgaaagt ggc 1200

ggt ggt ggt g gat t ggccgg aaaaagt ggt ggcggt ggt g gt ggt gaaaa t agaat agat 1260

gaagagaaat ct t cacagag agaagcagct t t gat gaagt t t cgt gaaaa gagaaaaaac 1320

cgat gct t t c aaaaaaaggt gcgat at caa aaccggaagc gact t gcaga agagaggcca 1380

cgt gt gagag gacaat t t gt aaagcaaact gggcaagaaa gt t caagt aa t ggt gaagac 1440

agat ag 1446

<210> 8
<211> 481
<212> PRT
<213> Cichorium endivia

<400> 8

Met Leu Val Glu Asn Asp Asp Cys Thr Arg His Ile Val Thr Ala Leu
1 5 10 15

Leu Arg Asn Cys Asn Tyr Glu Val Ile Glu Ala Ser Asn Gly Phe Gln
20 25 30

Ala Trp Lys Ile Leu Glu Asp Leu Ser Asn His Ile Asp Ile Val Leu
35 40 45

Thr Glu Val Val Met Pro Ser Phe Ser Gly Val Gly Leu Leu Cys Lys
50 55 60

```
Ile Met Ser His Lys Thr Arg Lys Asn Ile Pro Val Ile Met Met Ser
65                  70                  75                  80

Ser His Asp Ser Met Gly Leu Val Phe Lys Cys Leu Ser Lys Gly Ala
                85                  90                  95

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

Trp Gln His Ile Trp Arg Arg Cys His Ser Ser Ser Gly Ser Gly Ser
        115                 120                 125

Glu Ser Gly Thr Asn Ala Gln Lys Ser Val Asn Ser Lys Arg Ser Asp
    130                 135                 140

Asp Asn Val Ser Ile Ala Gly Asp Asp Asp Gly Ser Thr Arg Asp Gly
145                 150                 155                 160

Ser Asp Asp Gly Ser Gly Thr Gln Ser Ser Trp Thr Lys Gln Ala Glu
                165                 170                 175

Ala Ser Pro Cys Asp Ser Thr Cys Gly Leu Val Ile His Ser Ala Glu
            180                 185                 190

Thr Pro Thr Thr Lys Asp Phe Gln Val Gln Val Asp Glu Gly Asp Val
        195                 200                 205

Glu Met Asp Lys Glu Leu Glu Thr Gly Gly Ser Lys Asp Ser Lys Lys
    210                 215                 220

Ile Gly Glu Ser Glu Val Val Lys Asp Thr Asn Asn Ser Pro His Thr
225                 230                 235                 240

Asp Thr Lys Val Met Asn Glu Ser Lys Glu Thr Leu Ala Glu Phe Ser
                245                 250                 255

Leu Lys Arg Pro Arg Glu Val Thr Gln Val Gln Asn Gly Pro Asn Ile
            260                 265                 270

Leu Thr His Ser Glu Leu Ser Ala Phe Thr Arg Tyr Asn Thr Thr Ser
        275                 280                 285

Lys Gln Asp Val Pro Gly Lys Ser Ser Asp Ile Pro Gly Ala Asp Leu
    290                 295                 300

Val Gly Pro Pro Gln Val His His Ile His His His His His Val His
305                 310                 315                 320

His Tyr His Asn Ile Glu Ser Asn Gln Pro Pro Asp Asn His Glu Glu
                325                 330                 335
```

Ser Gly Leu Lys Asn Leu Ala Ala Ile Ala Pro His Cys Gly Ser Ser
340 345 350

Asn Val Leu Gly Gly Gly Gly Val Glu Gly Ile Ile Gly Asn Cys Ser
355 360 365

Leu Asn Gly Ser Gly Ser Gly Ser Lys His Gly Ser Asn Gly Gln Asn
370 375 380

Gly Ser Ser Ala Ala Val Asn Leu Glu Gly Lys Asn Val Glu Ser Gly
385 390 395 400

Gly Gly Gly Gly Leu Ala Gly Lys Ser Gly Gly Gly Gly Gly Gly Glu
405 410 415

Asn Arg Ile Asp Glu Glu Lys Ser Ser Gln Arg Glu Ala Ala Leu Met
420 425 430

Lys Phe Arg Glu Lys Arg Lys Asn Arg Cys Phe Gln Lys Lys Val Arg
435 440 445

Tyr Gln Asn Arg Lys Arg Leu Ala Glu Glu Arg Pro Arg Val Arg Gly
450 455 460

Gln Phe Val Lys Gln Thr Gly Gln Glu Ser Ser Ser Asn Gly Glu Asp
465 470 475 480

Arg

<210> 9
<211> 870
<212> DNA
<213> Saccharum officinarum

<400> 9
atgagtatgg tgtttaagtg tttgtcgaag ggagcagttg acttcttggt aaagccacta 60

cgtaagaatg agcttaagaa cctttggcag cacgtttgga ggcgatgcca cagttccagt 120

ggcagtggaa gtgaaagtgg catccagaca cagaagtgtg ccaaaccaaa tactggtgac 180

gagtatgaga acgacagtga cagcaatcat gatgatgaag aaaatgatga cgacgacgat 240

gacgacttca gtgtcggact caatgctagg gatggaagtg ataatggcag tggtactcaa 300

agctcatgga caaagcgtgc tgtggagatt gacagtccac aacctatgtc tcctgatcaa 360

ctagcagatc cacctgatag tacatgtgca caagtaattc accccaaatc agagatatgc 420

agtaacaagt ggctaccgac agcaaacaaa aggaatggca agaaacataa ggagaataaa 480

gatgaatcta tgggaagata cctagaaata ggtgctccta ggaactcaag tgcagaatat 540

caatcatctc tcaatgacgt atctgttaat ccaacagaaa aacgacatga gactcacatg 600

ccccaatgca aatccaaaaa gaaaatgatg gcagaagatg attgtacaga catacctagt 660

gaaacaaata ctgaaactgc tgatttaatt agctcaatag ccagaaacac agaaggccaa 720

caagcagtac gagccgttga tgcacctgat ggcccttcca agatgcccga tggaaatgat 780

aagaatcatg attctcatat cgaggtgaca ccccatgagt tgggttttgaa gagattgaga 840

acagatggag ctacagccga aaatccatga 870

<210> 10
<211> 289
<212> PRT
<213> Saccharum officinarum

<400> 10

Met Ser Met Val Phe Lys Cys Leu Ser Lys Gly Ala Val Asp Phe Leu
1 5 10 15

Val Lys Pro Leu Arg Lys Asn Glu Leu Lys Asn Leu Trp Gln His Val
20 25 30

Trp Arg Arg Cys His Ser Ser Ser Gly Ser Gly Ser Glu Ser Gly Ile
35 40 45

Gln Thr Gln Lys Cys Ala Lys Pro Asn Thr Gly Asp Glu Tyr Glu Asn
50 55 60

Asp Ser Asp Ser Asn His Asp Asp Glu Glu Asn Asp Asp Asp Asp Asp
65 70 75 80

Asp Asp Phe Ser Val Gly Leu Asn Ala Arg Asp Gly Ser Asp Asn Gly
85 90 95

Ser Gly Thr Gln Ser Ser Trp Thr Lys Arg Ala Val Glu Ile Asp Ser
100 105 110

Pro Gln Pro Met Ser Pro Asp Gln Leu Ala Asp Pro Pro Asp Ser Thr
115 120 125

Cys Ala Gln Val Ile His Pro Lys Ser Glu Ile Cys Ser Asn Lys Trp
130 135 140

Leu Pro Thr Ala Asn Lys Arg Asn Gly Lys Lys His Lys Glu Asn Lys
145 150 155 160

Asp Glu Ser Met Gly Arg Tyr Leu Glu Ile Gly Ala Pro Arg Asn Ser
165 170 175

Ser Ala Glu Tyr Gln Ser Ser Leu Asn Asp Val Ser Val Asn Pro Thr
180 185 190

Glu Lys Arg His Glu Thr His Met Pro Gln Cys Lys Ser Lys Lys Lys
195 200 205

```
Met Met Ala Glu Asp Asp Cys Thr Asp Ile Pro Ser Glu Thr Asn Thr
    210                 215                 220

Glu Thr Ala Asp Leu Ile Ser Ser Ile Ala Arg Asn Thr Glu Gly Gln
225                 230                 235                 240

Gln Ala Val Arg Ala Val Asp Ala Pro Asp Gly Pro Ser Lys Met Pro
                245                 250                 255

Asp Gly Asn Asp Lys Asn His Asp Ser His Ile Glu Val Thr Pro His
            260                 265                 270

Glu Leu Gly Leu Lys Arg Leu Arg Thr Asp Gly Ala Thr Ala Glu Asn
        275                 280                 285

Pro
```

<210> 11
<211> 2298
<212> DNA
<213> Sorghum bicolor

<400> 11

```
atgggtagcg ctt gccaagc t ggcatggac gggccttccc gcaaggatgt gtt ggggata      60

gggaatgtcg ccttagagaa t ggccaccat gaggttggag ctgatgcaga t gaatggagg      120

gaaaaggaag aggacttggc caatgggcac agtgcgccac cgggcatgca gcaggtggat      180

gagcaggagc aacaaggaca aagcattcac t gggagaggt t cctacctgt gaagacactg      240

agagtcatgc t ggtggagaa t gatgactct actcgtcagg t ggtcagtgc cctgctccgt      300

aagtgctgct atgaagttat ccctgctgaa aatggttcac atgcatggcg atatcttgaa      360

gatctgcaga acaacattga ccttgtattg actgaggttt tcatgccttg tctatctggc      420

atcggtctgc ttagcaaaat cacaagtcac aaaatttgca aggacattcc tgtgatttatg      480

atgtcttcaa atgactctat gagtatggtg tttaagtgtt tgtcgaaggg agcagttgac      540

ttcttggtaa agccactacg taagaatgag cttaagaacc tttggcagca cgttttggagg      600

cgatgccaca gttccagtgg cagtggaagt gaaagcggca tccagacaca gaagtgtgcc      660

aaaccaaata ctggtgatga gtatgagaac gacagtgaca gcaatcatga tgatgaagaa      720

aatgatgaag acgacgacga tgacttcagt gtcggactca atgctaggga tggaagtgat      780

aatggcagtg gtactcaaag ctcatggaca aaacgtgctg tggagattga cagtccagaa      840

cctatgtctc ctgatcaact agcagatcca cctgatagta catgtgcaca agtaattcac      900

cccaaatcag agatatgcag taacaagtgg ctaccgacag caaacaaaag gaatggcaag      960

aaacataagg agaataaaga tgaatctatg ggaagatact tagaaatagg tgctcctagg     1020

aactcaagtg cagaatatca atcatctctc aatgacgtat ctgttaatcc aacagaaaaa     1080

cgtcatgaga ctcacatgcc ccaatgcaaa tccaaaaaga aaatgatggc agaagatgat     1140
```

tgtacagaca tacctagtga aataaaatact gaaactgctg attttgattag ctcaatagcc 1200

agaaacacag aaggccaaca agcagtacga gctgttgatg cacctgatgg ccctttccaag 1260

atgcccgatg gaaatgataa gaatcatgat tctcatatcg tggtgacacc ccatgagttg 1320

ggttttgaaga gattgagaac agatggagct gcagatgaaa tccatgatga gcgaaatatt 1380

ctcaaaagat cagatcagtc agccttcacc aggtaccata catctgtggc ttccaatcaa 1440

ggtggagcaa gatgtgggga aagctcttca ccacaagata acagttctga ggctgtgaaa 1500

acagactcta catgcaagat gaagtcaaat tcagatgctg ctccaataaa gcagggctcc 1560

aatggcagta gcaacaacga tgtgggctcc agtacaaaga atgttattgc aaagccttca 1620

gctaacaggg agagagtaac gtcaccatca gccatcaaat ctacccagca tgcctcagca 1680

tttcatacta tacagaatca aacatcacct gctaatctgg ttggtaaaga caaagctgat 1740

gaaggaatt t ccaatgcagt gaaaatgagc cacccaacag aggttccaca aagctgcgtc 1800

cagcatcatc accacgtgca ttattacctc catgttatga cacagaaaca gtcatcaatc 1860

gaccgtggat catcagatgt tcagtgtggt tcgtcaaatg tgttgatcc tcctgttgaa 1920

ggacatgctg caaactatag tgtgaatggg ggtgtctcag ttggtcataa tgggtgcaat 1980

ggccagaatg gaacgagcac tgtccccaat attgcaagac caaacataga gagtgttaat 2040

ggtaccgtga gccaaaatat cgctggaggt ggcattgtaa gtgggagtgg gagtggcaat 2100

gatgtgtatc agaatcgatt ccccccaacga gaagctgcat tgaacaaatt cagactgaag 2160

cggaaagatc ggaacttttgg taaaaaggtt cgctaccaaa gcaggaagag gcttgctgag 2220

cagcggcctc gggtccgtgg acagttttgtg cgacaatctg ggcaagaaga tcaagcagca 2280

caaggttcag aaagatga 2298

<210> 12
<211> 765
<212> PRT
<213> Sorghum bicolor

<400> 12

Met Gly Ser Ala Cys Gln Ala Gly Met Asp Gly Pro Ser Arg Lys Asp
1 5 10 15

Val Leu Gly Ile Gly Asn Val Ala Leu Glu Asn Gly His His Glu Val
20 25 30

Gly Ala Asp Ala Asp Glu Trp Arg Glu Lys Glu Glu Asp Leu Ala Asn
35 40 45

Gly His Ser Ala Pro Pro Gly Met Gln Gln Val Asp Glu Gln Glu Gln
50 55 60

Gln Gly Gln Ser Ile His Trp Glu Arg Phe Leu Pro Val Lys Thr Leu
65 70 75 80

```
Arg Val Met Leu Val Glu Asn Asp Asp Ser Thr Arg Gln Val Val Ser
                85                  90                  95

Ala Leu Leu Arg Lys Cys Cys Tyr Glu Val Ile Pro Ala Glu Asn Gly
            100                 105                 110

Ser His Ala Trp Arg Tyr Leu Glu Asp Leu Gln Asn Asn Ile Asp Leu
        115                 120                 125

Val Leu Thr Glu Val Phe Met Pro Cys Leu Ser Gly Ile Gly Leu Leu
    130                 135                 140

Ser Lys Ile Thr Ser His Lys Ile Cys Lys Asp Ile Pro Val Ile Met
145                 150                 155                 160

Met Ser Ser Asn Asp Ser Met Ser Met Val Phe Lys Cys Leu Ser Lys
                165                 170                 175

Gly Ala Val Asp Phe Leu Val Lys Pro Leu Arg Lys Asn Glu Leu Lys
            180                 185                 190

Asn Leu Trp Gln His Val Trp Arg Arg Cys His Ser Ser Ser Gly Ser
        195                 200                 205

Gly Ser Glu Ser Gly Ile Gln Thr Gln Lys Cys Ala Lys Pro Asn Thr
    210                 215                 220

Gly Asp Glu Tyr Glu Asn Asp Ser Asp Ser Asn His Asp Asp Glu Glu
225                 230                 235                 240

Asn Asp Glu Asp Asp Asp Asp Asp Phe Ser Val Gly Leu Asn Ala Arg
                245                 250                 255

Asp Gly Ser Asp Asn Gly Ser Gly Thr Gln Ser Ser Trp Thr Lys Arg
            260                 265                 270

Ala Val Glu Ile Asp Ser Pro Glu Pro Met Ser Pro Asp Gln Leu Ala
        275                 280                 285

Asp Pro Pro Asp Ser Thr Cys Ala Gln Val Ile His Pro Lys Ser Glu
    290                 295                 300

Ile Cys Ser Asn Lys Trp Leu Pro Thr Ala Asn Lys Arg Asn Gly Lys
305                 310                 315                 320

Lys His Lys Glu Asn Lys Asp Glu Ser Met Gly Arg Tyr Leu Glu Ile
                325                 330                 335

Gly Ala Pro Arg Asn Ser Ser Ala Glu Tyr Gln Ser Ser Leu Asn Asp
            340                 345                 350
```

```
Val Ser Val Asn Pro Thr Glu Lys Arg His Glu Thr His Met Pro Gln
        355                 360                 365

Cys Lys Ser Lys Lys Lys Met Met Ala Glu Asp Asp Cys Thr Asp Ile
    370                 375                 380

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

Arg Asn Thr Glu Gly Gln Gln Ala Val Arg Ala Val Asp Ala Pro Asp
                405                 410                 415

Gly Pro Ser Lys Met Pro Asp Gly Asn Asp Lys Asn His Asp Ser His
            420                 425                 430

Ile Val Val Thr Pro His Glu Leu Gly Leu Lys Arg Leu Arg Thr Asp
        435                 440                 445

Gly Ala Ala Asp Glu Ile His Asp Glu Arg Asn Ile Leu Lys Arg Ser
    450                 455                 460

Asp Gln Ser Ala Phe Thr Arg Tyr His Thr Ser Val Ala Ser Asn Gln
465                 470                 475                 480

Gly Gly Ala Arg Cys Gly Glu Ser Ser Ser Pro Gln Asp Asn Ser Ser
                485                 490                 495

Glu Ala Val Lys Thr Asp Ser Thr Cys Lys Met Lys Ser Asn Ser Asp
            500                 505                 510

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

Gly Ser Ser Thr Lys Asn Val Ile Ala Lys Pro Ser Ala Asn Arg Glu
    530                 535                 540

Arg Val Thr Ser Pro Ser Ala Ile Lys Ser Thr Gln His Ala Ser Ala
545                 550                 555                 560

Phe His Thr Ile Gln Asn Gln Thr Ser Pro Ala Asn Leu Val Gly Lys
                565                 570                 575

Asp Lys Ala Asp Glu Gly Ile Ser Asn Ala Val Lys Met Ser His Pro
            580                 585                 590

Thr Glu Val Pro Gln Ser Cys Val Gln His His His His Val His Tyr
        595                 600                 605

Tyr Leu His Val Met Thr Gln Lys Gln Ser Ser Ile Asp Arg Gly Ser
    610                 615                 620
```

Ser Asp Val Gln Cys Gly Ser Ser Asn Val Phe Asp Pro Pro Val Glu
625 630 635 640

Gly His Ala Ala Asn Tyr Ser Val Asn Gly Gly Val Ser Val Gly His
645 650 655

Asn Gly Cys Asn Gly Gln Asn Gly Thr Ser Thr Val Pro Asn Ile Ala
660 665 670

Arg Pro Asn Ile Glu Ser Val Asn Gly Thr Val Ser Gln Asn Ile Ala
675 680 685

Gly Gly Gly Ile Val Ser Gly Ser Gly Ser Gly Asn Asp Val Tyr Gln
690 695 700

Asn Arg Phe Pro Gln Arg Glu Ala Ala Leu Asn Lys Phe Arg Leu Lys
705 710 715 720

Arg Lys Asp Arg Asn Phe Gly Lys Lys Val Arg Tyr Gln Ser Arg Lys
725 730 735

Arg Leu Ala Glu Gln Arg Pro Arg Val Arg Gly Gln Phe Val Arg Gln
740 745 750

Ser Gly Gln Glu Asp Gln Ala Ala Gln Gly Ser Glu Arg
755 760 765

<210> 13
<211> 1764
<212> DNA
<213> Zea mays

<400> 13
atgtctacga atgattctat gagtatggtg tttaagtgtt tgtcgaaggg agcagttgat 60

ttcttggtaa aaccactacg taagaatgag cttaagaacc tttggcagca tgtttggagg 120

cgatgccaca gttccagtgg aagtgaaagt ggcatccaga cacagaagtg tgccaaacta 180

aatactggcg acgagtatga gaacggcagt gacagcaatc atgatgatga agaaaatgat 240

gacggcgacg atgacgactt cagtgttgga ctcaatgcta gggatggaag tgacaatggc 300

agtggtactc aaagctcatg gacaaagcgt gctgtggaga ttgacagccc acaacctata 360

tctcccgatc aactagttga tccacctgat agtacatgtg cacaagtaat tcaccctaga 420

tcagagatat gcagtaacaa gtggttaccg acagcaaaca aaaggaatgt caagaaacag 480

aaggagaata aagatgaatc tatgggaaga tacttaggaa taggtgctcc taggaactca 540

agtgcagaat atcaatcatc tctcaatgat gtatctgtta atccaataga aaaaggacat 600

gagaatcaca tgtccaaatg caaatctaaa aaggaaacaa tggcagaaga tgattgtaca 660

aacatgccta gtgcaacaaa tgctgaaact gctgatttga ttagctcaat agccagaaac 720

acagaaggcc aacaagcagt acaagccgtt gacgcaccag atggccttc caaaatggct 780

aatggaaatg ataagaatca tgattctcat atcgaagtga caccccatga gttgggttg 840

aagagatcga gaacaaatgg agctacagcg gaaatccatg atgagcgaaa tattctgaaa 900

agatcagatc agtcagcctt caccaggtac catacatctg tggcttccaa tcaaggtgga 960

gcaagatatg gggaaagctc ttcaccacaa gataacagtt ctgaggccat gaaaacggac 1020

tctacatgca agatgaagtc aaattcagat gctgctccaa taaagcaggg ctccaatggc 1080

agtagcaata acgatgtggg atccagtaca aagaatgttg ctgcaaggcc ttcgggtgac 1140

agggagagag tagcgtcacc attagccatc aaatctaccc agcatgcctc agcattttcat 1200

actatacaga atcaaacgtc accagctaat ctgattgggg aagacaaagc tgatgaagga 1260

atttccaata cagtgaaaat gagccaccca acagaggttc cacaaggctg cgtccagcat 1320

catcatcatg tgcattatta cctccatgtt atgacacaga aacagccatc aacagaccgt 1380

ggatcatcag atgttcactg tggttcgtca aatgtgttg atcctcctgt tgaaggacat 1440

gctgcaaact acagtgtgaa tgggggtgtc tcagttggtc ataatgggtg caatgggcag 1500

aatggaagta gcgctgtccc caatattgca agaccaaaca tagagagtat taatggtacc 1560

atgagccaaa atattgccgg aggtggcatt gtaagtggga gtgggagtgg caatgacatg 1620

tatcagaatc ggttcctgca acgagaagct gcattgaaca aattcagact gaagcggaaa 1680

gatcggaact ttggtaaaaa ggtagcctgt tttcagttac atgcctgttg tagctatgac 1740

ttgaaatcag taatagttat ttga 1764

<210> 14
<211> 587
<212> PRT
<213> Zea mays

<400> 14

Met Ser Thr Asn Asp Ser Met Ser Met Val Phe Lys Cys Leu Ser Lys
1 5 10 15

Gly Ala Val Asp Phe Leu Val Lys Pro Leu Arg Lys Asn Glu Leu Lys
20 25 30

Asn Leu Trp Gln His Val Trp Arg Arg Cys His Ser Ser Ser Gly Ser
35 40 45

Glu Ser Gly Ile Gln Thr Gln Lys Cys Ala Lys Leu Asn Thr Gly Asp
50 55 60

Glu Tyr Glu Asn Gly Ser Asp Ser Asn His Asp Asp Glu Glu Asn Asp
65 70 75 80

Asp Gly Asp Asp Asp Asp Phe Ser Val Gly Leu Asn Ala Arg Asp Gly
85 90 95

Ser Asp Asn Gly Ser Gly Thr Gln Ser Ser Trp Thr Lys Arg Ala Val

```
                100                 105                 110

Glu Ile Asp Ser Pro Gln Pro Ile Ser Pro Asp Gln Leu Val Asp Pro
        115                 120                 125

Pro Asp Ser Thr Cys Ala Gln Val Ile His Pro Arg Ser Glu Ile Cys
    130                 135                 140

Ser Asn Lys Trp Leu Pro Thr Ala Asn Lys Arg Asn Val Lys Lys Gln
145                 150                 155                 160

Lys Glu Asn Lys Asp Glu Ser Met Gly Arg Tyr Leu Gly Ile Gly Ala
                165                 170                 175

Pro Arg Asn Ser Ser Ala Glu Tyr Gln Ser Ser Leu Asn Asp Val Ser
            180                 185                 190

Val Asn Pro Ile Glu Lys Gly His Glu Asn His Met Ser Lys Cys Lys
        195                 200                 205

Ser Lys Lys Glu Thr Met Ala Glu Asp Asp Cys Thr Asn Met Pro Ser
    210                 215                 220

Ala Thr Asn Ala Glu Thr Ala Asp Leu Ile Ser Ser Ile Ala Arg Asn
225                 230                 235                 240

Thr Glu Gly Gln Gln Ala Val Gln Ala Val Asp Ala Pro Asp Gly Pro
                245                 250                 255

Ser Lys Met Ala Asn Gly Asn Asp Lys Asn His Asp Ser His Ile Glu
            260                 265                 270

Val Thr Pro His Glu Leu Gly Leu Lys Arg Ser Arg Thr Asn Gly Ala
        275                 280                 285

Thr Ala Glu Ile His Asp Glu Arg Asn Ile Leu Lys Arg Ser Asp Gln
    290                 295                 300

Ser Ala Phe Thr Arg Tyr His Thr Ser Val Ala Ser Asn Gln Gly Gly
305                 310                 315                 320

Ala Arg Tyr Gly Glu Ser Ser Ser Pro Gln Asp Asn Ser Ser Glu Ala
                325                 330                 335

Met Lys Thr Asp Ser Thr Cys Lys Met Lys Ser Asn Ser Asp Ala Ala
            340                 345                 350

Pro Ile Lys Gln Gly Ser Asn Gly Ser Ser Asn Asn Asp Val Gly Ser
        355                 360                 365

Ser Thr Lys Asn Val Ala Ala Arg Pro Ser Gly Asp Arg Glu Arg Val
```

370 375 380

Ala Ser Pro Leu Ala Ile Lys Ser Thr Gln His Ala Ser Ala Phe His
385 390 395 400

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

Ala Asp Glu Gly Ile Ser Asn Thr Val Lys Met Ser His Pro Thr Glu
420 425 430

Val Pro Gln Gly Cys Val Gln His His His His Val His Tyr Tyr Leu
435 440 445

His Val Met Thr Gln Lys Gln Pro Ser Thr Asp Arg Gly Ser Ser Asp
450 455 460

Val His Cys Gly Ser Ser Asn Val Phe Asp Pro Pro Val Glu Gly His
465 470 475 480

Ala Ala Asn Tyr Ser Val Asn Gly Gly Val Ser Val Gly His Asn Gly
485 490 495

Cys Asn Gly Gln Asn Gly Ser Ser Ala Val Pro Asn Ile Ala Arg Pro
500 505 510

Asn Ile Glu Ser Ile Asn Gly Thr Met Ser Gln Asn Ile Ala Gly Gly
515 520 525

Gly Ile Val Ser Gly Ser Gly Ser Gly Asn Asp Met Tyr Gln Asn Arg
530 535 540

Phe Leu Gln Arg Glu Ala Ala Leu Asn Lys Phe Arg Leu Lys Arg Lys
545 550 555 560

Asp Arg Asn Phe Gly Lys Lys Val Ala Cys Phe Gln Leu His Ala Cys
565 570 575

Cys Ser Tyr Asp Leu Lys Ser Val Ile Val Ile
580 585

<210> 15
<211> 1725
<212> DNA
<213> Zea mays

<400> 15
atgtctacga atgattctat gagtatggtg tttaagtgtt tgtcgaaggg agcagttgat 60

ttcttggtaa aaccactacg taagaatgag cttaagaacc tttggcagca tgtttggagg 120

cgatgccaca gttccagtgg aagtgaaagt ggcatccaga cacagaagtg tgccaaacta 180

aatactggcg acgagtatga gaacggcagt gacagcaatc atgatgatga agaaaatgat 240

gacggcgacg atgacgactt cagtgttgga ctcaatgcta gggatggaag tgacaatggc 300

agtggtactc aaagctcatg gacaaagcgt gctgtggaga ttgacagccc acaacctata 360

tctcccgatc aactagttga tccacctgat agtacatgtg cacaagtaat tcaccctaga 420

tcagagatat gcagtaacaa gtggttaccg acagcaaaca aaaggaatgt caagaaacag 480

aaggagaata aagatgaatc tatgggaaga tacttaggaa taggtgctcc taggaactca 540

agtgcagaat atcaatcatc tctcaatgat gtatctgtta atccaataga aaaaggacat 600

gagaatcaca tgtccaaatg caaatctaaa aaggaaacaa tggcagaaga tgattgtaca 660

aacatgccta gtgcaacaaa tgctgaaact gctgattiga ttagctcaat agccagaaac 720

acagaaggcc aacaagcagt acaagccgtt gacgcaccag atggccctic caaaatggct 780

aatggaaatg ataagaatca tgattctcat atcgaagtga caccccatga gttgggttig 840

aagagatcga gaacaaatgg agctacagcg gaaatccatg atgagcgaaa tattctgaaa 900

agatcagatc agtcagcctt caccaggtac catacatctg tggcttccaa tcaaggtgga 960

gcaagatatg gggaaagctc ttcaccacaa gataacagtt ctgaggccat gaaaacggac 1020

tctacatgca agatgaagtc aaattcagat gctgctccaa taaagcaggg ctccaatggc 1080

agtagcaata acgatgtggg atccagtaca aagaatgttg ctgcaaggcc ttcgggtgac 1140

agggagagag tagcgtcacc attagccatc aaatctaccc agcatgcctc agcattticat 1200

actatacaga atcaaacgtc accagcttaat ctgattgggg aagacaaagc tgatgaagga 1260

attticcaata cagtgaaaat gagccaccca acagaggttc cacaaggctg cgtccagcat 1320

catcatcatg tgcattatta cctccatgtt atgacacaga aacagccatc aacagaccgt 1380

ggatcatcag atgttcactg tggttcgtca aatgtgttig atcctcctgt tgaaggacat 1440

gctgcaaact acagtgtgaa tgggggtgtc tcagttggtc ataatgggtg caatgggcag 1500

aatggaagta gcgctgtccc caatattgca agaccaaaca tagagagtat taatggtacc 1560

atgagccaaa atattgccgg aggtggcatt gtaagtggga gtgggagtgg caatgacatg 1620

tatcagaatc ggttcctgca acgagaagct gcattgaaca aattcagact gaagcggaaa 1680

gatcggaact ttggttcgct accaaagcag gaagaggctt gctga 1725

<210> 16
<211> 574
<212> PRT
<213> Zea mays

<400> 16

Met Ser Thr Asn Asp Ser Met Ser Met Val Phe Lys Cys Leu Ser Lys
1 5 10 15

Gly Ala Val Asp Phe Leu Val Lys Pro Leu Arg Lys Asn Glu Leu Lys
20 25 30

Asn Leu Trp Gln His Val Trp Arg Arg Cys His Ser Ser Ser Gly Ser

35 40 45

Glu Ser Gly Ile Gln Thr Gln Lys Cys Ala Lys Leu Asn Thr Gly Asp
50 55 60

Glu Tyr Glu Asn Gly Ser Asp Ser Asn His Asp Asp Glu Glu Asn Asp
65 70 75 80

Asp Gly Asp Asp Asp Asp Phe Ser Val Gly Leu Asn Ala Arg Asp Gly
85 90 95

Ser Asp Asn Gly Ser Gly Thr Gln Ser Ser Trp Thr Lys Arg Ala Val
100 105 110

Glu Ile Asp Ser Pro Gln Pro Ile Ser Pro Asp Gln Leu Val Asp Pro
115 120 125

Pro Asp Ser Thr Cys Ala Gln Val Ile His Pro Arg Ser Glu Ile Cys
130 135 140

Ser Asn Lys Trp Leu Pro Thr Ala Asn Lys Arg Asn Val Lys Lys Gln
145 150 155 160

Lys Glu Asn Lys Asp Glu Ser Met Gly Arg Tyr Leu Gly Ile Gly Ala
165 170 175

Pro Arg Asn Ser Ser Ala Glu Tyr Gln Ser Ser Leu Asn Asp Val Ser
180 185 190

Val Asn Pro Ile Glu Lys Gly His Glu Asn His Met Ser Lys Cys Lys
195 200 205

Ser Lys Lys Glu Thr Met Ala Glu Asp Asp Cys Thr Asn Met Pro Ser
210 215 220

Ala Thr Asn Ala Glu Thr Ala Asp Leu Ile Ser Ser Ile Ala Arg Asn
225 230 235 240

Thr Glu Gly Gln Gln Ala Val Gln Ala Val Asp Ala Pro Asp Gly Pro
245 250 255

Ser Lys Met Ala Asn Gly Asn Asp Lys Asn His Asp Ser His Ile Glu
260 265 270

Val Thr Pro His Glu Leu Gly Leu Lys Arg Ser Arg Thr Asn Gly Ala
275 280 285

Thr Ala Glu Ile His Asp Glu Arg Asn Ile Leu Lys Arg Ser Asp Gln
290 295 300

Ser Ala Phe Thr Arg Tyr His Thr Ser Val Ala Ser Asn Gln Gly Gly

305 310 315 320

Ala Arg Tyr Gly Glu Ser Ser Ser Pro Gln Asp Asn Ser Ser Glu Ala
325 330 335

Met Lys Thr Asp Ser Thr Cys Lys Met Lys Ser Asn Ser Asp Ala Ala
340 345 350

Pro Ile Lys Gln Gly Ser Asn Gly Ser Ser Asn Asn Asp Val Gly Ser
355 360 365

Ser Thr Lys Asn Val Ala Ala Arg Pro Ser Gly Asp Arg Glu Arg Val
370 375 380

Ala Ser Pro Leu Ala Ile Lys Ser Thr Gln His Ala Ser Ala Phe His
385 390 395 400

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

Ala Asp Glu Gly Ile Ser Asn Thr Val Lys Met Ser His Pro Thr Glu
420 425 430

Val Pro Gln Gly Cys Val Gln His His His His Val His Tyr Tyr Leu
435 440 445

His Val Met Thr Gln Lys Gln Pro Ser Thr Asp Arg Gly Ser Ser Asp
450 455 460

Val His Cys Gly Ser Ser Asn Val Phe Asp Pro Pro Val Glu Gly His
465 470 475 480

Ala Ala Asn Tyr Ser Val Asn Gly Gly Val Ser Val Gly His Asn Gly
485 490 495

Cys Asn Gly Gln Asn Gly Ser Ser Ala Val Pro Asn Ile Ala Arg Pro
500 505 510

Asn Ile Glu Ser Ile Asn Gly Thr Met Ser Gln Asn Ile Ala Gly Gly
515 520 525

Gly Ile Val Ser Gly Ser Gly Ser Gly Asn Asp Met Tyr Gln Asn Arg
530 535 540

Phe Leu Gln Arg Glu Ala Ala Leu Asn Lys Phe Arg Leu Lys Arg Lys
545 550 555 560

Asp Arg Asn Phe Gly Ser Leu Pro Lys Gln Glu Glu Ala Cys
565 570

<210> 17

<211> 2232
<212> DNA
<213> Zea mays

<400> 17
atgggcagtg cttgccaagc tggcacagac gggccttccc gcaaggatgt gttagggata 60
gggaatgccg ccttagagaa tggccaccat caggctgaag ctgacgcaga tgaatggagg 120
gaaaaggaag aggacttggc caacaacggg cacagtgcgc caccgccagg catgcagcag 180
gtggatgagc ataaggagga acaaagacaa agcattcact gggagaggtt cctacctgtg 240
aagacactga gagtcttgct ggtggagaat gatgactcta ctcgtcaggt ggtcagtgcc 300
ctgctccgta agtgctgcta tgaagttatt cctgctgaaa atggttttgca tgcatggcga 360
tatcttgaag atctgcagaa caacatcgac cttgtattga ctgaggtttt catgccttgt 420
ctatctggta tcggtctgct tagcaaaatc acaagtcaca aaatttgcaa agacattcct 480
gtgattatga tgtctacgaa tgattctatg agtatggtgt ttaagtgttt gtcgaaggga 540
gcagttgatt tcttggtaaa accactacgt aagaatgagc ttaagaacct ttggcagcat 600
gttttggaggc gatgccacag tgtaagctgt ttgttttttgt ccagtggaag tgaaagtggc 660
atccagacac agaagtgtgc caaactaaat actggcgacg agtatgagaa cggcagtgac 720
agcaatcatg atgatgaaga aaatgatgac ggcgacgatg acgacttcag tgttggactc 780
aatgctaggg atggaagtga caatggcagt ggtactcaaa gctcatggac aaagcgtgct 840
gtggagattg acagcccaca acctatatct cccgatcaac tagttgatcc acctgatagt 900
acatgtgcac aagtaattca ccctagatca gagatatgca gtaacaagtg gttaccgaca 960
gcaaacaaaa ggaatgtcaa gaaacagaag gagaataaag atgaatctat gggaagatac 1020
ttaggaatag gtgctcctag gaacttcaagt gcagaatatc aatcatctct caatgatgta 1080
tctgttaatc caatagaaaa aggacatgag aatcacatgt ccaaatgcaa atctaaaaag 1140
gaaacaatgg cagaagatga ttgtacaaac atgcctagtg caacaaatgc tgaaactgct 1200
gattttgatta gctcaatagc cagaaacaca gaaggccaac aagcagtaca agccgttgac 1260
gcaccagatg gccctttccaa aatggctaat ggaaatgata agaatcatga ttctcatatc 1320
gaagtgacac cccatgagtt gggttttgaag agatcgagaa caaatggagc tacagcggaa 1380
atccatgatg agcgaaatat tctgaaaaga tcagatcagt cagcctttcac caggtaccat 1440
acatctgtgg ctttccaatca aggtggagca agatatgggg aaagctcttc accacaagat 1500
aacagttctg aggccatgaa aacggactct acatgcaaga tgaagtcaaa ttcagatgct 1560
gctccaataa agcagggctc caatggcagt agcaataacg atgtgggatc cagtacaaag 1620
aatgttgctg caaggccttc gggtgacagg gagagagtag cgtcaccatt agccatcaaa 1680
tctacccagc atgcctcagc atttcatact atacagaatc aaacgtcacc agctaatctg 1740
attggggaag acaaagctga tgaaggaatt tccaatacag tgaaaatgag ccacccaaca 1800
gaggttccac aaggctgcgt ccagcatcat catcatgtgc atttatttacct ccatgttatg 1860
acacagaaac agccatcaac agaccgtgga tcatcagatg ttcactgtgg ttcgtcaaat 1920

gtgtttgatc ctcctgttga aggacatgct gcaaactaca gtgtgaatgg gggtgtctca 1980

gttggtcata atgggtgcaa tgggcagaat ggaagtagcg ctgtccccaa tattgcaaga 2040

ccaaacatag agagtattaa tggtaccatg agccaaaata ttgccggagg tggcattgta 2100

agtgggagtg ggagtggcaa tgacatgtat cagaatcggt tcctgcaacg agaagctgca 2160

ttgaacaaat tcagactgaa gcggaaagat cggaactttg gttcgctacc aaagcaggaa 2220

gaggcttgct ga 2232

<210> 18
<211> 743
<212> PRT
<213> Zea mays

<400> 18

Met Gly Ser Ala Cys Gln Ala Gly Thr Asp Gly Pro Ser Arg Lys Asp
1 5 10 15

Val Leu Gly Ile Gly Asn Ala Ala Leu Glu Asn Gly His His Gln Ala
20 25 30

Glu Ala Asp Ala Asp Glu Trp Arg Glu Lys Glu Glu Asp Leu Ala Asn
35 40 45

Asn Gly His Ser Ala Pro Pro Pro Gly Met Gln Gln Val Asp Glu His
50 55 60

Lys Glu Glu Gln Arg Gln Ser Ile His Trp Glu Arg Phe Leu Pro Val
65 70 75 80

Lys Thr Leu Arg Val Leu Leu Val Glu Asn Asp Asp Ser Thr Arg Gln
85 90 95

Val Val Ser Ala Leu Leu Arg Lys Cys Cys Tyr Glu Val Ile Pro Ala
100 105 110

Glu Asn Gly Leu His Ala Trp Arg Tyr Leu Glu Asp Leu Gln Asn Asn
115 120 125

Ile Asp Leu Val Leu Thr Glu Val Phe Met Pro Cys Leu Ser Gly Ile
130 135 140

Gly Leu Leu Ser Lys Ile Thr Ser His Lys Ile Cys Lys Asp Ile Pro
145 150 155 160

Val Ile Met Met Ser Thr Asn Asp Ser Met Ser Met Val Phe Lys Cys
165 170 175

Leu Ser Lys Gly Ala Val Asp Phe Leu Val Lys Pro Leu Arg Lys Asn
180 185 190

```
Glu Leu Lys Asn Leu Trp Gln His Val Trp Arg Arg Cys His Ser Val
        195                 200                 205

Ser Cys Leu Phe Leu Ser Ser Gly Ser Glu Ser Gly Ile Gln Thr Gln
    210                 215                 220

Lys Cys Ala Lys Leu Asn Thr Gly Asp Glu Tyr Glu Asn Gly Ser Asp
225                 230                 235                 240

Ser Asn His Asp Asp Glu Glu Asn Asp Asp Gly Asp Asp Asp Asp Phe
                245                 250                 255

Ser Val Gly Leu Asn Ala Arg Asp Gly Ser Asp Asn Gly Ser Gly Thr
            260                 265                 270

Gln Ser Ser Trp Thr Lys Arg Ala Val Glu Ile Asp Ser Pro Gln Pro
        275                 280                 285

Ile Ser Pro Asp Gln Leu Val Asp Pro Pro Asp Ser Thr Cys Ala Gln
    290                 295                 300

Val Ile His Pro Arg Ser Glu Ile Cys Ser Asn Lys Trp Leu Pro Thr
305                 310                 315                 320

Ala Asn Lys Arg Asn Val Lys Lys Gln Lys Glu Asn Lys Asp Glu Ser
                325                 330                 335

Met Gly Arg Tyr Leu Gly Ile Gly Ala Pro Arg Asn Ser Ser Ala Glu
            340                 345                 350

Tyr Gln Ser Ser Leu Asn Asp Val Ser Val Asn Pro Ile Glu Lys Gly
        355                 360                 365

His Glu Asn His Met Ser Lys Cys Lys Ser Lys Lys Glu Thr Met Ala
    370                 375                 380

Glu Asp Asp Cys Thr Asn Met Pro Ser Ala Thr Asn Ala Glu Thr Ala
385                 390                 395                 400

Asp Leu Ile Ser Ser Ile Ala Arg Asn Thr Glu Gly Gln Gln Ala Val
                405                 410                 415

Gln Ala Val Asp Ala Pro Asp Gly Pro Ser Lys Met Ala Asn Gly Asn
            420                 425                 430

Asp Lys Asn His Asp Ser His Ile Glu Val Thr Pro His Glu Leu Gly
        435                 440                 445

Leu Lys Arg Ser Arg Thr Asn Gly Ala Thr Ala Glu Ile His Asp Glu
    450                 455                 460
```

```
Arg Asn Ile Leu Lys Arg Ser Asp Gln Ser Ala Phe Thr Arg Tyr His
465                 470                 475                 480

Thr Ser Val Ala Ser Asn Gln Gly Gly Ala Arg Tyr Gly Glu Ser Ser
                485                 490                 495

Ser Pro Gln Asp Asn Ser Ser Glu Ala Met Lys Thr Asp Ser Thr Cys
            500                 505                 510

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

Gly Ser Ser Asn Asn Asp Val Gly Ser Ser Thr Lys Asn Val Ala Ala
    530                 535                 540

Arg Pro Ser Gly Asp Arg Glu Arg Val Ala Ser Pro Leu Ala Ile Lys
545                 550                 555                 560

Ser Thr Gln His Ala Ser Ala Phe His Thr Ile Gln Asn Gln Thr Ser
                565                 570                 575

Pro Ala Asn Leu Ile Gly Glu Asp Lys Ala Asp Glu Gly Ile Ser Asn
            580                 585                 590

Thr Val Lys Met Ser His Pro Thr Glu Val Pro Gln Gly Cys Val Gln
        595                 600                 605

His His His His Val His Tyr Tyr Leu His Val Met Thr Gln Lys Gln
    610                 615                 620

Pro Ser Thr Asp Arg Gly Ser Ser Asp Val His Cys Gly Ser Ser Asn
625                 630                 635                 640

Val Phe Asp Pro Pro Val Glu Gly His Ala Ala Asn Tyr Ser Val Asn
                645                 650                 655

Gly Gly Val Ser Val Gly His Asn Gly Cys Asn Gly Gln Asn Gly Ser
            660                 665                 670

Ser Ala Val Pro Asn Ile Ala Arg Pro Asn Ile Glu Ser Ile Asn Gly
        675                 680                 685

Thr Met Ser Gln Asn Ile Ala Gly Gly Gly Ile Val Ser Gly Ser Gly
    690                 695                 700

Ser Gly Asn Asp Met Tyr Gln Asn Arg Phe Leu Gln Arg Glu Ala Ala
705                 710                 715                 720

Leu Asn Lys Phe Arg Leu Lys Arg Lys Asp Arg Asn Phe Gly Ser Leu
                725                 730                 735
```

Pro Lys Gln Glu Glu Ala Cys
740

<210> 19
<211> 2301
<212> DNA
<213> Zea mays

<400> 19
atgggcagtg cttgccaagc tggcacagac gggccttccc gcaaggatgt gttagggata 60
gggaatgccg ccttagagaa tggccaccat caggctgaag ctgacgcaga tgaatggagg 120
gaaaaggaag aggacttggc caacaacggg cacagtgcgc caccgccagg catgcagcag 180
gtggatgagc ataaggagga acaaagacaa agcattcact gggagaggtt cctacctgtg 240
aagacactga gagtcttgct ggtggagaat gatgactcta ctcgtcaggt ggtcagtgcc 300
ctgctccgta agtgctgcta tgaagttatt cctgctgaaa atggttttgca tgcatggcga 360
tatcttgaag atctgcagaa caacatcgac ctttgtattga ctgaggtttt catgccttgt 420
ctatctggta tcggtctgct tagcaaaatc acaagtcaca aaatttgcaa agacattcct 480
gtgattatga tgtctacgaa tgattctatg agtatggtgt ttaagtgttt gtcgaaggga 540
gcagttgatt tctttggtaaa accactacgt aagaatgagc ttaagaacct ttggcagcat 600
gttttggaggc gatgccacag ttccagtgga agtgaaagtg gcatccagac acagaagtgt 660
gccaaactaa atactggcga cgagtatgag aacggcagtg acagcaatca tgatgatgaa 720
gaaaatgatg acggcgacga tgacgacttc agtgttggac tcaatgctag ggatggaagt 780
gacaatggca gtggtactca aagctcatgg acaaagcgtg ctgtggagat tgacagccca 840
caacctatat ctcccgatca actagttgat ccacctgata gtacatgtgc acaagtaatt 900
cacctagat cagagatatg cagtaacaag tggttaccga cagcaaacaa aaggaatgtc 960
aagaaacaga aggagaataa agatgaatct atgggaagat acttaggaat aggtgctcct 1020
aggaactcaa gtgcagaata tcaatcatct ctcaatgatg tatctgttaa tccaatagaa 1080
aaaggacatg agaatcacat gtccaaatgc aaatctaaaa aggaaacaat ggcagaagat 1140
gattgtacaa acatgcctag tgcaacaaat gctgaaactg ctgatttgat tagctcaata 1200
gccagaaaca cagaaggcca acaagcagta caagccgttg acgcaccaga tggccttcc 1260
aaaatggcta atggaaatga taagaatcat gattctcata tcgaagtgac accccatgag 1320
ttgggttga agagatcgag aacaaatgga gctacagcgg aaatccatga tgagcgaaat 1380
attctgaaaa gatcagatca gtcagcctc accaggtacc atacatctgt ggcttccaat 1440
caaggtggag caagatatgg ggaaagctct tcaccacaag ataacagttc tgaggccatg 1500
aaaacggact ctacatgcaa gatgaagtca aattcagatg ctgctccaat aaagcagggc 1560
tccaatggca gtagcaataa cgatgtggga tccagtacaa agaatgttgc tgcaaggcct 1620
tcgggtgaca gggagagagt agcgtcacca ttagccatca aatctaccca gcatgcctca 1680
gcattttcata ctatacagaa tcaaacgtca ccagctaatc tgattgggga agacaaagct 1740

gatgaaggaa tttccaatac agtgaaaatg agccacccaa cagaggttcc acaaggctgc 1800

gtccagcatc atcatcatgt gcattattac ctccatgtta tgacacagaa acagccatca 1860

acagaccgtg gatcatcaga tgttcactgt ggttcgtcaa atgtgtttga tcctcctgtt 1920

gaaggacatg ctgcaaacta cagtgtgaat gggggtgtct cagttggtca taatgggtgc 1980

aatgggcaga atggaagtag cgctgtcccc aatattgcaa gaccaaacat agagagtatt 2040

aatggtacca tgagccaaaa tattgccgga ggtggcattg taagtgggag tgggagtggc 2100

aatgacatgt atcagaatcg gttcctgcaa cgagaagctg cattgaacaa attcagactg 2160

aagcggaaag atcggaactt tggtaaaaag gttcgctacc aaagcaggaa gaggcttgct 2220

gagcagcggc cacgggtccg aggacagttt gtgcgacaat ctgagcaaga agatcaaaca 2280

gcgcaaggtt cagaaagatg a 2301

<210> 20
<211> 766
<212> PRT
<213> Zea mays

<400> 20

Met Gly Ser Ala Cys Gln Ala Gly Thr Asp Gly Pro Ser Arg Lys Asp
1 5 10 15

Val Leu Gly Ile Gly Asn Ala Ala Leu Glu Asn Gly His His Gln Ala
20 25 30

Glu Ala Asp Ala Asp Glu Trp Arg Glu Lys Glu Glu Asp Leu Ala Asn
35 40 45

Asn Gly His Ser Ala Pro Pro Pro Gly Met Gln Gln Val Asp Glu His
50 55 60

Lys Glu Glu Gln Arg Gln Ser Ile His Trp Glu Arg Phe Leu Pro Val
65 70 75 80

Lys Thr Leu Arg Val Leu Leu Val Glu Asn Asp Asp Ser Thr Arg Gln
85 90 95

Val Val Ser Ala Leu Leu Arg Lys Cys Cys Tyr Glu Val Ile Pro Ala
100 105 110

Glu Asn Gly Leu His Ala Trp Arg Tyr Leu Glu Asp Leu Gln Asn Asn
115 120 125

Ile Asp Leu Val Leu Thr Glu Val Phe Met Pro Cys Leu Ser Gly Ile
130 135 140

Gly Leu Leu Ser Lys Ile Thr Ser His Lys Ile Cys Lys Asp Ile Pro
145 150 155 160

```
Val Ile Met Met Ser Thr Asn Asp Ser Met Ser Met Val Phe Lys Cys
                165                 170                 175

Leu Ser Lys Gly Ala Val Asp Phe Leu Val Lys Pro Leu Arg Lys Asn
            180                 185                 190

Glu Leu Lys Asn Leu Trp Gln His Val Trp Arg Arg Cys His Ser Ser
        195                 200                 205

Ser Gly Ser Glu Ser Gly Ile Gln Thr Gln Lys Cys Ala Lys Leu Asn
    210                 215                 220

Thr Gly Asp Glu Tyr Glu Asn Gly Ser Asp Ser Asn His Asp Asp Glu
225                 230                 235                 240

Glu Asn Asp Asp Gly Asp Asp Asp Asp Phe Ser Val Gly Leu Asn Ala
                245                 250                 255

Arg Asp Gly Ser Asp Asn Gly Ser Gly Thr Gln Ser Ser Trp Thr Lys
            260                 265                 270

Arg Ala Val Glu Ile Asp Ser Pro Gln Pro Ile Ser Pro Asp Gln Leu
        275                 280                 285

Val Asp Pro Pro Asp Ser Thr Cys Ala Gln Val Ile His Pro Arg Ser
    290                 295                 300

Glu Ile Cys Ser Asn Lys Trp Leu Pro Thr Ala Asn Lys Arg Asn Val
305                 310                 315                 320

Lys Lys Gln Lys Glu Asn Lys Asp Glu Ser Met Gly Arg Tyr Leu Gly
                325                 330                 335

Ile Gly Ala Pro Arg Asn Ser Ser Ala Glu Tyr Gln Ser Ser Leu Asn
            340                 345                 350

Asp Val Ser Val Asn Pro Ile Glu Lys Gly His Glu Asn His Met Ser
        355                 360                 365

Lys Cys Lys Ser Lys Lys Glu Thr Met Ala Glu Asp Asp Cys Thr Asn
    370                 375                 380

Met Pro Ser Ala Thr Asn Ala Glu Thr Ala Asp Leu Ile Ser Ser Ile
385                 390                 395                 400

Ala Arg Asn Thr Glu Gly Gln Gln Ala Val Gln Ala Val Asp Ala Pro
                405                 410                 415

Asp Gly Pro Ser Lys Met Ala Asn Gly Asn Asp Lys Asn His Asp Ser
            420                 425                 430
```

```
His Ile Glu Val Thr Pro His Glu Leu Gly Leu Lys Arg Ser Arg Thr
        435                 440                 445

Asn Gly Ala Thr Ala Glu Ile His Asp Glu Arg Asn Ile Leu Lys Arg
    450                 455                 460

Ser Asp Gln Ser Ala Phe Thr Arg Tyr His Thr Ser Val Ala Ser Asn
465                 470                 475                 480

Gln Gly Gly Ala Arg Tyr Gly Glu Ser Ser Ser Pro Gln Asp Asn Ser
                485                 490                 495

Ser Glu Ala Met Lys Thr Asp Ser Thr Cys Lys Met Lys Ser Asn Ser
            500                 505                 510

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

Val Gly Ser Ser Thr Lys Asn Val Ala Ala Arg Pro Ser Gly Asp Arg
    530                 535                 540

Glu Arg Val Ala Ser Pro Leu Ala Ile Lys Ser Thr Gln His Ala Ser
545                 550                 555                 560

Ala Phe His Thr Ile Gln Asn Gln Thr Ser Pro Ala Asn Leu Ile Gly
                565                 570                 575

Glu Asp Lys Ala Asp Glu Gly Ile Ser Asn Thr Val Lys Met Ser His
            580                 585                 590

Pro Thr Glu Val Pro Gln Gly Cys Val Gln His His His His Val His
        595                 600                 605

Tyr Tyr Leu His Val Met Thr Gln Lys Gln Pro Ser Thr Asp Arg Gly
    610                 615                 620

Ser Ser Asp Val His Cys Gly Ser Ser Asn Val Phe Asp Pro Pro Val
625                 630                 635                 640

Glu Gly His Ala Ala Asn Tyr Ser Val Asn Gly Gly Val Ser Val Gly
                645                 650                 655

His Asn Gly Cys Asn Gly Gln Asn Gly Ser Ser Ala Val Pro Asn Ile
            660                 665                 670

Ala Arg Pro Asn Ile Glu Ser Ile Asn Gly Thr Met Ser Gln Asn Ile
        675                 680                 685

Ala Gly Gly Gly Ile Val Ser Gly Ser Gly Ser Gly Asn Asp Met Tyr
    690                 695                 700
```

Gln Asn Arg Phe Leu Gln Arg Glu Ala Ala Leu Asn Lys Phe Arg Leu
705 710 715 720

Lys Arg Lys Asp Arg Asn Phe Gly Lys Lys Val Arg Tyr Gln Ser Arg
725 730 735

Lys Arg Leu Ala Glu Gln Arg Pro Arg Val Arg Gly Gln Phe Val Arg
740 745 750

Gln Ser Glu Gln Glu Asp Gln Thr Ala Gln Gly Ser Glu Arg
755 760 765

<210> 21
<211> 1812
<212> DNA
<213> Zea mays

<400> 21
atgtctacga atgattctat gagtatggtg tttaagtgtt tgtcgaaggg agcagttgat 60

ttcttggtaa aaccactacg taagaatgag cttaagaacc tttggcagca tgtttggagg 120

cgatgccaca gttccagtgg aagtgaaagt ggcatccaga cacagaagtg tgccaaacta 180

aatactggcg acgagtatga gaacggcagt gacagcaatc atgatgatga agaaaatgat 240

gacggcgacg atgacgactt cagtgttgga ctcaatgcta gggatggaag tgacaatggc 300

agtggtactc aaagctcatg gacaaagcgt gctgtggaga ttgacagccc acaacctata 360

tctcccgatc aactagttga tccacctgat agtacatgtg cacaagtaat tcaccctaga 420

tcagagatat gcagtaacaa gtggttaccg acagcaaaca aaaggaatgt caagaaacag 480

aaggagaata aagatgaatc tatgggaaga tacttaggaa taggtgctcc taggaactca 540

agtgcagaat atcaatcatc tctcaatgat gtatctgtta atccaataga aaaaggacat 600

gagaatcaca tgtccaaatg caaatctaaa aaggaaacaa tggcagaaga tgattgtaca 660

aacatgccta gtgcaacaaa tgctgaaact gctgatttga ttagctcaat agccagaaac 720

acagaaggcc aacaagcagt acaagccgtt gacgcaccag atggccttc caaaatggct 780

aatggaaatg ataagaatca tgattctcat atcgaagtga cacccatga gttgggttg 840

aagagatcga gaacaaatgg agctacagcg gaaatccatg atgagcgaaa tattctgaaa 900

agatcagatc agtcagcctt caccaggtac catacatctg tggcttccaa tcaaggtgga 960

gcaagatatg gggaaagctc ttcaccacaa gataacagtt ctgaggccat gaaaacggac 1020

tctacatgca agatgaagtc aaattcagat gctgctccaa taaagcaggg ctccaatggc 1080

agtagcaata acgatgtggg atccagtaca aagaatgttg ctgcaaggcc ttcgggtgac 1140

agggagagag tagcgtcacc attagccatc aaatctaccc agcatgcctc agcattttcat 1200

actatacaga atcaaacgtc accagctaat ctgattgggg aagacaaagc tgatgaagga 1260

attttccaata cagtgaaaat gagccaccca acagaggttc cacaaggctg cgtccagcat 1320

catcatcatg tgcattatta cctccatgtt atgacacaga aacagccatc aacagaccgt 1380

ggatcatcag atgttcactg tggttcgtca aatgtgtttg atcctcctgt tgaaggacat 1440

gctgcaaact acagtgtgaa tgggggtgtc tcagttggtc ataatgggtg caatgggcag 1500

aatggaagta gcgctgtccc caatattgca agaccaaaca tagagagtat taatggtacc 1560

atgagccaaa atattgccgg aggtggcatt gtaagtggga gtgggagtgg caatgacatg 1620

tatcagaatc ggttcctgca acgagaagct gcattgaaca aattcagact gaagcggaaa 1680

gatcggaact ttggtaaaaa ggttcgctac caaagcagga agaggcttgc tgagcagcgg 1740

ccacgggtcc gaggacagtt tgtgcgacaa tctgagcaag aagatcaaac agcgcaaggt 1800

tcagaaagat ga 1812

<210> 22
<211> 603
<212> PRT
<213> Zea mays

<400> 22

Met Ser Thr Asn Asp Ser Met Ser Met Val Phe Lys Cys Leu Ser Lys
1 5 10 15

Gly Ala Val Asp Phe Leu Val Lys Pro Leu Arg Lys Asn Glu Leu Lys
20 25 30

Asn Leu Trp Gln His Val Trp Arg Arg Cys His Ser Ser Ser Gly Ser
35 40 45

Glu Ser Gly Ile Gln Thr Gln Lys Cys Ala Lys Leu Asn Thr Gly Asp
50 55 60

Glu Tyr Glu Asn Gly Ser Asp Ser Asn His Asp Asp Glu Glu Asn Asp
65 70 75 80

Asp Gly Asp Asp Asp Asp Phe Ser Val Gly Leu Asn Ala Arg Asp Gly
85 90 95

Ser Asp Asn Gly Ser Gly Thr Gln Ser Ser Trp Thr Lys Arg Ala Val
100 105 110

Glu Ile Asp Ser Pro Gln Pro Ile Ser Pro Asp Gln Leu Val Asp Pro
115 120 125

Pro Asp Ser Thr Cys Ala Gln Val Ile His Pro Arg Ser Glu Ile Cys
130 135 140

Ser Asn Lys Trp Leu Pro Thr Ala Asn Lys Arg Asn Val Lys Lys Gln
145 150 155 160

Lys Glu Asn Lys Asp Glu Ser Met Gly Arg Tyr Leu Gly Ile Gly Ala
165 170 175

```
Pro Arg Asn Ser Ser Ala Glu Tyr Gln Ser Ser Leu Asn Asp Val Ser
            180                 185                 190

Val Asn Pro Ile Glu Lys Gly His Glu Asn His Met Ser Lys Cys Lys
        195                 200                 205

Ser Lys Lys Glu Thr Met Ala Glu Asp Asp Cys Thr Asn Met Pro Ser
    210                 215                 220

Ala Thr Asn Ala Glu Thr Ala Asp Leu Ile Ser Ser Ile Ala Arg Asn
225                 230                 235                 240

Thr Glu Gly Gln Gln Ala Val Gln Ala Val Asp Ala Pro Asp Gly Pro
                245                 250                 255

Ser Lys Met Ala Asn Gly Asn Asp Lys Asn His Asp Ser His Ile Glu
            260                 265                 270

Val Thr Pro His Glu Leu Gly Leu Lys Arg Ser Arg Thr Asn Gly Ala
        275                 280                 285

Thr Ala Glu Ile His Asp Glu Arg Asn Ile Leu Lys Arg Ser Asp Gln
    290                 295                 300

Ser Ala Phe Thr Arg Tyr His Thr Ser Val Ala Ser Asn Gln Gly Gly
305                 310                 315                 320

Ala Arg Tyr Gly Glu Ser Ser Ser Pro Gln Asp Asn Ser Ser Glu Ala
                325                 330                 335

Met Lys Thr Asp Ser Thr Cys Lys Met Lys Ser Asn Ser Asp Ala Ala
            340                 345                 350

Pro Ile Lys Gln Gly Ser Asn Gly Ser Ser Asn Asn Asp Val Gly Ser
        355                 360                 365

Ser Thr Lys Asn Val Ala Ala Arg Pro Ser Gly Asp Arg Glu Arg Val
    370                 375                 380

Ala Ser Pro Leu Ala Ile Lys Ser Thr Gln His Ala Ser Ala Phe His
385                 390                 395                 400

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

Ala Asp Glu Gly Ile Ser Asn Thr Val Lys Met Ser His Pro Thr Glu
            420                 425                 430

Val Pro Gln Gly Cys Val Gln His His His His Val His Tyr Tyr Leu
        435                 440                 445
```

His Val Met Thr Gln Lys Gln Pro Ser Thr Asp Arg Gly Ser Ser Asp
450 455 460

Val His Cys Gly Ser Ser Asn Val Phe Asp Pro Pro Val Glu Gly His
465 470 475 480

Ala Ala Asn Tyr Ser Val Asn Gly Gly Val Ser Val Gly His Asn Gly
485 490 495

Cys Asn Gly Gln Asn Gly Ser Ser Ala Val Pro Asn Ile Ala Arg Pro
500 505 510

Asn Ile Glu Ser Ile Asn Gly Thr Met Ser Gln Asn Ile Ala Gly Gly
515 520 525

Gly Ile Val Ser Gly Ser Gly Ser Gly Asn Asp Met Tyr Gln Asn Arg
530 535 540

Phe Leu Gln Arg Glu Ala Ala Leu Asn Lys Phe Arg Leu Lys Arg Lys
545 550 555 560

Asp Arg Asn Phe Gly Lys Lys Val Arg Tyr Gln Ser Arg Lys Arg Leu
565 570 575

Ala Glu Gln Arg Pro Arg Val Arg Gly Gln Phe Val Arg Gln Ser Glu
580 585 590

Gln Glu Asp Gln Thr Ala Gln Gly Ser Glu Arg
595 600

<210> 23
<211> 489
<212> DNA
<213> Saccharum officinarum

<400> 23
atgggtagcg cttgccaagc tggcacggac gggtcttccc gcaaggatgt gttggggata 60

gggaatgtca ccttagataa tggccaccat gaggctgaag ctgatgcaga tgaatggagg 120

gaaaaggaag atgacttagc caatgggcgc agtgcgccac cgggcatgca gcaggtggat 180

gaacagaagg agcaacaagg acaaagcatt cactgggaga ggttcctacc tgtgaagaca 240

ctgagagtct tgctggtgga gaatgatgac tctactcgtc aggtggtcag tgccctgctc 300

cgtaagtgct gctatgaagt tatccctgct gaaaatggtt cacatgcatg gcgatatctt 360

gaaaatctgc agaacaacat tgaccttgta ttgactgagg ttttcatgcc ttgtctatct 420

ggcatcggtc tgcttagcaa aatcactagt cacaaaattt gcaaggacat tcctgtgatt 480

agtaagtag 489

<210> 24
<211> 162

<212> PRT
<213> Saccharum officinarum

<400> 24

Met Gly Ser Ala Cys Gln Ala Gly Thr Asp Gly Ser Ser Arg Lys Asp
1 5 10 15

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

Glu Ala Asp Ala Asp Glu Trp Arg Glu Lys Glu Asp Asp Leu Ala Asn
35 40 45

Gly Arg Ser Ala Pro Pro Gly Met Gln Gln Val Asp Glu Gln Lys Glu
50 55 60

Gln Gln Gly Gln Ser Ile His Trp Glu Arg Phe Leu Pro Val Lys Thr
65 70 75 80

Leu Arg Val Leu Leu Val Glu Asn Asp Asp Ser Thr Arg Gln Val Val
85 90 95

Ser Ala Leu Leu Arg Lys Cys Cys Tyr Glu Val Ile Pro Ala Glu Asn
100 105 110

Gly Ser His Ala Trp Arg Tyr Leu Glu Asn Leu Gln Asn Asn Ile Asp
115 120 125

Leu Val Leu Thr Glu Val Phe Met Pro Cys Leu Ser Gly Ile Gly Leu
130 135 140

Leu Ser Lys Ile Thr Ser His Lys Ile Cys Lys Asp Ile Pro Val Ile
145 150 155 160

Ser Lys

<210> 25
<211> 1797
<212> DNA
<213> Panicum virgatum

<400> 25
atgtcttcga atgactctat gagtatggtg tttaagtgtt tgtcaaaggg agcagttgac 60

ttcttagtaa agccactacg gaagaatgag cttaagaacc tttggcagca cgttttggagg 120

cggtgtcaca gttccagtgg tagtagaagt gaaagtggca tccagacaca aaagtgtgcc 180

ataccaaata ctggtgacaa gtatgagaac aacagtaaca gcaatcatga tgatgacgaa 240

gatgatgact tgagtgtcag actcaatgct agggatggaa gtgataatgg cagtggtact 300

caaagttcat ggacaaagcg tgctgtggag attgacagtc cacaacccat gtctcctgat 360

cagctagttg atccacctga tagtacatgt gcacaagtaa ttcaccccaa atctgagata 420

tgcagtaaca agtggttacc aggtgcaaac aaaaggaata gcaagaaaaa gaaggagaat 480

aaagatgaat ctatggggaa atattagaa ataggtgctc ctaggaattc gagtgtagaa 540

tatcaatcat ctctcaatga tacctctgtt aatccaacag gaaaaagaca tgagattcac 600

attcctcaat tcaaatctaa aaagaaagtg atggcagatg atgattgcac aaacatgatg 660

agtgaaccaa atactgaaac ggctgattg attagctcaa tagccagaaa tacagaaggc 720

cagcaagcag tacaagttgc tgatgcacct ggttgccctt ccagaatacc cgatggaaat 780

gataagaatc atgattccca tatccaagtg acaccacatg agttgggttt gaagagattg 840

aaaacaaatg gagctaccac ggaaatccat gatgagcgaa atattctgag aagatcagat 900

ttgtcagcct tcaccaggta ccatacatct gtggcttcca atcaaggtgg agcaagattt 960

ggggaaagct cttcgccgca agataacagt tctgaggctg ttaaaacaga ctctacctgc 1020

aagatgaagt caaattcaga tgctcctcca ataaagcagg gctccaatgg cagtagcaac 1080

aacaatgaca tgggctccag tacaaagaac gttgtcgtaa agcctt cggg taacagggag 1140

agagtaacgt caccatcagc tgccaaatct attcaacata cctcagcatt ccatcctatg 1200

ccacatcaaa catcaccagc taatgtagtt gggaaagaca aaactgatga aggaattttcc 1260

aatgcagtga aagtgggcca gacagaggta ccacaaagct gtgtgcaaca tcatcatcac 1320

gtccactact acctccatgt tatgacacag aaacagccat caattgacca tggatcatca 1380

gacgctcagt gtggttcatc caatgtgttt gatcctcctg ttgaaggaca cgctgcaaac 1440

tacagtgtga atggaggtgt ctctgttggt catAatgggt gcaatgggca gaatggaagt 1500

agcgctgctc ccaatattgc aagaccaaac atggagtgtg ttaatggcac catgagtaaa 1560

aatgtggctg gaggtggtag tggaagtgga agtggcaatg acatgtatca gaatcggttc 1620

cctcaacgag aagctgcatt gaacaaattc agattgaagc ggaaagatcg gaacttcgat 1680

aaaaaggttc gataccaaag caggaagagg ctggctgagc agcggccacg ggtccgtggg 1740

cagttttgtgc gacaatctgg acaagaagat aaagcaggcc aggattcaga gagatga 1797

<210> 26
<211> 598
<212> PRT
<213> Panicum virgatum

<400> 26

Met Ser Ser Asn Asp Ser Met Ser Met Val Phe Lys Cys Leu Ser Lys
1 5 10 15

Gly Ala Val Asp Phe Leu Val Lys Pro Leu Arg Lys Asn Glu Leu Lys
20 25 30

Asn Leu Trp Gln His Val Trp Arg Arg Cys His Ser Ser Ser Gly Ser
35 40 45

Arg Ser Glu Ser Gly Ile Gln Thr Gln Lys Cys Ala Ile Pro Asn Thr

```
     50                  55                  60

Gly Asp Lys Tyr Glu Asn Asn Ser Asn Ser Asn His Asp Asp Asp Glu
65                  70                  75                  80

Asp Asp Asp Leu Ser Val Arg Leu Asn Ala Arg Asp Gly Ser Asp Asn
                85                  90                  95

Gly Ser Gly Thr Gln Ser Ser Trp Thr Lys Arg Ala Val Glu Ile Asp
            100                 105                 110

Ser Pro Gln Pro Met Ser Pro Asp Gln Leu Val Asp Pro Pro Asp Ser
        115                 120                 125

Thr Cys Ala Gln Val Ile His Pro Lys Ser Glu Ile Cys Ser Asn Lys
    130                 135                 140

Trp Leu Pro Gly Ala Asn Lys Arg Asn Ser Lys Lys Lys Lys Glu Asn
145                 150                 155                 160

Lys Asp Glu Ser Met Gly Lys Tyr Leu Glu Ile Gly Ala Pro Arg Asn
                165                 170                 175

Ser Ser Val Glu Tyr Gln Ser Ser Leu Asn Asp Thr Ser Val Asn Pro
            180                 185                 190

Thr Gly Lys Arg His Glu Ile His Ile Pro Gln Phe Lys Ser Lys Lys
        195                 200                 205

Lys Val Met Ala Asp Asp Asp Cys Thr Asn Met Met Ser Glu Pro Asn
    210                 215                 220

Thr Glu Thr Ala Asp Leu Ile Ser Ser Ile Ala Arg Asn Thr Glu Gly
225                 230                 235                 240

Gln Gln Ala Val Gln Val Ala Asp Ala Pro Gly Cys Pro Ser Arg Ile
                245                 250                 255

Pro Asp Gly Asn Asp Lys Asn His Asp Ser His Ile Gln Val Thr Pro
            260                 265                 270

His Glu Leu Gly Leu Lys Arg Leu Lys Thr Asn Gly Ala Thr Thr Glu
        275                 280                 285

Ile His Asp Glu Arg Asn Ile Leu Arg Arg Ser Asp Leu Ser Ala Phe
    290                 295                 300

Thr Arg Tyr His Thr Ser Val Ala Ser Asn Gln Gly Gly Ala Arg Phe
305                 310                 315                 320

Gly Glu Ser Ser Ser Pro Gln Asp Asn Ser Ser Glu Ala Val Lys Thr
```

325 330 335

Asp Ser Thr Cys Lys Met Lys Ser Asn Ser Asp Ala Pro Pro Ile Lys
340 345 350

Gln Gly Ser Asn Gly Ser Ser Asn Asn Asn Asp Met Gly Ser Ser Thr
355 360 365

Lys Asn Val Val Val Lys Pro Ser Gly Asn Arg Glu Arg Val Thr Ser
370 375 380

Pro Ser Ala Ala Lys Ser Ile Gln His Thr Ser Ala Phe His Pro Met
385 390 395 400

Pro His Gln Thr Ser Pro Ala Asn Val Val Gly Lys Asp Lys Thr Asp
405 410 415

Glu Gly Ile Ser Asn Ala Val Lys Val Gly Gln Thr Glu Val Pro Gln
420 425 430

Ser Cys Val Gln His His His His Val His Tyr Tyr Leu His Val Met
435 440 445

Thr Gln Lys Gln Pro Ser Ile Asp His Gly Ser Ser Asp Ala Gln Cys
450 455 460

Gly Ser Ser Asn Val Phe Asp Pro Pro Val Glu Gly His Ala Ala Asn
465 470 475 480

Tyr Ser Val Asn Gly Gly Val Ser Val Gly His Asn Gly Cys Asn Gly
485 490 495

Gln Asn Gly Ser Ser Ala Ala Pro Asn Ile Ala Arg Pro Asn Met Glu
500 505 510

Cys Val Asn Gly Thr Met Ser Lys Asn Val Ala Gly Gly Gly Ser Gly
515 520 525

Ser Gly Ser Gly Asn Asp Met Tyr Gln Asn Arg Phe Pro Gln Arg Glu
530 535 540

Ala Ala Leu Asn Lys Phe Arg Leu Lys Arg Lys Asp Arg Asn Phe Asp
545 550 555 560

Lys Lys Val Arg Tyr Gln Ser Arg Lys Arg Leu Ala Glu Gln Arg Pro
565 570 575

Arg Val Arg Gly Gln Phe Val Arg Gln Ser Gly Gln Glu Asp Lys Ala
580 585 590

Gly Gln Asp Ser Glu Arg

595

<210> 27
<211> 489
<212> DNA
<213> Saccharum officinarum

<400> 27
atgggtagcg cttgccaagc tggcacggac gggtcttccc gcaaggatgt gttggggata 60

gggaatgtca ccttagataa tggccaccat gaggctgaag ctgatgcaga tgaatggagg 120

gaaaaggaag atgacttaac caatgggcgc agtgcgccac cgggcatgca gcaggtggat 180

gaacagaagg agcaacaagg acaaagcatt cactgggaga ggttcctacc tgtgaagaca 240

ctgagagtct tgctggtgga gaatgatgac tctactcgtc aggtggtcag tgccctgctc 300

cgtaagtgct gctatgaagt tatccctgct gaaaatggtt cacatgcatg gcgatatctt 360

gaaaatctgc agaacaacat tgaccttgta ttgactgagg ttttcatgcc ttgtctatct 420

ggcatcggtc tgcttagcaa aatcactagt cacaaaattt gcaaggacat tcctgtgatt 480

agtaagtag 489

<210> 28
<211> 162
<212> PRT
<213> Saccharum officinarum

<400> 28

Met Gly Ser Ala Cys Gln Ala Gly Thr Asp Gly Ser Ser Arg Lys Asp
1 5 10 15

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

Glu Ala Asp Ala Asp Glu Trp Arg Glu Lys Glu Asp Asp Leu Thr Asn
35 40 45

Gly Arg Ser Ala Pro Pro Gly Met Gln Gln Val Asp Glu Gln Lys Glu
50 55 60

Gln Gln Gly Gln Ser Ile His Trp Glu Arg Phe Leu Pro Val Lys Thr
65 70 75 80

Leu Arg Val Leu Leu Val Glu Asn Asp Asp Ser Thr Arg Gln Val Val
85 90 95

Ser Ala Leu Leu Arg Lys Cys Cys Tyr Glu Val Ile Pro Ala Glu Asn
100 105 110

Gly Ser His Ala Trp Arg Tyr Leu Glu Asn Leu Gln Asn Asn Ile Asp
115 120 125

Leu Val Leu Thr Glu Val Phe Met Pro Cys Leu Ser Gly Ile Gly Leu

130 135 140

Leu Ser Lys Ile Thr Ser His Lys Ile Cys Lys Asp Ile Pro Val Ile
145 150 155 160

Ser Lys

<210> 29
<211> 771
<212> DNA
<213> Saccharum officinarum

<400> 29
atgggtagcg cttgccaagc tggcacggac gggtcttccc gcaaggatgt gttgggtata 60

gggaatgtca ccttagataa tggccaccat gaggctgaag ctgatgcaga tgaatggagg 120

gaaaaggaag atgacttagc caatgggcac agtgcgccgc cgggcatgca gcaggtggat 180

gagcagaagg agcaacaagg acaaagcatt cactgggaga ggttcctacc tgtgaagaca 240

ctgagagtct tgctggtgga gaatgatgac tctactcgtc aggtggtcag tgccctgctc 300

cgtaagtgct gctatgaagt tatccctgct gaaaatggtt cacatgcatg gcgatatctt 360

gaaaatctgc agaacaacat tgaccttgta ttgactgagg ttttcatgcc ttgtctatct 420

ggcatcggtc tgcttagcaa aatcactagt cacaaaattt gcaaggacat tcctgtgatt 480

atgatgtctt caaatgactc tatgagtatg gtgtttaagt gttttgtcga gggagcagtt 540

gacttcttgg taaagccact acgtaagaat gagcttaaga acctttggca gcacgtttgg 600

aggcgatgcc acagttccag tggcagtgga agtgaaagtg gcatccagac acagaagtgt 660

gccaaaccaa atactggtga cgagtatgag aacgacagtg acagcaatca tgatgatgaa 720

gaaaatgatg aacacaacga tgacaacttc agtgttcgac ttcaatgcta g 771

<210> 30
<211> 256
<212> PRT
<213> Saccharum officinarum

<400> 30

Met Gly Ser Ala Cys Gln Ala Gly Thr Asp Gly Ser Ser Arg Lys Asp
1 5 10 15

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

Glu Ala Asp Ala Asp Glu Trp Arg Glu Lys Glu Asp Asp Leu Ala Asn
35 40 45

Gly His Ser Ala Pro Pro Gly Met Gln Gln Val Asp Glu Gln Lys Glu
50 55 60

Gln Gln Gly Gln Ser Ile His Trp Glu Arg Phe Leu Pro Val Lys Thr

65 70 75 80

Leu Arg Val Leu Leu Val Glu Asn Asp Asp Ser Thr Arg Gln Val Val
85 90 95

Ser Ala Leu Leu Arg Lys Cys Cys Tyr Glu Val Ile Pro Ala Glu Asn
100 105 110

Gly Ser His Ala Trp Arg Tyr Leu Glu Asn Leu Gln Asn Asn Ile Asp
115 120 125

Leu Val Leu Thr Glu Val Phe Met Pro Cys Leu Ser Gly Ile Gly Leu
130 135 140

Leu Ser Lys Ile Thr Ser His Lys Ile Cys Lys Asp Ile Pro Val Ile
145 150 155 160

Met Met Ser Ser Asn Asp Ser Met Ser Met Val Phe Lys Cys Leu Ser
165 170 175

Lys Gly Ala Val Asp Phe Leu Val Lys Pro Leu Arg Lys Asn Glu Leu
180 185 190

Lys Asn Leu Trp Gln His Val Trp Arg Arg Cys His Ser Ser Ser Gly
195 200 205

Ser Gly Ser Glu Ser Gly Ile Gln Thr Gln Lys Cys Ala Lys Pro Asn
210 215 220

Thr Gly Asp Glu Tyr Glu Asn Asp Ser Asp Ser Asn His Asp Asp Glu
225 230 235 240

Glu Asn Asp Glu His Asn Asp Asp Asn Phe Ser Val Arg Leu Gln Cys
245 250 255

<210> 31
<211> 873
<212> DNA
<213> Panicum virgatum

<400> 31
atgggtagcg cctgccaagc tggcttggat gggcctcaca aggatgtgag agggatcgca 60

aatgttgcca ctgagaatgg ccaccatggg gccgaggctg atgcggatga atggagagaa 120

aaagaagagg acttgcccaa tgggcacagt gcgccaccgg gcgcacagca gttggatgag 180

cagaaggagc aacagactat tctgtgggag aggttcctcc ctgtgaagac actgagagtc 240

ttgctggtgg agaacgatga ctctactcgt caggtggtca gtgctctgct ccgtaagtgc 300

tgctatgaag ttatccctgc tgaaaatggc ttacatgcat ggcaacatct tgaagacctg 360

cagaacaaca ttgaccttgt attgactgag gttttcatgc ctcgtctatc tggcatcggt 420

ctgcttagca aaatcactag tcacaaagtc tgcaaggata ttcctgtgat tatgatgtct 480

tcgaatgact ctatgagtat ggtgtttaag tgtttgtcaa agggagcagt tgacttctta 540

gtaaagccac tacggaagaa tgagcttaag aacctttggc agcacgtttg gaggcggtgt 600

cacagtggct ttttgttttt ggtctcgctt ttttgttgga tattaattct gcagaaccat 660

gatgcattga tgcatagttt cttttactct acttcttgct ctggcttttc tttgccatta 720

ctcttagttg gatttgcctc tttggtgctc agtccagtgg tagtggaagt gaaagtggca 780

tccagacaca aaagtgtgcc aaaccaaata ctggtgacaa gtatgagaac aacagtaaca 840

gcaatcatga tgatgacgaa gatgatgact tga 873

<210> 32
<211> 290
<212> PRT
<213> Panicum virgatum

<400> 32

Met Gly Ser Ala Cys Gln Ala Gly Leu Asp Gly Pro His Lys Asp Val
1 5 10 15

Arg Gly Ile Ala Asn Val Ala Thr Glu Asn Gly His His Gly Ala Glu
20 25 30

Ala Asp Ala Asp Glu Trp Arg Glu Lys Glu Glu Asp Leu Pro Asn Gly
35 40 45

His Ser Ala Pro Pro Gly Ala Gln Gln Leu Asp Glu Gln Lys Glu Gln
50 55 60

Gln Thr Ile Leu Trp Glu Arg Phe Leu Pro Val Lys Thr Leu Arg Val
65 70 75 80

Leu Leu Val Glu Asn Asp Asp Ser Thr Arg Gln Val Val Ser Ala Leu
85 90 95

Leu Arg Lys Cys Cys Tyr Glu Val Ile Pro Ala Glu Asn Gly Leu His
100 105 110

Ala Trp Gln His Leu Glu Asp Leu Gln Asn Asn Ile Asp Leu Val Leu
115 120 125

Thr Glu Val Phe Met Pro Arg Leu Ser Gly Ile Gly Leu Leu Ser Lys
130 135 140

Ile Thr Ser His Lys Val Cys Lys Asp Ile Pro Val Ile Met Met Ser
145 150 155 160

Ser Asn Asp Ser Met Ser Met Val Phe Lys Cys Leu Ser Lys Gly Ala
165 170 175

Val Asp Phe Leu Val Lys Pro Leu Arg Lys Asn Glu Leu Lys Asn Leu

180 185 190

Trp Gln His Val Trp Arg Arg Cys His Ser Gly Phe Leu Phe Leu Val
195 200 205

Ser Leu Phe Cys Trp Ile Leu Ile Leu Gln Asn His Asp Ala Leu Met
210 215 220

His Ser Phe Phe Tyr Ser Thr Ser Cys Ser Gly Phe Ser Leu Pro Leu
225 230 235 240

Leu Leu Val Gly Phe Ala Ser Leu Val Leu Ser Pro Val Val Val Glu
245 250 255

Val Lys Val Ala Ser Arg His Lys Ser Val Pro Asn Gln Ile Leu Val
260 265 270

Thr Ser Met Arg Thr Thr Val Thr Ala Ile Met Met Met Thr Lys Met
275 280 285

Met Thr
290

<210> 33
<211> 879
<212> DNA
<213> Panicum virgatum

<400> 33
atgggtagca cctgccaagc cggctcggat gggcctcaca aggatgtgag agggatcgca 60

aatggcgcca ctgagaatgg ccaccatggg gccgaggctg atgcggatga atggagagaa 120

aaagaagagg acttacccaa tgggcacagt gcgccactgg gcgcacagca gttggatgag 180

cagaaggagc aacagactat tcagtgggag aggttcctcc ctgtgaagac actgagagtc 240

ttgctggtgg agaatgatga atctactcgt caggtggtca gtgccctgct ccgtaagtgc 300

tgctatgaag ttatccctgc tgaaaatggc ttacatgcat ggcaacatct tgaagatctg 360

cagaacaaca ttgaccttgt attgactgag gttttcatgc cttgtctatc tggcatcggt 420

ctgcttagca aaatcacaag tcacaaagtc tgcaaggaca ttcctgtgat tatgatgtct 480

tcgaatgact ctatgagtat ggtgttaag tgttttgtcaa agggagcagt tgatttctta 540

gtaaagccac tacggaagaa tgagcttaag aaccttggc agcacgttg gaggcggtgt 600

cacagttcca gtggcagtgg aagtgaaagt gccatccaga cacaaaagtg tgccaaacca 660

aatactagtg atgagtatga gaacaacagt aacagcaatc atgatgatga tgaaaacgat 720

gatgacttga gtgtcggact caatgctagg gatggaagtg ataacggcag tggtactcaa 780

agttcatgga caaagcgtgc tgtggagatt gacagtccac aacccatgtc tcctgatcag 840

ctagttgatc cacctgatag tacatgtgca caagtaatt 879

<210> 34
<211> 293
<212> PRT
<213> Panicum virgatum

<400> 34

Met Gly Ser Thr Cys Gln Ala Gly Ser Asp Gly Pro His Lys Asp Val
1 5 10 15

Arg Gly Ile Ala Asn Gly Ala Thr Glu Asn Gly His His Gly Ala Glu
20 25 30

Ala Asp Ala Asp Glu Trp Arg Glu Lys Glu Glu Asp Leu Pro Asn Gly
35 40 45

His Ser Ala Pro Leu Gly Ala Gln Gln Leu Asp Glu Gln Lys Glu Gln
50 55 60

Gln Thr Ile Gln Trp Glu Arg Phe Leu Pro Val Lys Thr Leu Arg Val
65 70 75 80

Leu Leu Val Glu Asn Asp Glu Ser Thr Arg Gln Val Val Ser Ala Leu
85 90 95

Leu Arg Lys Cys Cys Tyr Glu Val Ile Pro Ala Glu Asn Gly Leu His
100 105 110

Ala Trp Gln His Leu Glu Asp Leu Gln Asn Asn Ile Asp Leu Val Leu
115 120 125

Thr Glu Val Phe Met Pro Cys Leu Ser Gly Ile Gly Leu Leu Ser Lys
130 135 140

Ile Thr Ser His Lys Val Cys Lys Asp Ile Pro Val Ile Met Met Ser
145 150 155 160

Ser Asn Asp Ser Met Ser Met Val Phe Lys Cys Leu Ser Lys Gly Ala
165 170 175

Val Asp Phe Leu Val Lys Pro Leu Arg Lys Asn Glu Leu Lys Asn Leu
180 185 190

Trp Gln His Val Trp Arg Arg Cys His Ser Ser Ser Gly Ser Gly Ser
195 200 205

Glu Ser Ala Ile Gln Thr Gln Lys Cys Ala Lys Pro Asn Thr Ser Asp
210 215 220

Glu Tyr Glu Asn Asn Ser Asn Ser Asn His Asp Asp Asp Glu Asn Asp
225 230 235 240

Asp Asp Leu Ser Val Gly Leu Asn Ala Arg Asp Gly Ser Asp Asn Gly

245 250 255

Ser Gly Thr Gln Ser Ser Trp Thr Lys Arg Ala Val Glu Ile Asp Ser
260 265 270

Pro Gln Pro Met Ser Pro Asp Gln Leu Val Asp Pro Pro Asp Ser Thr
275 280 285

Cys Ala Gln Val Ile
290

<210> 35
<211> 811
<212> DNA
<213> Triticum aestivum

<400> 35
atggtgagcg ccggtcaagc tggcgcggac ggaccttcca ccagtgatat taggggaact 60

ggaaacggcg ccgtagagaa tggccatgcc ctcaaggcaa acgaggacaa ggaatggagg 120

ggcggcagca aggaagagga ctggcccagc acgcacagtg cgccgccagg cttggacgag 180

cacaagcagc agcaaggccg ggtcatccgc tgggagaagt tcctgccggt ggagacacta 240

agggtcttgc tggtggagaa cgatgactgt acccgacatg tcgtccgtgc tctgctccgt 300

aagtgtggct atgaagttat cgctgctgag aatggattgc atgcatggca ttatcttgaa 360

gatgtgcaaa accgtattga ccttgtatta actgaggtcg ccatgccttg tctatctggc 420

atcggtctac tcagtaagat cacgagtcac agtatttgca agggcattcc tgtgatcatg 480

atgtctaaga atgactcgat gagtacagtc tttaggtgtc tatcaaaggg agcagttgac 540

ttcttagtga agccgatacg gaagaatgaa cttaagaccc tttggcagca catatggagg 600

cgatgccaca gttccagtgg aagtgaaagt ggcatccata cacaaaaatg ttccaaaccg 660

aaggctggtg atgaatatga gaacaacagt catgatgacg atgacgattg cggcagtcat 720

gatgacgatg acgatgacga tgatgatgcc gatgacgact ttagtgttgg gcccaatgct 780

agggatggca gtgataatgg cagtggcact c 811

<210> 36
<211> 270
<212> PRT
<213> Triticum aestivum

<400> 36

Met Val Ser Ala Gly Gln Ala Gly Ala Asp Gly Pro Ser Thr Ser Asp
1 5 10 15

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

Ala Asn Glu Asp Lys Glu Trp Arg Gly Gly Ser Lys Glu Glu Asp Trp
35 40 45

```
Pro Ser Thr His Ser Ala Pro Pro Gly Leu Asp Glu His Lys Gln Gln
    50                  55                  60

Gln Gly Arg Val Ile Arg Trp Glu Lys Phe Leu Pro Val Glu Thr Leu
65                  70                  75                  80

Arg Val Leu Leu Val Glu Asn Asp Asp Cys Thr Arg His Val Val Arg
                85                  90                  95

Ala Leu Leu Arg Lys Cys Gly Tyr Glu Val Ile Ala Ala Glu Asn Gly
            100                 105                 110

Leu His Ala Trp His Tyr Leu Glu Asp Val Gln Asn Arg Ile Asp Leu
        115                 120                 125

Val Leu Thr Glu Val Ala Met Pro Cys Leu Ser Gly Ile Gly Leu Leu
    130                 135                 140

Ser Lys Ile Thr Ser His Ser Ile Cys Lys Gly Ile Pro Val Ile Met
145                 150                 155                 160

Met Ser Lys Asn Asp Ser Met Ser Thr Val Phe Arg Cys Leu Ser Lys
                165                 170                 175

Gly Ala Val Asp Phe Leu Val Lys Pro Ile Arg Lys Asn Glu Leu Lys
            180                 185                 190

Thr Leu Trp Gln His Ile Trp Arg Arg Cys His Ser Ser Ser Gly Ser
        195                 200                 205

Glu Ser Gly Ile His Thr Gln Lys Cys Ser Lys Pro Lys Ala Gly Asp
    210                 215                 220

Glu Tyr Glu Asn Asn Ser His Asp Asp Asp Asp Asp Cys Gly Ser His
225                 230                 235                 240

Asp Asp Asp Asp Asp Asp Asp Asp Asp Ala Asp Asp Asp Phe Ser Val
                245                 250                 255

Gly Pro Asn Ala Arg Asp Gly Ser Asp Asn Gly Ser Gly Thr
            260                 265                 270
```

<210> 37
<211> 945
<212> DNA
<213> Hordeum vulgare

<400> 37

```
atggtgagcg ccggtcaagc tggcgcggac ggaccttcca gcagtgatat taggggatc      60

ggaaacggcg ccgtagagaa tcagaacggc catgccctca agccaaacga ggacaaggaa     120

tggaggggcg gcagcaagga agaggactgg cccagtacgc acagtgcgcc gccaggcttg     180
```

gacgagcaca agcagcagca gcaagaccgg gtcatccgct gggagaagtt cctgccggtc 240

aagacactaa gggtcttgct agtggagaac gatgactgta cccgacatgt tgtccgtgct 300

ctgctccgta agtgtggcta tgaagttatc tctgctgaga atggattgga tgcatggcaa 360

tatcttgaag atgtgcaaaa ccgtattgac cttgtattaa ctgaggtcgc catgccttgt 420

ctatctggca ttggtctgct cagtaagatc acgagtcaca gtatttgtaa gggcattcct 480

gtgatcatga tgtctaagaa tgactcaatg agtacagtct ttaagtgtct atcaaaggga 540

gcagttgact tcttagtgaa gccgataagg aagaatgaac ttaagaccct ttggcagcac 600

atatggagga gatgccacag ttccagtgga agtgaaagtg gcatccatat acaaaagtgt 660

tccaaaccaa agactggtga tgaatatgcg aaaaacagtg gcggcagtca tgatgacgat 720

gacgatgatg atgctgatga tgacttttagt gttgggccca atgctaggga tggcagtgat 780

aatggcagtg gcactcagag ttcatggacg aagcgtgctg tggagattga tagcccacaa 840

cttgtgtctt ctgaccatct atcagattca cctgatagta cttgtgcaca agtaattcac 900

cccagatcag agataggcag caataggggt gccgactgca aataa 945

<210> 38
<211> 314
<212> PRT
<213> Hordeum vulgare

<400> 38

Met Val Ser Ala Gly Gln Ala Gly Ala Asp Gly Pro Ser Ser Ser Asp
1 5 10 15

Ile Arg Gly Ile Gly Asn Gly Ala Val Glu Asn Gln Asn Gly His Ala
20 25 30

Leu Lys Pro Asn Glu Asp Lys Glu Trp Arg Gly Gly Ser Lys Glu Glu
35 40 45

Asp Trp Pro Ser Thr His Ser Ala Pro Pro Gly Leu Asp Glu His Lys
50 55 60

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

Lys Thr Leu Arg Val Leu Leu Val Glu Asn Asp Asp Cys Thr Arg His
85 90 95

Val Val Arg Ala Leu Leu Arg Lys Cys Gly Tyr Glu Val Ile Ser Ala
100 105 110

Glu Asn Gly Leu Asp Ala Trp Gln Tyr Leu Glu Asp Val Gln Asn Arg
115 120 125

Ile Asp Leu Val Leu Thr Glu Val Ala Met Pro Cys Leu Ser Gly Ile

```
    130                 135                 140

Gly Leu Leu Ser Lys Ile Thr Ser His Ser Ile Cys Lys Gly Ile Pro
145                 150                 155                 160

Val Ile Met Met Ser Lys Asn Asp Ser Met Ser Thr Val Phe Lys Cys
                165                 170                 175

Leu Ser Lys Gly Ala Val Asp Phe Leu Val Lys Pro Ile Arg Lys Asn
            180                 185                 190

Glu Leu Lys Thr Leu Trp Gln His Ile Trp Arg Arg Cys His Ser Ser
        195                 200                 205

Ser Gly Ser Glu Ser Gly Ile His Ile Gln Lys Cys Ser Lys Pro Lys
    210                 215                 220

Thr Gly Asp Glu Tyr Ala Lys Asn Ser Gly Gly Ser His Asp Asp Asp
225                 230                 235                 240

Asp Asp Asp Asp Ala Asp Asp Asp Phe Ser Val Gly Pro Asn Ala Arg
                245                 250                 255

Asp Gly Ser Asp Asn Gly Ser Gly Thr Gln Ser Ser Trp Thr Lys Arg
            260                 265                 270

Ala Val Glu Ile Asp Ser Pro Gln Leu Val Ser Ser Asp His Leu Ser
        275                 280                 285

Asp Ser Pro Asp Ser Thr Cys Ala Gln Val Ile His Pro Arg Ser Glu
    290                 295                 300

Ile Gly Ser Asn Arg Gly Ala Asp Cys Lys
305                 310
```

<210> 39
<211> 750
<212> DNA
<213> Triticum aestivum

<400> 39

```
atggtcatgc cctcgagtca aacgaggaca aggtatggag cggcggcatc aaggaagagg        60

actggcccaa cacgcacagt gcgccagacg ggcttggacg agcagaagca gcagcaagac       120

cgggttatcc ggtgggagaa gttcctgccg gtgaagacac taagggtctt gctggtggag       180

aacgatgact gtacccgaca tgttgtccgt gctctgctcc gtaagtgtgg ctatgaagtt       240

atctctgctg agaatggatt gcatgcatgg caatatcttg aagatgtgca aaaccgtatt       300

gacctggtat taaccgaggt cgccatgcct tgtctatctg gcattggtct gctcagtaag       360

atcacgagtc acagtatttg caagggcatt cctgtgatca tgatgtctaa gaatgactcg       420

atgagtacag tctttaagtg tctatcaaag ggagcagttg acttcttagt gaagccgata       480
```

cggaagaatg aacttaagac cctttggcag cacatatgga ggcgatgcca cagttccagt 540

ggaagtgaaa gtggcatcca tacacaaaaa tgttccaaac caaaggctgg tgatgaatat 600

gagaacaaca gtggcggcag tcatgatgat gatgacgatg acgatgatga tgctgacgac 660

gactttagtg ttgggcccaa tgctagggat ggcagtgata atggcagtgg cactcagagt 720

tcatggacga gcgtgctgtg gagattgata 750

<210> 40
<211> 251
<212> PRT
<213> Triticum aestivum

<400> 40

Met Val Met Pro Ser Ser Gln Thr Arg Thr Arg Tyr Gly Ala Ala Ala
1 5 10 15

Ser Arg Lys Arg Thr Gly Pro Thr Arg Thr Val Arg Gln Thr Gly Leu
20 25 30

Asp Glu Gln Lys Gln Gln Gln Asp Arg Val Ile Arg Trp Glu Lys Phe
35 40 45

Leu Pro Val Lys Thr Leu Arg Val Leu Leu Val Glu Asn Asp Asp Cys
50 55 60

Thr Arg His Val Val Arg Ala Leu Leu Arg Lys Cys Gly Tyr Glu Val
65 70 75 80

Ile Ser Ala Glu Asn Gly Leu His Ala Trp Gln Tyr Leu Glu Asp Val
85 90 95

Gln Asn Arg Ile Asp Leu Val Leu Thr Glu Val Ala Met Pro Cys Leu
100 105 110

Ser Gly Ile Gly Leu Leu Ser Lys Ile Thr Ser His Ser Ile Cys Lys
115 120 125

Gly Ile Pro Val Ile Met Met Ser Lys Asn Asp Ser Met Ser Thr Val
130 135 140

Phe Lys Cys Leu Ser Lys Gly Ala Val Asp Phe Leu Val Lys Pro Ile
145 150 155 160

Arg Lys Asn Glu Leu Lys Thr Leu Trp Gln His Ile Trp Arg Arg Cys
165 170 175

His Ser Ser Ser Gly Ser Glu Ser Gly Ile His Thr Gln Lys Cys Ser
180 185 190

Lys Pro Lys Ala Gly Asp Glu Tyr Glu Asn Asn Ser Gly Gly Ser His

195 200 205

Asp Asp Asp Asp Asp Asp Asp Asp Asp Ala Asp Asp Asp Phe Ser Val
210 215 220

Gly Pro Asn Ala Arg Asp Gly Ser Asp Asn Gly Ser Gly Thr Gln Ser
225 230 235 240

Ser Trp Thr Ser Val Leu Trp Arg Leu Ile Val
245 250

<210> 41
<211> 202
<212> DNA
<213> Brachypodium distachyon

<400> 41
atgccgtgtc tatctggcat cagtctgctc agtaagatca tgagtcacaa aatctgcaag 60

gacattcctg tgattatgat gtcaaagaat gactctatgg gtacagtctt taaatgcttg 120

tcaaagggag ctgttgactt tttagtgaag cctatacgaa agaatgagct taaaaacctt 180

tggcagcaca tatggaggcg at 202

<210> 42
<211> 67
<212> PRT
<213> Brachypodium distachyon

<400> 42

Met Pro Cys Leu Ser Gly Ile Ser Leu Leu Ser Lys Ile Met Ser His
1 5 10 15

Lys Ile Cys Lys Asp Ile Pro Val Ile Met Met Ser Lys Asn Asp Ser
20 25 30

Met Gly Thr Val Phe Lys Cys Leu Ser Lys Gly Ala Val Asp Phe Leu
35 40 45

Val Lys Pro Ile Arg Lys Asn Glu Leu Lys Asn Leu Trp Gln His Ile
50 55 60

Trp Arg Arg
65

<210> 43
<211> 50
<212> PRT
<213> Artificial sequence

<220>
<223> motif 1

<220>
<221> VARIANT

<222> (3)..(3)
<223> Xaa can be Ser or Thr

<220>
<221> VARIANT
<222> (26)..(26)
<223> Xaa can be Leu or Ile

<220>
<221> VARIANT
<222> (38)..(38)
<223> Xaa can be Val or Ile

<220>
<221> VARIANT
<222> (49)..(49)
<223> Xaa can be Gly or Glu

<400> 43

Met Ser Xaa Asn Asp Ser Met Ser Met Val Phe Lys Cys Leu Ser Lys
1 5 10 15

Gly Ala Val Asp Phe Leu Val Lys Pro Xaa Arg Lys Asn Glu Leu Lys
20 25 30

Asn Leu Trp Gln His Xaa Trp Arg Arg Cys His Ser Ser Ser Gly Ser
35 40 45

Xaa Ser
50

<210> 44
<211> 50
<212> PRT
<213> Artificial sequence

<220>
<223> motif 2

<220>
<221> VARIANT
<222> (2)..(2)
<223> Xaa can be Leu or Val

<220>
<221> VARIANT
<222> (3)..(3)
<223> Xaa can be Val or Asp

<220>
<221> VARIANT
<222> (4)..(4)
<223> Xaa can be Glu or Pro

<220>
<221> VARIANT
<222> (5)..(5)
<223> Xaa can be Asn or Pro

<220>
<221> VARIANT
<222> (7)..(7)

<223> Xaa can be Asp or Ser

<220>
<221> VARIANT
<222> (8)..(8)
<223> Xaa can be Ser, Cys or Thr

<220>
<221> VARIANT
<222> (9)..(9)
<223> Xaa can be Thr or Cys

<220>
<221> VARIANT
<222> (10)..(10)
<223> Xaa can be Arg or Ala

<220>
<221> VARIANT
<222> (11)..(11)
<223> Xaa can be Gln or His

<220>
<221> VARIANT
<222> (13)..(13)
<223> Xaa can be Val or Ile

<220>
<221> VARIANT
<222> (14)..(14)
<223> Xaa can be Ser or His

<220>
<221> VARIANT
<222> (15)..(15)
<223> Xaa can be Ala or Pro

<220>
<221> VARIANT
<222> (17)..(17)
<223> Xaa can be Leu or Ser

<220>
<221> VARIANT
<222> (18)..(18)
<223> Xaa can be Arg or Glu

<220>
<221> VARIANT
<222> (19)..(19)
<223> Xaa can be Lys or Ile

<220>
<221> VARIANT
<222> (21)..(21)
<223> Xaa can be Cys or Ser

<220>
<221> VARIANT
<222> (22)..(22)
<223> Xaa can be Tyr or Asn

<220>
<221> VARIANT
<222> (23)..(23)
<223> Xaa can be Glu or Lys

<220>

<221> VARIANT
<222> (24)..(24)
<223> Xaa can be Val or Trp

<220>
<221> VARIANT
<222> (25)..(25)
<223> Xaa can be Ile or Leu

<220>
<221> VARIANT
<222> (27)..(27)
<223> Xaa can be Ala or Thr

<220>
<221> VARIANT
<222> (28)..(28)
<223> Xaa can be Glu or Ala

<220>
<221> VARIANT
<222> (30)..(30)
<223> Xaa can be Gly or Lys

<220>
<221> VARIANT
<222> (31)..(31)
<223> Xaa can be Ser, Leu or Arg

<220>
<221> VARIANT
<222> (32)..(32)
<223> Xaa can be His or Asn

<220>
<221> VARIANT
<222> (34)..(34)
<223> Xaa can be Trp or Lys

<220>
<221> VARIANT
<222> (35)..(35)
<223> Xaa can be Arg, Lys or Gln

<220>
<221> VARIANT
<222> (37)..(37)
<223> Xaa can be Leu or Lys

<220>
<221> VARIANT
<222> (39)..(39)
<223> Xaa can be Asn or Asp

<220>
<221> VARIANT
<222> (40)..(40)
<223> Xaa can be Leu or Lys

<220>
<221> VARIANT
<222> (41)..(41)
<223> Xaa can be Gln or Asp

<220>
<221> VARIANT
<222> (42)..(42)
<223> Xaa can be Asn or Glu

<220>
<221> VARIANT
<222> (43)..(43)
<223> Xaa can be Asn or Ser

<220>
<221> VARIANT
<222> (44)..(44)
<223> Xaa can be Ile or Met

<220>
<221> VARIANT
<222> (45)..(45)
<223> Xaa can be Asp or Gly

<220>
<221> VARIANT
<222> (46)..(46)
<223> Xaa can be Leu or Arg

<220>
<221> VARIANT
<222> (47)..(47)
<223> Xaa can be Val or Tyr

<220>
<221> VARIANT
<222> (50)..(50)
<223> Xaa can be Glu or Ile

<400> 44

```
Leu Xaa Xaa Xaa Xaa Asp Xaa Xaa Xaa Xaa Xaa Val Xaa Xaa Xaa Leu
1               5                   10                  15


Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Pro Xaa Xaa Asn Xaa Xaa Xaa
            20                  25                  30


Ala Xaa Xaa Tyr Xaa Glu Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Leu
        35                  40                  45


Thr Xaa
    50
```

<210> 45
<211> 50
<212> PRT
<213> Artificial sequence

<220>
<223> motif 3

<220>
<221> VARIANT
<222> (9)..(9)
<223> Xaa can be Asp or Gly

<220>
<221> VARIANT
<222> (47)..(47)
<223> Xaa can be Met or Ile

<400> 45

His Asp Asp Glu Glu Asn Asp Asp Xaa Asp Asp Asp Asp Phe Ser Val
1 5 10 15

Gly Leu Asn Ala Arg Asp Gly Ser Asp Asn Gly Ser Gly Thr Gln Ser
20 25 30

Ser Trp Thr Lys Arg Ala Val Glu Ile Asp Ser Pro Gln Pro Xaa Ser
35 40 45

Pro Asp
50

<210> 46
<211> 2194
<212> DNA
<213> Oryza sativa

<400> 46
aatccgaaaa gttttctgcac cgttttcacc ccctaactaa caatataggg aacgtgtgct 60

aaatataaaa tgagaccttа tatatgtagc gctgataact agaactatgc aagaaaaact 120

catccacctа ctttagtggc aatcgggcta aataaaaaag agtcgctaca ctagtttcgt 180

tttccttagt aatttaagtgg gaaaatgaaa tcattattgc ttagaatata cgttcacatc 240

tctgtcatga agttaaatta ttcgaggtag ccataattgt catcaaactc ttcttgaata 300

aaaaaatctt tctagctgaa ctcaatgggt aaagagagag attttttta aaaaaataga 360

atgaagatat tctgaacgta ttggcaaaga tttaaacata taattatata attttatagt 420

ttgtgcattc gtcatatcgc acatcattaa ggacatgtct tactccatcc caatttttat 480

ttagtaatta aagacaattg acttatttt attatttatc ttttttcgat tagatgcaag 540

gtacttacgc acacactttg tgctcatgtg catgtgtgag tgcacctcct caatacacgt 600

tcaactagca acacatctct aatatcactc gcctatttaa tacatttagg tagcaatatc 660

tgaatttcaag cactccacca tcaccagacc acttttaata atatctaaaa tacaaaaaat 720

aattttacag aatagcatga aaagtatgaa acgaactatt taggttttc acatacaaaa 780

aaaaaaagaa ttttgctcgt gcgcgagcgc caatctccca tattgggcac acaggcaaca 840

acagagtggc tgcccacaga acaacccaca aaaaacgatg atctaacgga ggacagcaag 900

tccgcaacaa cctttaaca gcaggctttg cggccaggag agaggaggag aggcaaagaa 960

aaccaagcat cctcctctc ccatctataa attcctcccc cctttcccc tctctatata 1020

ggaggcatcc aagccaagaa gagggagagc accaaggaca cgcgactagc agaagccgag 1080

cgaccgcctt ctcgatccat atcttccggt cgagttcttg gtcgatctct tccctcctcc 1140

acctcctcct cacagggtat gtgcctccct tcggttgttc ttggatttat tgttctaggt 1200

tgtgtagtac gggcgttgat gttaggaaag gggatctgta tctgtgatga ttcctgttct 1260

tggatttggg atagaggggt tcttgatgtt gcatgttatc ggttcggttt gatttagtagt 1320

atggttttca atcgtctgga gagctctatg gaaatgaaat ggtttaggga tcggaatctt 1380

gcgattttgt gagtaccttt tgtttgaggt aaaatcagag caccggtgat tttgcttggt 1440

gtaataaagt acggttgttt ggtcctcgat tctggtagtg atgcttctcg atttgacgaa 1500

gctatccttt gtttattccc tattgaacaa aaataatcca actttgaaga cggtcccgtt 1560

gatgagattg aatgatttgat tcttaagcct gtccaaaatt tcgcagctgg cttgtttaga 1620

tacagtagtc cccatcacga aattcatgga aacagttata atcctcagga acaggggatt 1680

ccctgttctt ccgatttgct ttagtcccag aatttttttt cccaaatatc ttaaaaagtc 1740

actttctggt tcagttcaat gaattgattg ctacaaataa tgcttttata gcgttatcct 1800

agctgtagtt cagttaatag gtaatacccc tatagtttag tcaggagaag aacttatccg 1860

atttctgatc tccatttttа attatatgaa atgaactgta gcataagcag tattcatttg 1920

gattatttt tttattagct ctcacccctt cattattctg agctgaaagt ctggcatgaa 1980

ctgtcctcaa ttttgttttc aaattcacat cgattatcta tgcattatcc tcttgtatct 2040

acctgtagaa gtttctttt ggttattcct tgactgcttg attacagaaa gaaatttatg 2100

aagctgtaat cgggatagtt atactgcttg ttcttatgat tcatttcctt tgtgcagttc 2160

ttggtgtagc ttgccacttt caccagcaaa gttc 2194

<210> 47
<211> 51
<212> DNA
<213> Artificial sequence

<220>
<223> primer: prm15559

<400> 47
ggggacaagt ttgtacaaaa aagcaggctt aaacaatggg tagcgcttgc c 51

<210> 48
<211> 50
<212> DNA
<213> Artificial sequence

<220>
<223> primer: prm15560

<400> 48
ggggaccact ttgtacaaga aagctgggta cgaagatgcc atgttgtatt 50