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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
7 September 2012 (07.09.2012)

WIPO | PCT

(10) International Publication Number
WO 2012/117368 A1

(51) International Patent Classification:
C12N 15/82 (2006.01) *C12N 15/29* (2006.01)
A01H 5/00 (2006.01) *C07K 14/415* (2006.01)

(21) International Application Number:
PCT/IB2012/050969

(22) International Filing Date:
1 March 2012 (01.03.2012)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

61/447,811	1 March 2011 (01.03.2011)	US
61/447,797	1 March 2011 (01.03.2011)	US
11156495.1	1 March 2011 (01.03.2011)	EP
11156500.8	1 March 2011 (01.03.2011)	EP
61/468,616	29 March 2011 (29.03.2011)	US
61/478,975	26 April 2011 (26.04.2011)	US
11163740.1	26 April 2011 (26.04.2011)	EP
11172381.3	1 July 2011 (01.07.2011)	EP

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

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

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **REUZEAU, Christophe** [FR/FR]; Rue de Cimetière, 24350 La Chapelle-Gonaguet (FR). **HATZFELD, Yves** [FR/FR]; 6B, rue du Chemin de Fer, 59000 Lille (FR). **FRANKARD, Valerie** [BE/BE]; Avenue des Croix du Feu 84, 1410 Waterloo (BE).

(74) Agent: **DICK, Alexander**; Herzog Fiesser & Partner, Patentanwälte, Dudenstrasse 46, 68167 Mannheim (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, 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, 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, MD, 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 2012/117368 A1

(54) **Title**: PLANTS HAVING ENHANCED YIELD-RELATED TRAITS AND PRODUCING METHODS THEREOF

(57) **Abstract**: Provided is a method for enhancing yield-related traits in plants by modulating expression in a plant of a nucleic acid encoding a TLP (Tify like protein) polypeptide, a PMP22 polypeptide (22 kDa peroxisomal membrane like polypeptide), a RTF (REM-like transcription factor) polypeptide, or a BPl (Bigger plant 1) polypeptide. Also provided are plants having modulated expression of a nucleic acid encoding a TLP, PMP22, RTF or BP1 polypeptide, which plants have enhanced yield-related traits compared to control plants.

PLANTS HAVING ENHANCED YIELD-RELATED TRAITS AND PRODUCING METHODS THEREOF

The present application claims priority of the following applications: EP 11 156 500.8 filed on March 1, 2011, US 61/447811 filed on March 1, 2011, EP 11 156 495.1 filed on March 1, 2011, US 61/447797 filed on March 1, 2011, US 61/468676 filed on March 29, 2011, EP 11 163 740.1 filed on April 26, 2011, US 61/478975 filed on April 26, 2011, and EP 11 172 381.3 filed July 1, 2011 all of which are herewith incorporated by reference with respect to the entire disclosure content.

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 TLP (Tify like protein) polypeptide, a PMP22 polypeptide (22 kDa peroxisomal membrane like polypeptide), a RTF (REM-like transcription factor) polypeptide, or a BP1 (Bigger plant 1) polypeptide. The present invention also concerns plants having modulated expression of a nucleic acid encoding a TLP, PMP22, RTF or BP1 polypeptide, which plants have enhanced yield-related traits compared to corresponding wild type plants or other control plants. The invention also provides constructs useful in the methods, plants, harvestable parts and products 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.

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 TLP (Tify like protein) polypeptide in a plant.

It further has been found that modulating expression of a nucleic acid encoding a PMP22 polypeptide as defined herein gives plants having enhanced yield-related traits relative to control plants.

Moreover, 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 RTF (REM-like transcription factor) polypeptide in a plant.

Finally, 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 BP1 (Bigger plant 1) polypeptide in a plant.

A. Background

A-1. TLP (Tify like protein) polypeptide

The TIFY family is a novel plant-specific gene family involved in the regulation of diverse plant-specific biologic processes, such as development and responses to phytohormones, in Arabidopsis (Ye et al., Plant Mol Biol. 2009, 71(3)291-305). The function of the TIFY proteins in not fully understood, however it has been proposed that TIFY proteins are transcription factor (see Vanholme et al., Trends Plant Sci. 2007 12(6):239-44), Ye et al. (loc. cit.) reported that there are 20 TIFY genes in rice, the model monocot species. Sequence analysis indicated that rice TIFY proteins have conserved motifs beyond the TIFY domain as was previously shown in Arabidopsis. Most of the OsTIFY genes were predominantly expressed in leaf. Nine OsTIFY genes were responsive to jasmonic acid and wounding treatments. Almost, all the OsTIFY genes were responsive to one or more abiotic stresses including drought, salinity, and low temperature. Moreover, it is also assumed that TIFY proteins might be involved in developmental processes (see Vanholme et al., loc. cit.).

Surprisingly, it has now been found that modulating expression of a nucleic acid encoding a TLP polypeptide as defined herein gives plants having enhanced yield-related traits, in particular increased biomass and/or increased seed yield relative to control plants.

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 TLP polypeptide as defined herein.

A-2. PMP22 polypeptide (22 kDa peroxisomal membrane like polypeptide)

22 kDa Peroxisomal membrane proteins are major components of the peroxisome membranes. In humans, members of in this family are involved in the pore-forming activity and may contribute to the organelle membrane permeability. Mpv17 is a closely related peroxisomal protein involved in the development of early-onset glomerulosclerosis. In *Saccharomyces cerevisiae* (Baker's yeast), a member of this family was identified as an integral membrane protein of the inner mitochondrial membrane and suggested to play a role in mitochondrial function during heat shock. They are targeted to the peroxisome by specific targeting peptides. In plants, direct sorting to peroxisome without ER transit has been shown (Murphy et al. (2003). Plant Physiology 133:813-828. Characterization of the Targeting Signal of the Arabidopsis 22-kD Integral Peroxisomal Membrane Protein).

Peroxisomes play multiple roles at various stages of plant development. For example, they are known to participate in seed germination, leaf senescence, fruit maturation, response to abiotic and biotic stress, photomorphogenesis, biosynthesis of the plant hormones jasmonic acid and auxin, and in cell signaling by reactive oxygen and nitrogen species (ROS and RNS, respectively) (Baker et al. (2010). Biochem Soc Trans. 38(3):807-16. Peroxisome biogenesis and positioning, Del Rio (2010). Arch Biochem Biophys. Nov 3 (epub ahead of print) Peroxisomes as a cellular source of reactive nitrogen species signal molecules.). It becomes apparent that a peroxisome can be a source and sensor of molecules that can affect plant growth and development. Furthermore, biochemical and molecular studies have shown that multiple essential metabolic functions take place in the peroxisomes (e.g. Eubel et al. (2008). Plant Physiology 148:1809-1829.Novel Proteins, Putative Membrane Transporters, and an Integrated Metabolic Network Are Revealed by Quantitative Proteomic Analysis of Arabidopsis Cell Culture Peroxisomes; Reumann et al. (2009). Plant Physiology 150:125-143.In-Depth Proteome Analysis of Arabidopsis Leaf Peroxisomes Combined with in Vivo Subcellular Targeting Verification Indicates Novel Metabolic and Regulatory Functions of Peroxisomes.). These processes rely on several transport systems supporting the flux of metabolites in and out of the peroxisomes (for review see Visset et al. (2007). Biochem J. 401(2):365-75.Metabolite transport across the peroxisomal membrane.).

Peroxisomes play an important role in plant productivity as it is intimately involved in photorespiratory pathway, along with the chloroplast and mitochondria (for review see Peterhansel and Maurino (2011). Plant Physiol. 155:49-55.Photorespiration redesigned.). Interestingly, Arabidopsis plants with low photorespiration showed enhanced photosynthesis and growth. This effect was mainly driven by high catalase activity, but unfortunately it could not be stabilized over generations.

Surprisingly, it has now been found that modulating expression of a nucleic acid encoding a PMP22 polypeptide as defined herein gives plants having enhanced yield-related traits, in particular increased yield relative to control plants.

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 PMP22 polypeptide as defined herein.

A-3. RTF (REM-like transcription factor) polypeptide

The B3 DNA binding domain is a conserved domain which is only found in transcription factors from higher plants. A B3 binding domain usually consists of 100-120 residues, and includes seven beta strands and two alpha helices which form a DNA-binding pseudobarrel protein fold which is thought to interact with the major groove of the DNA. Five different protein families were shown to comprise B3 domains: auxin response factor (ARF), abscisic acid-insensitive3 (ABI3), high level expression of sugar inducible (HSI), related to ABI3/VP1 (RAV) and reproductive meristem (REM). Among the B3 families, ABI3, HSI, RAV and ARF are most structurally conserved, whereas the REM family has experienced a rapid divergence. This explains the variety of sequences found in the REM subgroup and the variety observed between plant species (Romanel EA, Schrago CG, Couñago RM, Russo CA, Alves-Ferreira M. (2009). Evolution of the B3 DNA binding superfamily: new insights into REM family gene diversification. PLoS One. 2009 Jun 8;4(6):e5791).

Swaminathan et al. (Swaminathan K, Peterson K, Jack T. 2008. The plant B3 superfamily. Trends Plant Sci. 2008 Dec;13(12):647-55) provides an overview on REM genes in Arabidopsis. According to Swaminathan, there are a total of 76 REM genes in Arabidopsis which can be divided into six subgroups, subgroups A to F.

REM10 (At2G24700) belongs to subgroup C according to the classification of Swaminathan. Subgroup C consists of 18 members (REM1 to REM18) of which REM1 to REM14 are clustered in the Arabidopsis genome. REM10 to REM14 are tightly linked on the a 35kb region of chromosome 2.

The members of the subgroup have the unusual feature that they comprise more than one B3 domain (except for REM18). For example, REM10 comprises four B3 domains. Not much information is available on the function of REM genes belonging to subgroup C. According to Swaminathan, no loss of function mutants have been reported.

Surprisingly, it has now been found that modulating expression of a nucleic acid encoding a RTF polypeptide as defined herein gives plants having enhanced yield-related traits, in particular increased yield, in particular relative to control plants.

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 RTF polypeptide as defined herein.

A-4. BP1 (Bigger plant 1) polypeptide

Os09g25410 is a protein expressed in rice. It shows homology to gene that is upregulated after anthesis in wheat (Genbank Accession Number CA611178, Ruuska SA, Lewis DC, Kennedy G, Furbank RT, Jenkins CL, Tabe LM, Large scale transcriptome analysis of the effects of nitrogen nutrition on accumulation of stem carbohydrate reserves in reproductive stage wheat. Plant Mol. Biol. 66: 15-32 (2008). In the prior art, no information regarding the function of this protein is available.

Surprisingly, it has now been found that modulating expression of a nucleic acid encoding a BP1 polypeptide as defined herein gives plants having enhanced yield-related traits, in particular increased yield relative to control plants.

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 BP1 polypeptide as defined herein.

The section captions and headings in this specification are for convenience and reference purpose only and should not affect in any way the meaning or interpretation of this specification.

B. Definitions

The following definitions will be used throughout the present specification.

Polypeptide(s)/Protein(s)

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

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.

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; insertions will usually be of the order of about 1 to 10 amino acid residues. 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 well 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.

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).

Orthologue(s)/Paralogue(s)

Orthologues and paralogues 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.

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) & The Pfam protein families database: R.D. Finn, J. Mistry, J. Tate, P. Coggill, A. Heger, J.E. Pollington, O.L. Gavin, P. Gunesekaran, G. Ceric, K. Forslund, L. Holm, E.L. Sonnhammer, S.R. Eddy, A. Bateman Nucleic Acids Research (2010) Database Issue 38:211-222). 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 A1 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):

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

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. If high strigency hybridization conditions are applied, the hybridization may be also followed by washing at 65°C in 0.1x 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. Preferably, the solution used for hybridization and washing also comprises 0.1% SDS. 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

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 pro-

moter 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 ami-notransferase	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 ltr1 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 pyro-phosphorylase	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.

In one embodiment of the invention an “isolated” nucleic acid sequence is located in a non-native chromosomal surrounding.

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” or the term “modulating expression” 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.

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.

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).

Throughout this application a plant, plant part, seed or plant cell transformed with - or interchangeably transformed by - a construct or transformed with or by a nucleic acid is to be understood as meaning a plant, plant part, seed or plant cell that carries said construct or said nucleic acid as a transgene due the result of an introduction of said construct or said nucleic acid by biotechnological means. The plant, plant part, seed or plant cell therefore comprises said recombinant construct or said recombinant nucleic acid. Any plant, plant part, seed or plant cell that no longer contains said recombinant construct or said recombinant nucleic acid after introduction in the past, is termed null-segregant, nullizygote or null control, but is not considered a plant, plant part, seed or plant cell transformed with said construct or with said nucleic acid within the meaning of this application.

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 Traits

Yield related traits are traits or features which are 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, increased growth rate, improved agronomic traits, such as e.g. increased tolerance to submergence (which leads to increased yield in rice), improved Water Use Efficiency (WUE), improved Nitrogen Use Efficiency (NUE), etc.

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 dry mature seed up to the stage where the plant has produced dry 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. "Mild stresses" are the everyday biotic and/or abiotic (environmental) stresses to which a plant is exposed. 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 partly inserted in or in physical contact with the 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.

In a preferred embodiment throughout this application any reference to "root" as biomass or harvestable parts or as organ of increased sugar content is to be understood as a reference to harvestable parts partly inserted in or in physical contact with the ground such as but not limited to beets and other hypocotyl areas of a plant, rhizomes, stolons or creeping rootstalks, but not including leaves, as well as harvestable parts belowground, such as but not limited to root, taproot, tubers or bulbs.

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, Erian-*

thus 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.

With respect to the sequences of the invention, a nucleic acid or a polypeptide sequence of plant origin has the characteristic of a codon usage optimised for expression in plants, and of the use of amino acids and regulatory sites common in plants, respectively. The plant of origin may be any plant, but preferably those plants as described in the previous paragraph.

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 (also called null control plants) are individuals missing the transgene by segregation. Further, a control plant has been grown under equal growing conditions to the growing conditions of the plants of the invention. Typically the control plant is grown under equal growing conditions and hence in the vicinity of the plants of the invention and at the same time. A “control plant” as used herein refers not only to whole plants, but also to plant parts, including seeds and seed parts.

C. Detailed description of the invention

C-1. TLP (Tify like protein) polypeptide

Surprisingly, it has now been found that modulating expression in a plant of a nucleic acid encoding a TLP 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 TLP 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 enhancing 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 TLP 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 TLP polypeptide is by introducing and expressing in a plant a nucleic acid encoding a TLP polypeptide.

Any reference hereinafter in section C-1 to a "protein useful in the methods of the invention" is taken to mean a TLP 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 TLP polypeptide. In one embodiment any reference to a protein or nucleic acid "useful in the methods of the invention" is to be understood to mean proteins or nucleic acids "useful in the methods, constructs, plants, harvestable parts and products of the invention". 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 "*TLP* nucleic acid" or "*TLP* gene".

A "TLP polypeptide" as defined herein refers, preferably, to any polypeptide comprising a Pfam domain having the Pfam accession number PF06200 (TIFY), or a Pfam domain having the accessing number PF09425 (CCT_2). More preferably, it refers to any polypeptide comprising a Pfam domain having the Pfam accession number PF06200 (TIFY) and a Pfam domain having the accessing number PF09425 (CCT_2).

Preferably, the PF06200 Pfam domain and PF09425 Pfam domain are separated in an increasing order of preference by at least 10, at least 25, at least 50, at least 75, at least 100 amino acids.

Preferably, the PF06200 Pfam domain is located in the central part of the protein. Preferably, the PF09425 Pfam domain is located in the C-terminal part of the polypeptide.

Preferably, the Pfam domain having the Pfam accession number PF06200 (also referred to as "PF06200 pfam domain" herein) comprises a sequence having 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%, 99% or 100% sequence identity to the conserved domain starting with amino acid 144 up to amino acid 178 in SEQ ID NO:2. Preferably, the Pfam domain having the Pfam accession number PF09425 (also referred to as "PF09425 pfam domain" herein) comprises a sequence having 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%, 99% or 100% sequence identity to the conserved domain starting with amino acid 282 up to amino acid 306 in SEQ ID NO:2.

Additionally or alternatively, the "TLP polypeptide" as defined herein refers, preferably, to any polypeptide comprising an Interpro domain IPR010399 (TIFY), or an Interpro domain IPR018467 (CO, COL, TOC1). More preferably, it refers to any polypeptide comprising an Interpro domain IPR010399 (TIFY) and an Interpro domain IPR018467 (CO, COL, TOC1).

The Interpro domains as referred to herein are, preferably, based on the InterPro database, Release 31.0 (9th February 2011).

The Pfam domains as referred to herein are, preferably, based on the Pfam database, Release 24.0 (Pfam 24.0, October 2009), see also The Pfam protein families database: R.D. Finn, J. Mistry, J. Tate, P. Coggill, A. Heger, J.E. Pollington, O.L. Gavin, P. Gunesekaran, G. Ceric, K. Forslund, L. Holm, E.L. Sonnhammer, S.R. Eddy, A. Bateman Nucleic Acids Research (2010) Database Issue 38:D211-222.

Preferably, the TLP polypeptide additionally or alternatively comprises one or more of the following motifs (see also Fig. 1):

Motif 1-1 (SEQ ID NO: 35): QLTIFY[AG]G[SM]V[NC]V[YF][DE][DN][IV]S[PA]EKAQ[AE][IL]M

Motif 2-1 (SEQ ID NO: 37): PQARKASLARFLEKRKERV[MT][NST][TAL][AS]PY

Motif 3-1 (SEQ ID NO: 39):
MERDF[LM]GL[NGSI][IS]K[DEN][PS][LP][LA][VT][VI]K[DE]Exxx[SD][SG], wherein "X", preferably, represents any amino acid

Motif 4-1 (SEQ ID NO: 40):

Q[LM]TIFY[AG]G[SMATL]V[NCS][VI][YF][DEN][DN][IV][STP][PAV][ED][KQ]A[QK][AE][IL]MFLA[GS][HNR]
Motif 5-1 (SEQ ID NO: 43):RFLEKRKE
Motif 6-1 (SEQ ID NO: 44): QLTIFY[AG]G
Motif 7-1 (SEQ ID NO: 45):MERDF[LM]GL

Instead of Motif 1-1, the TLP polypeptide may, preferably, comprise Motif 1-1a):
QLTIFYGGMV[NC]V[YF]E[DN][IV]S[PA]EKAQ[AE][IL]M (SEQ ID NO: 36)

Instead of Motif 2-1, the TLP polypeptide may, preferably, comprise Motif 2-1a):
PQARKASLARFLEKRKERV[MT][NST]L[AS]PY (SEQ ID NO: 38)

Instead of Motif 4-1, the TLP polypeptide may, preferably, comprise Motif 4-1a):
Q[LM]TIFY[AG]G[SMATL]V[NCS][VI][YF][DEN][DN][IV][STP][PAV][ED] (SEQ ID NO: 41),
and/or Motif 4-1b):
[KQ]A[QK][AE][IL]MFLA[GS][HNR] (SEQ ID NO: 42), preferably both, i. e. Motif 4-1a and 4b. Preferably, the order is Motif 4-1a and then Motif 4-1b). Preferably, Motifs 4-1a) and 4-1b) are separated in an increasing order of preference, by 20, 19, 18, 17, 16, 15, or 14 amino acids.

Preferably, Motif 1-1 (and/or Motif 1-1a and/or 4-1, respectively) is comprised by the PF06200 Pfam domain and/or IPR010399 domain. Preferably, Motif 2-1 (and/or Motif 2-1a and/or 5-1, respectively) is comprised by the PF09425 Pfam domain and/or IPR018467 domain.

The term “TLP” or “TLP polypeptide” as used herein also intends to include homologues as defined hereunder of “TLP polypeptide”.

Motifs 1-1 to 7-1 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.

More preferably, the TLP polypeptide comprises in increasing order of preference, at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, or all 7 motifs.

The following combinations of motifs are particularly preferred: Motif 1-1 and Motif 2-1; Motif 1-1 and Motif 3-1; Motif 2-1 and Motif 3-1; Motif 1-1, 2-1 and 3-1; Motif 1-1 and Motif 7-1; Motif 2-1 and Motif 7-1; Motif 1-1, 2-1 and 7-1; Motif 4-1 and Motif 2-1; Motif 4-1 and Motif 3-1; Motif 4-1, 2-1 and 3-1; Motif 4-1 and Motif 7-1; Motif 4-1, 2-1 and 7-1; Motif 1-1 and Motif 5-1; Motif 5-1 and Motif 3-1; Motif 1-1, 5-1 and 3-1. In the aforementioned list, Motif 1-1 may be replaced by

Motif 1-1a), Motif 2-1 by Motif 2-1a), and Motif 4-1 by (Motif 4-1a) and/or Motif 4-1b)), see above.

Thus, the TLP polypeptide, preferably, may comprise

a. all of the following motifs:
(i) Motif 1-1: (SEQ ID NO: 35):
QLTIFY[AG]G[SM]V[NC]V[YF][DE][DN][IV]S[PA]EKAQ[AE][IL]M,
(ii) Motif 2-1: (SEQ ID NO: 37):
PQARKASLARFLEKRKERV[MT][NST][TAL][AS]PY,
(iii) Motif 3-1: (SEQ ID NO: 39):
MERDF[LM]GL[NGSI][IS]K[DEN][PS][LP][LA][VT][VI]K[DE]Exxx[SD][SG],
(iv) Motif 4-1 (SEQ ID NO: 40)
Q[LM]TIFY[AG]G[SMATL]V[NCS][VI][YF][DEN][DN][IV][STP][PAV][ED][KQ]A[QK][AE][IL]MFLA[GS][HNR],
(v) Motif 5-1 (SEQ ID NO: 43):RFLEKRKE
(vi) Motif 6-1 (SEQ ID NO: 44): QLTIFY[AG]G
(vii) Motif 7-1 (SEQ ID NO: 45):MERDF[LM]GL;
b. or all of the motifs 2-1 to 7-1 as defined in a. above, and in addition the Motif 1-1a) (SEQ ID NO: 36):
QLTIFYGGMV[NC]V[YF]E[DN][IV]S[PA]EKAQ[AE][IL]M ; or
c. all of the motifs 1-1 and 3-1 to 7-1 as defined in a. above, and in addition the Motif 2-1a) (SEQ ID NO: 38)
PQARKASLARFLEKRKERV[MT][NST]L[AS]PY ; or
d. all of the motifs 1-1 to 7-1 as defined in a. above, wherein motif 4-1 is replaced by the Motif 4-1a) (SEQ ID NO: 41)
Q[LM]TIFY[AG]G[SMATL]V[NCS][VI][YF][DEN][DN][IV][STP][PAV][ED],
and/or Motif 4b) (SEQ ID NO: 42):
[KQ]A[QK][AE][IL]MFLA[GS][HNR] ; or
e. all of the motifs 1-1a), 2-1a), 3-1, 4-1a) and 4-1b), 5-1 to 7-1 as defined in a. to d above; or
f. any three, preferably any four , more preferably any 5 motifs as defined in a. to d. above; or
g. any combination of motifs as defined in f. wherein Motifs 1-1, 2-1 and 4-1 are not present; or
h. any motif as defined in a. to d. above.

Additionally or alternatively, the TLP protein, or the homologue of a TLP protein, preferably, 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 represented by SEQ ID NO: 2. Preferably, said TLP protein or said homologous protein comprises any one or more of the conserved motifs or domains, preferably 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 of a nucleic acid sequence is determined by comparison of the nucleic acid sequence over the entire length of the coding sequence of the sequence of SEQ ID NO: 1.

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 TLP 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 1-1 to 7-1 as defined herein above (including Motifs 1a, 2a, 4a and 4b).

In a preferred embodiment a method is provided wherein said TLP polypeptide comprises a conserved domain 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 144 up to amino acid 178 in SEQ ID NO:2 and/or (preferably and) a conserved domain 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 282 up to amino acid 306 in SEQ ID NO:2.

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

In a preferred embodiment, the TLP polypeptide is selected from the group consisting of:

a) a polypeptide comprising a sequence, or consisting of a sequence as shown in SEQ ID NO: 2,

b) a polypeptide having, in increasing order of preference at least 50%, at least 60%, at least 70%, at least 80%, or at least 90% sequence identity to a polypeptide as represented by SEQ ID NO: 2,

c) a polypeptide encoded by a polynucleotide which hybridizes under stringent conditions to a polynucleotide having a sequence as shown in SEQ ID NO: 1, or with

a complementary sequence of such a polynucleotide having a sequence as shown in SEQ ID NO: 1,

d) a polypeptide with the biological activity of the polypeptide as shown in SEQ ID NO: 2 or substantially the same biological activity of the polypeptide as shown in SEQ ID NO: 2; and

e) any combination of a.) to d) above.

Preferable, the TLP polypeptide comprises the domain and/or motifs as set forth herein above.

Preferably, the TLP polypeptide sequence which when used in the construction of a phylogenetic tree, such as the one depicted in Figure 3, clusters within the sequences not more than 4, 3, or 2 hierarchical branch points away from the amino acid sequence represented by SEQ ID NO: 2 rather than with any other group.

Preferably, TLP polypeptides, when expressed in rice, increases yield-related traits according to the methods of the present invention as outlined in Example XI-1.

Accordingly, TLP polypeptides (at least in their native form) when expressed in a plant, in particular in a monocot plant such as rice, maize, wheat or sugarcane, preferably, increase at least one of the yield related traits selected from the group consisting of aboveground biomass, total seed yield, number of filled seeds, number of flowers per panicle, thousand kernel weight, seedling biomass, and plant height (as compared to a control plant). Preferably, said increase is an increase of at least 1%, of at least 2%, more preferably, of at least 3% and, most preferably, of at least 5%. Tools and techniques for measuring whether the yield related traits are increased are described in the Examples.

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 TLP-encoding nucleic acid or TLP polypeptide as defined herein.

Examples of nucleic acids encoding TLP polypeptides are given in Table A1 of the Examples section herein. Such nucleic acids are useful in performing the methods of the invention. The amino acid sequences given in Table A1 of the Examples section are example sequences of orthologues and paralogues of the TLP 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 tomato sequences.

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 A1 of the Examples section, the terms "homologue" and "derivative" being as defined herein. Also useful in the methods, constructs, plants, harvestable parts and products 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 A1 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 TLP polypeptides, nucleic acids hybridising to nucleic acids encoding TLP polypeptides, splice variants of nucleic acids encoding TLP polypeptides, allelic variants of nucleic acids encoding TLP polypeptides and variants of nucleic acids encoding TLP polypeptides obtained by gene shuffling. The terms hybridising sequence, splice variant, allelic variant and gene shuffling are as described herein.

In one embodiment of the present invention the function of the nucleic acid sequences of the invention is to confer information for a protein that increases yield or yield related traits, when a nucleic acid sequence of the invention is transcribed and translated in a living plant cell.

Nucleic acids encoding TLP polypeptides need not be full-length nucleic acids, since performance of the methods of the invention does not 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 A1 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 A1 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, constructs, plants, harvestable parts and products of the invention, encode a TLP polypeptide as defined herein, and have substantially the same biological activity as the amino acid sequences given in Table A1 of the Examples section. Preferably, the portion is a portion of any one of the nucleic acids given in Table A1 of the Examples section, or is a portion of a nucleic acid encoding an orthologue or paralogue of any one of the ami-

no acid sequences given in Table A1 of the Examples section. Preferably the portion is at least 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150 or 1190 consecutive nucleotides in length, the consecutive nucleotides being of any one of the nucleic acid sequences given in Table A1 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 A1 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 i) at least one motif from Motif 1-1 to 7-1 as specified elsewhere herein; and/or ii) a PF06200 Pfam domain and/or PF09425 Pfam domain; and/or iii) an Interpro domain IPR010399 and/or an Interpro domain IPR018467; and/or iii) has, in increasing order of preference, at least 70, 80, 90, or 95% sequence identity to SEQ ID NO: 2.

Another nucleic acid variant useful in the methods, constructs, plants, harvestable parts and products of the invention is a nucleic acid capable of hybridising, under reduced stringency conditions, preferably under stringent conditions, with a nucleic acid encoding a TLP 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 any one of the nucleic acids given in Table A1 of the Examples section, or comprising introducing and expressing in a plant a nucleic acid capable of hybridising to a nucleic acid encoding an orthologue, paralogue or homologue of any of the nucleic acid sequences given in Table A1 of the Examples section.

Hybridising sequences useful in the methods, constructs, plants, harvestable parts and products of the invention encode a TLP polypeptide as defined herein, having substantially the same biological activity as the amino acid sequences given in Table A1 of the Examples section. Preferably, the hybridising sequence is capable of hybridising to the complement of any one of the nucleic acids given in Table A1 of the Examples section, or to a portion of any of these sequences, a portion being as defined above, 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 A1 of the Examples section. Most preferably, the hybridising sequence is capable of hybridising to the complement of a nucleic acid as represented by SEQ ID NO: 1 or to a portion thereof.

Preferably, the hybridising sequence encodes a polypeptide with an amino acid sequence which comprises i) at least one motif from Motif 1-1 to 7-1 as specified elsewhere herein; and/or ii) a PF06200 Pfam domain and/or PF09425 Pfam domain; and/or iii) an Interpro domain IPR010399 and/or an Interpro domain IPR018467; and/or iii) has, in increasing order of preference, at least 70, 80, 90, or 95% sequence identity to SEQ ID NO: 2.

In one embodiment the hybridising sequence is capable of hybridising to the complement of a nucleic acid as represented by SEQ ID NO: 1 or to a portion thereof under conditions of medium or high stringency, preferably high stringency as defined above. In another embodiment the hybridising sequence is capable of hybridising to the complement of a nucleic acid as represented by SEQ ID NO: 1 under stringent conditions.

Another nucleic acid variant useful in the methods, constructs, plants, harvestable parts and products of the invention is a splice variant encoding a TLP polypeptide as defined hereinabove, a splice variant being 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 splice variant of any one of the nucleic acid sequences given in Table A1 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 A1 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 i) at least one motif from Motif 1-1 to 7-1 as specified elsewhere herein; and/or ii) a PF06200 Pfam domain and/or PF09425 Pfam domain; and/or iii) an Interpro domain IPR010399 and/or an Interpro domain IPR018467; and/or iii) has, in increasing order of preference, at least 70, 80, 90, or 95% sequence identity to SEQ ID NO: 2.

Another nucleic acid variant useful in performing the methods of the invention is an allelic variant of a nucleic acid encoding a TLP polypeptide as defined hereinabove, an allelic variant being 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 an allelic variant of any one of the nucleic acids given in Table A1 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 A1 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 TLP polypeptide of SEQ ID NO: 2 and any of the amino acids depicted in Table A1 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 i) at least one motif from Motif 1-1 to 7-1 as

specified elsewhere herein; and/or ii) a PF06200 Pfam domain and/or PF09425 Pfam domain; and/or iii) an Interpro domain IPR010399 and/or an Interpro domain IPR018467; and/or iii) has, in increasing order of preference, at least 70, 80, 90, or 95% sequence identity to SEQ ID NO: 2.

Gene shuffling or directed evolution may also be used to generate variants of nucleic acids encoding TLP polypeptides as defined above; the term "gene shuffling" being 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 variant of any one of the nucleic acid sequences given in Table A1 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 A1 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 i) at least one motif from Motif 1-1 to 7-1 as specified elsewhere herein; and/or ii) a PF06200 Pfam domain and/or PF09425 Pfam domain; and/or iii) an Interpro domain IPR010399 and/or an Interpro domain IPR018467; and/or iii) has, in increasing order of preference, at least 70, 80, 90, or 95% 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.).

Nucleic acids encoding TLP 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 TLP polypeptide-encoding nucleic acid is from a plant, further preferably from a dicotyledonous plant, more preferably from the family Solanaceae, even more preferably the nucleic acid is from the genus Solanum most preferably the nucleic acid is from Solanum lycopersicum (frequently, also referred to as Lycopersicum esculentum). .

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, i.e. said nucleic acid is not in the chromosomal DNA in its native surrounding. Said recombinant chromosomal DNA may be a chromosome of native origin, with said nucleic acid inserted by recombinant means, or it may be a mini-chromosome or a non-native chromosomal structure, e.g. or an artificial chromosome. The nature of the chromosomal DNA may vary, as long it allows for stable passing on to successive generations of the recombinant nucleic acid useful in the methods, con-

structs, plants, harvestable parts and products of the invention, and allows for expression of said nucleic acid in a living plant cell resulting in increased yield or increased yield related traits of the plant cell or a plant comprising the plant cell.

In a further embodiment the recombinant chromosomal DNA of the invention is comprised in a plant cell. DNA comprised within a cell, particularly a cell with cell walls like a plant cell, is better protected from degradation than a bare nucleic acid sequence. The same holds true for a DNA construct comprised in a host cell, for example 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 seed yield relative to control plants. The terms “yield” and “seed yield” are described in more detail in the “definitions” section herein.

Reference herein to enhanced yield-related traits is taken to mean an increase 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 roots such as taproots, stems, beets, tubers, leaves, flowers or seeds , and performance of the methods of the invention results in plants having increased seed yield relative to the seed yield of control plants, and/or increased aboveground biomass, in particular stem biomass relative to the aboveground biomass, and in particular stem biomass of control plants, and/or increased root biomass relative to the root biomass of control plants and/or increased beet biomass relative to the beet biomass of control plants. Moreover, it is particularly contemplated that the sugar content (in particular the sucrose content) in the above ground parts, particularly stem (in particular of sugar cane plants) and/or in the be-lowground parts, in particular in roots including taproots and tubers, and/or in beets (in particular in sugar beets) is increased relative to the sugar content (in particular the sucrose con-tent) in corresponding part(s) of the control plant.

The present invention provides a method for increasing yield related traits, in particular aboveground biomass, total seed yield, number of filled seeds, number of flowers per panicle, thousand kernel weight, seedling biomass, and plant height relative to control plants, which method comprises modulating expression in a plant of a nucleic acid encoding a TLP 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 TLP 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 relative to control plants grown under comparable conditions. Therefore, according to the present invention, there is provided a method for increasing yield 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 TLP polypeptide.

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

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

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

The invention also provides genetic constructs and vectors to facilitate introduction and/or expression in plants of nucleic acids encoding TLP polypeptides. The gene constructs may be inserted into vectors, which may be commercially available, suitable for transforming into plants 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 TLP 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 TLP polypeptide is as defined above. The term “control sequence” and “termination sequence” are as defined herein.

The invention furthermore provides plants 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.

Plants are transformed with a vector comprising any of the nucleic acids described above. The skilled artisan is well aware of the genetic elements that must be present on the vector 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) in the vectors of the invention.

In one embodiment the plants of the invention are transformed with an expression cassette comprising any of the nucleic acids described above. The skilled artisan is well aware of the genetic elements that must be present on the expression cassette in order to successfully transform, select and propagate host cells containing the sequence of interest. In the expression cassettes of the invention the sequence of interest is operably linked to one or more control sequences (at least to a promoter). The promoter in such an expression cassette may be a non-native promoter to the nucleic acid described above, i.e. a promoter not regulating the expression of said nucleic acid in its native surrounding.
In one embodiment the terms expression cassettes of the invention, genetic construct and constructs of the invention are used exchangeably.
In a further embodiment the expression cassettes of the invention confer increased yield or yield related traits(s) to a living plant cell when they have been introduced into said plant cell and result in expression of the nucleic acid as defined above, comprised in the expression cassette(s).
The promoter in such expression cassettes may be a non-native promoter to the nucleic acid described above, i.e. a promoter not regulating the expression of said nucleic acid in its native surrounding.
The expression cassettes of the invention may be comprised in a host cell, plant cell, seed, agricultural product or plant.

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. Preferably the constitutive promoter is a ubiquitous constitutive promoter of medium strength. See the “Definitions” section herein for definitions of the various promoter types. Also useful in the methods of the invention is a root-specific promoter (e.g. when sugar beets are transformed).

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

The constitutive promoter is preferably a medium strength promoter. More preferably it is a plant derived promoter, 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.

According to another preferred feature of the invention, the nucleic acid encoding a TLP polypeptide is operably linked to a root-specific promoter. The root-specific promoter is preferably an RCc3 promoter (Plant Mol Biol. 1995 Jan;27(2):237-48).

In another preferred embodiment, the polynucleotide encoding the TLP polypeptide as used in the plants, constructs and methods of the present invention is linked to a promoter which allows for the expression, preferably the strongest expression in the aboveground parts of the plant as compared to the expression in other parts of the plant. This applies, in particular, if the plant is a monocot. As set forth elsewhere herein, preferred monocots are maize, wheat, rice, or sugarcane. In another preferred embodiment of the present invention, the polynucleotide encoding the TLP polypeptide as used in the plants, constructs and methods of the present invention is preferably linked to a promoter which allows for the expression, preferably the strongest expression in the belowground parts or beets of the plant as compared to the expression in other parts of the plant. This applies, in particular, if the plant is a dicot. Preferred dicots are sugar beet and potato. For example, if the plant is a sugar beet, the promoter, preferably, allows for the strongest expression in the taproot or beet as compared to the expression in other parts of the plant. In one embodiment the promoter used in for expression in sugar beets is, preferably a root specific, more preferably a taproot or beet specific promoter.
Optionally, one or more terminator sequences may be used in the construct introduced into a plant. 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 TLP polypeptide. More preferably, the construct comprises a zein terminator (t-zein) linked to the 3' end of the TLP coding sequence. Most preferably, the expression cassette comprises a sequence having in increasing order of preference at least 95%, at least 96%, at least 97%, at least 98%, at least 99% identity to the sequence represented by the pPRO::TLP::t-zein sequence (Fig. 5) comprised by the expression vector having the sequence as shown in SEQ ID NO: 47 (pPRO::TLP::t-zein 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 TLP polypeptide is by introducing and expressing in a plant a nucleic acid encoding a TLP 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, in particular aboveground biomass, total seed yield, number of filled seeds, number of flowers per panicle, thousand kernel weight, seedling biomass, and/or plant height, relative to control plants, comprising introduction and expression in a plant of any nucleic acid encoding a TLP polypeptide as defined hereinabove.

More specifically, the present invention provides a method for the production of transgenic plants having enhanced yield-related traits, particularly increased seed and biomass yield, more preferably aboveground biomass, total seed yield, number of filled seeds, number of flowers per panicle, thousand kernel weight, seedling biomass, and/or plant height,
which method comprises:
(i) introducing and expressing in a plant or plant cell a TLP polypeptide-encoding nucleic acid or a genetic construct comprising a TLP polypeptide-encoding nucleic acid; and
(ii) cultivating the plant cell under conditions promoting plant growth and development.

Cultivating the plant cell under conditions promoting plant growth and development, may or may not include regeneration and or growth to maturity.

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

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 by transformation. The term “transformation” is described in more detail in the “definitions” section herein.

In one embodiment the present invention clearly 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 parts thereof comprise a nucleic acid transgene encoding a TLP polypeptide as defined above. 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.

The present invention also extends in another embodiment to transgenic plant cells and seed comprising the nucleic acid molecule of the invention in a plant expression cassette or a plant expression construct.

In a further embodiment the seed of the invention recombinantly comprise the expression cassettes of the invention, the (expression) constructs of the invention, the nucleic acids described above and/or the proteins encoded by the nucleic acids as described above.

A further embodiment of the present invention extends to plant cells comprising the nucleic acid as described above in a recombinant plant expression cassette.

In yet another embodiment the plant cells of the invention are non-propagative cells, e.g. the cells can not be used to regenerate a whole plant from this cell as a whole using standard cell culture techniques, this meaning cell culture methods but excluding in-vitro nuclear, organelle or chromosome transfer methods. 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. One example are plant cells that do not sustain themselves through photosynthesis by synthesizing carbohydrate and protein from such inorganic substances as water, carbon dioxide and mineral salt.

In another embodiment the plant cells of the invention are plant cells that do not sustain themselves through photosynthesis by synthesizing carbohydrate and protein from such inorganic substances as water, carbon dioxide and mineral salt, i.e. they may be deemed non-plant variety. In a further embodiment the plant cells of the invention are non-plant variety and non-propagative.

The invention also includes host cells containing an isolated nucleic acid encoding a TLP polypeptide as defined hereinabove. Host cells of the invention may be any cell selected from the group consisting of bacterial cells, such as E.coli or Agrobacterium species cells, yeast cells, fungal, algal or cyanobacterial cells or plant cells. In one embodiment host cells according to the invention are plant cells, yeasts, bacteria or fungi. Host plants for the nucleic acids or the vector used in the method according to the invention, the expression cassette or construct or vector are, in principle, advantageously all plants, which are capable of synthesizing the polypeptides used in the inventive method.

In one embodiment the plant cells of the invention overexpress the nucleic acid molecule of the

invention.

The invention also includes methods for the production of 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, including seeds, of these plants. In a further embodiment the methods comprises steps a) growing the plants of the invention, b) removing the harvestable parts as defined above from the plants and c) producing said product from or by the harvestable parts of the invention.
Examples of such methods would be growing corn plants of the invention, harvesting the corn cobs and remove the kernels. These may be used as feedstuff or processed to starch and oil as agricultural products.
The product may be produced at the site where the plant has been grown, or the plants or parts thereof may be removed from the site where the plants have been grown to produce the product. Typically, the plant is grown, the desired harvestable parts are removed from the plant, if feasible in repeated cycles, and the product made from the harvestable parts of the plant. The step of growing the plant may be performed only once each time the methods of the invention is performed, while allowing repeated times the steps of product production e.g. by repeated removal of harvestable parts of the plants of the invention and if necessary further processing of these parts to arrive at the product. It is also possible that the step of growing the plants of the invention is repeated and plants or harvestable parts are stored until the production of the product is then performed once for the accumulated plants or plant parts. Also, the steps of growing the plants and producing the product may be performed with an overlap in time, even simultaneously to a large extend, or sequentially. Generally the plants are grown for some time before the product is produced.
Advantageously the methods of the invention are more efficient than the known methods, because the plants of the invention have increased yield and/or stress tolerance to an environmental stress compared to a control plant used in comparable methods.
In one embodiment the products produced by said 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. Foodstuffs are regarded as compositions used for nutrition or for supplementing nutrition. Animal feedstuffs and animal feed supplements, in particular, are regarded as foodstuffs.
In another embodiment the inventive methods for the 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.
It is possible that a plant product consists of one ore more agricultural products to a large extent.

In yet another embodiment the polynucleotide sequences or the polypeptide sequences or the construct of the invention are comprised in an agricultural product.

In a further embodiment the nucleic acid sequences and protein sequences of the invention may be used as product markers, for example for an agricultural product 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 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 one embodiment the plants of the invention or 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.

In another preferred embodiment of the present invention the plants of the invention and the plants used in the methods of the invention are sugarbeet plants with increased biomass and/or increased sugar content of the beets. In a further preferred embodiment of the present invention the plants of the invention and the plants used in the methods of the invention are sugarcane plants with increased biomass and/or increased sugar content of the stem.

The invention also extends to harvestable parts of a plant such as, but not limited to seeds, leaves, roots, stems, fruits, flowers, stems, roots, rhizomes, tubers and bulbs, which harvestable parts comprise a recombinant nucleic acid encoding a TLP polypeptide. The invention furthermore relates to products derived or produced, preferably directly derived or directly produced, from a harvestable part of such a plant, such as dry pellets or powders, oil, fat and fatty acids, sugars (in particular sucrose) starch or proteins. In one embodiment the product comprises a recombinant nucleic acid encoding a TLP polypeptide and/or a recombinant TLP polypeptide for example as an indicator of the particular quality of the product.

The present invention also encompasses use of nucleic acids encoding TLP polypeptides as described herein and use of these TLP polypeptides in enhancing any of the aforementioned yield-related traits in plants. For example, nucleic acids encoding TLP polypeptide described herein, or the TLP polypeptides themselves, may find use in breeding programmes in which a DNA marker is identified which may be genetically linked to a TLP polypeptide-encoding gene. The nucleic acids/genes, or the TLP 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 hereinabove in the methods of the invention. Furthermore, allelic variants of a TLP polypeptide-encoding nucleic acid/gene may find use in marker-assisted breeding programmes. Nucleic acids encoding TLP 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.

In one embodiment any comparison to determine sequence identity percentages is performed

- in the case of a comparison of nucleic acids over the entire coding region of SEQ ID NO: 1, or
- in the case of a comparison of polypeptide sequences over the entire length of SEQ ID NO: 2.

For example, a sequence identity of 50% sequence identity in this embodiment means that over the entire coding region of SEQ ID NO: 1, 50 percent of all bases are identical between the sequence of SEQ ID NO: 1 and the related sequence. Similarly, in this embodiment a polypeptide sequence is 50 % identical to the polypeptide sequence of SEQ ID NO: 2, when 50 percent of the amino acids residues of the sequence as represented in SEQ ID NO: 2, are found in the polypeptide tested when comparing from the starting methionine to the end of the sequence of SEQ ID NO: 2.

In a further embodiment the nucleic acid sequence employed in the invention are those sequences that are not the polynucleotides encoding the proteins selected from the group consisting of the proteins listed in Table A1, and those of at least 60, 70, 75, 80, 85, 90, 93, 95, 98 or 99% nucleotide identity when optimally aligned to the sequences encoding the proteins listed in Table A1.

In one embodiment, the sequence of the nucleic acid encoding said TLP polypeptide or the sequence of the TLP polypeptide is, preferably, not the sequence as shown in SEQ ID NO: 63278 as disclosed in US2007/061916, SEQ ID NO: 214797 as disclosed in US20040214272, SEQ ID NO 51042 as disclosed in US20040172684, and/or SEQ ID NO 70406 as disclosed in US20040034888.

C-2. PMP22 polypeptide (22 kDa peroxisomal membrane like polypeptide)

Surprisingly, it has now been found that modulating expression in a plant of a nucleic acid encoding a PMP22 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 PMP22 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 enhancing 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 PMP22 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 PMP22 polypeptide is by introducing and expressing in a plant a nucleic acid encoding a PMP22 polypeptide.

Any reference hereinafter in section C-2 to a "protein useful in the methods of the invention" is taken to mean a PMP22 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 PMP22 polypeptide. In one embodiment any reference to a protein or nucleic acid "useful in the methods of the invention" is to be understood to mean proteins or nucleic acids "useful in the methods, constructs, plants, harvestable parts and products of the invention". 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 "*PMP22* nucleic acid" or "*PMP22* gene".

A "PMP22 polypeptide" as defined herein, preferably, refers to any polypeptide comprising an Interpro domain having the Interpro Accession number IPR007248 (Mpv17/PMP22). "PMP22" is the abbreviation for "22 kDa Peroxisomal Membrane like protein". Thus, a PMP22 polypeptide is, preferably, a 22 kDa Peroxisomal Membrane like protein. More preferably, it is a 22 kDa Peroxisomal Membrane protein.

Additionally or alternatively, a "PMP22 polypeptide", preferably, refers to any polypeptide comprising a Pfam domain having the Pfam accession number PF04117 (PF04117, Mpv17/PMP22 domain).

The Pfam domains as referred to herein are, preferably, based on the Pfam database, Release 24.0 (Pfam 24.0, October 2009), see also The Pfam protein families database: R.D. Finn, J. Mistry, J. Tate, P. Coggill, A. Heger, J.E. Pollington, O.L. Gavin, P. Gunesekaran, G. Ceric, K.

Forslund, L. Holm, E.L. Sonnhammer, S.R. Eddy, A. Bateman Nucleic Acids Research (2010) Database Issue 38:D211-222.

Preferably, the Pfam domain having the Pfam accession number PF04117 (also referred to as "PF04117 pfam domain" or "PF04117 domain" herein) comprises a sequence having 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%, 99% or 100% sequence identity to the conserved domain starting with amino acid 283 up to amino acid 348 in SEQ ID NO:51.

The Interpro- and Pfam-domains as referred to herein are, preferably, based on the InterPro database, Release 31.0 (9th February 2011).

As set forth above, the PMP22 is the abbreviation for "22kDa Peroxisomal Membrane protein". However, it is envisaged that the PMP22 polypeptide in the context of the present invention may have a molecular weight that differs from 22 kDa.

Preferably, the PMP22 polypeptide additionally or alternatively comprises one or more of the following motifs (see also Fig. 6):

Motif 1-2 (SEQ ID NO: 126):
GDWIAQC[YF]EGKPLFE[FI]DR[AT]RM[FL]RSGLVGFTLHGSLSHYYY[QH]FCE[AE]LFPF[QKE]
Motif 2-2 (SEQ ID NO: 127):
LTID[HQ]DYWHGWT[LI][FY]EILRY[AM]P[QE]HNW[VSI]AYE[EQ]ALK[RTA]NPVLAKM
Motif 3-2 (SEQ ID NO: 128): [DE]WWVVP[AV]KVAFDQT[VA]W[SA]A[IV]WN
Motif 4-2 (SEQ ID NO: 129):
LVGFTLHGSLSHYYY[QH][FIL]CEALFPF[QKE][DE]WWVVP[AV]KVAFDQT[VI]WSAIWNSIYF
Motif 5-2 (SEQ ID NO: 130):
RY[AM]P[EQ]HNW[ISV]AYE[EQ]ALK[AR]NPVLAKM[VAM]ISG[VI]VYS[LIV]GDWIAQCYEGKP[LI]F[ED][FI]D
Motif 6-2 (SEQ ID NO: 131): AHL[IV]TYG[VL][IV]PVEQRLLWVDC
Motif 7-2 (SEQ ID NO: 132):
RYAPQHNW[IV]AYEEALK[RQ]NPVLAKMVISGVVYS[VL]GDWIAQCYEGKPLF[ED][IF]D
Motif 8-2 (SEQ ID NO: 133):
GFTLHGSLSH[YF]YYQFCE[AE]LFPF[QE]DWWVVP[VA]KVAFDQTVWSAIWNSIY[FY]TV
Motif 9-2 (SEQ ID NO: 134):
F[LW]PMLTAGWKLWPFAHLITYG[VL][VI]PVEQRLLWVDCVEL[IV]WVTILSTYSNEK

The term "PMP22" or "PMP22 polypeptide" as used herein also intends to include homologues as defined hereunder of "PMP22 polypeptide".

Motifs 1-2 to 9-2 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. Motifs 1-2 to 3-2 were derived when using MEME for all polypeptides shown in Table A2 (Cluster A, B and C). Motifs 4-2 to 6-2 were derived when using MEME for all polypeptides with SEQ ID NO: 51 to 97 (Cluster A and B) shown in Table A2. Motifs 4-2 to 6-2 were derived when using MEME for all polypeptides with SEQ ID NO: 51 to 65 (Cluster A) shown in Table A2.

In one preferred embodiment, the PMP22 polypeptide comprises one or more motifs selected from Motif 1-2, Motif 2-2, and Motif 3-2. Preferably, the PMP22 polypeptide comprises Motifs 1-2 and 2-2, or Motifs 2-2 and 3-2, or Motifs 1-2 and 3-2, or Motifs 1-2, 2-2 and 3-2.

In a further preferred embodiment, the PMP22 polypeptide comprises one or more motifs selected from Motif 4-2, Motif 5-2, and Motif 6-2. Preferably, the PMP22 polypeptide comprises Motifs 4-2 and 5-2, or Motifs 5-2 and 6-2, or Motifs 4-2 and 6-2, or, more preferably, Motifs 4-2, 5-2 and 6-2.

In an even further preferred embodiment, the PMP22 polypeptide comprises one or more motifs selected from Motif 7-2, Motif 8-2, and Motif 9-2. Preferably, the PMP22 polypeptide comprises Motifs 7-2 and 8-2, or Motifs 8-2 and 9-2, or Motifs 7-2 and 9-2, or, more preferably, Motifs 7-2, 8-2 and 9-2.

More preferably, the PMP22 polypeptide comprises in increasing order of preference, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, or all 9 motifs.

Thus, the PMP22 polypeptide, preferably, may comprise:

a. all of the following motifs:

Motif 1-2 (SEQ ID NO: 126):
GDWIAQC[YF]EGKPLFE[FI]DR[AT]RM[FL]RSGLVGFTLHGSLSHYYY[QH]FCE[AE]LFPF[QKE]

Motif 2-2 (SEQ ID NO: 127):
LTID[HQ]DYWHGWT[LI][FY]EILRY[AM]P[QE]HNW[VSI]AYE[EQ]ALK[RTA]NPVLAKM

Motif 3-2 (SEQ ID NO: 128): [DE]WWVVP[AV]KVAFDQT[VA]W[SA]A[IV]WN

Motif 4-2 (SEQ ID NO: 129):
LVGFT-
LHGSLSHYYY[QH][FIL]CEALFPF[QKE][DE]WWVVP[AV]KVAFDQT[VI]WSAIWNSIYF

Motif 5-2 (SEQ ID NO: 130): RY[AM]P[EQ]HNW[ISV]AYE[EQ]ALK[AR]NPVLAKM[VAM]ISG[VI]VYS[LIV]GD WIAQCYEGKP[LI]F[ED][FI]D
Motif 6-2 (SEQ ID NO: 131): AHL[IV]TYG[VL][IV]PVEQRLLWVDC
Motif 7-2 (SEQ ID NO: 132): RYAPQHNW[IV]AYEEALK[RQ]NPVLAKMVISGVVYS[VL]GDWIAQCYEGKPLF[ED][IF]D
Motif 8-2 (SEQ ID NO: 133): GFT-LHGSLSH[YF]YYQFCE[AE]LFPF[QE]DWWVVP[VA]KVAFDQTVWSAIWNSIY[FY]TV
Motif 9-2 (SEQ ID NO: 134): F[LW]PMLTAGWKLWPFAHLITYG[VL][VI]PVEQRLLWVDCVEL[IV]WVTILSTY SNEK;
or

b. at least one of the Motifs 7-2 to 9-2, preferably any two of Motifs 7-2 to 9-2, more preferably all three of Motifs 7-2 to 9-2 as defined in a. above; or

c. at least one of the Motifs 4-2 to 6-2, preferably any two of the Motifs 4-2 to 6-2, more preferably all three of the Motifs 4-2 to 6-2 as defined in a. above; or

d. at least one of the Motifs 1-2 to 3-2, preferably any two of the Motifs 1-2 to 3-2, more preferably all three of the Motifs 1-2 to 3-2 as defined in a. above; or

e. any four of the Motifs 1-2 to 9-2, preferably any five of the Motifs 1-2 to 9-2 as defined in a. above; or

f. any six of the Motifs 1-2 to 9-2, preferably any seven of the Motifs 1-2 to 9-2, more preferably any eight of the Motifs 1-2 to 9-2 as defined in a. above.

Additionally or alternatively, the PMP22 polypeptide or the homologue of a PMP22 protein, preferably, 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 represented by SEQ ID NO: 51. Preferably, said PMP22 polypeptide comprises the Pfam domain, and/or the Interpro domain and/or one or more 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: 51. In another embodiment the sequence identity level of a nucleic acid sequence is determined by comparison of the nucleic acid sequence over the entire length of the coding sequence of the sequence of SEQ ID NO: 50.

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 PMP22 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: 126 to SEQ ID NO: 134 (Motifs 1-2 to 9-2).

In other words, in another embodiment a method is provided wherein said PMP22 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 PF04117 domain. Preferably, said conserved PF04117 domain is starting with amino acid 283 up to amino acid 348 in SEQ ID NO:51.

In a preferred embodiment of the present invention, the PMP22 polypeptide to be used in the context of the present invention is selected from the group consisting of:

(i) a polypeptide comprising a sequence, or consisting of a sequence as shown in SEQ ID NO: 51, 57, 91 or 105,

(ii) a polypeptide having, in an increasing order of preference, at least 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 a polypeptide as represented by SEQ ID NO: 51, 57, 91 or 105 when compared over the entire length of the amino acid sequence as represented by SEQ ID NO: 51, 57, 91 or 105, respectively,

(iii) a polypeptide encoded by a polynucleotide which hybridizes under stringent conditions to a polynucleotide having a sequence as shown in SEQ ID NO: 50, 56, 90, or 104 or with a complementary sequence of such a polynucleotide having a sequence as shown in SEQ ID NO: 50, 56, 90, or 104;

(vi) a polypeptide with the biological activity of the polypeptide as shown in SEQ ID NO: 51, 57, 91 or 105 or substantially the same biological activity of the polypeptide as shown in SEQ ID NO: 51, 57, 91 or 105; and

(v) any combination of (i) to (iv) above.

Preferalby, the PMP22 polypeptide comprise the domains and/or motifs as set forth herein above.

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

In one embodiment, the sequence of the nucleic acid encoding said PMP22 polypeptide or the sequence of the PMP22 polypeptide is not the sequence as shown in SEQ ID NO: 20 as disclosed in WO2004/035798, as shown in SEQ ID NO: 5180 as disclosed in EP 1 586 645 A2, as shown in SEQ ID NO: 277535 as disclosed in US2004031072, as shown in SEQ ID NO: 42604 as disclosed in JP2005185101, as shown in SEQ ID NO: 302211 as disclosed in US2004214272, SEQ ID NO: 6940 as disclosed in US2009019601, or SEQ ID NO: 69977 or SEQ ID NO: 51830 as disclosed in US2007011783. Moreover, said sequence is, preferably, not SEQ ID NO: 34117 as disclosed in CA2300693, and/or not SEQ ID NO: 91119 as disclosed in US20070061916.

In one embodiment, the sequence of the nucleic acid encoding said PMP22 polypeptide or the sequence of the PMP22 polypeptide is, preferably, not the sequence as shown in SEQ ID NO 40059 as disclosed in US20080148432, SEQ ID NO: 168858 as disclosed in US20040123343, and/or SEQ ID NO: 168851 as disclosed in US20040123343.

Preferably, the polypeptide sequence which when used in the construction of a phylogenetic tree, such as the one depicted in Figure 8, clusters with the group of PMP22 polypeptides comprising the amino acid sequence represented by SEQ ID NO: 51 rather than with any other group (Cluster A).

In addition, PMP22 polypeptides, when expressed in a monocot plant such as rice, maize, wheat or sugarcane according to the methods of the present invention as outlined in Examples 7 and 8, give plants having increased yield related traits, in particular under-non stress conditions aboveground biomass (AreaMax), number of flowers per panicle (flowerperpan), thousand kernel weight (TKW) and/or under nitrogen deficiency increased seed fillrate (number of filled seeds over the number of florets), number of flowers per panicle (flowerperpan), and thousand kernel weight (TKW).

The present invention is illustrated by transforming plants with the nucleic acid sequence represented by SEQ ID NO: 50, encoding the polypeptide sequence of SEQ ID NO: 51. However, performance of the invention is not restricted to these sequences; the methods of the invention may advantageously be performed using any PMP22-encoding nucleic acid or PMP22 polypeptide as defined herein.

Examples of nucleic acids encoding PMP22 polypeptides are given in Table A2 of the Examples section herein. Such nucleic acids are useful in performing the methods of the invention. The

amino acid sequences given in Table A2 of the Examples section are example sequences of orthologues and paralogues of the PMP22 polypeptide represented by SEQ ID NO: 51, 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: 50 or SEQ ID NO: 51, the second BLAST (back-BLAST) would be against Lycopersicon esculentum sequences.

The invention also provides hitherto unknown PMP22-encoding nucleic acids and PMP22 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: 50, 56, 90, or 104;

(ii) the complement of a nucleic acid represented by SEQ ID NO: 50, 56, 90, or 104;

(iii) a nucleic acid encoding a PMP22 polypeptide having in increasing order of preference at least 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%, 99%, or 100% sequence identity to the amino acid sequence represented by SEQ ID NO: 51, 57, 91 or 105, 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.

Preferably, said PMP22 polypeptide encoded by said nucleic acid comprises a Pfam domain having the accession number PF04117. Additionally or alternatively, said PMP22 polypeptide comprises an Interpro domain having the accession number IPR007248. It is also preferred that said PMP22 polypeptide comprises –additionally or alternatively- one or more of Motifs 1-2 to 9-2. Preferred combinations of Motifs 1-2 to 9-2 are disclosed herein above.

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: 57, 91 or 105;

(ii) an amino acid sequence having, in increasing order of preference, at least 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 se-

quence represented by SEQ ID NO: 57, 91 or 105, and preferably conferring enhanced yield-related traits relative to control plants; and

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

Preferably, said polypeptide comprises a Pfam domain having the accession number PF04117. Additionally or alternatively, said polypeptide comprises an Interpro domain having the accession number IPR007248. It is also preferred that said polypeptide comprises –additionally or alternatively- one or more of Motifs 1-2 to 9-2. Preferred combinations of Motifs 1-2 to 9-2 are disclosed herein above.

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 any one of SEQ ID NO: 56, 90, and 104;

(ii) the complement of a nucleic acid represented by (any one of) SEQ ID NO: 56, 90, and 104;

(iii) a nucleic acid encoding the polypeptide as represented by any one of SEQ ID NO: 57, 91, and 105, preferably as a result of the degeneracy of the genetic code, said isolated nucleic acid can be derived from a polypeptide sequence as represented by any one of SEQ ID NO: 57, 91, and 105 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 any of the nucleic acid sequences of Table A2 and further preferably conferring enhanced yield-related traits relative to control plants;

(v) a nucleic acid molecule which hybridizes with 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 a PMP22 polypeptide having, in increasing order of preference, at least 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 any one of SEQ ID NO: 57, 91 and 105 and any of the other amino acid sequences in Table A2 and preferably conferring enhanced yield-related traits relative to control plants.

Preferably, said polypeptide encoded by said nucleic acid comprises a Pfam domain having the accession number PF04117. Additionally or alternatively, said polypeptide comprises an Interpro domain having the accession number IPR007248. It is also preferred that said polypeptide comprises –additionally or alternatively- one or more of Motifs 1-2 to 9-2. Preferred combinations of Motifs 1-2 to 9-2 are disclosed herein above.

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 any one of SEQ ID NO: 57, 91 and 105;

(ii) an amino acid sequence having, in increasing order of preference, at least 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 any one of SEQ ID NO: 57, 91 and 105 and any of the other amino acid sequences in Table A2 and 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.

Preferably, said polypeptide comprises a Pfam domain having the accession number PF04117. Additionally or alternatively, said polypeptide comprises an Interpro domain having the accession number IPR007248. It is also preferred that said polypeptide comprises –additionally or alternatively- one or more of Motifs 1-2 to 9-2. Preferred combinations of Motifs 1-2 to 9-2 are disclosed herein 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 A2 of the Examples section, the terms “homologue” and “derivative” being as defined herein. Also useful in the methods, constructs, plants, harvestable parts and products 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 A2 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 PMP22 polypeptides, nucleic acids hybridising to nucleic acids encoding PMP22 polypeptides, splice variants of nucleic acids encoding PMP22 polypeptides, allelic variants of nucleic acids encoding PMP22 polypeptides and variants of nucleic acids encoding PMP22 polypeptides obtained by gene shuffling. The terms hybridising sequence, splice variant, allelic variant and gene shuffling are as described herein.

In one embodiment of the present invention the function of the nucleic acid sequences of the invention is to confer information for a protein that increases yield or yield related traits, when a nucleic acid sequence of the invention is transcribed and translated in a living plant cell.

Nucleic acids encoding PMP22 polypeptides need not be full-length nucleic acids, since performance of the methods of the invention does not 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 A2 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 A2 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, constructs, plants, harvestable parts and products of the invention, encode a PMP22 polypeptide as defined herein, and have substantially the same biological activity as the amino acid sequences given in Table A2 of the Examples section. Preferably, the portion is a portion of any one of the nucleic acids given in Table A2 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 A2 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 or 1302 consecutive nucleotides in length, the consecutive nucleotides being of any one of the nucleic acid sequences given in Table A2 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 A2 of the Examples section. Most preferably the portion is a portion of the nucleic acid of SEQ ID NO: 50. Preferably, the portion encodes a fragment of an amino acid sequence which, when used in the construction of a phylogenetic tree, such as the one depicted in Figure 8, clusters with the group of PMP22 polypeptides comprising the amino acid sequence represented by SEQ ID NO: 51 (cluster A), rather than with any other group, and/or comprises motifs 1-2 to 9-2, and/or has at least 70% sequence identity to SEQ ID NO: 51.

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 PMP22 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 any one of the nucleic acids given in Table A2 of the Examples section, or comprising introducing and expressing in a plant a nucleic acid capable of hybridising to a nucleic acid encoding

an orthologue, paralogue or homologue of any of the nucleic acid sequences given in Table A2 of the Examples section.

Hybridising sequences useful in the methods, constructs, plants, harvestable parts and products of the invention encode a PMP22 polypeptide as defined herein, having substantially the same biological activity as the amino acid sequences given in Table A2 of the Examples section. Preferably, the hybridising sequence is capable of hybridising to the complement of any one of the nucleic acids given in Table A2 of the Examples section, or to a portion of any of these sequences, a portion being as defined above, 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 A2 of the Examples section. Most preferably, the hybridising sequence is capable of hybridising to the complement of a nucleic acid as represented by SEQ ID NO: 50 or to a portion thereof.

Preferably, the hybridising sequence encodes a polypeptide with an amino acid sequence which, when full-length and used in the construction of a phylogenetic tree, such as the one depicted in Figure 8, clusters with the group of PMP22 comprising the amino acid sequence represented by SEQ ID NO: 51 (cluster A) rather than with any other group, and/or comprises a PF04117 or IPR007248 domain, and/or comprises at least one motif from motifs 1-2 to 9-2 as specified elsewhere herein, and/or has at least 70% sequence identity to SEQ ID NO: 51.

In one embodiment the hybridising sequence is capable of hybridising to the complement of a nucleic acid as represented by SEQ ID NO: 50 or to a portion thereof under conditions of medium or high stringency, preferably high stringency as defined above. In another embodiment the hybridising sequence is capable of hybridising to the complement of a nucleic acid as represented by SEQ ID NO: 50 under stringent conditions.

Another nucleic acid variant useful in the methods, constructs, plants, harvestable parts and products of the invention is a splice variant encoding a PMP22 polypeptide as defined hereinabove, a splice variant being 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 splice variant of any one of the nucleic acid sequences given in Table A2 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 A2 of the Examples section.

Preferred splice variants are splice variants of a nucleic acid represented by SEQ ID NO: 50, or a splice variant of a nucleic acid encoding an orthologue or paralogue of SEQ ID NO: 51. Preferably, the amino acid sequence encoded by the splice variant, when used in the construction

of a phylogenetic tree, such as the one depicted in Figure 8, clusters with the group of PMP22 comprising the amino acid sequence represented by SEQ ID NO: 51 (cluster A) rather than with any other group, and/or comprises a PF04117 or IPR007248 domain, and/or comprises at least one motif from motifs 1-2 to 9-2 as specified elsewhere herein, and/or has at least 70% sequence identity to SEQ ID NO: 51.

Another nucleic acid variant useful in performing the methods of the invention is an allelic variant of a nucleic acid encoding a PMP22 polypeptide as defined hereinabove, an allelic variant being 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 an allelic variant of any one of the nucleic acids given in Table A2 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 A2 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 PMP22 polypeptide of SEQ ID NO: 51 and any of the amino acids depicted in Table A2 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: 50 or an allelic variant of a nucleic acid encoding an orthologue or paralogue of SEQ ID NO: 51. Preferably, the amino acid sequence encoded by the allelic variant, when used in the construction of a phylogenetic tree, such as the one depicted in Figure 8, clusters with the group of PMP22 comprising the amino acid sequence represented by SEQ ID NO: 51 (cluster A) rather than with any other group, and/or comprises a PF04117 or IPR007248 domain, and/or comprises at least one motif from motifs 1-2 to 9-2 as specified elsewhere herein, and/or has at least 70% sequence identity to SEQ ID NO: 51.

Gene shuffling or directed evolution may also be used to generate variants of nucleic acids encoding PMP22 polypeptides as defined above; the term “gene shuffling” being 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 variant of any one of the nucleic acid sequences given in Table A2 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 A2 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, when used in the construction of a phylogenetic tree such as the one depicted in Figure 8, clusters with the group of PMP22 comprising the amino acid sequence represented by SEQ ID NO: 51 (cluster A) rather than with any other group, and/or comprises a PF04117 or IPR007248 domain, and/or comprises at least one motif from motifs 1-2 to 9-2 as specified elsewhere herein, and/or has at least 70% sequence identity to SEQ ID NO: 51.

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.).

Nucleic acids encoding PMP22 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 PMP22 polypeptide-encoding nucleic acid is from a plant, further preferably from a dicotyledonous plant, more preferably from the family Solanaceae, even further preferably from the genus Solanum, and most preferably the nucleic acid is from *S. lycopersicum (which is the same of Lycopersicum esculentum).*

In another embodiment the present invention extends to recombinant chromosomal DNA comprising a nucleic acid sequence useful in the methods, constructs, plants, harvestable parts and products of the invention, wherein said nucleic acid is present in the chromosomal DNA as a result of recombinant methods, i.e. said nucleic acid is not in the chromosomal DNA in its native surrounding. Said recombinant chromosomal DNA may be a chromosome of native origin, with said nucleic acid inserted by recombinant means, or it may be a mini-chromosome or a non-native chromosomal structure, e.g. or an artificial chromosome. The nature of the chromosomal DNA may vary, as long it allows for stable passing on to successive generations of the recombinant nucleic acid useful in the methods, constructs, plants, harvestable parts and products of the invention, and allows for expression of said nucleic acid in a living plant cell resulting in increased yield or increased yield related traits of the plant cell or a plant comprising the plant cell.

In a further embodiment the recombinant chromosomal DNA of the invention is comprised in a plant cell. DNA comprised within a cell, particularly a cell with cell walls like a plant cell, is better protected from degradation than a bare nucleic acid sequence. The same holds true for a DNA construct comprised in a host cell, for example 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 seed yield relative to control plants. The terms “yield” and “seed yield” are described in more detail in the “definitions” section herein.

Reference herein to enhanced yield-related traits is taken to mean an increase 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. Preferably, such harvestable parts are seeds, leafs, roots and shoots. In particular, such harvestable parts are roots such as taproots, stems, seeds, and performance of the methods of the invention results in plants having increased seed yield relative to the seed yield of control plants, and/or increased stem biomass relative to the stem biomass of control plants, and/or increased root biomass relative to the root biomass and/or increased beet biomass relative to the beet biomass and/or increased tuber biomass relative to the tuber biomass of control plants. Moreover, it is particularly contemplated that the sugar content (in particular the sucrose content) in the stem (in particular of sugar cane plants) and/or in the belowground parts, in particular in roots including taproots, tubers and/or beets (in particular in sugar beets) is increased relative to the sugar content (in particular the sucrose content) in corresponding part(s)of the control plant.

The present invention provides a method for increasing yield-related traits, especially biomass yield or seed yield of plants, relative to control plants, which method comprises modulating expression in a plant of a nucleic acid encoding a PMP22 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 PMP22 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 relative to control plants grown under comparable conditions. Therefore, according to the present invention, there is provided a method for increasing yield 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 PMP22 polypeptide.

Performance of the methods of the invention gives plants grown under conditions of drought, increased yield relative to control plants grown under comparable conditions. Therefore, according to the present invention, there is provided a method for increasing yield in plants grown under conditions of drought which method comprises modulating expression in a plant of a nucleic acid encoding a PMP22 polypeptide.

Performance of the methods of the invention gives plants grown under conditions of nutrient deficiency, particularly under conditions of nitrogen deficiency, increased yield relative to control plants grown under comparable conditions. Therefore, according to the present invention, there is provided a method for increasing yield in plants grown under conditions of nutrient deficiency, which method comprises modulating expression in a plant of a nucleic acid encoding a PMP22 polypeptide.

Performance of the methods of the invention gives plants grown under conditions of salt stress, increased yield relative to control plants grown under comparable conditions. Therefore, according to the present invention, there is provided a method for increasing yield in plants grown under conditions of salt stress, which method comprises modulating expression in a plant of a nucleic acid encoding a PMP22 polypeptide.

The invention also provides genetic constructs and vectors to facilitate introduction and/or expression in plants of nucleic acids encoding PMP22 polypeptides. The gene constructs may be inserted into vectors, which may be commercially available, suitable for transforming into plants 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 PMP22 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 PMP22 polypeptide is as defined above. The term “control sequence” and “termination sequence” are as defined herein.

The invention furthermore provides plants 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.

Plants are transformed with a vector comprising any of the nucleic acids described above. The skilled artisan is well aware of the genetic elements that must be present on the vector 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) in the vectors of the invention.

In one embodiment the plants of the invention are transformed with an expression cassette comprising any of the nucleic acids described above. The skilled artisan is well aware of the genetic elements that must be present on the expression cassette in order to successfully trans-

form, select and propagate host cells containing the sequence of interest. In the expression cassettes of the invention the sequence of interest is operably linked to one or more control sequences (at least to a promoter). The promoter in such an expression cassette may be a non-native promoter to the nucleic acid described above, i.e. a promoter not regulating the expression of said nucleic acid in its native surrounding.

In one embodiment the terms expression cassettes of the invention, genetic construct and constructs of the invention are used exchangeably.

In a further embodiment the expression cassettes of the invention confer increased yield or yield related traits(s) to a living plant cell when they have been introduced into said plant cell and result in expression of the nucleic acid as defined above, comprised in the expression cassette(s). The promoter in such expression cassettes may be a non-native promoter to the nucleic acid described above, i.e. a promoter not regulating the expression of said nucleic acid in its native surrounding.

The expression cassettes of the invention may be comprised in a host cell, plant cell, seed, agricultural product or plant.

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. Preferably the constitutive promoter is a ubiquitous constitutive promoter of medium strength. See the “Definitions” section herein for definitions of the various promoter types. Also useful in the methods, constructs, plants, harvestable parts and products of the invention is a tissue specific promoter such as a seed or root-specific promoter.

It should be clear that the applicability of the present invention is not restricted to the PMP22 polypeptide-encoding nucleic acid represented by SEQ ID NO: 50, nor is the applicability of the invention restricted to expression of a PMP22 polypeptide-encoding nucleic acid when driven by a constitutive promoter.

The constitutive promoter is preferably a medium strength promoter. 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: 135, most preferably the constitutive promoter is as represented by SEQ ID NO: 135. See the “Definitions” section herein for further examples of constitutive promoters.

In a preferred embodiment, the polynucleotide encoding the PMP22 polypeptide as used in the plants, constructs and methods of the present invention is linked to a promoter which allows for the expression, preferably the strongest expression in the aboveground parts of the plant as

compared to the expression in other parts of the plant. This applies, in particular, if the plant is a monocot. As set forth elsewhere herein, preferred monocots are maize, wheat, rice, or sugarcane. In another preferred embodiment of the present invention, the polynucleotide encoding the PMP22 polypeptide as used in the plants, constructs and methods of the present invention is preferably linked to a promoter which allows for the expression, preferably the strongest expression in the belowground parts of the plant as compared to the expression in other parts of the plant. This applies, in particular, if the plant is a dicot. Preferred dicots are sugar beet and potato. For example, if the plant is a sugar beet, the promoter, preferably, allows for the strongest expression in the taproot as compared to the expression in other parts of the plant. In one embodiment the promoter used for expression in sugar beets is, preferably a root specific, more preferably a taproot or beet specific promoter.

Optionally, one or more terminator sequences may be used in the construct introduced into a plant. Preferably, the construct comprises an expression cassette comprising a GOS2 promoter, substantially similar to SEQ ID NO: 135, operably linked to the nucleic acid encoding the PMP22 polypeptide. More preferably, the construct comprises a zein terminator (t-zein) linked to the 3' end of the PMP22 coding sequence. Most preferably, the expression cassette comprises a sequence having in increasing order of preference at least 95%, at least 96%, at least 97%, at least 98%, at least 99% identity to the sequence represented by pPRO::PMP22::t-zein sequence as comprised by the expression vector having a sequence as shown in SEQ ID NO:136. 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 (and, thus over-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 PMP22 polypeptide is by introducing and expressing in a plant a nucleic acid encoding a PMP22 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 PMP22 polypeptide as defined hereinabove.

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

(i) introducing and expressing in a plant or plant cell a PMP22 polypeptide-encoding nucleic acid or a genetic construct comprising a PMP22 polypeptide-encoding nucleic acid; and

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

Cultivating the plant cell under conditions promoting plant growth and development, may or may not include regeneration and or growth to maturity.

The nucleic acid of (i) may be any of the nucleic acids capable of encoding a PMP22 polypeptide as defined herein.

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 by transformation. The term “transformation” is described in more detail in the “definitions” section herein.

In one embodiment the present invention clearly 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 parts thereof comprise a nucleic acid transgene encoding a PMP22 polypeptide as defined above. 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.

The present invention also extends in another embodiment to transgenic plant cells and seed comprising the nucleic acid molecule of the invention in a plant expression cassette or a plant expression construct.

In a further embodiment the seed of the invention recombinantly comprise the expression cassettes of the invention, the (expression) constructs of the invention, the nucleic acids described above and/or the proteins encoded by the nucleic acids as described above.

A further embodiment of the present invention extends to plant cells comprising the nucleic acid as described above in a recombinant plant expression cassette.

In yet another embodiment the plant cells of the invention are non-propagative cells, e.g. the cells can not be used to regenerate a whole plant from this cell as a whole using standard cell

culture techniques, this meaning cell culture methods but excluding in-vitro nuclear, organelle or chromosome transfer methods. 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. One example are plant cells that do not sustain themselves through photosynthesis by synthesizing carbohydrate and protein from such inorganic substances as water, carbon dioxide and mineral salt.

In another embodiment the plant cells of the invention are plant cells that do not sustain themselves through photosynthesis by synthesizing carbohydrate and protein from such inorganic substances as water, carbon dioxide and mineral salt, i.e. they may be deemed non-plant variety. In a further embodiment the plant cells of the invention are non-plant variety and non-propagative.

The invention also includes host cells containing an isolated nucleic acid encoding a PMP22 polypeptide as defined hereinabove. Host cells of the invention may be any cell selected from the group consisting of bacterial cells, such as E.coli or Agrobacterium species cells, yeast cells, fungal, algal or cyanobacterial cells or plant cells. In one embodiment host cells according to the invention are plant cells, yeasts, bacteria or fungi. Host plants for the nucleic acids or the vector used in the method according to the invention, the expression cassette or construct or vector are, in principle, advantageously all plants, which are capable of synthesizing the polypeptides used in the inventive method.

In one embodiment the plant cells of the invention overexpress the nucleic acid molecule of the invention.

The invention also includes methods for the production of 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, including seeds, of these plants. In a further embodiment the methods comprises steps a) growing the plants of the invention, b) removing the harvestable parts as defined above from the plants and c) producing said product from or by the harvestable parts of the invention.
Examples of such methods would be growing corn plants of the invention, harvesting the corn cobs and remove the kernels. These may be used as feedstuff or processed to starch and oil as agricultural products.
The product may be produced at the site where the plant has been grown, or the plants or parts thereof may be removed from the site where the plants have been grown to produce the product. Typically, the plant is grown, the desired harvestable parts are removed from the plant, if feasible in repeated cycles, and the product made from the harvestable parts of the plant. The step of growing the plant may be performed only once each time the methods of the invention is

performed, while allowing repeated times the steps of product production e.g. by repeated removal of harvestable parts of the plants of the invention and if necessary further processing of these parts to arrive at the product. It is also possible that the step of growing the plants of the invention is repeated and plants or harvestable parts are stored until the production of the product is then performed once for the accumulated plants or plant parts. Also, the steps of growing the plants and producing the product may be performed with an overlap in time, even simultaneously to a large extend, or sequentially. Generally the plants are grown for some time before the product is produced.

Advantageously the methods of the invention are more efficient than the known methods, because the plants of the invention have increased yield and/or stress tolerance to an environmental stress compared to a control plant used in comparable methods.

In one embodiment the products produced by said 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. Foodstuffs are regarded as compositions used for nutrition or for supplementing nutrition. Animal feedstuffs and animal feed supplements, in particular, are regarded as foodstuffs.

In another embodiment the inventive methods for the 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.

It is possible that a plant product consists of one ore more agricultural products to a large extent.

In yet another embodiment the polynucleotide sequences or the polypeptide sequences of the invention are comprised in an agricultural product.

in a further embodiment the nucleic acid sequences and protein sequences of the invention may be used as product markers, for example for an agricultural product 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 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, constructs, plants, harvestable parts and products 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 one embodiment the plants used of the invention or 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.
In another embodiment of the present invention the plants of the invention and the plants used in the methods of the invention are sugarbeet plants with increased biomass and/or increased sugar content of the beets. In another embodiment of the present invention the plants of the invention and the plants used in the methods of the invention are sugarcane plants with increased biomass and/or increased sugar content.

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 PMP22 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 or powders, oil, fat and fatty acids, starch or proteins. In one embodiment the product comprises a recombinant nucleic acid encoding a PMP22 polypeptide and/or a recombinant PMP22 polypeptide for example as an indicator of the particular quality of the product.

The present invention also encompasses use of nucleic acids encoding PMP22 polypeptides as described herein and use of these PMP22 polypeptides in enhancing any of the aforementioned yield-related traits in plants. For example, nucleic acids encoding PMP22 polypeptide described herein, or the PMP22 polypeptides themselves, may find use in breeding programmes in which a DNA marker is identified which may be genetically linked to a PMP22 polypeptide-encoding gene. The nucleic acids/genes, or the PMP22 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 hereinabove in the methods of the invention. Furthermore, allelic variants of a PMP22 polypeptide-encoding nucleic acid/gene may find use in marker-assisted breeding programmes. Nucleic acids encoding PMP22 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.

In one embodiment any comparison to determine sequence identity percentages is performed
- in the case of a comparison of nucleic acids over the entire coding region of SEQ ID

NO: 50, or

- in the case of a comparison of polypeptide sequences over the entire length of SEQ ID NO: 51.

For example, a sequence identity of 50% sequence identity in this embodiment means that over the entire coding region of SEQ ID NO: 50, 50 percent of all bases are identical between the sequence of SEQ ID NO: 50 and the related sequence. Similarly, in this embodiment a polypeptide sequence is 50 % identical to the polypeptide sequence of SEQ ID NO: 51, when 50 percent of the amino acids residues of the sequence as represented in SEQ ID NO: 51, are found in the polypeptide tested when comparing from the starting methionine to the end of the sequence of SEQ ID NO: 51.

In one embodiment the nucleic acid sequences employed in the methods, constructs, plants, harvestable parts and products of the invention are sequences encoding PMP22.

C-3. RTF (REM-like transcription factor) polypeptide

Surprisingly, it has now been found that modulating expression in a plant of a nucleic acid encoding a RTF 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 RTF 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 enhancing 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 RTF 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 RTF polypeptide is by introducing and expressing in a plant a nucleic acid encoding a RTF polypeptide.

Any reference hereinafter in section C-3 to a “protein useful in the methods of the invention” is taken to mean a RTF 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 RTF polypeptide. In one embodiment any reference to a protein or nucleic acid “useful in the methods of the invention” is to be understood to mean proteins or nucleic acids “useful in the methods, constructs, plants, harvestable parts and products of the invention”. The nucleic acid to be introduced into a plant (and therefore useful in performing the methods of the inven-

tion) is any nucleic acid encoding the type of protein which will now be described, hereafter also named "*RTF* nucleic acid" or "*RTF* gene". "RTF" is the abbreviation for REM (Reproductive meristem)-like transcription factor.

A "RTF polypeptide" as used herein, preferably, refers to a polypeptide comprising at least two B3 domains. Preferably, the RTF polypeptide also comprises an IPR015300 domain (DNA-binding pseudobarrel domain).

Preferably, the RTF polypeptide applied in the context of the present invention is encoded by a nucleic acid selected from

(i) a nucleic acid represented by any one of SEQ ID NO: 139, 141, 143, 145, 147, 149, 151, 153, 155, 157, 159, 161, or 163;

(ii) the complement of a nucleic acid represented by any one of SEQ ID NO: 139, 141, 143, 145, 147, 149, 151, 153, 155, 157, 159, 161, or 163;

(iii) a nucleic acid encoding the polypeptide as represented by any one of SEQ ID NO: 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, or 164, preferably as a result of the degeneracy of the genetic code, said isolated nucleic acid can be deduced from a polypeptide sequence as represented by any one of SEQ ID NO: 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, or 164;

(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 any of the nucleic acid sequences of SEQ ID NO SEQ ID NO: 139, 141, 143, 145, 147, 149, 151, 153, 155, 157, 159, 161, or 163,

(v) a nucleic acid which hybridizes with the nucleic acid molecule of (i) to (iv) under stringent hybridization conditions, and

(vi) a nucleic acid encoding a polypeptide having, in increasing order of preference, at least 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 any one of SEQ ID NO: 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, or 164.

Preferably, the RTF polypeptide encoded by the nucleic acid as set forth above confers-when expressed in a plant - enhanced yield-related traits relative to control plants, in particular, increased biomass (in particular increased aboveground and increased root biomass), and/or improved early vigor

Preferably, the B3 domains comprised by the "RTF polypeptide" are domains having the PFAM accession number pfam02362. More preferably, the B3 domains comprised by the "RTF polypeptide" are domains having the Interpro accession number IPR003340.

The Interpro domain IPR003340, preferably, corresponds to the IPR003340 domain of the InterPro database, Release 31.0 (9th February 2011). The Interpro domain IPR015300 is a DNA-binding pseudobarrel domain. Preferably, the domain, preferably, corresponds to the IPR015300 domain of the InterPro database, Release 35.0, 15 December, 2011.

The Pfam domain pfam02362, preferably, corresponds to PFAM domain with the accession number pfam02362 in the Pfam database, Release 24.0 (Pfam 24.0, October 2009), see also The Pfam protein families database: R.D. Finn, J. Mistry, J. Tate, P. Coggill, A. Heger, J.E. Pollington, O.L. Gavin, P. Gunesekaran, G. Ceric, K. Forslund, L. Holm, E.L. Sonnhammer, S.R. Eddy, A. Bateman Nucleic Acids Research (2010) Database Issue 38:D211-222.

In a preferred embodiment of the present invention, the RTF polypeptide comprises three B3 domains, in particular three domains having the PFAM accession number pfam02362 or having the Interpro accession number IPR003340. In an even more preferred embodiment, the RTF polypeptide comprises four B3 domains, in particular four domains having the PFAM accession number pfam02362 or having the Interpro accession number IPR003340. Is it also preferred that the RTF polypeptide comprises five, six, seven or eight B3 domains. Preferably, the RTF polypeptide further comprises an IPR015300 domain (DNA-binding pseudobarrel domain).

The B3 domains comprised by the RTF polypeptide are, preferably, separated by 10 to 150 amino acids, more preferably by 15 to 120 amino acids, even more preferably, by 20 to 200 amino acids, even more preferably, by 25 to 95 amino acids, and most preferably by 29 to 92 amino acids.

As set forth above, the RTF polypeptide comprises preferably a four B3 domains: a first, second, third and fourth B3 domain.

Preferably, the first B3 domain comprises a sequence having, 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%, 99%, or 100% sequence identity to a conserved domain from amino acid 13 to 105 in SEQ ID NO: 140. Preferably, the second B3 domain comprises a sequence having, 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%, 99%, or 100% sequence identity to a conserved domain from amino acid 150 to 247 in SEQ ID NO: 140. Pref-

erably, the third B3 domain comprises a sequence having, 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%, 99%, or 100% sequence identity to a conserved domain from amino acid 276 to 372 in SEQ ID NO: 140. Preferably, the fourth B3 domain comprises a sequence having, 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%, 99%, or 100% sequence identity to a conserved domain from amino acid 464 to 555 in SEQ ID NO:140).

Preferably, the order within the RTF polypeptide is as follows (from the N- to the C-terminus) is as follows: first B3 domain, second B3 domain, a third B3 domain and fourth B3 domain. Preferably the B3 domains are separated by 10 to 150 amino acids, and, more preferably, by 25 to 95 amino acids. Is particularly preferred that the first and second B3 domains are separated by 40 to 60 amino acids, that the second and third second B3 domains are separated by 20 to 50 amino acids, and that the third and fourth second B3 domains are separated by 80 to 120 amino acids.

Preferably, the degree of sequence identity is determined over the entire length of the aforementioned domains.

The B3 domains comprised by the RTF polypeptide, preferably, have a structure as decribed by Swaminathan et al. ((2008) The plant B3 superfamily. Trends Plant Sci. 2008 Dec;13(12):647-55, see Fig. 4). Accordingly, the B3 domain, preferably, comprises seven beta-strands which form an open beta barrel and two alpha helices.

Preferably, the RTF polypeptide comprises at least one Motif selected from
Motif 1-3 (SEQ ID NO: 165): PVAFF, and
Motif 2-3 (SEQ ID NO: 166): HDLRVGDIVVF.

It is particularly preferred that the RTF polypeptide comprises both Motif 1-3 and Motif 2-3.

When comprised by the plant cell, the RTF polypeptide is, preferably, located in the nucleus of a plant cell.

The term “RTF” or “RTF polypeptide” as used herein also intends to include homologues as defined hereunder of “RTF polypeptide”.

Additionally or alternatively, the RTF polypeptide or homologue of thereof 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 represented by SEQ ID NO: 140 provided that the homologous protein comprises at least two B3 domains, in particular three or four B3 domains as described above. Preferably, said RTF polypeptide or homologue of thereof comprised Motif 1-3 and/or Motif 2-3 (preferably Motif 1-3 and Motif 2-3). 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).

Preferably, the sequence identity is determined by comparison of the polypeptide sequences over the entire length of the sequence of SEQ ID NO: 140. Also, the sequence identity level of a nucleic acid sequence is, preferably, determined by comparison of the nucleic acid sequence over the entire length of the coding sequence of the sequence of SEQ ID NO: 139.

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 RTF 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: 165, an/or SEQ ID NO: 166 (Motifs 1-3 or 2-3).

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

In a preferred embodiment the RTF polypeptide/nucleic acid employed in the methods, constructs, plants, harvestable parts and products of the invention does not comprise the following sequence:
SEQ ID NOs: 43550, 43565, 43576, 43568, 43548, 43575, 193877, 93871, 43560, 93863, 43562, 93879, 43570, 43558, 43578, 93869, 43556, 43572, and 93875 as disclosed in EP2090662A2,
SEQ ID NOs: 312 and 2527 as disclosed in WO02/16655;
SEQ ID NO: 72 as disclosed in EP 2154956A2, and
SEQ ID NOs: 931, 584 , 838, 1764 as disclosed in WO2009014665;
SEQ ID NOs 10289 and 10291 as disclosed in US20060107345, and
SEQ ID NOs: 20362 and 20364 as disclosed in US20060150283.
In one embodiment, the sequence of the nucleic acid encoding said RTF polypeptide or the sequence of the RTF polypeptide is, preferably, not the sequence as shown in SEQ ID NO 237 as disclosed in WO2008/122980 and US20100154077, respectively, and the sequence as shown in SEQ ID NO: 931 as disclosed in WO2009/014665.

Preferably, the polypeptide sequence which when used in the construction of a phylogenetic tree, such as the one depicted in Figure 13, clusters with the RTF polypeptdide comprising the amino acid sequence represented by SEQ ID NO: 140 rather than with any other group.

Furthermore, the RTF polypeptide (at least its native form) preferably, binds to DNA, and, thus, has DNA binding activity. In particular, the RTF polypeptide shall bind to the major groove. Tools and techniques for assessing whether a polypeptide binds to DNA are well known in the art.

In addition, RTF polypeptides, when expressed in plant, in particular in monocots such as rice, maize, wehat or sugarcane, according to the methods of the present invention as outlined in the Examples section (see, e.g. Example XI-3), give plants having increased yield related traits, in particular increased biomass (in particular increased aboveground and increased root biomass), and improved early vigor. Preferably, said increased yield related traits obtained under non stress conditions.

The present invention is illustrated by transforming plants with the nucleic acid sequence represented by SEQ ID NO: 139, encoding the polypeptide sequence of SEQ ID NO: 140. However, performance of the invention is not restricted to these sequences; the methods of the invention may advantageously be performed using any RTF-encoding nucleic acid or RTF polypeptide as defined herein.

Examples of preferred nucleic acids encoding RTF polypeptides are given in Table A3 of the Examples section herein. Such nucleic acids are useful in performing the methods of the invention. The amino acid sequences given in Table A3 of the Examples section are example sequences of orthologues and paralogues of the RTF polypeptide represented by SEQ ID NO: 140, 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: 139 or SEQ ID NO: 140, the second BLAST (back-BLAST) would be against Arabidopsis thaliana sequences.

In the context of the present invention, the nucleic acid encoding the RTF polypeptide is, preferably, selected from:

(i) a nucleic acid represented by SEQ ID NO: 139;

(ii) the complement of a nucleic acid represented by SEQ ID NO: 139;

(iii) a nucleic acid encoding a RTF polypeptide having in increasing order of preference at least 20%, 30%, 40%, 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: 140.

(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.

Preferably, the polypeptide encoded by the said nucleic acid comprises at least 2 (in particular, 2, 3 or 4) B3 domains as described herein above. Preferably, said polypeptide also comprises Motif 1-3 and/or Motif 2-3 (preferably, both). Moreover, said polypeptide, preferably, confers enhanced yield-related traits relative to control plants, in particular, increased biomass (in particular increased aboveground and increased root biomass), and improved early vigor (when expressed in a plant).

Preferably, the RTF 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 any one of SEQ ID NO: 139, 141, 143, 145, 147, 149, 151, 153, 155, 157, 159, 161, or 163;

(ii) the complement of a nucleic acid represented by any one of SEQ ID NO: 139, 141, 143, 145, 147, 149, 151, 153, 155, 157, 159, 161, or 163;

(iii) a nucleic acid encoding the polypeptide as represented by any one of SEQ ID NO: 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, or 164, preferably as a result of the degeneracy of the genetic code, said isolated nucleic acid can be deduced from a polypeptide sequence as represented by any one of SEQ ID NO: 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, or 164 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 any of the nucleic acid sequences of SEQ ID NO SEQ ID NO: 139, 141, 143, 145, 147, 149, 151, 153, 155, 157, 159, 161, or 163, and further preferably conferring enhanced yield-related traits relative to control plants,

(v) a first nucleic acid molecule which hybridizes with a second 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 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 any one of SEQ ID NO: 140, 142, 144, 146, 148, 150, 152, 154, 156,

158, 160, 162, or 164 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.

Most preferably, the RTF polypeptide is selected from:

(i) an amino acid sequence represented by SEQ ID NO: 140;

(ii) an amino acid sequence having, in increasing order of preference, at least 20%, 30%, 40%, 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%, 99% or 100% sequence identity to the amino acid sequence represented by SEQ ID NO: 140;

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

(iv) a polypeptide with the biological activity of the polypeptide as shown in SEQ ID NO: 140 or substantially the same biological activity of the polypeptide as shown in SEQ ID NO: 140; and

(v) any combination of i) to iv) above.

Preferably, the polypeptide comprises at least 2 (in particular, 2, 3 or 4) B3 domains as described herein above. Preferably, said polypeptide also comprises Motif 1-3 and/or Motif 2-3 (preferably, both). Moreover, said polypeptide, preferably, confers enhanced yield-related traits relative to control plants, in particular, increased biomass (in particular increased aboveground and increased root biomass), and improved early vigor (when expressed in a plant).

Further preferred RTF polypeptides to be applied in the context of the present invention are the Arabidopsis thaliana transcription factors AtREM1 (At4g31610, NM_119310.3, NP_567880.1), AtREM2 (At4g31615, NM_1483872, NP_680753.2) AtREM3 (At4g31620, NM_119311.3, NP_194890.2) AtREM4 (At4g31630, NM_119312.1, NP_194891.1), AtREM5 (At4g31640, NM_119313.2, NP_194892.1) AtREM6 (At4g31650, NM_119314.1, NP_194893.1) AtREM8 (At4g31680, NM_119317.3, NP_194896.2), AtREM9, (At4g31690, NM_119318.1, NP_194897).1, AtREM7 (At4g31660, NM_119315.6, NP_194894.2), AtREM18 (At2g46730 NM_130238.1, NP_566083.1), AtREM13 (At2g24650, NM_001161059.1, NP_001154531.1), AtREM11 (At2g24690, NM_128030.4, NP_180045.4), AtREM12 and (At2g24680, NM_128029.1, NP_180044.1). The first number in the brackets is the TAIR Accession number (The Arabidopsis Information Resource (TAIR), see Swarbreck D, Wilks C, Lamesch P, Berardini TZ, Garcia-Hernandez M, Foerster H, Li D, Meyer T, Muller R, Ploetz L, Radenbaugh A, Singh S, Swing V, Tissier C, Zhang P, Huala E.(2008). The Arabidopsis Information Resource (TAIR): gene structure and function annotation. Nucleic Acids Research, 2008, Vol. 36, Database issue D1009–D1014). The second and third number in the brackets represent the GenBank-Accession-Number of the preferred RTF-polynucleotides (full length CDS) and poly-

peptide, respectively. Further preferred RTF polynucleotides are from rice and are selected from the group of Os04g27960, Os04g27990, Os06g42630, Os08g30500, and Os03g11370 (for the sequences, see e.g. Conte MG, Gaillard S, Lanau N, Rouard M, Périn C (2008). GreenPhylDB: a database for plant comparative genomics. Nucleic Acids Research. 2008 January; 36 D991–D998).

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 A3 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 A3 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 RTF polypeptides, nucleic acids hybridising to nucleic acids encoding RTF polypeptides, splice variants of nucleic acids encoding RTF polypeptides, allelic variants of nucleic acids encoding RTF polypeptides and variants of nucleic acids encoding RTF polypeptides obtained by gene shuffling. The terms hybridising sequence, splice variant, allelic variant and gene shuffling are as described herein.

In one embodiment of the present invention the function of the nucleic acid sequences of the invention is to confer information for a protein that increases yield or yield related traits, when a nucleic acid sequence of the invention is transcribed and translated in a living plant cell.

Nucleic acids encoding RTF polypeptides need not be full-length nucleic acids, since performance of the methods of the invention does not 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 A3 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 A3 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, constructs, plants, harvestable parts and productsof the invention, encode a RTF polypeptide as defined herein, and have substantially the same biological activity as the amino acid sequences given in Table A3 of the Examples section. Preferably, the portion is a portion of any one of the nucleic acids given in Table A3 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 A3 of the Examples section. Preferably the portion is at least 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, 2100, 2200, 2300, 2400, 2500, 2600, 2700 or 2763 consecutive nucleotides in length, the consecutive nucleotides being of any one of the nucleic acid sequences given in Table A3 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 A3 of the Examples section. Most preferably the portion is a portion of the nucleic acid of SEQ ID NO: 139. Preferably, the portion encodes a fragment of an amino acid sequence which, when used in the construction of a phylogenetic tree, such as the one depicted in Figure 13, clusters with the group of polypeptides which comprises the polypeptide having an amino acid sequence as shown in SEQ ID NO: 140 rather than with any other group, and/or comprises at least two B3 domains (in particular four B3 domains) as outlined herein above), and/or has DNA binding activity, and/or has at least 70% sequence identity to SEQ ID NO: 140.

Another nucleic acid variant useful in the methods, constructs, plants, harvestable parts and products of the invention is a nucleic acid capable of hybridising, under reduced stringency conditions, preferably under stringent conditions, with a nucleic acid encoding a RTF 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 any one of the nucleic acids given in Table A3 of the Examples section, or comprising introducing and expressing in a plant a nucleic acid capable of hybridising to a nucleic acid encoding an orthologue, paralogue or homologue of any of the nucleic acid sequences given in Table A3 of the Examples section.

Hybridising sequences useful in the methods, constructs, plants, harvestable parts and products of the invention encode a RTF polypeptide as defined herein, having substantially the same biological activity as the amino acid sequences given in Table A3 of the Examples section. Preferably, the hybridising sequence is capable of hybridising to the complement of any one of the nucleic acids given in Table A3 of the Examples section, or to a portion of any of these sequences, a portion being as defined above, 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 A3 of the Examples section. Most preferably, the hybrid-

ising sequence is capable of hybridising to the complement of a nucleic acid as represented by SEQ ID NO: 139 or to a portion thereof.

Preferably, the hybridising sequence encodes a polypeptide with an amino acid sequence which, when full-length and used in the construction of a phylogenetic tree, when used in the construction of a phylogenetic tree, such as the one depicted in Figure 13, clusters with the group of polypeptides which comprises the polypeptide having an amino acid sequence as shown in SEQ ID NO: 140 rather than with any other group, and/or comprises at least two B3 domains (in particular four B3 domains) as outlined herein above), and/or has DNA binding activity, and/or has at least 70% sequence identity to SEQ ID NO: 140.

In one embodiment the hybridising sequence is capable of hybridising to the complement of a nucleic acid as represented by SEQ ID NO: 139 or to a portion thereof under conditions of medium or high stringency, preferably high stringency as defined above. In another embodiment the hybridising sequence is capable of hybridising to the complement of a nucleic acid as represented by SEQ ID NO: 139 under stringent conditions.

Another nucleic acid variant useful in the methods, constructs, plants, harvestable parts and products of the invention is a splice variant encoding a RTF polypeptide as defined hereinabove, a splice variant being 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 splice variant of any one of the nucleic acid sequences given in Table A3 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 A3 of the Examples section.

Preferred splice variants are splice variants of a nucleic acid represented by SEQ ID NO: 139, or a splice variant of a nucleic acid encoding an orthologue or paralogue of SEQ ID NO: 140. Preferably, the amino acid sequence encoded by the splice variant, when used in the construction of a phylogenetic tree, such as the one depicted in Figure 13, clusters with the group of polypeptides which comprises the polypeptide having an amino acid sequence as shown in SEQ ID NO: 140 rather than with any other group, and/or comprises at least two B3 domains (in particular four B3 domains) as outlined herein above), and/or has DNA binding activity, and/or has at least 70% sequence identity to SEQ ID NO: 140.

Another nucleic acid variant useful in performing the methods of the invention is an allelic variant of a nucleic acid encoding a RTF polypeptide as defined hereinabove, an allelic variant being 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 an allelic variant of any one of the nucleic acids given in Table A3 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 A3 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 RTF polypeptide of SEQ ID NO: 140 and any of the amino acids depicted in Table A3 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: 139 or an allelic variant of a nucleic acid encoding an orthologue or paralogue of SEQ ID NO: 140. Preferably, the amino acid sequence encoded by the allelic variant, when used in the construction of a phylogenetic tree, such as the one depicted in Figure 13, clusters with the group of polypeptides which comprises the polypeptide having an amino acid sequence as shown in SEQ ID NO: 140 rather than with any other group, and/or comprises at least two B3 domains (in particular four B3 domains) as outlined herein above), and/or has DNA binding activity, and/or has at least 70% sequence identity to SEQ ID NO: 140.

Gene shuffling or directed evolution may also be used to generate variants of nucleic acids encoding RTF polypeptides as defined above; the term “gene shuffling” being 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 variant of any one of the nucleic acid sequences given in Table A3 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 A3 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, when used in the construction of a phylogenetic tree when used in the construction of a phylogenetic tree, such as the one depicted in Figure 13, clusters with the group of polypeptides which comprises the polypeptide having an amino acid sequence as shown in SEQ ID NO: 140 rather than with any other group, and/or comprises at least two B3 domains (in particular four B3 domains) as outlined herein above), and/or has DNA binding activity, and/or has at least 70% sequence identity to SEQ ID NO: 140.

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.).

Nucleic acids encoding RTF 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 RTF polypeptide-encoding nucleic acid is from a plant, further preferably from a from a dicotyledonous plant, further preferably from the family Brassicaceae, more preferably from the genus Arabidopsis, most preferably from Arabidopsis thaliana.

In another embodiment the present invention extends to recombinant chromosomal DNA comprising a nucleic acid sequence useful in the methods, constructs, plants, harvestable parts and products of the invention, wherein said nucleic acid is present in the chromosomal DNA as a result of recombinant methods, i.e. said nucleic acid is not in the chromosomal DNA in its native surrounding. Said recombinant chromosomal DNA may be a chromosome of native origin, with said nucleic acid inserted by recombinant means, or it may be a mini-chromosome or a non-native chromosomal structure, e.g. or an artificial chromosome. The nature of the chromosomal DNA may vary, as long it allows for stable passing on to successive generations of the recombinant nucleic acid useful in the methods, constructs, plants, harvestable parts and products of the invention, and allows for expression of said nucleic acid in a living plant cell resulting in increased yield or increased yield related traits of the plant cell or a plant comprising the plant cell.

In a further embodiment the recombinant chromosomal DNA of the invention is comprised in a plant cell. DNA comprised within a cell, particularly a cell with cell walls like a plant cell, is better protected from degradation than a bare nucleic acid sequence. The same holds true for a DNA construct comprised in a host cell, for example 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 seed yield relative to control plants. The terms “yield” and “seed yield” are described in more detail in the “definitions” section herein.

Reference herein to enhanced yield-related traits is taken to mean an increase 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 roots such as taproots, stems, seeds, and performance of the methods of the invention results in plants having increased seed yield relative to the seed yield of control plants, and/or increased stem biomass relative to the stem biomass of control plants, and/or increased root biomass relative to the root biomass and/or increased beet biomass relative to the beet biomass and/or increased tuber biomass relative to the tuber biomass of control plants. Moreover, it is particularly contemplated that the sugar content (in particular the sucrose content) in the stem (in particular of sugar cane plants) and/or in the belowground parts, in particular in roots including taproots, tubers and/or

beets (in particular in sugar beets) is increased relative to the sugar content (in particular the sucrose content) in the corresponding part(s) of the control plant.

The present invention provides a method for increasing yield-related traits, in particular increased biomass (in particular increased aboveground and increased root biomass), and improved early vigor, relative to control plants, which method comprises modulating expression in a plant of a nucleic acid encoding a RTF polypeptide as defined herein. Preferably, said increased yield related traits obtained under non stress conditions.

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 RTF 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 relative to control plants grown under comparable conditions. Therefore, according to the present invention, there is provided a method for increasing yield 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 RTF polypeptide.

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

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

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

The invention also provides genetic constructs and vectors to facilitate introduction and/or expression in plants of nucleic acids encoding RTF polypeptides. The gene constructs may be inserted into vectors, which may be commercially available, suitable for transforming into plants 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 RTF 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 RTF polypeptide is as defined above. The term "control sequence" and "termination sequence" are as defined herein.

The invention furthermore provides plants 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.

Plants are transformed with a vector comprising any of the nucleic acids described above. The skilled artisan is well aware of the genetic elements that must be present on the vector 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) in the vectors of the invention. In one embodiment the plants of the invention are transformed with an expression cassette comprising any of the nucleic acids described above. The skilled artisan is well aware of the genetic elements that must be present on the expression cassette in order to successfully transform, select and propagate host cells containing the sequence of interest. In the expression cassettes of the invention the sequence of interest is operably linked to one or more control sequences (at least to a promoter). The promoter in such an expression cassette may be a non-native promoter to the nucleic acid described above, i.e. a promoter not regulating the expression of said nucleic acid in its native surrounding.

In one embodiment the terms expression cassettes of the invention, genetic construct and constructs of the invention are used exchangeably.

In a further embodiment the expression cassettes of the invention confer increased yield or yield related traits(s) to a living plant cell when they have been introduced into said plant cell and result in expression of the nucleic acid as defined above, comprised in the expression cassette(s).

The promoter in such expression cassettes may be a non-native promoter to the nucleic acid described above, i.e. a promoter not regulating the expression of said nucleic acid in its native surrounding.

The expression cassettes of the invention may be comprised in a host cell, plant cell, seed, ag-

ricultural product or plant.

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. Preferably the constitutive promoter is a ubiquitous constitutive promoter of medium strength, in particular, the GOS2 promoter. See the "Definitions" section herein for definitions of the various promoter types. Also useful in the methods, constructs, plants, harvestable parts and products of the invention is a root-specific promoter.

It should be clear that the applicability of the present invention is not restricted to the RTF polypeptide-encoding nucleic acid represented by SEQ ID NO: 139, nor is the applicability of the invention restricted to expression of a RTF polypeptide-encoding nucleic acid when driven by a constitutive promoter, or when driven by a root-specific promoter.

The constitutive promoter is preferably a medium strength promoter. 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: 167, most preferably the constitutive promoter is as represented by SEQ ID NO:167. See the "Definitions" section herein for further examples of constitutive promoters.

In a preferred embodiment, the polynucleotide encoding the RTF polypeptide as used in the plants, constructs and methods of the present invention is linked to a promoter which allows for the expression, preferably the strongest expression in the aboveground parts of the plant as compared to the expression in other parts of the plant. This applies, in particular, if the plant is a monocot. As set forth elsewhere herein, preferred monocots are maize, wheat, rice, or sugarcane. In another preferred embodiment of the present invention, the polynucleotide encoding the RTF polypeptide as used in the plants, constructs and methods of the present invention is preferably linked to a promoter which allows for the expression, preferably the strongest expression in the belowground parts of the plant as compared to the expression in other parts of the plant. This applies, in particular, if the plant is a dicot. Preferred dicots are sugar beet and potato. For example, if the plant is a sugar beet, the promoter, preferably, allows for the strongest expression in the taproot as compared to the expression in other parts of the plant. In one embodiment the promoter used in for expression in sugar beets is, preferably a root specific, more preferably a taproot or beet specific promoter.

Optionally, one or more terminator sequences may be used in the construct introduced into a plant. Preferably, the construct comprises an expression cassette comprising a GOS2 promot-

er, substantially similar to SEQ ID NO: 167, operably linked to the nucleic acid encoding the RTF polypeptide. More preferably, the construct comprises a zein terminator (t-zein) linked to the 3' end of the HAB1 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 RTF polypeptide is by introducing and expressing in a plant a nucleic acid encoding a RTF 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 RTF polypeptide as defined hereinabove.

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

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

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

Particularly increased yield related traits are increased biomass (in particular increased aboveground and increased root biomass), and improved early vigor. Preferably, said increased yield related traits obtained under non stress conditions.

Cultivating the plant cell under conditions promoting plant growth and development, may or may not include regeneration and or growth to maturity.

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

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 by transformation. The term "transformation" is described in more detail in the "definitions" section herein.

In one embodiment the present invention clearly 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 parts thereof comprise a nucleic acid transgene encoding a RTF polypeptide as defined above. 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.

The present invention also extends in another embodiment to transgenic plant cells and seed comprising the nucleic acid molecule of the invention in a plant expression cassette or a plant expression construct.

In a further embodiment the seed of the invention recombinantly comprise the expression cassettes of the invention, the (expression) constructs of the invention, the nucleic acids described above and/or the proteins encoded by the nucleic acids as described above.

A further embodiment of the present invention extends to plant cells comprising the nucleic acid as described above in a recombinant plant expression cassette.

In yet another embodiment the plant cells of the invention are non-propagative cells, e.g. the cells can not be used to regenerate a whole plant from this cell as a whole using standard cell culture techniques, this meaning cell culture methods but excluding in-vitro nuclear, organelle or chromosome transfer methods. 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. One example are plant cells that do not sustain themselves through photosynthesis by synthesizing carbohydrate and protein from such inorganic substances as water, carbon dioxide and mineral salt.

In another embodiment the plant cells of the invention are plant cells that do not sustain themselves through photosynthesis by synthesizing carbohydrate and protein from such inorganic substances as water, carbon dioxide and mineral salt, i.e. they may be deemed non-plant variety. In a further embodiment the plant cells of the invention are non-plant variety and non-propagative.

The invention also includes host cells containing an isolated nucleic acid encoding a RTF polypeptide as defined hereinabove. Host cells of the invention may be any cell selected from the

group consisting of bacterial cells, such as E.coli or Agrobacterium species cells, yeast cells, fungal, algal or cyanobacterial cells or plant cells. In one embodiment host cells according to the invention are plant cells, yeasts, bacteria or fungi. Host plants for the nucleic acids or the vector used in the method according to the invention, the expression cassette or construct or vector are, in principle, advantageously all plants, which are capable of synthesizing the polypeptides used in the inventive method.

In one embodiment the plant cells of the invention overexpress the nucleic acid molecule of the invention.

The invention also includes methods for the production of 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, including seeds, of these plants. In a further embodiment the methods comprises steps a) growing the plants of the invention, b) removing the harvestable parts as defined above from the plants and c) producing said product from or by the harvestable parts of the invention.

Examples of such methods would be growing corn plants of the invention, harvesting the corn cobs and remove the kernels. These may be used as feedstuff or processed to starch and oil as agricultural products.

The product may be produced at the site where the plant has been grown, or the plants or parts thereof may be removed from the site where the plants have been grown to produce the product. Typically, the plant is grown, the desired harvestable parts are removed from the plant, if feasible in repeated cycles, and the product made from the harvestable parts of the plant. The step of growing the plant may be performed only once each time the methods of the invention is performed, while allowing repeated times the steps of product production e.g. by repeated removal of harvestable parts of the plants of the invention and if necessary further processing of these parts to arrive at the product. It is also possible that the step of growing the plants of the invention is repeated and plants or harvestable parts are stored until the production of the product is then performed once for the accumulated plants or plant parts. Also, the steps of growing the plants and producing the product may be performed with an overlap in time, even simultaneously to a large extend, or sequentially. Generally the plants are grown for some time before the product is produced.

Advantageously the methods of the invention are more efficient than the known methods, because the plants of the invention have increased yield and/or stress tolerance to an environmental stress compared to a control plant used in comparable methods.

In one embodiment the products produced by said 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. Foodstuffs are regarded as compositions used for nutrition or for supplementing nutrition. Animal feedstuffs and animal feed supplements, in particular, are regarded as foodstuffs.

In another embodiment the inventive methods for the 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.

It is possible that a plant product consists of one ore more agricultural products to a large extent.

In yet another embodiment the polynucleotide sequences or the polypeptide sequences of the invention are comprised in an agricultural product.
in a further embodiment the nucleic acid sequences and protein sequences of the invention may be used as product markers, for example for an agricultural product 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 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 one preferred embodiment the plants of the invention or 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. Especially preferred plants are sugar beet and sugarcane.

In one embodiment of the present invention the plants of the invention and the plants used in the methods of the invention are sugarbeet plants with increased biomass and/or increased sugar content of the beets. In another embodiment of the present invention the plants of the invention and the plants used in the methods of the invention are sugarcane plants with increased biomass and/or increased sugar content of the stems.

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 com-

prise a recombinant nucleic acid encoding a RTF polypeptide. The invention furthermore relates to products derived or produced, preferably directly derived or directly produced, from a harvestable part of such a plant, such as dry pellets or powders, oil, fat and fatty acids, starch or proteins. In one embodiment the product comprises a recombinant nucleic acid encoding a RTF polypeptide and/or a recombinant RTF polypeptide for example as an indicator of the particular quality of the product.

The present invention also encompasses use of nucleic acids encoding RTF polypeptides as described herein and use of these RTF polypeptides in enhancing any of the aforementioned yield-related traits in plants. For example, nucleic acids encoding RTF polypeptide described herein, or the RTF polypeptides themselves, may find use in breeding programmes in which a DNA marker is identified which may be genetically linked to a RTF polypeptide-encoding gene. The nucleic acids/genes, or the RTF 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 hereinabove in the methods of the invention. Furthermore, allelic variants of a RTF polypeptide-encoding nucleic acid/gene may find use in marker-assisted breeding programmes. Nucleic acids encoding RTF 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.

Any comparison to determine sequence identity percentages is, preferably performed

- in the case of a comparison of nucleic acids over the entire coding region of SEQ ID NO: 139, or
- in the case of a comparison of polypeptide sequences over the entire length of SEQ ID NO: 140.

For example, a sequence identity of 50% sequence identity in this embodiment means that over the entire coding region of SEQ ID NO: 139, 50 percent of all bases are identical between the sequence of SEQ ID NO: 139 and the related sequence. Similarly, in this embodiment a polypeptide sequence is 50 % identical to the polypeptide sequence of SEQ ID NO: 140, when 50 percent of the amino acids residues of the sequence as represented in SEQ ID NO: 140, are found in the polypeptide tested when comparing from the starting methionine to the end of the sequence of SEQ ID NO: 140.

In a further embodiment the nucleic acid sequence employed in methods, constructs, plants, harvestable parts and products of the invention are those sequences that are not the polynucleotides encoding the proteins selected from the group consisting of the proteins listed in Table A3, and those of at least 60, 70, 75, 80, 85, 90, 93, 95, 98 or 99% nucleotide identity when optimally aligned to the sequences encoding the proteins listed in Table A3.

C-4. BP1 (Bigger plant 1) polypeptide

Surprisingly, it has now been found that modulating expression in a plant of a nucleic acid encoding a BP1 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 BP1 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 enhancing 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 BP1 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 BP1 polypeptide is by introducing and expressing in a plant a nucleic acid encoding a BP1 polypeptide. Preferably, said nucleic acid is over-expressed.

Any reference hereinafter in section C-4 to a "protein useful in the methods of the invention" is taken to mean a BP1 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 BP1 polypeptide. In one embodiment any reference to a protein or nucleic acid "useful in the methods of the invention" is to be understood to mean proteins or nucleic acids "useful in the methods, constructs, plants, harvestable parts and products of the invention". 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 "*BP1* nucleic acid" or "*BP1* gene".

A "BP1 polypeptide" as defined herein, preferably, refers to a polypeptide comprising one or more of the following motifs:

(i) Motif 1-4:
LNQ[DG]SXXND[EV]X[NS]DX[QP]G[HQ]X[GN]H[LP]EXXKX[DE][QE][VA][GE]VXE[DE]X[MI][TA][AP]DV[KN]LS[VA]CRDTG[NE] (SEQ ID NO: 276),

(ii) Motif 2-4:
L[WR]RDYXD[LV][LV][QK][ED][TN]EXK[KR][KR]XLXSX[KN][RK][RT][KS]L[AV]LL[AS]EVKFL[RQ][RK]K[YL]XSF[AKLP]K[GN][GDN]SQ[QK] (SEQ ID NO: 277),

(iii) Motif 3-4:
[DE][DG]KRX[VI][PS]WQD[RQ]XALK (SEQ ID NO: 278),

(iv) Motif 4-4 as disclosed as SEQ ID NO: 279;
(v) Motif 5-4 as disclosed as SEQ ID NO: 280;
(vi) Motif 6-4 as disclosed as SEQ ID NO: 281;

"X", preferably, represents any amino acid. Particularly preferred amino acid residues for the amino acids indicated with "X" are given in SEQ ID NO: 279 and SEQ ID NO: 292 for Motif 1-4, in SEQ ID NO: 280 and SEQ ID NO: 293 for Motif 2-4, and in SEQ ID NO: 281 and SEQ ID NO: 294 for Motif 3-4. Accordingly, in a preferred embodiment of the present invention, Motif 1-4 has a sequence as shown in SEQ ID NO: 279, Motif 2-4 has a sequence as shown in SEQ ID NO: 280, and Motif 3-4 has a sequence as shown in SEQ ID NO: 281. In a even more preferred embodiment motif 2-4 has a sequence starting with amino acid 40 up to amino acid 88 in SEQ ID NO: 171, motif 1-4 has a sequence starting with amino acid 129 up to amino acid 178 in SEQ ID NO: 171, and motif 3-4 has a sequence starting with amino acid 183 up to amino acid 197 in SEQ ID NO: 171.

The sequence of Motif 1-4 is also shown in SEQ ID NO: 289. The sequence of Motif 2-4 is also shown in SEQ ID NO: 290. The sequence of Motif 3-4 is also shown in SEQ ID NO: 291.

In a preferred embodiment, the BP1 polypeptide as set forth in the context of the present invention comprises:

a) all of the following motifs:
(i) Motif 1-4:
LNQ[DG]SXXND[EV]X[NS]DX[QP]G[HQ]X[GN]H[LP]EXXKX[DE][QE][VA][GE]V XE[DE]X[MI][TA][AP]DV[KN]LS[VA]CRDTG[NE] (SEQ ID NO: 276),
(ii) Motif 2-4:
L[WR]RDYXD[LV][LV][QK][ED][TN]EXK[KR][KR]XLXSX[KN][RK][RT][KS]L[AV] LL[AS]EVKFL[RQ][RK]K[YL]XSF[AKLP]K[GN][GDN]SQ[QK] (SEQ ID NO: 277), and
(iii) Motif 3-4: [DE][DG]KRX[VI][PS]WQD[RQ]XALK (SEQ ID NO: 278);
(iv) Motif 4-4 as disclosed as SEQ ID NO: 279;
(v) Motif 5-4 as disclosed as SEQ ID NO: 280;
(vi) Motif 6-4 as disclosed as SEQ ID NO: 281;
b) any two of the Motifs 1-4 to 6-4, preferably any two of Motifs 4-4 to motif 6-4 as defined in a) above; or
c) any three of the Motifs 1-4 to 6-4, preferably all three of Motifs 4-4 to motif 6-4 as defined in a) above; or
d) any one of the Motifs 1-4 to 6-4, preferably any two of Motifs 4-4 to motif 6-4 as defined in a) above.

Alternatively or additionally the "BP1 polypeptide" as defined herein, preferably, refers to any polypeptide comprising one or more of the following motifs:

(i) a motif comprising in increasing order of preference at least 70%, 75%, 80%, 85%, 90%, 95%, or more sequence identity to Motif 1-4 or 4-4, preferably, as represented by SEQ ID NO: 276 or by SEQ ID NO: 279, more preferably when compared to motif 4-4,

(ii) a motif comprising in increasing order of preference at least 70%, 75%, 80%, 85%, 90%, 95%, or more sequence identity to Motif 2-4 or 5-4, preferably, as represented by SEQ ID NO: 277 or by SEQ ID NO: 280, more preferably when compared to motif 5-4,

(iii) a motif comprising in increasing order of preference at least 70%, 75%, 80%, 85%, 90%, 95%, or more sequence identity to Motif 3-4 or 6-4, preferably, as represented by SEQ ID NO: 278 or by SEQ ID NO: 281, more preferably when compared to motif 6-4.

Preferred combinations of motifs are given herein above.

Preferably, the BP1 polypeptide as used in the context of the present invention is selected from the group consisting of:

(i) a polypeptide comprising a sequence, or consisting of a sequence as shown in SEQ ID NO: 171,

(ii) a polypeptide which has, in an increasing order of preference, at least 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: 171,

(iii) a polypeptide encoded by a polynucleotide which hybridizes under stringent conditions to a polynucleotide having a sequence as shown in SEQ ID NO: 170, or with a complementary sequence of such a polynucleotide having a sequence as shown in SEQ ID NO: 170,

(iv) a polypeptide with the biological activity of the polypeptide as shown in SEQ ID NO: 171or substantially the same biological activity of the polypeptide as shown in SEQ ID NO: 171; and

(v) v) any combination of i) to iv) above

Preferably, the BP1 polypeptide comprises the motifs, combinations of motifs as set forth herein above.

The term "BP1" or "BP1 polypeptide" as used herein also intends to include homologues as defined hereunder of "BP1 polypeptide".

Motifs 1-4 to 3-4 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) using the BP1 polypeptides from Oryza sativa (Os09g25410, SEQ ID NO: 171), Panicum virgatum (TC30704, SEQ ID NO: 239), Sorghum bicolor (Sb02g024920, SEQ ID NO: 243), and Zea mays (GRMZM2G093731_T02, SEQ ID NO: 267), see also Table A4. Motifs 4-4, 5-4 and 6-4 were derived manually. The motifs were adjusted in order to bring them in compliance with SEQ ID NO: 171. 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.

More preferably, the BP1 polypeptide comprises in increasing order of preference, at least 1, at least 2, or all 3 motifs of either motifs 1-4 to 3-4 or motifs 4-4 to 6-4. Thus, the BP1 polypeptide preferably comprises Motif 4-4, Motif 5-4 or Motif 6-4. More preferably, the BP1 polypeptide comprises Motifs 4-4 and 5-4, Motifs 5-4 and 6-4 or Motifs 4-4 and 6-4. Most preferably, the BP1 polypeptide comprises Motifs 4-4, 5-4 and 6-4.

Additionally or alternatively, the BP1 polypeptide as set forth herein or the homologue thereof, preferably, 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 represented by SEQ ID NO: 171. Preferably, the BP1 protein or homologue protein therof comprises any one or more of the conserved motifs - i.e. of motifs 1-4, 2-4 or 3-4 or motifs 4-4 to 6-4, or the variants having in increasing order of preference at least 70%, 75%, 80%, 85%, 90%, 95%, or more sequence identity to Motif 1-4, 2-4 or 3-4 or or motifs 4-4 to 6-4, preferably to motifs 4-4 to 6-4 - as outlined above. Preferred combinations of motifs are given herein above.

In another embodiment, a “BP1 polypeptide” as defined herein, preferably, refers to a BP-like polypeptide comprising one or more of the following motifs: Motif 7-4 as disclosed as SEQ ID NO: 282, motif 8-4 as disclosed as SEQ ID NO: 283, motif 9-4 as disclosed as SEQ ID NO: 284.

The overall sequence identity is, preferably, 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).

Preferably, the sequence identity level is determined by comparison of the polypeptide sequences over the entire length of the sequence of SEQ ID NO: 171. In another embodiment the sequence identity level of a nucleic acid sequence is determined by comparison of the nucleic

acid sequence over the entire length of the coding sequence of the sequence of SEQ ID NO: 170.

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 BP1 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: 276 to SEQ ID NO: 278 (Motifs 1-4 to 3-4) or motifs 4-4 to 6-4 as represented by SEQ ID NO: 279 to 281. Moreover, preferably, the motifs in a BP1 polypeptide have 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 motif starting with amino acid 40 up to amino acid 88 in SEQ ID NO: 171, and/or to the motif starting with amino acid 129 up to amino acid 178 in SEQ ID NO: 171, and/or to the motif starting with amino acid 183 up to amino acid 197 in SEQ ID NO: 171

Preferably, the polypeptide sequence which when used in the construction of a phylogenetic tree/circular phylogram, such as the one depicted in Figure 18, clusters with the group of BP1 polypeptides, particularly with the polypeptide comprising the amino acid sequence represented by SEQ ID NO: 171 (see Fig. 18, Os09g25410), rather than with the other groups (such as the "outgroup" in Fig. 18, or the group of BP1-like polypeptides in Fig.3). Preferably, said polypeptide comprises one ore more of :motifs 1-4 to 3-4, or motifs 4-4 to 6-4, preferably motifs 4-4 to 6-4 as outlined above, and/or has at least 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: 171.

Preferably, BP1 polypeptides, when expressed in a monocot plant such as rice, maize, wheat or sugarcane according to the methods of the present invention as outlined in the Examples, give plants having at least one increased yield related trait.

Accordingly, a BP1 polypeptide when expressed in a plant, in particular in a monocot plant such as rice, maize, wheat or sugarcane, preferably, increase at least one of the yield related traits selected from the group consisting of aboveground biomass, root biomass, total seed yield per plant, flowers per panicle, number of filled seeds per plant, increased nitrogen use efficiency and number of thick roots (as compared to a control plant not expressing said BP1 polypeptide). Preferably, said increase of said at least one of the yield related traits is an increase of at least 1%, of at least 2%, more preferably, of at least 3% and, most preferably, of at least 5%. Tools and techniques for measuring whether a yield related traits are increased are described in the Example. Preferably, said increase of said at least one yield related trait is under nitrogen deficiency.

The present invention is illustrated by transforming plants with the nucleic acid sequence represented by SEQ ID NO: 170, encoding the polypeptide sequence of SEQ ID NO: 171. However, performance of the invention is not restricted to these sequences; the methods of the invention may advantageously be performed using any BP1-encoding nucleic acid or BP1 polypeptide as defined herein.

Examples of nucleic acids encoding BP1 polypeptides are given in Table A4 of the Examples section herein. Such nucleic acids are useful in performing the methods of the invention. The amino acid sequences given in Table A4 of the Examples section are example sequences of orthologues and paralogues of the BP1 polypeptide represented by SEQ ID NO: 171, 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: 170 or SEQ ID NO: 171, the second BLAST (back-BLAST) would be against Oryza sativa sequences.

Particularly preferred BP1 polypeptide are selected from the BP1 polypeptide from Oryza sativa having an amino acid sequence as shown SEQ ID NO: 171 (see Table A4, Os09g25410), from Panicum virgatum having an amino acid sequence as shown SEQ ID NO: 239 (TC30704), from Sorghum bicolor having an amino acid sequence as shown SEQ ID NO: 243 (Sb02g024920), and from Zea mays having an amino acid sequence as shown SEQ ID NO: 267 (GRMZM2G093731_T02).

In another preferred embodiment the nucleic acid molecules useful in the methods, uses, transgenic plants, host cells, expression cassettes, vectors and/or products of the invention are nucleic acid molecules encoding the BP1 polypeptide selected from the group consisting of

(i) a nucleic acid represented by SEQ ID NO: 170;

(ii) the complement of a nucleic acid represented by SEQ ID NO: 170;

(iii) a nucleic acid encoding a BP1 polypeptide having in increasing order of preference at least 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: 170,

(iv) a nucleic acid molecule which hybridizes with a nucleic acid molecule of (i) to (iii) under high stringency hybridization conditions.

The polypeptide encoded by said nucleic acid, preferably, comprises 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 1-4 to 3-

4 preferably one or more of the motifs 4-4 to 6-4 as outlined elsewhere herein (e.g. shown in SEQ ID NO: 276 to SEQ ID NO: 278). Moreover, it shall preferably confer enhanced yield-related traits relative to control plants.

Preferably, the sequence identity level of a nucleic acid sequence is determined by comparison of the nucleic acid sequence over the entire length of the coding sequence of the sequence of SEQ ID NO: 170.

In another preferred embodiment BP1 polypeptides useful in the methods, uses, transgenic plants, host cells, expression cassettes, vectors and/or products of the invention are polypeptides selected from the group consisting of:

(i) a polypeptide having an amino acid sequence represented by SEQ ID NO: 171;

(ii) a polypeptide having an amino acid sequence having, in increasing order of preference, at least 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: 171, 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 1-4 to 3-4 as outlined above (e.g. having a sequence given in SEQ ID NO: 276 to SEQ ID NO: 278), or more prefably one or more of the motifs 4-4 to 6-4 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 A4 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 A4 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 BP1 polypeptides, nucleic acids hybridising to nucleic acids encoding BP1 polypeptides, splice variants of nucleic acids encoding BP1 polypeptides, allelic variants of

nucleic acids encoding BP1 polypeptides and variants of nucleic acids encoding BP1 polypeptides obtained by gene shuffling. The terms hybridising sequence, splice variant, allelic variant and gene shuffling are as described herein.

In one embodiment of the present invention the function of the nucleic acid sequences of the invention is to confer information for a protein that increases yield or yield related traits, when a nucleic acid sequence of the invention is transcribed and translated in a living plant cell.

Nucleic acids encoding BP1 polypeptides need not be full-length nucleic acids, since performance of the methods of the invention does not 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 A4 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 A4 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, constructs, plants, harvestable parts and productsof the invention, encode a BP1 polypeptide as defined herein, and have substantially the same biological activity as the amino acid sequences given in Table A4 of the Examples section. Preferably, the portion is a portion of any one of the nucleic acids given in Table A4 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 A4 of the Examples section. Preferably the portion is at least 500, 550, 600, 650, 700, 750, 800, 850, or 909 consecutive nucleotides in length, the consecutive nucleotides being of any one of the nucleic acid sequences given in Table A4 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 A4 of the Examples section. Most preferably the portion is a portion of the nucleic acid of SEQ ID NO: 170. Preferably, the portion encodes a fragment of an amino acid sequence which, when used in the construction of a phylogenetic tree/ circular phylogram, such as the one depicted in Figure 18, clusters with the group of BP1 polypeptides comprising the amino acid sequence represented by SEQ ID NO: 171 (and, thus, preferably, with the BP1-proteins in Fig. 18) rather than with any other group, and/or comprises one ore more of motifs 1-4 to 3-4 preferably one or more of the motifs 4-4 to 6-4 as outlined elsewhere herein, and/or has at least 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: 171.

Another nucleic acid variant useful in the methods, constructs, plants, harvestable parts and products of the invention is a nucleic acid capable of hybridising, under reduced stringency conditions, preferably under stringent conditions, with a nucleic acid encoding a BP1 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 any one of the nucleic acids given in Table A4 of the Examples section, or comprising introducing and expressing in a plant a nucleic acid capable of hybridising to a nucleic acid encoding an orthologue, paralogue or homologue of any of the nucleic acid sequences given in Table A4 of the Examples section.

Hybridising sequences useful in the methods, constructs, plants, harvestable parts and products of the invention encode a BP1 polypeptide as defined herein, having substantially the same biological activity as the amino acid sequences given in Table A4 of the Examples section. Preferably, the hybridising sequence is capable of hybridising to the complement of any one of the nucleic acids given in Table A4 of the Examples section, or to a portion of any of these sequences, a portion being as defined above, 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 A4 of the Examples section. Most preferably, the hybridising sequence is capable of hybridising to the complement of a nucleic acid as represented by SEQ ID NO: 170 or to a portion thereof.

Preferably, the hybridising sequence encodes a polypeptide with an amino acid sequence which, when full-length and when used in the construction of a phylogenetic tree/ circular phylogram, such as the one depicted in Figure 18, clusters with the group of BP1 polypeptides comprising the amino acid sequence represented by SEQ ID NO: 171 (and, thus, preferably, with the BP1-proteins in Fig. 18) rather than with any other group, and/or comprises one ore more of motifs 1-4 to 3-4 preferably one or more of the motifs 4-4 to 6-4 as outlined elsewhere herein, and/or has at least 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: 171.

In one embodiment the hybridising sequence is capable of hybridising to the complement of a nucleic acid as represented by SEQ ID NO: 170 or to a portion thereof under conditions of medium or high stringency, preferably high stringency as defined above. In another embodiment

the hybridising sequence is capable of hybridising to the complement of a nucleic acid as represented by SEQ ID NO: 170 under stringent conditions.

Another nucleic acid variant useful in the methods, constructs, plants, harvestable parts and products of the invention is a splice variant encoding a BP1 polypeptide as defined hereinabove, a splice variant being 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 splice variant of any one of the nucleic acid sequences given in Table A4 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 A4 of the Examples section.

Preferred splice variants are splice variants of a nucleic acid represented by SEQ ID NO: 170, or a splice variant of a nucleic acid encoding an orthologue or paralogue of SEQ ID NO: 171. Preferably, the amino acid sequence encoded by the splice variant, when used in the construction of a phylogenetic tree/ circular phylogram, such as the one depicted in Figure 18, clusters with the group of BP1 polypeptides comprising the amino acid sequence represented by SEQ ID NO: 171 (and, thus, preferably, with the BP1-proteins in Fig. 18) rather than with any other group, comprises one or more of motifs 1-4 to 3-4, preferably one or more of the motifs 4-4 to 6-4 as outlined elsewhere herein, and/or has at least 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: 171.

Another nucleic acid variant useful in performing the methods of the invention is an allelic variant of a nucleic acid encoding a BP1 polypeptide as defined hereinabove, an allelic variant being 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 an allelic variant of any one of the nucleic acids given in Table A4 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 A4 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 BP1 polypeptide of SEQ ID NO: 171 and any of the amino acids depicted in Table A4 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: 170 or an allelic var-

iant of a nucleic acid encoding an orthologue or paralogue of SEQ ID NO: 171. Preferably, the amino acid sequence encoded by the allelic variant, when used in the construction of a phylogenetic tree/ circular phylogram, such as the one depicted in Figure 18, clusters with the group of BP1 polypeptides comprising the amino acid sequence represented by SEQ ID NO: 171 (and, thus, preferably, with the BP1-proteins in Fig. 18) rather than with any other group, comprises one or more of motifs 1-4 to 3-4, preferably one or more of the motifs 4-4 to 6-4 as outlined elsewhere herein, and/or has at least 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: 171.

Gene shuffling or directed evolution may also be used to generate variants of nucleic acids encoding BP1 polypeptides as defined above; the term “gene shuffling” being 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 variant of any one of the nucleic acid sequences given in Table A4 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 A4 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, when used in the construction of a phylogenetic tree/ circular phylogram, such as the one depicted in Figure 18, clusters with the group of BP1 polypeptides comprising the amino acid sequence represented by SEQ ID NO: 171 (and, thus, preferably, with the BP1-proteins in Fig. 18) rather than with any other group, and/or comprises one ore more of motifs 1-4 to 3-4, preferably one or more of the motifs 4-4 to 6-4 as outlined elsewhere herein, and/or has at least 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: 171.

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.).

Nucleic acids encoding BP1 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 BP1 polypeptide-encoding nucleic acid is from a plant, further preferably from a monocotyledonous plant, more preferably from the

family Poaceae, more preferably from the genus Oryza most preferably the nucleic acid is from Oryza sativa.

In another embodiment the present invention extends to recombinant chromosomal DNA comprising a nucleic acid sequence useful in the methods, constructs, plants, harvestable parts and products of the invention, wherein said nucleic acid is present in the chromosomal DNA as a result of recombinant methods, i.e. said nucleic acid is not in the chromosomal DNA in its native surrounding. Said recombinant chromosomal DNA may be a chromosome of native origin, with said nucleic acid inserted by recombinant means, or it may be a mini-chromosome or a non-native chromosomal structure, e.g. or an artificial chromosome. The nature of the chromosomal DNA may vary, as long it allows for stable passing on to successive generations of the recombinant nucleic acid useful in the methods of the invention, and allows for expression of said nucleic acid in a living plant cell resulting in increased yield or increased yield related traits of the plant cell or a plant comprising the plant cell.
In a further embodiment the recombinant chromosomal DNA of the invention is comprised in a plant cell. DNA comprised within a cell, particularly a cell with cell walls like a plant cell, is better protected from degradation than a bare nucleic acid sequence. The same holds true for a DNA construct comprised in a host cell, for example 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 seed yield or increased biomass relative to control plants. The terms “yield” and “seed yield” are described in more detail in the “definitions” section herein.

Reference herein to enhanced yield-related traits is taken to mean an increase seed yield 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 roots such as taproots, stems, seeds, and performance of the methods of the invention results in plants having increased seed yield relative to the seed yield of control plants, and/or increased stem biomass relative to the stem biomass of control plants, and/or increased root biomass relative to the root biomass and/or increased beet biomass relative to the beet biomass and/or increased tuber biomass relative to the tuber biomass of control plants. Moreover, it is particularly contemplated that the sugar content (in particular the sucrose content) in the stem (in particular of sugar cane plants) and/or in the belowground parts, in particular in roots including taproots, tubers and/or beets (in particular in sugar beets) is increased relative to the sugar content (in particular the sucrose content) in corresponding part(s)of the control plant.
The present invention provides a method for increasing yield related traits, especially seed yield and/o of plants, relative to control plants, which method comprises modulating expression in a plant of a nucleic acid encoding a BP1 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 BP1 polypeptide as defined herein. Preferably, by modulating expression of the BP1 polypeptide at least one of the yield-related trait selected from aboveground biomass, root biomass, root thickness, root length is increased. In particular, the increased trait is aboveground or root biomass. Preferably, the yield-related traits are increased under nitrogen limiting conditions, in particular under nitrogen deficient conditions.

Performance of the methods of the invention gives plants grown under non-stress conditions or under mild drought conditions increased yield relative to control plants grown under comparable conditions. Therefore, according to the present invention, there is provided a method for increasing yield 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 BP1 polypeptide.

Performance of the methods of the invention gives plants grown under conditions of nutrient deficiency, particularly under conditions of nitrogen deficiency, increased yield relative to control plants grown under comparable conditions. Therefore, according to the present invention, there is provided a method for increasing yield in plants grown under conditions of nutrient deficiency, which method comprises modulating expression in a plant of a nucleic acid encoding a BP1 polypeptide.

Performance of the methods of the invention gives plants grown under conditions of salt stress, increased yield relative to control plants grown under comparable conditions. Therefore, according to the present invention, there is provided a method for increasing yield in plants grown under conditions of salt stress, which method comprises modulating expression in a plant of a nucleic acid encoding a BP1 polypeptide.

The invention also provides genetic constructs and vectors to facilitate introduction and/or expression in plants of nucleic acids encoding BP1 polypeptides. The gene constructs may be inserted into vectors, which may be commercially available, suitable for transforming into plants 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 BP1 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 BP1 polypeptide is as defined above. The term "control sequence" and "termination sequence" are as defined herein.

The invention furthermore provides plants 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.

Plants are transformed with a vector comprising any of the nucleic acids described above. The skilled artisan is well aware of the genetic elements that must be present on the vector 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) in the vectors of the invention.

In one embodiment the plants of the invention are transformed with an expression cassette comprising any of the nucleic acids described above. The skilled artisan is well aware of the genetic elements that must be present on the expression cassette in order to successfully transform, select and propagate host cells containing the sequence of interest. In the expression cassettes of the invention the sequence of interest is operably linked to one or more control sequences (at least to a promoter). The promoter in such an expression cassette may be a non-native promoter to the nucleic acid described above, i.e. a promoter not regulating the expression of said nucleic acid in its native surrounding.
In one embodiment the terms expression cassettes of the invention, genetic construct and constructs of the invention are used exchangeably.
In a further embodiment the expression cassettes of the invention confer increased yield or yield related traits(s) to a living plant cell when they have been introduced into said plant cell and result in expression of the nucleic acid as defined above, comprised in the expression cassette(s). The promoter in such expression cassettes may be a non-native promoter to the nucleic acid described above, i.e. a promoter not regulating the expression of said nucleic acid in its native surrounding.
The expression cassettes of the invention may be comprised in a host cell, plant cell, seed, agricultural product or plant.

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. Preferably the constitutive promoter is a ubiquitous constitutive promoter of medium strength. See the "Definitions" section herein for definitions of the various promoter types.

It should be clear that the applicability of the present invention is not restricted to the BP1 polypeptide-encoding nucleic acid represented by SEQ ID NO: 170, nor is the applicability of the invention restricted to expression of a BP1 polypeptide-encoding nucleic acid when driven by a constitutive promoter.

The constitutive promoter is preferably a medium strength promoter. 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: 285, most preferably the constitutive promoter is as represented by SEQ ID NO: 285. See the "Definitions" section herein for further examples of constitutive promoters.

According to another preferred feature of the invention, the nucleic acid encoding a BP1 polypeptide is operably linked to a root-specific promoter.

In a preferred embodiment, the polynucleotide encoding the BP1 polypeptide as used in the plants, constructs and methods of the present invention is linked to a promoter which allows for the expression, preferably the strongest expression in the aboveground parts of the plant as compared to the expression in other parts of the plant. This applies, in particular, if the plant is a monocot. As set forth elsewhere herein, preferred monocots are maize, wheat, rice, or sugarcane. In another preferred embodiment of the present invention, the polynucleotide encoding the BP1 polypeptide as used in the plants, constructs and methods of the present invention is preferably linked to a promoter which allows for the expression, preferably the strongest expression in the belowground parts of the plant as compared to the expression in other parts of the plant. This applies, in particular, if the plant is a dicot. Preferred dicots are sugar beet and potato. For example, if the plant is a sugar beet, the promoter, preferably, allows for the strongest expression in the taproot as compared to the expression in other parts of the plant. In one embodiment the promoter used in for expression in sugar beets is, preferably a root specific, more preferably a taproot or beet specific promoter.

Optionally, one or more terminator sequences may be used in the construct introduced into a plant. Preferably, the construct comprises an expression cassette comprising a GOS2 promoter, substantially similar to SEQ ID NO: 285 operably linked to the nucleic acid encoding the BP1 polypeptide. More preferably, the construct comprises a zein terminator (t-zein) linked to the 3' end of the coding sequence for the BP1 polypeptide. Most preferably, the expression cassette comprises a sequence having in increasing order of preference at least 95%, at least 96%, at least 97%, at least 98%, at least 99% identity to the pGOS2::BP1::t-zein sequence comprised by the expression vector having a sequence as shown in SEQ ID NO: 286 (see also Fig. 19).

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 BP1 polypeptide is by introducing and expressing in a plant a nucleic acid encoding a BP1 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 BP1 polypeptide as defined hereinabove.

More specifically, the present invention provides a method for the production of transgenic plants having enhanced yield-related traits, preferably increased biomass or increased yield, more preferably, enhances yield related traits as described in Example XI-4, which method comprises:

(i) introducing and expressing in a plant or plant cell a BP1 polypeptide-encoding nucleic acid or a genetic construct comprising a BP1 polypeptide-encoding nucleic acid; and

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

Cultivating the plant cell under conditions promoting plant growth and development, may or may not include regeneration and or growth to maturity.

The nucleic acid of (i) may be any of the nucleic acids capable of encoding a BP1 polypeptide as defined herein.

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 by transformation. The term “transformation” is described in more detail in the “definitions” section herein.

In one embodiment the present invention clearly 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 parts thereof comprise a nucleic acid transgene

encoding a BP1 polypeptide as defined above. 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.

The present invention also extends in another embodiment to transgenic plant cells and seed comprising the nucleic acid molecule of the invention in a plant expression cassette or a plant expression construct.

In a further embodiment the seed of the invention recombinantly comprise the expression cassettes of the invention, the (expression) constructs of the invention, the nucleic acids described above and/or the proteins encoded by the nucleic acids as described above.

A further embodiment of the present invention extends to plant cells comprising the nucleic acid as described above in a recombinant plant expression cassette.

In yet another embodiment the plant cells of the invention are non-propagative cells, e.g. the cells can not be used to regenerate a whole plant from this cell as a whole using standard cell culture techniques, this meaning cell culture methods but excluding in-vitro nuclear, organelle or chromosome transfer methods. 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. One example are plant cells that do not sustain themselves through photosynthesis by synthesizing carbohydrate and protein from such inorganic substances as water, carbon dioxide and mineral salt.

In another embodiment the plant cells of the invention are plant cells that do not sustain themselves through photosynthesis by synthesizing carbohydrate and protein from such inorganic substances as water, carbon dioxide and mineral salt, i.e. they may be deemed non-plant variety. In a further embodiment the plant cells of the invention are non-plant variety and non-propagative.

The invention also includes host cells containing an isolated nucleic acid encoding a BP1 polypeptide as defined hereinabove. Host cells of the invention may be any cell selected from the group consisting of bacterial cells, such as E.coli or Agrobacterium species cells, yeast cells, fungal, algal or cyanobacterial cells or plant cells. In one embodiment host cells according to the invention are plant cells, yeasts, bacteria or fungi. Host plants for the nucleic acids or the vector used in the method according to the invention, the expression cassette or construct or vector

are, in principle, advantageously all plants, which are capable of synthesizing the polypeptides used in the inventive method.

In one embodiment the plant cells of the invention overexpress the nucleic acid molecule of the invention, i.e. nucleic acid molecule encoding the BP1 polypeptide.

The invention also includes methods for the production of 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, including seeds, of these plants. In a further embodiment the methods comprises steps a) growing the plants of the invention, b) removing the harvestable parts as defined above from the plants and c) producing said product from or by the harvestable parts of the invention. Examples of such methods would be growing corn plants of the invention, harvesting the corn cobs and remove the kernels. These may be used as feedstuff or processed to starch and oil as agricultural products.
The product may be produced at the site where the plant has been grown, or the plants or parts thereof may be removed from the site where the plants have been grown to produce the product. Typically, the plant is grown, the desired harvestable parts are removed from the plant, if feasible in repeated cycles, and the product made from the harvestable parts of the plant. The step of growing the plant may be performed only once each time the methods of the invention is performed, while allowing repeated times the steps of product production e.g. by repeated removal of harvestable parts of the plants of the invention and if necessary further processing of these parts to arrive at the product. It is also possible that the step of growing the plants of the invention is repeated and plants or harvestable parts are stored until the production of the product is then performed once for the accumulated plants or plant parts. Also, the steps of growing the plants and producing the product may be performed with an overlap in time, even simultaneously to a large extend, or sequentially. Generally the plants are grown for some time before the product is produced.
Advantageously the methods of the invention are more efficient than the known methods, because the plants of the invention have increased yield and/or stress tolerance to an environmental stress compared to a control plant used in comparable methods.
In one embodiment the products produced by said 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. Foodstuffs are regarded as compositions used for nutrition or for supplementing nutrition. Animal feedstuffs and animal feed supplements, in particular, are regarded as foodstuffs.
In another embodiment the inventive methods for the 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.
It is possible that a plant product consists of one ore more agricultural products to a large extent.

In yet another embodiment the polynucleotide sequences or the polypeptide sequences of the invention are comprised in an agricultural product.

in a further embodiment the nucleic acid sequences and protein sequences of the invention may be used as product markers, for example for an agricultural product 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 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 one embodiment the plants of the invention or 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.

In another embodiment of the present invention the plants of the invention and the plants used in the methods of the invention are sugarbeet plants with increased biomass and/or increased sugar content of the beets, or sugar cane plants with increased sugar content.

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 BP1 polypeptide. The invention furthermore relates to products derived or produced, preferably directly derived or directly produced, from a harvestable part of such a plant, such as dry pellets or powders, oil, fat and fatty acids, starch or proteins. In one embodiment the product comprises a recombinant nucleic acid encoding a BP1 polypeptide and/or a recombinant BP1 polypeptide. In one embodiment the product comprises a recombinant nucleic acid encoding a BP1 polypeptide and/or a recombinant BP1 polypeptide for example as an indicator of the particular quality of the product.

The present invention also encompasses use of nucleic acids encoding BP1 polypeptides as described herein and use of these BP1 polypeptides in enhancing any of the aforementioned yield-related traits in plants. For example, nucleic acids encoding BP1 polypeptide described herein, or the BP1 polypeptides themselves, may find use in breeding programmes in which a DNA marker is identified which may be genetically linked to a BP1 polypeptide-encoding gene. The nucleic acids/genes, or the BP1 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 hereinabove in the methods of the invention. Furthermore, allelic variants of a BP1 polypeptide-encoding nucleic acid/gene may find use in marker-assisted breeding programmes. Nucleic acids encoding BP1 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.

In one embodiment any comparison to determine sequence identity percentages is performed

- in the case of a comparison of nucleic acids over the entire coding region of SEQ ID NO: 170, or
- in the case of a comparison of polypeptide sequences over the entire length of SEQ ID NO: 171.

For example, a sequence identity of 50% sequence identity in this embodiment means that over the entire coding region of SEQ ID NO: 170, 50 percent of all bases are identical between the sequence of SEQ ID NO: 170 and the related sequence. Similarly, in this embodiment a polypeptide sequence is 50 % identical to the polypeptide sequence of SEQ ID NO: 171, when 50 percent of the amino acids residues of the sequence as represented in SEQ ID NO: 171, are found in the polypeptide tested when comparing from the starting methionine to the end of the sequence of SEQ ID NO: 171.

In a further embodiment the nucleic acid sequence employed in the invention are those sequences that are not the polynucleotides encoding the proteins selected from the group consisting of the proteins listed in Table A4, and those of at least 60, 70, 75, 80, 85, 90, 93, 95, 98 or 99% nucleotide identity when optimally aligned to the sequences encoding the proteins listed in Table A4.

In one embodiment, the sequence of the nucleic acid encoding said BP1 polypeptide or the sequence of the BP1 polypeptide is, preferably, not the sequence as shown in SEQ ID NO: 1907, 30374, 19675, and/or 48067 as disclosed in US20060123505, and, preferably, not the sequence as shown in SEQ ID NO: 75649 and/or 178132 as disclosed in US20030135870.

Moreover, the sequence of the nucleic acid encoding said BP1 polypeptide or the sequence of the BP1 polypeptide is, preferably, not the sequence as shown in SEQ ID NO: 75649 as dis-

closed in US2004123343, and, preferably, not the sequence as shown in SEQ ID NO: 64503 as disclosed in WO2009/091518, SEQ ID NO: 53534 as disclosed in US2004/172684.

D. ITEMS

In the following, the expression "as defined in claim/item X" is meant to direct the artisan to apply the definition as disclosed in item/claim X. For example, "a nucleic acid as defined in item 1" has to be understood so that the definition of the nucleic acid as in item 1 is to be applied to the nucleic acid. In consequence the term " as defined in item" or " as defined in claim" may be replaced with the corresponding definition of that item or claim, respectively.

D-1. TLP (Tify like protein) polypeptide- Items

The explanations and definitions given herein above in section C-1 apply mutatis mutandis to the following items (in D-1).

Item D-1-1 to D-1-24

1. A method for enhancing yield-related traits in plants relative to control plants, comprising modulating expression, preferably increasing expression, in a plant of a nucleic acid encoding a TLP polypeptide,
   (i) wherein said TLP polypeptide comprises a domain having the Pfam accession number PF06200 and/or a Pfam domain having the accessing number PF09425, preferably both domains, and/or
   (ii) wherein said polypeptide comprises an Interpro domain having the Interpro accession number IPR010399 and/or an Interpro domain having the Interpro accession number IPR018467, preferably both domains.

2. Method according to Item 1, wherein said modulated expression is effected by introducing and expressing in a plant said nucleic acid encoding said TLP polypeptide.

3. Method according to Item 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 Items 1 to 3, wherein said enhanced yield-related traits are obtained under non-stress conditions.

5. Method according to any one of Items 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 Items 1 to 4, wherein said TLP polypeptide comprises one or more of the following motifs:
   (i) Motif 1-1: (SEQ ID NO: 35):
   QLTIFY[AG]G[SM]V[NC]V[YF][DE][DN][IV]S[PA]EKAQ[AE][IL]M,
   (ii) Motif 2-1: (SEQ ID NO: 37):
   PQARKASLARFLEKRKERV[MT][NST][TAL][AS]PY,
   (iii) Motif 3-1: (SEQ ID NO: 39):
   MERDF[LM]GL[NGSI][IS]K[DEN][PS][LP][LA][VT][VI]K[DE]Exxx[SD][SG]
   (iv) Motif 4-1 (SEQ ID NO: 40)
   Q[LM]TIFY[AG]G[SMATL]V[NCS][VI][YF][DEN][DN][IV][STP][PAV][ED][KQ]A[QK][AE][IL]MFLA[GS][HNR].
   (v) any of the Motifs 1-1a, 2-1a, 4-1a, 4-1b, 5-1, 6-1 or 7-1 as defiend herein above.

7. Method according to any one of Items 1 to 6, wherein said nucleic acid encoding a TLP is of plant origin, preferably from a dicotyledonous plant, further preferably from the family Solanceae, more preferably from the genus Solanum, most preferably from Solanum lycopersicum.

8. Method according to any one of Items 1 to 7, wherein said nucleic acid encoding a TLP encodes any one of the polypeptides listed in Table A1 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 Items 1 to 8, wherein said nucleic acid sequence encodes an orthologue or paralogue of any of the polypeptides given in Table A1.

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

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

12. Plant, plant part thereof, including seeds, or plant cell, obtainable by a method according to any one of Items 1 to 11, wherein said plant, plant part or plant cell comprises a recombinant nucleic acid encoding a TLP polypeptide as defined in any of Items 1 and 6 to 11.

13. Construct comprising:
   (i) nucleic acid encoding a TLP as defined in any of Items 1 and 6 to 11;

(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 Item 13, wherein one of said control sequences is a constitutive promoter, preferably a medium strength constitutive promoter, preferably to a plant promoter, more preferably a GOS2 promoter, most preferably a GOS2 promoter from rice.

15. Use of a construct according to Item 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 Item 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 a TLP polypeptide as defined in any of Items 1 and 6 to 11; 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 a TLP polypeptide as defined in any of Items 1 and 6 to 11 or a transgenic plant cell derived from said transgenic plant.

19. Transgenic plant according to Item 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 Item 19, wherein said harvestable parts are preferably shoot biomass and/or seeds.

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

22. Use of a nucleic acid encoding a TLP polypeptide as defined in any of Items 1 and 6 to 11 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 the production of a product comprising the steps of growing the plants according to Item 12, 16 or 18 and producing said product from or by
    a. said plants; or
    b. parts, including seeds, of said plants.

24. Construct according to Item 13 or 14 comprised in a plant cell

Other particularly preferred embodiments

Item D-1-A to D-1-W:

A. A method for enhancing yield in plants relative to control plants, comprising modulating expression, preferably increasing expression, in a plant of a nucleic acid molecule encoding a polypeptide, preferably a TLP polypeptide,
   i) wherein said polypeptide comprises at least one PF06200 Pfam domain and/or PF09425 Pfam domain, preferably both, and/or
   ii) wherein said polypeptide comprises an Interpro domain IPR010399 and/or an Interpro domain IPR018467, preferably both.

B. Method according to item A, wherein said polypeptide comprises one or more of the following motifs:
   Motif 1-1 (SEQ ID NO: 35): QLTI-FY[AG]G[SM]V[NC]V[YF][DE][DN][IV]S[PA]EKAQ[AE][IL]M;
   Motif 2-1 (SEQ ID NO: 37): PQARKASLARFLEKRKERV[MT][NST][TAL][AS]PY;
   Motif 3-1 (SEQ ID NO: 39): MERDF[LM]GL[NGSI][IS]K[DEN][PS][LP][LA][VT][VI]K[DE]Exxx[SD][SG]; wherein "X", preferably, represents any amino acid,
   Motif 4-1 (SEQ ID NO: 40): Q[LM]TIFY[AG]G[SMATL]V[NCS][VI][YF][DEN][DN][IV][STP][PAV][ED][KQ]A[QK][AE][IL]MFLA[GS][HNR];
   Motif 5-1 (SEQ ID NO: 43):RFLEKRKE;
   Motif 6-1 (SEQ ID NO: 44): QLTIFY[AG]G;
   Motif 7-1 (SEQ ID NO: 45):MERDF[LM]GL;
   or any of motifs 1-1a, 2-1a, 4-1a or 4-1b as defined herein above

C. Method according to item A or B, wherein said modulated expression is effected by introducing and expressing in a plant a nucleic acid molecule encoding a TLP polypeptide.

D. Method according to any one of items A to C, wherein said polypeptide is encoded by a nucleic acid molecule comprising a nucleic acid molecule selected from the group con-

sisting of:

(i) a nucleic acid represented by any one of SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, or 33 ;

(ii) the complement of a nucleic acid represented by any one of SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, or 33 ;

(iii) a nucleic acid encoding the polypeptide as represented by any one of SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, or 34, preferably as a result of the degeneracy of the genetic code, said isolated nucleic acid can be deduced from a polypeptide sequence as represented by any one of SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, or 34, and further preferably confers enhanced yield-related traits relative to control plants (as described herein elsewhere);

(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 any of the nucleic acid sequences of SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, or 33;

(v) a first nucleic acid molecule which hybridizes with a second 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 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 any one of SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, or 34, 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.

E. Method according to any item A to D, wherein said enhanced yield-related traits comprise increased yield, preferably seed yield and/or shoot biomass relative to control plants.

F. Method according to any one of items A to E, wherein said enhanced yield-related traits are obtained under non-stress conditions.

G. Method according to any one of items A to E, wherein said enhanced yield-related traits are obtained under conditions of drought stress, salt stress or nitrogen deficiency.

H. Method according to any one of items A to G, wherein said nucleic acid is operably linked to a constitutive promoter, preferably to a GOS2 promoter, most preferably to a GOS2 promoter from rice.

I. Method according to any one of items A to H, wherein said nucleic acid molecule or said polypeptide, respectively, is of plant origin, preferably from a dicotyledonous plant, further preferably from the family Solanaceae, more preferably from the genus Solanum, most preferably from Solanum lycopersicum.

J. Plant or part thereof, including seeds, obtainable by a method according to any one of items A to I, wherein said plant or part thereof comprises a recombinant nucleic acid encoding said polypeptide as defined in any one of items A to I.

K. Construct comprising:

(i) nucleic acid encoding said polypeptide as defined in any one of items A to H;

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

(iii) a transcription termination sequence.

L. Construct according to item K, wherein one of said control sequences is a constitutive promoter, preferably a GOS2 promoter, most preferably a GOS2 promoter from rice.

M. Use of a construct according to item K or L in a method for making plants having increased yield, particularly seed yield and/or shoot biomass and/or root biomass relative to control plants relative to control plants.

N. Plant, plant part or plant cell transformed with a construct according to item K or L or obtainable by a method according to any one of items A to 9, wherein said plant or part thereof comprises a recombinant nucleic acid encoding said polypeptide as defined in any one of items A to J.

O. Method for the production of a transgenic plant having increased yield, particularly increased biomass and/or increased seed yield relative to control plants, comprising:

(i) introducing and expressing in a plant a nucleic acid encoding said polypeptide as defined in any one of items A to H; and

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

P. Plant having increased yield, particularly increased biomass and/or increased seed yield, relative to control plants, resulting from modulated expression of a nucleic acid encoding said polypeptide, or a transgenic plant cell originating from or being part of said transgenic plant.

Q. A method for the production of a product comprising the steps of growing the plants of the invention and producing said product from or by
   a. the plants of the invention; or
   b. parts, including seeds, of these plants.

R. Plant according to item J, N, or P, or a transgenic plant cell originating thereof, or a method according to item Q, wherein said plant is a crop plant, preferably a dicot such as sugar beet, alfalfa, trefoil, chicory, carrot, cassava, cotton, soybean, canola or a monocot, such as sugarcane, or a cereal, such as rice, maize, wheat, barley, millet, rye, triticale, sorghum emmer, spelt, secale, einkorn, teff, milo and oats.

S. Harvestable parts of a plant according to item J, wherein said harvestable parts are preferably shoot and/or root biomass and/or seeds.

T. Products produced from a plant according to item J and/or from harvestable parts of a plant according to item R.

U. Use of a nucleic acid encoding a polypeptide as defined in any one of items A to H in increasing yield, particularly seed yield and/or shoot biomass relative to control plants.

V. Construct according to item K or L comprised in a plant cell.

W. Recombinant chromosomal DNA comprising the construct according to item K or L.

D-2. PMP22 polypeptide (22 kDa peroxisomal membrane like polypeptide)- ITEMS

The explanations and definitions given herein above in section C-2 apply mutatis mutandis to the following items (in D-2).

Items D-2-1 to D-2-27

1. A method for enhancing yield-related traits in plants relative to control plants, comprising modulating expression, preferably increasing expression, in a plant of a nucleic acid encoding PMP22 (22 kDa Peroxisomal Membrane protein) polypeptide, wherein said PMP22 polypeptide comprises a Pfam domain having the pfam accession number PF04117, and/or an Interpro domain having the Interpro Accession number IPR007248.

2. Method according to Item 1, wherein said modulated expression is effected by introducing and expressing in a plant said nucleic acid encoding said PMP22 polypeptide.

3. Method according to Item 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 Items 1 to 3, wherein said enhanced yield-related traits are obtained under non-stress conditions.

5. Method according to any one of Items 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 Items 1 to 5, wherein said PMP22 polypeptide comprises one or more of the following motifs:
   (i) Motif 1-2:
   GDWIAQC[YF]EGKPLFE[FI]DR[AT]RM[FL]RSGLVGFTLHGSLSHYYY[QH]FCE[AE]LFPF[QKE] (SEQ ID NO: 126),
   (ii) Motif 2-2:
   LTID[HQ]DYWHGWT[LI][FY]EILRY[AM]P[QE]HNW[VSI]AYE[EQ]ALK[RTA]NPVLAKM (SEQ ID NO: 127),
   (iii) Motif 3-2:
   [DE]WWVVP[AV]KVAFDQT[VA]W[SA]A[IV]WN (SEQ ID NO: 128);
   (iv) or any of the motifs 4-2 to 9-2 as defined herein above.

7. Method according to any one of Items 1 to 6, wherein said nucleic acid encoding a PMP22 is of plant origin, preferably from a dicotyledonous plant, further preferably from the family Solanaceae, more preferably from the genus Solanum, most preferably from Solanum lycopersicum.

8. Method according to any one of Items 1 to 7, wherein said nucleic acid encoding a PMP22 encodes any one of the polypeptides listed in Table A2 or is a portion of such a nucleic acid, or a nucleic acid capable of hybridising with the complementary sequence of such a nucleic acid.

9. Method according to any one of Items 1 to 8, wherein said nucleic acid sequence encodes an orthologue or paralogue of any of the polypeptides given in Table A2.

10. Method according to any one of Items 1 to 9, wherein said nucleic acid encodes the polypeptide represented by SEQ ID NO: 51.

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

12. Plant, plant part thereof, including seeds, or plant cell, obtainable by a method according to any one of Items 1 to 11, wherein said plant, plant part or plant cell comprises a recombinant nucleic acid encoding a PMP22 polypeptide as defined in any of Items 1 and 6 to 11.

13. Construct comprising:
    (i) nucleic acid encoding a PMP22 as defined in any of Items 1 and 6 to 11;
    (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 Item 13, wherein one of said control sequences is a constitutive promoter, preferably a medium strength constitutive promoter, preferably to a plant promoter, more preferably a GOS2 promoter, most preferably a GOS2 promoter from rice.

15. Use of a construct according to Item 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 Item 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 a PMP22 polypeptide as defined in any of Items 1 and 6 to 11; 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 a PMP22 polypeptide as defined in any of Items 1 and 6 to 11or a transgenic plant cell derived from said transgenic plant.

19. Transgenic plant according to Item 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 Item 19, wherein said harvestable parts are preferably shoot biomass and/or seeds.

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

22. Use of a nucleic acid encoding a PMP22 polypeptide as defined in any of Items 1 and 6 to 11 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 the production of a product comprising the steps of growing the plants according to Item 12, 16 or 18 and producing said product from or by
    (i) said plants; or
    (ii) parts, including seeds, of said plants.

24. Construct according to Item 13 or 14 comprised in a plant cell.

25. Any of the preceding Items, wherein the sequence of the nucleic acid encodes said PMP22 polypeptide or the sequence of the PMP22 polypeptide is not the sequence as shown in SEQ ID NO: 20 as disclosed in WO2004/035798, as shown in SEQ ID NO: 5180 as disclosed in EP 1 586 645 A2, as shown in SEQ ID NO: 277535 as disclosed in

US2004031072, as shown in SEQ ID NO: 42604 as disclosed in JP2005185101, as shown in SEQ ID NO: 302211 as disclosed in US2004214272, SEQ ID NO: 6940 as disclosed in US2009019601, or SEQ ID NO: 69977 or SEQ ID NO: 51830 as disclosed in US2007011783.

26. An isolated nucleic acid molecule selected from:
    (i) a nucleic acid represented by SEQ ID NO: 56, 90, or 104;
    (ii) the complement of a nucleic acid represented by SEQ ID NO: 56, 90, or 104;
    (iii) a nucleic acid encoding a PMP22 polypeptide having in increasing order of preference at least 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%, 99%, or 100% sequence identity to the amino acid sequence represented by SEQ ID NO: 57, 91 or 105, 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.

27. An isolated polypeptide selected from:
    (i) an amino acid sequence represented by SEQ ID NO: 57, 91 or 105;
    (ii) an amino acid sequence having, in increasing order of preference, at least 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: 57, 91 or 105, and preferably conferring enhanced yield-related traits relative to control plants; and
    (iii) derivatives of any of the amino acid sequences given in (i) or (ii) above.

Other embodiments

Item D-2-A to D-2-W:

A. A method for enhancing yield in plants relative to control plants, comprising modulating expression, preferably increasing expression, in a plant of a nucleic acid molecule encoding a polypeptide, preferalby, a PMP22 protein, wherein said polypeptide comprises comprising a Pfam domain having the Pfam accession number PF04117 and/or an Interpro domain having the accession number IPR007248.

B. Method according to item A, wherein said polypeptide comprises one or more of the following motifs:

Motif 1-2 (SEQ ID NO: 126):
GDWIAQC[YF]EGKPLFE[FI]DR[AT]RM[FL]RSGLVGFTLHGSLSHYYY[QH]FCE[AE]LFPF[QKE];
Motif 2-2 (SEQ ID NO: 127):
LTID[HQ]DYWHGWT[LI][FY]EILRY[AM]P[QE]HNW[VSI]AYE[EQ]ALK[RTA]NPVLAKM;
Motif 3-2 (SEQ ID NO: 128): [DE]WWVVP[AV]KVAFDQT[VA]W[SA]A[IV]WN;
Motif 4-2 (SEQ ID NO: 129):
LVGFT-LHGSLSHYYY[QH][FIL]CEALFPF[QKE][DE]WWVVP[AV]KVAFDQT[VI]WSAIWNSIYF;
Motif 5-2 (SEQ ID NO: 130):
RY[AM]P[EQ]HNW[ISV]AYE[EQ]ALK[AR]NPVLAKM[VAM]ISG[VI]VYS[LIV]GDWIAQCYEGKP[LI]F[ED][FI]D;
Motif 6-2 (SEQ ID NO: 131): AHL[IV]TYG[VL][IV]PVEQRLLWVDC;
Motif 7-2 (SEQ ID NO: 132):
RYAPQHNW[IV]AYEEALK[RQ]NPVLAKMVISGVVYS[VL]GDWIAQCYEGKPLF[ED][IF]D;
Motif 8-2 (SEQ ID NO: 133):
GFT-LHGSLSH[YF]YYQFCE[AE]LFPF[QE]DWWVVP[VA]KVAFDQTVWSAIWNSIY[FY]TV;
and
Motif 9-2 (SEQ ID NO: 134):
F[LW]PMLTAGWKLWPFAHLITYG[VL][VI]PVEQRLLWVDCVEL[IV]WVTILSTYSNEK.

C. Method according to item A or B, wherein said modulated expression is effected by introducing and expressing in a plant a nucleic acid molecule encoding a PMP22 protein.

D. Method according to any one of items A to C, 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 any one of SEQ ID NO: 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, or 124;

(ii) the complement of a nucleic acid represented by any one of SEQ ID NO: 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, or 124;

(iii) a nucleic acid encoding the polypeptide as represented by any one of SEQ ID NO: 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, or 125, preferably as a result of the degeneracy of the genetic code, said isolated nucleic acid can be deduced from a polypeptide sequence as represented by (any one of) SEQ ID NO: 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, or 125 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 any of the nucleic acid sequences of SEQ ID NO: 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, or 124, and further preferably conferring enhanced yield-related traits relative to control plants,

(v) a first nucleic acid molecule which hybridizes with a second 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 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 (any one of) SEQ ID NO: 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, or 125 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.

E. Method according to any item A to D, wherein said enhanced yield-related traits comprise increased yield, preferably seed yield and/or biomass relative to control plants.

F. Method according to any one of items A to E, wherein said enhanced yield-related traits are obtained under non-stress conditions.

G. Method according to any one of items A to E, wherein said enhanced yield-related traits are obtained under conditions of drought stress, salt stress or nitrogen deficiency.

H. Method according to any one of items A to G, wherein said nucleic acid is operably linked to a constitutive promoter, preferably to a GOS2 promoter, most preferably to a GOS2 promoter from rice.

I. Method according to any one of items A to H, wherein said nucleic acid molecule or said polypeptide, respectively, is of plant origin, preferably from a dicotyledonous plant, further preferably from the family Solanceae, more preferably from the genus Solanum,

most preferably from *Solanum lycopersicum.*

J. Plant or part thereof, including seeds, obtainable by a method according to any one of items A to I, wherein said plant or part thereof comprises a recombinant nucleic acid encoding said polypeptide as defined in any one of items A to I.

K. Construct comprising:

(i) nucleic acid encoding said polypeptide as defined in any one of items A to H;

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

(iii) a transcription termination sequence.

L. Construct according to item K, wherein one of said control sequences is a constitutive promoter, preferably a GOS2 promoter, most preferably a GOS2 promoter from rice.

M. Use of a construct according to item K or L in a method for making plants having increased yield, particularly seed yield and/or biomass relative to control plants relative to control plants.

N. Plant, plant part or plant cell transformed with a construct according to item K or L or obtainable by a method according to any one of items A to I, wherein said plant or part thereof comprises a recombinant nucleic acid encoding said polypeptide as defined in any one of items A to J.

O. Method for the production of a transgenic plant having increased yield, particularly increased biomass and/or increased seed yield relative to control plants, comprising:

(i) introducing and expressing in a plant a nucleic acid encoding said polypeptide as defined in any one of items A to H; and

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

P. Plant having increased yield, particularly increased biomass and/or increased seed yield, relative to control plants, resulting from modulated expression of a nucleic acid encoding said polypeptide, or a transgenic plant cell originating from or being part of said transgenic plant.

Q. A method for the production of a product comprising the steps of growing the plants of the invention and producing said product from or by

a. the plants of the invention; or

b. parts, including seeds, of these plants.

R. Plant according to item J, N, or P, or a transgenic plant cell originating thereof, or a

method according to item Q, wherein said plant is a crop plant, preferably a dicot such as sugar beet, alfalfa, trefoil, chicory, carrot, cassava, cotton, soybean, canola or a monocot, such as sugarcane, or a cereal, such as rice, maize, wheat, barley, millet, rye, triticale, sorghum emmer, spelt, secale, einkorn, teff, milo and oats.

S. Harvestable parts of a plant according to item J, wherein said harvestable parts are preferably shoot and/or root biomass and/or seeds.

T. Products produced from a plant according to item J and/or from harvestable parts of a plant according to item R.

U. Use of a nucleic acid encoding a polypeptide as defined in any one of items A to H in increasing yield, particularly seed yield and/or shoot biomass relative to control plants.

V. Construct according to item K or L comprised in a plant cell.

W. Recombinant chromosomal DNA comprising the construct according to item K or L.

D-3. RTF (REM-like transcription factor) polypeptide - ITEMS

The explanations and definitions given herein above in section C-3 apply mutatis mutandis to the following items (in D-3).

Items D-3-1 to D-3-22

1. A method for enhancing yield-related traits in plants relative to control plants, comprising modulating expression, preferably increasing expression, in a plant of a nucleic acid encoding a RTF (REM-like transcription factor) polypeptide, wherein said nucleic acid is selected from

(i) a nucleic acid represented by any one of SEQ ID NO: 139, 141, 143, 145, 147, 149, 151, 153, 155, 157, 159, 161, or 163;

(ii) the complement of a nucleic acid represented by any one of SEQ ID NO: 139, 141, 143, 145, 147, 149, 151, 153, 155, 157, 159, 161, or 163;

(iii) a nucleic acid encoding the polypeptide as represented by any one of SEQ ID NO: 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, or 164, preferably as a result of the degeneracy of the genetic code, said isolated nucleic acid can be deduced from a polypeptide sequence as represented by any one of SEQ ID NO: 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, or 164;

(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 any of the nucleic acid sequences of SEQ ID NO SEQ ID NO: 139, 141, 143, 145, 147, 149, 151, 153, 155, 157, 159, 161, or 163,

(v) a nucleic acid which hybridizes with the nucleic acid molecule of (i) to (iv) under stringent hybridization conditions, and

(vi) a nucleic acid encoding a polypeptide having, in increasing order of preference, at least 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 any one of SEQ ID NO: 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, or 164.

2. Method according to Item 1, wherein the RTF polypeptide comprises at least two B3 PFAM domains, in particular 4 B3 domains, having the PFAM accession number pfam02362.

3. Method according to Item 1, wherein said modulated expression is effected by introducing and expressing in a plant said nucleic acid encoding said RTF polypeptide.

4. Method according to any one of Items 1 to 3, wherein said enhanced yield-related traits comprise improved early vigour and increased yield, in particular increased biomass relative to control plants.

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

6. Method according to any of Items 1 to 5, wherein said RTF polypeptide comprises one or both of the following motifs:
(i) Motif 1-3: PVAFF (SEQ ID NO: 165),
(ii) Motif 2-3: HDLRVGDIVVF (SEQ ID NO: 166).

7. Method according to any one of Items 1 to 6, wherein said nucleic acid encoding a RTF polypeptide is of plant origin, preferably from a dicotyledonous plant, further preferably from the family Brassicaceae, more preferably from the genus Arabidopsis, most preferably from Arabidopsis thaliana.

8. Method according to any one of Items 1 to 7, wherein said nucleic acid encoding a RTF encodes any one of the polypeptides listed in Table A3 or is a portion of such a nucleic acid, or a nucleic acid capable of hybridising with a complementary sequence of such a nucleic acid.

9. Method according to any one of Items 1 to 8, wherein said nucleic acid sequence encodes an orthologue or paralogue of any of the polypeptides given in Table A3.

10. Method according to any one of Items 1 to 9, wherein said nucleic acid encodes the polypeptide represented by SEQ ID NO: 140.

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

12. Plant, plant part thereof, including seeds, or plant cell, obtainable by a method according to any one of Items 1 to 11, wherein said plant, plant part or plant cell comprises a recombinant nucleic acid encoding a RTF polypeptide as defined in any of Items 1, 2 and 6 to 10.

13. Construct comprising:
    (i) nucleic acid encoding a RTF as defined in any of Items1, 2 and 6 to 10;
    (ii) one or more control sequences capable of driving expression of the nucleic acid sequence of (i); and optionally
    (i) a transcription termination sequence.

14. Construct according to Item 12, wherein one of said control sequences is a constitutive promoter, preferably a medium strength constitutive promoter, preferably to a plant promoter, more preferably a GOS2 promoter, most preferably a GOS2 promoter from rice.

15. Use of a construct according to Item 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 Item 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 a RTF polypeptide as defined in any of Items 1, 2 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 a RTF polypeptide as defined in any of Items 1, 2 and 6 to 10or a transgenic plant cell derived from said transgenic plant.

19. Transgenic plant according to Item 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 Item 19, wherein said harvestable parts are preferably shoot biomass and/or seeds.

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

22. Use of a nucleic acid encoding a RTF polypeptide as defined in any of Items 1, 2 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.

Other embodiments

Item D-3-A to D-3- X:

A. A method for enhancing yield in plants relative to control plants, comprising modulating expression, preferably increasing expression, in a plant of a nucleic acid molecule encoding a RTF polypeptide, wherein said polypeptide comprises at least two, in particular three or four B3 domains, having the PFAM accession number pfam02362, and/or having the Interpro Accession number IPR003340, and/or wherein the RTF polypeptide comprises an IPR015300 domain (DNA-binding pseudobarrel domain).

B. Method according to item A, wherein said polypeptide comprises one or both of the following motifs:
(i) Motif 1-3: PVAFF (SEQ ID NO: 165),
(ii) Motif 2-3: HDLRVGDIVVF (SEQ ID NO: 166).

C. Method according to item A or B, wherein said modulated expression is effected by introducing and expressing in a plant a nucleic acid molecule encoding said RTF polypeptide..

D. Method according to any one of items A to C, 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 any one of SEQ ID NO: 139, 141, 143, 145, 147, 149, 151, 153, 155, 157, 159, 161, or 163;
(ii) the complement of a nucleic acid represented by any one of SEQ ID NO: 139, 141, 143, 145, 147, 149, 151, 153, 155, 157, 159, 161, or 163;;
(iii) a nucleic acid encoding the polypeptide as represented by any one of SEQ ID NO: 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, or 164, preferably as a result of the degeneracy of the genetic code, said isolated nucleic acid can be deduced from a polypeptide sequence as represented by any one of SEQ ID NO: 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, or 164 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 any of the nucleic acid sequences of SEQ ID NO SEQ ID NO: 139, 141, 143, 145, 147, 149, 151, 153, 155, 157, 159, 161, or 163, and further preferably conferring enhanced yield-related traits relative to control plants,

(v) a first nucleic acid molecule which hybridizes with a second 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 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 any one of SEQ ID NO: 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, or 164 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.

E. Method according to any item A to D, wherein said enhanced yield-related traits comprise increased yield, preferably aboveground biomass or improved early vigourr relative to control plants.

F. Method according to any one of items A to E, wherein said enhanced yield-related traits are obtained under non-stress conditions.

G. Method according to any one of items A to E, wherein said enhanced yield-related traits are obtained under conditions of drought stress, salt stress or nitrogen deficiency.

H. Method according to any one of items A to G, wherein said nucleic acid is operably linked to a constitutive promoter, preferably to a GOS2 promoter, most preferably to a GOS2 promoter from rice.

I. Method according to any one of items A to H, wherein said nucleic acid molecule or said polypeptide, respectively, is of plant origin, preferably from a dicotyledonous plant, further preferably from the family Brassicaceae, more preferably from the genus Arabidopsis, most preferably from Arabidopsis thaliana.

J. Plant or part thereof, including seeds, obtainable by a method according to any one of items A to I, wherein said plant or part thereof comprises a recombinant nucleic acid encoding said polypeptide as defined in any one of items A to I.

K. Construct comprising:

(i) nucleic acid encoding said polypeptide as defined in any one of items A to H;

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

(iii) a transcription termination sequence.

L. Construct according to item K, wherein one of said control sequences is a constitutive promoter, preferably a GOS2 promoter, most preferably a GOS2 promoter from rice.

M. Use of a construct according to item K or L in a method for making plants having increased yield, particularly seed yield and/or shoot biomass relative to control plants relative to control plants.

N. Plant, plant part or plant cell transformed with a construct according to item K or L or obtainable by a method according to any one of items A to I, wherein said plant or part thereof comprises a recombinant nucleic acid encoding said polypeptide as defined in any one of items A to J.

O. Method for the production of a transgenic plant having increased yield, particularly increased biomass and/or increased seed yield relative to control plants, comprising:

(i) introducing and expressing in a plant a nucleic acid encoding said polypeptide as defined in any one of items A to H; and

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

P. Plant having increased yield, particularly increased biomass and/or increased seed yield, relative to control plants, resulting from modulated expression of a nucleic acid encoding said polypeptide, or a transgenic plant cell originating from or being part of said transgenic plant.

Q. A method for the production of a product comprising the steps of growing the plants of the invention and producing said product from or by

a. the plants of the invention; or

b. parts, including seeds, of these plants.

R. Plant according to item J, N, or P, or a transgenic plant cell originating thereof, or a method according to item Q, wherein said plant is a crop plant, preferably a dicot such as sugar beet, alfalfa, trefoil, chicory, carrot, cassava, cotton, soybean, canola or a monocot, such as sugarcane, or a cereal, such as rice, maize, wheat, barley, millet, rye, triticale, sorghum emmer, spelt, secale, einkorn, teff, milo and oats.

S. Harvestable parts of a plant according to item J, wherein said harvestable parts are preferably shoot and/or root biomass and/or seeds.

T. Products produced from a plant according to item J and/or from harvestable parts of a plant according to item R.

U. Use of a nucleic acid encoding a polypeptide as defined in any one of items A to H in increasing yield, particularly seed yield and/or shoot biomass relative to control plants.

V. Construct according to item K or L comprised in a plant cell.

W. Recombinant chromosomal DNA comprising the construct according to item K or L.

X. Any of the preceding items A to U, wherein the nucleic acid encodes the RTF polypeptide or polynucleotide does not have a sequence as shown in
SEQ ID NOs: 43550, 43565, 43576, 43568, 43548, 43575, 193877, 93871, 43560, 93863, 43562, 93879, 43570, 43558, 43578, 93869, 43556, 43572, and 93875 as disclosed in EP2090662A2,
SEQ ID NOs: 312 and 2527 as disclosed in WO02/16655; and
SEQ ID NO: 72 as disclosed in EP 2154956A2.

D-4. BP1 (Bigger plant 1) polypeptide - ITEMS

The explanations and definitions given herein above in section C-4 apply mutatis mutandis to the following items (in D-4).

Items D-4-1 to D-4-24

1. A method for enhancing yield-related traits in plants relative to control plants, comprising modulating expression, preferably increasing expression, in a plant of a nucleic acid encoding a BP1 polypeptide, wherein said BP1 polypeptide comprises the following motifs:
(i) Motif 1-4:
LNQ[DG]SXXND[EV]X[NS]DX[QP]G[HQ]X[GN]H[LP]EXXKX[DE][QE][VA][GE]VXE[DE]X[MI][TA][AP]DV[KN]LS[VA]CRDTG[NE] (SEQ ID NO: 276),
(ii) Motif 2-4:
L[WR]RDYXD[LV][LV][QK][ED][TN]EXK[KR][KR]XLXSX[KN][RK][RT][KS]L[AV]LL[AS]EVKFL[RQ][RK]K[YL]XSF[AKLP]K[GN][GDN]SQ[QK] (SEQ ID NO: 277), and
(iii) Motif 3-4:
[DE][DG]KRX[VI][PS]WQD[RQ]XALK (SEQ ID NO: 278),

(iv) Or any of the motifs 4-4 to 9-4, preferably any one or more of motifs 4-4 to 6-4 as defined herein above.

2. A method for enhancing yield-related traits in plants relative to control plants,, comprising modulating expression, preferably increasing expression, in a plant of a nucleic acid encoding a BP1 polypeptide, wherein said BP1 polypeptide has, in an increasing order of preference, at least 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: 171.

3. The method according to Item 2, wherein said BP1 polypeptide comprises one, two or three motifs of motifs 4-4, 5-4 and 6-4 as defined in Item 1.

4. Method according to any one of Items 1 to 3, wherein said enhanced yield-related traits comprise increased yield relative to control plants, and preferably comprise increased biomass, increased shoot biomass, increased root biomass, increased NUE (nitrogen use efficiency) and/or increased seed yield relative to control plants.

5. Method according to any one of Items 1 to 4, wherein said enhanced yield-related traits are obtained under non-stress conditions.

6. Method according to any one of Items 1 to 4, wherein said enhanced yield-related traits are obtained under conditions of drought stress, salt stress or nitrogen deficiency.

7. Method according to any one of Items 1 to 5, wherein said nucleic acid encoding a BP1 is of plant origin, preferably from a monocotyledonous plant, further preferably from the family Poaceae, more preferably from the genus Oryza, most preferably from Oryza sativa.

8. Method according to any one of Items 1 to 7, wherein said nucleic acid encoding a BP1 polypeptide encodes any one of the polypeptides listed in Table A4, preferably, a polypeptide represented by SEQ ID NO: 171, 70, 74 or 98, 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 Items 1 to 8, wherein said nucleic acid sequence encodes an orthologue or paralogue of any of the polypeptides given in Table A4.

10. Method according to any one of Items 1 to 9, wherein said nucleic acid encodes the polypeptide represented by SEQ ID NO: 171.

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

12. Plant, plant part thereof, including seeds, or plant cell, obtainable by a method according to any one of Items 1 to 11, wherein said plant, plant part or plant cell comprises a recombinant nucleic acid encoding a BP1 polypeptide as defined in any of Items 1 to 3, and 7 to 10.

13. Construct comprising:
    (i) nucleic acid encoding a BP1 as defined in any of Items 1 to 3 and 7 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 Item 13, wherein one of said control sequences is a constitutive promoter, preferably a medium strength constitutive promoter, preferably to a plant promoter, more preferably a GOS2 promoter, most preferably a GOS2 promoter from rice.

15. Use of a construct according to Item 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 Item 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 a BP1 polypeptide as defined in any of Items 1 to 3 and 7 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 a BP1 polypeptide as defined in any of Items 1 to 3 and 7 to 10 or a transgenic plant cell originating from said transgenic plant and comprising a nucleic acid encoding a BP1 polypeptide as defined in any of Items 1 to 3 and 7 to 10.

19. Transgenic plant according to Item 12, 16 or 18, or a transgenic plant cell originating therefrom and comprising a nucleic acid encoding a BP1 polypeptide as defined in any of Items 1 to 3 and 7 to 10, 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 Item 19, wherein said harvestable parts are preferably shoot biomass and/or seeds.

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

22. Use of a nucleic acid encoding a polypeptide listed in Table A4 or a nucleic acid encoding a BP1 polypeptide as defined in any of Items 1 to 3 and 7 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 the production of a product comprising the steps of growing the plants according to Item 12, 16 or 18 and producing said product from or by
    (i) said plants; or
    (ii) parts, including seeds, of said plants.

24. Construct according to Item 13 or 14 comprised in a plant cell.

Other preferred items

Item D-4-A to D-4-X:

A. A method for enhancing yield in plants relative to control plants, comprising modulating expression, preferably increasing expression, in a plant of a nucleic acid molecule encoding a BP1 polypeptide, wherein said BP1 polypeptide comprises one or more of the following motifs:
    (i) Motif 1-4:
    LNQ[DG]SXXND[EV]X[NS]DX[QP]G[HQ]X[GN]H[LP]EXXKX[DE][QE][VA][GE]VXE[DE]X[MI][TA][AP]DV[KN]LS[VA]CRDTG[NE] (SEQ ID NO: 276 or SEQ ID NO: 289),
    (ii) Motif 2-4:
    L[WR]RDYXD[LV][LV][QK][ED][TN]EXK[KR][KR]XLXSX[KN][RK][RT][KS]L[AV]LL[AS]EVKFL[RQ][RK]K[YL]XSF[AKLP]K[GN][GDN]SQ[QK] (SEQ ID NO: 277 or SEQ ID NO: 290), and
    (iii) Motif 3-4:
    [DE][DG]KRX[VI][PS]WQD[RQ]XALK (SEQ ID NO: 278 or SEQ ID NO: 291)

(iv) Or any of the motifs 4-4 to 9-4, preferably any one or mmore of motifs 4-4 to 6-4 as defined herein above.

B. A method for enhancing yield in plants relative to control plants, comprising modulating expression, preferably increasing expression, in a plant of a nucleic acid molecule encoding a BP1 polypeptide, wherein said BP1 polypeptide 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 SEQ ID NO: 171, and wherein said BP1 polypeptide comprises one or more of the motifs as defined in item A.

C. Method according to item A or B, wherein said modulated expression is effected by introducing and expressing in a plant a nucleic acid molecule encoding a BP1 polypeptide.

D. Method according to any one of items A to C, 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 any one of SEQ ID NO: 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 208, 210, 212, 214, 216, 218, 220, 222, 224, 226, 228, 230, 232, 234, 236, 238, 240, 242, 244, 246, 248, 250, 252, 254, 256, 258, 260, 262, 264, 266, 268, 270, 272, or 274,

(ii) the complement of a nucleic acid represented by any one of SEQ ID NO: 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 208, 210, 212, 214, 216, 218, 220, 222, 224, 226, 228, 230, 232, 234, 236, 238, 240, 242, 244, 246, 248, 250, 252, 254, 256, 258, 260, 262, 264, 266, 268, 270, 272, or 274;

(iii) (iii) a nucleic acid encoding the polypeptide as represented by any one of SEQ ID NO: 171, 173, 175, 177, 179, 181, 183, 185, 187, 189, 191, 193, 195, 197, 199, 201, 203, 205, 207, 209, 211, 213, 215, 217, 219, 221, 223, 225, 227, 229, 231, 233, 235, 237, 239, 241, 243, 245, 247, 249, 251, 253, 255, 257, 259, 261, 263, 265, 267, 269, 271, 273, or 275,, preferably as a result of the degeneracy of the genetic code, said isolated nucleic acid can be deduced from a polypeptide sequence as represented by any one of SEQ ID NO: 171, 173, 175, 177, 179, 181, 183, 185, 187, 189, 191, 193, 195, 197, 199, 201, 203, 205, 207, 209, 211, 213, 215, 217, 219, 221, 223, 225, 227, 229, 231, 233, 235, 237, 239, 241, 243, 245, 247, 249, 251, 253, 255, 257,

259, 261, 263, 265, 267, 269, 271, 273, or 275 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 any of the nucleic acid sequences of SEQ ID NO: 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 208, 210, 212, 214, 216, 218, 220, 222, 224, 226, 228, 230, 232, 234, 236, 238, 240, 242, 244, 246, 248, 250, 252, 254, 256, 258, 260, 262, 264, 266, 268, 270, 272, or 274,

(v) a first nucleic acid molecule which hybridizes with a second 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 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 any one of SEQ ID NO: 171, 173, 175, 177, 179, 181, 183, 185, 187, 189, 191, 193, 195, 197, 199, 201, 203, 205, 207, 209, 211, 213, 215, 217, 219, 221, 223, 225, 227, 229, 231, 233, 235, 237, 239, 241, 243, 245, 247, 249, 251, 253, 255, 257, 259, 261, 263, 265, 267, 269, 271, 273, or 275, 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.

E. Method according to any item A to D, wherein said enhanced yield-related traits comprise increased yield, preferably seed yield, root biomass, aboveground biomass and/or shoot biomass relative to control plants.

F. Method according to any one of items A to E, wherein said enhanced yield-related traits are obtained under non-stress conditions.

G. Method according to any one of items A to E, wherein said enhanced yield-related traits are obtained under conditions of drought stress, salt stress or, more preferably, unter nitrogen deficiency.

H. Method according to any one of items A to G, wherein said nucleic acid is operably linked to a constitutive promoter, preferably to a GOS2 promoter, most preferably to a GOS2 promoter from rice.

I. Method according to any one of items A to H, wherein said nucleic acid molecule or said polypeptide, respectively, is of plant origin, preferably from a monocotyledonous plant, further preferably from the family Poaceae, more preferably from the genus Oryza, most preferably from Oryza sativa.

J. Plant or part thereof, including seeds, obtainable by a method according to any one of items A to I, wherein said plant or part thereof comprises a recombinant nucleic acid encoding said polypeptide as defined in any one of items A to I.

K. Construct comprising:
(i) nucleic acid encoding said polypeptide as defined in any one of items A to H;
(ii) one or more control sequences capable of driving expression of the nucleic acid sequence of (a); and optionally
(iii) a transcription termination sequence.

L. Construct according to item K, wherein one of said control sequences is a constitutive promoter, preferably a GOS2 promoter, most preferably a GOS2 promoter from rice.

M. Use of a construct according to item K or L in a method for making plants having increased yield, particularly seed yield and/or shoot biomass relative to control plants relative to control plants.

N. Plant, plant part or plant cell transformed with a construct according to item K or L or obtainable by a method according to any one of items A to H, wherein said plant or part thereof comprises a recombinant nucleic acid encoding said polypeptide as defined in any one of items A to J.

O. Method for the production of a transgenic plant having increased yield, particularly increased biomass and/or increased seed yield relative to control plants, comprising:
(i) introducing and expressing in a plant a nucleic acid encoding said polypeptide as defined in any one of items A to H; and
(ii) cultivating the plant cell under conditions promoting plant growth and development.

P. Plant having increased yield, particularly increased biomass and/or increased seed yield, relative to control plants, resulting from modulated expression of a nucleic acid encoding

a BP1 polypeptide, or a transgenic plant cell originating from or being part of said transgenic plant.

Q. A method for the production of a product comprising the steps of growing the plants of the invention and producing said product from or by
   a. the plants of the invention; or
   b. parts, including seeds, of these plants.

R. Plant according to item J, N, or P, or a transgenic plant cell originating thereof, or a method according to item Q, wherein said plant is a crop plant, preferably a dicot , preferably sugar beet, alfalfa, trefoil, chicory, carrot, cassava, cotton, soybean, canola or a monocot, preferably, sugarcane, or a cereal, such as rice, maize, wheat, barley, millet, rye, triticale, sorghum emmer, spelt, secale, einkorn, teff, milo and oats.

S. Harvestable parts of a plant according to item J, wherein said harvestable parts are preferably shoot and/or root biomass and/or seeds.

T. Products produced from a plant according to item J and/or from harvestable parts of a plant according to item R.

U. Use of a nucleic acid encoding a polypeptide as defined in any one of items A to H in increasing yield, preferably seed yield, root biomass, aboveground biomass and/or shoot biomass relative to control plants.

V. Construct according to item K or L comprised in a plant cell.

W. Recombinant chromosomal DNA comprising the construct according to item K or L.

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 domains and motifs. The conserved domains are indicated in bold. PF06200 is located in the central part, PF09425 is located in the C-terminal part of the protein. The motifs 1 to 6 are indicated with dashed lines (arabic numbers). Motifs 1 to 6 as indicated in Figure 1 correspond to motifs 1-1 to 6-1 as set forth in section C-1.

Fig. 2 represents a multiple alignment of various TLP polypeptides. The asterisks indicate identical amino acids among the various protein sequences, colons represent highly conserved amino acid substitutions, and the dots represent less conserved amino acid substitution; on

other positions there is no sequence conservation. These alignments can be used for defining further motifs or signature sequences, when using conserved amino acids. For the alignments, the sequences of the TLP polypeptides from Solanum lycopersicum UNK LLR (SEQ ID NO: 2), Arabidopsis thaliana_AT3G17860.1 (SEQ ID NO: 4), Brassica napus (SEQ ID NO: 6), Glycine max (SEQ ID NO: 8), Glycine max (SEQ ID NO: 10), Glycine max (SEQ ID NO: 12), Hordeum vulgare (SEQ ID NO: 14), Medicago truncatula (SEQ ID NO: 16), Medicago truncatula (SEQ ID NO: 18), Populus trichocarpa (SEQ ID NO: 20), Populus trichocarpa (SEQ ID NO: 22), Populus trichocarpa (SEQ ID NO: 24), Populus trichocarpa (SEQ ID NO: 26), Solanum lycopersicum (SEQ ID NO: 28), Triticum aestivum (SEQ ID NO: 30), Oryza sativa (SEQ ID NO: 32), Zea mays (SEQ ID NO: 34).

Fig. 3 shows phylogenetic tree of TLP polypeptides. The phylogenetic tree was constructed by aligning PMP22 sequences using MAFFT (Katoh and Toh (2008) - Briefings in Bioinformatics 9:286-298). A neighbour-joining tree was calculated using Quick-Tree (Howe et al. (2002), Bioinformatics 18(11): 1546-7), 100 bootstrap repetitions.

Fig. 4 shows the MATGAT table of Example III-1.

Fig. 5 represents the binary vector used for increased expression in Oryza sativa of a TLP-encoding nucleic acid under the control of a rice GOS2 promoter (pGOS2).

Fig. 6 represents the domain structure of SEQ ID NO: 51 with conserved domains and motifs. The conserved Pfam domain PF04117 is indicated in bold. The motifs 1 to 9 are indicated with dashed lines (arabic numbers). Motifs 1 to 9 as indicated in Figure 6 correspond to motifs 1-2 to 9-2 as set forth in section C-2.

Fig. 7 represents a multiple alignment of various PMP22 polypeptides. The asterisks indicate identical amino acids among the various protein sequences, colons represent highly conserved amino acid substitutions, and the dots represent less conserved amino acid substitution; on other positions there is no sequence conservation. These alignments can be used for defining further motifs or signature sequences, when using conserved amino acids. Fig. 7A shows an alignment of polypeptides comprised by cluster A. Fig 7B shows an alignment of polypeptides comprised by clusters A and B. Fig 7C shows an alignment of polypeptides comprised by clusters A, B and C (for the Clusters, see phylogenetic tree in Fig. 8). For the alignments, the sequences of the PMP polypeptides from Lycopersicon esculentum L450 PMP22 (SEQ ID NO: 51), Arabidopsis lyrata (SEQ ID NO: 53), Arabidopsis thaliana AT1G52870.2 (SEQ ID NO: 55), Brassica napus BN06MC04723 42271568 (SEQ ID NO: 57), Ipomoea nil (SEQ ID NO: 59), Nicotiana benthamiana (SEQ ID NO: 61), Nicotiana tabacum (SEQ ID NO: 63), Solanum tuberosum (SEQ ID NO:65), Arabidopsis lyrata (SEQ ID NO: 67), Arabidopsis thaliana (SEQ ID NO: 69), Glycine max (SEQ ID NO: 71), Glycine max (SEQ ID NO: 73), Glycine max (SEQ ID NO: 75), Helianthus annuus (SEQ ID NO: 77), Helianthus paradoxus (SEQ ID NO: 79), Malus domestica (SEQ ID NO: 81), Oryza sativa (SEQ ID NO: 83), Physcomitrella patens (SEQ ID NO: 85), Panicum virgatum (SEQ ID NO: 87), Sorghum bicolor (SEQ ID NO: 89), Zea mays ZM07MC32543 BFb0296A02@32446 (SEQ ID NO: 91), Zea mays (SEQ ID NO: 93), Triphysaria sp (SEQ ID NO: 95), Vitis vinifera (SEQ ID NO: 97), Aquilegia sp (SEQ ID NO: 99), Glycine

max (SEQ ID NO: 101), Glycine max (SEQ ID NO: 103), Glycine max GM06MC03382 49802960@3354 (SEQ ID NO: 105), Gossypium raimondii (SEQ ID NO: 107), Helianthus argophyllus (SEQ ID NO: 109), Lactuca sativa (SEQ ID NO: 111), Prunus persica (SEQ ID NO: 113), Poncirus trifoliata (SEQ ID NO: 115), Phaseolus vulgaris (SEQ ID NO: 117), Theobroma cacao (SEQ ID NO: 119), Vitis vinifera (SEQ ID NO: 121), Cichorium intybus (SEQ ID NO: 123), Gossypium hirsutum (SEQ ID NO: 125). For the full name of the proteins, see Table A2.

Fig. 8 shows phylogenetic tree of PMP22 polypeptides. The phylogenetic tree was constructed by aligning PMP22 sequences using MAFFT (Katoh and Toh (2008) - Briefings in Bioinformatics 9:286-298). A neighbour-joining tree was calculated using Quick-Tree (Howe et al. (2002), Bioinformatics 18(11): 1546-7), 100 bootstrap repetitions.

Fig. 9 shows the MATGAT table of Example III-2.

Fig. 10-2 represents the binary vector used for increased expression in Oryza sativa of a PMP22-encoding nucleic acid under the control of a rice GOS2 promoter (pGOS2).

Fig. 11 represents the domain structure of SEQ ID NO: 140 with conserved motifs 1 and 2. The four B3 domains are shown in bold. Motifs 1 and 2 as indicated in Figure 11 correspond to motifs 1-3 and 2-3 as set forth in section C-3.

Fig. 12 represents a multiple alignment of various RTF polypeptides. The asterisks indicate identical amino acids among the various protein sequences, colons represent highly conserved amino acid substitutions, and the dots represent less conserved amino acid substitution; on other positions there is no sequence conservation. These alignments can be used for defining further motifs or signature sequences, when using conserved amino acids. The aligned polypeptide sequences are as follows: Arabidopsis thaliana_AT2G24700.1#1, SEQ ID NO: 140; Arabidopsis thaliana_AT4G00260.1#1, SEQ ID NO: 142; Arabidopsis thaliana_AT2G24650.1#1, SEQ ID NO: 144; Arabidopsis thaliana_AT2G24650.2#1, SEQ ID NO: 146; Arabidopsis thaliana_AT1G26680.1#1, SEQ ID NO: 148; Brassica napus_CD826203#1, SEQ ID NO: 150; Brassica napus_TC73539#1, SEQ ID NO: 152; Solanum lycopersicum_TC201533#1, SEQ ID NO:154; Arabidopsis thaliana_ At4g31680, SEQ ID NO: 156; Arabidopsis thaliana_At4g31640, SEQ ID NO: 158; Arabidopsis thaliana_At4g31660, SEQ ID NO: 160; Arabidopsis thaliana_At4g31650, SEQ ID NO: 162; Arabidopsis thaliana_At4g31690, SEQ ID NO: 164.

Fig. 13 shows phylogenetic tree of RTF polypeptides (see Examples).

Fig. 14 shows the MATGAT table of Example III-3.

Fig. 15 represents the binary vector used for increased expression in Oryza sativa of a RTF-encoding nucleic acid under the control of a rice GOS2 promoter (pGOS2).

Fig. 16 represents the domain structure of SEQ ID NO: 171 with conserved motifs 1to 6. The location of the motifs is indicated with dashed lines , arabic number “1” represents motifs 1 and 4, “2” the location of motifs 2 and 5 and “3” the location of motifs 3 and 6. Motifs 1 to 6 as indicated in Figure 16 correspond to motifs 1-4 to 6-4 as set forth in section C-4.

Fig. 17 represents a multiple alignment of various BP1 polypeptides. The asterisks indicate identical amino acids among the various protein sequences, colons represent highly conserved amino acid substitutions, and the dots represent less conserved amino acid substitution; on other positions there is no sequence conservation. These alignments can be used for defining further motifs or signature sequences, when using conserved amino acids. The aligned sequences are the following (SEQ ID NOs are given in brackets): Oryza sativa LOC Os09g25410 (SEQ ID NO: 171); Arabidopis lyrata 944925 (SEQ ID NO: 173); Arabidpsis thaliana AT4G30630 (SEQ ID NO: 175); Brassica napus TC91202 (SEQ ID NO: 177); Capsicum annuum TC15926 (SEQ ID NO: 179); Cichorium intybus TA970 13427 (SEQ ID NO: 181); Centaurea maculosa TA1609 215693 (SEQ ID NO: 183); Centaurea maculosa TA3603 215693 (SEQ ID NO: 185); Carthamus tinctorius TA4044 4222 (SEQ ID NO: 187); Euphorbia esula TC2982 (SEQ ID NO: 189); Fragaria vesca TA11867 57918 (SEQ ID NO: 191); Gossypium hirsutum TC136942 (SEQ ID NO: 193); Glycine max Glyma02g36620 (SEQ ID NO: 195); Glycine max Glyma17g08070 (SEQ ID NO: 197); Helianthus annuus TC40508 (SEQ ID NO: 199); Helianthus argophyllus TA3915 73275 (SEQ ID NO: 201); Helianthus tuberosus EL464130 (SEQ ID NO: 203); Hordeum vulgare TC185682 (SEQ ID NO: 205); Lotus japonicus TC37963 (SEQ ID NO: 207); Lactuca saligna TA2168 75948 (SEQ ID NO: 4209); Lactuca sativa DW131501 (SEQ ID NO: 211); Lactuca sativa TC20185 (SEQ ID NO: 213); Lactuca serriola BU011148 (SEQ ID NO: 215); Lactuca virosa TA2198 75947 (SEQ ID NO: 217); Mesembryanthemum crystallinum TC9929 (SEQ ID NO: 219); Medicago truncatula AC150446 9.5 (SEQ ID NO: 221); Oryza sativa LOC Os08g16930 (SEQ ID NO: 223); Passiflora edulis FP092509 (SEQ ID NO: 225); Picea glauca BT104710 (SEQ ID NO: 227); Picea sitchensis TA13795 3332 (SEQ ID NO: 229); Pinus taeda TA10646 3352 (SEQ ID NO: 231); Populus trichocarpa 578729 (SEQ ID NO: 233); Populus trichocarpa scaff VI.1304 (SEQ ID NO: 235); Panicum virgatum TC29094 (SEQ ID NO: 237); Panicum virgatum TC30704 (SEQ ID NO: 239); Phaseolus vulgaris TC10046 (SEQ ID NO: 241); Sorghum bicolor Sb02g024920 (SEQ ID NO: 243); Sorghum bicolor Sb07g011060 (SEQ ID NO: 245); Solanum chacoense TA1669 4108 (SEQ ID NO: 247); Saruma henryi DT604565 (SEQ ID NO: 249); Solanum lycopersicum TC195266 (SEQ ID NO: 251); Solanum lycopersicum TC206342 (SEQ ID NO: 253); Solanum tuberosum AM908388 (SEQ ID NO: 255); Solanum tuberosum NP13064295 (SEQ ID NO: 257); Solanum lycopersicum 16878 (SEQ ID NO: 259); Theobroma cacao TC4923 (SEQ ID NO: 261); Tagetes erecta 417 (SEQ ID NO:263); Zea mays GRMZM2G075851 T01 (SEQ ID NO: 265); Zea mays GRMZM2G093731 T02 (SEQ ID NO: 267); Zea mays GRMZM2G371316 T01 (SEQ ID NO: 269); Zingiber officinale TA4076 94328 (SEQ ID NO: 271); Zingiber officinale TA6335 94328 (SEQ ID NO: 273); Zingiber officinale TA6947 94328 (SEQ ID NO: 275).

Fig. 18 shows phylogenetic tree of BP1 polypeptides. A phylogenetic tree of BP1 polypeptides was constructed by aligning BP1 sequences using MAFFT (Katoh and Toh (2008) - Briefings in Bioinformatics 9:286-298). A neighbour-joining tree was calculated using Quick-Tree (Howe et al. (2002), Bioinformatics 18(11): 1546-7), 100 bootstrap repetitions. The circular dendrogram (Fig. 18) was drawn using Dendroscope (Huson et al. (2007), BMC Bioinformatics 8(1):460). Confidence levels for 100 bootstrap repetitions are indicated for major branchings.

Fig. 19 represents the binary vector used for increased expression in Oryza sativa of a BP1-encoding nucleic acid under the control of a rice GOS2 promoter (pGOS2).
Fig. 20 shows the MATGAT table of Example III-4.

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.

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 I

I-1. TLP (Tify like protein) polypeptide

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 A1 provides a list of nucleic acid sequences related to SEQ ID NO: 1 and SEQ ID NO: 2.

Table A1: Examples of TLP nucleic acids and polypeptides:

Plant Source	Nucleic acid SEQ ID NO:	Protein SEQ ID NO:
Solanum lycopersicum_UNK LLR	1	2
Arabidopsis thaliana_AT3G17860.1	3	4
Brassica napus_TC100484	5	6
Glycine max_Glyma05g34960.1	7	8
Glycine max_Glyma08g04770.1	9	10
Glycine max_Glyma09g30460.1	11	12
Hordeum_vulgare_subsp_vulgare_AK250045	13	14
Medicago truncatula_TC118189	15	16
Medicago truncatula_TC119816	17	18
Populus trichocarpa_scaff_VIII.1248	19	20
Populus trichocarpa_scaff_X.1014	21	22
Populus trichocarpa_scaff_XV.429	23	24
Populus trichocarpa_TC117359	25	26
Solanum lycopersicum_TC207683	27	28
Triticum aestivum_TC290287	29	30
Oryza sativa_LOC_Os08g33160.1	31	32
Zea_mays_GRMZM2G126507_T01	33	34

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.

I-2. PMP22 polypeptide (22 kDa peroxisomal membrane like polypeptide)

Sequences (full length cDNA, ESTs or genomic) related to SEQ ID NO: 50 and SEQ ID NO: 51 were identified as described herein above under I-1.

Table A2 provides a list of nucleic acid sequences related to SEQ ID NO: 50 and SEQ ID NO: 51.

Table A2: Examples of PMP22 nucleic acids and polypeptides:

Plant Source	Nucleic acid SEQ ID NO:	Protein SEQ ID NO:

Lycopersicon esculentum L450 PMP22	50	51
Arabidopsis lyrata 474411	52	53
Arabidopsis thaliana AT1G52870.2	54	55
Brassica napus BN06MC04723 42271568	56	57
Ipomoea nil TC25	58	59
Nicotiana benthamiana TC11792	60	61
Nicotiana tabacum TC57360	62	63
Solanum tuberosum TC167881	64	65
Arabidopsis lyrata 490222	66	67
Arabidopsis thaliana AT4G03410.1	68	69
Glycine max Glyma12g32920.1	70	71
Glycine max Glyma13g37540.1	72	73
Glycine max TC302765	74	75
Helianthus annuus TC45682	76	77
Helianthus paradoxus TA2132 73304	78	79
Malus domestica TC51384	80	81
Oryza sativa LOC Os03g38730.1	82	83
Physcomitrella patens TC32628	84	85
Panicum virgatum TC39638	86	87
Sorghum bicolor Sb01g015680.1	88	89
Zea mays ZM07MC32543 BFb0296A02@32446	90	91
Zea mays GRMZM2G011269 T01	92	93
Triphysaria sp TC3527	94	95
Vitis vinifera GSVIVT00025998001	96	97
Aquilegia sp TC27124	98	99
Glycine max Glyma12g34950.1	100	101
Glycine max Glyma13g35620.1	102	103
Glycine max GM06MC03382 49802960@3354	104	105
Gossypium raimondii TC2516	106	107
Helianthus argophyllus TA2174 73275	108	109
Lactuca sativa TC17162	110	111
Prunus persica TC12602	112	113
Poncirus trifoliata TA5129 37690	114	115
Phaseolus vulgaris TC9799	116	117
Theobroma cacao TC9688	118	119
Vitis vinifera GSVIVT00027620001	120	121
Cichorium intybus TA545 13427	122	123
Gossypium hirsutum TC165523	124	125

I-3. RTF (REM-like transcription factor) polypeptide

Sequences (full length cDNA, ESTs or genomic) related to SEQ ID NO: 139 and SEQ ID NO: 140 were identified as described herein above under I-1.

Table A3 provides a list of nucleic acid sequences related to SEQ ID NO: 139 and SEQ ID NO: 140.

Table A3: Examples of RTF nucleic acids and polypeptides:

Plant Source	Nucleic acid SEQ ID NO:	Protein SEQ ID NO:
Arabidopsis thaliana AT2G24700.1#1	139	140
Arabidopsis thaliana AT4G00260.1#1	141	142
Arabidopsis thaliana AT2G24650.1#1	143	144
Arabidopsis thaliana AT2G24650.2#1	145	146
Arabidopsis thaliana_AT1G26680.1#1	147	148
Brassica napus_CD826203#1	149	150
Brassica napus_TC73539#1	151	152
Solanum lycopersicum_TC201533#1	153	154
Arabidopsis thaliana_ At4g31680	155	156
Arabidopsis thaliana_At4g31640	157	158
Arabidopsis thaliana_At4g31660	159	160
Arabidopsis thaliana_At4g31650	161	162
Arabidopsis thaliana_At4g31690	163	164

I-4. BP1 (Bigger plant 1) polypeptide

Sequences (full length cDNA, ESTs or genomic) related to SEQ ID NO: 170 and SEQ ID NO: 171 were identified as described herein above under I-1.

Table A4 provides a list of nucleic acid sequences related to SEQ ID NO: 170 and SEQ ID NO: 171.

Table A4: Examples of BP1 nucleic acids and polypeptides and other related sequences

Plant Source	**Nucleic acid SEQ ID NO:**	**Protein SEQ ID NO:**
Oryza sativa Os09g25410	170	171
Arabidopis lyrata 944925	172	173
Arabidpsis thaliana AT4G30630	174	175

Brassica napus TC91202	176	177
Capsicum annuum TC15926	178	179
Cichorium intybus TA970 13427	180	181
Centaurea maculosa TA1609 215693	182	183
Centaurea maculosa TA3603 215693	184	185
Carthamus tinctorius TA4044 4222	186	187
Euphorbia esula TC2982	188	189
Fragaria vesca TA11867 57918	190	191
Gossypium hirsutum TC136942	192	193
Glycine max Glyma02g36620	194	195
Glycine max Glyma17g08070	196	197
Helianthus annuus TC40508	198	199
Helianthus argophyllus TA3915 73275	200	201
Helianthus tuberosus EL464130	202	203
Hordeum vulgare TC185682	204	205
Lotus japonicus TC37963	206	207
Lactuca saligna TA2168 75948	208	209
Lactuca sativa DW131501	210	211
Lactuca sativa TC20185	212	213
Lactuca serriola BU011148	214	215
Lactuca virosa TA2198 75947	216	217
Mesembryanthemum crystallinum TC9929	218	219
Medicago truncatula AC150446 9.5	220	221
Oryza sativa LOC Os08g16930	222	223
Passiflora edulis FP092509	224	225
Picea glauca BT104710	226	227
Picea sitchensis TA13795 3332	228	229
Pinus taeda TA10646 3352	230	231
Populus trichocarpa 578729	232	233
Populus trichocarpa scaff VI.1304	234	235
Panicum virgatum TC29094	236	237
Panicum virgatum TC30704	238	239
Phaseolus vulgaris TC10046	240	241
Sorghum bicolor Sb02g024920	242	243
Sorghum bicolor Sb07g011060	244	245

Solanum chacoense TA1669 4108	246	247
Saruma henryi DT604565	248	249
Solanum lycopersicum TC195266	250	251
Solanum lycopersicum TC206342	252	253
Solanum tuberosum AM908388	254	255
Solanum tuberosum NP13064295	256	257
Solanum lycopersicum 16878	258	259
Theobroma cacao TC4923	260	261
Tagetes erecta 417	262	263
Zea mays GRMZM2G075851 T01	264	265
Zea mays GRMZM2G093731 T02	266	267
Zea mays GRMZM2G371316 T01	268	269
Zingiber officinale TA4076 94328	270	271
Zingiber officinale TA6335 94328	272	273
Zingiber officinale TA6947 94328	274	275

Example II: Alignment of TLP polypeptide sequences

Alignment of polypeptide sequences was performed using the ClustalW 2.0 algorithm of progressive alignment (Thompson et al. (1997) Nucleic Acids Res 25:4876-4882; Chenna et al. (2003); CLUSTAL 2.0.11). 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.

II-1. TLP (Tify like protein) polypeptide

The TLP polypeptides are aligned in Figure 2.

A phylogenetic tree of TLP polypeptides (Figure 3) was constructed by aligning TLP sequences using MAFFT (Katoh and Toh (2008) - Briefings in Bioinformatics 9:286-298). A neighbour-joining tree was calculated using Quick-Tree (Howe et al. (2002), Bioinformatics 18(11): 1546-7), 100 bootstrap repetitions.

II-2. PMP22 polypeptide (22 kDa peroxisomal membrane like polypeptide)

The PMP22 polypeptides are aligned in Figure 7.

A phylogenetic tree of PMP22 polypeptides (Figure 8) was constructed by aligning PMP22 sequences using MAFFT (Katoh and Toh (2008) - Briefings in Bioinformatics 9:286-298). A neigh-

bour-joining tree was calculated using Quick-Tree (Howe et al. (2002), Bioinformatics 18(11): 1546-7), 100 bootstrap repetitions.

II-3. RTF (REM-like transcription factor) polypeptide

The RTF polypeptides are aligned in Figure 12.

A phylogenetic tree of RTF polypeptides (Figure 13) was constructed by aligning RTF sequences using MAFFT (Katoh and Toh (2008) - Briefings in Bioinformatics 9:286-298). A neighbour-joining tree was calculated using Quick-Tree (Howe et al. (2002), Bioinformatics 18(11): 1546-7), 100 bootstrap repetitions. The dendrogram was drawn using Dendroscope (Huson et al. (2007), BMC Bioinformatics 8(1):460). Confidence levels for 100 bootstrap repetitions are indicated for major branchings.

II-4. BP1 (Bigger plant 1) polypeptide

The BP1 polypeptides are aligned in Figure 17.

A phylogenetic tree of BP1 polypeptides was constructed by aligning BP1 sequences using MAFFT (Katoh and Toh (2008) - Briefings in Bioinformatics 9:286-298). A neighbour-joining tree was calculated using Quick-Tree (Howe et al. (2002), Bioinformatics 18(11): 1546-7), 100 bootstrap repetitions. The dendrogram (Fig. 18) was drawn using Dendroscope (Huson et al. (2007), BMC Bioinformatics 8(1):460). Confidence levels for 100 bootstrap repetitions are indicated for major branchings.

Example III: 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 one of the methods available in the art, the 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 software 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 (with a gap opening penalty of 12, and a gap extension penalty of 2), calculates similarity and identity using for example Blosum 62 (for polypeptides), and then places the results in a distance matrix.

III-1. TLP (Tify like protein) polypeptide

Results of the analysis are shown in Figure 4 for the global similarity and identity 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 comparison were: Scoring matrix: Blosum62, First Gap: 12, Extending Gap: 2. The sequence identity (in %) between the TLP polypeptide sequences useful in performing the methods of the invention can be as is generally higher than 54.3 % compared to SEQ ID NO: 2.

III-2. PMP22 polypeptide (22 kDa peroxisomal membrane like polypeptide)

Results of the analysis are shown in Figure 9 for the global similarity and identity 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 comparison were: Scoring matrix: Blosum62, First Gap: 12, Extending Gap: 2. The sequence identity (in %) between the PMP22 polypeptide sequences useful in performing the methods of the invention can be as low as 35 %, and, thus, is generally higher than 35% compared to SEQ ID NO: 51.

III-3. RTF (REM-like transcription factor) polypeptide

Results of the analysis are shown in Figure 14 for the global similarity and identity 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 comparison were: Scoring matrix: Blosum62, First Gap: 12, Extending Gap: 2.

III-4. BP1 (Bigger plant 1) polypeptide

Results of the analysis are shown in Figure 20 for the global similarity and identity 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 comparison were: Scoring matrix: Blosum62, First Gap: 12, Extending Gap: 2.

Example IV: 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.

IV-1. TLP (Tify like protein) polypeptide

The results of the InterPro scan (see Zdobnov E.M. and Apweiler R.; "InterProScan - an integration platform for the signature-recognition methods in InterPro."; Bioinformatics, 2001, 17(9): 847-8; InterPro database, Release 31.0, 9th February 2011) of the polypeptide sequence as represented by SEQ ID NO: 2 are presented in Table B1.

In an embodiment a TLP polypeptide comprises a conserved domain 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 144 to 178 in SEQ ID NO:2 and/or 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 282 to 306 in SEQ ID NO:2.

IV-2. PMP22 polypeptide (22 kDa peroxisomal membrane like polypeptide)

The results of the InterPro scan (see Zdobnov E.M. and Apweiler R.; "InterProScan - an integration platform for the signature-recognition methods in InterPro."; Bioinformatics, 2001, 17(9): 847-8, InterPro database, Release 31.0, 9th February 2011) of the polypeptide sequence as represented by SEQ ID NO: 51 are presented in Table B2.

In an embodiment a PMP22 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 123 to 367 in SEQ ID NO:51, or from amino acid from 283 to 348 in SEQ ID NO: 51.

IV-3. RTF (REM-like transcription factor) polypeptide

The results of the InterPro scan (see Zdobnov E.M. and Apweiler R.; "InterProScan - an integration platform for the signature-recognition methods in InterPro."; Bioinformatics, 2001, 17(9): 847-8; InterPro database, Release 31.0, 9th February 2011) of the polypeptide sequence as represented by SEQ ID NO: 140 are presented in Table B3.

In a preferred embodiment of the present invention, the RTF polypeptide comprises a first, a second, a third and fourth B3 domain. Preferably, the first B3 domain comprises a sequence having, 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%, 99%, or 100% sequence identity to a conserved domain from amino acid 13 to 105 in SEQ ID NO: 140). Preferably, the second B3 domain comprises a se-

quence having, 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%, 99%, or 100% sequence identity to a conserved domain from amino acid 150 to 247 in SEQ ID NO: 140). Preferably, the third B3 domain comprises a sequence having, 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%, 99%, or 100% sequence identity to a conserved domain from amino acid 276 to 372 in SEQ ID NO: 140). Preferably, the fourth B3 domain comprises a sequence having, 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%, 99%, or 100% sequence identity to a conserved domain from amino acid 464 to 555 in SEQ ID NO: 140). Preferably, the order within the RTF polypeptide is as follows (from the N- to the C-terminus): first B3 domain, second B3 domain, third B3 domain and fourth B3 domain. Preferably the B3 domains are separated by 10 to 150 amino acids, and, more preferably, by 25 to 95 amino acids.

Table B1: 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**
Interpro	IPR010399	Tify	
Interpro	IPR018467	CO/COL/TOC1	
Pfam	PF06200	tify	144 to 178
Pfam	PF09425	CCT_2	282 to 306

Table B2: InterPro scan results (major accession numbers) of the polypeptide sequence as represented by SEQ ID NO: 51.

Database	**Accession number**	**Accession name**	**Amino acid coordinates of SEQ ID NO 51**
Interpro	IPR007248	Mpv17/PMP22	
PANTHER	PTHR11266	PEROXISOMAL MEMBRANE PROTEIN 2, PXMP2 (MPV17)	123-367
PFAM	PF04117	Mpv17_PMP22	283-348

Table B3: InterPro scan results (major accession numbers) of the polypeptide sequence as represented by SEQ ID NO: 140.

Database	**Accession number**	**Accession name**	**Amino acid coordinates on SEQ ID NO 140**
Interpro	IPR003340	Transcriptional factor B3	
	IPR015300	DNA-binding pseudobarrel domain	
Pfam	pfam02362	B3 DNA binding domain	13 to 105 150 to 247 276 to 372 464 to 555

Example V: Topology prediction of the TLP 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. TargetP is maintained at the server of the Technical University of Denmark.

For the sequences predicted to contain an N-terminal presequence a potential cleavage site can also be predicted.

A number of parameters were selected, 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).

V-1. TLP (Tify like protein) polypeptide

The results of TargetP 1.1 analysis of the polypeptide sequence as represented by SEQ ID NO: 2 are presented in Table C1. 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 cytoplasmic and/or nuclear.

Table C1: TargetP 1.1 analysis of the polypeptide sequence as represented by SEQ ID NO: 2

Length (AA)	338
Chloroplastic transit peptide	0.127
Mitochondrial transit peptide	0.113
Secretory pathway signal peptide	0.035
Other subcellular targeting	0.894
Predicted Location	/
Reliability class	2

V-2. PMP22 polypeptide (22 kDa peroxisomal membrane like polypeptide)

The results of TargetP 1.1 analysis of the polypeptide sequence as represented by SEQ ID NO: 51 are presented Table C2. 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: 51 may be Chloroplast. Thus the PMP22

polypeptide as set forth herein in preferably located in the Chloroplast. More preferably, it is located in the peroxisomal membrane.

Table C2: TargetP 1.1 analysis of the polypeptide sequence as represented by SEQ ID NO: 51

Length (AA)	376
Chloroplastic transit peptide	0.915
Mitochondrial transit peptide	0.098
Secretory pathway signal peptide	0.013
Other subcellular targeting	0.142
Predicted Location	Chloroplast
Reliability class	2
Predicted transit peptide length	/

V-3. RTF (REM-like transcription factor) polypeptide

The results of TargetP 1.1 analysis of the polypeptide sequence as represented by SEQ ID NO: 140 are presented Table C3. 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: 140 may be the nucleus. Accordingly, the RTF polypeptide is, preferably, located in the nucleus.

Table C3: TargetP 1.1 analysis of the polypeptide sequence as represented by SEQ ID NO: 140

Length (AA)	555
Chloroplastic transit peptide	
Mitochondrial transit peptide	0.079
Secretory pathway signal peptide	0.252
Other subcellular targeting	0.737
Predicted Location	/
Reliability class	3
Predicted transit peptide length	/

V-4. BP1 (Bigger plant 1) polypeptide

The analysis shows that subcellular localization of the polypeptide sequence as represented by SEQ ID NO: 171, is most likely the nucleus. Accordingly, the BP1 polypeptide as set forth herein in the context of the method of the present invention is, preferably, localized in the nucleus.

An analysis using PSORT (URL: psort.org) also indicated that the polypeptide having a sequence as shown in SEQ ID NO: 171 is located in the nucleus (0.91). A further analysis using ChloroP 1.1 hosted on the server of the Technical University of Denmark indicated that said polypeptide is non-chloroplastic.

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 VII:

VII-1. TLP (Tify like protein) polypeptide

Cloning of the TLP encoding nucleic acid sequence

The nucleic acid sequence was amplified by PCR using as template a custom-made *Solanum lycopersicum* 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 prm16282 (SEQ ID NO: 48; sense, start codon in bold):

5' ggggacaagtttgtacaaaaaagcaggcttaaaca**atg**gagagggactttatggga 3'

and prm16282 (SEQ ID NO: 49; reverse, complementary):

5' ggggaccactttgtacaagaaagctgggtgtggaagttgcagagaaacca3',

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", pTLP. Plasmid pDONR201 was purchased from Invitrogen, as part of the Gateway® technology.

The cDNA library used for cloning was custom made from different tissues (eg leaves, roots) of Solanum lycopersicum seedlings grown from seeds obtained in Belgium.

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::TLP (Figure 5) was transformed into Agrobacterium strain LBA4044 according to methods well known in the art.

VII-2. PMP22 polypeptide (22 kDa peroxisomal membrane like polypeptide)

Cloning of the PMP22 encoding nucleic acid sequence

The nucleic acid sequence was amplified by PCR using as template a custom-made *Solanum lycopersicum* 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 primer16396 (SEQ ID NO: 137; sense, start codon in bold):

5' ggggacaagtttgtacaaaaaagcaggcttaaacaa**atg**gcgaccatcaatgg 3'

and primer16397 (SEQ ID NO: 138; reverse complementary):

5' ggggaccactttgtacaagaaagctgggtaattttggttgtgtcattgct 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", pPMP22. Plasmid pDONR201 was purchased from Invitrogen, as part of the Gateway® technology.

The cDNA library used for cloning was custom made from different tissues (eg leaves, roots) of Solanum lycopersicum seedlings grown from seeds obtained in Belgium.

The entry clone comprising SEQ ID NO: 50 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: 135) for constitutive expression was located upstream of this Gateway cassette.

After the LR recombination step, the resulting expression vector pGOS2::PMP22 (Figure 10) was transformed into *Agrobacterium* strain LBA4044 according to methods well known in the art.

VII-3. RTF (REM-like transcription factor) polypeptide

Cloning of the RTF encoding nucleic acid sequence

The nucleic acid sequence was amplified by PCR using as template a custom-made Arabidopsis thaliana seedlings cDNA library. PCR was performed using a commercially available proof-

reading Taq DNA polymerase in standard conditions, using 200 ng of template in a 50 µl PCR mix. The primers used were prm15379 (SEQ ID NO: 168; sense, start codon in bold):
5' ggggacaagtttgtacaaaaaagcaggcttaaacaa**atg**gctaatccacttctctat 3'
and prm15380 (SEQ ID NO: 169; reverse, complementary):
5' ggggaccactttgtacaagaaagctgggtcgatgatcctagacattctta 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", pRTF. Plasmid pDONR201 was purchased from Invitrogen, as part of the Gateway® technology.

The cDNA library used for cloning was custom made from different tissues (eg leaves, roots) of Arabidopsis thaliana Col-0 seedlings grown from seeds obtained in Belgium.

The entry clone comprising SEQ ID NO: 139 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: 167) for constitutive expression was located upstream of this Gateway cassette.

After the LR recombination step, the resulting expression vector pGOS2::RTF (Figure 15) was transformed into Agrobacterium strain LBA4044 according to methods well known in the art.

VII-4. BP1 (Bigger plant 1) polypeptide

Cloning of the BP1 encoding nucleic acid sequence

The nucleic acid sequence was amplified by PCR using as template a custom-made *Oryza sativa* 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 prm16202 (SEQ ID NO: 284; sense, start codon in bold):
5' ggggacaagtttgtacaaaaaagcaggcttaaacaa**atg**gactacggcgacg 3'
and prm16203 (SEQ ID NO: 285; reverse, complementary):
5' ggggaccactttgtacaagaaagctgggtaaggatgttatttcatagcca 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", pBP1. Plasmid pDONR201 was purchased from Invitrogen, as part of the Gateway® technology.

The cDNA library used for cloning was custom made from different tissues (eg leaves, roots) of Oryza sativa seedlings.The entry clone comprising SEQ ID NO: 170 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: 282) for constitutive expression was located upstream of this Gateway cassette.

After the LR recombination step, the resulting expression vector pGOS2::BP1 (Figure 19) was transformed into *Agrobacterium* strain LBA4044 according to methods well known in the art.

Example VIII: 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 minutes in 0.2% $HgCl_2$, followed by a 6 times 15 minutes wash with sterile distilled water. The sterile seeds were then germinated on a medium containing 2,4-D (callus induction medium). After incubation in the dark for four weeks, embryogenic, scutellum-derived calli were excised and propagated on the same medium. After two weeks, the calli were multiplied or propagated by subculture on the same medium for another 2 weeks. Embryogenic callus pieces were sub-cultured on fresh medium 3 days before co-cultivation (to boost cell division activity).

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 suspension was then transferred to a Petri dish and the calli immersed in the suspension for 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. Co-cultivated calli were grown on 2,4-D-containing medium for 4 weeks in the dark at 28°C in the presence of a selection agent. During this period, rapidly growing resistant callus islands developed. After transfer of this material to a regeneration medium and incubation in the light, the embryogenic potential was released and shoots developed in the next four to five 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.

Approximately 35 independent T0 rice transformants were generated for one the TLP and RTF constructs. Approximately 35 to 90 independent T0 rice transformants were generated for one the PMP22 and BP1 constructs. 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).

Alternatively the following method may be used:
The *Agrobacterium* containing the expression vector is used to transform *Oryza sativa* plants. Mature dry seeds of the rice japonica cultivar Nipponbare are dehusked. Sterilization is 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 ish with sterile distilled water. The sterile seeds are 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 is used for co-cultivation. *Agrobacterium* is inoculated on AB medium with the appropriate antibiotics and cultured for 3 days at 28°C. The bacteria are then collected and suspended in liquid co-cultivation medium to a density (OD_{600}) of about 1. The calli are immersed in the suspension for 1 to 15 minutes. The callus tissues are 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 ishing away the *Agrobacterium,* the calli are 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 is released and shoots developed in the next four to six weeks. Shoots are excised from the calli and incubated for 2 to 3 weeks on an auxin-containing medium from which they are transferred to soil. Hardened shoots are 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 are generated for one construct. The primary transformants are 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 are kept for harvest of T1 seed. Seeds are then harvest-

ed 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 IX: 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 (see Murashige, T., and Skoog, ., 1962. A revised medium for rapid growth and bioassays with tobacco tissue cultures. Physiol. Plant, vol. 15, 473-497) including B5 vitamins (Gamborg et al.; Nutrient requirements of suspension cultures of soybean root cells. 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. Clonal propagation of sugarbeet plants and the formation of polylTLPds by tissue culture. 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. Transformation of sugarbeet (Beta vulgaris) by Agrobacterium tumefaciens. 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 (see Arencibia A., at al., 1998. An efficient protocol for sugarcane (Saccharum spp. L.) transformation mediated by Agrobacterium tumefaciens. Transgenic Research, vol. 7, 213-22; Enriquez-Obregon G., et al. , 1998. Herbicide-resistant sugarcane (Saccharum officinarum L.) plants by Agrabacterium-mediated transformation. 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. A revised medium for rapid growth and bioassays with tobacco tissue cultures. Physiol. Plant, vol. 15, 473-497) based medium incl. B5 vitamins (Gamborg, O., et al., 1968. Nutrient requirements of suspension cultures of soybean root cells. 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 calli 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 ished with sterile water followed by a non-selective period on similar medium containing 500 mg/l cefotaxime for eliminating the Agrobacterium. 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 international application published as WO2010/151634A and the granted European patent EP1831378.

Example X: Phenotypic evaluation procedure

10.1 Evaluation setup

Approximately 35 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 shorts 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.

10.2 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.

Because two experiments with overlapping events were carried out, a combined analysis was performed. This is useful to check consistency of the effects over the two experiments, and if this is the case, to accumulate evidence from both experiments in order to increase confidence in the conclusion. The method used was a mixed-model approach that takes into account the multilevel structure of the data (i.e. experiment - event - segregants). P values were obtained by comparing likelihood ratio test to chi square distributions.

10.3 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.

Early seedling vigour is the seedling aboveground area post-germination (plantlets of about 4 cm high).

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 XI: Results of the phenotypic evaluation of the transgenic plants

XI-1. TLP (Tify like protein) polypeptide

The results of the evaluation of transgenic rice plants in the T2 generation and expressing a nucleic acid comprising the longest Open Reading Frame in SEQ ID NO: 1 under non-stress conditions are presented below. See previous Examples for details on the generations of the transgenic plants.

The results of the evaluation of transgenic rice plants under non-stress conditions are presented below. An increase of 5 % or more was observed for aboveground biomass (AreaMax), total seed yield (Totalwgseeds) number of filled seeds (nrfilledseed), number of flowers per panicle (flowerperpan), an increase of 3% was observed for thousand kernel weight (TKW). In addition, plants expressing a TLP nucleic acid showed an increased height, an increased height of the gravity center, increased seedling biomass, an increased proportion of filled seed over the number of florets.

Table D1: Data summary for transgenic rice plants; for each parameter, the overall percent increase is shown for the confirmation (T2 generation), for each parameter the p-value is <0.05.

Parameter	Overall increase
AreaMax	5.4
totalwgseeds	9.7
flowerperpan	5.0
TKW	3.0
nrfilledseed	6.7

XI-2. PMP22 polypeptide (22 kDa peroxisomal membrane like polypeptide)

The results of the evaluation of transgenic rice plants in the T2 generation and expressing a nucleic acid as shown in SEQ ID NO: 50 under non-stress conditions are presented below. See previous Examples for details on the generations of the transgenic plants.

The results of the evaluation of transgenic rice plants under non-stress conditions are presented below (see Table D2a). An increase of more than 5 % was observed for aboveground biomass (AreaMax), number of flowers per panicle (flowerperpan), and of more than 3% for thousand kernel weight (TKW). In addition, plants expressing the PMP22 nucleic acid showed increased seed yield per plant (totalwgseeds), increased seed yield per leafy biomass (harvestindex), and a heightened gravity centre (GravitYMax).

Table D2a (results of phenotypic evaluation under non-stress conditions). Data summary for transgenic rice plants; for each parameter, the overall percent increase is shown for the confirmation (T2 generation), for each parameter the p-value is <0.05.

Parameter	Overall increase
AreaMax	5.1
flowerperpan	6.9
TKW	3.1

The results of the evaluation of transgenic rice plants expressing a PMP22 nucleic acid under nitrogen deficiency are presented hereunder (Table D2b). An increase of yield related parameters was observed. In particular, an increase was observed for the seed fillrate (number of filled seeds over the number of florets), number of flowers per panicle (flowerperpan), and thousand kernel weight (TKW). Moreover, increases were observed for the following parameters: final biomass (AreaCycle), number of total seeds (nrtotalseed), number of filled seeds per plant (nrfilledseed), seed yield per plant (totalwgseeds), seed yield per leafy biomass (harvestindex). Also the gravity centre of the plants was heightened (GravitYMax).

Table D2b (results of phenotypic evaluation under nitrogen deficient conditions). Data summary for transgenic rice plants; for each parameter, the overall percent increase is shown for the confirmation (T2 generation), for each parameter the p-value is <0.05.

Parameter	Overall increase
fillrate	7.4

XI-3. RTF (REM-like transcription factor) polypeptide

The results of the evaluation of transgenic rice plants in the T2 generation and expressing a nucleic acid encoding the RTF polypeptide of SEQ ID NO: 140 under non stress conditions are presented below in Table D3. When grown under non-stress conditions, an increase of at least 5 % was observed for aboveground biomass (AreaMax), root biomass (RootMax and RootThickMax). Moreover, a significantly improved early vigour (Emervigor) was observed since the aboveground area of the leafy biomass was increased by more than 10% as compared to control plants. In addition, plants expressing a RTF nucleic acid had an earlier start point of flowering (a shorter time (in days) needed between sowing and the emergence of the first pani-

cle (timetoflower). Also the number of seeds (nrtotalseeds), the total weight of the seeds (totalwgseeds), and the number of filled seeds (nrfilledseeds) was increased. Moreover, they had an later start point of flowering (a longer time (in days) needed between sowing and the emergence of the first panicle (timetoflower) Moreover, the plants had an increased height as compared to control plants (GravityYMax) in three events

Table D3: Data summary for transgenic rice plants; for each parameter, the overall percent increase is shown for the confirmation (T2 generation), for each parameter the p-value is <0.05.

Parameter	Overall increase [%]
AreaMax	6.4
EmerVigor	10.7
RootMax	7.5
RootThickMax	6.1

XI-4. BP1 (Bigger plant 1) polypeptide

Transgenic rice plants expressing a BP1 (SEQ ID NO: 170) nucleic acid under nitrogen deficient conditions showed an increase for the following yield related parameters: aboveground biomass (Areamax), root biomass (RootMax), total seed yield per plant (totalwgseeds), flowers per panicle (flowersperpan), number of filled seeds per plant (nrfilledseed), and number of thick roots (RootThickMax). For example, the Areamax was increased by 7 to 10 % (with a p-value between or equal to 0.2 and 0.1), and Rootmax values were increased from 10% to 13% with a p-value between or equal to to 0.2 and 0.1. In at least two events the RootThickMax value was increased around 10% with a p-value between or equal to 0.2 and 0.1.

Claims

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 TLP (Tify like protein) polypeptide, wherein said TLP polypeptide is selected from the group consisting of:
   a) a polypeptide comprising a sequence as shown in SEQ ID NO: 2,
   b) a polypeptide having at least 50% sequence identity to a polypeptide as represented by SEQ ID NO: 2, and
   c) a polypeptide encoded by a polynucleotide which hybridizes under stringent conditions to a polynucleotide having a sequence as shown in SEQ ID NO: 1, or to a complementary sequence of such a polynucleotide having a sequence as shown in SEQ ID NO: 1.

2. The method according to claim 1, wherein
   (i) said TLP polypeptide comprises a Pfam domain having the Pfam accession number PF06200 and/or a Pfam domain having the accessing number PF09425, preferably both domains, and/or
   (ii) the TLP polypeptide comprises an Interpro domain having the Interpro accession number IPR010399 and/or an Interpro domain having the Interpro accession number IPR018467, preferably both domains.

3. Method according to claims 1 and 2, wherein said TLP polypeptide comprises
   a. all of the following motifs:
   (i) Motif 1-1: (SEQ ID NO: 35):
   QLTIFY[AG]G[SM]V[NC]V[YF][DE][DN][IV]S[PA]EKAQ[AE][IL]M,
   (ii) Motif 2-1: (SEQ ID NO: 37):
   PQARKASLARFLEKRKERV[MT][NST][TAL][AS]PY,
   (iii) Motif 3-1: (SEQ ID NO: 39):
   MERDF[LM]GL[NGSI][IS]K[DEN][PS][LP][LA][VT][VI]K[DE]Exxx[SD][SG],
   (iv) Motif 4-1 (SEQ ID NO: 40)
   Q[LM]TIFY[AG]G[SMATL]V[NCS][VI][YF][DEN][DN][IV][STP][PAV][ED][KQ]A[QK][AE][IL]MFLA[GS][HNR],
   (v) Motif 5-1 (SEQ ID NO: 43):RFLEKRKE
   (vi) Motif 6-1 (SEQ ID NO: 44): QLTIFY[AG]G
   (vii) Motif 7-1 (SEQ ID NO: 45):MERDF[LM]GL; or
   b. all of the motifs 2-1 to 7-1 as defined in a. above, and in addition the Motif 1-1a) (SEQ ID NO: 36):
   QLTIFYGGMV[NC]V[YF]E[DN][IV]S[PA]EKAQ[AE][IL]M ; or
   c. all of the motifs 1-1 and 3-1 to 7-1 as defined in a. above, and in addition the Motif 2-1a) (SEQ ID NO: 38)
   PQARKASLARFLEKRKERV[MT][NST]L[AS]PY ; or

d. all of the motifs 1-1 to 7-1 as defined in a. above, wherein motif 4-1 is replaced by the Motif 4-1a) (SEQ ID NO: 41)
Q[LM]TIFY[AG]G[SMATL]V[NCS][VI][YF][DEN][DN][IV][STP][PAV][ED],
and/or Motif 4b) (SEQ ID NO: 42):
[KQ]A[QK][AE][IL]MFLA[GS][HNR] ; or
e. all of the motifs 1-1a), 2-1a), 3-1, 4-1a) and 4-1b), 5-1 to 7-1 as defined in a. to d. above; or
f. any three, preferably any four , more preferably any 5 motifs as defined in a. to d. above; or
g. any combination of motifs as defined in f. wherein Motifs 1-1, 2-1 and 4-1 are not present; or
h. any motif as defined in a. to d above.

4. Method according to any one of claims 1 to 3, wherein said modulated expression is effected by introducing, preferably recombinantly introducing and expressing in a plant said nucleic acid encoding said TLP polypeptide.

5. Method according to claim 1 or 4, 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.

6. Method according to any one of claims 1 to 4, wherein said enhanced yield-related traits are obtained under non-stress conditions and/or wherein said enhanced yield-related traits are obtained under conditions of drought stress, salt stress and/or nitrogen deficiency.

7. Method according to any one of claims 1 to 6, wherein said nucleic acid encoding a TLP encodes any one of the polypeptides listed in Table A1.

8. Method according to any one of claims 1 to 7, wherein said nucleic acid encodes the polypeptide represented by SEQ ID NO: 2.

9. Construct comprising:
(i) nucleic acid encoding a TLP as defined in any of claims 1 to 3, 7 and 8;
(ii) one or more control sequences capable of driving expression of the nucleic acid sequence of (i); and optionally
(iii) a transcription termination sequence.

10. Construct according to claim 9, wherein one of said control sequences is a constitutive promoter, preferably a medium strength constitutive promoter, preferably to a plant promoter, more preferably a GOS2 promoter, most preferably a GOS2 promoter from rice.

11. Use of a construct according to claim 9 or 10 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.

12. The construct according to claim 9 or 10 comprised in a plant, plant part or plant cell.

13. 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 a TLP polypeptide as defined in any of claims 1 to 3, 7 and 8; and
(ii) cultivating said plant cell or plant under conditions promoting plant growth and development.

14. 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 recombinantly increased expression of a nucleic acid encoding a TLP polypeptide as defined in any of claims 1 to 3, 7 and 8 or a transgenic plant cell derived from said transgenic plant.

15. The transgenic plant according to claim 14 or a transgenic plant cell derived therefrom, wherein said plant is a crop plant, preferably
- a dicot crop plant such as soybean, cotton, oilseed rape including canola, beet, sugarbeet or alfalfa; or
- a monocotyledonous crop plant such as sugarcane or a cereal, such as rice, maize, wheat, barley, millet, rye, triticale, sorghum, emmer, spelt, einkorn, teff, milo or oats.

16. Use of a nucleic acid encoding a TLP polypeptide as defined in any of claim 1 to 3, 7 and 8 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.

17. Isolated polynucleotide selected from the group consisting of:
a) a polynucleotide encoding a polypeptide comprising a sequence as shown in SEQ ID NO: 2,
b) a polynucleotide encoding a polypeptide having at least 50% sequence identity to a polypeptide as represented by SEQ ID NO: 2, and
c) a polynucleotide comprising a sequence as shown in SEQ ID NO: 1
d) a polynucleotide having at least 50% sequence identity to a polynucleotide as represented by SEQ ID NO: 1, and

e) a polynucleotide which hybridizes under stringent conditions to a polynucleotide having a sequence as shown in SEQ ID NO: 1, or to a complementary sequence of such a polynucleotide having a sequence as shown in SEQ ID NO: 1.

18. Isolated polypeptide selected from the group consisting of:

a) a polypeptide comprising a sequence as shown in SEQ ID NO: 2,

b) a polypeptide having at least 50% sequence identity to a polypeptide as represented by SEQ ID NO: 2, and

c) a polypeptide encoded by a polynucleotide which hybridizes under stringent conditions to a polynucleotide having a sequence as shown in SEQ ID NO: 1, or to a complementary sequence of such a polynucleotide having a sequence as shown in SEQ ID NO: 1.

19. 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 PMP22 (22 kDa Peroxisomal Membrane like protein) polypeptide, an RTF (REM-like transcription factor) polypeptide, or a BP1 polypeptide,

wherein said PMP22 polypeptide is selected from the group consisting of:

(i) a polypeptide comprising a sequence as shown in SEQ ID NO: 51, 57, 91 or 105,

(ii) a polypeptide having, in an increasing order of preference, at least 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 a polypeptide as represented by SEQ ID NO: 51, 57, 91 or 105 when compared over the entire length of the amino acid sequence as represented by SEQ ID NO: 51, 57, 91 or 105, respectively, and

(iii) a polypeptide encoded by a polynucleotide which hybridizes under stringent conditions to a polynucleotide having a sequence as shown in SEQ ID NO: 50, 56, 90, or 104 or to a complementary sequence of such a polynucleotide having a sequence as shown in SEQ ID NO: 50, 56, 90, or 104 ; and

wherein said RTF polypeptide is encoded by a nucleic acid selected from the group consisting of:

(i) a nucleic acid represented by any one of SEQ ID NO: 139, 141, 143, 145, 147, 149, 151, 153, 155, 157, 159, 161, or 163;

(ii) the complement of a nucleic acid represented by any one of SEQ ID NO: 139, 141, 143, 145, 147, 149, 151, 153, 155, 157, 159, 161, or 163;

(iii) a nucleic acid encoding the polypeptide as represented by any one of SEQ ID NO: 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, or 164, preferably as a result of the degeneracy of the genetic code, said isolated nucleic acid can be deduced from a polypeptide sequence as represented by any one of SEQ ID NO: 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, or 164;

(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 any of the nucleic acid sequences of SEQ ID NO SEQ ID NO: 139, 141, 143, 145, 147, 149, 151, 153, 155, 157, 159, 161, or 163,

(v) a nucleic acid which hybridizes with the nucleic acid molecule of (i) to (iv) or with a nucleic acid complementary thereto under stringent hybridization conditions, and

(vi) a nucleic acid encoding a polypeptide having, in increasing order of preference, at least 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 any one of SEQ ID NO: 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, or 164, and

wherein the expression of said nucleic acid encoding a BP1 polypeptide is modulated by recombinant insertion of an expression cassette comprising a nucleic acid encoding a BP1 polypeptide, and said BP1 polypeptide is selected from the group consisting of:

(i) a polypeptide comprising a sequence as shown in SEQ ID NO: 171,

(ii) a polypeptide which has, in an increasing order of preference, at least 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: 171, wherein the BP1 polypeptide is, preferably, not the polypeptide encoded by the nucleic acid sequence disclosed as SEQ ID NO: 75649 in the patent application published as US20030135870, and

(iii) a polypeptide encoded by a polynucleotide which hybridizes under stringent conditions to a polynucleotide having a sequence as shown in SEQ ID NO: 170, or to a complementary sequence of such a polynucleotide having a sequence as shown in SEQ ID NO: 170, wherein the BP1 polypeptide is, preferably, not the polypeptide encoded by the nucleic acid sequence disclosed as SEQ ID NO: 75649 in the patent application published as US20030135870.

20. The method of claim 19, wherein the polypeptide is the PMP22 polypeptide, and wherein said PMP22 polypeptide comprises a PFAM domain having the PFAM accession number PF04117 and/or an Interpro domain having the Interpro Accession number IPR007248.

21. The method of claims 19 and 20, wherein the polypeptide is the PMP22 polypeptide, and wherein said PMP22 polypeptide comprises

a. all of the following motifs:
Motif 1-2 (SEQ ID NO: 126):
GDWIAQC[YF]EGKPLFE[FI]DR[AT]RM[FL]RSGLVGFTLHGSLSHYYY[QH]FC E[AE]LFPF[QKE]
Motif 2-2 (SEQ ID NO: 127):
LTID[HQ]DYWHGWT[LI][FY]EILRY[AM]P[QE]HNW[VSI]AYE[EQ]ALK[RTA]NP VLAKM
Motif 3-2 (SEQ ID NO: 128): [DE]WWVVP[AV]KVAFDQT[VA]W[SA]A[IV]WN
Motif 4-2 (SEQ ID NO: 129):
LVGFT-
LHGSLSHYYY[QH][FIL]CEALFPF[QKE][DE]WWVVP[AV]KVAFDQT[VI]WSAIW NSIYF
Motif 5-2 (SEQ ID NO: 130):
RY[AM]P[EQ]HNW[ISV]AYE[EQ]ALK[AR]NPVLAKM[VAM]ISG[VI]VYS[LIV]GD WIAQCYEGKP[LI]F[ED][FI]D
Motif 6-2 (SEQ ID NO: 131): AHL[IV]TYG[VL][IV]PVEQRLLWVDC
Motif 7-2 (SEQ ID NO: 132):
RYAPQHNW[IV]AYEEALK[RQ]NPVLAKMVISGVVYS[VL]GDWIAQCYEGKPLF[ED][IF]D
Motif 8-2 (SEQ ID NO: 133):
GFT-
LHGSLSH[YF]YYQFCE[AE]LFPF[QE]DWWVVP[VA]KVAFDQTVWSAIWNSIY[FY]TV
Motif 9-2 (SEQ ID NO: 134):
F[LW]PMLTAGWKLWPFAHLITYG[VL][VI]PVEQRLLWVDCVEL[IV]WVTILSTY SNEK;
or

b. at least one of the Motifs 7-2 to 9-2, preferably any two of Motifs 7-2 to 9-2, more preferably all three of Motifs 7-2 to 9-2 as defined in a. above; or

c. at least one of the Motifs 4-2 to 6-2, preferably any two of the Motifs 4-2 to 6-2, more preferably all three of the Motifs 4-2 to 6-2 as defined in a. above; or

d. at least one of the Motifs 1-2 to 3-2, preferably any two of the Motifs 1-2 to 3-2, more preferably all three of the Motifs 1-2 to 3-2 as defined in a. above; or

e. any four of the Motifs 1-2 to 9-2, preferably any five of the Motifs 1-2 to 9-2 as defined in a. above; or

f. any six of the Motifs 1-2 to 9-2, preferably any seven of the Motifs 1-2 to 9-2, more preferably any eight of the Motifs 1-2 to 9-2 as defined in a. above.

22. The method of claim 19, wherein the polypeptide is the RTF polypeptide, and wherein the RTF polypeptide comprises at least two B3 PFAM domains, in particular four B3 PFAM domains, having the PFAM accession number Pfam02362.

23. The method of claims 19 and 22, wherein the polypeptide is the RTF polypeptide, wherein said RTF polypeptide comprises one or both of the following motifs:
    (i) Motif 1: PVAFF (SEQ ID NO: 165),
    (ii) Motif 2: HDLRVGDIVVF (SEQ ID NO: 166).

24. The method of any one of claims 19, 22 and 23, and wherein the polypeptide is the RTF polypeptide, and wherein the polypeptide comprises both Motifs 1 and 2 as defined in claim 25, and four B3 PFAM domains having the PFAM accession number Pfam02362.

25. The method of claim 19, wherein the polypeptide is the BP1 polypeptide, wherein said BP1 polypeptide comprises

a) all of the following motifs:
    (i) Motif 1-4: LNQ[DG]SXXND[EV]X[NS]DX[QP]G[HQ]X[GN]H[LP]EXXKX[DE][QE][VA][GE]VXE[DE]X[MI][TA][AP]DV[KN]LS[VA]CRDTG[NE] (SEQ ID NO: 276),
    (ii) Motif 2-4: L[WR]RDYXD[LV][LV][QK][ED][TN]EXK[KR][KR]XLXSX[KN][RK][RT][KS]L[AV]LL[AS]EVKFL[RQ][RK]K[YL]XSF[AKLP]K[GN][GDN]SQ[QK] (SEQ ID NO: 277), and
    (iii) Motif 3-4: [DE][DG]KRX[VI][PS]WQD[RQ]XALK (SEQ ID NO: 278);
    (iv) Motif 4-4 as disclosed as SEQ ID NO: 279;
    (v) Motif 5-4 as disclosed as SEQ ID NO: 280;
    (vi) Motif 6-4 as disclosed as SEQ ID NO: 281; or
b) any two of the Motifs 1-4 to 6-4, preferably any two of Motifs 4-4 to motif 6-4 as defined in a) above; or
c) any three of the Motifs 1-4 to 6-4, preferably all three of Motifs 4-4 to motif 6-4 as defined in a) above; or
d) any one of the Motifs 1-4 to 6-4, preferably any two of Motifs 4-4 to motif 6-4 as defined in a) above.

26. The method of claims 19 or 25, wherein the expression of the nucleic acid encoding a BP1 polypeptide is increased.

27. The method according to any one of claims 19 to 26, wherein said modulated expression is effected by introducing, preferably recombinantly introducing, and expressing in a plant said nucleic acid encoding said PMP22, RTF, or BP1 polypeptide.

28. Method according to any one of claims 19 to 21, wherein the polypeptide is the PMP22 polypeptide, and 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.

29. Method according to any one of claims 19 and 22 to 24, wherein the polypeptide is the RTF polypeptide, and wherein said enhanced yield-related traits comprise improved early vigour and increased yield, in particular increased biomass relative to control plants.

30. Method according to any one of claims 19, 25 and 26, wherein the polypeptide is the BP1 polypeptide, wherein said enhanced yield-related traits comprise increased yield relative to control plants, and preferably comprise increased biomass, increased shoot biomass, increased root biomass, increased NUE (nitrogen use efficiency) and/or increased seed yield relative to control plants.

31. Method according to any one of claims 19 to 30, wherein said enhanced yield-related traits are obtained under non-stress conditions, and/or are obtained under conditions of drought stress, salt stress and/or nitrogen deficiency.

32. Method according to any one of claims 19 to 31, wherein
    (i) the nucleic acid encoding the PMP22 polypeptide encodes any one of the polypeptides listed in Table A2
    (ii) the nucleic acid encoding the RTF polypeptide encodes any one of the polypeptides listed in Table A3, or
    (iii) the nucleic acid encoding the BP1 polypeptide encodes any one of the polypeptides listed in Table A4, preferably a polypeptide represented by SEQ ID NO: 171, 239, 243 or 267

33. Method according to any one of claims 19 to 32, wherein
    (i) the polypeptide is the PMP22 polypeptide as represented by SEQ ID NO: 51,
    (ii) the polypeptide is the RTF polypeptide as represented by SEQ ID NO: 140, or
    (iii) the polypeptide is the BP1 polypeptide as represented by SEQ ID NO: 171.

34. Plant, plant part thereof, including seeds, or plant cell, obtainable by a method according to any one of claims 19 to 33, wherein said plant, plant part or plant cell comprises a recombinant nucleic acid encoding a PMP22, RTF or BP1 polypeptide as defined in any of claims 19 to 25, 31 and 32.

35. Construct comprising:

(i) nucleic acid encoding a PMP22, RTF, or BP1 polypeptide as defined in any of claims 19 to 25, 31 and 32;

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

(iii) a transcription termination sequence.

36. Construct according to claim 35, wherein one of said control sequences is a constitutive promoter, preferably a medium strength constitutive promoter, preferably to a plant promoter, more preferably a GOS2 promoter, most preferably a GOS2 promoter from rice.

37. The construct according to claims 35 and 36 comprised in a plant, plant part or plant cell.

38. Use of a construct according to claims 35 and 36 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.

39. Method for 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 a PMP22 or BP1 polypeptide as defined in any of claims 19 to 21, 25, 32 and 33; and

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

40. Method for the production of a transgenic plant having enhanced yield-related traits relative to control plants, preferably improved early vigor or increased yield relative to control plants, and more preferably increased biomass relative to control plants, comprising:

(i) introducing and expressing in a plant cell or plant a nucleic acid encoding a RTF polypeptide as defined in any of claims 19, 22 to 24, 32 and 33; and

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

41. Use of a nucleic acid encoding a PMP22, RTF, or BP1 polypeptide as defined in any one of claims 19 to 25, 32 and 33 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.

42. 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, preferably resulting from recombinantly increased expression

of a nucleic acid encoding a PMP22, RTF, or BP1 polypeptide as defined in any of claims 19 to 25, 32 and 33 or a transgenic plant cell originating from said transgenic plant and comprising a nucleic acid encoding a PMP22, RTF, or BP1 polypeptide as defined in any of claims 19 to 25, 32 and 33

43. The use of claims 38 or 41, the plant of plant part or plant cell of claim 34, or the transgenic plant, or plant cell of claim 42, wherein
    (i) the polypeptide is the PMP22 polypeptide and wherein the 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,
    (ii) the polypeptide is the RTF polypeptide, and wherein the polypeptide is the RTF polypeptide, and wherein said enhanced yield-related traits comprise improved early vigour and increased yield, in particular increased biomass relative to control plants, or
    (iii) the polypeptide is the BP1 polypeptide, and wherein said enhanced yield-related traits comprise increased yield relative to control plants, and preferably comprise increased biomass, increased shoot biomass, increased root biomass, increased NUE (nitrogen use efficiency) and/or increased seed yield relative to control plants.

44. An isolated nucleic acid molecule selected from:
    (i) a nucleic acid represented by SEQ ID NO:50, 56, 90, or 104;
    (ii) the complement of a nucleic acid represented by SEQ ID NO: 50, 56, 90, or 104;
    (iii) a nucleic acid encoding a PMP22 polypeptide having in increasing order of preference at least 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%, 99%, or 100% sequence identity to the amino acid sequence represented by SEQ ID NO: 51, 57, 91 or 105, 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.

45. An isolated polypeptide selected from:
    (i) an amino acid sequence represented by SEQ ID NO:51, 57, 91 or 105;
    (ii) an amino acid sequence having, in increasing order of preference, at least 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: 51, 57, 91 or 105, and preferably conferring enhanced yield-related traits relative to control plants; and

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

MERDFMGLNIKDSLLVVKDEPVESSKDSGFRWPMSSKVGVPHFMSLNSAQDENTFKALSATDGVDAGL
----------3-------------
----7---

KRQPGELQNVHAMHLPYDVKMLPFNMNNPSYKTHFGGIGQMKQVLGGIPVTAPHSMLPSRGSVAGTTE

PWFNSKG**SAAPAQLTIFYGGMVNVFEDISPEKAQAIMFLAGH**GCAPPNVVQPRFQLQASASKPAAADG
---------1---------------
-----------4------------------
--------4a---------
----6--- ----4b-----

VCVNQTPNMLPASGLSSTMSVSSHPIGQSDGSSGNKDDMKMSKTANISVTPHVKLDTSKIVTSLGPVG

ATTIMTAAV**PQARKASLARFLEKRKERVMNLAPY**GLSKKSPECSTPESNGVGFSATSTPLLAGKET
---------2---------------
---5----

Fig. 1

```
G.max_Glyma05g34960.1                  ------------MERDFMGLNLKEP-LAVVKEEMN-NDG----------
G.max_Glyma08g04770.1                  ------------MERDFMGLNSKEP-LAVVKEEMH-VYG----------
M.truncatula_TC119816                  ------------MERDFLSLCSKE-----SSPEIN-NEG----------
P.trichocarpa_656962                   ------------MERDFLGLSSREP-LAVVKEEVN-ADG----------
P.trichocarpa_scaff_VIII.1248          ------------MERDFLGLSSREP-LAVVKEEVQGCSGPSLTRSPPLV
P.trichocarpa_scaff_X.1014             ------------MERDFLGLSSKKP-SAVVKEEIS-SDG----------
P.trichocarpa_TC117359                 ------------MERDFLGLSSKKP-SAVVKEEIS-SDG----------
S.lycopersicum_TC193363                ------------MERDFMGLNIKDS-LLVVKDEPVES------------
Le_UNK                                 ------------MERDFMGLNIKDS-LLVVKDEPVES------------
A.thaliana_AT3G17860.1                 ------------MERDFLGLGSKNS-PITVKEETSESSR----------
B.napus_TC100484                       ------------MERDFLGLGSKNS-PITVKEETSESSR----------
P.trichocarpa_scaff_XV.429             ------------MERDFLGLGSKNN-PVTIKEEATDTPLK---------
S.lycopersicum_TC207683                ------------MERDFMGL--------TVKQEVLEEPI-----------
G.max_Glyma09g30460.1                  ------------MEREFFGLSSKNGAWTTMKDDAVNKSR----------
M.truncatula_TC118189                  ------------MEREFLGFSSKNSPWTTMKEDASNKPK----------
Hordeum_vulgare_subsp_vulgare_         ------------MERDFLGAIGHEQQQLLMQQQQQRQRAAAEDDAARKE
T.aestivum_TC290287                    ------------MERDFLGTIGHEQLQQ-QQQQQQRQRAAAEDAAARKE
Zea_mays_GRMZM2G126507_T01             MAKSGASFPESSWMERDFLAAIGKE------QQHPHKEEAGAE------E
O.sativa_LOC_Os08g33160.1              ------------MERDFLGAIWRKE-----EAAGKPEEHSVS-----RD
                                                   ***:*:.          .

G.max_Glyma05g34960.1                  -----CKNSGFKKGRIAQWPFSNKVSALP---HLMSFKASQDDKTKNTVS
G.max_Glyma08g04770.1                  -----FEDSGFTKGRIAQWPFSNKVSALP---HLMSFKASQDDKTKKIVS
M.truncatula_TC119816                  -----SKNSGFSNVSAVKWPFLNKVAVHS---YLTPFKVSEDDKAKMISS
P.trichocarpa_656962                   -----CKESG----SGMQWPFSNKVSTP----------------------
P.trichocarpa_scaff_VIII.1248          TPRLKSRQGGTHF-SLTAYPVQHDVHS----------------------
P.trichocarpa_scaff_X.1014             -----CKDIGFTKGSGMHWPFSNKVSTLH---NLMSFKAAQEDKTKTIES
P.trichocarpa_TC117359                 -----CKDIGFTKGSGMHWPFSNKVSTLH---NLMSFKVAQEDKTKTIES
S.lycopersicum_TC193363                -----SKDSGFR------WPMSSKVGVPH---FMSLNSAQDENTFKALSA
Le_UNK                                 -----SKDSGFR------WPMSSKVGVPH---FMSLNSAQDENTFKALSA
A.thaliana_AT3G17860.1                 ---------DSAPNRGMNWSFSNKVSASSS-QFLSFRPT-QEDRHRKSGN
B.napus_TC100484                       ---------DSAPNRGMNWSFSKKGSAASS-QFLSFRPS-QDDRHRKPGN
P.trichocarpa_scaff_XV.429             -------DSVPMRGSGMQWSFSNKVSAIP--QFLSFKSS-MEDKPRKAVH
S.lycopersicum_TC207683                -------DPAPLRSSAMQWSFTNNVTAHP--QYLSFKSA-PEDKP-KIGF
G.max_Glyma09g30460.1                  ---------DQVRSSGMQWSFPNKVSALP--QFLSFKTN-QEDKPRKTIL
M.truncatula_TC118189                  ---------DQVRSSGMQWSFSNKVSNLP--QFMSFKNNTHEDRSRNNVM
Hordeum_vulgare_subsp_vulgare_         SAYF-GGGG----VPPMDWSF-AGRAGAAP-AVMSLRS---AAREEQQQQ
T.aestivum_TC290287                    SAYF-GGGG----VPPMDWSF-AGRAGAAP-AVMSFRS---APREEQR--
Zea_mays_GRMZM2G126507_T01             SAYF-GGAGAAAAAPAMDWSF-ASKPGAAP-ALMSFRS---ASFPQFS--
O.sativa_LOC_Os08g33160.1              ADYR-GGGGG--ASAAMQWQFPATKVGAASSAFMSFRSSAAAAREEDPKE
                                                          : .

G.max_Glyma05g34960.1                  DTLSSS---GFMSILSQEAFDTSQKRSAGEPQKSFNHDGQ-----GDLHF
G.max_Glyma08g04770.1                  DPLTSA---GFMSILSQEAFDSSQKRCAGEPQKSFNHDGQ-----GRLHF
M.truncatula_TC119816                  G-------------------------------------------------
P.trichocarpa_656962                   ------------------------RHSTAEIQ-----------------
P.trichocarpa_scaff_VIII.1248          ------------------------MHHPHDVK-----------------
P.trichocarpa_scaff_X.1014             DALVSS---GFMSVLSADACDPGQKRSAAEIQKLFNLNRQ-----GKTHF
P.trichocarpa_TC117359                 DALVSS---GFMSVLSADACDPGQKRSAAEIQ-----------------
S.lycopersicum_TC193363                TDGVDA---G-------------LKRQPGELQ-----------------
Le_UNK                                 TDGVDA---G-------------LKRQPGELQNV----------------
A.thaliana_AT3G17860.1                 YHLPHS--GSFMPSSVADVYDSTRKAPYSSVQG-------------VRM
B.napus_TC100484                       YHLTHS--GSFMPSSVADGYDSNRNTPYSSVQG-------------ARM
P.trichocarpa_scaff_XV.429             DPMASSS-SGYMSISTADAFDSNQKSYSALIQKNMALDKQ-AGNHYAMTT
S.lycopersicum_TC207683                DSLAS---TGLVTITTTTEAVDSSHRTYSDVTQKNMMLERQ-GGTHYTTTT
G.max_Glyma09g30460.1                  EPLAS---SGYMAMSTQYAFDSNQKSFLGLTNRNLSISKHAAGNKQGMTV
M.truncatula_TC118189                  DPVAS---SGYMTISTKDAFDSNQKSFLGVTQENLAIKQ--VGNKHGITI
Hordeum_vulgare_subsp_vulgare_         QGELAFSKQQASR-------------VLTPQRSFGAENHGSVQYAAAAR
T.aestivum_TC290287                    -GELAYPKQQASR-------------VLTPQRSFGAESHGSVQYAAAAR
Zea_mays_GRMZM2G126507_T01             --SFDGAKNPAPR-------------ILTHQRSFGPD---STHYAAAHR
O.sativa_LOC_Os08g33160.1              AAVFDRFSLSGFRPPPRPSPGDAFDGAAAMKQRQFGFN--GRQQYAAAAQ

G.max_Glyma05g34960.1                  SLTSYPVQHDVN---HPHDVKMFSVPNQAISVSLGNPFLKNHFAAAGQ--
G.max_Glyma08g04770.1                  SLTPYPVQHDVNSVNRPHDVKMLSVPNQAISLSLGNPFLKNDFATAGQ--
M.truncatula_TC119816                  ------------------------------------FIQNINGANAK--
P.trichocarpa_656962                   --------------------MFPVSNHAIPISMGNHFFKNHYPATGQNM
P.trichocarpa_scaff_VIII.1248          --------------------MFPVSNHAIPISMGNHFFKNHYPATGQNM
```

Fig. 2

```
P.trichocarpa_scaff_X.1014          SPTAFPAQHDVHSVHHPHDMKMFPVANHAISISTGNPFFNNHYPATGQNM
P.trichocarpa_TC117359              ----------MFPVANHAISISTGNPFFNNHYPATGQNM
S.lycopersicum_TC193363             ----------
Le_UNK                              ----------HAMHLPYDVKMLPFNMNNPSYKTHFGGIGQ--
A.thaliana_AT3G17860.1              FPNSN-------QHEETNAVSMSMPGFQSH----------HYAPGGRSF
B.napus_TC100484                    FPNSH-------QQQESITVSMARPGLQS------------HYPPGGKSF
P.trichocarpa_scaff_XV.429          YGKQHFDAYFANRPQDMRMFPISSQQNQTINVSMSSPILQSLFPPTGHSM
S.lycopersicum_TC207683             FSPHHYDAHSMHRSHGVRVLPLASPTNQ-ISVSMTMPGHKSFISPVGQNL
G.max_Glyma09g30460.1               YPLQCCDAQ----SEEARIFSVSNQSNQ------VSPVLQSNLASTGLNM
M.truncatula_TC118189               YPIQSSDAQSVC-NQEARTISVSNQSNH------V---------ITGINM
Hordeum_vulgare_subsp_vulgare_      AAYGGQPPQQ------QHAPNGARVIPMSSPFNPNNPMFRVQSSPNLPNG
T.aestivum_TC290287                 AAYGGQPPQQH-----QHAPNGARVIPMSSPFNPNNPMFRVQSSPNLPNG
Zea_mays_GRMZM2G126507_T01          T---------------QHALNGARVTPVSSPFNQNSPMFRVQSSPSLPN-
O.sativa_LOC_Os08g33160.1           HGHREQGVDSYGVAAPHHFPSPSPSPRHPVPFGHANPMLRVHSLPNVAG-

G.max_Glyma05g34960.1               -------KPLLGGIPVTTS-HSVLPSAVAVAGMTESCNS---VKP---SA
G.max_Glyma08g04770.1               -------KLLLGGIPVTTP-HSVLPSAVAVAGMTESCNT---VKPSAPSA
M.truncatula_TC119816               -------QSLLGGLPVTAP-HSVLPIVGTVAGLVEPCE-----KPSAPAP
P.trichocarpa_656962                AGT-TTKPQLLGGIPVTAP-HSILPMVGSVAGVTDSSV-----RASGSPA
P.trichocarpa_scaff_VIII.1248       AGT-TTKPQLLGGIPVTAP-HSILPMVGSVAGVTDSSV-----RASGSPA
P.trichocarpa_scaff_X.1014          FGT-TMKPQLLGGFPVTAP-HSILPMVGPVAGVTDSSV-----KAYGSPA
P.trichocarpa_TC117359              FGT-TMKPQLLGGFPVTAP-HSILPMVGPVAGVTDSSV-----KAYGSPA
S.lycopersicum_TC193363             ------MKQVLGGIPVTAP-HSMLPSRGSVAGTTEPWFN---SKGSAAPA
Le_UNK                              ------MKQVLGGIPVTAP-HSMLPSRGSVAGTTEPWFN---SKGSAAPA
A.thaliana_AT3G17860.1              MNN-NNNSQPLVGVPIMAPPISILPPPGSIVGTTDIR---SSSKPIGSPA
B.napus_TC100484                    MSS-GINSQPFVGVPIMAPPISVLPAPGSIVGTTDIR---SSSKPLGSPA
P.trichocarpa_scaff_XV.429          ITSNSIVSKPLGGVPVITP-ASALPTPSSVIGTTDLR--DVAKSS-GAPA
S.lycopersicum_TC207683             ITT--VNQLPGAGALVVSP-ISAVPSSSIVVGTTDLR--GAPKTP-PGPA
G.max_Glyma09g30460.1               VNS-VIKPQPFGSKSSGTP-LSILPSIGSIVGSTDLR--NNSKSS-TMPT
M.truncatula_TC118189               VNS-VTNSQTFGSKSSATP-LSVLPSKGPIVGSTDLRSRNCSKSNGTTPA
Hordeum_vulgare_subsp_vulgare_      VAA---------GSPFKQP--PFAMNN-AVAASTVGVYK-SRDMPKPKTA
T.aestivum_TC290287                 VAA---------GSPFKQP--PFVMNN-AVAASTVGVYK-SRDTPKPKTA
Zea_mays_GRMZM2G126507_T01          ------------GTAFKQP--PFAINNNAAASSTVGFYG-TRDVVRPKTA
O.sativa_LOC_Os08g33160.1           ------------GSPYRNQ--SFSVGN-SVAGSTVGVYGGPRDLQNPKVT
                                                .       .       . .             .

G.max_Glyma05g34960.1               QLTIFYAGTVNIFDDISAE--------------KAQAIMLLAGN-SLSAA
G.max_Glyma08g04770.1               QLTIFYAGTVNVFDDISAE--------------KAQAIMLLAGN-SLSAG
M.truncatula_TC119816               QLTMFYGGTVNIFNDITPE--------------KAQAIMLLAGS-GLSAA
P.trichocarpa_656962                QLTIFYAGSVNVYDDISPE--------------KAQAIMFLAGN-GPSIS
P.trichocarpa_scaff_VIII.1248       QLTIFYAGSVNVYDDISPE--------------KAQAIMFLAGN-GPSIS
P.trichocarpa_scaff_X.1014          QLTIFYAGAVNVYDDISPEKMAILNLVSLRNPIQAQAIMFLAGN-GSSIS
P.trichocarpa_TC117359              QLTIFYAGAVNVYDDISPE--------------KAQAIMFLAGN-GSSIS
S.lycopersicum_TC193363             QLTIFYGGMVNVFEDISPE--------------KAQAIMFLAGH-G-CAP
Le_UNK                              QLTIFYGGMVNVFEDISPE--------------KAQAIMFLAGH-G-CAP
A.thaliana_AT3G17860.1              QLTIFYAGSVCVYDDISPE--------------KAKAIMLLAGN-GSSMP
B.napus_TC100484                    QLTVFYAGSVCVYDDISPD--------------KAKAIMLLAGN-GSSMP
P.trichocarpa_scaff_XV.429          QLTIFYAGSVSVYDDVSPE---------------KAQAIMLLAGNGGSSGT
S.lycopersicum_TC207683             QLTIFYGGSVCVYDNVSPE--------------KAQAIMLLAGN-APPVT
G.max_Glyma09g30460.1               QLTIFYAGSVCVYDDISPE--------------KAKAIMLMAGN-GYTPT
M.truncatula_TC118189               QLTIFYGGTVCVYDDISPE--------------KAKAIMLLAGN-GTKMQ
Hordeum_vulgare_subsp_vulgare_      QLTIFYAGSVNVFNNVSAE--------------KAQELMFLASRGSLPAA
T.aestivum_TC290287                 QLTIFYAGSVNVFNNVSAE--------------KAQELMFLASRGSLPTA
Zea_mays_GRMZM2G126507_T01          QLTIFYAGSVNVFDNVSAE--------------KAQELMLLASRGSLPSS
O.sativa_LOC_Os08g33160.1           QMTIFYDGLVNVFDNIPVE--------------KAQELMLLASRASIPSP
                                    *:*:** * * ::::::. :               :*: :*::*.  .

G.max_Glyma05g34960.1               SNMAQPNVQ------------VPI-SKLGAG-AGVPVS------------
G.max_Glyma08g04770.1               SNMAQPKVQ------------VLV-SKLAVG-DGVPVS------------
M.truncatula_TC119816               SNRAQPEVQ------------ASS-SKFASGDDGLPIS------------
P.trichocarpa_656962                SNLAQPIVQ------------VQASSSKPAAADLSPVN------------
P.trichocarpa_scaff_VIII.1248       SNLAQPIVQ------------VQASSSKPAAADLSPVN------------
P.trichocarpa_scaff_X.1014          SKSAQPKVQ------------VQAFSSKPAAADVSPVN------------
P.trichocarpa_TC117359              SKSAQPKVQ------------VQAFSSKPAAADVSPVN------------
S.lycopersicum_TC193363             PNVVQPRFQ------------LQASASKPAAADGVCVN------------
Le_UNK                              PNVVQPRFQ------------LQASASKPAAADGVCVN------------
A.thaliana_AT3G17860.1              QVFSPPQTH------------QQVVHHTRASVDS----------------
```

Fig. 2

```
B.napus_TC100484                    QAFTPPQTH-----------QQVVHHARASVDS----------
P.trichocarpa_scaff_XV.429          QNKPISTPQ----------AQAQAPIPGPPVGDI----------
S.lycopersicum_TC207683             PN--ATSTL----------SPVQAPIPKSSAIDS----------
G.max_Glyma09g30460.1               EKMELPTVK-----------LQPAISIPSKDD----------
M.truncatula_TC118189               QEISIPSKKDNFIISQPYPSPLPSPIPMTSHASTQPRGGSSSNNEVTIIR
Hordeum_vulgare_subsp_vulgare_      PSTVTRSPDA--------TFFTPAKLAAPELSPAKQMLPQ----------
T.aestivum_TC290287                 PTTVTRSPDA--------TFFTPAKLAAPEASPAKQMLAQ----------
Zea_mays_GRMZM2G126507_T01          -APVARKPEA--------PILAPAKVTAPEVLHATQMLFQ----------
O.sativa_LOC_Os08g33160.1           -PSAARKSDS--------PISAAAKLTVPEALPARQIVVQ----------

G.max_Glyma05g34960.1               ---QPAN-TSPGS-----GLSSPLSVSSHTGVQSGSGLTSTDEFLAAKTT
G.max_Glyma08g04770.1               ---QPAN-TSPGS-----GLSSPLSVSSHTGIQSGSGSTSTD-----KTT
M.truncatula_TC119816               ---PPVY-IPPCS-----GISSPLSVSSHTGPQPGSGSSSSDEFMAAKTS
P.trichocarpa_656962                ---QPIMSTPPCS-----RLSSP----SHTGAQSGSGSTSTEEIMATKTT
P.trichocarpa_scaff_VIII.1248       ---QPIMSTPPCS-----RLSSP----SHTGAQSGSGSTSTEEIMATKTT
P.trichocarpa_scaff_X.1014          ---QPIMSTPPCS-----SLSSP----SHTGAQSGSGSTSTEEIMATKTT
P.trichocarpa_TC117359              ---QPIMSTPPCS-----SLSSP----SHTGAQSGSGSTSTEEIMATKTT
S.lycopersicum_TC193363             ---QTPN-MLPAS-----GLSSPMSVSSHPIGQSDGSSGNKDDMKMSKTA
Le_UNK                              ---QTPN-MLPAS-----GLSSTMSVSSHPIGQSDGSSGNKDDMKMSKTA
A.thaliana_AT3G17860.1              -------SAMP---------PSFMPTISYLSPE---AGSSTNGLGATKAT
B.napus_TC100484                    -------SAVP---------PSFMPTASYLSHE---GGSSTYGLGAVKAT
P.trichocarpa_scaff_XV.429          FVGNKINTTAPCS-----GMPSPISVT----------SSSTNDLAIVKPV
S.lycopersicum_TC207683             FVVNQCRNTTPT-------LASPISITSHGGAQAAKVSTTTNGVTIIKSI
G.max_Glyma09g30460.1               --GFMISQSYP-----PSTFPTPLPLTSHVNSQPGGGSSSNKEISIIRQV
M.truncatula_TC118189               TLGHSIAKSSHNDLSHLSSLPSPLPMTSHASSYPKGGSSSNNEVTIIRSL
Hordeum_vulgare_subsp_vulgare_      ---MPQRVSPPLS-----GISKPISMVSQAACVPKSASSSNIDSTAPKSS
T.aestivum_TC290287                 ---IPQRVSPPLP-----AISKPMSIMSQAACLPKSTSSSNTDSAVPKSS
Zea_mays_GRMZM2G126507_T01          ---KPQHVSPPSS-----AISKPIPGILQAASLPRSASSSNLDSPFPKSS
O.sativa_LOC_Os08g33160.1           ---KPEASVPLVS-----GVSNPITIVSQAVTLPKSFSSSN-DSAGPKSG
                                                      ..                  ..      :

G.max_Glyma05g34960.1               G--VPNTPICNVEPPK----------VVSATTMLTS----------
G.max_Glyma08g04770.1               G--VPTTPVCNVEPPK----------IVSATTMLTS----------
M.truncatula_TC119816               RGPTPTTSACKVVTPK----------VVNATTMIPS----------
P.trichocarpa_656962                G--ALTTHVTKPEHTKT----ANVVGSVTTTTMIPS----------
P.trichocarpa_scaff_VIII.1248       G--ALTTHVTKPEHTKT----ANVVGSVTTTTMIPS----------
P.trichocarpa_scaff_X.1014          G--PVTIPVIKPDHPKT----GNVVGSVATTTMIPSGMISLFSLESGVLI
P.trichocarpa_TC117359              G--PVTIPVIKPDHPKT----GNVVGSVATTTMIPS----------
S.lycopersicum_TC193363             N--ISVTPHVKLDTSKI----VTSLGPVGATTIMTA----------
Le_UNK                              N--ISVTPHVKLDTSKI----VTSLGPVGATTIMTA----------
A.thaliana_AT3G17860.1              R--GLTSTYHNNQANGSNINCPVPVSCSTNVMAPTV----------
B.napus_TC100484                    T--GFTSTYHNNQT----------VKPQTV----------
P.trichocarpa_scaff_XV.429          V--NLAPSVKHIEPT----KPASSVGPTSATLVPAV----------
S.lycopersicum_TC207683             G--VLPSPSLKAKPS----KVTSSVGSFPASLVPS----------
G.max_Glyma09g30460.1               G--PSTAPTNHLESP----IIG-SIGSASKEKAQPV----------
M.truncatula_TC118189               G--PSNAPTNHLESP----IVATSVGLTPTNVIQPV----------
Hordeum_vulgare_subsp_vulgare_      -GQLVVPPTSQPSSSTHP----VTLASTTAASIMPR----------
T.aestivum_TC290287                 -GQLVVPPTSQPSSS--------TLASTTAASIMPR----------
Zea_mays_GRMZM2G126507_T01          -VPFPVSPVSQAPRAQP-----ATIAATTAAAIMPR----------
O.sativa_LOC_Os08g33160.1           GLPLAVTPLSQASPSQP-----IPVATTNASAIMPR----------
                                              :                     .

G.max_Glyma05g34960.1               --------------------------AVPQARKASLARFLEKRKERVMS
G.max_Glyma08g04770.1               --------------------------AVPQARKASLARFLEKRKERLVW
M.truncatula_TC119816               --------------------------AIPQARKASLARFLEKRKERVMS
P.trichocarpa_656962                --------------------------VPQARKASLARFLEKRKERVMN
P.trichocarpa_scaff_VIII.1248       --------------------------VPQARKASLARFLEKRKESDVM
P.trichocarpa_scaff_X.1014          YIASMKMFFFFFLSSSAMYNQCPISVTAVPQARKASLARFLEKRKERATN
P.trichocarpa_TC117359              --------------------------VPQARKASLARFLEKRKERATN
S.lycopersicum_TC193363             --------------------------AVPQARKASLARFLEKRKERVMN
Le_UNK                              --------------------------AVPQARKASLARFLEKRKERVMN
A.thaliana_AT3G17860.1              --------------------------ALPLARKASLARFLEKRKERVTS
B.napus_TC100484                    --------------------------ALPQARKASLARFLEKRKERVTS
P.trichocarpa_scaff_XV.429          --------------------------AVPQARKASLARFLEKRKERVMQ
S.lycopersicum_TC207683             --------------------------AVPQARKASLARFLEKRKERVIS
G.max_Glyma09g30460.1               --------------------------CLPQARKASLARFLEKRKGRMMR
M.truncatula_TC118189               --------------------------GLPQARKASLTRFLEKRKERAMS
```

Fig. 2

```
Hordeum_vulgare_subsp_vulgare_          ----------------------------AVPQARKASLARFLEKRKERVTT
T.aestivum_TC290287                     ----------------------------AVPQARKASLARFLEKRKERVTT
Zea_mays_GRMZM2G126507_T01              ----------------------------AVPQARKASLARFLEKRKERVTT
O.sativa_LOC_Os08g33160.1               ----------------------------AVPQARKASLARFLEKRKERVSS
                                                                    :* ******:*******

G.max_Glyma05g34960.1                   AAPYN-LNKKSEECATAEYAGVNFSATN----TVLAKQG-----------
G.max_Glyma08g04770.1                   MQYFS-GWVLSDTIYISPKK-EGVVVII----HWLTT-------------
M.truncatula_TC119816                   TAPYN-LNKKSEDAQMPNSMGANISATTGTANMLVANQG-----------
P.trichocarpa_656962                    AAPYN-LNKKSPHFTNPEPY-------------------------------
P.trichocarpa_scaff_VIII.1248           DATIS-STKSSGS--------------------------------------
P.trichocarpa_scaff_X.1014              AEPYN-LSKKSPDFANPEPY-------------------------------
P.trichocarpa_TC117359                  AEPYN-LSKKSPDFANPEPY-------------------------------
S.lycopersicum_TC193363                 LAPYG-LSKKSPECSTPESNGVGFSATSTPLLAGKET-------------
Le_UNK                                  LAPYG-LSKKSPECSTPESNGVGFSATSTPLLAGKET-------------
A.thaliana_AT3G17860.1                  VSPYCLDKKSSTDCRRSMSECISSSLSSAT--------------------
B.napus_TC100484                        ASPYCLDKKSPTDCRTPISECISSSFSSAT--------------------
P.trichocarpa_scaff_XV.429              TSPYN-GSKKSPEGGAHRFDGMSLSMSTSSSFPLPASN------------
S.lycopersicum_TC207683                 ASPYPLNSKQSPECSTPELGSRSLSMNSSGSCPPHIISLVK---------
G.max_Glyma09g30460.1                   TSPYLYMSKKSPECSSSGSDSVSFSLNFSGSCSLPATN------------
M.truncatula_TC118189                   TSPY-YMSKISPECS-TGSDNASFSIDFSGSSTQPPTNLPLRRTCMEVIR
Hordeum_vulgare_subsp_vulgare_          TAPYPSAKSPMESSDTVGSANDNNSKSSSCTEIAFSSN-HEESLRLGGRP
T.aestivum_TC290287                     TAPYPSAKSPMESSDTVGSANDNNSKSSSCTEIAFSSN-HEESLRLG-RP
Zea_mays_GRMZM2G126507_T01              AAPYPSAKSPMESSDTFGSGSANDK--SSCTDIALSSN-HEESLCLG-QP
O.sativa_LOC_Os08g33160.1               VAPYPSSKSPLESSDTIGSPSTPSK--SSCTDITPSTNNCEDSLCLG-QP

G.max_Glyma05g34960.1                   ------------------
G.max_Glyma08g04770.1                   ------------------
M.truncatula_TC119816                   ------------------
P.trichocarpa_656962                    ------------------
P.trichocarpa_scaff_VIII.1248           ------------------
P.trichocarpa_scaff_X.1014              ------------------
P.trichocarpa_TC117359                  ------------------
S.lycopersicum_TC193363                 ------------------
Le_UNK                                  ------------------
A.thaliana_AT3G17860.1                  ------------------
B.napus_TC100484                        ------------------
P.trichocarpa_scaff_XV.429              ------------------
S.lycopersicum_TC207683                 ------------------
G.max_Glyma09g30460.1                   ------------------
M.truncatula_TC118189                   ------------------
Hordeum_vulgare_subsp_vulgare_          RNISFS-GESPSTKLHI
T.aestivum_TC290287                     RNISFS-GESPSTKLHI
Zea_mays_GRMZM2G126507_T01              RNISFI-QESPSTKLQI
O.sativa_LOC_Os08g33160.1               RNISFSSQEPPSTKLQI
```

Fig. 2

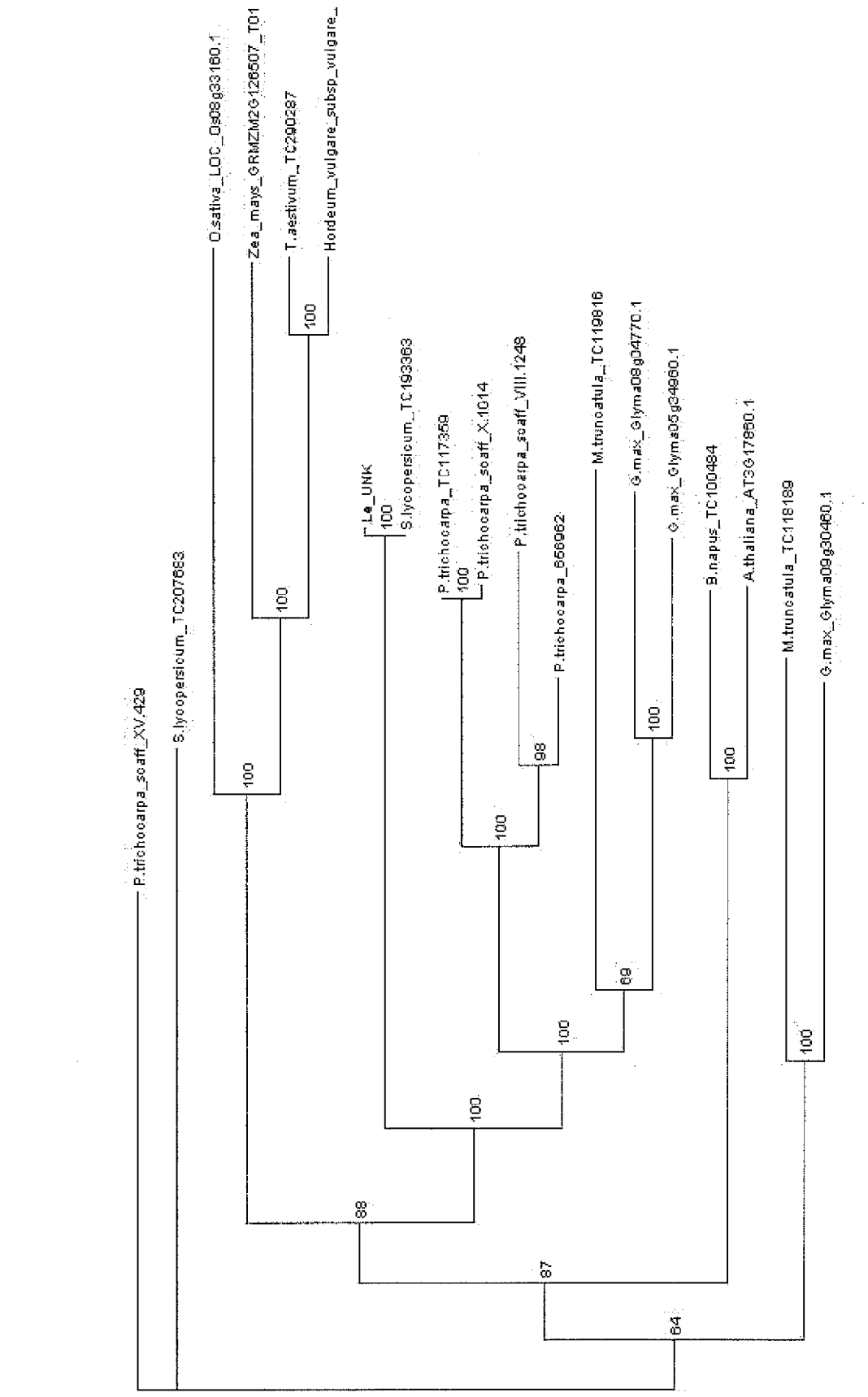

P.trichocarpa_scaff_XV.429
S.lycopersicum_TC207683
O.sativa_LOC_Os08g33160.1
Zea_mays_GRMZM2G126507_T01
T.aestivum_TC290287
Hordeum_vulgare_subsp_vulgare_
La_UNK
S.lycopersicum_TC193363
P.trichocarpa_TC117359
P.trichocarpa_scaff_X.1014
P.trichocarpa_scaff_VIII.1248
P.trichocarpa_856962
M.truncatula_TC119816
G.max_Glyma08g04770.1
G.max_Glyma05g34960.1
B.napus_TC100484
A.thaliana_AT3G17950.1
M.truncatula_TC118189
G.max_Glyma09g30460.1
100
100
100
100
100
100
100
98
100
100
88
69
100
100
87
64
0.01

Fig. 3

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1. G.max_Glyma05g34960.1		88.6	74.3	77.1	77.1	53.1	74.3	71.4	71.4	71.4	65.7	73.5	67.6	67.6	71.4	60.0	60.0	65.7	54.3
2. G.max_Glyma08g04770.1	91.4		74.3	80.0	80.0	55.1	77.1	77.1	77.1	74.3	68.6	80.0	71.4	71.4	74.3	62.9	62.9	68.6	57.1
3. M.truncatula_TC119816	85.7	91.4		62.9	62.9	42.9	60.0	71.4	71.4	57.1	54.3	65.7	62.9	57.1	62.9	48.6	48.6	48.6	48.6
4. P.trichocarpa_656962	88.6	88.6	85.7		100.0	67.3	94.3	80.0	80.0	88.6	82.9	85.7	80.0	77.1	77.1	62.9	62.9	62.9	51.4
5. P.trichocarpa_scaff_VIII.1248	88.6	88.6	85.7	100.0		67.3	94.3	80.0	80.0	88.6	82.9	85.7	80.0	77.1	77.1	62.9	62.9	62.9	51.4
6. P.trichocarpa_scaff_X.1014	59.2	59.2	57.1	69.4	69.4		71.4	55.1	55.1	61.2	57.1	59.2	55.1	53.1	55.1	42.9	42.9	42.9	36.7
7. P.trichocarpa_TC117359	82.9	82.9	80.0	97.1	97.1	71.4		77.1	77.1	85.7	80.0	82.9	77.1	74.3	77.1	60.0	60.0	60.0	51.4
8. S.lycopersicum_TC193363	82.9	85.7	82.9	91.4	91.4	63.3	88.6		100.0	68.6	62.9	71.4	68.6	62.9	71.4	60.0	60.0	57.1	51.4
9. Le_UNK	82.9	85.7	82.9	91.4	91.4	63.3	88.6	100.0		68.6	62.9	71.4	68.6	62.9	71.4	60.0	60.0	57.1	51.4
10. A.thaliana_AT3G17860.1	82.9	82.9	80.0	94.3	94.3	67.3	94.3	85.7	85.7		91.4	85.7	82.9	85.7	85.7	54.3	54.3	60.0	48.6
11. B.napus_TC100484	82.9	82.9	80.0	94.3	94.3	67.3	94.3	85.7	85.7	100.0		80.0	77.1	80.0	80.0	48.6	48.6	54.3	42.9
12. P.trichocarpa_scaff_XV.429	88.2	91.4	91.4	97.1	97.1	69.4	97.1	88.6	88.6	94.3	94.3		85.3	79.4	80.0	60.0	60.0	65.7	48.6
13. S.lycopersicum_TC207683	82.4	82.9	88.6	88.6	88.6	63.3	88.6	85.7	85.7	91.4	91.4	88.2		76.5	80.0	62.9	62.9	68.6	54.3
14. G.max_Glyma09g30460.1	85.3	88.6	85.7	88.6	88.6	61.2	85.7	82.9	82.9	88.6	88.6	88.2	88.2		82.9	51.4	51.4	57.1	50.0
15. M.truncatula_TC118189	80.0	80.0	82.9	88.6	88.6	61.2	85.7	85.7	85.7	91.4	91.4	88.6	91.4	88.6		48.6	48.6	54.3	48.6
16. Hordeum_vulgare_subsp_vulgare_	82.9	82.9	71.4	77.1	77.1	55.1	77.1	68.6	68.6	74.3	74.3	77.1	74.3	74.3	71.4		100.0	91.4	68.6
17. T.aestivum_TC290287	82.9	82.9	71.4	77.1	77.1	55.1	77.1	68.6	68.6	74.3	74.3	77.1	74.3	74.3	71.4	100.0		91.4	68.6
18. Zea_mays_GRMZM2G126507_T01	82.9	82.9	71.4	77.1	77.1	55.1	77.1	71.4	71.4	74.3	74.3	77.1	74.3	74.3	71.4	100.0	100.0		74.3
19. O.sativa_LOC_Os08g33160.1	68.6	68.6	68.6	68.6	68.6	46.9	65.7	71.4	71.4	62.9	62.9	65.7	65.7	68.6	62.9	80.0	80.0	80.0	

Fig. 4

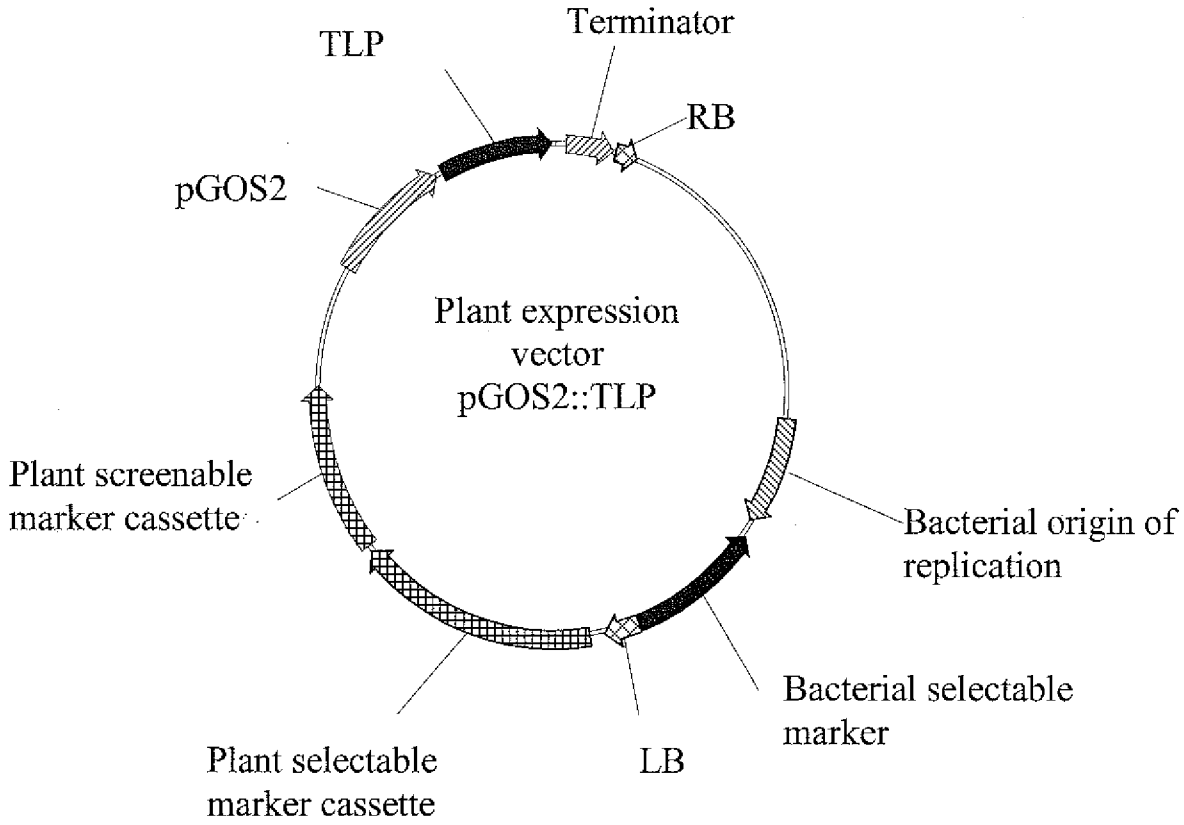

TLP
Terminator
RB
pGOS2
Plant expression
vector
pGOS2::TLP
Plant screenable
marker cassette
Bacterial origin of
replication
Bacterial selectable
marker
Plant selectable
marker cassette
LB

Fig. 5

MATINGLTSPHPFLSLQKSNSKPFLSLSKPMKSQFLFKGLKQQNQQLRDWAVVGSVSNEADEIRVQSS

DVMDQQNGVVIGLEREPKLGGGDMGIVNQVVGGFGNEGRLSFEGGGGFSSASGVGGSESKEEDVEKLI

DRTINATIVLAAGSFAITKLLTIDHDYWHGWTLFEILRYAPQHNWVAYEEALKRNPVLAKMVISGVVY
----------2----------
----------5------
----------7------

SVGDWIAQCYEGKPLFEIDRARMLRSGLVGFTLHGSLSHYYYQFCEALFPFEDWWVVPVKVAFDQTVW
----------1----------
----------3---
----------4----------
----------8----------

SAIWNSIYFT**VLGILRLESPLAIFSELKATFLPMLTAGWKLWPFAHLITYGVIPVEQRLLWVDCVELV**
----------6----------
-----9----------

WVTILSTYSNEKSESRVSEVIVEAEAQPPSVSPPQE

Fig. 6

```
CLUSTAL 2.0.11 multiple sequence alignment

L.esculentum_L450_PM22                --MATINGLTSPHPFLSLQKSNSKP-FLSLSKPMKSQFLFKGLKQQNQQ-
S.tuberosum_TC167881#1                --MATINGLTSPQPFLSLQKSNSKP-FLSLSKPMKSQFLFKGLKQQNQQ-
N.benthamiana_TC11792#1               --MATINGLTSPQPFLSLPKPKRKS-SHSHFKPLKPQILSKGLKQQNQQ-
N.tabacum_TC57360#1                   --MSTINGLTSPQPFLSLPKPKPKP-SYSHFKPLKPRILSKGLKQQNQQN
I.nil_TC25#1                          --MATINGLPTPQTLRSLSKPNPKTGSPSHFKPTKLQLFPRPVSKQFAA-
A.lyrata_474411#1                     MAAASLHTSISPRSFLPLSKPSLKP--------HRSQIFLRNKQR-----
A.thaliana_AT1G52870.2#1              MAAASLHTSISPRSFLPLSKPSLKP--------HRSQILLRNKQRNC---
B.napus_BN06MC04723_42271568@4        MAASSLNTSISPRSFLPLSKPSPKP--------HRSQVFLRNKQRSC---
                                         ::::   :*:.: .* *.. *.          : :.: :   .:

L.esculentum_L450_PM22                --LRDWAVVGSVSNEADEIRVQSSDVMDQQNGVVIGLEREPKLGGGDMG-
S.tuberosum_TC167881#1                --LRDWAVVGSVTNEADVIPVQSTDVTDQQNGVVIGLEREPELGGGEMD-
N.benthamiana_TC11792#1               --LRDWAVVGSVDNETDLIPVQSDDTTDQQSGVVIGLEREPE--GGDMD-
N.tabacum_TC57360#1                   QQLRDWAVVGSVANETDLIPVQSNDTTDQQRGVVIGLEREPE--GGDMD-
I.nil_TC25#1                          ----------AAAKELDVIPVQSSDSTDQQGGVVDATEREAE-GGGDIDS
A.lyrata_474411#1                     -------TCALIRDEIDVIPVQSRDRTDHEEGSVVVMSTETE-----RD-
A.thaliana_AT1G52870.2#1              ------VSCALIRDEIDLIPVQSRDRTDHEEGSVVVMSTETA-----VD-
B.napus_BN06MC04723_42271568@4        ------VSCALINDETDVIPVKS-----HDDGSVVVMSNETE-----RD-
                                                .* * * *:*     :: * *   . *.        .

L.esculentum_L450_PM22                IVNQ--VVGGFG---NEGRLSFE-----GGG-GFSSASGVGGSESKEEDV
S.tuberosum_TC167881#1                IVNQ--VVGGFG---NEGRLSFE-----GAG-GFSSASGVGGS-----DV
N.benthamiana_TC11792#1               LVKQ--VVGGFG---NEGRLSFE-----GAARGFSSSSSGSNGKLEQEDM
N.tabacum_TC57360#1                   MVNQ--VVGGFG---NEGRLSFE-----GAAGGFSSSSSGSNGKSEQEDM
```

Fig. 7A

```
I.nil_TC25#1                          IVNQVVVVGGFG---NEGRLSFEGPTGFGSSSSSSSAASSSGGAGEELGM
A.lyrata_474411#1                     -VNES-VVVGFSAATSEGQLSLEG----FPSS---GGDLGDEKRGENEER
A.thaliana_AT1G52870.2#1              -GNES-VVVGFSAATSEGQLSLEG----FPSSSSSGADLGDEKRRENEEM
B.napus_BN06MC04723_42271568@4        -VNEP-VVVG--------QLSFEG----FPSSSP-AADMGDDKSRESEEM
                                        ::   ** *         :**:*        .     ..    .

L.esculentum_L450_PM22                EKLIDRTINATIVLAAGSFAITKLLTIDHDYWHGWTLFEILRYAPQHNWV
S.tuberosum_TC167881#1                EKLIDRAINATIVLAAGSFAITKLLTIDHDYWHGWTLFEILRYAPQHNWV
N.benthamiana_TC11792#1               NRVIDRAINATIVLAAGSFAITKLLTIDHDYWHGWTVYEIFRYAPQHNWI
N.tabacum_TC57360#1                   DRVIDRAINATIVLAAGSFAITKLLTIDHDYWHGWTIYEILRYAPQHNWI
I.nil_TC25#1                          EKVVDRAINATIVLAAGTFAITKLLTIDHNYWHGWTLFEILRYAPQHNWS
A.lyrata_474411#1                     EKMIDRTINATIVLAAGSYAITKLLTIDHDYWHVWTLLEILRYAPQHNWI
A.thaliana_AT1G52870.2#1              EKMIDRTINATIVLAAGSYAITKLLTIDHDYWHGWTLFEILRYAPQHNWI
B.napus_BN06MC04723_42271568@4        DKMIDRSINATIVLAAGTYAITKLLTIDHDYWHGWTLFEILRYAPQHNWL
                                      ::::**:**********::**********:*** **: **:********

L.esculentum_L450_PM22                AYEEALKRNPVLAKMVISGVVYSVGDWIAQCYEGKPLFEIDRARMLRSGL
S.tuberosum_TC167881#1                AYEEALKRNPVLAKMVISGVVYSVGDWIAQCYEGKPLFEIDRARMLRSGL
N.benthamiana_TC11792#1               AYEEALKRNPVLAKMVISGVVYSIGDWIAQCYEGKPLFDFDRARMLRSGF
N.tabacum_TC57360#1                   AYEEALKRNPVLAKMVISGVVYSLGDWIAQCYEGKPLFDFDRARMLRSGL
I.nil_TC25#1                          AYEEALKENPVLAKMVISGVVYSLGDWIAQCYEGKPLFEFDRARMFRSGL
A.lyrata_474411#1                     AYEEALKQNPVLAKMVISGVVYSVGDWIAQCYEGKPLFEIDRARTLRSGL
A.thaliana_AT1G52870.2#1              AYEEALKQNPVLAKMVISGVVYSVGDWIAQCYEGKPLFEIDRARTLRSGL
B.napus_BN06MC04723_42271568@4        AYEEALKRNPVLAKMVISGVVYSVGDWIAQCYEGKPLFEIDRARTLRSGL
                                      *******.***************:***************::**** :***:

L.esculentum_L450_PM22                VGFTLHGSLSHYYYQFCEALFPFEDWWVVPVKVAFDQTVWSAIWNSIYFT
```

Fig. 7A

```
S.tuberosum_TC167881#1                VGFTLHGSLSHYYYQFCEALFPFEDWWVVPVKVAFDQTVWSAIWNSIYFT
N.benthamiana_TC11792#1               VGFTLHGSLSHYYYQFCEALFPFQDWWVVPVKVAFDQTVWSAIWNSIYYT
N.tabacum_TC57360#1                   VGFTLHGSLSHYYYQICEALFPFQDWWVVPVKVAFDQTVWSAIWNSIYYT
I.nil_TC25#1                          TGFTLHGSLSHYYYHFCEALFPFDDWWVVPAKVAFDQTVWSAIWNSIYFT
A.lyrata_474411#1                     VGFTLHGSLSHFYYQFCEELFPFQDWWVVPVKVVFDQTVWSAIWNSIYFT
A.thaliana_AT1G52870.2#1              VGFTLHGSLSHFYYQFCEELFPFQDWWVVPVKVAFDQTVWSAIWNSIYFT
B.napus_BN06MC04723_42271568@4        VGFTLHGSLSHFYYQFCEELFPFQDWWVVPAKVAFDQTVWSAIWNSIYFT
                                      .**********:**:** ****:******.**.**************:*

L.esculentum_L450_PM22                VLGILRLESPLAIFSELKATFLPMLTAGWKLWPFAHLITYGVIPVEQRLL
S.tuberosum_TC167881#1                VLGILRFESPLAIFSELKATFLPMLTAGWKLWPFAHLITYGVIPVEQRLL
N.benthamiana_TC11792#1               VLGVLRFESPLGIFSELKATFWPMLTAGWKLWPFAHLITYGVVPVEQRLL
N.tabacum_TC57360#1                   VLGVLRFESPLAIFSELKATFWPMLTAGWKLWPFAHLITYGVVPVEQRLL
I.nil_TC25#1                          VLGFLRLESPTSIFSELKATFWPMLTAGWKLWPFAHLITYGVIPVEQRLL
A.lyrata_474411#1                     VLGFLRFESPLSIFKELKATFLPMLTAGWKLWPFAHLITYGLVPVEQRLL
A.thaliana_AT1G52870.2#1              VLGFLRFESPISIFKELKATFLPMLTAGWKLWPFAHLITYGLVPVEQRLL
B.napus_BN06MC04723_42271568@4        VLGFLRLESPLSIFKELKATFLPMLTAGWKLWPFAHLITYGLVPVEQRLL
                                      ***.**:*** .**.****** ********************::*******

L.esculentum_L450_PM22                WVDCVELVWVTILSTYSNEKSESRVSEVIVEAEAQ-PPSVSPPQE
S.tuberosum_TC167881#1                WVDCVELIWVTILSTYSNEKSESRVSEVIVEAEAQ-PPSISPPQE
N.benthamiana_TC11792#1               WVDCVELIWVTILSTYSNEKSEASL--------------------
N.tabacum_TC57360#1                   WVDCVELIWVTILSTYSNEKSEARVSEASVEAEMQ-PPSIGPPQE
I.nil_TC25#1                          WVDCVELIWVTILSTYSNEKSEARISEESVEATSN-PPSVGPSQK
A.lyrata_474411#1                     WVDCVELIWVTILSTYSNEKSEARISESVIETSSSSTTSIDPSKE
A.thaliana_AT1G52870.2#1              WVDCVELIWVTILSTYSNEKSEARISESVIETSSSSTTTIDPSKE
B.napus_BN06MC04723_42271568@4        WVDCVELIWVTILSTYSNEKSEARISESVIENPSSSTTPIDPSKE
                                      *******:***************: :
```

Fig. 7A

```
CLUSTAL 2.0.11 multiple sequence alignment

Z.mays_ZM07MC32543_BFb0296A02@          ---MAAAAVATVA-------------------------TPSP-ASRRLFP
Zea_mays_GRMZM2G011269_T01#1            ---MAAAAVATVA-------------------------SPSPAASRRMLP
P.virgatum_TC39638#1                    --MAAAAAVATVS-------------------------SPSP-ASRRLLP
S.bicolor_Sb01g015680.1#1               -MAAASASVATVA-------------------------SPSSPASRRLLL
O.sativa_LOC_Os03g38730.1#1             -MAAMAAAVASSP-------------------------SPSP-ASTRLLR
L.esculentum_L450_PM22                  -----MATINGLT-------------------------SPHPFLSLQKSN
S.tuberosum_TC167881#1                  -----MATINGLT-------------------------SPQPFLSLQKSN
N.benthamiana_TC11792#1                 -----MATINGLT-------------------------SPQPFLSLPKPK
N.tabacum_TC57360#1                     -----MSTINGLT-------------------------SPQPFLSLPKPK
I.nil_TC25#1                            -----MATINGLP-------------------------TPQTLRSLSKPN
A.lyrata_474411#1                       ---MAAASLHTSI-------------------------SPRSFLPLSKPS
A.thaliana_AT1G52870.2#1                ---MAAASLHTSI-------------------------SPRSFLPLSKPS
B.napus_BN06MC04723_42271568@4          ---MAASSLNTSI-------------------------SPRSFLPLSKPS
M.domestica_TC51384#1                   -----MATLNTI--------------------------AAHNLPSPSTSS
V.vinifera_GSVIVT00025998001#1          -----MATLNTI--------------------------TAHRLFPLPKSI
A.lyrata_490222#1                       -----MAALCC---------------------------PTTIIVSGKLSS
A.thaliana_AT4G03410.1#1                -----MAALCCC--------------------------PTTIIVSGKVSS
G.max_Glyma13g37540.1#1                 -----MATVSS---------------------------PCTVFPILTSTL
G.max_TC302765#1                        -----MATVSS---------------------------PCTVFPILTSTL
G.max_Glyma12g32920.1#1                 -----MATVSS---------------------------LCAVFPILTPTL
H.annuus_TC45682#1                      -----MATVNNHVVF-----------------------FSGHKISPPTNPI
H.paradoxus_TA2132_73304#1              -----MMSY-----F-----------------------FSGHKVSPPTNPI
Triphysaria_sp_TC3527#1                 ---------------------------------------------------
P.patens_TC32628#1                      MAGVGPACVRSVASQGFGAPCVGLPQRNRGVRGLSGCGLPERVSSGKFCS
```

Fig. 7B

```
Z.mays_ZM07MC32543_BFb0296A02@      ----------------SAGPSLLRLP-RPTKRLR----------RALRF
Zea_mays_GRMZM2G011269_T01#1        ----------------SAAPSFLRLP-RPTSRLR----------RATQV
P.virgatum_TC39638#1                SS--------------SSAVPSLLRLP-RPGRRLR----------RALRV
S.bicolor_Sb01g015680.1#1           PLP-------------SSAAPSLLRLPSRPARRLRRA--------RALRV
O.sativa_LOC_Os03g38730.1#1         GHHP----------PPSCSAPSLLRLT-RSSRRLR------------LRA
L.esculentum_L450_PM22              SKP-----------FLSLSKPMKSQFLFKGLKQQNQQ---LRDWAVVGSV
S.tuberosum_TC167881#1              SKP-----------FLSLSKPMKSQFLFKGLKQQNQQ---LRDWAVVGSV
N.benthamiana_TC11792#1             RKS-----------SHSHFKPLKPQILSKGLKQQNQQ---LRDWAVVGSV
N.tabacum_TC57360#1                 PKP-----------SYSHFKPLKPRILSKGLKQQNQQNQQLRDWAVVGSV
I.nil_TC25#1                        PKTG----------SPSHFKPTKLQLFPRPVSKQFAA-----------AA
A.lyrata_474411#1                   LKP-----------------HRSQIFLRNKQR-----------TCALI
A.thaliana_AT1G52870.2#1            LKP-----------------HRSQILLRNKQRNC---------VSCALI
B.napus_BN06MC04723_42271568@4      PKP-----------------HRSQVFLRNKQRSC---------VSCALI
M.domestica_TC51384#1               FAT----------------QSHKPNSLSPFFCTQTSS---KPHPPRLPAF
V.vinifera_GSVIVT00025998001#1      TES----------------NNSYTVCISTSSSQLFLN---KSNPSLNSTF
A.lyrata_490222#1                   RIK----------------TACQLPRINPFKNRKR--------------L
A.thaliana_AT4G03410.1#1            RIK----------------TAGPLPRINPFKNQKR--------------L
G.max_Glyma13g37540.1#1             SKP----------------HLAPLSCRTQKRSRVS--------------I
G.max_TC302765#1                    SKP----------------HLAPLSCRTQKRSRVS--------------I
G.max_Glyma12g32920.1#1             SKP----------------RLAPLSCRIPKRRRIS--------------I
H.annuus_TC45682#1                  SYS----------------TTFSISPPTKLFSQSR---------------
H.paradoxus_TA2132_73304#1          SYS----------------TTFSISPPTKLFSQSK---------------
Triphysaria_sp_TC3527#1             --------------------------------------------------
P.patens_TC32628#1                  SSGKGGFFAGKKDGFSIRCSSLPGDGVAAGTPKTVIKSSCSKILRWSEED


Z.mays_ZM07MC32543_BFb0296A02@      AAAGDEADVLPGPGAEGEAVVP------GRLEEQPDEQ-LGGSQLDIGGL
```

Fig. 7B

```
Zea_mays_GRMZM2G011269_T01#1          AAAGGEADVLPGPGAEGEAAVP------GRLEEQRDEP-LAGSQLDIGGL
P.virgatum_TC39638#1                  VSAGGEADVLPGPGGEAEAAVP------GRLEEPRDEP-LGGSQLDIGGL
S.bicolor_Sb01g015680.1#1             AAAGDEADVLPGPGAEGEAAVPGT----GRLEEQRDEPPLGGSQLDIGGL
O.sativa_LOC_Os03g38730.1#1           AAAAEEADVLPGPGAEGEMATG------GRLEEQPEGP-IGGSQVDIGGL
L.esculentum_L450_PM22                SNEADEIRVQSSDVMDQQNGVV-------IGLEREPKLGGGDMG-IVNQ-
S.tuberosum_TC167881#1                TNEADVIPVQSTDVTDQQNGVV-------IGLEREPELGGGEMD-IVNQ-
N.benthamiana_TC11792#1               DNETDLIPVQSDDTTDQQSGVV-------IGLEREPE--GGDMD-LVKQ-
N.tabacum_TC57360#1                   ANETDLIPVQSNDTTDQQRGVV-------IGLEREPE--GGDMD-MVNQ-
I.nil_TC25#1                          AKELDVIPVQSSDSTDQQGGVV-------DATEREAE-GGGDIDSIVNQV
A.lyrata_474411#1                     RDEIDVIPVQSRDRTDHEEGSV-------VVMSTETE-----RD--VNES
A.thaliana_AT1G52870.2#1              RDEIDLIPVQSRDRTDHEEGSV-------VVMSTETA-----VD--GNES
B.napus_BN06MC04723_42271568@4        NDETDVIPVKS-----HDDGSV-------VVMSNETE-----RD--VNEP
M.domestica_TC51384#1                 PTRKKRSWIVRSITEDREVVPL-------KKK-DTGEQDSSLL---LNGS
V.vinifera_GSVIVT00025998001#1        SKKKRLNWVINAVAEDQDLAPA-------QTSGSKVHQDEDLPRSSLDGS
A.lyrata_490222#1                     ITERR-NLIVKSIIEDREAVDV-------KNENFKPEE------------
A.thaliana_AT4G03410.1#1              ITERR-NLIVKSIIEDREAIDV-------KNDNFKAEE------------
G.max_Glyma13g37540.1#1               ITSVAEDREIVPVSEDRGIRLN-------EVDGFQPSE------------
G.max_TC302765#1                      ITSVAEDREIVPVSEDRGIRLN-------EVDGFQPSE------------
G.max_Glyma12g32920.1#1               ITSVAEDREIVPVSEGRGTRLN-------EVEGFQPSE------------
H.annuus_TC45682#1                    -KNHKVYPVMKSIAEDKEVVAD-------EQNGSERES------------
H.paradoxus_TA2132_73304#1            -QNHKVYPVMKSIAEDKEVVAD-------EQNGSERES------------
Triphysaria_sp_TC3527#1               ----------------------------------------------------
P.patens_TC32628#1                    VVSSGGVPVVQALPDDAETHSGANGMLVELRETDRPQPIGSGTYGSGNQL


Z.mays_ZM07MC32543_BFb0296A02@        AFQG----------------DVGGG-FTGGG--AGSGASGGGG-GNKMLD
Zea_mays_GRMZM2G011269_T01#1          AFQG----------------DVGGG-FTGGG--AGSGASGGGGDGNKMLD
```

Fig. 7B

```
P.virgatum_TC39638#1                    AFQG----------------DVGGG-FAGGGGGVGSGASGGGG-GNKTLD
S.bicolor_Sb01g015680.1#1               AFQG----------------DVGGG-FTGGGSGSGSGASGGGD-GNKMLD
O.sativa_LOC_Os03g38730.1#1             AFQG----------------DMGGGGFAGGS----GGAGAGGGDGNKMLD
L.esculentum_L450_PM22                  -VVGGFG---NEGRLSFE-----GGG-GFSSASGVGGSESKEEDVEKLID
S.tuberosum_TC167881#1                  -VVGGFG---NEGRLSFE-----GAG-GFSSASGVGGS-----DVEKLID
N.benthamiana_TC11792#1                 -VVGGFG---NEGRLSFE-----GAARGFSSSSSGSNGKLEQEDMNRVID
N.tabacum_TC57360#1                     -VVGGFG---NEGRLSFE-----GAAGGFSSSSSGSNGKSEQEDMDRVID
I.nil_TC25#1                            VVVGGFG---NEGRLSFEGPTGFGSSSSSSSAASSSGGAGEELGMEKVVD
A.lyrata_474411#1                       -VVVGFSAATSEGQLSLEG----FPSS---GGDLGDEKRGENEEREKMID
A.thaliana_AT1G52870.2#1                -VVVGFSAATSEGQLSLEG----FPSSSSSGADLGDEKRRENEEMEKMID
B.napus_BN06MC04723_42271568@4          -VVVG--------QLSFEG----FPSSSP-AADMGDDKSRESEEMDKMID
M.domestica_TC51384#1                   ------------------------GEFQALSTP---SEEKGEGDDLDKLTS
V.vinifera_GSVIVT00025998001#1          ------------------------EDSEGLSSSSVSSQGKGDNYEFDRLRS
A.lyrata_490222#1                       -----------------------------------EQT----EDTDQLMS
A.thaliana_AT4G03410.1#1                -----------------------------------ELSEDKVEDTDRLMS
G.max_Glyma13g37540.1#1                 -----------------------------------PHTESDVVVPRLTSS
G.max_TC302765#1                        -----------------------------------PHTESDVVVPRLTSS
G.max_Glyma12g32920.1#1                 -----------------------------------PHTESEVVVSSLTSS
H.annuus_TC45682#1                      -----------------------------------GVVGEGSEGIDKFAS
H.paradoxus_TA2132_73304#1              -----------------------------------GVVGAGSEGIDKFAS
Triphysaria_sp_TC3527#1                 -------------------------------------------------MLL
P.patens_TC32628#1                      SMEMAVN-------------EAQTMTATMYRPKLLWFTKSEDEESRLLD

Z.mays_ZM07MC32543_BFb0296A02@          RGINTAIVLGASTYALTKLLTVDQDYWHGWTIFEILRYMPEHNWSAYEEA
Zea_mays_GRMZM2G011269_T01#1            RGINTAIVLAASTYALTKLLTVDQDYWHGWTIFEILRYMPEHNWSAYEEA
P.virgatum_TC39638#1                    RGINTAIVLGASTFALTKLLTVDQDYWHGWTIFEILRYMPEHNWSAYEEA
```

Fig. 7B

```
S.bicolor_Sb01g015680.1#1          RGINTAIVLGASTYALTKLLTVDQDYWHGWTIFEILRYMPEHNWSAYEEA
O.sativa_LOC_Os03g38730.1#1        RGINTAIVLGASTYALTKLLTVDHDYWHGWTIFEILRYMPEHNWSAYEEA
L.esculentum_L450_PM22             RTINATIVLAAGSFAITKLLTIDHDYWHGWTLFEILRYAPQHNWVAYEEA
S.tuberosum_TC167881#1             RAINATIVLAAGSFAITKLLTIDHDYWHGWTLFEILRYAPQHNWVAYEEA
N.benthamiana_TC11792#1            RAINATIVLAAGSFAITKLLTIDHDYWHGWTVYEIFRYAPQHNWIAYEEA
N.tabacum_TC57360#1                RAINATIVLAAGSFAITKLLTIDHDYWHGWTIYEILRYAPQHNWIAYEEA
I.nil_TC25#1                       RAINATIVLAAGTFAITKLLTIDHNYWHGWTLFEILRYAPQHNWSAYEEA
A.lyrata_474411#1                  RTINATIVLAAGSYAITKLLTIDHDYWHVWTLLEILRYAPQHNWIAYEEA
A.thaliana_AT1G52870.2#1           RTINATIVLAAGSYAITKLLTIDHDYWHGWTLFEILRYAPQHNWIAYEEA
B.napus_BN06MC04723_42271568@4     RSINATIVLAAGTYAITKLLTIDHDYWHGWTLFEILRYAPQHNWLAYEEA
M.domestica_TC51384#1              RAINALIVLGFGTFAVSKLLTIDHDYWHGWTLYEILRYVPEHNWIAYEQA
V.vinifera_GSVIVT00025998001#1      KTINATIVLAGGTLAITRLLTIDHDYWHGWTLYEVLRYAPEHNWVAYEEA
A.lyrata_490222#1                  RGINAAIVLAAGTVAVTKLLTIDHDYWQGWTLYEILRYAPEHNWVAYEQI
A.thaliana_AT4G03410.1#1           RGINAAIVLAAGTVAVTKLLTIDHDYWQGWTLYEILRYAPEHNWFAYEQI
G.max_Glyma13g37540.1#1            STVNAIIVLGFGTFAVTKLLTIDHDYWHGWTLFEIVRYIPEHNWIAYEQA
G.max_TC302765#1                   STVNAIIVLGFGTFAVTKLLTIDHDYWHGWTLFEIVRYIPEHNWIAYEQA
G.max_Glyma12g32920.1#1            SSVNAIIVLGFGTFAVTKLLTIDHDYWHGWTLYEIVRYIPEHNWIAYEQA
H.annuus_TC45682#1                 KAINASIVLGFGTLAVTRLLTIDHEYWHGWTLYEILRYAPEHNWIAYEEA
H.paradoxus_TA2132_73304#1         KAINASIVLGFGTLAVTRLLTIDHEYWHGWTLYEILRYAPEHNWIAYEEA
Triphysaria_sp_TC3527#1            LCWLLDLLPLPSCLPLIMIIGMDG------PFFEILRYAPQHNWSAYEEA
P.patens_TC32628#1                 RTFNALLAGAAISYAVTKAVTVDHDVWHGWTVFEVLKYAPLHNWHAYEEF
                                         :       .:   : :*       .. *:.:* * *** ***:

Z.mays_ZM07MC32543_BFb0296A02@     LKANPVLAKMMISGVVYSLGDWIAQCYEGKPIFDFDRARMFRSGLVGFTL
Zea_mays_GRMZM2G011269_T01#1       LKANPVLAKMMISGVVYSLGDWIAQCYEGKPIFDFDRARMFRSGLVGFTL
P.virgatum_TC39638#1               LKANPVLAKMMISGVVYSLGDWIAQCYEGKPIFDFDRARMFRSGLVGFTL
S.bicolor_Sb01g015680.1#1          LKANPVLAKMMISGVVYSLGDWIAQCYEGKPIFDFDRARMFRSGLVGFTL
```

Fig. 7B

```
O.sativa_LOC_Os03g38730.1#1          LKTNPVLAKMMISGVVYSLGDWIAQCYEGKPIFEFDRARMFRSGLVGFTL
L.esculentum_L450_PM22               LKRNPVLAKMVISGVVYSVGDWIAQCYEGKPLFEIDRARMLRSGLVGFTL
S.tuberosum_TC167881#1               LKRNPVLAKMVISGVVYSVGDWIAQCYEGKPLFEIDRARMLRSGLVGFTL
N.benthamiana_TC11792#1              LKRNPVLAKMVISGVVYSIGDWIAQCYEGKPLFDFDRARMLRSGFVGFTL
N.tabacum_TC57360#1                  LKRNPVLAKMVISGVVYSLGDWIAQCYEGKPLFDFDRARMLRSGLVGFTL
I.nil_TC25#1                         LKENPVLAKMVISGVVYSLGDWIAQCYEGKPLFEFDRARMFRSGLTGFTL
A.lyrata_474411#1                    LKQNPVLAKMVISGVVYSVGDWIAQCYEGKPLFEIDRARTLRSGLVGFTL
A.thaliana_AT1G52870.2#1             LKQNPVLAKMVISGVVYSVGDWIAQCYEGKPLFEIDRARTLRSGLVGFTL
B.napus_BN06MC04723_42271568@4       LKRNPVLAKMVISGVVYSVGDWIAQCYEGKPLFEIDRARTLRSGLVGFTL
M.domestica_TC51384#1                LKANPVFAKMVISGVVYTLGDWIAQCFEGKPLLEFDRKRMFRSGLVGFTF
V.vinifera_GSVIVT00025998001#1       LKTNPVLAKMAISGAVYSIGDWIAQCYEGKPLFEFDLTRMLRSGLVGFSL
A.lyrata_490222#1                    LKTNPVLAKMAISGIVYSLGDWIAQCYEGKPLFEFDRARVLRSGLVGFTL
A.thaliana_AT4G03410.1#1             LKTNPVLAKMAISGIVYSLGDWIAQCYEGKPLFEFDRTRVLRSGLVGFTL
G.max_Glyma13g37540.1#1              LKANPVLAKMAISGIVYSIGDWIAQCYEGKPLFEFDRTRVLRSGLVGFTL
G.max_TC302765#1                     LKANPVLAKMAISGIVYSIGDWIAQCYEGKPLFEFDRTRVLRSGLVGFTL
G.max_Glyma12g32920.1#1              LKANPVLAKMAISGIVYSIGDWIAQCYEGKPLFEFDRTRVLRSGLVGFTL
H.annuus_TC45682#1                   LKENPVLAKMVISGIVYSLGDWIAQCYEGKPLFEFDRIRLFRSGLVGFTL
H.paradoxus_TA2132_73304#1           LKENPVLAKMVISGIVYSLGDWIAQCYEGKPLFEFDRTRLFRSGLVGFTL
Triphysaria_sp_TC3527#1              LKANPVLAKMVISGVVYSVGDWIAQCYEGKPIFEYDRARMFRSGLVGFTL
P.patens_TC32628#1                   LSANPVLAKMMISGVVYSIGDWIGQCVEGKPVLEFSRVRLLRSGLVGFCL
                                     *. ***:*** *** **::****.** ****::: .  * :***:.** :

Z.mays_ZM07MC32543_BFb0296A02@       HGSLSHYYYHICEALFPFKDWWVVPAKVAFDQTIWSAIWNSIYFVVLGFL
Zea_mays_GRMZM2G011269_T01#1         HGSLSHYYYHICEALFPFKDWWVVPAKVAFDQTVWSAIWNSIYFVVLGFL
P.virgatum_TC39638#1                 HGSLSHYYYHICEALFPFKDWWVVPAKVAFDQTIWSAIWNSIYFVVLGFL
S.bicolor_Sb01g015680.1#1            HGSLSHYYYHICEALFPFKDWWVVPAKVAFDQTIWSAIWNSIYFVVLGFL
O.sativa_LOC_Os03g38730.1#1          HGSLSHYYYHFCEALFPFKDWWVVPAKVVFDQTAWSAIWNSIYFVVLGFL
```

Fig. 7B

```
L.esculentum_L450_PM22                  HGSLSHYYYQFCEALFPFEDWWVVPVKVAFDQTVWSAIWNSIYFTVLGIL
S.tuberosum_TC167881#1                  HGSLSHYYYQFCEALFPFEDWWVVPVKVAFDQTVWSAIWNSIYFTVLGIL
N.benthamiana_TC11792#1                 HGSLSHYYYQFCEALFPFQDWWVVPVKVAFDQTVWSAIWNSIYYTVLGVL
N.tabacum_TC57360#1                     HGSLSHYYYQICEALFPFQDWWVVPVKVAFDQTVWSAIWNSIYYTVLGVL
I.nil_TC25#1                            HGSLSHYYYHFCEALFPFDDWWVVPAKVAFDQTVWSAIWNSIYFTVLGFL
A.lyrata_474411#1                       HGSLSHFYYQFCEELFPFQDWWVVPVKVVFDQTVWSAIWNSIYFTVLGFL
A.thaliana_AT1G52870.2#1                HGSLSHFYYQFCEELFPFQDWWVVPVKVAFDQTVWSAIWNSIYFTVLGFL
B.napus_BN06MC04723_42271568@4          HGSLSHFYYQFCEELFPFQDWWVVPAKVAFDQTVWSAIWNSIYFTVLGFL
M.domestica_TC51384#1                   HGSLSHYYYQFCEALFPLEDWWVVPAKIAFDQTVWSAIWNSIYFVVLGFL
V.vinifera_GSVIVT00025998001#1           HGSLSHYYYQFCEALFPSKDWWVVPAKVVVDQTVWAAIWNSIYYVALGFL
A.lyrata_490222#1                       HGSLSHYYYQFCEALFPFQEWWVVPAKVAFDQTIWSAIWNSIYFTVLGLL
A.thaliana_AT4G03410.1#1                HGSLSHYYYQFCEALFPFQEWWVVPAKVAFDQTVWSAIWNSIYFTVLGLL
G.max_Glyma13g37540.1#1                 HGSLSHYYYQLCEALFPFQEWWVVPAKVAFDQTVWSAIWNSIYFVVLGLL
G.max_TC302765#1                        HGSLSHYYYQLCEALFPFQEWWVVPAKVAFDQTVWSAIWNSIYFVVLGLL
G.max_Glyma12g32920.1#1                 HGSLSHYYYQLCEALFPFQEWWVVPAKVAFDQTVWSAIWNSIYFVVLGLL
H.annuus_TC45682#1                      HGSLSHYYYQLCEALIPFKDWWVVPAKIAFDQTVWSAVWNSIYFDVLGLL
H.paradoxus_TA2132_73304#1              HGSLSHYYYRLCEALFPFKDWWVVPAKIVFDQTVWSAVWNSIYFVVLGLL
Triphysaria_sp_TC3527#1                 HGSLSHYYYQFCEALFPLHDWWVVPAKVIFDQTAWSAVWNSIYFTLLGVL
P.patens_TC32628#1                      HGSLSHYYYHVCEFLFPFQGWWVVPVKVAFDQTIWSAIWNSIYFITLGLL
                                        ******:**:.** *:* . *****.*: .*** *:*:*****:  **.*

Z.mays_ZM07MC32543_BFb0296A02@          RLESPTTIYNELKSTFWPMLTAGWKLWPFAHLVTYGVVPVEQRLLWVDCV
Zea_mays_GRMZM2G011269_T01#1            RLESPTTIYSELKSTFWPMLTAGWKLWPFAHLITYGVVPVEQRLLWVDCV
P.virgatum_TC39638#1                    RLESPTTIYKELKSTFWPMLTAGWKLWPFAHLITYGVIPVEQRLLWVDCV
S.bicolor_Sb01g015680.1#1               RLESPTTIYGELKSTFWPMLTAGWKLWPFAHLVTYGVVPVEQRLLWVDCV
O.sativa_LOC_Os03g38730.1#1             RLESPATISSELKSTFWPMLTAGWKLWPFAHLVTYGLVPVEQRLLWVDCV
L.esculentum_L450_PM22                  RLESPLAIFSELKATFLPMLTAGWKLWPFAHLITYGVIPVEQRLLWVDCV
```

Fig. 7B

```
S.tuberosum_TC167881#1                RFESPLAIFSELKATFLPMLTAGWKLWPFAHLITYGVIPVEQRLLWVDCV
N.benthamiana_TC11792#1               RFESPLGIFSELKATFWPMLTAGWKLWPFAHLITYGVVPVEQRLLWVDCV
N.tabacum_TC57360#1                   RFESPLAIFSELKATFWPMLTAGWKLWPFAHLITYGVVPVEQRLLWVDCV
I.nil_TC25#1                          RLESPTSIFSELKATFWPMLTAGWKLWPFAHLITYGVIPVEQRLLWVDCV
A.lyrata_474411#1                     RFESPLSIFKELKATFLPMLTAGWKLWPFAHLITYGLVPVEQRLLWVDCV
A.thaliana_AT1G52870.2#1              RFESPISIFKELKATFLPMLTAGWKLWPFAHLITYGLVPVEQRLLWVDCV
B.napus_BN06MC04723_42271568@4        RLESPLSIFKELKATFLPMLTAGWKLWPFAHLITYGLVPVEQRLLWVDCV
M.domestica_TC51384#1                 RFESPTKIFDELEGYICANADSGMETLAICSCG----------YVWPDPC
V.vinifera_GSVIVT00025998001#1        RRESPANIYGEVKSTFWPMLTAGWKLWPFAHLITYGVIPVEQRLLWVDCV
A.lyrata_490222#1                     RFQSPADIFSEIKTTFWPMLTAGWKLWPLAHLVTYGVIPVDQRLLWVDCI
A.thaliana_AT4G03410.1#1              RFQSPADIFSEIKTTFLPMLTAGWKLWPLAHLVTYGVIPVDQRLLWVDCI
G.max_Glyma13g37540.1#1               RFESLTNIYGELKSTFLPLLTAGWKLWPFAHLITYGVIPVEQRLLWVDCV
G.max_TC302765#1                      RFESLTNIYGELKSTFLPLLTAGWKLWPFAHLITYGVIPVEQRLLWVDCV
G.max_Glyma12g32920.1#1               RFESLTNIYGELKSTFLPLLTAGWKLWPFAHLITYGVIPVEQRLLWVDCV
H.annuus_TC45682#1                    RFESPANIFNELKTTFLPLLTAGWKLWPFAHLVTYGLIPVEQRLLWVDCV
H.paradoxus_TA2132_73304#1            RFESPSNIFNELKTTFLPLLTAGWKLWPFAHLVTYGLIPVEQRLLWVDCV
Triphysaria_sp_TC3527#1               RLESPVNIFNKWKATFFPMLTAGWKLWPFAHLITYGVIPVEQRLLWVDCV
P.patens_TC32628#1                    RFESPVRILKDLRETFFPLLTAGWKLWPFAHLITYGLVPVEQRLLWVDCV
                                      * :*   *  . .  : .   :* :  .:.              :* *

Z.mays_ZM07MC32543_BFb0296A02@        ELVWVTILSTYSNEKSEARNSDSTSTPAAS-KDNST--------------
Zea_mays_GRMZM2G011269_T01#1          ELVWVTILSTYSNEKSEARNSDSTSTPDAS-KDNSR--------------
P.virgatum_TC39638#1                  ELVWVTILSTYSNEKSEARNSDGTSTPDTS-KDNSR--------------
S.bicolor_Sb01g015680.1#1             ELVWVTILSTYSNEKSEARNSDSTSTPAASSKDNSR--------------
O.sativa_LOC_Os03g38730.1#1           ELIWVTILSTYSNEKSEARNSEDASTSNAS-NDNSI--------------
L.esculentum_L450_PM22                ELVWVTILSTYSNEKSESRVSEVIVEAEAQ-PPSVSPPQE----------
S.tuberosum_TC167881#1                ELIWVTILSTYSNEKSESRVSEVIVEAEAQ-PPSISPPQE----------
```

Fig. 7B

```
N.benthamiana_TC11792#1                ELIWVTILSTYSNEKSEASL----------------------------
N.tabacum_TC57360#1                    ELIWVTILSTYSNEKSEARVSEASVEAEMQ-PPSIGPPQE----------
I.nil_TC25#1                           ELIWVTILSTYSNEKSEARISEESVEATSN-PPSVGPSQK----------
A.lyrata_474411#1                      ELIWVTILSTYSNEKSEARISESVIETSSSSTTSIDPSKE----------
A.thaliana_AT1G52870.2#1               ELIWVTILSTYSNEKSEARISESVIETSSSSTTTIDPSKE----------
B.napus_BN06MC04723_42271568@4         ELIWVTILSTYSNEKSEARISESVIENPSSSTTPIDPSKE----------
M.domestica_TC51384#1                  ------------------------------------------------
V.vinifera_GSVIVT00025998001#1         ELIWVTILSTYSNEKSEARITEATQEANSRSSNNAEVSPVSEFRKQHHQA
A.lyrata_490222#1                      ELIWVTILSTYSNEKAEAQASGETNS--SSHSSEVCQLLERLNVLPVCSS
A.thaliana_AT4G03410.1#1               ELIWVTILSTYSNEKAEAQASEETNS--SSHSSED--------------
G.max_Glyma13g37540.1#1                ELIWVTILSTYSNEKSEARISEAASETGSSTSSENSKDIYTYTLHRSCIL
G.max_TC302765#1                       ELIWVTILSTYSNEKSEARISEAASETGSSTSSENSKE-----------
G.max_Glyma12g32920.1#1                ELIWVTILSTYSNEKSEARISEAASETGSSTSSENSKE-----------
H.annuus_TC45682#1                     ELVWVTILSTYSNEKSETRISDTSSEVDPNSPSKKH-------------
H.paradoxus_TA2132_73304#1             ELVWVTILSTYSNEKSETRISDTSSEVDPNSPSKKH-------------
Triphysaria_sp_TC3527#1                ELIWVTILSTYSNEKSEARISEAPVEANASLPPPEGPFE-----------
P.patens_TC32628#1                     EILWVTILSVFANEKAQQRMEISEVETVAVLSSGDAVEQKLPIEDSQ---

Z.mays_ZM07MC32543_BFb0296A02@         ------------------------------------------------
Zea_mays_GRMZM2G011269_T01#1           ------------------------------------------------
P.virgatum_TC39638#1                   ------------------------------------------------
S.bicolor_Sb01g015680.1#1              ------------------------------------------------
O.sativa_LOC_Os03g38730.1#1            ------------------------------------------------
L.esculentum_L450_PM22                 ------------------------------------------------
S.tuberosum_TC167881#1                 ------------------------------------------------
N.benthamiana_TC11792#1                ------------------------------------------------
```

Fig. 7B

```
N.tabacum_TC57360#1                 ----------------------------------------------------
I.nil_TC25#1                        ----------------------------------------------------
A.lyrata_474411#1                   ----------------------------------------------------
A.thaliana_AT1G52870.2#1            ----------------------------------------------------
B.napus_BN06MC04723_42271568@4      ----------------------------------------------------
M.domestica_TC51384#1               ----------------------------------------------------
V.vinifera_GSVIVT00025998001#1      LSIEHHAS------------------------------------------
A.lyrata_490222#1                   CACKYLTGLGDVSM------------------------------------
A.thaliana_AT4G03410.1#1            ----------------------------------------------------
G.max_Glyma13g37540.1#1             LAKINYSAIAAVKIGTLHSTLCSAMAKIEELNESIPKSKNKRNHGSTTFF
G.max_TC302765#1                    ----------------------------------------------------
G.max_Glyma12g32920.1#1             ----------------------------------------------------
H.annuus_TC45682#1                  ----------------------------------------------------
H.paradoxus_TA2132_73304#1          ----------------------------------------------------
Triphysaria_sp_TC3527#1             ----------------------------------------------------
P.patens_TC32628#1                  ----------------------------------------------------

Z.mays_ZM07MC32543_BFb0296A02@      -------------------------
Zea_mays_GRMZM2G011269_T01#1        -------------------------
P.virgatum_TC39638#1                -------------------------
S.bicolor_Sb01g015680.1#1           -------------------------
O.sativa_LOC_Os03g38730.1#1         -------------------------
L.esculentum_L450_PM22              -------------------------
S.tuberosum_TC167881#1              -------------------------
N.benthamiana_TC11792#1             -------------------------
N.tabacum_TC57360#1                 -------------------------
```

Fig. 7B

```
I.nil_TC25#1                    ----------------------------------------
A.lyrata_474411#1               ----------------------------------------
A.thaliana_AT1G52870.2#1        ----------------------------------------
B.napus_BN06MC04723_42271568@4  ----------------------------------------
M.domestica_TC51384#1           ----------------------------------------
V.vinifera_GSVIVT00025998001#1  ----------------------------------------
A.lyrata_490222#1               ----------------------------------------
A.thaliana_AT4G03410.1#1        ----------------------------------------
G.max_Glyma13g37540.1#1         VFVDYLFILFFLGFLCFTVFKILSI
G.max_TC302765#1                ----------------------------------------
G.max_Glyma12g32920.1#1         ----------------------------------------
H.annuus_TC45682#1              ----------------------------------------
H.paradoxus_TA2132_73304#1      ----------------------------------------
Triphysaria_sp_TC3527#1         ----------------------------------------
P.patens_TC32628#1              ----------------------------------------
```

Fig. 7B

```
CLUSTAL 2.0.11 multiple sequence alignment

L.esculentum_L450_PM22                --MATINGLTSPHPFLSLQ-----KSNSKP-FLSLSKPMKSQFLFKGLKQ
S.tuberosum_TC167881#1                --MATINGLTSPQPFLSLQ-----KSNSKP-FLSLSKPMKSQFLFKGLKQ
N.benthamiana_TC11792#1               --MATINGLTSPQPFLSLP-----KPKRKS-SHSHFKPLKPQILSKGLKQ
N.tabacum_TC57360#1                   --MSTINGLTSPQPFLSLP-----KPKPKP-SYSHFKPLKPRILSKGLKQ
I.nil_TC25#1                          --MATINGLPTPQTLRSLS-----KPNPKTGSPSHFKPTKLQLFPRPVSK
A.lyrata_474411#1                     MAAASLHTSISPRSFLPLS-----KPSLKP--------HRSQIFLRNKQR
A.thaliana_AT1G52870.2#1              MAAASLHTSISPRSFLPLS-----KPSLKP--------HRSQILLRNKQR
B.napus_BN06MC04723_42271568@4        MAASSLNTSISPRSFLPLS-----KPSPKP--------HRSQVFLRNKQR
G.max_Glyma12g34950.1#1               --MASVHSTVAPRTFLPSIP-KPRAPLHAAKFVAG-GAHNFRRLSRNLIL
G.max_GM06MC03382_49802960@335        --MASVPSTVAPRTFLPSIP-KPRAPLHAAKSVAS-AARNFPRLSHNLIL
G.max_Glyma13g35620.1#1               --MASVPSTVAPRTFLPSIP-KPRAPLHAAKSVAS-AARNFPRLSHNLIL
P.vulgaris_TC9799#1                   --MTSVHSTVAPRAFLPSIS-KPKPRAPHTKSFAGGVVHDFPRFSRSCIF
G.raimondii_TC2516#1                  --MASLHPTISPQSLLSLT--KPRKPISRH---LSASSLTTSKLPEGLAF
G.hirsutum_TC165523#1                 --MASLHPTISPQSLLSLT--KPRKPISRH---LSASSLTTSKLPEGLAF
T.cacao_TC9688#1                      --MASLHTTITPQGLLYLS--KSGNPISRQ---ASASPLTSSKLPESLTF
P.trifoliata_TA5129_37690#1           --MASLS----PQTLLSIS--KPNKPNSKQ---LPSNACLLSKLSQNPNF
P.persica_TC12602#1                   --MASIHS-IAPQSFLPLPNSKPRKPTNAHPKPILSSNLCNSKLSKNQTF
L.sativa_TC17162#1                    --MAAASISTTFLPLYKPIK-----------SFVPISQAPKFFTKTKHG
C.intybus_TA545_13427#1               --MAAASISTTFLPLHKPIKDF---------KSFVPASQKPKFLTKTKHG
H.argophyllus_TA2174_73275#1          --MAAASINTSFKPLYKPTKDF---------KSFVPISPTPKLLTKSKNL
Aquilegia_sp_TC27124#1                --MAASVNSLTPQRFLHPLSSKPINKESKLSKSNFPLKGISKNSTFLRTN
V.vinifera_GSVIVT00027620001#1        ------------MGFWRELRWK--------------------------K
Z.mays_ZM07MC32543_BFb0296A02@        -------MAAAAVATVATPSP-ASRRLFP-------SAGPSLLRLP-RPTK
Zea_mays_GRMZM2G011269_T01#1          -------MAAAAVATVASPSPAASRRMLP-------SAAPSFLRLP-RPTS
```

Fig. 7C

```
S.bicolor_Sb01g015680.1#1                ----MAAASASVATVASPSSPASRRLLLPL---PSSAAPSLLRLPSRPAR
P.virgatum_TC39638#1                     -----MAAAAAVATVSSPSP-ASRRLLPS----SSSAVPSLLRLP-RPGR
O.sativa_LOC_Os03g38730.1#1              ----MAAMAAAVASSPSPSP-ASTRLLRGHHPPPSCSAPSLLRLT-RSSR
P.patens_TC32628#1                       ---MAGVGPACVRSVASQGFGAPCVGLPQRN--RGVRGLSGCGLPERVSS
Triphysaria_sp_TC3527#1                  ----------------------------------------------------
A.lyrata_490222#1                        --MAALCC-----------------PTTIIVSGKLSSRIKTACQLPRIN
A.thaliana_AT4G03410.1#1                 --MAALCCC----------------PTTIIVSGKVSSRIKTAGPLPRIN
G.max_Glyma13g37540.1#1                  --MATVS--------------------SPCTVFPILTSTLSKPHLAPLSC
G.max_TC302765#1                         --MATVS--------------------SPCTVFPILTSTLSKPHLAPLSC
G.max_Glyma12g32920.1#1                  --MATVS--------------------SLCAVFPILTPTLSKPRLAPLSC
H.annuus_TC45682#1                       --MATVNNHVVF---------FSGHKISPPTNPISYSTTFSISPPTKLFS
H.paradoxus_TA2132_73304#1               --MMSY-----F---------FSGHKVSPPTNPISYSTTFSISPPTKLFS
M.domestica_TC51384#1                    --MATLNTIAAHNLPSPSTSSFATQSHKPNSLSPFFCTQTSSKPHPPRLP
V.vinifera_GSVIVT00025998001#1           --MATLNTITAHRLFPLPKSITESNNSYTVCISTSSSQLFLNKSNPSLNS

L.esculentum_L450_PM22                   QNQQ--------LRDWAVVGSVSNEAD-------EIRVQSSDVMDQQ---
S.tuberosum_TC167881#1                   QNQQ--------LRDWAVVGSVTNEAD-------VIPVQSTDVTDQQ---
N.benthamiana_TC11792#1                  QNQQ--------LRDWAVVGSVDNETD-------LIPVQSDDTTDQQ---
N.tabacum_TC57360#1                      QNQQNQ-----QLRDWAVVGSVANETD-------LIPVQSNDTTDQQ---
I.nil_TC25#1                             QFAA----------------AAAKELD-------VIPVQSSDSTDQQ---
A.lyrata_474411#1                        -----------------TCALIRDEID-------VIPVQSRDRTDHE---
A.thaliana_AT1G52870.2#1                 NC--------------VSCALIRDEID-------LIPVQSRDRTDHE---
B.napus_BN06MC04723_42271568@4           SC--------------VSCALINDETD-------VIPVKS-----HD---
G.max_Glyma12g34950.1#1                  SGNKRA---------VAANSAA-EEFD-------VISVQSDDITDQQ---
G.max_GM06MC03382_49802960@335           SGNKRA---------VAVNSAS-EEFD-------VISVQSEDITDQQ---
G.max_Glyma13g35620.1#1                  SGNKRA---------VAVNSAS-EEFD-------VISVQSEDITDQQ---
```

Fig. 7C

```
P.vulgaris_TC9799#1                SGNNCA---------VAANAAG-DDLN-------VISVKSDDITDQQ---
G.raimondii_TC2516#1               SGTKQK--NKRANSVVVVKSLA-EELD-------VIPVQSEDVTDMQ---
G.hirsutum_TC165523#1              SGTKQK--NKRENSVVVVKSLA-EELD-------VIPVQSEDVTDIQ---
T.cacao_TC9688#1                   SRNKWRKNSNRKANSVVVNSLA-EELD-------VIPVQSEDVTDMQ---
P.trifoliata_TA5129_37690#1        SRNKQR----KLNSWVVNSVVN-EELD-------VLPVQSQDLTDMQ---
P.persica_TC12602#1                LRNKRK------DRWILKSLVDQEECD-------VIPVQSTDCTDQQ---
L.sativa_TC17162#1                 AGAG-------------TVRAVSKEQD-------VIPVQSNDFTDHQ---
C.intybus_TA545_13427#1            AGAV-------------TVKAVSKEQD-------VIPLQSNDFTDHQ---
H.argophyllus_TA2174_73275#1       AGAG--------------AVKAVAKEQD-------VIPVKS----DDQ---
Aquilegia_sp_TC27124#1             NRNW-------------IIGSIKEEID-------VIPVQSNDLIDQQ---
V.vinifera_GSVIVT00027620001#1     RGNW-------------WVGSLVK-------------------------
Z.mays_ZM07MC32543_BFb0296A02@     RLR--R----------ALRFAAAGDEAD-------VLPGPGAEG-------
Zea_mays_GRMZM2G011269_T01#1       RLR--R----------ATQVAAAGGEAD-------VLPGPGAEG-------
S.bicolor_Sb01g015680.1#1          RLRRAR----------ALRVAAAGDEAD-------VLPGPGAEG-------
P.virgatum_TC39638#1               RLR--R----------ALRVVSAGGEAD--------VLPGPGGEA-------
O.sativa_LOC_Os03g38730.1#1        RLR-------------LRAAAAAEEAD-------VLPGPGAEG-------
P.patens_TC32628#1                 GKFCSS----------SGKGGFFAGKKDGFSIRCSSLPGDGVAAGTPKTVI
Triphysaria_sp_TC3527#1            ------------------------------------------------
A.lyrata_490222#1                  PFKNRKR--LITERRNLIVKSIIEDRE-------AVDVKNENFKPEEEQT
A.thaliana_AT4G03410.1#1           PFKNQKR--LITERRNLIVKSIIEDRE-------AIDVKNDNFKAEEELS
G.max_Glyma13g37540.1#1            RTQKRS--------RVS-IITSVAEDRE-------IVPVSEDRGIRLNE--
G.max_TC302765#1                   RTQKRS--------RVS-IITSVAEDRE-------IVPVSEDRGIRLNE--
G.max_Glyma12g32920.1#1            RIPKRR--------RIS-IITSVAEDRE-------IVPVSEGRGTRLNE--
H.annuus_TC45682#1                 QSRKNH--------KVYPVMKSIAEDKE-------VVADEQNGSERESG--
H.paradoxus_TA2132_73304#1         QSKQNH--------KVYPVMKSIAEDKE-------VVADEQNGSERESG--
M.domestica_TC51384#1              AFPTRK--------KRSWIVRSITEDRE-------VVPLKKK-DTGEQDSS
V.vinifera_GSVIVT00025998001#1     TFSKKK--------RLNWVINAVAEDQD-------LAPAQTSGSKVHQDED
```

Fig. 7C

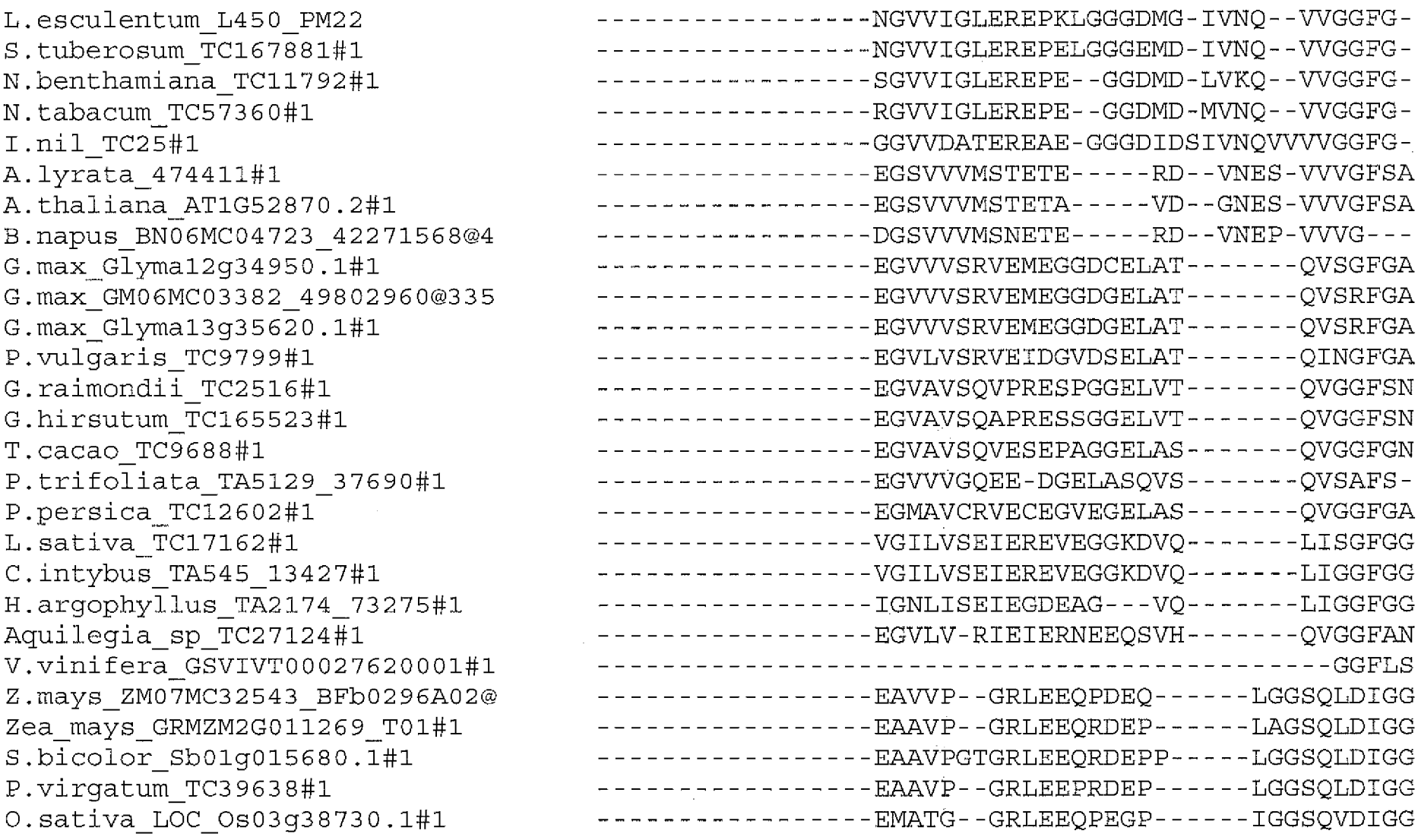

L.esculentum_L450_PM22 ----------NGVVIGLEREPKLGGGDMG-IVNQ--VVGGFG-
S.tuberosum_TC167881#1 ----------NGVVIGLEREPELGGGEMD-IVNQ--VVGGFG-
N.benthamiana_TC11792#1 ----------SGVVIGLEREPE--GGDMD-LVKQ--VVGGFG-
N.tabacum_TC57360#1 ----------RGVVIGLEREPE--GGDMD-MVNQ--VVGGFG-
I.nil_TC25#1 ----------GGVVDATEREAE-GGGDIDSIVNQVVVVGGFG-
A.lyrata_474411#1 ----------EGSVVVMSTETE-----RD--VNES-VVVGFSA
A.thaliana_AT1G52870.2#1 ----------EGSVVVMSTETA-----VD--GNES-VVVGFSA
B.napus_BN06MC04723_42271568@4 ----------DGSVVVMSNETE-----RD--VNEP-VVVG---
G.max_Glyma12g34950.1#1 ----------EGVVVSRVEMEGGDCELAT-------QVSGFGA
G.max_GM06MC03382_49802960@335 ----------EGVVVSRVEMEGGDGELAT-------QVSRFGA
G.max_Glyma13g35620.1#1 ----------EGVVVSRVEMEGGDGELAT-------QVSRFGA
P.vulgaris_TC9799#1 ----------EGVLVSRVEIDGVDSELAT-------QINGFGA
G.raimondii_TC2516#1 ----------EGVAVSQVPRESPGGELVT-------QVGGFSN
G.hirsutum_TC165523#1 ----------EGVAVSQAPRESSGGELVT-------QVGGFSN
T.cacao_TC9688#1 ----------EGVAVSQVESEPAGGELAS-------QVGGFGN
P.trifoliata_TA5129_37690#1 ----------EGVVVGQEE-DGELASQVS-------QVSAFS-
P.persica_TC12602#1 ----------EGMAVCRVECEGVEGELAS-------QVGGFGA
L.sativa_TC17162#1 ----------VGILVSEIEREVEGGKDVQ-------LISGFGG
C.intybus_TA545_13427#1 ----------VGILVSEIEREVEGGKDVQ-------LIGGFGG
H.argophyllus_TA2174_73275#1 ----------IGNLISEIEGDEAG---VQ-------LIGGFGG
Aquilegia_sp_TC27124#1 ----------EGVLV-RIEIERNEEQSVH-------QVGGFAN
V.vinifera_GSVIVT00027620001#1 ----------GGFLS
Z.mays_ZM07MC32543_BFb0296A02@ ----------EAVVP--GRLEEQPDEQ------LGGSQLDIGG
Zea_mays_GRMZM2G011269_T01#1 ----------EAAVP--GRLEEQRDEP------LAGSQLDIGG
S.bicolor_Sb01g015680.1#1 ----------EAAVPGTGRLEEQRDEPP-----LGGSQLDIGG
P.virgatum_TC39638#1 ----------EAAVP--GRLEEPRDEP------LGGSQLDIGG
O.sativa_LOC_Os03g38730.1#1 ----------EMATG--GRLEEQPEGP------IGGSQVDIGG

Fig. 7C

```
P.patens_TC32628#1                KSSCSKILRWSEEDVVSSGGVPVVQALPDDAETHSG----ANGMLVELRE
Triphysaria_sp_TC3527#1           ----MLLLCWLLD------LLPLPSCLP----------------LIMIIG
A.lyrata_490222#1                 --------------------------------------------------
A.thaliana_AT4G03410.1#1          ED------------------------------------------------
G.max_Glyma13g37540.1#1           --------------------VDGFQPSEPHTES-----------------
G.max_TC302765#1                  --------------------VDGFQPSEPHTES-----------------
G.max_Glyma12g32920.1#1           --------------------VEGFQPSEPHTES-----------------
H.annuus_TC45682#1                --------------------VVGEG-SEG---------------------
H.paradoxus_TA2132_73304#1        --------------------VVGAG-SEG---------------------
M.domestica_TC51384#1             LL------------------LNGSGEFQALSTP------------SEEK-
V.vinifera_GSVIVT00025998001#1    LPR---------------SSLDGSEDSEGLSSSSVS---------SQGK-

L.esculentum_L450_PM22            --NEGRLSFE------GGG-GFSSASGVGGSE----------------SK
S.tuberosum_TC167881#1            --NEGRLSFE------GAG-GFSSASGVGGS-----------------
N.benthamiana_TC11792#1           --NEGRLSFE------GAARGFSSSSSGSNGK----------------LE
N.tabacum_TC57360#1               --NEGRLSFE------GAAGGFSSSSSGSNGK----------------SE
I.nil_TC25#1                      --NEGRLSFEGPT-GFGSSSSSSSAASSSGGA----------------GE
A.lyrata_474411#1                 ATSEGQLSLEG-----FPSS---GGDLGDEKR----------------GE
A.thaliana_AT1G52870.2#1          ATSEGQLSLEG-----FPSSSSSGADLGDEKR----------------RE
B.napus_BN06MC04723_42271568@4    -----QLSFEG-----FPSSSP-AADMGDDKS----------------RE
G.max_Glyma12g34950.1#1           N--EGLLSLEG-----FSSSSSSSSS-LVGNE----------------SE
G.max_GM06MC03382_49802960@335    N--EGLLSLEG-----FSSSSSSSSSSLVGNE----------------SE
G.max_Glyma13g35620.1#1           N--EGLLSLEG-----FSSSSSPSSSSLIGSE----------------S-
P.vulgaris_TC9799#1               K--DGLLSLEG-----FSSSSSTSTSALVGSE----------------SE
G.raimondii_TC2516#1              ---EGMLSFEG-----VSSSGSPSGNGFGDGQ----------------GS
G.hirsutum_TC165523#1             ---EGMLSFEG-----VSSSGSPSGNGFGDGQ----------------GS
```

Fig. 7C

```
T.cacao_TC9688#1                    ---EGRLSFEG-----FSASGSASGSGFGDGE----------------GG
P.trifoliata_TA5129_37690#1         ---DGTLSFDG-----FSSASSAS----ADER----------------RG
P.persica_TC12602#1                 S--EGRLSFEG-----AGGFGSSG----VGNE----------------RE
L.sativa_TC17162#1                  SE--GRLSFEGG----FSSASSSGDGNQVVE-----------------G
C.intybus_TA545_13427#1             SE--GRLSFEGG----FSSASSSGDGNQMVE-----------------G
H.argophyllus_TA2174_73275#1        SD--GRLSFEGG----FSSAAS-GSGGEVVE-----------------G
Aquilegia_sp_TC27124#1              EG--RFLFEGAGEINGFSSSLSSSSSASVVSEV-------------IGNE
V.vinifera_GSVIVT00027620001#1      G----LFFFGQG-------------------------------------E
Z.mays_ZM07MC32543_BFb0296A02@      LAFQGDVGGG------FTGGGAGSG--------------------ASGG
Zea_mays_GRMZM2G011269_T01#1        LAFQGDVGGG------FTGGGAGSG--------------------ASGG
S.bicolor_Sb01g015680.1#1           LAFQGDVGGG------FTGGGSGSGS-------------------GASG
P.virgatum_TC39638#1                LAFQGDVGGG------FAGGGGGVGS-------------------GASG
O.sativa_LOC_Os03g38730.1#1         LAFQGDMGGGG-----FAGGSGGAG----------------------AG
P.patens_TC32628#1                  TDRPQPIGSGT-----YGSGNQLSMEMAVNEAQTMTATMYRPKLLWFTKS
Triphysaria_sp_TC3527#1             MDGP----------------------------------------------
A.lyrata_490222#1                   ----------------------------------------------------
A.thaliana_AT4G03410.1#1            ----------------------------------------------------K
G.max_Glyma13g37540.1#1             ----------------------------------------------------
G.max_TC302765#1                    ----------------------------------------------------
G.max_Glyma12g32920.1#1             ----------------------------------------------------
H.annuus_TC45682#1                  ----------------------------------------------------
H.paradoxus_TA2132_73304#1          ----------------------------------------------------
M.domestica_TC51384#1               ----------------------------------------------------GE
V.vinifera_GSVIVT00025998001#1      ----------------------------------------------------GD


L.esculentum_L450_PM22              EEDVEKLID-RTINATIVLAAGSFAITKLLTIDHDYWHGWTLFEILRYAP
```

Fig. 7C

```
S.tuberosum_TC167881#1                  --DVEKLID-RAINATIVLAAGSFAITKLLTIDHDYWHGWTLFEILRYAP
N.benthamiana_TC11792#1                 QEDMNRVID-RAINATIVLAAGSFAITKLLTIDHDYWHGWTVYEIFRYAP
N.tabacum_TC57360#1                     QEDMDRVID-RAINATIVLAAGSFAITKLLTIDHDYWHGWTIYEILRYAP
I.nil_TC25#1                            ELGMEKVVD-RAINATIVLAAGTFAITKLLTIDHNYWHGWTLFEILRYAP
A.lyrata_474411#1                       NEEREKMID-RTINATIVLAAGSYAITKLLTIDHDYWHVWTLLEILRYAP
A.thaliana_AT1G52870.2#1                NEEMEKMID-RTINATIVLAAGSYAITKLLTIDHDYWHGWTLFEILRYAP
B.napus_BN06MC04723_42271568@4          SEEMDKMID-RSINATIVLAAGTYAITKLLTIDHDYWHGWTLFEILRYAP
G.max_Glyma12g34950.1#1                 -EDMEKLID-RTINATIVLAAGTFAVTKLLTIDSDYWHGWTLYEILRYAP
G.max_GM06MC03382_49802960@335          -EDMEKLID-RTINATIVLAAGTFAVTKLLTIDSDYWHGWTLYEILRYAP
G.max_Glyma13g35620.1#1                 -VDMEKLID-RTINATIVLAAGTFAVTKLLTIDSDYWHGWTLYEILRYAP
P.vulgaris_TC9799#1                     -EDVEKLID-RTINATIVLAAGAFAITKLLTIDSDYWHGWTLYEILRYAP
G.raimondii_TC2516#1                    QEELDRLID-RTINATIVLAAGTYAITKLLTIDQNYWQGWTIYEIVRYAP
G.hirsutum_TC165523#1                   QEELDRLID-RTINATIVLAAGTYAITKLLTIDQNYWQGWTIYEIVRYAP
T.cacao_TC9688#1                        AEDMEKLID-RTINATIVLAAGTFALTKLLTIDHNYWQGWTLYEILRYAP
P.trifoliata_TA5129_37690#1             NEEIERLID-RTINATIVLAVGTFAITKLLTIDQDYWHGWTIFEIVRYAP
P.persica_TC12602#1                     SEEFERLVD-RTINATIVLAAGTFAITKLLTIDQDYWHGWTLYEILRYAP
L.sativa_TC17162#1                      DN-LDKLID-RTINATIVLAAGTFGITKLLTIDYDYWHGWTIFEILRYAP
C.intybus_TA545_13427#1                 EN-LDKLMD-RTINATIVLAAGTFGITKLLTIDYDYWHGWTIFEILRYAP
H.argophyllus_TA2174_73275#1            ETSVEKLID-RGINAAIVLAAGTFGITKMLTIDYDYWHGWTLFEILRYAP
Aquilegia_sp_TC27124#1                  EEDVEMLID-RAINAVIVLAAGTYAITKLLTIDRDYWHGWTIFEILRYAP
V.vinifera_GSVIVT00027620001#1          AEDVERLID-RSINATIVLAAGSFAVTKLLTIDADYWHGWTIFEILRYAP
Z.mays_ZM07MC32543_BFb0296A02@          GG-GNKMLD-RGINTAIVLGASTYALTKLLTVDQDYWHGWTIFEILRYMP
Zea_mays_GRMZM2G011269_T01#1            GGDGNKMLD-RGINTAIVLAASTYALTKLLTVDQDYWHGWTIFEILRYMP
S.bicolor_Sb01g015680.1#1               GGDGNKMLD-RGINTAIVLGASTYALTKLLTVDQDYWHGWTIFEILRYMP
P.virgatum_TC39638#1                    GGGGNKTLD-RGINTAIVLGASTFALTKLLTVDQDYWHGWTIFEILRYMP
O.sativa_LOC_Os03g38730.1#1             GGDGNKMLD-RGINTAIVLGASTYALTKLLTVDHDYWHGWTIFEILRYMP
P.patens_TC32628#1                      EDEESRLLD-RTFNALLAGAAISYAVTKAVTVDHDVWHGWTVFEVLKYAP
```

Fig. 7C

```
Triphysaria_sp_TC3527#1                  ----------------------------------------FFEILRYAP
A.lyrata_490222#1                        -EDTDQLMS-RGINAAIVLAAGTVAVTKLLTIDHDYWQGWTLYEILRYAP
A.thaliana_AT4G03410.1#1                 VEDTDRLMS-RGINAAIVLAAGTVAVTKLLTIDHDYWQGWTLYEILRYAP
G.max_Glyma13g37540.1#1                  DVVVPRLTSSSTVNAIIVLGFGTFAVTKLLTIDHDYWHGWTLFEIVRYIP
G.max_TC302765#1                         DVVVPRLTSSSTVNAIIVLGFGTFAVTKLLTIDHDYWHGWTLFEIVRYIP
G.max_Glyma12g32920.1#1                  EVVVSSLTSSSSVNAIIVLGFGTFAVTKLLTIDHDYWHGWTLYEIVRYIP
H.annuus_TC45682#1                       ---IDKFAS-KAINASIVLGFGTLAVTRLLTIDHEYWHGWTLYEILRYAP
H.paradoxus_TA2132_73304#1               ---IDKFAS-KAINASIVLGFGTLAVTRLLTIDHEYWHGWTLYEILRYAP
M.domestica_TC51384#1                    GDDLDKLTS-RAINALIVLGFGTFAVSKLLTIDHDYWHGWTLYEILRYVP
V.vinifera_GSVIVT00025998001#1           NYEFDRLRS-KTINATIVLAGGTLAITRLLTIDHDYWHGWTLYEVLRYAP
                                                                              . *:.:* *

L.esculentum_L450_PM22                   QHNWVAYEEALKRNPVLAKMVISGVVYSVGDWIAQCYEGKPLFEIDRARM
S.tuberosum_TC167881#1                   QHNWVAYEEALKRNPVLAKMVISGVVYSVGDWIAQCYEGKPLFEIDRARM
N.benthamiana_TC11792#1                  QHNWIAYEEALKRNPVLAKMVISGVVYSIGDWIAQCYEGKPLFDFDRARM
N.tabacum_TC57360#1                      QHNWIAYEEALKRNPVLAKMVISGVVYSLGDWIAQCYEGKPLFDFDRARM
I.nil_TC25#1                             QHNWSAYEEALKENPVLAKMVISGVVYSLGDWIAQCYEGKPLFEFDRARM
A.lyrata_474411#1                        QHNWIAYEEALKQNPVLAKMVISGVVYSVGDWIAQCYEGKPLFEIDRART
A.thaliana_AT1G52870.2#1                 QHNWIAYEEALKQNPVLAKMVISGVVYSVGDWIAQCYEGKPLFEIDRART
B.napus_BN06MC04723_42271568@4           QHNWLAYEEALKRNPVLAKMVISGVVYSVGDWIAQCYEGKPLFEIDRART
G.max_Glyma12g34950.1#1                  QHNWSAYEEALKTNPVLAKMMISGIVYSIGDWIAQCFEGKPLFEFDRARM
G.max_GM06MC03382_49802960@335           QHNWSAYEEALKTNPVLAKMMISGIVYSLGDWIAQCFEGKPLFEFDRARM
G.max_Glyma13g35620.1#1                  QHNWSAYEEALKTNPVLAKMMISGIVYSLGDWIAQCVEGKPLFEFDRARM
P.vulgaris_TC9799#1                      QHNWSAYEEALKTNPVLAKMMISGIVYSVGDWIAQCFEGKPLFEFDRARM
G.raimondii_TC2516#1                     QHNWSAYEEALRTNPVLAKMVISGIVYSIGDWIAQCFEGKPLFEFDRMRM
G.hirsutum_TC165523#1                    QHNWSAYEEALRTNPVLAKMVISGIVYSIGDWIAQCFEGKPLFEFDRMRM
T.cacao_TC9688#1                         QHNWSAYEEVLKTNPVLAKMVISGVVYSLGDWIAQCFEGKPLFEFDRTRM
```

Fig. 7C

```
P.trifoliata_TA5129_37690#1          LHNWTAYEEALKTNPVLAKMVISGVVYSLGDWIAQCFEGKPLFEFDRTRM
P.persica_TC12602#1                  QHNWSAYEEALKTNPVLAKMVISGVVYSVGDWIAQCFEGKPLFEFDRTRM
L.sativa_TC17162#1                   QHNWSAYEEALKENPLLAKMMISGVVYSVGDWIAQCCEGKPLLEFDRTRM
C.intybus_TA545_13427#1              QHNWSAYEEALKENPLLAKMMISGVVYSIGDWIAQCCEGKPLLEFDRTRM
H.argophyllus_TA2174_73275#1         QHNWSAYEEALKQNPLLAKMIISGVVYSIGDWIAQCCEGKPLLEFDRTRM
Aquilegia_sp_TC27124#1               LHNWAAYEEALKENPVLAKMVISGVVYSLGDWIAQCYEGKPLFDFDRTRM
V.vinifera_GSVIVT00027620001#1       QHNWSAYEEALKTNPVFAKMVISGVVYSLGDWIAQCYEGKPLFEFDRARM
Z.mays_ZM07MC32543_BFb0296A02@       EHNWSAYEEALKANPVLAKMMISGVVYSLGDWIAQCYEGKPIFDFDRARM
Zea_mays_GRMZM2G011269_T01#1         EHNWSAYEEALKANPVLAKMMISGVVYSLGDWIAQCYEGKPIFDFDRARM
S.bicolor_Sb01g015680.1#1            EHNWSAYEEALKANPVLAKMMISGVVYSLGDWIAQCYEGKPIFDFDRARM
P.virgatum_TC39638#1                 EHNWSAYEEALKANPVLAKMMISGVVYSLGDWIAQCYEGKPIFDFDRARM
O.sativa_LOC_Os03g38730.1#1          EHNWSAYEEALKTNPVLAKMMISGVVYSLGDWIAQCYEGKPIFEFDRARM
P.patens_TC32628#1                   LHNWHAYEEFLSANPVLAKMMISGVVYSIGDWIGQCVEGKPVLEFSRVRL
Triphysaria_sp_TC3527#1              QHNWSAYEEALKANPVLAKMVISGVVYSVGDWIAQCYEGKPIFEYDRARM
A.lyrata_490222#1                    EHNWVAYEQILKTNPVLAKMAISGIVYSLGDWIAQCYEGKPLFEFDRARV
A.thaliana_AT4G03410.1#1             EHNWFAYEQILKTNPVLAKMAISGIVYSLGDWIAQCYEGKPLFEFDRTRV
G.max_Glyma13g37540.1#1              EHNWIAYEQALKANPVLAKMAISGIVYSIGDWIAQCYEGKPLFEFDRTRV
G.max_TC302765#1                     EHNWIAYEQALKANPVLAKMAISGIVYSIGDWIAQCYEGKPLFEFDRTRV
G.max_Glyma12g32920.1#1              EHNWIAYEQALKANPVLAKMAISGIVYSIGDWIAQCYEGKPLFEFDRTRV
H.annuus_TC45682#1                   EHNWIAYEEALKENPVLAKMVISGIVYSLGDWIAQCYEGKPLFEFDRIRL
H.paradoxus_TA2132_73304#1           EHNWIAYEEALKENPVLAKMVISGIVYSLGDWIAQCYEGKPLFEFDRTRL
M.domestica_TC51384#1                EHNWIAYEQALKANPVFAKMVISGVVYTLGDWIAQCFEGKPLLEFDRKRM
V.vinifera_GSVIVT00025998001#1       EHNWVAYEEALKTNPVLAKMAISGAVYSIGDWIAQCYEGKPLFEFDLTRM
                                      *** ***: *  **::*** *** **::****.** ****::: .  *

L.esculentum_L450_PM22               LRSGLVGFTLHGSLSHYYYQFCEALFPFEDWWVVPVKVAFDQTVWSAIWN
S.tuberosum_TC167881#1               LRSGLVGFTLHGSLSHYYYQFCEALFPFEDWWVVPVKVAFDQTVWSAIWN
```

Fig. 7C

```
N.benthamiana_TC11792#1                 LRSGFVGFTLHGSLSHYYYQFCEALFPFQDWWVVPVKVAFDQTVWSAIWN
N.tabacum_TC57360#1                     LRSGLVGFTLHGSLSHYYYQICEALFPFQDWWVVPVKVAFDQTVWSAIWN
I.nil_TC25#1                            FRSGLTGFTLHGSLSHYYYHFCEALFPFDDWWVVPAKVAFDQTVWSAIWN
A.lyrata_474411#1                       LRSGLVGFTLHGSLSHFYYQFCEELFPFQDWWVVPVKVVFDQTVWSAIWN
A.thaliana_AT1G52870.2#1                LRSGLVGFTLHGSLSHFYYQFCEELFPFQDWWVVPVKVAFDQTVWSAIWN
B.napus_BN06MC04723_42271568@4          LRSGLVGFTLHGSLSHFYYQFCEELFPFQDWWVVPAKVAFDQTVWSAIWN
G.max_Glyma12g34950.1#1                 FRSGLVGFTLHGSLSHFYYQFCEELFPYKEWWVVPAKVAFDQTAWSALWN
G.max_GM06MC03382_49802960@335          FRSGLVGFTLHGSLSHFYYQFCEELFPYKEWWVVPAKVAFDQTAWSALWN
G.max_Glyma13g35620.1#1                 FRSGLVGFTLHGSLSHFYYQFCEELFPYKEWWVVPAKVAFDQTAWSALWN
P.vulgaris_TC9799#1                     FRSGIVGFTLHGSLSHYYYYFCEDLFPYKEWWVVPAKVAFDQTAWSAIWN
G.raimondii_TC2516#1                    FRSGLVGFTLHGSLSHYYYQFCEELFPFQDWWVVPVKVAFDQTAWAAVWN
G.hirsutum_TC165523#1                   FRSGLVGFTLHGSLSHYYYQFCEELFPFQDWWVVPVKVAFDQTAWAAVWN
T.cacao_TC9688#1                        FRSGLVGFTLHGSLSHYYYQFCEELFPSQDWWVVPAKVVFDQTAWAAVWN
P.trifoliata_TA5129_37690#1             FRSGLVGFTLHGSLSHYYYQFCEELFPFQDWWVVPAKVAFDQTAWAAAWN
P.persica_TC12602#1                     LRSGLVGFTLHGSLSHYYYHPCEELIPFQDWWVVPAKVAFDQTVWAAIWN
L.sativa_TC17162#1                      FRSGLVGFTLHGSLSHYYYYFCEALFPFQDWWVVPVKVAFDQTAWSAVWN
C.intybus_TA545_13427#1                 FRSGLVGFTLHGSLSHYYYYFCEALFPFQDWWVVPVKVAFDQTAWSAIWN
H.argophyllus_TA2174_73275#1            FRSGLVGFTLHGSLSHYYYYFCEALFPFQDWWVVPVKVAFDQTAWAAVWN
Aquilegia_sp_TC27124#1                  FRSGLVGFSLHGSLSHYYYQFCEALFPFQDWWVVPVKVVFDQTAWAALWN
V.vinifera_GSVIVT00027620001#1          LRSGLVGFTLHGSLSHYYYQFCEALFPFQDWWVVPAKVAFDQTLWAAVWN
Z.mays_ZM07MC32543_BFb0296A02@          FRSGLVGFTLHGSLSHYYYHICEALFPFKDWWVVPAKVAFDQTIWSAIWN
Zea_mays_GRMZM2G011269_T01#1            FRSGLVGFTLHGSLSHYYYHICEALFPFKDWWVVPAKVAFDQTVWSAIWN
S.bicolor_Sb01g015680.1#1               FRSGLVGFTLHGSLSHYYYHICEALFPFKDWWVVPAKVAFDQTIWSAIWN
P.virgatum_TC39638#1                    FRSGLVGFTLHGSLSHYYYHICEALFPFKDWWVVPAKVAFDQTIWSAIWN
O.sativa_LOC_Os03g38730.1#1             FRSGLVGFTLHGSLSHYYYHFCEALFPFKDWWVVPAKVVFDQTAWSAIWN
P.patens_TC32628#1                      LRSGLVGFCLHGSLSHYYYHVCEFLFPFQGWWVVPVKVAFDQTIWSAIWN
Triphysaria_sp_TC3527#1                 FRSGLVGFTLHGSLSHYYYQFCEALFPLHDWWVVPAKVIFDQTAWSAVWN
```

Fig. 7C

```
A.lyrata_490222#1                    LRSGLVGFTLHGSLSHYYYQFCEALFPFQEWWVVPAKVAFDQTIWSAIWN
A.thaliana_AT4G03410.1#1             LRSGLVGFTLHGSLSHYYYQFCEALFPFQEWWVVPAKVAFDQTVWSAIWN
G.max_Glyma13g37540.1#1              LRSGLVGFTLHGSLSHYYYQLCEALFPFQEWWVVPAKVAFDQTVWSAIWN
G.max_TC302765#1                     LRSGLVGFTLHGSLSHYYYQLCEALFPFQEWWVVPAKVAFDQTVWSAIWN
G.max_Glyma12g32920.1#1              LRSGLVGFTLHGSLSHYYYQLCEALFPFQEWWVVPAKVAFDQTVWSAIWN
H.annuus_TC45682#1                   FRSGLVGFTLHGSLSHYYYQLCEALIPFKDWWVVPAKIAFDQTVWSAVWN
H.paradoxus_TA2132_73304#1           FRSGLVGFTLHGSLSHYYYRLCEALFPFKDWWVVPAKIVFDQTVWSAVWN
M.domestica_TC51384#1                FRSGLVGFTFHGSLSHYYYQFCEALFPLEDWWVVPAKIAFDQTVWSAIWN
V.vinifera_GSVIVT00025998001#1        LRSGLVGFSLHGSLSHYYYQFCEALFPSKDWWVVPAKVVVDQTVWAAIWN
                                     :***:.** :******:**  ** *:* . *****.*: .*** *:* **

L.esculentum_L450_PM22               SIYFTVLGILRLESPLAIFSELKATFLPMLTAGWKLWPFAHLITYGVIPV
S.tuberosum_TC167881#1               SIYFTVLGILRFESPLAIFSELKATFLPMLTAGWKLWPFAHLITYGVIPV
N.benthamiana_TC11792#1              SIYYTVLGVLRFESPLGIFSELKATFWPMLTAGWKLWPFAHLITYGVVPV
N.tabacum_TC57360#1                  SIYYTVLGVLRFESPLAIFSELKATFWPMLTAGWKLWPFAHLITYGVVPV
I.nil_TC25#1                         SIYFTVLGFLRLESPTSIFSELKATFWPMLTAGWKLWPFAHLITYGVIPV
A.lyrata_474411#1                    SIYFTVLGFLRFESPLSIFKELKATFLPMLTAGWKLWPFAHLITYGLVPV
A.thaliana_AT1G52870.2#1             SIYFTVLGFLRFESPISIFKELKATFLPMLTAGWKLWPFAHLITYGLVPV
B.napus_BN06MC04723_42271568@4       SIYFTVLGFLRLESPLSIFKELKATFLPMLTAGWKLWPFAHLITYGLVPV
G.max_Glyma12g34950.1#1              SIYYTVVALLRRDPPMSILNELKATFFPMLTAGWKLWPFAHLITYGVIPV
G.max_GM06MC03382_49802960@335       SIYYTVVALLRRDPPMSILNELKATFFPMLTAGWKLWPFAHLITYGVIPV
G.max_Glyma13g35620.1#1              SIYYTVVALLRLDPPMSILNELKATFFPMLTAGWKLWPFAHLITYGVIPV
P.vulgaris_TC9799#1                  SIYYTVVALLRLDPPISILNELKATFFPMLTAGWKLWPFAHLITYGVIPV
G.raimondii_TC2516#1                 SIYFVVLGFLRLESPISIFNEWKATFLPMLTAGWKLWPFAHLITYGFIPV
G.hirsutum_TC165523#1                SIYFVVLGFLRLESPISIFNEWKATFLPMLTAGWKLWPFAHLITYGFIPV
T.cacao_TC9688#1                     SIYFTVLGFLRLESPISIFNELKATFWPMLTAGWKLWPFAHLITYGVIPV
P.trifoliata_TA5129_37690#1          SIYFTVLGLLRLESPFSIFSELKATFWPMLTAGWKLWPFAHLVTYGLVPV
```

Fig. 7C

```
P.persica_TC12602#1                 SIYFTALGFLRFESPIDIFSELKATFWPMLTAGWKLWPIAHLGYLWGEPL
L.sativa_TC17162#1                  SIYYMVIGLLRFESPVTIFSELKSTFWPMLTAGWKLWPFAHIITYGVVPV
C.intybus_TA545_13427#1             SIYYMVIGLLRFESPVTIFNELRATFWPMLTAGWKLWPFAHIITYGVVPV
H.argophyllus_TA2174_73275#1        SIYFVVVALLRFESPLTIFNELKATFFPMLTAGWKLWPFAHIITYGVVPI
Aquilegia_sp_TC27124#1              SIYYVVLGLLRFESPASIFTELKATFIPMLTAGWKLWPFAHLITYGVVPV
V.vinifera_GSVIVT00027620001#1      SIYYTVVGFLRFDSPANVFGELRATFWPMLTAGWKLWPFAHLITYGVIPV
Z.mays_ZM07MC32543_BFb0296A02@      SIYFVVLGFLRLESPTTIYNELKSTFWPMLTAGWKLWPFAHLVTYGVVPV
Zea_mays_GRMZM2G011269_T01#1        SIYFVVLGFLRLESPTTIYSELKSTFWPMLTAGWKLWPFAHLITYGVVPV
S.bicolor_Sb01g015680.1#1           SIYFVVLGFLRLESPTTIYGELKSTFWPMLTAGWKLWPFAHLVTYGVVPV
P.virgatum_TC39638#1                SIYFVVLGFLRLESPTTIYKELKSTFWPMLTAGWKLWPFAHLITYGVIPV
O.sativa_LOC_Os03g38730.1#1         SIYFVVLGFLRLESPATISSELKSTFWPMLTAGWKLWPFAHLVTYGLVPV
P.patens_TC32628#1                  SIYFITLGLLRFESPVRILKDLRETFFPLLTAGWKLWPFAHLITYGLVPV
Triphysaria_sp_TC3527#1             SIYFTLLGVLRLESPVNIFNKWKATFFPMLTAGWKLWPFAHLITYGVIPV
A.lyrata_490222#1                   SIYFTVLGLLRFQSPADIFSEIKTTFWPMLTAGWKLWPLAHLVTYGVIPV
A.thaliana_AT4G03410.1#1            SIYFTVLGLLRFQSPADIFSEIKTTFLPMLTAGWKLWPLAHLVTYGVIPV
G.max_Glyma13g37540.1#1             SIYFVVLGLLRFESLTNIYGELKSTFLPLLTAGWKLWPFAHLITYGVIPV
G.max_TC302765#1                    SIYFVVLGLLRFESLTNIYGELKSTFLPLLTAGWKLWPFAHLITYGVIPV
G.max_Glyma12g32920.1#1             SIYFVVLGLLRFESLTNIYGELKSTFLPLLTAGWKLWPFAHLITYGVIPV
H.annuus_TC45682#1                  SIYFDVLGLLRFESPANIFNELKTTFLPLLTAGWKLWPFAHLVTYGLIPV
H.paradoxus_TA2132_73304#1          SIYFVVLGLLRFESPSNIFNELKTTFLPLLTAGWKLWPFAHLVTYGLIPV
M.domestica_TC51384#1               SIYFVVLGFLRFESPTKIFDELEGYICANADSGMETLAICSCG-------
V.vinifera_GSVIVT00025998001#1      SIYYVALGFLRRESPANIYGEVKSTFWPMLTAGWKLWPFAHLITYGVIPV
                                    ***:  :..** :.   :  . .  : .   :* :  .:.

L.esculentum_L450_PM22              EQRLLWVDCVELVWVTILSTYSNEKSESRVSEVIVEAEAQ-PPSVSPPQE
S.tuberosum_TC167881#1              EQRLLWVDCVELIWVTILSTYSNEKSESRVSEVIVEAEAQ-PPSISPPQE
N.benthamiana_TC11792#1             EQRLLWVDCVELIWVTILSTYSNEKSEASL--------------------
```

Fig. 7C

```
N.tabacum_TC57360#1                    EQRLLWVDCVELIWVTILSTYSNEKSEARVSEASVEAEMQ-PPSIGPPQE
I.nil_TC25#1                           EQRLLWVDCVELIWVTILSTYSNEKSEARISEESVEATSN-PPSVGPSQK
A.lyrata_474411#1                      EQRLLWVDCVELIWVTILSTYSNEKSEARISESVIETSSSSTTSIDPSKE
A.thaliana_AT1G52870.2#1               EQRLLWVDCVELIWVTILSTYSNEKSEARISESVIETSSSSTTTIDPSKE
B.napus_BN06MC04723_42271568@4         EQRLLWVDCVELIWVTILSTYSNEKSEARISESVIENPSSSTTPIDPSKE
G.max_Glyma12g34950.1#1                EQRLLWVDTIELIWVTILSTFSNEKSEARNSQSMVPSEVKSTS-VYPPEE
G.max_GM06MC03382_49802960@335         EQRLLWVDTIELIWVTILSTFSNEKSEARNSQSMVPSEVKSTS-VYPPEE
G.max_Glyma13g35620.1#1                EQRLLWVDTIELIWVTILSTFSNEKSEARNSQSMVPSEVKSTTYVHPPEE
P.vulgaris_TC9799#1                    EQRLLWVDTVELVWVTILSTFSNEKSEAKSSESMMPTEVISTTSVQPPEE
G.raimondii_TC2516#1                   EQRLLWVDCVELIWVTILSTYSNEKSEARISEAAPAEASSILPPVGPPEL
G.hirsutum_TC165523#1                  EQRLLWVDCVELIWVTILSTYSNEKSEARISEAAPAEASSILPPVGPPEE
T.cacao_TC9688#1                       EQRLLWVDCVELIWVTILSTYSNEKSEARIAEA-PAEANSSLPPVGPSEE
P.trifoliata_TA5129_37690#1            EQRLLWVDCVELIWVTILSTYSNEKSEARIAEA-PAEVKPCLPDISPPEE
P.persica_TC12602#1                    SNK-------GLIGWTV----------------------------------
L.sativa_TC17162#1                     EQRLLWVDCVELIWVTILSTLSNEKSEARVLDAPIDQSSS--SSFEPSEE
C.intybus_TA545_13427#1                EQRLLWVDCVELIWVTILSTLSNEKSEARIIDVPADTNQS--SSLEHLEE
H.argophyllus_TA2174_73275#1           EQRLLWVDCVELIWVTILSTLSNEKSEARISDAPTDPNTI--SDFEPSEE
Aquilegia_sp_TC27124#1                 EQRLLWVDCVELIWVTILSTLSNEKSEARTSEASLEVNSS--TLSINAEE
V.vinifera_GSVIVT00027620001#1         EQRLLWVDCVELIWVTILSTYSNEKSEARVSEASAEAASN--SPPTGSPE
Z.mays_ZM07MC32543_BFb0296A02@         EQRLLWVDCVELVWVTILSTYSNEKSEARNSDSTSTPAAS-KDNST----
Zea_mays_GRMZM2G011269_T01#1           EQRLLWVDCVELVWVTILSTYSNEKSEARNSDSTSTPDAS-KDNSR----
S.bicolor_Sb01g015680.1#1              EQRLLWVDCVELVWVTILSTYSNEKSEARNSDSTSTPAASSKDNSR----
P.virgatum_TC39638#1                   EQRLLWVDCVELVWVTILSTYSNEKSEARNSDGTSTPDTS-KDNSR----
O.sativa_LOC_Os03g38730.1#1            EQRLLWVDCVELIWVTILSTYSNEKSEARNSEDASTSNAS-NDNSI----
P.patens_TC32628#1                     EQRLLWVDCVEILWVTILSVFANEKAQQRMEISEVETVAVLSSGDAVEQK
Triphysaria_sp_TC3527#1                EQRLLWVDCVELIWVTILSTYSNEKSEARISEAPVEANASLPP-----PE
A.lyrata_490222#1                      DQRLLWVDCIELIWVTILSTYSNEKAEAQASGETNSSSHSSEVCQLLERL
```

Fig. 7C

```
A.thaliana_AT4G03410.1#1            DQRLLWVDCIELIWVTILSTYSNEKAEAQASEETNSSSHSSED-------
G.max_Glyma13g37540.1#1             EQRLLWVDCVELIWVTILSTYSNEKSEARISEAASETGSSTSSENSKDIY
G.max_TC302765#1                    EQRLLWVDCVELIWVTILSTYSNEKSEARISEAASETGSSTSSENSKE--
G.max_Glyma12g32920.1#1             EQRLLWVDCVELIWVTILSTYSNEKSEARISEAASETGSSTSSENSKE--
H.annuus_TC45682#1                  EQRLLWVDCVELVWVTILSTYSNEKSETRISDTSSEVDPNSPSKKH----
H.paradoxus_TA2132_73304#1          EQRLLWVDCVELVWVTILSTYSNEKSETRISDTSSEVDPNSPSKKH----
M.domestica_TC51384#1               ---YVWPDPC----------------------------------------
V.vinifera_GSVIVT00025998001#1      EQRLLWVDCVELIWVTILSTYSNEKSEARITEATQEANSRSSNNAEVSPV

L.esculentum_L450_PM22              ----------------------------------------------------
S.tuberosum_TC167881#1              ----------------------------------------------------
N.benthamiana_TC11792#1             ----------------------------------------------------
N.tabacum_TC57360#1                 ----------------------------------------------------
I.nil_TC25#1                        ----------------------------------------------------
A.lyrata_474411#1                   ----------------------------------------------------
A.thaliana_AT1G52870.2#1            ----------------------------------------------------
B.napus_BN06MC04723_42271568@4      ----------------------------------------------------
G.max_Glyma12g34950.1#1             ----------------------------------------------------
G.max_GM06MC03382_49802960@335      ----------------------------------------------------
G.max_Glyma13g35620.1#1             ----------------------------------------------------
P.vulgaris_TC9799#1                 ----------------------------------------------------
G.raimondii_TC2516#1                E---------------------------------------------------
G.hirsutum_TC165523#1               ----------------------------------------------------
T.cacao_TC9688#1                    QVMR------------------------------------------------
P.trifoliata_TA5129_37690#1         ----------------------------------------------------
P.persica_TC12602#1                 ----------------------------------------------------
```

Fig. 7C

```
L.sativa_TC17162#1                  --------------------------------------------------
C.intybus_TA545_13427#1             --------------------------------------------------
H.argophyllus_TA2174_73275#1        --------------------------------------------------
Aquilegia_sp_TC27124#1              --------------------------------------------------
V.vinifera_GSVIVT00027620001#1      VCQLNVLF----------------------------------------
Z.mays_ZM07MC32543_BFb0296A02@      --------------------------------------------------
Zea_mays_GRMZM2G011269_T01#1        --------------------------------------------------
S.bicolor_Sb01g015680.1#1           --------------------------------------------------
P.virgatum_TC39638#1                --------------------------------------------------
O.sativa_LOC_Os03g38730.1#1         --------------------------------------------------
P.patens_TC32628#1                  LPIEDSQ-----------------------------------------
Triphysaria_sp_TC3527#1             GPFE-------------------------------------------
A.lyrata_490222#1                   NVLPVCSSCACKYLTGLGDVSM--------------------------
A.thaliana_AT4G03410.1#1            --------------------------------------------------
G.max_Glyma13g37540.1#1             TYTLHRSCILLAKINYSAIAAVKIGTLHSTLCSAMAKIEELNESIPKSKN
G.max_TC302765#1                    --------------------------------------------------
G.max_Glyma12g32920.1#1             --------------------------------------------------
H.annuus_TC45682#1                  --------------------------------------------------
H.paradoxus_TA2132_73304#1          --------------------------------------------------
M.domestica_TC51384#1               --------------------------------------------------
V.vinifera_GSVIVT00025998001#1      SEFRKQHHQALSIEHHAS------------------------------

L.esculentum_L450_PM22              ----------------------------------
S.tuberosum_TC167881#1              ----------------------------------
N.benthamiana_TC11792#1             ----------------------------------
N.tabacum_TC57360#1                 ----------------------------------
```

Fig. 7C

```
I.nil_TC25#1                                  ----------
A.lyrata_474411#1                             ----------
A.thaliana_AT1G52870.2#1                      ----------
B.napus_BN06MC04723_42271568@4                ----------
G.max_Glyma12g34950.1#1                       ----------
G.max_GM06MC03382_49802960@335                ----------
G.max_Glyma13g35620.1#1                       ----------
P.vulgaris_TC9799#1                           ----------
G.raimondii_TC2516#1                          ----------
G.hirsutum_TC165523#1                         ----------
T.cacao_TC9688#1                              ----------
P.trifoliata_TA5129_37690#1                   ----------
P.persica_TC12602#1                           ----------
L.sativa_TC17162#1                            ----------
C.intybus_TA545_13427#1                       ----------
H.argophyllus_TA2174_73275#1                  ----------
Aquilegia_sp_TC27124#1                        ----------
V.vinifera_GSVIVT00027620001#1                ----------
Z.mays_ZM07MC32543_BFb0296A02@                ----------
Zea_mays_GRMZM2G011269_T01#1                  ----------
S.bicolor_Sb01g015680.1#1                     ----------
P.virgatum_TC39638#1                          ----------
O.sativa_LOC_Os03g38730.1#1                   ----------
P.patens_TC32628#1                            ----------
Triphysaria_sp_TC3527#1                       ----------
A.lyrata_490222#1                             ----------
A.thaliana_AT4G03410.1#1                      ----------
```

Fig. 7C

```
G.max_Glyma13g37540.1#1          KRNHGSTTFFVFVDYLFILFFLGFLCFTIVFKILSI
G.max_TC302765#1                 ------------------------------------
G.max_Glyma12g32920.1#1          ------------------------------------
H.annuus_TC45682#1               ------------------------------------
H.paradoxus_TA2132_73304#1       ------------------------------------
M.domestica_TC51384#1            ------------------------------------
V.vinifera_GSVIVT00025998001#1   ------------------------------------
```

Fig. 7C

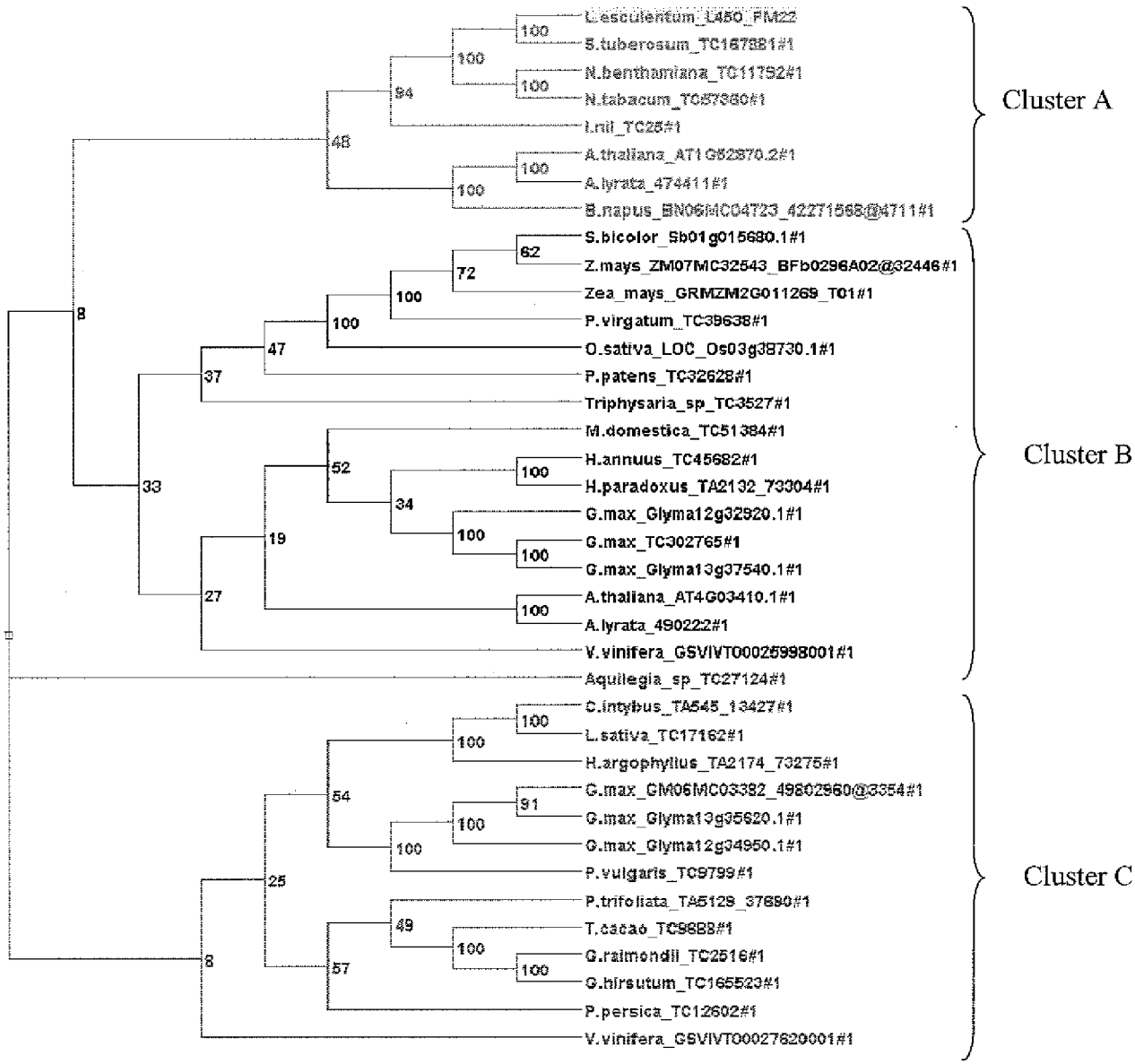
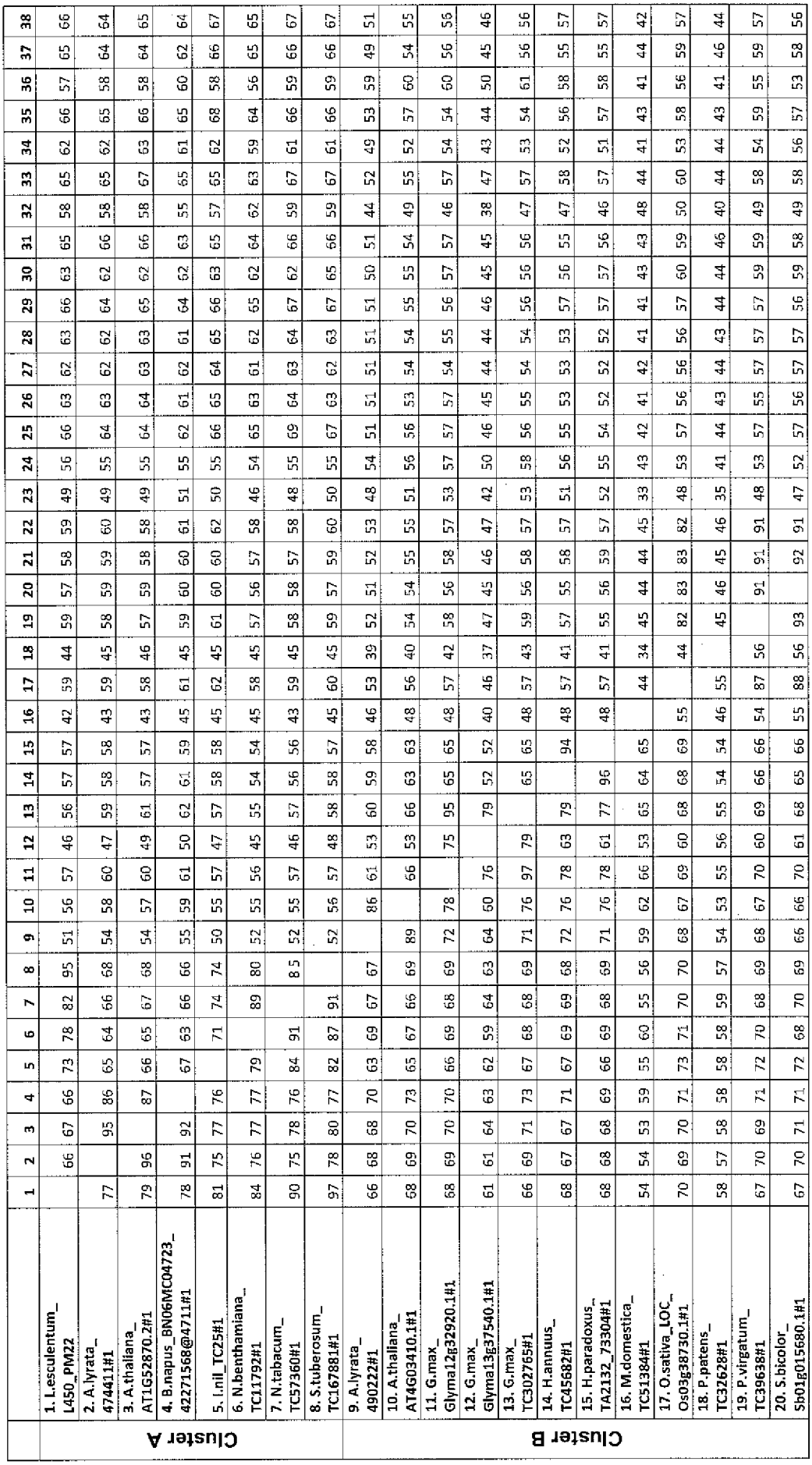

L.esculentum_L450_FM22
S.tuberosum_TC167881#1
N.benthamiana_TC11752#1
N.tabacum_TC57360#1
I.nil_TC25#1
A.thaliana_AT1G52870.2#1
A.lyrata_474411#1
B.napus_BN06MC04723_42271568@4711#1
Cluster A
S.bicolor_Sb01g015680.1#1
Z.mays_ZM07MC32543_BFb0296A02@32446#1
Zea_mays_GRMZM2G011269_T01#1
P.virgatum_TC39638#1
O.sativa_LOC_Os03g38730.1#1
P.patens_TC32628#1
Triphysaria_sp_TC3527#1
M.domestica_TC51384#1
H.annuus_TC45682#1
H.paradoxus_TA2132_73304#1
G.max_Glyma12g32920.1#1
G.max_TC302765#1
G.max_Glyma13g37540.1#1
A.thaliana_AT4G03410.1#1
A.lyrata_490222#1
V.vinifera_GSVIVT00025998001#1
Cluster B
Aquilegia_sp_TC27124#1
C.intybus_TA545_13427#1
L.sativa_TC17162#1
H.argophyllus_TA2174_73275#1
G.max_GM06MC03382_49802960@3354#1
G.max_Glyma13g35620.1#1
G.max_Glyma12g34950.1#1
P.vulgaris_TC9799#1
P.trifoliata_TA5128_37690#1
T.cacao_TC9888#1
G.raimondii_TC2516#1
G.hirsutum_TC165523#1
P.persica_TC12602#1
V.vinifera_GSVIVT00027620001#1
Cluster C

Fig. 8



		1	2	3	4	5	6	7	8	9	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
Cluster A	1. L.esculentum_ L450_PM22		66	67	66	73	78	82	95	51	56	57	46	56	57	57	42	59	44	59	57	58	59	49	56	66	63	62	63	66	63	65	65	66	62	66	57	65	66
	2. A.lyrata_ 474411#1	77		95	86	65	64	66	68	54	58	60	47	59	58	58	43	59	45	58	59	59	60	49	55	64	63	62	62	64	62	66	58	65	62	65	58	64	64
	3. A.thaliana_ AT1G52870.2#1	79	96		87	66	65	67	68	54	57	60	49	61	57	57	43	58	46	57	59	58	58	49	55	64	64	63	63	65	62	66	58	67	63	66	58	64	65
	4. B.napus_BN06MC04723_ 42271568@4711#1	78	91	92		67	63	66	66	55	59	61	50	62	61	59	45	61	45	59	60	60	61	51	55	62	61	62	61	64	62	63	55	65	61	65	60	62	64
	5. I.nil_TC25#1	81	75	77	76		71	74	74	50	55	57	47	57	58	58	45	62	45	61	60	60	62	50	55	66	65	64	65	66	63	65	57	65	62	68	58	66	67
	6. N.benthamiana_ TC11792#1	84	76	77	77	79		89	80	52	55	56	45	55	54	54	45	58	45	57	56	57	58	46	54	65	63	61	62	65	62	64	62	63	59	64	56	65	65
	7. N.tabacum_ TC57360#1	90	75	78	76	84	91		85	52	55	57	46	57	56	56	43	59	45	58	58	57	58	48	55	69	64	63	64	67	62	66	59	67	61	66	59	66	67
	8. S.tuberosum_ TC167881#1	97	78	80	77	82	87	91		52	56	57	48	58	58	57	45	60	45	59	57	60	59	50	55	67	63	62	63	67	65	66	59	67	61	66	59	66	67
Cluster B	9. A.lyrata_ 490222#1	66	68	68	70	63	69	67	67		86	61	53	60	59	58	46	53	48	52	51	52	53	48	54	51	51	51	51	51	50	51	44	52	49	53	59	49	51
	10. A.thaliana_ AT4G03410.1#1	68	69	70	73	65	67	66	69	89		66	53	66	63	63	48	56	40	54	54	55	55	51	56	56	53	54	54	55	55	54	49	55	52	57	60	54	55
	11. G.max_ Glyma12g32920.1#1	68	69	70	70	66	69	68	69	72	78		75	95	65	65	48	57	42	58	56	58	57	53	57	57	57	54	55	56	57	57	46	57	54	54	60	56	56
	12. G.max_ Glyma13g37540.1#1	61	61	64	63	62	59	64	63	64	60	76		79	52	52	40	46	37	47	45	46	47	42	50	46	45	44	44	46	45	45	38	47	43	44	50	45	46
	13. G.max_ TC302765#1	66	69	71	73	67	68	68	69	71	76	97	79		65	65	48	57	43	59	56	58	57	53	58	56	55	54	54	56	56	56	47	57	53	54	61	56	56
	14. H.annuus_ TC45682#1	68	67	67	71	67	69	69	68	72	76	78	63	79		94	48	57	41	57	55	58	57	51	56	55	53	53	57	56	55	47	58	52	56	58	55	57
	15. H.paradoxus_ TA2132_73304#1	68	68	68	69	66	69	68	69	71	76	78	61	77	96		48	57	41	55	56	59	57	52	55	54	52	52	57	57	56	46	57	51	57	58	55	57
	16. M.domestica_ TC51384#1	54	54	53	59	55	60	55	56	59	62	66	53	65	64	65		44	34	45	44	44	45	33	43	42	41	42	41	41	43	43	48	44	41	43	41	44	42
	17. O.sativa_LOC_ Os03g38730.1#1	70	69	70	71	73	71	70	70	68	67	69	60	68	68	69	55		44	82	83	83	82	48	53	57	56	56	56	57	60	59	50	60	53	58	56	59	57
	18. P.patens_ TC32628#1	58	57	58	58	58	58	59	57	54	53	55	56	55	54	54	46	55		45	46	45	46	35	41	44	43	44	43	44	44	46	40	44	44	43	41	46	44
	19. P.virgatum_ TC39638#1	67	70	69	71	72	70	68	69	68	67	70	60	69	66	66	54	87	56		91	91	91	48	53	57	55	57	57	57	59	59	49	58	54	59	55	59	57
	20. S.bicolor_ Sb01g015680.1#1	67	70	71	71	72	68	70	69	66	66	70	61	68	65	66	55	88	56	93		92	91	47	52	57	56	57	57	56	59	58	49	58	56	57	53	58	56

Fig. 9

	21. Z.mays_ZM07MC32543_ BFb0296A02@32446#1	69	69	69	72	72	71	69	73	68	70	69	59	69	68	70	56	88	55	93	93		94	49	51	57	55	57	57	56	60	59	49	57	55	56	55	60	56
	22. Zea_mays_ GRMZM2G011269_T01#1	69	71	69	73	73	69	69	70	68	69	70	60	69	67	69	55	87	56	94	93	96		49	52	57	57	57	57	56	61	61	49	59	55	56	56	61	55
	23. Triphysaria_sp_ TC3527#1	53	55	55	58	55	54	53	54	57	61	61	48	60	59	61	45	55	42	54	53	55	55		46	45	45	47	44	48	49	48	37	49	46	49	61	48	48
	24. V.vinifera_ GSVIVT00025998001#1	71	69	68	68	69	68	71	70	67	68	70	67	69	68	68	56	65	55	62	65	64	64	52		56	54	53	53	56	54	54	48	57	52	58	53	54	56
Cluster C	25. Aquilegia_sp_ TC27124#1	78	74	76	73	77	73	78	78	64	67	67	64	69	68	66	55	68	58	67	69	69	69	52	71		63	63	63	64	64	67	58	68	61	65	57	66	65
	26. G.max_ Glyma12g34950.1#1	76	74	77	74	77	73	78	76	64	64	66	60	66	64	65	54	70	58	65	68	69	70	52	68	76		94	96	68	62	64	59	66	84	67	55	63	67
	27. G.max_ Glyma13g35620.1#1	74	74	76	74	76	73	76	74	64	66	65	59	65	65	64	55	69	59	68	68	70	69	54	68	76	96		97	67	61	63	57	68	82	68	54	63	66
	28. G.max_GM06MC03382_ 49802960@3354#1	75	74	77	74	77	74	76	75	65	65	66	59	65	66	64	54	69	60	68	69	70	70	52	69	76	97	99		67	61	63	59	68	82	67	55	63	67
	29. G.raimondii_ TC2516#1	77	74	76	73	77	75	79	77	66	66	67	63	66	67	66	55	69	60	69	70	69	68	54	69	77	81	81	81		63	64	60	72	64	82	57	65	99
	30. H.argophyllus_ TA2174_73275#1	76	76	75	77	74	79	77	77	68	68	70	60	69	71	71	55	74	58	73	72	72	74	55	67	75	75	75	75	74		81	52	64	63	65	56	81	63
	31. L.sativa_ TC17162#1	77	78	79	77	76	79	78	79	65	68	69	61	69	67	68	56	71	57	70	71	71	73	56	68	77	77	75	76	76	91		55	64	62	64	57	92	65
	32. P.persica_ TC12602#1	67	71	71	69	67	73	70	69	64	65	60	52	62	63	61	61	63	50	62	61	63	63	47	61	68	70	69	70	70	65	67		61	56	61	48	56	59
	33. P.trifoliata_ TA5129_37690#1	77	77	80	78	77	76	78	78	67	67	69	62	67	68	67	56	71	56	70	72	70	70	55	72	76	81	82	83	81	74	76	71		64	73	58	64	72
	34. P.vulgaris_ TC9799#1	75	74	78	75	76	73	77	74	63	63	65	59	64	63	62	52	68	59	66	69	68	68	54	66	75	92	90	91	78	75	74	69	79		65	53	61	64
	35. T.cacao_ TC9688#1	76	75	77	74	77	74	77	76	68	67	65	62	65	66	66	53	70	57	68	70	68	69	55	71	76	78	79	79	88	76	76	68	81	77		58	65	81
	36. V.vinifera_ GSVIVT00027620001#1	65	68	67	68	65	66	66	66	70	73	72	59	72	72	73	59	64	50	64	64	65	65	69	64	63	63	63	63	65	64	67	61	65	62	65		57	57
	37. C.intybus_ TA545_13427#1	78	76	78	77	76	79	78	79	67	68	69	61	68	68	68	56	71	58	70	71	72	72	54	68	76	76	75	76	76	90	96	67	77	74	77	64		66
	38. G.hirsutum_ TC165523#1	77	74	76	73	77	75	79	77	66	67	66	62	66	66	66	55	69	58	69	70	70	68	54	69	77	80	80	81	99	75	76	70	81	78	88	65	77	

Fig. 9

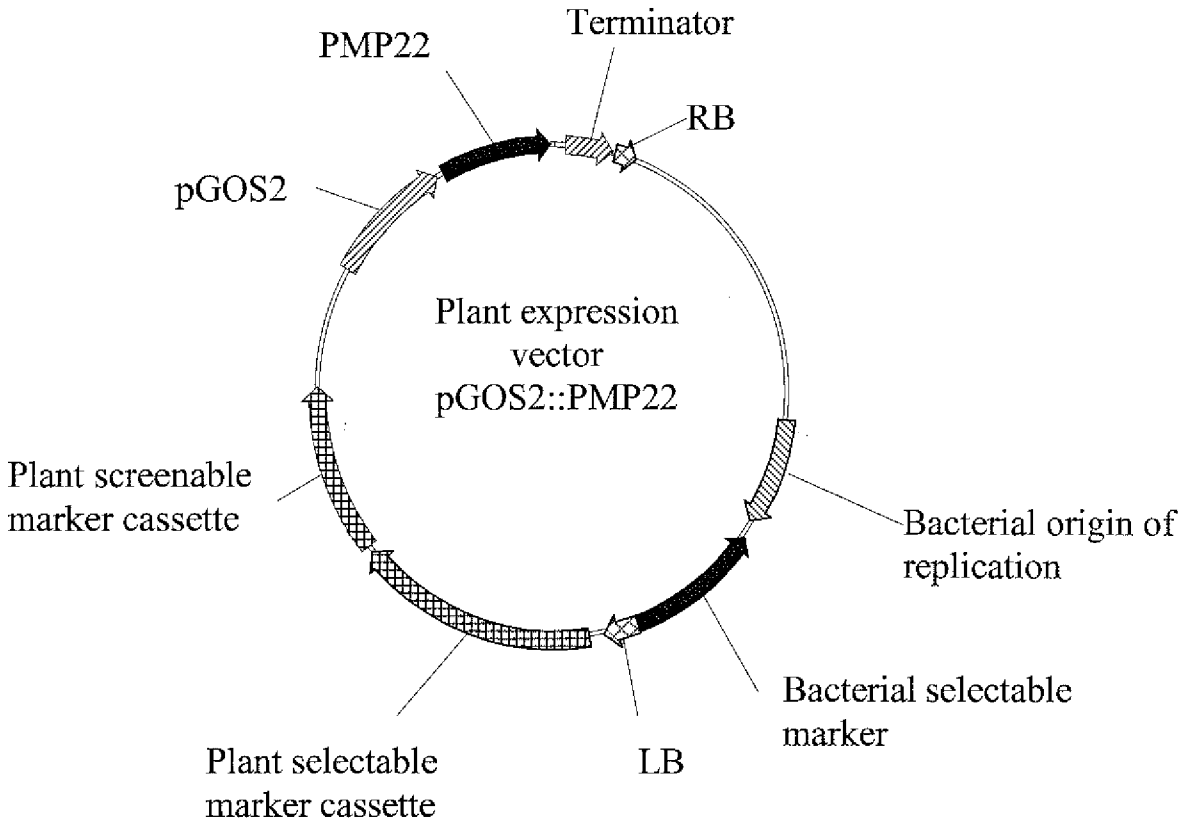

PMP22
Terminator
RB
pGOS2
Plant expression vector pGOS2::PMP22
Plant screenable marker cassette
Bacterial origin of replication
Bacterial selectable marker
Plant selectable marker cassette
LB

Fig. 10

MANPLLYSPINPQ**FFQPLLPGFTNHLDIPVAFFLKYLVGTNVGKTAELRSDASEMTWKVKIDGRRLSN**

--1--

GWEDFTIAHDLRVGDIVVFRQEGELVFHVTALGPSCCEIQYGEDTLEEDKIEKLCGTENVSSKKKSLK

-----2-----

REAESAPDNSLDSC**FVATVTGSNLKRDTLYIPKEFALSNGLMNKYQIVLMNEEGESWKIDLRREAYNY**

GRFYMRRGWRSFCIANGKKPGDVFAFKLVKNEETPMIQLFPMTIEDLDKLQSLPRHKIRKTEAAPSSP

DLSS**FVATVTASNLSRDRLYLPKTFIMSNGLLKKFQMCLMNEEGESWTIDVKHEAHTGRFLTIRGWRR**

FCVANGKKPGDLLKFKLVHNEETPVLQLLPLNSEDLHKLNPSNDTRHGQSLKVTKKEFLGMEATENEF

LGEEVYCNDSFKASEKDTLPFAEPINEDIRQGQCSQTIKQEYVSTEEKNSTSQNR**FVTLTLTHSSKLN**

LPFEFMKRNGIKKAGKITMVDRYDAKWRTSLLMDKIGTMSLGRGSKGFCEVNGVEMNESFILELIWED

TVPLLKFCSKV

FIG. 11

```
A.thaliana_AT2G24650.1#1        MANSRIYPQFFHTLVPSFHTHLMIPEDFFSEYIEGRSVAELKLDFSDKSW 50
A.thaliana_AT2G24650.2#1        --------------------------------------------------
A.thaliana_AT4G00260.1#1        --------------------------------------------------
B.napus_TC73539#1               --------------------------------------------------
A.thaliana_AT1G26680.1#1        --------------------------------------------------
B.napus_CD826203#1              --------------------------------------------------
A.thaliana_AT2G24700.1#1        --------------------------------------------------
S.lycopersicum_TC201533#1       --------------------------------------------------

A.thaliana_AT2G24650.1#1        EVKLSDRRITDGWEEFVVANDFRIGDVVAFRYVGNLVFHVSNLGPNYYEI 100
A.thaliana_AT2G24650.2#1        --------------------------------------------------
A.thaliana_AT4G00260.1#1        --------------------------------------------------
B.napus_TC73539#1               --------------------------------------------------
A.thaliana_AT1G26680.1#1        --------------------------------------------------
B.napus_CD826203#1              --------------------------------------------------
A.thaliana_AT2G24700.1#1        --------------------------------------------------
S.lycopersicum_TC201533#1       --------------------------------------------------

A.thaliana_AT2G24650.1#1        EHNDELPKEKKAKTNSEEADAVSSSSSADKSCFMAIITALDLTTDTLYLP 150
A.thaliana_AT2G24650.2#1        --------------------------------------------------
A.thaliana_AT4G00260.1#1        --------------------------------------------------
B.napus_TC73539#1               --------------------------------------------------
A.thaliana_AT1G26680.1#1        --------------------------------------------------
B.napus_CD826203#1              --------------------------------------------------
A.thaliana_AT2G24700.1#1        --------------------------------------------------
S.lycopersicum_TC201533#1       --------------------------------------------------

A.thaliana_AT2G24650.1#1        LHFTSANGLTRKNREIILTDGGERSRVLDLRFDESSGTFYISRGWRNFCD 200
A.thaliana_AT2G24650.2#1        --------------------------------------------------
A.thaliana_AT4G00260.1#1        --------------------------------------------------
B.napus_TC73539#1               --------------------------------------------------
A.thaliana_AT1G26680.1#1        --------------------------------------------------
B.napus_CD826203#1              --------------------------------------------------
A.thaliana_AT2G24700.1#1        --------------------------------------------------
S.lycopersicum_TC201533#1       --------------------------------------------------

A.thaliana_AT2G24650.1#1        ENGQKACGFFLFKLVGKGETLVLSFCPTESINGEENITREDSKDECSSLD 250
A.thaliana_AT2G24650.2#1        --------------------------------------------------
A.thaliana_AT4G00260.1#1        --------------------------------------------------
B.napus_TC73539#1               --------------------------------------------------
A.thaliana_AT1G26680.1#1        --------------------------------------------------
B.napus_CD826203#1              --------------------------------------------------
A.thaliana_AT2G24700.1#1        --------------------------------------------------
S.lycopersicum_TC201533#1       --------------------------------------------------

A.thaliana_AT2G24650.1#1        SLMNIVEKKKYIPKPRGSPYSSYSPSHKQFVTFTLPPDYARIGKLSLSAP 300
A.thaliana_AT2G24650.2#1        --------------------------------------------------
A.thaliana_AT4G00260.1#1        --------------------------------------------------
B.napus_TC73539#1               --------------------------------------------------
A.thaliana_AT1G26680.1#1        --------------------------------------------------
B.napus_CD826203#1              --------------------------------------------------
A.thaliana_AT2G24700.1#1        --------------------------------------------------
S.lycopersicum_TC201533#1       --------------------------------------------------

A.thaliana_AT2G24650.1#1        FVRENGINKPGEICLLDKHGRKWLTSLLLDSKGTMSLGKGWKEFVKANSL 350
A.thaliana_AT2G24650.2#1        --------------------------------------------------
A.thaliana_AT4G00260.1#1        --------------------------------------------------
B.napus_TC73539#1               --------------------------------------------------
A.thaliana_AT1G26680.1#1        --------------------------------------------------
B.napus_CD826203#1              --------------------------------------------------
A.thaliana_AT2G24700.1#1        --------------------------------------------------
S.lycopersicum_TC201533#1       --------------------------------------------------
```

FIG. 12

```
A.thaliana_AT2G24650.1#1        ETGFTLKLIWEETTPVLSLCSPESNSDREQEEISKAIEKHSLFIDPSNRD 400
A.thaliana_AT2G24650.2#1        --------------------------------------------------
A.thaliana_AT4G00260.1#1        --------------------------------------------------
B.napus_TC73539#1               --------------------------------------------------
A.thaliana_AT1G26680.1#1        --------------------------------------------------
B.napus_CD826203#1              --------------------------------------------------
A.thaliana_AT2G24700.1#1        --------------------------------------------------
S.lycopersicum_TC201533#1       --------------------------------------------------

A.thaliana_AT2G24650.1#1        KISNNDKEENMSWERKKDHLKSRDSTLSSQKQFVTITITPSSDRLVSLSN 450
A.thaliana_AT2G24650.2#1        --------------------------------------------------
A.thaliana_AT4G00260.1#1        --------------------------------------------------
B.napus_TC73539#1               --------------------------------------------------
A.thaliana_AT1G26680.1#1        --------------------------------------------------
B.napus_CD826203#1              --------------------------------------------------
A.thaliana_AT2G24700.1#1        --------------------------------------------------
S.lycopersicum_TC201533#1       --------------------------------------------------

A.thaliana_AT2G24650.1#1        DSCLVVVSLLYFDMRLPKVFTRENGINKPGRITLLGKDGIKQQTNLLFDK 500
A.thaliana_AT2G24650.2#1        --------------------------------------------------
A.thaliana_AT4G00260.1#1        --------------------------------------------------
B.napus_TC73539#1               --------------------------------------------------
A.thaliana_AT1G26680.1#1        --------------------------------------------------
B.napus_CD826203#1              --------------------------------------------------
A.thaliana_AT2G24700.1#1        --------------------------------------------------
S.lycopersicum_TC201533#1       --------------------------------------------------

A.thaliana_AT2G24650.1#1        ANGAMSLGHGWKDFVKDNGLKTGDSFTLKLIWEDQTPVLSLCPADCSIDR 550
A.thaliana_AT2G24650.2#1        --------------------------------------------------
A.thaliana_AT4G00260.1#1        --------------------------------------------------
B.napus_TC73539#1               --------------------------------------------------
A.thaliana_AT1G26680.1#1        --------------------------------------------------
B.napus_CD826203#1              --------------------------------------------------
A.thaliana_AT2G24700.1#1        --------------------------------------------------
S.lycopersicum_TC201533#1       --------------------------------------------------

A.thaliana_AT2G24650.1#1        EAGGGRSETNQKKSLPIEPSTCKKIRKDVNIKDDNSKEKNDKEESKSVDG 600
A.thaliana_AT2G24650.2#1        --------------------------------------------------
A.thaliana_AT4G00260.1#1        --------------------------------------------------
B.napus_TC73539#1               --------------------------------------------------
A.thaliana_AT1G26680.1#1        --------------------------------------------------
B.napus_CD826203#1              --------------------------------------------------
A.thaliana_AT2G24700.1#1        --------------------------------------------------
S.lycopersicum_TC201533#1       --------------------------------------------------

A.thaliana_AT2G24650.1#1        ERKYLRGTYLTPSSQKHFVTLTITPSSIKKDRLILSPQFARKNNIDKPGM 650
A.thaliana_AT2G24650.2#1        --------------------------------------------------
A.thaliana_AT4G00260.1#1        --------------------------------------------------
B.napus_TC73539#1               --------------------------------------------------
A.thaliana_AT1G26680.1#1        --------------------------------------------------
B.napus_CD826203#1              --------------------------------------------------
A.thaliana_AT2G24700.1#1        --------------------------------------------------
S.lycopersicum_TC201533#1       --------------------------------------------------

A.thaliana_AT2G24650.1#1        IYLLDTDGTKWLISLQRDKKGTMSLGKGWKEFAEANDFKLGESFTMELVW 700
A.thaliana_AT2G24650.2#1        --------------------------------------------------
A.thaliana_AT4G00260.1#1        --------------------------------------------------
B.napus_TC73539#1               --------------------------------------------------
A.thaliana_AT1G26680.1#1        --------------------------------------------------
B.napus_CD826203#1              --------------------------------------------------
A.thaliana_AT2G24700.1#1        --------------------------------------------------
S.lycopersicum_TC201533#1       --------------------------------------------------
```

FIG. 12 (Cont)

```
A.thaliana_AT2G24650.1#1        EDTTPMLSLLRTEFRSSKANEKESISSEHKTRESSPTIKNRIVTPALTPE 750
A.thaliana_AT2G24650.2#1        --------------------------------------------------
A.thaliana_AT4G00260.1#1        --------------------------------------------------
B.napus_TC73539#1               --------------------------------------------------
A.thaliana_AT1G26680.1#1        --------------------------------------------------
B.napus_CD826203#1              --------------------------------------------------
A.thaliana_AT2G24700.1#1        --------------------------------------------------
S.lycopersicum_TC201533#1       --------------------------------------------------

A.thaliana_AT2G24650.1#1        DVKACKLILPSQFMKKIRTVDKERNHLKGRDLNPSCQKQFILSAQFAREN 800
A.thaliana_AT2G24650.2#1        --------------------------------------------------
A.thaliana_AT4G00260.1#1        --------------------------------------------------
B.napus_TC73539#1               --------------------------------------------------
A.thaliana_AT1G26680.1#1        --------------------------------------------------
B.napus_CD826203#1              --------------------------------------------------
A.thaliana_AT2G24700.1#1        --------------------------------------------------
S.lycopersicum_TC201533#1       --------------------------------------------------

A.thaliana_AT2G24650.1#1        NINQPGTIYLLDTDGRKWLTTVKRDKKGTMSLGKGWKEFADTKDLKSGDS 850
A.thaliana_AT2G24650.2#1        --------------------------------------------------
A.thaliana_AT4G00260.1#1        --------------------------------------------------
B.napus_TC73539#1               --------------------------------------------------
A.thaliana_AT1G26680.1#1        --------------------------------------------------
B.napus_CD826203#1              --------------------------------------------------
A.thaliana_AT2G24700.1#1        --------------------------------------------------
S.lycopersicum_TC201533#1       --------------------------------------------------

A.thaliana_AT2G24650.1#1        FTMELIWEDTNPVLSLLRTKFSSSKSNKEESIFLEPKSRDSSSPTIVNRF 900
A.thaliana_AT2G24650.2#1        --------------------------------------------------
A.thaliana_AT4G00260.1#1        --------------------------------------------------
B.napus_TC73539#1               --------------------------------------------------
A.thaliana_AT1G26680.1#1        --------------------------------------------------
B.napus_CD826203#1              --------------------------------------------------
A.thaliana_AT2G24700.1#1        --------------------------------------------------
S.lycopersicum_TC201533#1       --------------------------------------------------

A.thaliana_AT2G24650.1#1        VTLALTPEDVTACKLILPSQFMKANGINNKLGKITLLGENGVEWPGYMLS 950
A.thaliana_AT2G24650.2#1        ---------------MANQHFFKP-------------------------- 9
A.thaliana_AT4G00260.1#1        ---------------MAHQHFFKP-------------------------- 9
B.napus_TC73539#1               ---------------MANQHFFKP-------------------------- 9
A.thaliana_AT1G26680.1#1        ------MADQSLLHSPINPHFFQP-------------------------- 18
B.napus_CD826203#1              ------MADDAVLQSPTNPHFFQP-------------------------- 18
A.thaliana_AT2G24700.1#1        ------MAN-PLLYSPINPQFFQP-------------------------- 17
S.lycopersicum_TC201533#1       -------MQSSSPELPISRRFFKV-------------------------- 17
                                             :*::

A.thaliana_AT2G24650.1#1        LDGTLALGNGWEGFCEANGVKLGQTFTLEFVNEQDTTTTPRIPVAFFLKN 1000
A.thaliana_AT2G24650.2#1        ----LLPG-----------------------------FHTHLRIPVAFFLKN 28
A.thaliana_AT4G00260.1#1        ----LLPG-----------------------------FHASLTIPVAFFLKY 28
B.napus_TC73539#1               ----LLPG-----------------------------FHSHLKIPVAFFSKH 28
A.thaliana_AT1G26680.1#1        ----ILTE-----------------------------SRTHLNIPVAFFSKH 37
B.napus_CD826203#1              ----LLPG-----------------------------YHSHMNVPVAFFSKH 37
A.thaliana_AT2G24700.1#1        ----LLPG-----------------------------FTNHLDIPVAFFLKY 36
S.lycopersicum_TC201533#1       ----MVPG-----------------------------FHSKLTIPPAFFPK- 35
                                    :                                    :* *** *

A.thaliana_AT2G24650.1#1        IEGRYEQK-TAELRSDASKITWEVKID-GQRLTDGWKEFALSHDLRIGDI 1048
A.thaliana_AT2G24650.2#1        IEGRYEQK-TAELRSDASKITWEVKID-GQRLTDGWKEFALSHDLRIGDI 76
A.thaliana_AT4G00260.1#1        IEGRYEQK-TAKLRSDASKRTWEVKID-GQRLTDGWKEFAVSHDLRIGDI 76
B.napus_TC73539#1               IEGRNEHKNTAKLRSDTSEITWKVKIEDGLRLTEGWKEFALAHDLRVGDI 78
A.thaliana_AT1G26680.1#1        VEGRNNQNKTVTLRSDASDKTWLVKMD-GLKLTDGWEDFAFAHDLRTGDI 86
B.napus_CD826203#1              VQGRNG-DKTARLRSDASDTTWEVVIN-GRRLTGGWKEFVKAHDLRVGDV 85
A.thaliana_AT2G24700.1#1        LVGTNV-GKTAELRSDASEMTWKVKID-GRRLSNGWEDFTIAHDLRVGDI 84
S.lycopersicum_TC201533#1       LKGDNLEKGTIIKGKDIWNVEINRSEKGTISFDKGWEEFVQNHDLRVGDF 85
                                : *      *    .*  .       .    :  **::*.  **** **.
```

FIG. 12 (Cont)

```
A.thaliana_AT2G24650.1#1        VVFRQERDMSFHVTMLGPSCCEIQYGSCSDEERNLEKKK-----N----- 1088
A.thaliana_AT2G24650.2#1        VVFRQERDMSFHVTMLGPSCCEIQYGSCSDEERNLEKKK-----N----- 116
A.thaliana_AT4G00260.1#1        VVFRQESDLAFHVTLLGPSCCGIQYGSCSVEKNNLGDEKKKVKEN----- 121
B.napus_TC73539#1               VVFRQENDMAFHVTLFGPSCCEIQYGSCLDDKNKLVTIQ---------- 117
A.thaliana_AT1G26680.1#1        VVFRLEGEMVFHVTALGPSCCEIQYHTSSHNINDDDRNDQINIASRNSSR 136
B.napus_CD826203#1              LVFRHEGELVFHVTALGSSCCEVEYAT----LDDGDDNKETVLAS----- 126
A.thaliana_AT2G24700.1#1        VVFRQEGELVFHVTALGPSCCEIQYGEDTLEEDKIEKLCGTENVS----- 129
S.lycopersicum_TC201533#1       AVFEHLGDMRFSVTLLDSTGCDKKLLEKSEVVPSQEKKKKVKSAQP---- 131
                                 **.   :: * ** :.: *  :          .

A.thaliana_AT2G24650.1#1        --------------------------------------------------
A.thaliana_AT2G24650.2#1        --------------------------------------------------
A.thaliana_AT4G00260.1#1        --------------------------------------------------
B.napus_TC73539#1               --------------------------------------------------
A.thaliana_AT1G26680.1#1        VKKNPRKKVESSLDHSRFVAKVSAWCLSNDRLYIPLSFARLNGLNKINSK 186
B.napus_CD826203#1              --------------------------------------------------
A.thaliana_AT2G24700.1#1        --------------------------------------------------
S.lycopersicum_TC201533#1       --------------------------------------------------

A.thaliana_AT2G24650.1#1        --------------------------------------------------
A.thaliana_AT2G24650.2#1        --------------------------------------------------
A.thaliana_AT4G00260.1#1        --------------------------------------------------
B.napus_TC73539#1               --------------------------------------------------
A.thaliana_AT1G26680.1#1        KIYLQNEEGRSWKLVLRHDKSGMQTFVQSGWRRFCSENGIRQGQYTFKLV 236
B.napus_CD826203#1              --------------------------------------------------
A.thaliana_AT2G24700.1#1        --------------------------------------------------
S.lycopersicum_TC201533#1       --------------------------------------------------

A.thaliana_AT2G24650.1#1        --------------------------------------------------
A.thaliana_AT2G24650.2#1        --------------------------------------------------
A.thaliana_AT4G00260.1#1        --------------------------------------------------
B.napus_TC73539#1               --------------------------------------------------
A.thaliana_AT1G26680.1#1        RKSAPPVIRLCRAKAKPKQRSVAEYSSDHSCFEGSVTPSSLRNDLLYLPR 286
B.napus_CD826203#1              -------------DIKP---------------------------------- 130
A.thaliana_AT2G24700.1#1        ------------SKKKS---------------------------------- 134
S.lycopersicum_TC201533#1       --------------------------------------------------

A.thaliana_AT2G24650.1#1        --------------------------------------------------
A.thaliana_AT2G24650.2#1        --------------------------------------------------
A.thaliana_AT4G00260.1#1        --------------------------------------------------
B.napus_TC73539#1               --------------------------------------------------
A.thaliana_AT1G26680.1#1        SFVNSNRLDKRCSEIVLKNEQGVKWPLVLKRFKSVTYLPKGWTSFCQVNR 336
B.napus_CD826203#1              --------------------------------------------------
A.thaliana_AT2G24700.1#1        --------------------------------------------------
S.lycopersicum_TC201533#1       --------------------------------------------------

A.thaliana_AT2G24650.1#1        --------------------------------------------------
A.thaliana_AT2G24650.2#1        --------------------------------------------------
A.thaliana_AT4G00260.1#1        --------------------------------------------------
B.napus_TC73539#1               --------------------------------------------------
A.thaliana_AT1G26680.1#1        IKAGDSFKFKLVGTWKKPVLSLCPTQSNNHKTPLECSEGNKSEESEEDCL 386
B.napus_CD826203#1              --------------------------------------------------
A.thaliana_AT2G24700.1#1        --------------------------------------------------
S.lycopersicum_TC201533#1       --------------------------------------------------

A.thaliana_AT2G24650.1#1        ----------------------------------PNGEAKSSSLDPSC 1102
A.thaliana_AT2G24650.2#1        ----------------------------------PNGEAKSSSLDPSC 130
A.thaliana_AT4G00260.1#1        ----------------------------------PNGEAESSSRDPSC 135
B.napus_TC73539#1               --------------------------------------------------
A.thaliana_AT1G26680.1#1        EVKKKKYWSRCRASVENMDDDQTNIGNSSRKKRVSKNPREKVESSSSDHSS 436
B.napus_CD826203#1              ----------------------------MKKRVKKNPR---------- 140
A.thaliana_AT2G24700.1#1        ----------------------------LKREAESAPDNSLDS-----C 150
S.lycopersicum_TC201533#1       --------------------------------------------------
```

FIG. 12 (Cont)

```
A.thaliana_AT2G24650.1#1        FSANVAPSSLRYDLMLFPMGFVRENGVVGSGKIVLMNEKGRSWNFNLRQK 1152
A.thaliana_AT2G24650.2#1        FSANVAPSSLRYDLMLFPMGFVRENGVVGSGKIVLMNEKGRSWNFNLRQK 180
A.thaliana_AT4G00260.1#1        FVANVAPSSLRYDLMRFPRGFVRDNGVVGSGEIVLMNEKGRSWNFNLRQK 185
B.napus_TC73539#1               ------------------------------SSKKKMKNTK---------- 127
A.thaliana_AT1G26680.1#1        FVGSVNPSSLYKDQLYLPRNFVSSNFLDKRCSEIVLKNE-RGEKRTLVLK 485
B.napus_CD826203#1              -------------------------------------------------- 
A.thaliana_AT2G24700.1#1        FVATVTGSNLKRDTLYIPKEFALSNGLMNKYQIVLMNEEGESWKIDLRRE 200
S.lycopersicum_TC201533#1       ----------------------------GIEKRKVEGDQ---------- 142

A.thaliana_AT2G24650.1#1        P-SCGTVYVRGGWVSFCDANGLQAGDIYTFKLIKRGGTLVLRLLP----- 1196
A.thaliana_AT2G24650.2#1        P-SCGTVYVRGGWVSFCDANGLQAGDIYTFKLIKRGGTLVLRLLP----- 224
A.thaliana_AT4G00260.1#1        P-SNGTVYVRGGWVSFCDANGLKAGDNYTFKLIKRAGTLVLRLLPNEPKE 234
B.napus_TC73539#1               -------------------------------------------------- 
A.thaliana_AT1G26680.1#1        HFKKDLTFLKKGWTSFCQVNRIKAGDSFKFKLVGTWNKPVLSLCPTETNY 535
B.napus_CD826203#1              --KKD----------------------------------------------- 143
A.thaliana_AT2G24700.1#1        AYNYGRFYMRRGWRSFCIANGKKPGDVFAFKLVKNEETPMIQLFPMTIED 250
S.lycopersicum_TC201533#1       -------------------------------------------------- 

A.thaliana_AT2G24650.1#1        -------------------------------------------------- 
A.thaliana_AT2G24650.2#1        -------------------------------------------------- 
A.thaliana_AT4G00260.1#1        EANEVSLPEEPES------------------------------------D 248
B.napus_TC73539#1               -------------------------------------------------- 
A.thaliana_AT1G26680.1#1        HKTPLACSEGNKSEESEEEGTEDKNTSQDCLEVKKRKYWSTCRASAENID 585
B.napus_CD826203#1              -----SCS---------------------------------------- 146
A.thaliana_AT2G24700.1#1        LDKLQSLP---------------------------------------- 258
S.lycopersicum_TC201533#1       -------------------------------------------------- 

A.thaliana_AT2G24650.1#1        ------------------KGAESCSLDPSCFVANVAPSTLRYDT--LYL 1225
A.thaliana_AT2G24650.2#1        ------------------KGAESCSLDPSCFVANVAPSTLRYDT--LYL 253
A.thaliana_AT4G00260.1#1        AERNLEKIQRKEKVKKNVTREAESSSQDPSCFVANVSPSSLRYDT--LYL 296
B.napus_TC73539#1               ------------------REAESSTLEPSCYVANVTPSSLRYDM--LNI 156
A.thaliana_AT1G26680.1#1        DDQTNIGNSSKEKRVKKNPVKKAESSSDHSSFVANVTASSLNYDR--LYL 633
B.napus_CD826203#1              ------------------------------CFLANVTVSSLHEDK--LYV 164
A.thaliana_AT2G24700.1#1        -------------RHKIRKTEAAPSSPDLSSFVATVTASNLSRDR--LYL 293
S.lycopersicum_TC201533#1       --------------------PHYQCMKGSSEFTARIKEYNVRKRSPYMHI 172
                                                    : * :   .:      : :

A.thaliana_AT2G24650.1#1        PKRFMRENGVDKRRGE-MILMNEKGKSWTLDLKLKKSCGTSLIRRGWRSF 1274
A.thaliana_AT2G24650.2#1        PKRFMRENGVDKRRGE-MILMNEKGKSWTLDLKLKKSCGTSLIRRGWRSF 302
A.thaliana_AT4G00260.1#1        PKRFMRENGVDKRCGE-MILINEKGKSWTLDLKVKKSSGTSLIKRGWRSF 345
B.napus_TC73539#1               PKSFARANGIESRSGEIIVLMNEKGTSWTLILKRKNSCGTMYITRGWRRF 206
A.thaliana_AT1G26680.1#1        PLSFVSSNGLDKMNGKEIVLLNEEGLSWKFNLKYNQAGKHTYVRPGWSRF 683
B.napus_CD826203#1              PVSFVRSNGLSNTYCK-IVLLNEKGRSWKLS------------------- 194
A.thaliana_AT2G24700.1#1        PKTFIMSNGLLKKFQM--CLMNEEGESWTIDVKHEAHTGRFLTIRGWRRF 341
S.lycopersicum_TC201533#1       PTEFCQSNALFQNTTMTLTGPSGKSCPVSLRICNGGKTLYACITRGWHDF 222
                                *  *   *.: .          . :. . .:

A.thaliana_AT2G24650.1#1        CSANGLRAGSIITFKLIKKSGTLVLCLLSNEPEEEVCSEANEVESLSTDQ 1324
A.thaliana_AT2G24650.2#1        CSANGLRAGSIITFKLIKKSGTLVLCLLSNEPEEEVCSEANEVESLSTDQ 352
A.thaliana_AT4G00260.1#1        CSANGLRAGSIITLKLIKKRATLVLRLIPNEPE-----EANEVVSLSTEQ 390
B.napus_TC73539#1               CRANGLKAGCFFTFKLIQRGGT----------------------SFST-- 232
A.thaliana_AT1G26680.1#1        CDANGMSQGQQFTFKLVQKHGPPVMYLS--LSEHRPKSESSSSHRSYFVGS 731
B.napus_CD826203#1              -------------------------------------------------- 
A.thaliana_AT2G24700.1#1        CVANGKKPGDLLKFKLVHNEETPVLQLLPLNSEDLHKLNPSNDTRHGQSL 391
S.lycopersicum_TC201533#1       FSSNKLKVGDVCLFQLDRSKSD--------------------SNSIAIDV 252

A.thaliana_AT2G24650.1#1        ESHEESSHNEKISRRKEKKGRMIWKASSSPSENRFVTLNLTPYN------ 1368
A.thaliana_AT2G24650.2#1        ESHEESSHNEKISRRKEKKGRMIWKASSSPSENRFVTLNLTPYN------ 396
A.thaliana_AT4G00260.1#1        ESDEESIHDEKISRRK-----------SLLSENRFVTLTLTPYT------ 423
B.napus_TC73539#1               -------------------------------------------------- 
A.thaliana_AT1G26680.1#1        VTASSIKKDKLYLWKSFVSSNGLDKGCKKILKNKWGREWKLVLKHYKSN 781
B.napus_CD826203#1              -------------------------------------------------- 
A.thaliana_AT2G24700.1#1        KVTKKEFLGMEATENEFLGEEVYCNDSFKASEKDTLPFAEPINEDIRQGQ 441
S.lycopersicum_TC201533#1       RVL---------------------------------------------- 255
```

FIG. 12 (Cont)

```
A.thaliana_AT2G24650.1#1        ----------------------------ILRSAL----------RLP 1377
A.thaliana_AT2G24650.2#1        ----------------------------ILRSAL----------RLP 405
A.thaliana_AT4G00260.1#1        ----------------------------IQSSLLNENLLCESMFQRLP 443
B.napus_TC73539#1               --------------------------------------------------
A.thaliana_AT1G26680.1#1        CFTIIKRGWTSFCQGNGLKAGDSFKFKLVGTGEKPVLSLCPAESSHEKIP 831
B.napus_CD826203#1              --------------------------------------------------
A.thaliana_AT2G24700.1#1        CSQTIKQEYVSTEEKNS------------TSQNRFVTLTLTHSSKLNLP 478
S.lycopersicum_TC201533#1       --------------------------------------------------

A.thaliana_AT2G24650.1#1        IPFTR-----MNGINEETKMTLLDKHGMKWLTTLRLVDYKRKRLGMVGGW 1422
A.thaliana_AT2G24650.2#1        IPFTR-----MNGINEETKMTLLDKHGMKWLTTLRLVDYKRKRLGMVGGW 450
A.thaliana_AT4G00260.1#1        VPFTR-----MNGINEETKMTLLDKHGVKWLTTLRFEDDKRKRLRMVGGW 488
B.napus_TC73539#1               --------------------------------------------------
A.thaliana_AT1G26680.1#1        LECPEGIDDVNSLSSNPSSGDDSSRSEESEEENMEDKNISQDCLETKK-- 879
B.napus_CD826203#1              --------------------------------------------------
A.thaliana_AT2G24700.1#1        FEFMK-----RNGIKKAGKITMVDRYDAKWRTSLLMDKIGTMSLGRGS-- 521
S.lycopersicum_TC201533#1       --------------------------------------------------

A.thaliana_AT2G24650.1#1        KGFIQANGVKANESIMLELIWEEETSCVLK-FCSKVKLAIK 1462
A.thaliana_AT2G24650.2#1        KGFIQANGVKANESIMLELIWEEETSCVLK-FCSKVKLAIK 490
A.thaliana_AT4G00260.1#1        QGFIQANDVKANESIMLELIWEEETSCVLK-FCSKVKLEIK 528
B.napus_TC73539#1               ------------------------------------------
A.thaliana_AT1G26680.1#1        RKYCSSSSYSQNRFVTLTLTRSAFQTYKLVSFFNNYASRLS 920
B.napus_CD826203#1              ------------------------------------------
A.thaliana_AT2G24700.1#1        KGFCEVNGVEMNESFILELIWED--TVPLLKFCSKV----- 555
S.lycopersicum_TC201533#1       ------------------------------------------
```

FIG. 12 (Cont)

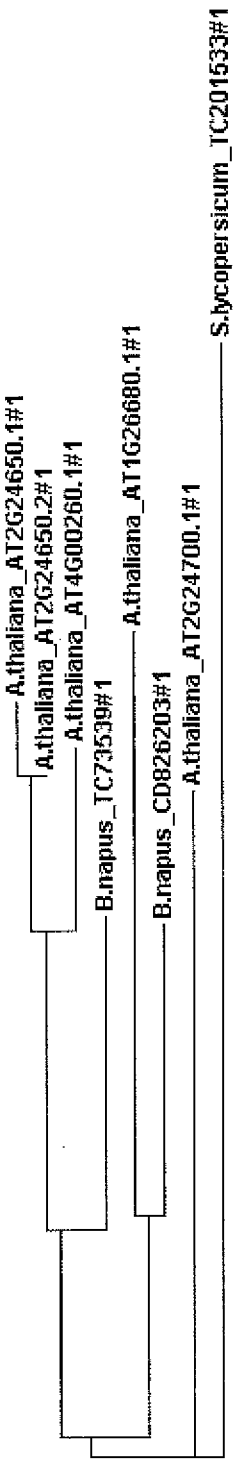

A.thaliana_AT2G24650.1#1
A.thallana_AT2G24650.2#1
A.thaliana_AT4G00260.1#1
B.napus_TC73539#1
A.thaliana_AT1G26680.1#1
B.napus_CD826203#1
A.thaliana_AT2G24700.1#1
S.lycopersicum_TC201533#1

FIG. 13

	1	2	3	4	5	6	7	8
1. A.thaliana_AT4G00260.1#1		40.5	26.6	75.2	25.0	19.2	29.9	17.3
2. A.thaliana_AT2G24700.1#1	59.8		16.1	42.9	26.9	19.7	24.4	16.2
3. A.thaliana_AT2G24650.1#1	29.4	24.3		33.1	33.9	7.3	10.5	8.0
4. A.thaliana_AT2G24650.2#1	82.2	58.9	33.2		25.2	20.0	32.5	17.7
5. A.thaliana_AT1G26680.1#1	35.4	38.2	36.0	34.5		13.6	11.6	10.8
6. B.napus_CD826203#1	26.9	27.6	10.1	28.6	17.3		43.0	25.3
7. B.napus_TC73539#1	35.8	31.4	12.7	38.8	18.8	57.8		31.6
8. S.lycopersicum_TC201533#1	25.9	26.1	12.0	27.8	15.7	42.0	49.4	

Fig. 14

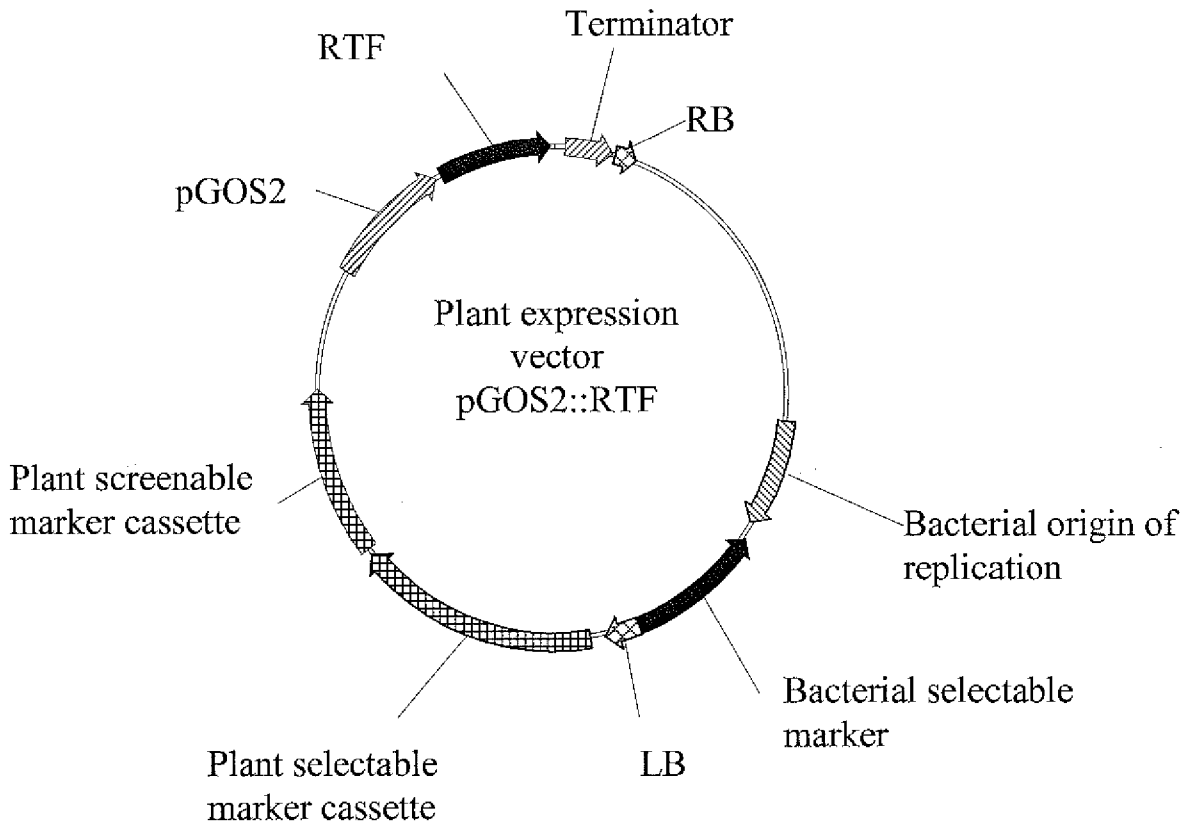

RTF
Terminator
RB
pGOS2
Plant expression
vector
pGOS2::RTF
Plant screenable
marker cassette
Bacterial origin of
replication
Bacterial selectable
marker
Plant selectable
marker cassette
LB

FIGURE 15

```
MDYGDGEVRLVRRKGKKRLAPPPPPPPAAEGGERDRLDELRRDYRDVLKDNEMKRRKLESINKRKLVL
                                                          ----------2--

LSEVKFLQKKLNSFKKNDSQQVRLKKKAPRVPSHVGINDASAFYGASTEVPSTSKRTDLDLNQDSAMN
                                                                  --------

--------------------

DELSDFPGHHNHLELKKAEQAGVDEDIMTADVNLSACRDTGNSPASDDKRSVSWQDRVALKV
------------1------------------------------
                                             --------3------
```

FIG. 16

```
                                  1                                                 50
A.lyrata_944925                   ~M..................................................
A.thaliana_AT4G30630              ~M..................................................
B.napus_TC91202                   ~M..................................................
C.annuum_TC15926                  ~M..................................................
S.lycopersicum_TC206342           ~M..................................................
S.lycopersicum_16878              ~M..................................................
C.intybus_TA970_13427             ~M..................................................
C.tinctorius_TA4044_4222          ~M..................................................
H.annuus_TC40508                  ~M..................................................
H.argophyllus_TA3915_73275        ~M..................................................
T.erecta_417                      MM..................................................
E.esula_TC2982                    MM..................................................
G.hirsutum_TC136942               ~M..................................................
T.cacao_TC4923                    ~M..................................................
P.trichocarpa_578729              ~M..................................................
P.trichocarpa_scaff_VI.1304       ~M..................................................
G.max_Glyma02g36620               ~M..................................................
G.max_Glyma17g08070               ~M..................................................
P.vulgaris_TC10046                ~M..................................................
M.truncatula_AC150446_9.5         MM..................................................
L.japonicus_TC37963               ~M..................................................
C.maculosa_TA3603_215693          ~M..................................................
S.henryi_DT604565                 ~M..................................................
H.vulgare_TC185682                ~M..................................................
O.sativa_LOC_Os08g16930           ~M..................................................
P.edulis_FP092509                 ~M..................................................
P.virgatum_TC29094                ~M..................................................
S.bicolor_Sb07g011060             ~M..................................................
Z.mays_GRMZM2G371316_T01          ~M..................................................
Z.mays_GRMZM2G075851_T01          ~M..................................................
Z.officinale_TA6335_94328         ~M..................................................
Z.officinale_TA6947_94328         ~M..................................................
O.sativa_LOC_Os09g25410           ~M..................................................
P.virgatum_TC30704                ~MDDIGGERGRKGD....ALLT.PPRGPGLPSSSPSPPGSRLLPSPSS..
S.bicolor_Sb02g024920             ~MDDIAGEPGRKGDDSSRAVLTPPPRGPGLLSPSSPPSGSGPIPSPPSAA
Z.mays_GRMZM2G093731_T02          ~MDDIAGERGRKGDDSSRAVLT.PPRGPGLLSPSS.PSRSDPIPSPPS..
P.glauca_BT104710                 ~M..................................................
P.sitchensis_TA13795_3332         ~M..................................................
P.taeda_TA10646_3352              ~M..................................................
L.saligna_TA2168_75948            ~M..................................................
L.sativa_TC20185                  ~M..................................................
S.chacoense_TA1669_4108           ~M..................................................
S.tuberosum_AM908388              ~M..................................................
S.lycopersicum_TC195266           ~M..................................................
S.tuberosum_NP13064295            ~M..................................................
C.maculosa_TA1609_215693          ~M..................................................
H.tuberosus_EL464130              ~M..................................................
L.sativa_DW131501                 ~M..................................................
L.virosa_TA2198_75947             ~M..................................................
L.serriola_BU011148               ~M..................................................
F.vesca_TA11867_57918             ~M..................................................
Z.officinale_TA4076_94328         ~M..................................................
M.crystallinum_TC9929             ~M..................................................
```

Fig. 17

```
                                  51                                                100
A.lyrata_944925                   ..................................K..ETNIVSSAS
A.thaliana_AT4G30630              ..................................K..KTNFVSSAS
B.napus_TC91202                   ..................................K..EITDLV...
C.annuum_TC15926                  .................................SK..KMKRGALEL
S.lycopersicum_TC206342           .................................SK..KMKRVALES
S.lycopersicum_16878              .................................SK..KMKRGALES
C.intybus_TA970_13427             .................................SK..KMKGVAFDS
C.tinctorius_TA4044_4222          .................................S...KMKGVAYDS
H.annuus_TC40508                  .................................SK..KMKGVAFDP
H.argophyllus_TA3915_73275        .................................SK..KMKGVAFDP
T.erecta_417                      .................................SN..KMKGVAFDS
E.esula_TC2982                    ..................................K..RLKGVAMPF
G.hirsutum_TC136942               ..................................K..KMKGVAAAA
T.cacao_TC4923                    ..................................K..KLKGVAAAV
P.trichocarpa_578729              ..................................K..KMKGVAAAA
P.trichocarpa_scaff_VI.1304       ..................................K..KMKAAVAAM
G.max_Glyma02g36620               ..................................K..KMKGV.VPM
G.max_Glyma17g08070               ..................................K..KMKGV.VSL
P.vulgaris_TC10046                ..................................K..KIKPV.VSM
M.truncatula_AC150446_9.5         ..................................K..KMKGVSASM
L.japonicus_TC37963               ..................................K..KMKGVVVSV
C.maculosa_TA3603_215693          .................................SK..KVKGNGFG.
S.henryi_DT604565                 .................................MK..KMKRVSMES
H.vulgare_TC185682                ..................................K..RANKVVLAQ
O.sativa_LOC_Os08g16930           ...................................KKAKVVLAQ
P.edulis_FP092509                 ..................................K..KGKKVVLAQ
P.virgatum_TC29094                ..................................AAGKKARVVLAQ
S.bicolor_Sb07g011060             ..................................AAIKKTKVVLAQ
Z.mays_GRMZM2G371316_T01          ..................................AAIKKTKIVLAQ
Z.mays_GRMZM2G075851_T01          ..................................AATKKTKIVLAQ
Z.officinale_TA6335_94328         ..................................K..RGEKLAVQS
Z.officinale_TA6947_94328         ..................................R..KGQKL....
O.sativa_LOC_Os09g25410           ........................DYGDGEVRLVR..RKGKKRLAP
P.virgatum_TC30704                .AAVLALPASRDPSPGAGLPAAPRLAGAMDGGREPLLVR..RSKDKKKKR
S.bicolor_Sb02g024920             AAAVLALPASA.TSQGAGLLASSRLAGEMDGGGETLLVR..RSKGKKK.R
Z.mays_GRMZM2G093731_T02          .AAVLALPAAA.TSPSAGLLASSRLTGEMDGGGETLLVR..RSKGKKRRQ
P.glauca_BT104710                 ..................................S..KKRGSDGE.
P.sitchensis_TA13795_3332         ..................................S..KKRGSDGE.
P.taeda_TA10646_3352              ..................................S..KKRGSDGE.
L.saligna_TA2168_75948            ..................................TLEDTMKVEVES
L.sativa_TC20185                  ..................................TLEDTMKVEVES
S.chacoense_TA1669_4108           ..................................GEEEAKKVETAE
S.tuberosum_AM908388              ..................................GEEEAKKVETDE
S.lycopersicum_TC195266           ..................................GEEEAKKVETAE
S.tuberosum_NP13064295            ..................................AELEAKKVEIVD
C.maculosa_TA1609_215693          ..................................A..........
H.tuberosus_EL464130              ..................................A..........
L.sativa_DW131501                 ..................................A..........
L.virosa_TA2198_75947             ..................................A..........
L.serriola_BU011148               ..................................A..........
F.vesca_TA11867_57918             ..................................A..........
Z.officinale_TA4076_94328         ..................................AKGG........
M.crystallinum_TC9929             ..................................A..........
```

Fig. 17(cont.)

```
                              101                                              150
A.lyrata_944925               .............................CGDP.M...............
A.thaliana_AT4G30630          .............................GGDP.M...............
B.napus_TC91202               ...............................DP.R...............
C.annuum_TC15926              .....SP...................YVMYEDT.K...............
S.lycopersicum_TC206342       .....SP...................YVVFEEN.N...............
S.lycopersicum_16878          .....SS...................YGVFEDT.K...............
C.intybus_TA970_13427         .....SP...................HVSYEDA.R...............
C.tinctorius_TA4044_4222      .....S..........................S.R...............
H.annuus_TC40508              .....S......................PYEDV.R...............
H.argophyllus_TA3915_73275    .....A......................PYEDV.R...............
T.erecta_417                  .....SS...................HGSFEDV.R...............
E.esula_TC2982                EASPSSY...................AAVYEDP.R...............
G.hirsutum_TC136942           AVMAYSP...................YDMYDDQ.R...............
T.cacao_TC4923                AAMEYSP...................YATYEDQ.R...............
P.trichocarpa_578729          DSPPPSY...................ATMYEDP.R...............
P.trichocarpa_scaff_VI.1304   DFSPPSY...................AAMYEDP.R...............
G.max_Glyma02g36620           E....PP...................YEVYQDQ.R...............
G.max_Glyma17g08070           D....PP...................YEVYQDQ.R...............
P.vulgaris_TC10046            D....PP...................YEAYQGQ.R...............
M.truncatula_AC150446_9.5     N....PP...................PDVYMDH.R...............
L.japonicus_TC37963           D....SP.....................VYEDQ.R...............
C.maculosa_TA3603_215693      .....SY...................PHEDDNG.R...............
S.henryi_DT604565             .....SP...................F.AIEEA.K...............
H.vulgare_TC185682            PAAAAAPPR...LLPRAPRPIR....GDGGEAY.R...............
O.sativa_LOC_Os08g16930       PAARAPPSP...LLPRAPGHSRGGGVGGGEEAY.R...............
P.edulis_FP092509             PVVAASSP.....LSRAARPVHG.GVGGGGEAY.R...............
P.virgatum_TC29094            PAAGALPPRPPALFSRAPGPVRG...SGEEEAY.R...............
S.bicolor_Sb07g011060         PAVGAPASRPP.FFSRAPGPVRG...AGDEAAY.R...............
Z.mays_GRMZM2G371316_T01      PAVGAQPSRPP.FFSRAPGSVRG...AGDEAAY.R...............
Z.mays_GRMZM2G075851_T01      PAVGAPTSRLP.FFSRAPGPVRG...AGDEAAH.R...............
Z.officinale_TA6335_94328     .............................SPAA.K...............
Z.officinale_TA6947_94328     .............................SEEV.K...............
O.sativa_LOC_Os09g25410       PPPPPPA.......................AEG.G...............
P.virgatum_TC30704            PAAAAAH.....................AERDA.G...............
S.bicolor_Sb02g024920         PQPA........................AERGSGG...............
Z.mays_GRMZM2G093731_T02      PAAP........................AERGS.G...............
P.glauca_BT104710             ..........................FGTEDLR.R...............
P.sitchensis_TA13795_3332     ..........................FGTEDLR.R...............
P.taeda_TA10646_3352          ..........................FGTEDLR.R...............
L.saligna_TA2168_75948        ECPPEPP......VAAVEKAIV..PVEPHVHEE.KPTDDTKALAIIEKPT
L.sativa_TC20185              ECPPEPP......VAAVEKAIV..PVEPHVHEE.KPTDDTKALAIIEKPT
S.chacoense_TA1669_4108       .............HVAEEKAIVLSTVPP.KELE.DKPDDSKALVIVEPET
S.tuberosum_AM908388          .............HVAEEKAIVLSTVPPSKELE.DKHDDSKALVIVEPET
S.lycopersicum_TC195266       .............HAAEEKAIVLSTVPPSEESK.DKPDDSKALVIVEPET
S.tuberosum_NP13064295        PAPAQEPVEAPKEVVADEKAIVEPAPPPPAEEK.EKPDDSK..VIVEPET
C.maculosa_TA1609_215693      ..................................................
H.tuberosus_EL464130          ..................................................
L.sativa_DW131501             ..................................................
L.virosa_TA2198_75947         ..................................................
L.serriola_BU011148           ..................................................
F.vesca_TA11867_57918         ..................................................
Z.officinale_TA4076_94328     ..................................................
M.crystallinum_TC9929         ..................................................
```

Fig. 17(cont.)

```
                                   151                                              200
A.lyrata_944925                    .....................IRF...TLMQQVLELEK...ELE......
A.thaliana_AT4G30630               .....................IRL...TFV.QVLELEK...ELE......
B.napus_TC91202                    .....................IMFKHQSLLQDYHELRK...ETE......
C.annuum_TC15926                   .....................ARLKHQTLFQDFQELHK...ETN......
S.lycopersicum_TC206342            .....................ARLKHQTLLQDYLELHK...DTN......
S.lycopersicum_16878               .....................DRLKHQTLLQDYQELQK...ETD......
C.intybus_TA970_13427              .....................ARFKHQTLLQDYLELEQ...ETE......
C.tinctorius_TA4044_4222           .....................ARFKHQTLLQDYLELQQ...ETE......
H.annuus_TC40508                   .....................ARFKHQTLVQDFLELQQ...ETE......
H.argophyllus_TA3915_73275         .....................ARFKHQTLVQDFLELQQ...ETE......
T.erecta_417                       .....................ARFKHQTLVQDYLELQQ...ETE......
E.esula_TC2982                     .....................IRMRHQSLMHDFEDLYK...ETE......
G.hirsutum_TC136942                .....................ARFRYPTHMEDIEDLHK...ETE......
T.cacao_TC4923                     .....................TRFKHQSLMQDFEDLHK...ETE......
P.trichocarpa_578729               .....................IMFKHQSLMQDYEELYK...ETE......
P.trichocarpa_scaff_VI.1304        .....................IMLKHQSLMQDYDELYK...ETE......
G.max_Glyma02g36620                .....................ARLRHQSLLQDYEDLHK...ETE......
G.max_Glyma17g08070                .....................ARLRHQSLLQDYEDLHK...ETE......
P.vulgaris_TC10046                 .....................TRFRHQSLLQDYEDLHK...ETE......
M.truncatula_AC150446_9.5          .....................SMMRHQSLLQDYEDLNK...ETE......
L.japonicus_TC37963                .....................TRLRHQSLLQDYEDLLE...DTN......
C.maculosa_TA3603_215693           .....................PRFKHQTLVQNYLDLQK...ETE......
S.henryi_DT604565                  .....................NRFKLQNLIRDFDELQK...ETE......
H.vulgare_TC185682                 .....................TSQRYQSLLQDYKELLK...ETA......
O.sativa_LOC_Os08g16930            .....................ARAKYKNLLQDYKELLE...ETE......
P.edulis_FP092509                  .....................ARLRYQSLLQDYQELVK...ETE......
P.virgatum_TC29094                 .....................ARLRYQALLHDYQELVK...EAQ......
S.bicolor_Sb07g011060              .....................ASLRYRALLQDYQELIK...ETQ......
Z.mays_GRMZM2G371316_T01           .....................ASLRYRALHQDYQDLIK...EIQ......
Z.mays_GRMZM2G075851_T01           .....................ASLRYRSLLQDYQELIK...ETQ......
Z.officinale_TA6335_94328          .....................ARFRHQSLLQDYEDLVK...ETD......
Z.officinale_TA6947_94328          .....................ARSRHQSLLWDYGDLVE...ETE......
O.sativa_LOC_Os09g25410            .....................ERDRLDELRRDYRDVLK...DNE......
P.virgatum_TC30704                 .....................AGDRSRSLWRDYNDLVQ...ETE......
S.bicolor_Sb02g024920              .....................SGDRFRTLWRDYHDLLQ...ETE......
Z.mays_GRMZM2G093731_T02           .....................SGDRFRALWRDYHDLLQ...ETE......
P.glauca_BT104710                  .....................MAMKHASLKQEYNELQR...ETR......
P.sitchensis_TA13795_3332          .....................MAMKHASLKQEYNELQR...ETR......
P.taeda_TA10646_3352               .....................MAMKHASLKQEYNELQR...ETR......
L.saligna_TA2168_75948             EVCEEKPKEGSINRDDVLAKVATEKKLALITAWEESEKSKAENK......
L.sativa_TC20185                   EVCEEKPKEGSINRDDVLAKVATEKKLALITAWEESEKSKAENK......
S.chacoense_TA1669_4108            KALVPVEKKGSIDRDATLARLTTEKRLSLIKAWEESEKSKAENK......
S.tuberosum_AM908388               KALVPVEKKGSIDRDATLARLTTEKRLLLIKAWEESEKSKAENK......
S.lycopersicum_TC195266            KALVPVEKKGSIDRDATLARLTTEKRLSLIKAWEESEKAKAENK......
S.tuberosum_NP13064295             KALVPVEKKGSIDRDATLARLTTEKRLSLIKAWEESEKSKAENK......
C.maculosa_TA1609_215693           .....................TKPSNRVILEQLRSGVA...HFELVSSPV
H.tuberosus_EL464130               .....................AKRSNGAILEDLRTGLA...EFELVSSPV
L.sativa_DW131501                  .....................AKPSNGIILEELRNGVA...EFELVSSPV
L.virosa_TA2198_75947              .....................AKPSNGVILEELRNGVA...EFELVSSPV
L.serriola_BU011148                .....................AKPSNGIILEELRNGVA...EFELVSSPV
F.vesca_TA11867_57918              .....................HKLSNEAVLEQLKNGFA...QFELVSSPL
Z.officinale_TA4076_94328          .....................RLPSNEVLLEQLRNGTA...QFELLPTPV
M.crystallinum_TC9929              .....................ATFSNDSILQQLRNGAA...QFELVTSLA
```

Fig. 17(cont.)

```
                                201                                              250
A.lyrata_944925                 ........................................VKRKK...LE
A.thaliana_AT4G30630            ........................................IKRKR...LE
B.napus_TC91202                 ........................................FKMRK...LE
C.annuum_TC15926                ........................................DMRDK...LE
S.lycopersicum_TC206342         ........................................NTRNK...LE
S.lycopersicum_16878            ........................................AARHK...LE
C.intybus_TA970_13427           ........................................AARNK...LE
C.tinctorius_TA4044_4222        ........................................AARNK...LE
H.annuus_TC40508                ........................................AARNQ...LE
H.argophyllus_TA3915_73275      ........................................AARNQ...LE
T.erecta_417                    ........................................AARNK...LE
E.esula_TC2982                  ........................................AQRDK...IE
G.hirsutum_TC136942             ........................................SMRKK...LQ
T.cacao_TC4923                  ........................................AMRKK...LQ
P.trichocarpa_578729            ........................................AKKRK...LQ
P.trichocarpa_scaff_VI.1304     ........................................AKKRK...LQ
G.max_Glyma02g36620             ........................................AMRRK...LQ
G.max_Glyma17g08070             ........................................AMRRK...LQ
P.vulgaris_TC10046              ........................................AVRRK...LQ
M.truncatula_AC150446_9.5       ........................................AMRRK...LE
L.japonicus_TC37963             ........................................SMRMK...LQ
C.maculosa_TA3603_215693        ........................................AARNK...LE
S.henryi_DT604565               ........................................DIKEN...LE
H.vulgare_TC185682              ........................................AKKNR...LH
O.sativa_LOC_Os08g16930         ........................................AKKKR...LQ
P.edulis_FP092509               ........................................AKKKL...LH
P.virgatum_TC29094              ........................................AKKRR...LH
S.bicolor_Sb07g011060           ........................................AKKKR...LH
Z.mays_GRMZM2G371316_T01        ........................................AKKKR...LH
Z.mays_GRMZM2G075851_T01        ........................................AKKER...LY
Z.officinale_TA6335_94328       ........................................AKRKN...LH
Z.officinale_TA6947_94328       ........................................AKRKK...LQ
O.sativa_LOC_Os09g25410         ........................................MKRRK...LE
P.virgatum_TC30704              ........................................AKKKR...LL
S.bicolor_Sb02g024920           ........................................GKKKM...LA
Z.mays_GRMZM2G093731_T02        ........................................AKKRR...LA
P.glauca_BT104710               ........................................AMKKR...LQ
P.sitchensis_TA13795_3332       ........................................AMKKR...LQ
P.taeda_TA10646_3352            ........................................AMKKR...LQ
L.saligna_TA2168_75948          ........................................AQQKLSSIIA
L.sativa_TC20185                ........................................AQQKLSSIIA
S.chacoense_TA1669_4108         ........................................AQKKRSEILA
S.tuberosum_AM908388            ........................................AQKKRSEILA
S.lycopersicum_TC195266         ........................................AQKKIAEILA
S.tuberosum_NP13064295          ........................................AQKKRSEILA
C.maculosa_TA1609_215693        TSICTNAIQSQSQSLFNNQNAHRFFALIGPTM..QSSAAAAALSK...VE
H.tuberosus_EL464130            TTIAT...RNSVQSVFLTDHTHRFFARIGPKL..EVS...PALSK...VE
L.sativa_DW131501               TSISTNVYRNPIRFVFITDHTHRFFARIGQKM..EVS...PALSK...VE
L.virosa_TA2198_75947           TSISTNVYRNPIRSVFITDHTHRFFARIGQKM..EVS...PALSK...VE
L.serriola_BU011148             TSISTNVYRNPIRSVFITDHTHRFFARIGQKM..EVS...PALSK...VE
F.vesca_TA11867_57918           LSISASI.SQPNATPFTGDHSHRFFARIGPPLG.RGS...SAMRK...VE
Z.officinale_TA4076_94328       PLSAPDV...............RFFARIGSSLGGGVS...PATKK...VE
M.crystallinum_TC9929           PSVSESKAFASPLQFLQNNNTAGFFARIGPSLGAAAST..SSAKK...GE
```

Fig. 17(cont.)

	251 300
A.lyrata_944925	MIKQKRLTLKSEVRFLRRRYEHLKQDQTLETSPKMLRLSE..SGGL.EIP
A.thaliana_AT4G30630	MIKQKRLTLKSEVRFLRRRYEHLKQDQTLETSPEMLRLSE..SGGL.EVP
B.napus_TC91202	MMTQKRLALDAQVRFLRRRYKHLKQDQTLETSPDTLRLSE..SGG..DVK
C.annuum_TC15926	DAKMRKQRLLAEVCFLRRRHKYLLQMKSSSHPE.QQERAA..LPNS.EFY
S.lycopersicum_TC206342	DMKMRKQKLLAEVHFLRRRHKYLLQMKSSGHLE.QQERAA..LLNT.ELY
S.lycopersicum_16878	DSKLRKLRLQAEVRFLRQRHKHLLQSKSLTTPK.GKELVS..LPNV.ETY
C.intybus_TA970_13427	AMKQKKQTLEAEVRFLRRRHKFLLKTKSATPQQ.....VK..SHNLETTQ
C.tinctorius_TA4044_4222	AMKQKKQTLEAEVRFLRRRHKFLLKTKSSTPQQ.....LK..SQNFETTQ
H.annuus_TC40508	AMKQKKQTLEAEVRFLRRRRKFLLKTRSSTSQE.....VI..LPNIETTQ
H.argophyllus_TA3915_73275	AMKQKKQTLEAEVRFLRRRRKFLLKTRSSTSQE.....VI..LPNIETTQ
T.erecta_417	AMKQKKHTLEAEVRFLRRRHRFLMRTNSSAPQQ.....VK..LHNVETTQ
E.esula_TC2982	IMKKKKMTLLAEVRFLKQRYNYLKQNQSQKKPVSEKKPVQ..RSKL.VNR
G.hirsutum_TC136942	MMKEQRLTLLAEVKFLKRRHKFLMQDQSSKPQA.ERHFLQ..PQDI.VIR
T.cacao_TC4923	MMKEQKLTLLAEVRFLKRRHKFLMQNQSSNTPA.QKNIVQ..PQNL.VIR
P.trichocarpa_578729	MVRQKKLTLMAEVRFLRRRYKYLTQNKPKKAPM.ERSFVQ..PQNL.VPA
P.trichocarpa_scaff_VI.1304	MMRQKKLTLMAEVRFLRRRYKYLTQNKSQKPPM.EQNFVQ..PQNL.VTA
G.max_Glyma02g36620	ATKQKKFMLEDEVRFLRQRYNYLLKHPILKPQP.KQQVVK..PQKL.KIQ
G.max_Glyma17g08070	ATRQKKLMLEDEVRFLRQRYNYLLKHPILKPHP..KQVVK..PQKL.KLQ
P.vulgaris_TC10046	ATEQKRLILEDEVRFLRHRFKYLLKHPIPKPQP.KKEVVK..SQKL.KIQ
M.truncatula_AC150446_9.5	AVKQKKLILSAEVRFLRQRYKYLLKHPLPKPQP.NQEVLK..PQKL.KIK
L.japonicus_TC37963	AAKQKRLTLSAEVSFLRKRYKYLLKNPSPKLQP.KQDISQ..PRKL.KIQ
C.maculosa_TA3603_215693	AMKLKKLTLEAQVRFLRRRHKFLVKNKSCIQQE..HTIVK..PQFL...E
S.henryi_DT604565	KMKQKRLTLLAEVSFLRRRYKHLVKNPSQKTPA..ESLTQ..SQK.....
H.vulgare_TC185682	LEKLKKQRLSAEVKFLRRRYQSMSGNPSQTSVYRLKNPAM.VPTTSRTTV
O.sativa_LOC_Os08g16930	MEKLKKQRLLAEVKFLRKRYKSMSENPSQTFVYRVKNPAL..PPTLRQPG
P.edulis_FP092509	MKKLKKQRLLGEVKFLRRRYKSMSENPSQTVVYRLKNPAM..PPALRATA
P.virgatum_TC29094	MERLNKQRLLAEVKFLRKRYKSMSENPSQTIVCRVRNPAM..RPARRTAA
S.bicolor_Sb07g011060	MERLRKQRLLAEVKFLRKRYKSMSENPSQTIVCRLRNPAM..SSASWTAA
Z.mays_GRMZM2G371316_T01	MKRLKKQRLLAEVKFLRKRYKSMSENPSQTIVCRVRNPAM..SSASWTAG
Z.mays_GRMZM2G075851_T01	MERLRKKRLLAEVKFLRKRYQSMAKNPSQTIACRVKNPAL..SSASWRAA
Z.officinale_TA6335_94328	IAKQKKLKLHAEVKFLRKRYKTLLQNRSGTVQFRLKTKPK..SRFFVAQP
Z.officinale_TA6947_94328	IANQKKLKLLAEVKFLQRRYKSLSQNPSGTKQLRLKKKSQ..SQSLIGQP
O.sativa_LOC_Os09g25410	SINKRKLVLLSEVKFLQKKLNSFKKNDSQQV..RLKKKAPRVPSHVGIND
P.virgatum_TC30704	SANRTKLALLAEVKFLRRKYGSFAKGNSQKTHYKLKKKARYIPSPLG..R
S.bicolor_Sb02g024920	STKRRSLALLAEVKFLRRKYQSFLKGGSQQKHYKLKKQARYMPSPLGSNK
Z.mays_GRMZM2G093731_T02	SKKRRSLALLAEVKFLRRKYQSFPKGGSQQTHYRLKKQARYMPSPLG...
P.glauca_BT104710	RAKLKKENLLAEVRFLRKRHRLLTKRGSQFSEGSPSRKQTLLAQAEYAVP
P.sitchensis_TA13795_3332	RAKLKKDNLLAEVRFLRKRHRLLTKRGSQFSEGSPSRKQTLLAQAEYAVP
P.taeda_TA10646_3352	RAKLKKGNLLAEVRFLRKRHRLLTKRGSQFNEGSPSQKQTFLAQAEYAMP
L.saligna_TA2168_75948	WENSKKADLEAE...LKKMEEDLEKKKAKYIEKMKNKIVL..........
L.sativa_TC20185	WENSKKADLEAE...LKKMEEDLEKKKAKYIEKMKNRIVL..........
S.chacoense_TA1669_4108	WENSKKASLEAE...LKRTEEQLLKKKAEYIEKLKNKIAL..........
S.tuberosum_AM908388	WENSKKASLEAE...LKRTEEQLLKKKAEYIEKLKNKIAL..........
S.lycopersicum_TC195266	WENSKKASLEAE...LKRTEEQLLKKKAEYIEKLKNKIAL..........
S.tuberosum_NP13064295	WENSKKASLEAE...LKRTEEQLLKKKAEYIEKLKNKIAL..........
C.maculosa_TA1609_215693	SFRVHKVTGDGRCMF.RALVKGMAINKSVTLSPREER.............
H.tuberosus_EL464130	SFRVHRVTGDGRCMF.RALVKGMAMNKSLTLSLREER.............
L.sativa_DW131501	SFRVHKVTGDGRCMF.RALVKGMAMNKSVNLSLREER.............
L.virosa_TA2198_75947	SFRVHKVTGDGRCMF.RALVKGMAMNKSVNLSLREER.............
L.serriola_BU011148	SFRVHKVTGDGRCMF.RALVKGMAMNKSVNLSLREER.............
F.vesca_TA11867_57918	RYSVQRVTGDGRCLF.RALVKGMAANKGSPLSPREER.............
Z.officinale_TA4076_94328	RYAVHKITGDGRCMF.RALVKGMANYKGLALSLRQEK.............
M.crystallinum_TC9929	HYAVQKVAGDGRCMF.RALVKGMARNKGIPLTSWQER.............

Fig. 17(cont.)

```
                                  301                                       350
A.lyrata_944925                   RKPSGENKK.....HPGVSAPVPCFD.....
A.thaliana_AT4G30630              RKPSGERKK.....QSGVRASVPCFD.....
B.napus_TC91202                   VTTSGKRKK.....HSG.....PWFD.....
C.annuum_TC15926                  CKDGMKDRF.LSKKEAKQYKLPPLPVPKQKARIQAANEASRQKTPPDILV
S.lycopersicum_TC206342           GINGMNDRF.HSKKEVKQHKLPPLPRLKQKARIQVAKEASLQRTPSDILV
S.lycopersicum_16878              PRSRSRDQV.TRKKEAKLNKLPPLPGPKPNGRIQVAKVVASQNIP.DIHV
C.intybus_TA970_13427             FRKSKKENV.HPKKAPTLRNLPPIHN.....
C.tinctorius_TA4044_4222          FKKNKKEKV.HPKKSATLPNLPPIHN.....
H.annuus_TC40508                  YRKSKKEKVNNPKKGTTLTNLPPTPN.....
H.argophyllus_TA3915_73275        YRKSKKEKVNNPKKGATLTNLPPTPN.....
T.erecta_417                      YRKSKKDKV.NPKKVATLPNLPPIPN.....
E.esula_TC2982                    NRTMRKEKE.HSGKNAA.....AQID.....
G.hirsutum_TC136942               SKSNMKAKK.STGKGTSSNR.AMAFH.....
T.cacao_TC4923                    SKSNMKEKK.STGKERTMQRLATGFD.....
P.trichocarpa_578729              SKNLKKEKS.YSGNNAALRPPVPRFD.....
P.trichocarpa_scaff_VI.1304       SKKL.KEKN.SCGNNSSLRPPVPRFD.....
G.max_Glyma02g36620               APIISKGKN.YNKKEPNLRPHHPASH.....
G.max_Glyma17g08070               APIISKGKN.YSRKEPTLRP.HPTSH.....
P.vulgaris_TC10046                TPIISKGKS.YSRKDHTSRS.HSASH.....
M.truncatula_AC150446_9.5         EPKIAKGRN.YNRKESTLRP.HAASK.....
L.japonicus_TC37963               ANLIPKGQK.YNRKES.MQP.PVAPH.....
C.maculosa_TA3603_215693          CRKFSNEV..YSRNEVALQN.....
S.henryi_DT604565                 .....
H.vulgare_TC185682                WV..DHHRP.....
O.sativa_LOC_Os08g16930           WSHGEEYHT.....
P.edulis_FP092509                 WAHDEEHRP.....
P.virgatum_TC29094                WANDAQHRS.....
S.bicolor_Sb07g011060             SAGDAQHQS.....
Z.mays_GRMZM2G371316_T01          SAGDTWHQS.....
Z.mays_GRMZM2G075851_T01          SVSDAQHQL.....
Z.officinale_TA6335_94328         SNSNASHQ.....
Z.officinale_TA6947_94328         SNANAYPQ.....
O.sativa_LOC_Os09g25410           ASAFYG.....
P.virgatum_TC30704                ASAFEDHDV.....
S.bicolor_Sb02g024920             ATALGDHG.....
Z.mays_GRMZM2G093731_T02          ......HG.....
P.glauca_BT104710                 LKTPLATQAIPYEYD.....
P.sitchensis_TA13795_3332         LKTPLATQAIPYEYD.....
P.taeda_TA10646_3352              LKTSLTAQAIPYEYE.....
L.saligna_TA2168_75948            .....
L.sativa_TC20185                  .....
S.chacoense_TA1669_4108           .....
S.tuberosum_AM908388              .....
S.lycopersicum_TC195266           .....
S.tuberosum_NP13064295            .....
C.maculosa_TA1609_215693          .....
H.tuberosus_EL464130              .....
L.sativa_DW131501                 .....
L.virosa_TA2198_75947             .....
L.serriola_BU011148               .....
F.vesca_TA11867_57918             .....
Z.officinale_TA4076_94328         .....
M.crystallinum_TC9929             .....
```

Fig. 17(cont.)

```
                              351                                               400
A.lyrata_944925               .....................LKQKNTICN..EKE...............
A.thaliana_AT4G30630          .....................LKLKNTICN..EKE...............
B.napus_TC91202               .....................LKRKDTVCN..DNN...............
C.annuum_TC15926              H.........NTFLHSTVSS.LKHTSRVHS..GKE.V.............
S.lycopersicum_TC206342       NHKQALQIGKDAVHRSTVSSGLKHQSRVYS..GKE.M.............
S.lycopersicum_16878          NQKHRLGGRQESVFRNADLE.LNRKSRVYR..GKD.V.............
C.intybus_TA970_13427         .....................LNQRGRVYT..EKKNS.............
C.tinctorius_TA4044_4222      .....................LNQRGRVYT..EKKNM.............
H.annuus_TC40508              .....................VIQRGKMYT..KKKNV.............
H.argophyllus_TA3915_73275    .....................VIQRGKMYT..KKKNV.............
T.erecta_417                  .....................QFQRGKTYS..KKKNI.............
E.esula_TC2982                .....................LNKKKKVYS..EKE.S.............
G.hirsutum_TC136942           .....................LNQKGKTNS..VKG.T.............
T.cacao_TC4923                .....................LNQKGKTYS..EKE.T.............
P.trichocarpa_578729          .....................LNQKGKLYI..ERE.A.............
P.trichocarpa_scaff_VI.1304   .....................LNQKGKVYI..ERE.A.............
G.max_Glyma02g36620           .....................LNSNGRISN..VAD.V.............
G.max_Glyma17g08070           .....................LNSNGRISN..VAE.A.............
P.vulgaris_TC10046            .....................LNPNGKISN..VAE.V.............
M.truncatula_AC150446_9.5     .....................LNPKERVFNGRAVE.D.............
L.japonicus_TC37963           .....................WNSNGRISN..GAEIN.............
C.maculosa_TA3603_215693      .....................MNRNVKTYA..QKNNKNVLLGIPSHMLHG
S.henryi_DT604565             .......................QNSRMVR..GAATM.............
H.vulgare_TC185682            .....................IQAVG........................
O.sativa_LOC_Os08g16930       .....................VRAIG........................
P.edulis_FP092509             .....................VQAVG........................
P.virgatum_TC29094            .....................IQAIG........................
S.bicolor_Sb07g011060         .....................VHGVGSSQH....................
Z.mays_GRMZM2G371316_T01      .....................VRAAG........................
Z.mays_GRMZM2G075851_T01      .....................VRAVG........................
Z.officinale_TA6335_94328     ..................................................
Z.officinale_TA6947_94328     ..................................................
O.sativa_LOC_Os09g25410       ..................................................
P.virgatum_TC30704            ..................................................
S.bicolor_Sb02g024920         ..................................................
Z.mays_GRMZM2G093731_T02      ..................................................
P.glauca_BT104710             .....................YATTGTSSD....................
P.sitchensis_TA13795_3332     .....................YATTGTSSD....................
P.taeda_TA10646_3352          .....................YTTTGTSSD....................
L.saligna_TA2168_75948        .....................LHKKA........................
L.sativa_TC20185              .....................LHKKA........................
S.chacoense_TA1669_4108       .....................VHKSA........................
S.tuberosum_AM908388          .....................VHKSA........................
S.lycopersicum_TC195266       .....................VHKSA........................
S.tuberosum_NP13064295        .....................VHKSA........................
C.maculosa_TA1609_215693      ..................................................
H.tuberosus_EL464130          ..................................................
L.sativa_DW131501             ..................................................
L.virosa_TA2198_75947         ..................................................
L.serriola_BU011148           ..................................................
F.vesca_TA11867_57918         ..................................................
Z.officinale_TA4076_94328     ..................................................
M.crystallinum_TC9929         ..................................................
```

Fig. 17(cont.)

```
                                   401                                                450
A.lyrata_944925                    ....ALAKN.ASCDLEKKRR.....RSRGNDVLTIPVSLPDLNGEG....
A.thaliana_AT4G30630               ....ALANNVASCDLDKKRK.....RSRGSDVLTFPVSLPDLNGEG....
B.napus_TC91202                    ....A....................RSRGNEVLT...PLPDLN.DN....
C.annuum_TC15926                   ....LLRKAAPPFDLNQNDR..SFA..GNSSISRSTTPVFDLN.QE.TGH
S.lycopersicum_TC206342            ....LLQNAAPVFDLNQNDR..SFA..GNGFVLRSTIPVFDLN.QE.TDY
S.lycopersicum_16878               ....SSRKGAPAYDLKQDER..LCI..ADNAILQSSSPTFDLN.QD.AGH
C.intybus_TA970_13427              ....NPSRPPP...GDKNVNLHSV..........NLSPVTEFNRKE.SGF
C.tinctorius_TA4044_4222           ....VPSRPIPVFDFDQRVNLHGVN.LNQRVMGNNLSLVNEFNQKE.AAV
H.annuus_TC40508                   ....INNEPPNAVNLGRNVNLYGVD.LNQRVIGNNSFPVTEFIQKE.NGV
H.argophyllus_TA3915_73275         ....TNNEPPNVVNLGRNVNLYGVD.LNQRVIGNNSFPVTEFVQKE.NGV
T.erecta_417                       ....VPSQPLPVFDLDRNVNLQ..........GHNSFPV.....KE.SGF
E.esula_TC2982                     ....TL..........QMQI..SNP.GT..........IPDLNQKE.RVY
G.hirsutum_TC136942                ....CFTHPALMLDVNRKQQ..KISSGKEVT.LRSSLSVLDLNQRD.RLY
T.cacao_TC4923                     ....TFTHPSLMFDLNQKQQ..KILNGKEVTLLRSSLPVLDLNQRE.RVY
P.trichocarpa_578729               ....TLRNSTPIFDLNQKQM..THI.GKEAALRKTA.PIPDLNQKE.RIY
P.trichocarpa_scaff_VI.1304        ....TLRNPAPIFDLNKKQK..THI.GREVALLNST..IPDLNKKE.RTY
G.max_Glyma02g36620                ....PLKKTGHLFDLNLNAR..SSS.KKDASINISGPPVLNLNHKE.RIN
G.max_Glyma17g08070                ....PLKKTGHLFDLNLNAR..SSS.KKDASFNGSAPPVFELNHKE.RIH
P.vulgaris_TC10046                 ....PLQKTSHLFDLNQNARNFSSS.RKDSTIHCSASPARDLNHKE.RLH
M.truncatula_AC150446_9.5          ....TLQKTSHVFDLNQNGR..SLN.KKDASFHSSAAPVLDSNHKD.RVQ
L.japonicus_TC37963                ....TLQKKAPMFDLNQNAR..SLS.RKDPSFLSSAPP.LDLNHKD.MSH
C.maculosa_TA3603_215693           VMDDSLRRATQAYYPNQMAV.......................RTN.PSD
S.henryi_DT604565                  ............................................KE.RKH
H.vulgare_TC185682                 .........................................SSSKGHQPV
O.sativa_LOC_Os08g16930            .........................................SSSKG.PSA
P.edulis_FP092509                  .........................................SSSKG.PSV
P.virgatum_TC29094                 .........................................SSSRS.QLV
S.bicolor_Sb07g011060              ..................................QSVHAVRSSSRS.QLV
Z.mays_GRMZM2G371316_T01           .........................................SSSRS.QLV
Z.mays_GRMZM2G075851_T01           .........................................SSSRN.QLV
Z.officinale_TA6335_94328          .........................................LPAKD.QSS
Z.officinale_TA6947_94328          .........................................LPAKD.QRS
O.sativa_LOC_Os09g25410            ..................................................
P.virgatum_TC30704                 ..................................................
S.bicolor_Sb02g024920              ..................................................
Z.mays_GRMZM2G093731_T02           ..................................................
P.glauca_BT104710                  ...................................EDGIPQAVRKD.SSS
P.sitchensis_TA13795_3332          ...................................EDGIPQAVRKD.SSS
P.taeda_TA10646_3352               ...................................EDGIPQAARKD.SSS
L.saligna_TA2168_75948             ..................................................
L.sativa_TC20185                   ..................................................
S.chacoense_TA1669_4108            ..................................................
S.tuberosum_AM908388               ..................................................
S.lycopersicum_TC195266            ..................................................
S.tuberosum_NP13064295             ..................................................
C.maculosa_TA1609_215693           ..................................................
H.tuberosus_EL464130               ..................................................
L.sativa_DW131501                  ..................................................
L.virosa_TA2198_75947              ..................................................
L.serriola_BU011148                ..................................................
F.vesca_TA11867_57918              ..................................................
Z.officinale_TA4076_94328          ..................................................
M.crystallinum_TC9929              ..................................................
```

Fig. 17(cont.)

```
                                  451                                                 500
A.lyrata_944925                   ....NTSVTDKVPGFDLNQISR...EE...EE.P..EVNGEHMVV.....
A.thaliana_AT4G30630              ....NTSGTDKVPGFDLNQISR...EE...EE.P..EVNGEHMVA.....
B.napus_TC91202                   ....TLVVSSKVSGFDLNLVSR...EE...EEEP..EGNG..........
C.annuum_TC15926                  .GGKDVVLPSRAPVIDLNEISV...GE...EEPP..QANFEPLNF.....
S.lycopersicum_TC206342           .VVKDVVLPSRAPVIDLNEISI...GE...EE.P..QANFEPLNF.....
S.lycopersicum_16878              .SGKETALPTRAPVFDLNEISM...GE...EE.R..DGNSEQVQL.....
C.intybus_TA970_13427             .GGK....EARVPIFDLNQISM...EE...EE...................
C.tinctorius_TA4044_4222          ........QSRAPIFDLNQISM...EE...ED.V..QEGFEDQR......
H.annuus_TC40508                  .SGEKPLVQARAPIFDLNQISM...EE...EE.V..QESFDDQR......
H.argophyllus_TA3915_73275        .SGEKPLVQARAPIFDLNQISM...EE...EE.V..QESFDDQR......
T.erecta_417                      .SGKAAVVQPRAPIFDLNQISV...EE...EE.L..QESFEDQK......
E.esula_TC2982                    .SGKEAGVRNSTPIFDLNQISM...EE...EE.L..QANDEGMRI.....
G.hirsutum_TC136942               .GGKEAAAQSIKPIFDLNQISR...EE...EE.L..EAQTNSVRI.....
T.cacao_TC4923                    .SGKEATARSMTPIFDLNQISR...EE...EE.L..QASDNSMRI.....
P.trichocarpa_578729              .RGKEATVRNNTPIFDLNEIS..................VNGDMLRT.....
P.trichocarpa_scaff_VI.1304       .SVKEATVQNNTPIFDLNEIST...EE...EE....HVNGDMLRT.....
G.max_Glyma02g36620               .SSKEATKKSVTPFFDLNQISR...EE...EE.L..QGNSEPMGI.....
G.max_Glyma17g08070               .SSKEATKKSVTPFFDLNQISR...EE...EE.L..QGNSEAMRI.....
P.vulgaris_TC10046                .SSKEATKKSVAPFFDLNHISR...EE...EE.L..QGNSETMRI.....
M.truncatula_AC150446_9.5         .SGKEAAKKSVTPFFDLNQISR...EE...EE.L..EGDNQSMWI.....
L.japonicus_TC37963               .SMKEPAKKSITPFFDLNQIAR...EE...EE.L..L.VTEPVRI.....
C.maculosa_TA3603_215693          .VSKETPVKTRPPTFDLNQISR...EE...EEEL..QERYEPKDL.....
S.henryi_DT604565                 .KEKEVVKRNPHPSFDLNQISSPGDET...ME.F..HVVWEPLRM.....
H.vulgare_TC185682                QQQKQHLVPRASPIIDLNGACELSSSE.ETEE.F..HGYEETARA.....
O.sativa_LOC_Os08g16930           .HRRLNAAPRASPVIDLNEACEPSSEE..MEE.L..HGYQEPVRV.....
P.edulis_FP092509                 QQRRQHSAPRASPVIDLNEACEPSFEE..MEE.F..HGHQEPVRA.....
P.virgatum_TC29094                .QWRQDDSPRASPVIDLNEACEPGYEI.EMGD.H..HGYGEPLGI.....
S.bicolor_Sb07g011060             .HRRHGGSPRASPVIDLNEACELGYEEMEIDE.H..HGYRAPLGI.....
Z.mays_GRMZM2G371316_T01          .QRTHGGSPRASPVIDLNEACEPGYEEMEIEG.Q..HGYMSPLGV.....
Z.mays_GRMZM2G075851_T01          .QPRRGGSPRVYPAIDLNEACEPGYEMEIVEE.R..HGNRAPLGI.....
Z.officinale_TA6335_94328         .KLRDDASPSTSAIIDLNQVSLPIGED..MDK.Y..QFDAAPSKV.....
Z.officinale_TA6947_94328         .KLRDADTPCTSALIDLNQIGE...DT...EE.YPSSFSVEPMKS.....
O.sativa_LOC_Os09g25410           .ASTEVPSTSKRTDLDLNQDSAMNDEL...SD.F..PGHHNHLEL.....
P.virgatum_TC30704                .ARTEGPSSSKNPNVDLNQGSLVNDER...ND.C..QGHRGHLEP.....
S.bicolor_Sb02g024920             .ARTKVPSTSKNSNLHLNQDSVPNDVV...ND.H..QGQQGHPEV.....
Z.mays_GRMZM2G093731_T02          .AGTEVPSTSRNPNLDLNQDSVLNDVG...ND.C..QGQHGHPEV.....
P.glauca_BT104710                 .SMEFSHSPETKPLVT..............KE.F..QVFWEPLRTAASAG
P.sitchensis_TA13795_3332         .SMEFSHSPETKPLVT..............KE.F..QVFWEPLRTAASAG
P.taeda_TA10646_3352              .SMEFTHSPETKPLVT..............KE.F..QVFWEPLRTAASAG
L.saligna_TA2168_75948            ........................EE...KKAI..................
L.sativa_TC20185                  ........................EE...KKAI..................
S.chacoense_TA1669_4108           ........................EE...KRAI..................
S.tuberosum_AM908388              ........................EE...KRAI..................
S.lycopersicum_TC195266           ........................EE...KRAI..................
S.tuberosum_NP13064295            ........................EE...KRAI..................
C.maculosa_TA1609_215693          .ENADELRMAIKEVLCDNEK.....ERHQYEEALIAITVEEPLKRYCQRI
H.tuberosus_EL464130              .DNADELRMAIKEVLCDNAK.....ERHQYEEAMIAITVEESLKRYCQRI
L.sativa_DW131501                 .ENADELRMAIKEVLCDNEK.....ERHKYEEALIAITVEEPLKRYCLRI
L.virosa_TA2198_75947             .ENADELRMAIKEVLCDNEK.....ERHKYEEALIAITVEEPLKRYCQRI
L.serriola_BU011148               .ENADELRMAIKEVLCDNEK.....ERHKYEEALIAITVEEPLKRYCQRI
F.vesca_TA11867_57918             .DNADELRMAVKEVICENEE.....ERPQYEAALVAITVDEPLKRYCQRI
Z.officinale_TA4076_94328         .DDADELRMAVKEIICDNET.....ERRQYEEALIAITVDESLKQYCQRI
M.crystallinum_TC9929             .DDADELRTAVKEVICENDT.....ERRQYQEAVVAITVEEPLKRYCQRI
```

Fig. 17(cont.)

```
                                  501                                              550
A.lyrata_944925                   ........ESMKNAM..RDNRIS.................DL.HGERKLP
A.thaliana_AT4G30630              ........EAMKNAM..LDNRIS.................DL.HVERKLP
B.napus_TC91202                   ........EATKKVM..LGNGID.....................CELKLP
C.annuum_TC15926                  ........EEPKRGL..IQNTND.................DQ.HKDLKLS
S.lycopersicum_TC206342           ........EEPKRGL..IQNIND.................DQ.HRDLKLS
S.lycopersicum_16878              ........EEPKRSP..IQNLND......................DVKLL
C.intybus_TA970_13427             ........................................EE.QNDLKLS
C.tinctorius_TA4044_4222          ......................K.................EQ.HSDLKLS
H.annuus_TC40508                  ......................K.................EQ.QNDLMLS
H.argophyllus_TA3915_73275        ......................K.................EQ.QNDLMLS
T.erecta_417                      ......................K.................EV.QNDLKLS
E.esula_TC2982                    ........EEGKMISCVMRGGSD.................EQIHGDMKLS
G.hirsutum_TC136942               ........EEFKRSS..TRIGSD.................EQ.HQDVKIS
T.cacao_TC4923                    ........EEFKKSS..MRIGAD.................EQ.HNDIKIS
P.trichocarpa_578729              ........EELKISS..MRGGSA.................EL.HNDMKLS
P.trichocarpa_scaff_VI.1304       ........EEPKISL..MRGASD.................EM.HNERKLS
G.max_Glyma02g36620               ........EEPKRSS..QRVATD.................EQ.HGDIKLS
G.max_Glyma17g08070               ........EEPKRST..QRVATD.................EQ.HSDIKLS
P.vulgaris_TC10046                ........EEPKRST..HRVGSD.................EQ.HNDIKLS
M.truncatula_AC150446_9.5         ........EEQKKNA..QRVVNE.................EQ.INDIKLS
L.japonicus_TC37963               ........EEPKRIT..QRDASE.................EQ.LNDIMLS
C.maculosa_TA3603_215693          .................VRSVID.................EK.VNDLTLL
S.henryi_DT604565                 ........DKELKSP..SIEGGE.................DL.SSNLNLS
H.vulgare_TC185682                ........DKVNKYM..LQGNVV.................AG.PSDSKMS
O.sativa_LOC_Os08g16930           ........GRVMRYP..MEGDFA.................AG.PSDAKMA
P.edulis_FP092509                 ........EKAKKCP..LEGGVA.................TG.PGDAKIP
P.virgatum_TC29094                ........NNVRRYP..MEGVAA.................AG.PSEVRIP
S.bicolor_Sb07g011060             ........NKSKRYP..MEGDAA.................AG.PSQVRMP
Z.mays_GRMZM2G371316_T01          ........NKSKRYP..MEG.AA.................AS.PSQMRMH
Z.mays_GRMZM2G075851_T01          ........NKSKRYP..MESDAAAGNNSKRYPMEMGSDVVAG.PTQGRMP
Z.officinale_TA6335_94328         ........E..................................ATDIKLS
Z.officinale_TA6947_94328         ........DKLRKSL..VEDNSV....................ANDLNLS
O.sativa_LOC_Os09g25410           ........KKAEQAG..VDEDIM....................TADVNLS
P.virgatum_TC30704                ........DKFDQVG..VEEEMI....................AADVKLS
S.bicolor_Sb02g024920             ........EKFDEVG..VDEDMM....................TPDVKLS
Z.mays_GRMZM2G093731_T02          ........EKLDQVE..VNEDTM....................TPDVKLS
P.glauca_BT104710                 REEAPTMSETAKKK....PPPLV....................ESDLNLS
P.sitchensis_TA13795_3332         REEAPTMSETAKKKP...PPPLV....................ESDLNLS
P.taeda_TA10646_3352              REEAPTMTETAKRKP...PTSVV....................ESDLNLS
L.saligna_TA2168_75948            ........AEAKCGE.........................DLIKAEEVAA
L.sativa_TC20185                  ........AEAKCGE.........................DLIKAEEVAA
S.chacoense_TA1669_4108           ........TEAKRGE.........................DLLTAEEMAA
S.tuberosum_AM908388              ........TEAKRGE.........................DLLTAEEMAA
S.lycopersicum_TC195266           ........TEAKRGE.........................DLLTAEEMAA
S.tuberosum_NP13064295            ........TEAKRGE.........................DLLTAEEMAP
C.maculosa_TA1609_215693          GRPDFWGGE..................................SELLVLS
H.tuberosus_EL464130              GRPDFWGGE..................................SELLVLS
L.sativa_DW131501                 GRPDFWGGE..................................SELLVLS
L.virosa_TA2198_75947             GRPDFWGGE..................................SELLVLS
L.serriola_BU011148               GRPDFWGGE..................................SELLVLS
F.vesca_TA11867_57918             QRPDFWGGE..................................SELLVLS
Z.officinale_TA4076_94328         GRSDFWGGE..................................SELLVLS
M.crystallinum_TC9929             GQSDFWGGE..................................SELLVLS
```

Fig. 17(cont.)

```
                          551                                              600
A.lyrata_944925                .ICG.................................DVE.KE..L.NRA
A.thaliana_AT4G30630           .ICG.................................DVE.KE..L.NRA
B.napus_TC91202                .ICR.................................DVG.KE..I.SRA
C.annuum_TC15926               .LCR.................................NVG.EG..T.SRV
S.lycopersicum_TC206342        .ICR.................................NVG.EG..T.SHV
S.lycopersicum_16878           .VCR.................................NVG.GD..T.SRL
C.intybus_TA970_13427          .MCR.................................NVG.DG..STNRA
C.tinctorius_TA4044_4222       .MCR.................................NVG.DG..STSRS
H.annuus_TC40508               .VCR.................................NVG.DG..STNRS
H.argophyllus_TA3915_73275     .VCR.................................NVG.DG..STNRS
T.erecta_417                   .ICR.................................NVG.DG..STNRS
E.esula_TC2982                 GACR.................................SIG.NG..S.SRA
G.hirsutum_TC136942            .ACR.................................NTG.NG..P.NRT
T.cacao_TC4923                 .ACR.................................NTG.NG..P.NRV
P.trichocarpa_578729           .ACR.................................NVG.NG..S.GWA
P.trichocarpa_scaff_VI.1304    .ACR.................................NVG.NG..S.SRA
G.max_Glyma02g36620            AACR.................................SVG.DG..S.NRA
G.max_Glyma17g08070            AACR.................................SVG.DG..S.NRA
P.vulgaris_TC10046             .ACR.................................SVG.DG..S.NRA
M.truncatula_AC150446_9.5      .VCR.................................NIG.EG..S.SRA
L.japonicus_TC37963            .VCR.................................NDG.SG..S.SRT
C.maculosa_TA3603_215693       .TCG.................................NVGRDA..STSQA
S.henryi_DT604565              .MCR.................................DVG.NG..P.NRT
H.vulgare_TC185682             .AFW.................................DAR.NP..AVARA
O.sativa_LOC_Os08g16930        .AFW.................................DVR.NA..ASSRA
P.edulis_FP092509              .AFW.................................NVR.NA..A.GRA
P.virgatum_TC29094             .AFW.................................DAR.SL..E.GRA
S.bicolor_Sb07g011060          .VFW.................................DVQ.NP..A.GRS
Z.mays_GRMZM2G371316_T01       .VFW.................................DVQ.NP..A.GRS
Z.mays_GRMZM2G075851_T01       .VFW.................................DVQ.NP..A.GRS
Z.officinale_TA6335_94328      .ICR.................................DVG.HA..SSNQV
Z.officinale_TA6947_94328      .ICR.................................DVV.HE..SSNKV
O.sativa_LOC_Os09g25410        .ACR.................................DTG.N...SPASD
P.virgatum_TC30704             .VCR.................................DTG.N...SPTSD
S.bicolor_Sb02g024920          .VCR.................................DTG.N...SPASE
Z.mays_GRMZM2G093731_T02       .VCR.................................DTG.EY..SPVSE
P.glauca_BT104710              .IFK.................................DVP.NGFLLPNRT
P.sitchensis_TA13795_3332      .IFK.................................DVP.NGFLLPNRT
P.taeda_TA10646_3352           .IFK.................................DVP.NGFLLPNRT
L.saligna_TA2168_75948         .KCR.................................ATG.....ETP..
L.sativa_TC20185               .KCR.................................ATG.....ETP..
S.chacoense_TA1669_4108        .KCR.................................ATG.....SSPK.
S.tuberosum_AM908388           .KCR.................................ATG.....SSPK.
S.lycopersicum_TC195266        .KCR.................................ATG.....SSPK.
S.tuberosum_NP13064295         .KCR.................................ATG.....SSPK.
C.maculosa_TA1609_215693       KLCRQPIVVYIPEHEHTNSGYGSGFIPI....AEYGADFGKGSKKAKAKK
H.tuberosus_EL464130           KLCRQPIVVYIPEHEHTRAYNGSGFIPI....AEYGGEFGKSSKQEKPKK
L.sativa_DW131501              KLCRQPIVVYIPEHEHSRGGYGSGFIPI....AEYGADFGKGSRKEKGKK
L.virosa_TA2198_75947          KLCRQPIVVYIPEHEHSRGGYESGFIPI....AEYGADFGKGSRKEKGKK
L.serriola_BU011148            KLCRQPIVVYIPEHEHSRGGYGSGFIPI....AEYGADFGKGSRKEKARK
F.vesca_TA11867_57918          KLCRQPIIVYIPEHEHTNGGRGSGFIPI....AEYGREFSKGSRNRKPRK
Z.officinale_TA4076_94328      KLCRQPIIVYIPEHEGMPSTELHSLTPLHNLPCCNIQEI...........
M.crystallinum_TC9929          KLCCQPIIVYIPAHELLAQIHTQEAEKMHGN................PSV
```

Fig. 17(cont.)

```
                                601             618

A.lyrata_944925                 V.KRKVTWQ..DPVALSV
A.thaliana_AT4G30630            V.KRKVTWQ..DPVALSV
B.napus_TC91202                 V.KRKVSWQ..DPVALRV
C.annuum_TC15926                E.KRKISWQ..DPVALRV
S.lycopersicum_TC206342         E.KRKISWQ..DPVALRV
S.lycopersicum_16878            G.KQKLSWQ..DPVALRV
C.intybus_TA970_13427           G.KRKITWQ..DPVALRV
C.tinctorius_TA4044_4222        G.KRKISWQ..DPVALRV
H.annuus_TC40508                G.KRKISWQ..DPVALRV
H.argophyllus_TA3915_73275      G.KRKISWQ..DPVALRV
T.erecta_417                    G.KRKISWQ..DPVALRV
E.esula_TC2982                  G.KRKISWQ..DQVALRV
G.hirsutum_TC136942             G.KRKISWQ..DQVALRA
T.cacao_TC4923                  A.KRKITWQ..DPVALRV
P.trichocarpa_578729            G.MRKIIWQ..DQVALRV
P.trichocarpa_scaff_VI.1304     G.KRKITWQ..DQVALRV
G.max_Glyma02g36620             G.KRKISWQ..DQVALRV
G.max_Glyma17g08070             G.KRKISWQ..DQVALRV
P.vulgaris_TC10046              G.KRKISWQ..DQVALRV
M.truncatula_AC150446_9.5       G.KRKITWQ..DQVALRV
L.japonicus_TC37963             SVKRKISWQ..DQVALRV
C.maculosa_TA3603_215693        R.KRKISWQ..DPVALRV
S.henryi_DT604565               G.KRKITWE..DQVALRV
H.vulgare_TC185682              G.KRKISWQ..DQLALRV
O.sativa_LOC_Os08g16930         G.KRKISWQ..DQLALRV
P.edulis_FP092509               G.KRKISWQ..DQLALRV
P.virgatum_TC29094              G.KRKISWQ..DQLALRV
S.bicolor_Sb07g011060           G.KRKISWQ..DQLALRV
Z.mays_GRMZM2G371316_T01        G.KRKISWQ..DQLALRV
Z.mays_GRMZM2G075851_T01        G.KRKISCQ..DQR~~~~
Z.officinale_TA6335_94328       G.KRKITWQ..DQVALTV
Z.officinale_TA6947_94328       A.KRKITWQ..DQLALRV
O.sativa_LOC_Os09g25410         D.KRSVSWQ..DRVALKV
P.virgatum_TC30704              D.KRAIPWQ..DQMALKA
S.bicolor_Sb02g024920           G.KRTVPWQ..DRLALKA
Z.mays_GRMZM2G093731_T02        G.KRAVPWQ..DRFALKT
P.glauca_BT104710               G.KRKISWQ..DQMALKV
P.sitchensis_TA13795_3332       G.KRKISWQ..DQMALKV
P.taeda_TA10646_3352            G.KRKISWQ..DQMALKV
L.saligna_TA2168_75948          ..KKLLGW.........F
L.sativa_TC20185                ..KKLLGW.........F
S.chacoense_TA1669_4108         ..KPLLGC.........F
S.tuberosum_AM908388            ..KPLLGC.........F
S.lycopersicum_TC195266         ..KPLLGC.........F
S.tuberosum_NP13064295          ..KPLLGC.........F
C.maculosa_TA1609_215693        AVRLLYSGR..NHYDLLV
H.tuberosus_EL464130            AVRLLYSGS..NHYDLLV
L.sativa_DW131501               AVKLLYSGR..NHYDLLV
L.virosa_TA2198_75947           AVRLLYSGR..NHYDLLV
L.serriola_BU011148             ................L
F.vesca_TA11867_57918           AVRLLFSGR..NHYDLLV
Z.officinale_TA4076_94328       ..LQAYTWRMGQWIYSYS
M.crystallinum_TC9929           APKSSIFIRVAMALVLFR
```

Fig. 17(cont.)

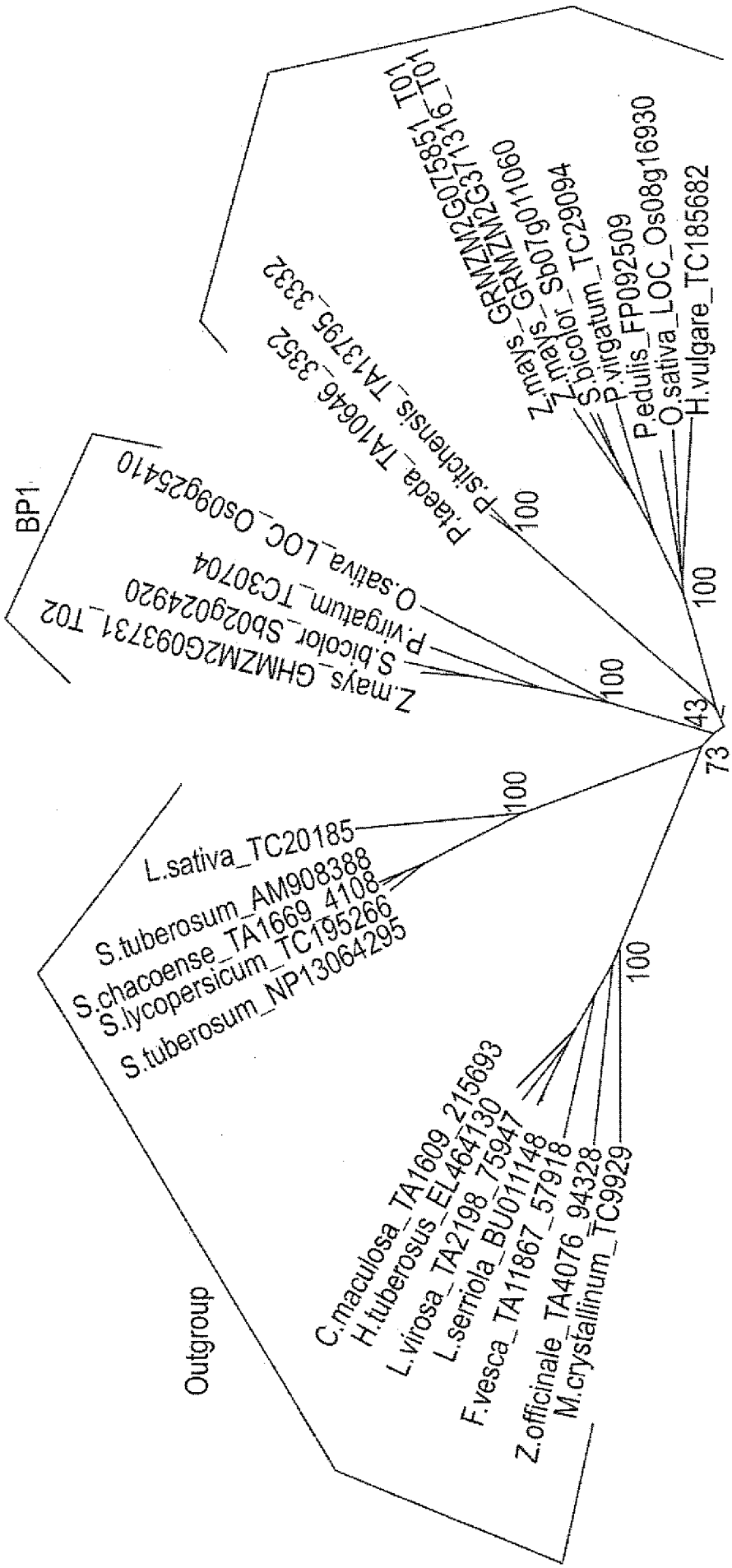

H.vulgare_TC185682
O.sativa_LOC_Os08g16930
P.edulis_FP092509
P.virgatum_TC29094
S.bicolor_TC29094
Z.mays_GRMZM2G071060
Z.mays_GRMZM2G371396_T01
P.laeda_TA10646_3352
P.sitchensis_TA13795_3332
O.sativa_LOC_Os09g25410
P.virgatum_TC307O4
S.bicolor_Sb02g024920
Z.mays_GHMZM2G093731_T02
BP1
S.tuberosum_NP13064295
S.lycopersicum_TC195266
S.chacoense_TA1669_4108
S.tuberosum_AM908388
L.sativa_TC20185
Outgroup
C.maculosa_TA1609_215693
H.tuberosus_EL464130
L.virosa_TA2198_75947
L.serriola_BU011148
F.vesca_TA11867_57918
Z.officinale_TA4076_94328
M.crystallinum_TC9929
100
100
100
43
73
100

Fig. 18

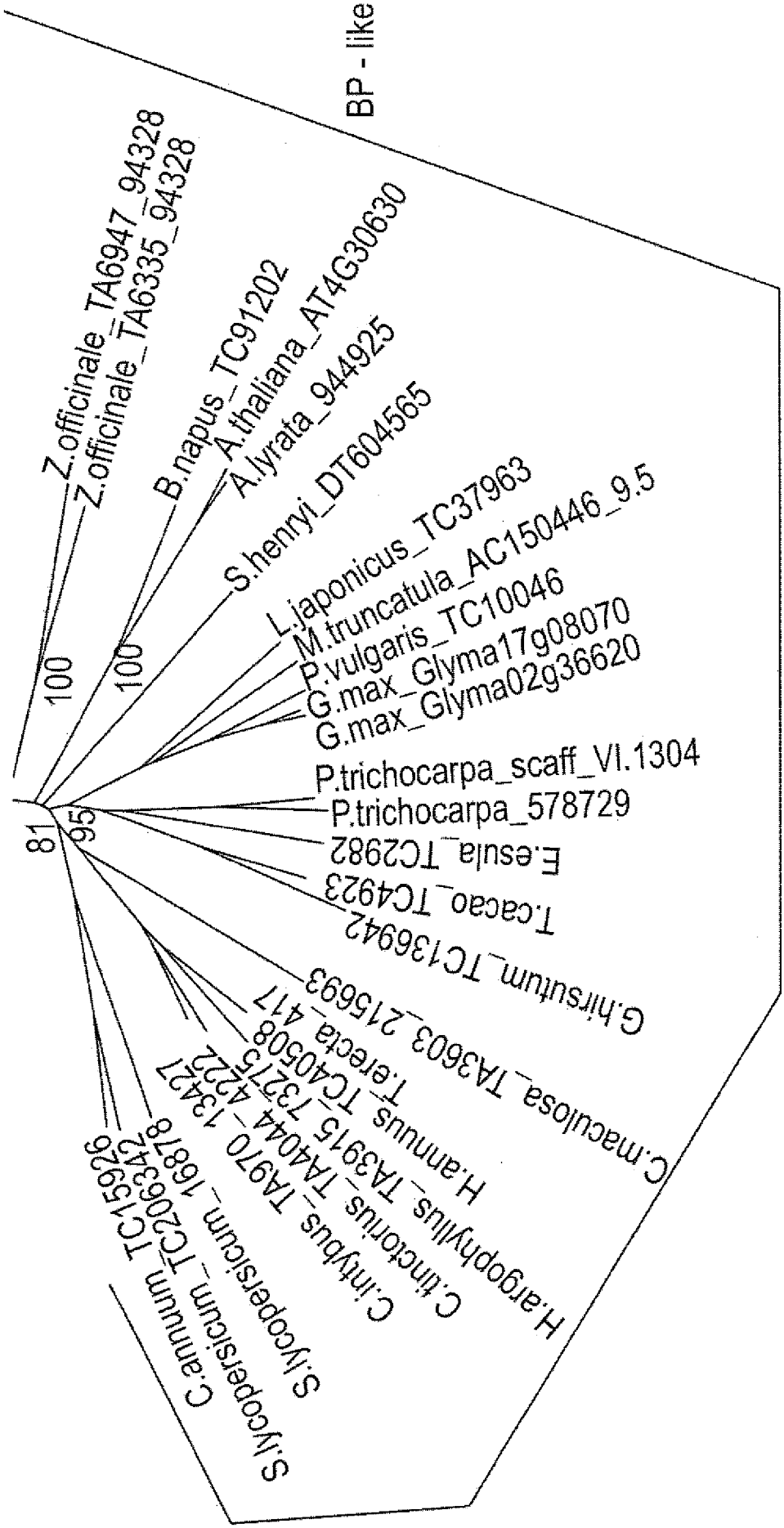
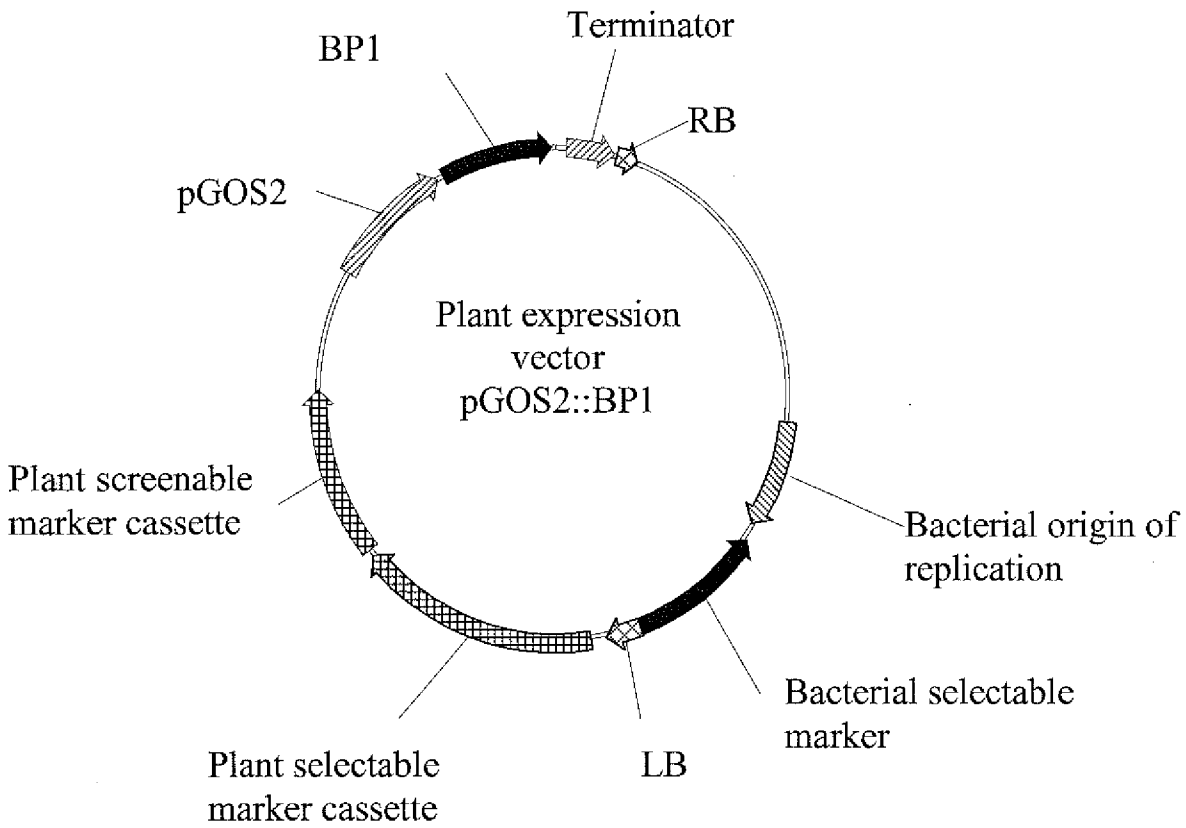

BP - like
Z.officinale_TA6947_94328
Z.officinale_TA6335_94328
B.napus_TC91202
A.thaliana_AT4G30630
A.lyrata_944925
S.henryi_DT604565
L.japonicus_TC37963
M.truncatula_AC150446_9.5
P.vulgaris_TC10046
G.max_Glyma17g08070
G.max_Glyma02g36620
P.trichocarpa_scaff_VI.1304
P.trichocarpa_578729
E.esula_TC2982
T.cacao_TC4923
G.hirsutum_TC136942
C.maculosa_TA3603_215693
H.annuus_TC40508
T.erecta_417
H.argophyllus_TA3915_TC32715
C.tinctorius_TA4044_4222
C.intybus_TA970_13421
S.lycopersicum_TC259263
S.lycopersicum_168782
C.annuum_TC159269
100
100
81
95

Fig. 18 (cont)


BP1
Terminator
RB
pGOS2
Plant expression vector pGOS2::BP1
Plant screenable marker cassette
Bacterial origin of replication
Bacterial selectable marker
Plant selectable marker cassette
LB

	1	2	3	4
1. SEQ ID NO: 171		39.5	40.2	39.6
2. SEQ ID NO: 239	51.7		65.6	64.9
3. SEQ ID NO: 243	52.5	77.5		80.6
4. SEQ ID NO: 267	50.6	77.1	85.5	

Fig. 20

INTERNATIONAL SEARCH REPORT

International application No.
PCT/IB2012/050969

A. CLASSIFICATION OF SUBJECT MATTER

See extra sheet

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC: C12N; A01H; C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNABS, CPRSABS, JPABS, CNMED, DWPI, SIPOABS, CPEA, CNTXT, TWTXT, WOTXT, EPTXT, USTXT, CATXT, CNKI, ISI WEB OF KNOWLEDGE, BAIDU, GOOGLE SCHOLAR and keywords: TLP, TIFY, ZIM, Jasmonate zim domain protein, JAZ3, TIFY6b, AT3G17860, PMP22, RTF, REM like transcription factor, AT2G24700, REM10, AtREM10, transcriptional factor B3 family protein, BP1, Os09g25410, Os09g0421500, yield, trait?, plant? et al.
GENBANK, EMBL: sequence search on SEQ ID NOs: 1-4, 50-51, 139-140 and 170-171

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2004035798A2 (CROPDESIGN NV ET AL.) 29 Apr. 2004(29.04.2004) see abstract, claims 1-39, page 17, lines 5-6, page 20, lines 15-19, page 29, line 32 - page 30, line 30, pages 69-76, example 12 and page 96, Table 4 of the description, and SEQ ID NOs: 19-20 and 1271-1272	1-21, 27-28, 31-39, 41-45
A	see the whole document	22-26, 29-30, 40

☒ Further documents are listed in the continuation of Box C. ☒ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim (S) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&"document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
25 Jun. 2012(25.06.2012)	**19 Jul. 2012 (19.07.2012)**
Name and mailing address of the ISA/CN The State Intellectual Property Office, the P.R.China 6 Xitucheng Rd., Jimen Bridge, Haidian District, Beijing, China 100088 Facsimile No. 86-10-62019451	Authorized officer LIN, Junkai Telephone No. (86-10)62412317

Form PCT/ISA /210 (second sheet) (July 2009)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/IB2012/050969

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
	JP 2008271805A (DOKURITSU GYOSEI HOJIN RIKAGAKU KENKYUS) 13 Nov. 2008(13.11.2008)	
Y	see abstract, claims 1-14	1-16
A	see the whole document	17-45
	CHUNG, H. S. et al. Top hits in contemporary JAZ: An update on jasmonate signaling. Phytochemistry. 1 October 2009(01.10.2009), vol. 70, nos. 13-14, pages 1547-1559, ISSN 0031-9422	
Y	see page 1549, left column, paragraph 2 - right column, paragraph 3, figure 1 and table 2	1-16
	GenBank Accession No. AK224691.2, 1 May 2010(01.05.2010), [retrieved on 25 Jun. 2012(25.06.2012)]. Retrieved from: GenBank database	
X	see the nucleic acid sequence	17-18
	GenBank Accession No. EU194561.1, 15 May 2008(15.05.2008), [retrieved on 25 Jun. 2012(25.06.2012)]. Retrieved from: GenBank database	
X	see the nucleic acid sequence and the amino acid sequence	17-18
Y	see the whole document	1-16
	GenBank Accession No. BT014218.1, 11 May 2004(11.05.2004), [retrieved on 25 Jun. 2012(25.06.2012)]. Retrieved from: GenBank database	
X	see the nucleic acid sequence	44-45
	GenBank Accession No. CAW52912.1, 07 Jan. 2009(07.01.2009), [retrieved on 25 Jun. 2012(25.06.2012)]. Retrieved from: GenBank database	
A	see the whole document	19, 22-24, 27, 29, 31-38, 40-43
	SWAMINATHAN, K. et al. The plant B3 superfamily. Trends in Plant Science. 3 November 2008(03.11.2008), vol. 13, no. 12, pages 647-655, ISSN 1360-1385	
A	see the whole document	19, 22-24, 27, 29, 31-38, 40-43

INTERNATIONAL SEARCH REPORT

International application No.
PCT/IB2012/050969

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
	GenBank Accession No. NP_001063204.1, 08 Jun. 2010(08.06.2010), [retrieved on 25 Jun. 2012(25.06.2012)]. Retrieved from: GenBank database	
A	see the whole document	19, 25-27, 30-39, 41-43
	YE, H. Y. et al. Identification and expression profiling analysis of TIFY family genes involved in stress and phytohormone responses in rice. Plant Mol Biol. 2009, vol. 71, No. 3, pages 291-305, ISSN: 0167-4412	
A	see the whole document	1-45
	WO 2010086221A1 (BASF PLANT SCI CO GMBH) 05 Aug. 2010(05.08.2010)	
A	see the whole document	1-45

INTERNATIONAL SEARCH REPORT

International application No.
PCT/IB2012/050969

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

See extra sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of additional fee.

3. ☒ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
1-7, 9-32, 34-45 (all partially), and 8, 33 (inventions 1-2, 18, 56 and 69)

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on protest

☒ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/IB2012/050969

Continuation of: Box No. III

Because the TLP polypeptides listed in Table A1 represent alternatives of a Markush grouping, the Authority thinks that the unity of inventions 1-17 of Group I shall be considered based on the "Markush practice". For similar reasons, the unity of inventions 18-55 of Group II, the unity of inventions 56-68 of Group III and the unity of inventions 69-121 of Group IV shall also be considered based on the "Markush practice".

The Authority finds that the application comprises 121 separate inventions:

Group I:

1: claims 1-7, 9-18 (all partially) and 8

a TLP (Tify like protein) polypeptide (SEQ ID NO: 2) and subject-matter relating thereto;

2-17: claims 1-7, 9-18 (all partially)

a TLP polypeptide according to SEQ ID NOs: 3-34 even numbers, respectively, and subject-matter relating thereto;

Group II:

18: claims 19-21, 27-28, 31-39 and 41-45 (all partially)

a PMP22 (22 kDa Peroxisomal Membrane like protein) polypeptide (SEQ ID NO: 51) and subject-matter relating thereto;

19-55: claims 19-21, 27-28, 31-39 and 41-45 (all partially)

a PMP22 polypeptide according to SEQ ID NOs: 52-125 odd numbers, respectively, and subject-matter relating thereto;

Group III:

56: claims 19, 22-24, 27, 29, 31-38 and 40-43 (all partially)

a RTF (REM-like transcription factor) polypeptide (SEQ ID NO: 140) and subject-matter relating thereto;

57-68: claims 19, 22-24, 27, 29, 31-38 and 40-43 (all partially)

a RTF polypeptide according to SEQ ID NOs: 141-164 even numbers, respectively, and subject-matter relating thereto;

Group IV:

69: claims 19, 25-27, 30-39 and 41-43 (all partially)

a BP1 polypeptide (SEQ ID NO: 171) and subject-matter relating thereto;

70-121: claims 19, 25-27, 30-39 and 41-43 (all partially)

a BP1 polypeptide according to SEQ ID NOs: 172-275 odd numbers, respectively, and subject-matter relating thereto.

Detailed lack of unity reasoning concerning the present application:

The present application concerns proteins potentially useful for enhancing yield-related traits in plants. In the light of the prior art (e.g. WO 2010086221A1), the underlying problem to be solved by the present application is seen in the provision of further proteins potentially useful for enhancing yield-related traits in plants. The problem is solved by providing the structurally completely different polypeptides of Groups I-IV.

However, this concept cannot be regarded as the "single general inventive concept" required by Rules 13.1 PCT because it is neither novel nor inventive, since such proteins potentially useful for enhancing yield-related traits in plants are already known in the art (e.g. WO 2010086221A1).

INTERNATIONAL SEARCH REPORT

International application No.
PCT/IB2012/050969

Continuation of: Box No. III

As for Group I, the TLP polypeptides listed in Table A1 represents the alternatives of "Markush practice", although all alternatives (i.e. TLP polypeptides) are alleged to have a common property or activity (i.e. modulating expression in a plant of a nucleic acid encoding a TLP polypeptide to enhancing yield-related traits in plants relative to control plants), they do not share a common chemical structure which occupies a large portion of their structure, and the description does not provide any evidence to show that the small portion of commonly shared structure thereof (i.e. the domain Q[LM]TIFY, see also claim 3) is essential to said common property or activity. Therefore, inventions 1-17 of Group I lack unity of invention.

As for Group II, the PMP22 polypeptides listed in Table A2 represents the alternatives of "Markush practice", although all alternatives (i.e. PMP22 polypeptides) are alleged to have a common property or activity (i.e. modulating expression in a plant of a nucleic acid encoding a PMP22 polypeptide to enhancing yield-related traits in plants relative to control plants), they do not share a common chemical structure which occupies a large portion of their structure, and the description does not provide any evidence to show that the small portion of commonly shared structure thereof (i.e. the motifs shared by the PMP22 polypeptides, see also claim 21) is essential to said common property or activity. Therefore, inventions 18-55 of Group II lack unity of invention.

As for Group III, the RTF polypeptides listed in Table A3 represents the alternatives of "Markush practice", although all alternatives (i.e. RTF polypeptides) are alleged to have a common property or activity (i.e. modulating expression in a plant of a nucleic acid encoding a RTF polypeptide to enhancing yield-related traits in plants relative to control plants), they do not share a common chemical structure which occupies a large portion of their structure, and the description does not provide any evidence to show that the small portion of commonly shared structure thereof (i.e. the motifs shared by the RTF polypeptides, see also claim 23) is essential to said common property or activity. Therefore, inventions 56-68 of Group III lack unity of invention.

As for Group IV, the BP1 polypeptides listed in Table A4 represents the alternatives of "Markush practice", although all alternatives (i.e. BP1 polypeptides) are alleged to have a common property or activity (i.e. modulating expression in a plant of a nucleic acid encoding a BP1 polypeptide to enhancing yield-related traits in plants relative to control plants), they do not share a common chemical structure which occupies a large portion of their structure, and the description does not provide any evidence to show that the small portion of commonly shared structure thereof (i.e. the motifs shared by the BP1 polypeptides, see also claim 25) is essential to said common property or activity. Therefore, inventions 69-121 of Group IV lack unity of invention.

In conclusion, the 121 groups of claims are not linked by common or corresponding special technical features and define 121 separate inventions not linked by a single general inventive concept. The application, hence does not meet the requirements of unity of invention as defined in Rules 13.1 and 13.2 PCT.

As only 4 additional search fees have been paid, this international search report covers only those inventions for which fees were paid, i.e. inventions 1-2, 18, 56 and 69.

Continuation of: CLASSIFICATION OF SUBJECT MATTER

C12N 15/82 (2006.01) i
A01H 5/00 (2006.01) i
C12N 15/29 (2006.01) i
C07K 14/415 (2006.01) i

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/IB2012/050969

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
WO 2004035798A2	29.04.2004	AU 2003298095A1	04.05.2004
		EP 1551983A2	13.07.2005
		US 2006021088A1	26.01.2006
		AU 2003298095A8	03.11.2005
		US 7847156B2	07.12.2010
		EP 2302062A1	30.03.2011
		EP 2316953A2	04.05.2011
		US 2011162107A1	30.06.2011
		EP 2302062A8	03.08.2011
		EP 2316953A3	05.10.2011
		WO 2004035798A3	04.11.2004
JP 2008271805A	13.11.2008	None	
WO 2010086221A1	05.08.2010	AU 2010209875A1	05.08.2010
		CA 2750007A1	05.08.2010
		MXPA 11007599A	31.08.2011
		AU 2010209875A2	05.08.2010
		US 2011283420A1	17.11.2011
		EP 2391719A1	07.12.2011