



(86) Date de dépôt PCT/PCT Filing Date: 2012/02/27
(87) Date publication PCT/PCT Publication Date: 2012/09/07
(85) Entrée phase nationale/National Entry: 2013/08/08
(86) N° demande PCT/PCT Application No.: IB 2012/050888
(87) N° publication PCT/PCT Publication No.: 2012/117330
(30) Priorités/Priorities: 2011/02/28 (EP11156182.5);
2011/02/28 (US61/447151); 2011/03/25 (EP11159809.0);
2011/03/25 (US61/467423); 2011/05/06 (EP11165187.3);
2011/05/06 (US61/483115)

(51) Cl.Int./Int.Cl. *C12N 15/82* (2006.01),
A01H 5/00 (2006.01), *C07K 14/415* (2006.01),
C12N 15/29 (2006.01), *C12N 9/00* (2006.01)
(71) Demandeur/Applicant:
BASF PLANT SCIENCE COMPANY GMBH, DE
(72) Inventeurs/Inventors:
FRANKARD, VALERIE, BE;
REUZEAU, CHRISTOPHE, FR;
HATZFELD, YVES, FR;
VANDENABEELE, STEVEN, BE;
MIRONOV, VLADIMIR, NO
(74) Agent: ROBIC

(54) Titre : PLANTES PRESENTANT DE MEILLEURES CARACTERISTIQUES ASSOCIEES AU RENDEMENT ET
PROCEDES D'OBTENTION DES PLANTES
(54) Title: PLANTS HAVING ENHANCED YIELD-RELATED TRAITS AND PRODUCING METHODS THEREOF

(57) **Abrégé/Abstract:**

Provided are a method for enhancing yield-related traits in plants by modulating expression in a plant of a nucleic acid encoding a WAK-like polypeptide, or a CDKB-RKA polypeptide, or a UPA20-like polypeptide, and plants having modulated expression of a nucleic acid encoding a WAK-like polypeptide, or a CDKB-RKA polypeptide, or a UPA20-like polypeptide, which plants have enhanced yield-related traits relative to control plants. Also provided are WAK-like-encoding nucleic acids, or CDKB-RKA-encoding nucleic acids, or UPA20-like-encoding nucleic acids, and constructs comprising the same, useful in performing the methods for enhancing yield-related traits in plants.



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

(19) World Intellectual Property

Organization

International Bureau



WIPO | PCT



(10) International Publication Number

WO 2012/117330 A1

(43) International Publication Date

7 September 2012 (07.09.2012)

(51) International Patent Classification:

C12N 15/82 (2006.01) C07K 14/415 (2006.01)

A01H 5/00 (2006.01) C12N 15/29 (2006.01)

C12N 9/00 (2006.01)

9700 Oudenaarde (BE). MIRONOV, Vladimir [BE/NO];

Nordahl Brunsvei 2A, N-7052 Trondheim (NO).

(74) Agent: POPP, Andreas; BASF SE, Global Intellectual Property GVX/B - C 006, 67056 Ludwigshafen (DE).

(21) International Application Number:

PCT/IB2012/050888

(22) International Filing Date:

27 February 2012 (27.02.2012)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

61/447151 28 February 2011 (28.02.2011) US

11156182.5 28 February 2011 (28.02.2011) EP

61/467423 25 March 2011 (25.03.2011) US

11159809.0 25 March 2011 (25.03.2011) EP

61/483115 6 May 2011 (06.05.2011) US

11165187.3 6 May 2011 (06.05.2011) EP

(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))

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

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

(72) Inventors; and

(75) Inventors/Applicants (for US only): FRANKARD, Valerie [BE/BE]; Avenue des Croix du Feu 84, 1410 Waterloo (BE). REUZEAU, Christophe [FR/FR]; rue de Cimetière, F-24350 La Chapelle Gonaguet (FR). HATZFELD, Yves [FR/FR]; 6B rue du Chemin de Fer, F-59000 Lille (FR). VANDENABEELE, Steven [BE/BE]; Meersport 23, B-

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

(57) Abstract: Provided are a method for enhancing yield-related traits in plants by modulating expression in a plant of a nucleic acid encoding a WAK-like polypeptide, or a CDKB-RKA polypeptide, or a UPA20-like polypeptide, and plants having modulated expression of a nucleic acid encoding a WAK-like polypeptide, or a CDKB-RKA polypeptide, or a UPA20-like polypeptide, which plants have enhanced yield-related traits relative to control plants. Also provided are WAK-like-encoding nucleic acids, or CDKB-RKA-encoding nucleic acids, or UPA20-like-encoding nucleic acids, and constructs comprising the same, useful in performing the methods for enhancing yield-related traits in plants.



WO 2012/117330 A1

PLANTS HAVING ENHANCED YIELD-RELATED TRAITS AND PRODUCING METHODS THEREOF

Background

The present invention relates generally to the field of molecular biology and concerns a method for enhancing yield-related traits in plants by modulating expression in a plant of a nucleic acid encoding a WAK-like (Wall-Associated Kinase-like) polypeptide, or a UPA20-like polypeptide. The present invention also concerns plants having modulated expression of a nucleic acid encoding a WAK-like polypeptide or a UPA20-like polypeptide, which plants have enhanced yield-related traits relative to corresponding wild type plants or other control plants. The invention also provides constructs useful in the methods of the invention.

Concerning CDKB-RKA polypeptides, the present invention relates generally to the field of molecular biology and concerns a method for enhancing various economically important yield-related traits in plants. More specifically, the present invention concerns a method for enhancing yield-related traits in plants by modulating expression in a plant of a nucleic acid encoding a CDKB-RKA polypeptide wherein said CDKB-RKA is a modified CDKB polypeptide which CDKB-RKA polypeptide has modified, preferably reduced, more preferably lacks catalytic kinase activity in comparison to a full length or unmodified CDKB polypeptide. More in particular, the present invention concerns a method for enhancing yield-related traits in plants by introducing and expressing in said plant of a nucleic acid encoding a CDKB-RKA polypeptide as defined herein. The present invention also concerns plants having modulated expression of a nucleic acid encoding said CDKB-RKA polypeptide, which plants have enhanced yield-related traits relative to control plants. The invention also provides constructs comprising the CDKB-RKA polypeptide as defined herein that are useful in performing the methods of the invention.

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

A trait of particular economic interest is increased yield. Yield is normally defined as the measurable produce of economic value from a crop. This may be defined in terms of

quantity and/or quality. Yield is directly dependent on several factors, for example, the number and size of the organs, plant architecture (for example, the number of branches), seed production, leaf senescence and more. Root development, nutrient uptake, stress tolerance and early vigour may also be important factors in determining yield. Optimizing the abovementioned factors may therefore contribute to increasing crop yield.

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

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

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

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

Concerning WAK polypeptides, WAK polypeptides belong to the receptor-like kinase (RLK) superfamily, with from N-terminus to C-terminus an EGF domain, a transmembrane domain and a tyrosine kinase domain. They are membrane proteins, usually targeted to the plasma membrane. WAK polypeptides are covalently bound to the cell wall and they transmit

information from the cell wall to the cell.

Zhang et al. (2005) identified that there are at least 125 WAKs in rice, compared to 26 WAK and WAK-like in Arabidopsis. Zhang et al. (2005) also identified via sequence analysis, 3 major classes of WAK and WAK-like polypeptides: WAK-RLK with the classical structure with from N-terminus to C-terminus an EGF domain, a transmembrane domain and a tyrosine kinase domain; WAK-RLCK which only has the kinase domain, but no EGF domain and no transmembrane domain; and WAK-RLP comprises the EGF domain and the transmembrane domain, but has no kinase domain.

Concerning CDKB-RKA polypeptides, the inherent growth mechanisms of a plant reside in a highly ordered sequence of events collectively known as the 'cell cycle'. Progression through the cell cycle is fundamental to the growth and development of all multicellular organisms and is crucial to cell proliferation. The major components of the cell cycle are highly conserved in yeast, mammals, and plants. The cell cycle is typically divided into the following sequential phases: G0 – G1 – S – G2 – M. DNA replication or synthesis generally takes place during the S phase ("S" is for DNA synthesis) and mitotic segregation of the chromosomes occurs during the M phase (the "M" is for mitosis), with intervening gap phases, G1 (during which cells grow before DNA replication) and G2 (a period after DNA replication during which the cell prepares for division). Cell division is completed after cytokinesis, the last step of the M phase. Cells that have exited the cell cycle and that have become quiescent are said to be in the G0 phase. Cells in this phase can be stimulated to reenter the cell cycle at the G1 phase. The "G" in G1, G2 and G0 stands for "gap". Completion of the cell cycle process allows each daughter cell during cell division to receive a full copy of the parental genome.

Cell division is controlled by two principal cell cycle events, namely initiation of DNA synthesis and initiation of mitosis. Each transition to each of these key events is controlled by a checkpoint represented by specific protein complexes (involved in DNA replication and division). The expression of genes necessary for DNA synthesis at the G1/S boundary is regulated by the E2F family of transcription factors in mammals and plant cells (La Thangue, 1994 (Curr. Opin. Cell Biol. 6 (3), 443 – 450); Muller *et al.*, 2001 (Genes Dev. 15 (3), 267 – 285); De Veylder *et al.*, 2002 (EMBO J. 21 (6), 1360 - 1368)). Entry into the cell cycle is regulated/triggered by an E2F/Rb complex that integrates signals and allows activation of transcription of cell cycle genes. The transition between the different phases of the cell cycle, and therefore progression through the cell cycle, is driven by the formation and activation of different heterodimeric serine/threonine protein kinases, generally referred to as cyclin-dependent kinases (CDKs). A prerequisite for activity of these kinases is the physical association with a specific cyclin, the timing of activation being largely dependent upon cyclin expression. Cyclin binding induces conformational changes in the N-terminal lobe of the associating CDK and contributes to the localisation and substrate specificity of the complex. Monomeric CDKs are activated when they are associated with cyclins and

thus have kinase activity. Cyclin protein levels fluctuate in the cell cycle and therefore represent a major factor in determining timing of CDK activation. The periodic activation of these complexes containing cyclins and CDK during cell cycle mediates the temporal regulation of cell-cycle transitions (checkpoints). Other factors regulating CDK activity include CDK inhibitors (CKIs or ICKs, KIPs, CIPs, INKs), CDK activating kinases (CAKs), a CDK phosphatase (Cdc25) and a CDK subunit (CKS) (Mironov *et al.* 1999 (Plant Cell 11 (4), 509 – 522); Reed 1996 (Progression Cell cycle Research 2, 15 – 27)).

In plants, two major classes of CDKs, known as A-type and B-type CDKs, have been studied to date. The A-type CDKs regulate both the G1-to-S and G2-to-M transitions, whereas the B-type CDKs seem to control the G2-to-M checkpoint only (Hemerly *et al.*, 1995 (EMBO J. 14 (16), 3925 – 3936); Magyar *et al.*, 1997 (Plant Cell 9 (2), 223 – 235); Porceddu *et al.*, 2001 (J. Biol. Chem. 276 (39) 36354 – 36360)). In addition, the presence of C-type CDKs and CDK-activating kinases (CAKs) has been reported (Magyar *et al.*, 1997 (Plant Cell 9 (2), 223 – 235); Umeda *et al.*, 1998 (Proc Natl Acad Sci U S A. 95 (9), 5021 - 5026.; Joubès *et al.*, 2001 (Plant Physiol. 126 (4), 1403 – 1415)), as has the presence of D-type, E-type and F-type CDKs (Vandepoele *et al.* 2002 (Plant Cell 14 (4), 903 – 916)).

Use of CDKB genes and CDKB polypeptides for increasing yield in plants has been disclosed in the prior art. For instance, US application US20070214517 discloses the use of various genes, including for instance a CDKB gene for increasing yield in plants. In another example, WO2005/024029 describes methods for improving growth characteristics of plants which comprise introducing and expressing in a plant genes encoding A or B-type cyclin-dependent kinases or variants thereof. B-type CDK nucleic acids or variants thereof are defined in WO 2005/024029 as nucleic acid/genes encoding a B-type CDK protein having: (i) a PPTALRE motif with no mismatches or with a mismatch at position 2 and/or 4 from left to right; (ii) a catalytic kinase domain; and (iii) a T-loop activation kinase domain (Magyar *et al.*, 1997 (Plant Cell 9 (2), 223 – 235)).

Concerning UPA20-like polypeptides, “UPA” stands for “Upreregulated by the bacterial effector protein AvrBs3”.

AvrBs3 is an effector protein produced by *Xanthomonas* bacteria. The *Xanthomonas* type III effector protein AvrBs3 was reported to modulate plant gene expression and to induce cell hypertrophy in a susceptible host plant by Marois *et al.* (2002, MPMI Vol. 15, No. 7, pp. 637–646). *Xanthomonas campestris* pv. *vesicatoria* bacteria expressing the type III effector protein AvrBs3 induce a hypersensitive response in pepper plants carrying the resistance gene *Bs3*. Marois *et al.* (2002) described that infection of susceptible pepper and tomato plants leads to an AvrBs3-dependent hypertrophy of the mesophyll tissue. *Agrobacterium*-mediated transient expression of the *avrBs3* gene in tobacco and potato plants resulted in a similar phenotype. Induction of hypertrophy was shown to depend on the repeat region, nuclear localization signals, and acidic transcription activation domain (AAD) of AvrBs3,

suggesting that the effector modulates the host's transcriptome. To search for host genes regulated by AvrBs3 in an AAD-dependent manner, the authors performed a cDNA-amplified fragment length polymorphism analysis of pepper mRNA populations. Thirteen AvrBs3-induced transcripts were identified and confirmed by reverse transcriptase-polymerase chain reaction. Sequence analysis revealed homologies to auxin-induced and expansin-like genes, which play a role in cell enlargement. These results suggested that some of the AvrBs3-induced genes may be involved in hypertrophy development and that *Xanthomonas* possess type III effectors that steer host gene expression.

Kay et al. (2007, Science, Vol 318, 648-651) showed that AvrBs3 induces the expression of a master regulator of cell size (a regulator of cell enlargement), *upa20*, which encodes a transcription factor containing a basic helix-loop-helix domain. They further reported that AvrBs3 binds to a conserved element in the *upa20* promoter via its central repeat region and induces gene expression through its activation domain. More in particular, to isolate *upa* genes that are direct targets of AvrBs3 Kay et al. (2007) infected susceptible pepper plants [cultivar Early Calwonder (ECW)] with *Xanthomonas* strain 85-10 expressing *avrBs3* or carrying an empty vector in the presence of cycloheximide, which blocks eukaryotic protein synthesis. cDNA fragments corresponding to AvrBs3-induced pepper genes were identified by the authors by suppression-subtractive hybridization and confirmed by reverse Northern analysis and reverse transcription polymerase chain reaction (RT-PCR). Among these cDNAs, the authors detected one gene, designated *upa20*, that encodes a putative transcription factor of the basic helix-loop-helix (bHLH) family. Kay et al (2007) described that *upa20* has a complex gene structure comprising eight exons and seven introns and encodes a putative 340-amino acid bHLH transcription factor with a predicted molecular mass of 37.8 kD. The bHLH domain, which generally serves as DNA binding and dimerization domain was reported to be located in the region from amino acids 167 to 225. *Agrobacterium*-mediated *upa20* expression demonstrated that Upa20 alone induces hypertrophy in *N. benthamiana* and other solanaceous plants (Kay et al. 2007). In *upa20*-expressing tissue of *N. benthamiana*, palisade and spongy parenchyma cells were strongly enlarged as compared with those in tissue of the control. The hypertrophy reaction was phenotypically similar to tissue transiently expressing *avrBs3* but was faster and stronger. In addition, N- and C-terminal green fluorescent protein (GFP) fusions of Upa20 were exclusively localized to the nucleus, where they often concentrated in distinct foci (Kay et al. 2007).

UPA20-like polypeptides are part of the family of the basic/helix-loop-helix (bHLH) proteins. Basic helix-loop-helix proteins (bHLH) are a group of eukaryotic transcription factors that exert a determinative influence in a variety of developmental pathways. These transcription factors are characterised by a highly evolutionary conserved bHLH domain that mediates specific dimerisation. They facilitate the conversion of inactive monomers to trans-activating dimers at appropriate stages of development. Toledo-Ortiz and Quail (2003, The Plant Cell, Vol. 15, 1749-1770) described that bHLH proteins are a superfamily of transcription factors

that bind as dimers to specific DNA target sites and that have been well characterized in non-plant eukaryotes as important regulatory components in diverse biological processes. Toledo-Ortiz and Quail (2003) performed a comprehensive computational analysis of the Arabidopsis genome sequence databases to define the scope and features of the bHLH family. Using a set of criteria derived from a previously defined consensus motif, the authors identified 147 bHLH protein-encoding genes, making this one of the largest transcription factor families in Arabidopsis. Phylogenetic analysis of the bHLH domain sequences permits classification of these genes into 21 subfamilies.

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 WAK-like polypeptide, or a UPA20-like polypeptide in a plant.

Concerning CDKB-RKA polypeptides, 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 CDKB-RKA polypeptide wherein said CDKB-RKA is a modified CDKB polypeptide which CDKB-RKA polypeptide has modified, preferably reduced, catalytic kinase activity in comparison to a full length or unmodified CDKB polypeptide.

Detailed description of the invention

1. WAK polypeptides and UPA20-like polypeptides

Concerning WAK-like polypeptides, the present invention shows that modulating expression in a plant of a nucleic acid encoding a WAK-like polypeptide gives plants having enhanced yield-related traits relative to control plants.

Concerning UPA20-like polypeptides, the present invention shows that modulating expression in a plant of a nucleic acid encoding a UPA20-like polypeptide gives plants having enhanced yield-related traits relative to control plants.

According to a first embodiment, the present invention provides a method for enhancing yield-related traits in plants relative to control plants, comprising modulating expression in a plant of a nucleic acid encoding a WAK-like polypeptide, or a UPA20-like polypeptide, and optionally selecting for plants having enhanced yield-related traits. According to another embodiment, the present invention provides a method for producing plants having

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 WAK-like polypeptide, or a UPA20-like polypeptide, as described herein and optionally selecting for plants having enhanced yield-related traits.

5

A preferred method for modulating, preferably increasing expression of a nucleic acid encoding a WAK-like polypeptide, or a UPA20-like polypeptide, is by introducing and expressing in a plant a nucleic acid encoding a WAK-like polypeptide, or a UPA20-like polypeptide.

10

Any reference hereinafter to a "protein useful in the methods of the invention" is taken to mean a WAK-like polypeptide, or a UPA20-like polypeptide, as defined herein. Any reference hereinafter to a "nucleic acid useful in the methods of the invention" is taken to mean a nucleic acid capable of encoding such a WAK-like polypeptide, or a UPA20-like polypeptide. The nucleic acid to be introduced into a plant (and therefore useful in performing the methods of the invention) is any nucleic acid encoding the type of protein which will now be described, hereafter also named "*WAK-like* nucleic acid", or "*UPA20-like* nucleic acid", or "*WAK-like* gene", or "*UPA20-like* gene". Concerning UPA20-like polypeptides, the term UPA20-like and UPA20 are used herein as synonyms.

20

A "WAK-like polypeptide" as defined herein refers to any polypeptide comprising at least an EGF domain and a transmembrane domain. The EGF domain can be defined as a Pattern-Scan PS01187 domain. Preferably, a WAK-like polypeptide according to the present invention, further comprises a tyrosine kinase domain.

25

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

Additionally or alternatively, the homologue of a WAK-like protein has in increasing order of preference at least 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% overall sequence identity to the amino acid represented by SEQ ID NO: 2 or 4, provided that the homologous protein comprises at least an EGF domain and a transmembrane 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 or 4.

40

Compared to overall sequence identity, the sequence identity will generally be higher when only conserved domains or motifs are considered.

5 A “UPA20-like polypeptide” as defined herein refers to any polypeptide comprising a bHLH domain. In a preferred embodiment, said UPA20-like gene encodes a transcription factor containing a basic helix-loop-helix domain.

10 A helix-loop-helix (HLH) DNA-binding domain consists of a closed bundle of four helices in a left-handed twist with two crossover connections. The HLH domain directs dimerisation, and is juxtaposed to basic regions to create a DNA interaction interface surface that recognises specific DNA sequences. Basic region/HLH (bHLH) proteins regulate diverse biological pathways.

15 In an embodiment, said bHLH domain comprises an amino acid sequence having at least 50% overall sequence identity, and for instance at least 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% to the amino acid represented by SEQ ID NO: 544. In another embodiment, said bHLH domain comprises an
20 amino acid sequence having SEQ ID NO: 544.

In a preferred embodiment a method according to the invention is provided wherein said UPA20-like polypeptide comprises one or more of the following motifs:

- 25 (i) Motif 2 : YIHVRARRGQATDSHSLAERVRR[KR]ISERM[KR][LIF]LQ[DL]LVPGC [ND]K[IV]TGKA[LV]ML (SEQ ID NO: 538),
- (ii) Motif 3: DEIINYVQSLQ[RN]QVEFLSMKLA[ST]V[NS]P[RLV]L[DY] (SEQ ID NO: 539),
- (iii) Motif 4: [MN][AS][RK]KRK[SA][RA]x[KG][FGS] (SEQ ID NO: 540).

30 In another preferred embodiment a method according to the invention is provided wherein said UPA20-like polypeptide comprises one or more of the following motifs:

- (i) Motif 5: RVRRE[KR]ISERM[RK][MIV]LQ[ARLS]LVPGCDK[IV]TGKAL[MI]LDEIIN YVQSLQNQVEFLSM (SEQ ID NO: 541),
- 35 (ii) Motif 6: [KRS]KV[HE][EGD]E[PAE]P[KAT]GYIHVRARRGQATDSHSLAE (SEQ ID NO: 542),
- (iii) Motif 7: [KR][LI]AS[LMV][SN]P[VLM][LF]Y[DG]FGMD[LSR] (SEQ ID NO: 543)

40 Motifs 2 to 7 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 UPA20-like polypeptide comprises in increasing order of preference, at least 2, at least 3, at least 4, at least 5, or all 6 motifs.

- 5 Alternatively or in addition thereto, an UPA20-like polypeptide according to the invention comprises a conserved domain (or motif) with 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% sequence identity to or
- 10 consists of a conserved domain selected from any one of:
1. amino acid 186 to 236 in SEQ ID NO: 2 (motif 8 - SEQ ID NO: 556);
 2. amino acid 174 to 231 in SEQ ID NO: 2 (motif 9 - SEQ ID NO: 557);
 3. amino acid 183 to 230 in SEQ ID NO: 2 (motif 10 - SEQ ID NO: 558);
 4. amino acid 181 to 231 in SEQ ID NO: 2 (motif 11 - SEQ ID NO: 559);
 - 15 5. amino acid 176 to 255 in SEQ ID NO: 2 (motif 12 - SEQ ID NO: 560); and
 6. amino acid 176 to 257 in SEQ ID NO: 2 (motif 13 - SEQ ID NO: 561).

In another embodiment, an UPA20-like polypeptide according to the invention comprises a polypeptide having any one or more of the following domains:

- 20 - a domain as determined with the HMMSmart database and having accession number SM00353;
- a domain as determined with the HMMPanther database and having accession number as PTHR12565:SF7 and/or PTHR12565;
- 25 - a domain as determined with the HMMPfam database and having accession number PF00010;
- a domain as determined with the Profile Scan database and having accession number PS50888;
- a domain as determined with the Superfamily database and having accession number SSF47459; and
- 30 - a domain as determined with the Gene3D database and having accession number G3DSA:4.10.280.10.

The term “UPA20-like” or “UPA20-like polypeptide” as used herein also intends to include homologues as defined hereunder of “UPA20-like polypeptide”.

- 35 Additionally or alternatively, the homologue of a UPA20-like protein has in increasing order of preference at least 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%,
- 40 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: 407, provided that

the homologous protein comprises any one or more of the conserved motifs as outlined above. The overall sequence identity is determined using a global alignment algorithm, such as the Needleman Wunsch algorithm in the program GAP (GCG Wisconsin Package, Accelrys), preferably with default parameters and preferably with sequences of mature proteins (i.e. without taking into account secretion signals or transit peptides). 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 UPA20-like polypeptide have, in increasing order of preference, at least 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to any one or more of the motifs represented by SEQ ID NO: 133, 134, 135, 136, 137, 138, 151, 152, 153, 154, 155, and 156 (Motifs 2 to 13).

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

Furthermore, WAK-like polypeptides, at least in their native form, typically have oligogalacturonides binding activity. Tools and techniques for measuring oligogalacturonides binding activity are well known in the art, as e.g. described in Decreux and Messiaen, 2005. In addition, WAK-like polypeptides, when expressed in rice according to the methods of the present invention as outlined in Examples 6 and 7, give plants having increased yield related traits, in particular increased seed yield and/or increased root biomass.

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

Concerning UPA20-like polypeptides, the polypeptide sequence which when used in the construction of a phylogenetic tree, such as the one depicted in Figure 9, preferably clusters with the group of UPA20-like polypeptides comprising the amino acid sequence represented by SEQ ID NO: 407 rather than with any other group. The group of UPA20-like polypeptides comprising the amino acid sequence represented by SEQ ID NO: 407 is also denoted as Clade A (see also Table A3).

Furthermore, UPA20-like polypeptides (at least in their native form) typically have DNA binding activity. Tools and techniques for measuring DNA binding activity are well known in the art and are therefore not detailed herein. In addition, UPA20-like polypeptides, when expressed in rice according to the methods of the present invention as outlined in Examples 6 and 7, give plants having increased yield related traits, in particular increased biomass, including above ground biomass (area max) and root biomass (root max and root thick max); and increased seed yield, including total weight of seeds, harvest index, number of

filled seeds, number of flowers per panicle and fill rate as compared to control plants.

Concerning WAK-like polypeptides, 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; as well as by transforming plants with the nucleic acid sequence represented by SEQ ID NO: 3, encoding the polypeptide sequence of SEQ ID NO: 4. However, performance of the invention is not restricted to these sequences; the methods of the invention may advantageously be performed using any WAK-like-encoding nucleic acid or WAK-like polypeptide as defined herein.

Examples of nucleic acids encoding WAK-like polypeptides are given in Table A1 of the Examples section herein. Nucleic acids useful in performing the methods of the invention can be chosen from Table A1, provided these nucleic acids encode polypeptides which comprise an EGF domain and a transmembrane domain. The amino acid sequences useful in performing the methods of the invention can be chosen from Table A1, provided these polypeptides comprise an EGF domain and a transmembrane domain and as such are example sequences of orthologues and paralogues of the WAK-like polypeptide represented by SEQ ID NO: 2, the terms "orthologues" and "paralogues" being as defined herein. Further orthologues and paralogues may readily be identified by performing a so-called reciprocal blast search as described in the definitions section; where the query sequence is SEQ ID NO: 1 or SEQ ID NO: 2, the second BLAST (back-BLAST) would be against rice sequences.

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

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

- (i) a nucleic acid represented by any one of SEQ ID NO: 3, 5, 7 or 9;
- (ii) the complement of a nucleic acid represented by any one of SEQ ID NO: 3, 5, 7 or 9;
- (iii) a nucleic acid encoding a WAK-like 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: 4, 6, 8 or 10, and additionally comprising an EGF domain and a transmembrane domain, and further preferably conferring enhanced yield-related traits relative to control plants.
- (iv) a nucleic acid molecule which hybridizes with a nucleic acid molecule of (i) to (iii)

under high stringency hybridization conditions and preferably confers enhanced yield-related traits relative to control plants.

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

- (i) an amino acid sequence represented by any one of SEQ ID NO: 4, 6, 8 or 10;
- (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: 4, 6, 8 or 10, and additionally comprising an EGF domain and a transmembrane domain, 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.

Concerning UP20-like polypeptides, the present invention is illustrated by transforming plants with the nucleic acid sequence represented by SEQ ID NO: 406, encoding the polypeptide sequence of SEQ ID NO: 407. However, performance of the invention is not restricted to these sequences; the methods of the invention may advantageously be performed using any UPA20-like -encoding nucleic acid or UPA20-like polypeptide as defined herein.

Examples of nucleic acids encoding UPA20-like 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 UPA20-like polypeptide represented by SEQ ID NO: 407, 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: 406 or SEQ ID NO: 407, the second BLAST (back-BLAST) would be against poplar sequences.

The invention also provides hitherto unknown UPA20-like encoding nucleic acids and UPA20-like polypeptides useful for conferring enhanced yield-related traits in plants relative to control plants.

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

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: 458, 460 and 496;
 - (ii) the complement of a nucleic acid represented by any one of SEQ ID NO: 458, 460 and 496;
 - (iii) a nucleic acid encoding a UPA20-like 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: 459, 461 and 497, and additionally or alternatively comprising one or more motifs having in increasing order of preference at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more sequence identity to any one or more of the motifs given in SEQ ID NO: 538, 539, 540, 541, 542, 543, 556, 557, 558, 559, 560, and 561 (motifs 2 to 13), 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.
- 20 According to another 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: 459, 461 and 497;
 - (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: 459, 461 and 497, and additionally or alternatively comprising one or more motifs having in increasing order of preference at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more sequence identity to any one or more of the motifs given in SEQ ID NO: 538, 539, 540, 541, 542, 543, 556, 557, 558, 559, 560, and 561 (motifs 2 to 13) 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 A1 of the Examples section, provided that these polypeptides comprise an EGF domain and a transmembrane domain, or 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 A1 of the Examples section, provided that these polypeptides comprise an EGF domain and a transmembrane domain, or 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 WAK-like polypeptides, or UPA20-like polypeptides, nucleic acids hybridising to nucleic acids encoding WAK-like polypeptides, or UPA20-like polypeptides, splice variants of nucleic acids encoding WAK-like polypeptides, or UPA20-like polypeptides, allelic variants of nucleic acids encoding WAK-like polypeptides, or UPA20-like polypeptides, and variants of nucleic acids encoding WAK-like polypeptides, or UPA20-like polypeptides, obtained by gene shuffling. The terms hybridising sequence, splice variant, allelic variant and gene shuffling are as described herein.

Nucleic acids encoding WAK-like polypeptides, or UPA20-like 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, provided that these polypeptides comprise an EGF domain and a transmembrane domain, or 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 A1 of the Examples section, provided that these polypeptides comprise an EGF domain and a transmembrane domain, or 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.

Concerning WAK-like polypeptides, portions useful in the methods of the invention, encode a WAK-like polypeptide as defined herein, and have substantially the same biological activity as the amino acid sequences given in Table A1 of the Examples section, provided these polypeptides comprise an EGF domain and a transmembrane domain. 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 amino acid sequences given in Table A1 of the Examples section, provided these nucleic acids encode polypeptides comprising an EGF domain and a transmembrane domain. Preferably the portion is at least 50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100 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, provided these nucleic acids encode polypeptides comprising an EGF domain and a transmembrane domain. Most preferably the portion is a portion of the nucleic acid of SEQ ID NO: 1, provided the nucleic acid encodes a polypeptide comprising an EGF domain and a transmembrane domain.

Concerning UPA20-like polypeptides, portions useful in the methods of the invention, encode a UPA20-like 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 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, 1450, 1500, 1550, 1600, 1650, 1700, 1750 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.

Preferably, the portion encodes a fragment of an amino acid sequence which has one or more of the following characteristics:

- when used in the construction of a phylogenetic tree, such as the one depicted in Figure 9, clusters with the group of polypeptides comprising the amino acid sequence represented by SEQ ID NO: 407 (Clade A) rather than with any other group;
- comprises a bHLH domain as defined herein,
- comprises any one or more of the motifs 2 to 13 as provided herein,
- has DNA binding activity, and
- has at least 30% sequence identity to SEQ ID NO: 407.

Most preferably the portion is a portion of the nucleic acid of SEQ ID NO: 406.

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 WAK-like polypeptide, or a UPA20-like polypeptide, as defined herein, or with a portion as defined herein.

According to the present invention, there is provided a method for enhancing yield-related traits in plants, comprising introducing and expressing in a plant a nucleic acid capable of hybridizing to any one of the nucleic acids given in Table A1 of the Examples section, provided that these polypeptides comprise an EGF domain and a transmembrane domain, or 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 A1 of the Examples section, provided that these polypeptides comprise an EGF domain and a transmembrane domain, or Table A3 of the Examples section.

Concerning WAK-like polypeptides, hybridising sequences useful in the methods of the invention encode a WAK-like polypeptide as defined herein, having substantially the same biological activity as the amino acid sequences given in Table A1 of the Examples section, provided these nucleic acids encode polypeptides comprising an EGF domain and a transmembrane domain. 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, provided these nucleic acids encode polypeptides comprising an EGF domain and a transmembrane domain. 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. In one embodiment, the hybridization conditions are of medium stringency, preferably of high stringency, as defined above.

Concerning UPA20-like polypeptides, hybridising sequences useful in the methods of the invention encode a UPA20-like 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 hybridising sequence is capable of hybridising to the complement of a nucleic acid as represented by SEQ ID NO: 406 or to a portion thereof.

Preferably, the hybridising sequence encodes a polypeptide with an amino acid sequence which has one or more of the following characteristics:

- when used in the construction of a phylogenetic tree, such as the one depicted in Figure 9, clusters with the group of polypeptides comprising the amino acid sequence represented by SEQ ID NO: 407 (Clade A) rather than with any other group;

- comprises a bHLH domain as defined herein,
- comprises any one or more of the motifs 2 to 13 as provided herein,
- has DNA binding activity, and
- has at least 30% sequence identity to SEQ ID NO: 407.

5

Another nucleic acid variant useful in the methods of the invention is a splice variant encoding a WAK-like polypeptide, or a UPA20-like polypeptide, as defined hereinabove, a splice variant being as defined herein.

- 10 In another embodiment, there is provided a method for enhancing yield-related traits in plants, comprising introducing and expressing in a plant a splice variant of a nucleic acid encoding any one of the proteins given in Table A1 of the Examples section, provided that these polypeptides comprise an EGF domain and a transmembrane domain, or Table A3 of the Examples section, or a splice variant of a nucleic acid encoding an orthologue,
- 15 paralogue or homologue of any of the amino acid sequences given in Table A1 of the Examples section, provided that these polypeptides comprise an EGF domain and a transmembrane domain, or Table A3 of the Examples section.

- Concerning WAK-like polypeptides, preferred splice variants are splice variants of a nucleic acid represented by SEQ ID NO: 1 or 3, or a splice variant of a nucleic acid encoding an orthologue or paralogue of SEQ ID NO: 2 or 4, provided these nucleic acids encode polypeptides comprising an EGF domain and a transmembrane domain.

- Concerning UPA20-like polypeptides, preferred splice variants are splice variants of a nucleic acid represented by SEQ ID NO: 406, or a splice variant of a nucleic acid encoding an orthologue or paralogue of SEQ ID NO: 407. Preferably, the amino acid sequence encoded by the splice variant, has one or more of the following characteristics:

- when used in the construction of a phylogenetic tree, such as the one depicted in Figure 9, clusters with the group of polypeptides comprising the amino acid sequence represented by SEQ ID NO: 407 (Clade A) rather than with any other group;
- comprises a bHLH domain as defined herein,
- comprises any one or more of the motifs 2 to 13 as provided herein,
- has DNA binding activity, and
- has at least 30% sequence identity to SEQ ID NO: 407.

- Another nucleic acid variant useful in performing the methods of the invention is an allelic variant of a nucleic acid encoding a WAK-like polypeptide, or a UPA20-like polypeptide, as defined hereinabove, an allelic variant being as defined herein.

40

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

encoding any one of the proteins given in Table A1 of the Examples section, provided that these polypeptides comprise an EGF domain and a transmembrane domain, or 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 A1 of the Examples section, provided that these polypeptides comprise an EGF domain and a transmembrane domain, or Table A3 of the Examples section.

Concerning WAK-like polypeptides, the polypeptides encoded by allelic variants useful in the methods of the present invention have substantially the same biological activity as the WAK-like polypeptide of SEQ ID NO: 2 or 4 and any of the amino acids depicted in Table A1 of the Examples section provided these nucleic acids encode polypeptides comprising an EGF domain and a transmembrane domain. 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 3 or an allelic variant of a nucleic acid encoding an orthologue or paralogue of SEQ ID NO: 2 or 4, provided these nucleic acids encode polypeptides comprising an EGF domain and a transmembrane domain.

Concerning UPA20-like polypeptides, the polypeptides encoded by allelic variants useful in the methods of the present invention have substantially the same biological activity as the UPA20-like polypeptide of SEQ ID NO: 407 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: 406 or an allelic variant of a nucleic acid encoding an orthologue or paralogue of SEQ ID NO: 407.

Preferably, the amino acid sequence encoded by the allelic variant, has one or more of the following characteristics:

- when used in the construction of a phylogenetic tree, such as the one depicted in Figure 9, clusters with the group of polypeptides comprising the amino acid sequence represented by SEQ ID NO: 407 (Clade A) rather than with any other group;
- comprises a bHLH domain as defined herein,
- comprises any one or more of the motifs 2 to 13 as provided herein,
- has DNA binding activity, and
- has at least 30% sequence identity to SEQ ID NO: 407.

Gene shuffling or directed evolution may also be used to generate variants of nucleic acids encoding WAK-like polypeptides, or UPA20-like polypeptides, as defined above; the term "gene shuffling" being as defined herein.

Concerning WAK-like polypeptides, 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 provided these nucleic acids encode polypeptides comprising an EGF domain and a transmembrane domain, 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, provided these nucleic acids encode polypeptides comprising an EGF domain and a transmembrane domain, which variant nucleic acid is obtained by gene shuffling.

Concerning UPA20-like polypeptides, 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, has one or more of the following characteristics:

- when used in the construction of a phylogenetic tree, such as the one depicted in Figure 9, clusters with the group of polypeptides comprising the amino acid sequence represented by SEQ ID NO: 407 (Clade A) rather than with any other group;
- comprises a bHLH domain as defined herein,
- comprises any one or more of the motifs 2 to 13 as provided herein,
- has DNA binding activity, and
- has at least 30% sequence identity to SEQ ID NO: 407.

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.).

WAK-like polypeptides differing from the sequence of SEQ ID NO: 2 or 4 by one or several amino acids (substitution(s), insertion(s) and/or deletion(s) as defined above) may equally be useful to increase the yield of plants in the methods and constructs and plants of the invention, provided these polypeptides comprise an EGF domain and a transmembrane domain.

Nucleic acids encoding WAK-like polypeptides may be derived from any natural or artificial source. The nucleic acid may be modified from its native form in composition and/or genomic environment through deliberate human manipulation. Preferably the WAK-like

polypeptide-encoding nucleic acid is from a plant, further preferably from a monocotyledonous plant, more preferably from the family Poaceae, most preferably the nucleic acid is from *Oryza sativa*.

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

10 Nucleic acids encoding UPA20-like polypeptides may be derived from any natural or artificial source. The nucleic acid may be modified from its native form in composition and/or genomic environment through deliberate human manipulation.

- 15 In a preferred embodiment, the UPA20-like polypeptide-encoding nucleic acid is from a plant, preferably from a dicotyledonous plant, further preferably from the family Salicaceae, more preferably from the genus *Populus*, most preferably the nucleic acid is from *Populus trichocarpa*. In another preferred embodiment, the UPA20-like polypeptide-encoding nucleic acid is from a plant, preferably from a dicotyledonous plant, further preferably from the family Brassicaceae, more preferably from the genus *Arabidopsis*, most preferably the nucleic acid is from *Arabidopsis thaliana*. In yet another preferred embodiment, the UPA20-like polypeptide-encoding nucleic acid is from a plant, preferably from a monocotyledonous plant, further preferably from the family Poaceae, more preferably from the genus *Oryza*, most preferably the nucleic acid is from *Oryza sativa*.

25 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 natural genetic environment. In a further embodiment the recombinant chromosomal DNA of the invention is comprised in a plant cell.

35 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 and/or increased root biomass relative to control plants. The terms "yield", "seed yield" and "root biomass" are described in more detail in the "definitions" section herein.

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

seeds, and performance of the methods of the invention results in plants having increased seed yield relative to the seed yield of control plants.

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

10 The present invention also provides a method for increasing yield-related traits, preferably for increasing yield more preferably for increasing biomass as defined herein and/or for increasing seed yield as defined herein, relative to control plants, which method comprises modulating expression in a plant of a nucleic acid encoding a UPA20-like polypeptide as defined herein.

15 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 WAK-like polypeptide, or a UPA20-like polypeptide, as defined herein.

20 Performance of the methods of the invention gives plants grown under non-stress conditions or under mild drought conditions increased yield-related traits relative to control plants grown under comparable conditions. Therefore, according to the present invention, there is provided a method for increasing yield-related traits in plants grown under non-stress conditions or under mild drought conditions, which method comprises modulating expression in a plant of a nucleic acid encoding a WAK-like polypeptide, or a UPA20-like polypeptide.

30 Performance of the methods of the invention gives plants grown under conditions of drought, increased yield-related traits relative to control plants grown under comparable conditions. Therefore, according to the present invention, there is provided a method for increasing yield-related traits in plants grown under conditions of drought which method comprises modulating expression in a plant of a nucleic acid encoding a WAK-like polypeptide, or a UPA20-like polypeptide.

35 Performance of the methods of the invention gives plants grown under conditions of nutrient deficiency, particularly under conditions of nitrogen deficiency, increased yield-related traits relative to control plants grown under comparable conditions. Therefore, according to the present invention, there is provided a method for increasing yield-related traits in plants grown under conditions of nutrient deficiency, which method comprises modulating expression in a plant of a nucleic acid encoding a WAK-like polypeptide, or a UPA20-like polypeptide.

40

Performance of the methods of the invention gives plants grown under conditions of salt stress, increased yield-related traits relative to control plants grown under comparable conditions. Therefore, according to the present invention, there is provided a method for increasing yield-related traits in plants grown under conditions of salt stress, which method comprises modulating expression in a plant of a nucleic acid encoding a WAK-like polypeptide, or a UPA20-like polypeptide.

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

More specifically, the present invention provides a construct comprising:

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

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

Concerning UPA20-like polypeptides, the invention also provides genetic constructs and vectors to facilitate introduction and/or expression in plants of nucleic acids encoding UPA20-like polypeptides. The gene constructs may be inserted into vectors, which may be commercially available, suitable for transforming into plants or host cells and suitable for expression of the gene of interest in the transformed cells. The invention also provides use of a gene construct as defined herein in the methods of the invention.

More specifically, the present invention provides a construct comprising:

- (a) a nucleic acid encoding a UPA20-like polypeptide as defined above;
- (b) one or more control sequences capable of driving expression of the nucleic acid sequence of (a); and optionally
- (c) a transcription termination sequence.

Preferably, the nucleic acid encoding a UPA20-like polypeptide is as defined above. The term “control sequence” and “termination sequence” are as defined herein.

The genetic construct of the invention may be comprised in a host cell, plant cell, seed, agricultural product or plant. Plants or host cells are transformed with a genetic construct

such as a vector or an expression cassette comprising any of the nucleic acids described above. Thus the invention furthermore provides plants or host cells transformed with a construct as described above. In particular, the invention provides plants transformed with a construct as described above, which plants have increased yield-related traits as described herein.

In one embodiment the genetic construct of the invention confers increased yield or yield related trait(s) to a plant when it has been introduced into said plant, which plant expresses the nucleic acid encoding the WAK-like polypeptide, or the UPA20-like polypeptide, comprised in the genetic construct. In another embodiment the genetic construct of the invention confers increased yield or yield related traits(s) to a plant comprising plant cells in which the construct has been introduced, which plant cells express the nucleic acid encoding the WAK-like polypeptide, or the UPA20-like polypeptide, comprised in the genetic construct.

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

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

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 GOS2 promoter from rice. Further preferably the constitutive promoter is represented by a nucleic acid sequence substantially similar to SEQ ID NO: 306, or SEQ ID NO: 545, most preferably the constitutive promoter is as represented by SEQ ID NO: 306, or SEQ ID NO: 545. See the "Definitions" section herein for further examples of constitutive promoters.

Concerning WAK-like polypeptides, it should be clear that the applicability of the present invention is not restricted to the WAK-like polypeptide-encoding nucleic acid represented by SEQ ID NO: 1 or 3, nor is the applicability of the invention restricted to expression of a WAK-like polypeptide-encoding nucleic acid when driven by a constitutive promoter.

Concerning UPA20-like polypeptides, it should be clear that the applicability of the present

invention is not restricted to the UPA20-like polypeptide-encoding nucleic acid represented by SEQ ID NO: 406, nor is the applicability of the invention restricted to expression of a UPA20-like polypeptide-encoding nucleic acid when driven by a constitutive promoter. According to another embodiment of the invention, the nucleic acid encoding a UPA20-like polypeptide can operably linked to a root-specific promoter. Examples of other root-specific promoters which may also be used to perform the methods of the invention are shown in Table 2b in the “Definitions” section above. According to yet another embodiment, to another embodiment of the invention, the nucleic acid encoding a UPA20-like polypeptide can operably linked to a seed specific promoter. For instance, the seed specific promoter is an endosperm/aleurone/embryo specific promoter. Examples of endosperm/aleurone/embryo specific promoters are given in Table 2 of the definitions section.

Concerning WAK-like polypeptides, 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: 306, operably linked to the nucleic acid encoding the WAK-like polypeptide. More preferably, the construct comprises a zein terminator (t-zein) linked to the 3' end of the WAK-like 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 any of the SEQ ID NO: 303, 304 or 305 (pPRO::WAK-like::t-zein sequence). Furthermore, one or more sequences encoding selectable markers may be present on the construct introduced into a plant.

Concerning UPA20-like polypeptides, 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: 545, operably linked to the nucleic acid encoding the UPA20-like polypeptide. More preferably, the construct comprises a zein terminator (t-zein) linked to the 3' end of the UPA20-like 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 SEQ ID NO: 546 (pGOS2::UPA20-like::t-zein sequence – expression vector 1). Furthermore, one or more sequences encoding selectable markers may be present on the construct introduced into a plant.

In another embodiment, the construct comprises an expression cassette comprising a GOS2 promoter, substantially similar to SEQ ID NO: 545, operably linked to the nucleic acid encoding the UPA20-like polypeptide. More preferably, the construct comprises a zein terminator (t-zein) linked to the 3' end of the UPA20-like 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 SEQ ID NO: 547 (pGOS2::UPA20-like::tzein sequence – expression vector

2). 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 WAK-like polypeptide, or a UPA20-like polypeptide, is by introducing and expressing in a plant a nucleic acid encoding a WAK-like polypeptide, or a UPA20-like polypeptide; however the effects of performing the method, i.e. enhancing yield-related traits may also be achieved using other well known techniques, including but not limited to T-DNA activation tagging, TILLING, homologous recombination. A description of these techniques is provided in the definitions section.

The invention also provides a method for the production of transgenic plants having enhanced yield-related traits relative to control plants, comprising introduction and expression in a plant of any nucleic acid encoding a WAK-like polypeptide, or a UPA20-like polypeptide, as defined herein.

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

- (i) introducing and expressing in a plant or plant cell a nucleic acid encoding a WAK-like polypeptide, or a UPA20-like polypeptide, or a genetic construct comprising a nucleic acid encoding a WAK-like polypeptide, or a UPA20-like polypeptide; and
- (ii) cultivating the plant cell under conditions promoting plant growth and development.

The nucleic acid of (i) may be any of the nucleic acids capable of encoding a WAK-like polypeptide, or a UPA20-like polypeptide, as defined herein.

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

In another embodiment the plant cells of the invention are plant cells that do not sustain themselves in an autotrophic way.

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

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

In another embodiment, the plant cells of the invention are plant cells that do not sustain themselves in an autotrophic way, such plant cells are not deemed to represent a plant
15 variety. In a further embodiment the plant cells of the invention are non-plant variety and non-propagative.

The present invention encompasses plants or parts thereof (including seeds) obtainable by the methods according to the present invention. The plants or plant parts or plant cells
20 comprise a nucleic acid transgene encoding a WAK-like polypeptide, or a UPA20-like polypeptide, as defined above, preferably in a genetic construct such as an expression cassette. The present invention extends further to encompass the progeny of a primary transformed or transfected cell, tissue, organ or whole plant that has been produced by any of the aforementioned methods, the only requirement being that progeny exhibit the same
25 genotypic and/or phenotypic characteristic(s) as those produced by the parent in the methods according to the invention.

In a further embodiment the invention extends to seeds comprising the expression cassettes of the invention, the genetic constructs of the invention, or the nucleic acids
30 encoding the WAK-like polypeptide, or the UPA20-like polypeptide, and/or the a WAK-like polypeptides, or the UPA20-like polypeptides, as described above.

The invention also includes host cells containing an isolated nucleic acid encoding a WAK-like polypeptide, or a UPA20-like polypeptide, as defined above. In one embodiment host
35 cells according to the invention are plant cells, yeasts, bacteria or fungi. Host plants for the nucleic acids, construct, expression cassette or the vector used in the method according to the invention are, in principle, advantageously all plants which are capable of synthesizing the polypeptides used in the inventive method. In a particular embodiment the plant cells of the invention overexpress the nucleic acid molecule of the invention.

40 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
5 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,
10 barley, millet, rye, triticale, sorghum, emmer, spelt, einkorn, teff, milo and oats. In a particular embodiment the plants used in the methods of the invention are selected from the group consisting of maize, wheat, rice, soybean, cotton, oilseed rape including canola, sugarcane, sugar beet and alfalfa. Advantageously the methods of the invention are more efficient than the known methods, because the plants of the invention have increased yield
15 and/or tolerance to an environmental stress compared to control plants used in comparable methods.

The invention also extends to harvestable parts of a plant such as, but not limited to seeds, leaves, fruits, flowers, stems, roots, rhizomes, tubers and bulbs, which harvestable parts
20 comprise a recombinant nucleic acid encoding a WAK-like polypeptide, or a UPA20-like polypeptide. The invention furthermore relates to products derived or produced, preferably directly derived or produced, from a harvestable part of such a plant, such as dry pellets, meal or powders, oil, fat and fatty acids, starch or proteins.

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

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

In yet another embodiment the polynucleotides or the polypeptides of the invention are comprised in an agricultural product. In a particular embodiment the nucleic acid sequences
40 and protein sequences of the invention may be used as product markers, for example where an agricultural product was produced by the methods of the invention. Such a marker can be used to identify a product to have been produced by an advantageous process

resulting not only in a greater efficiency of the process but also improved quality of the product due to increased quality of the plant material and harvestable parts used in the process. Such markers can be detected by a variety of methods known in the art, for example but not limited to PCR based methods for nucleic acid detection or antibody based methods for protein detection.

The present invention also encompasses use of nucleic acids encoding WAK-like polypeptides, or UPA20-like polypeptides, as described herein and use of these the WAK-like polypeptides, or the UPA20-like polypeptides, in enhancing any of the aforementioned yield-related traits in plants. For example, nucleic acids encoding WAK-like polypeptide, or UPA20-like polypeptide, described herein, or the WAK-like polypeptides, or the UPA20-like polypeptides themselves, may find use in breeding programmes in which a DNA marker is identified which may be genetically linked to gene encoding a WAK-like polypeptide, or a UPA20-like polypeptide. The nucleic acids/genes, or the WAK-like polypeptides, or UPA20-like polypeptides, themselves may be used to define a molecular marker. This DNA or protein marker may then be used in breeding programmes to select plants having enhanced yield-related traits as defined herein in the methods of the invention. Furthermore, allelic variants of a nucleic acid/gene encoding a WAK-like polypeptide, or a UPA20-like polypeptide, encoding may find use in marker-assisted breeding programmes. Nucleic acids encoding WAK-like polypeptides, or UPA20-like polypeptides, may also be used as probes for genetically and physically mapping the genes that they are a part of, and as markers for traits linked to those genes. Such information may be useful in plant breeding in order to develop lines with desired phenotypes.

2. CDKB-RKA polypeptides

The present invention is directed to the use of nucleic acid sequences encoding CDKB-RKA polypeptides as defined herein, for increasing yield, and preferably for increasing seed yield and/or for increasing biomass such as vegetative biomass in plants relative to control plants. The present invention is based on the use of modified forms of B-type cyclin-dependent kinases -herein also named "CDKB-RKA polypeptides" or "CDKB-RKA"- that have reduced kinase activity and preferably that lack kinase activity as compared to wild-type B-type cyclin-dependent kinases (CDKB) and on the use of genes encoding such CDKB-RKA polypeptides.

Surprisingly, it has now been found that modulating expression in a plant of a nucleic acid encoding a CDKB-RKA as defined herein, preferably by introducing and expressing in said plant of a nucleic acid encoding a CDKB-RKA as defined herein, gives plants having enhanced yield-related traits relative to control plants, in particular increased yield, and more particularly (i) increased vegetative biomass such as increased leaf and/or root biomass, and/or (ii) increased seed yield relative to control plants.

In particular embodiments, a CDKB-RKA as defined herein does not comprise or consist of

a full length CDKB polypeptide as described in WO 2005/024029 or as given in Example 1, Table A2, and in particular does not comprise or consist of a full length CDKB polypeptide as represented by SEQ ID NO: 320.

5 According to a first aspect, 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 CDKB-RKA as defined herein. A preferred method for modulating, preferably increasing expression of a nucleic acid encoding a CDKB-RKA polypeptide as defined herein is by introducing and expressing a nucleic acid encoding
10 such polypeptide in a plant. Hence, a method for increasing yield in a plant relative to a control plant is provided herein, wherein said method comprises introducing and expressing in said plant a nucleic acid molecule encoding a CDKB-RKA polypeptide as defined herein.

In another aspect, the present invention provides a method for producing plants having
15 enhanced yield-related traits relative to control plants, wherein said method comprises the steps of modulating expression in said plant of a nucleic acid encoding a CDKB-RKA polypeptide as defined herein and optionally selecting for plants having enhanced yield-related traits. In a preferred embodiment, the present invention provides a method for producing plants having enhanced yield-related traits relative to control plants, wherein said
20 method comprises the steps of introducing and expressing in said plant a nucleic acid molecule encoding a CDKB-RKA polypeptide as defined herein and optionally selecting for plants having enhanced yield-related traits.

Any reference hereinafter to a “protein useful in the methods of the invention” is taken to
25 mean a CDKB-RKA 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 a CDKB-RKA as defined herein. It shall further be noted that the terms “CDKB-RKA” and “CDKB-RKA polypeptide” and “CDKB-RKA protein” are used herein as synonyms.

30 The term “CDKB-RKA” as used herein encompasses modified forms of B-type cyclin-dependent kinases (CDKB), wherein said modified forms have at least reduced kinase activity compared to wild-type CDKB proteins. Preferably said CDKB-RKA lacks kinase activity. CDKB polypeptides are well known in the art and have for instance been described
35 in Van Leene et al., 2010, Mol Syst Biol 6, 397; Francis 2007 New Phytol 174(2):261-78; Doonan & Kitsios 2009 Mol Biotechnol 42(1):14-29. Examples of CDKB polypeptides are also shown in Example 1, Table A2.

More in particular, the term “CDKB-RKA” as used herein encompasses B-type cyclin-
40 dependent kinases (CDKB) that are mutated, which mutants have at least reduced kinase activity compared to wild-type proteins, and preferably lack kinase activity. In an example, term “CDKB-RKA” as used encompasses B-type cyclin-dependent kinases (CDKB) that are

mutated in the kinase domain. The term “reduced catalytic kinase activity” as used herein refers to a polypeptide which has reduced catalytic kinase activity as compared to a wild type polypeptide. The term “lacking catalytic kinase activity” as used herein refers to a polypeptide which has no or no detectable catalytic kinase activity as compared to a wild type polypeptide. The terms “catalytic kinase activity” and “kinase activity” are used herein as synonyms. Tools and techniques for detecting or determining catalytic kinase activity are well known in the art. Catalytic kinase activity can for instance be determined by means of a kinase assay, such as the one described in Example 4 of the present specification.

The term “mutated” in the context of the present invention intends to refer to one or more non-silent mutations, i.e. mutations that modify the amino sequence of the polypeptide as compared to a corresponding unmodified polypeptide. Non-limitative examples of such mutations may for instance include: insertions; deletions; amino acid substitutions; frameshift mutations caused by insertion or deletion of a number of nucleotides in a DNA acid sequence that is not evenly divisible by three, point mutations, nonsense mutations resulting for instance in a premature stop codon or a nonsense codon in the transcribed mRNA, and possibly resulting in a truncated protein product; missense mutations or nonsynonymous mutations which is a type of point mutations where a single nucleotide is changed to cause substitution of a different amino acid; etc. Techniques for inducing mutations such as insertions, substitutions, truncations, deletions, etc. are well known in the art, and will therefore not be described in detail herein.

The term “CDKB-RKA” as used herein also encompasses truncated forms of CDKB, wherein the truncations are preferably located in the kinase domain. The term “CDKB-RKA” as used herein also encompasses truncated forms of CDKB polypeptides in which no active kinase domain is present. An example thereof is represented by *Arabidopsis* sequence SEQ ID NO: 318, which is a truncated version of SEQ ID NO: 320. Preferably, a CDKB-RKA polypeptide useful in the methods of the present invention is a truncated form of a CDKB polypeptide with a deletion in the C-terminal half of the polypeptide that reduces, preferably substantially inactivates that kinase activity of the polypeptide. More preferably, a CDKB-RKA polypeptide useful in the methods of the present invention is a truncated form of a CDKB polypeptide with a deletion in the C-terminal half of the polypeptide that comprises or consists of substantially the complete kinase domain of that CDKB polypeptide or a part thereof.

In a preferred embodiment, a CDKB-RKA polypeptide as provided herein comprise a cyclin binding domain represented by a PPTALRE motif with no mismatches or with a mismatch at position 2 and/or 4 from left to right. Preferably a CDKB-RKA polypeptide as provided herein comprise a cyclin binding domain represented by SEQ ID NO: 403, and more preferably represented by SEQ ID NO: 404. In other words, the present invention relates to the use of a nucleic acid sequence encoding a CDKB-RKA polypeptide that is capable of binding cyclins and hence that has cyclin-binding activity. The PPTALRE motifs with no

mismatches or with a mismatch at position 2 and/or 4 from left to right, of a representative number of B-type CDK polypeptides can be seen on Figure 3.

Alternatively or additionally, the present invention relates to the use of a nucleic acid sequence encoding a CDKB-RKA polypeptide that

- (i) has cyclin-binding activity, and
- (ii) has reduced kinase activity and/or lacks kinase activity as defined herein.

In another embodiment, the present invention relates to the use of a nucleic acid sequence encoding a CDKB-RKA polypeptide that

- (i) has a cyclin-binding domain represented by a PPTALRE motif with no mismatches or with a mismatch at position 2 and/or 4 from left to right, and preferably represented by SEQ ID NO: 403, and more preferably represented by SEQ ID NO: 404, and
- (ii) has a mutated or truncated kinase domain or lacks a kinase domain.

In yet another embodiment, the present invention relates to the use of a nucleic acid sequence encoding a CDKB-RKA polypeptide that

- (i) has a cyclin-binding domain represented by a PPTALRE motif with no mismatches or with a mismatch at position 2 and/or 4 from left to right, and preferably represented by SEQ ID NO: 403, and more preferably represented by SEQ ID NO: 404, and
- (ii) has a mutated or truncated kinase domain or that lacks a kinase domain, and
- (iii) has at least about 30-40%, preferably at least about 41-50%, more preferably at least about 51-60%, 61-70%, 71-80%, 81-85%, 86-90%, 91-95%, 96-99% or more than 99% overall sequence identity to SEQ ID NO: 318.

In another embodiment, a CDKB-RKA polypeptide as provided herein comprises a protein kinase inhibitory domain. In a preferred embodiment, said protein kinase inhibitory domain has a protein kinase inhibitory function, and it is able to be phosphorylated by a kinase, such as by a Wee1 Kinase, to inhibit catalytic kinase activity.

In yet another embodiment, a CDKB-RKA polypeptide as provided herein lacks a T-loop activation kinase domain (as defined by Magyar et al. 1997; plant cell 9(2) 223-235).

Nucleic acids encoding a CDKB-RKA polypeptide as defined herein can be derived from any one of the CDKB nucleic acid sequences as given in Table A2 of the Examples section herein. Nucleic acids derived from nucleic acid sequences as given in Table A2 of the Examples section herein 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 a CDKB polypeptide represented by SEQ ID NO: 320, 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: 319 or SEQ ID NO: 320, the second BLAST (back-BLAST) would be against *Arabidopsis* sequences. In addition, in another embodiment, a nucleic acid sequence encoding a CDKB polypeptide as defined herein can also be derived from an orthologue or paralogue of any of the polypeptides given in Table A2. Other examples of nucleic acids encoding CDKB polypeptides from which nucleic acid sequences encoding CDKB-RKA polypeptides as defined herein can be derived are also disclosed for instance in WO 2005/024029 or US20070214517, which are incorporated herein by reference.

In a preferred example, a CDKB-RKA polypeptide useful in the methods of the present invention is a modified form such as a mutated form and/or a truncated form, of any of the CDKB polypeptides as represented in Table A2, and preferably as represented by SEQ ID NO: 320.

In another preferred example, a CDKB-RKA polypeptide useful in the methods of the present invention is a truncated form of any of the CDKB polypeptides as represented in Table A2, and preferably as represented by SEQ ID NO: 320, with a deletion in the C-terminal half of said polypeptide that reduces, preferably substantially inactivates that kinase activity of the polypeptide.

In yet another preferred example a CDKB-RKA polypeptide useful in the methods of the present invention is a truncated form of any of the CDKB polypeptides as represented in Table A2, and preferably as represented by SEQ ID NO: 320, with a deletion that consists of a part of or substantially the complete kinase domain of any of said polypeptides. For instance, a CDKB-RKA polypeptide useful in the methods of the present invention is a truncated form of the CDKB polypeptide as represented by SEQ ID NO: 320, with a deletion of substantially the complete kinase domain of said polypeptide, for instance as represented by the amino acid coordinates 140 to 152 on SEQ ID NO: 4 (IPR008271), or of only a part thereof, wherein said part may comprise 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 amino acids.

In a particularly preferred embodiment, a CDKB-RKA polypeptide useful in the methods of the present invention corresponds to an amino acid sequence represented by SEQ ID NO: 318 or a homologous, orthologous or paralogous sequence thereof.

The term "homologue of SEQ ID NO: 318" as defined herein refers to an amino acid sequence that has 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% overall sequence identity to SEQ ID NO: 318, and for instance has at least about

30-40%, preferably at least about 41-50%, more preferably at least about 51-60%, 61-70%, 71-80%, 81-85%, 86-90%, 91-95%, 96-99% or more than 99% overall sequence identity to SEQ ID NO: 318, provided that the amino acid sequence has

- (i) reduced catalytic kinase activity, and preferably lacks catalytic kinase activity, as defined above, and
- (ii) optionally a cyclin binding domain represented by a PPTALRE motif with no mismatches or with a mismatch at position 2 and/or 4 from left to right, and preferably represented by SEQ ID NO: 403, and more preferably represented by SEQ ID NO: 404.

The overall sequence identity can be 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: 318.

In one embodiment of the present invention the function of the nucleic acid sequences according to the invention is to confer information for synthesis of the CDKB-RKA polypeptides having reduced, and preferably lacking catalytic kinase activity and that increase yield or yield-related traits, when such a nucleic acid sequence of the invention is transcribed and translated in a living plant cell.

Nucleic acids encoding CDKB-RKA polypeptides as defined herein 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 CDKB-RKA polypeptide-encoding nucleic acid is from a plant, further preferably from a dicotyledonous plant, more preferably from the family *Brassicaceae*, most preferably the nucleic acid is from *Arabidopsis thaliana*.

In addition, CDKB-RKA polypeptides as defined herein, when expressed in rice according to the methods of the present invention as outlined in Examples 8 and 9, give plants having increased yield related traits, in particular any one or more aboveground biomass, root biomass, and seed yield including total weight of seeds and number of seeds.

The present invention is illustrated by transforming plants with the nucleic acid sequence represented by SEQ ID NO: 317, encoding the polypeptide sequence of SEQ ID NO: 318. However, performance of the invention is not restricted to these sequences; the methods of the invention may advantageously be performed using any CDKB-RKA-encoding nucleic acid or CDKB-RKA polypeptide or variants thereof as defined herein.

Nucleic acid variants may also be useful in practicing the methods of the invention. Examples of such nucleic acid variants include nucleic acids encoding homologues and

derivatives of any one of the amino acid sequences given in Table A2 of the Examples section or of SEQ ID NO: 318, the terms “homologue” and “derivative” being as defined herein, provided that said homologues and derivatives have reduced, and preferably lack catalytic kinase activity. 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 A2 of the Examples section or of SEQ ID NO: 318, provided that said homologues and derivatives of orthologues or paralogues have reduced, and preferably lack catalytic kinase activity as defined above. Further variants useful in practicing 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 practicing the methods of the invention include portions of nucleic acids encoding CDKB-RKA polypeptides, nucleic acids hybridising to nucleic acids encoding CDKB-RKA polypeptides, splice variants of nucleic acids encoding CDKB-RKA polypeptides, allelic variants of nucleic acids encoding CDKB-RKA polypeptides and variants of nucleic acids encoding CDKB-RKA polypeptides obtained by gene shuffling. The terms hybridising sequence, splice variant, allelic variant and gene shuffling are as described herein. These further nucleic acid variants according to the invention are all characterized in that they encode polypeptides that have reduced, and preferably lack catalytic kinase activity.

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 of SEQ ID NO: 317, 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 or of SEQ ID NO: 317.

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

Portions useful in the methods of the invention, encode a CDKB-RKA polypeptide as defined herein. Preferably, the portion is a portion of any one of the nucleic acids given in Table A2 of the Examples section or of SEQ ID NO: 317, 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 or of SEQ ID NO: 318. Preferably the portion is at least 50, 60, 70, 80, 90, 100, 110, 120, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700 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 SEQ ID NO: 317, 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 or of SEQ ID NO: 317. Preferably, the portion encodes a fragment of an amino acid sequence which has reduced, and preferably lacks kinase activity, and which (i) comprises a cyclin binding domain and/or (ii) has at least about 30-40%, preferably at least about 41-50%, more preferably at least about 51-60%, 61-70%, 71-80%, 81-85%, 86-90%, 91-95%, 96-99% or more than 99% overall sequence identity to SEQ ID NO: 318.

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 CDKB-RKA 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 to SEQ ID NO: 317 to a portion as defined herein, 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 or to SEQ ID NO: 317 or to a portion as defined herein.

Hybridising sequences useful in the methods of the invention encode a CDKB-RKA polypeptide as defined herein. 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 SEQ ID NO: 317, 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 or of SEQ ID NO: 317.

Most preferably, the hybridising sequence is capable of hybridising to the complement of a nucleic acid as represented by SEQ ID NO: 317 or to a portion thereof. In one embodiment, the hybridization conditions are of medium stringency, preferably of high stringency, as defined above.

Preferably, the hybridising sequence encodes a polypeptide with an amino acid sequence which has reduced, and preferably lacks kinase activity, and which (i) comprises a cyclin binding domain and/or (ii) has at least about 30-40%, preferably at least about 41-50%, more preferably at least about 51-60%, 61-70%, 71-80%, 81-85%, 86-90%, 91-95%, 96-99% or more than 99% overall sequence identity to SEQ ID NO: 318.

Another nucleic acid variant useful in the methods of the invention is a splice variant encoding a CDKB-RKA 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 of SEQ ID NO: 317, 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 or of SEQ ID NO: 317.

Preferred splice variants are splice variants of a nucleic acid represented by SEQ ID NO: 317, or a splice variant of a nucleic acid encoding an orthologue or paralogue of SEQ ID NO: 318. Preferably, the amino acid sequence encoded by the splice variant has reduced, and preferably lacks kinase activity, and (i) comprises a cyclin binding domain and/or (ii) has at least about 30-40%, preferably at least about 41-50%, more preferably at least about 51-60%, 61-70%, 71-80%, 81-85%, 86-90%, 91-95%, 96-99% or more than 99% overall sequence identity to SEQ ID NO: 318.

Another nucleic acid variant useful in performing the methods of the invention is an allelic variant of a nucleic acid encoding a CDKB-RKA 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 of SEQ ID NO: 317, 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 or of SEQ ID NO: 317.

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: 317 or an allelic variant of a nucleic acid encoding an orthologue or paralogue of SEQ ID NO: 318. Preferably, the amino acid sequence encoded by the allelic variant, has reduced, and preferably lacks kinase activity, and (i) comprises a cyclin binding domain and/or (ii) has at least about 30-40%, preferably at least about 41-50%, more preferably at least about 51-60%, 61-70%, 71-80%, 81-85%, 86-90%, 91-95%, 96-99% or more than 99% overall sequence identity to SEQ ID NO: 318.

Gene shuffling or directed evolution may also be used to generate variants of nucleic acids encoding CDKB-RKA 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 of SEQ ID NO: 317, 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 or of SEQ ID NO: 317, which variant nucleic acid is obtained by

gene shuffling.

Preferably, the amino acid sequence encoded by the variant nucleic acid obtained by gene shuffling, has reduced, and preferably lacks kinase activity, and (i) comprises a cyclin binding domain and/or (ii) has at least about 30-40%, preferably at least about 41-50%,
5 more preferably at least about 51-60%, 61-70%, 71-80%, 81-85%, 86-90%, 91-95%, 96-99% or more than 99% overall sequence identity to SEQ ID NO: 318.

Furthermore, nucleic acid variants may also be obtained by site-directed mutagenesis. Several methods are available to achieve site-directed mutagenesis, the most common
10 being PCR based methods (Current Protocols in Molecular Biology. Wiley Eds.). CDKB-RKA polypeptides differing from the sequence of SEQ ID NO: 318 by one or several amino acids (substitution(s), insertion(s) and/or deletion(s) as defined above) may equally be useful to increase the yield of plants in the methods and constructs and plants of the invention.

15 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 natural genetic environment. In a
20 further embodiment the recombinant chromosomal DNA of the invention is comprised in a plant cell.

Performance of the methods of the invention gives plants having enhanced yield-related traits. In particular performance of the methods of the invention gives plants having
25 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. In another embodiment, performance of the methods of the invention gives plants having increased yield, especially increased vegetative biomass, such as increased leaf and/or root biomass relative to control plants. The terms "vegetative biomass" is described in more detail in the
30 "definitions" section herein.

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

The present invention provides a method for increasing yield-related traits, and preferably
40 yield, and more preferably seed yield and biomass, especially increased vegetative biomass, such as increased leaf and/or root biomass, relative to control plants, which method comprises modulating expression in a plant of a nucleic acid encoding a CDKB-

RKA 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 CDKB-RKA 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 CDKB-RKA polypeptide as defined herein.

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 CDKB-RKA polypeptide as defined herein.

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 CDKB-RKA polypeptide as defined herein.

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 CDKB-RKA polypeptide as defined herein.

The invention also provides genetic constructs and vectors to facilitate introduction and/or expression in plants of nucleic acids encoding CDKB-RKA polypeptides as defined herein. 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 CDKB-RKA 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 CDKB-RKA polypeptide is as defined above. The term “control sequence” and “termination sequence” are as defined herein.

The genetic construct of the invention may be comprised in a host cell, plant cell, seed, agricultural product or plant. 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 genetic construct such as a vector or 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 genetic construct in order to successfully transform, select and propagate host cells containing the sequence of interest. The sequence of interest is operably linked to one or more control sequences (at least to a promoter).

In one embodiment the genetic construct of the invention confers increased yield or yield related traits(s) to a living plant cell when it has been introduced into said plant cell and express the nucleic acid encoding the CDKB-RKA polypeptide, comprised in the genetic construct.

It should be clear that the applicability of the present invention is not restricted to the CDKB-RKA polypeptide-encoding nucleic acid represented by SEQ ID NO: 317.

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 medium strength promoter. In another embodiment said constitutive promoter is an ubiquitous constitutive promoter of medium strength. See the “Definitions” section herein for definitions of the various promoter types.

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 medium strength constitutive promoter, operably linked to the nucleic acid encoding the CDKB-RKA polypeptide as defined herein. 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,
5 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 CDKB-RKA polypeptide as defined herein is by introducing and expressing in a plant a nucleic acid encoding a CDKB-RKA polypeptide; however the effects of performing
10 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
15 enhanced yield-related traits relative to control plants, comprising introduction and expression in a plant of any nucleic acid encoding a CDKB-RKA polypeptide as defined herein.

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

- (i) introducing and expressing in a plant or plant cell a CDKB-RKA polypeptide-
encoding nucleic acid as defined herein or a genetic construct comprising a
25 CDKB-RKA polypeptide-encoding nucleic acid as defined herein; 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
30 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 CDKB-RKA polypeptide as defined herein.

35 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.

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

invention encompasses plants or parts thereof (including seeds) obtainable by the methods according to the present invention. The plants or plant parts or plant cells comprise a nucleic acid transgene encoding a CDKB-RKA polypeptide as defined above, preferably in a genetic construct such as an expression cassette. The present invention extends further
5 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.

- 10 In a further embodiment invention extends to seeds comprising the expression cassettes of the invention, the genetic constructs of the invention, the nucleic acids encoding the CDKB-RKA and/or the CDKB-RKA encoded by the nucleic acids as described above.

In a particular embodiment the plant cells of the invention are non-propagative cells, i.e.
15 cells that are not capable to regenerate into a plant using cell culture techniques known in the art. While plants cells generally have the characteristic of totipotency, some plant cells can not be used to regenerate or propagate intact plants from said cells. In one embodiment of the invention the plant cells of the invention are such cells.

- 20 In another embodiment the plant cells of the invention are plant cells that do not sustain themselves in an autotrophic way, such plant cells are not deemed to represent a plant variety. In a further embodiment the plant cells of the invention are non-plant variety and non-propagative.

- 25 The invention also includes host cells containing an isolated nucleic acid encoding a CDKB-RKA polypeptide as defined hereinabove. 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
30 polypeptides used in the inventive method. In a particular embodiment the plant cells of the invention overexpress the nucleic acid molecule of the invention.

The 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
35 invention or parts, including seeds, of these plants. In a further embodiment the methods comprises the steps of 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 with the harvestable parts of the invention.

- 40 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 the methods of the invention are plant products such as, but not limited to, a foodstuff, feedstuff, a food supplement, feed supplement, fiber, cosmetic or pharmaceutical. In another embodiment the 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.

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

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

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 CDKB-RKA 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.

The present invention also encompasses use of nucleic acids encoding CDKB-RKA polypeptides as described herein and use of these CDKB-RKA polypeptides in enhancing

any of the aforementioned yield-related traits in plants. For example, nucleic acids encoding CDKB-RKA polypeptide described herein, or the CDKB-RKA polypeptides themselves, may find use in breeding programs in which a DNA marker is identified which may be genetically linked to a CDKB-RKA polypeptide-encoding gene. The nucleic acids/genes, or the CDKB-RKA polypeptides themselves may be used to define a molecular marker. This DNA or protein marker may then be used in breeding programs to select plants having enhanced yield-related traits as defined hereinabove in the methods of the invention. Furthermore, allelic variants of a CDKB-RKA polypeptide-encoding nucleic acid/gene may find use in marker-assisted breeding programs. Nucleic acids encoding CDKB-RKA 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.

Embodiments

1. WAK-like polypeptides

The present invention is further characterized by one or more of the following embodiments.

1. A method for enhancing yield-related traits in plants relative to control plants, comprising modulating expression in a plant of a nucleic acid encoding a WAK-like polypeptide, wherein said WAK-like polypeptide comprises an EGF domain and a transmembrane domain.
2. Method according to embodiment 1, wherein said modulated expression is effected by introducing and expressing in a plant said nucleic acid encoding said WAK-like polypeptide.
3. Method according to embodiment 1 or 2, wherein said enhanced yield-related traits comprise increased yield relative to control plants, and preferably comprise increased biomass and/or increased seed yield relative to control plants.
4. Method according to any one of embodiments 1 to 3, wherein said enhanced yield-related traits are obtained under non-stress conditions.
5. Method according to any one of embodiments 1 to 3, wherein said enhanced yield-related traits are obtained under conditions of drought stress or nitrogen deficiency.
6. Method according to any one of embodiments 1 to 5, wherein said nucleic acid encoding a WAK-like polypeptide 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*.
7. Method according to any one of embodiments 1 to 6, wherein said nucleic acid encoding a WAK-like 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, provided these nucleic acids encode polypeptides comprising an EGF domain and a transmembrane domain.
8. Method according to any one of embodiments 1 to 7, wherein said nucleic acid sequence encodes an orthologue or paralogue of any of the polypeptides given in

Table A1, provided these nucleic acids encode polypeptides comprising an EGF domain and a transmembrane domain.

9. Method according to any one of embodiments 1 to 8, wherein said nucleic acid encodes the polypeptide represented by SEQ ID NO: 2 or 4.

5 10. Method according to any one of embodiments 1 to 9, 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.

10 11. Plant, plant part thereof, including seeds, or plant cell, obtainable by a method according to any one of embodiments 1 to 10, wherein said plant, plant part or plant cell comprises a recombinant nucleic acid encoding a WAK-like polypeptide as defined in any of embodiments 1 and 4 to 9.

12. Construct comprising:

- 15 (i) nucleic acid encoding a WAK-like as defined in any of embodiments 1 and 4 to 9;
- (ii) one or more control sequences capable of driving expression of the nucleic acid sequence of (i); and optionally
- (i) a transcription termination sequence.

20 13. Construct according to embodiment 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.

25 14. Use of a construct according to embodiment 12 or 13 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.

15. Plant, plant part or plant cell transformed with a construct according to embodiment 12 or 13.

30 16. 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:

- 35 (i) introducing and expressing in a plant cell or plant a nucleic acid encoding a WAK-like polypeptide as defined in any of embodiments 1 and 4 to 9; and
- (ii) cultivating said plant cell or plant under conditions promoting plant growth and development.

40 17. 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 WAK-like polypeptide as defined in any of embodiments 1 and 4 to 9 or a transgenic plant cell derived from said transgenic plant.

18. Transgenic plant according to embodiment 11, 15 or 17, 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.

- 5 19. Harvestable parts of a plant according to embodiment 18, wherein said harvestable parts are preferably shoot biomass and/or seeds.
20. Products derived from a plant according to embodiment 18 and/or from harvestable parts of a plant according to embodiment 19.
21. Use of a nucleic acid encoding a WAK-like polypeptide as defined in any of
10 embodiments 1 and 4 to 9 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.
22. A method for the production of a product comprising the steps of growing the plants according to embodiment 11, 14, 15, 17 or 18 and producing said product from or by
15 (i) said plants; or
(ii) parts, including seeds, of said plants.
23. Construct according to embodiment 12 or 13 comprised in a plant cell.
24. Recombinant chromosomal DNA comprising the construct according to embodiment 12 or 13.

20

2. CDKB-RKA polypeptides

The present invention is characterized by one or more of the following embodiments.

1. A method for increasing yield in a plant relative to a control plant, comprising introducing and expressing in said plant a nucleic acid molecule encoding a CDKB-RKA polypeptide, wherein said CDKB-RKA polypeptide is a modified CDKB polypeptide, and wherein said CDKB-RKA polypeptide has reduced catalytic kinase activity, as compared to said CDKB polypeptide.
25
2. Method according to embodiment 1, wherein said CDKB-RKA polypeptide lacks catalytic kinase activity.
- 30 3. Method according to embodiment 1 or 2, wherein said CDKB-RKA polypeptide is a mutated CDKB polypeptide.
4. Method according to embodiment 1, 2 or 3, wherein said CDKB-RKA polypeptide is a CDKB polypeptide comprising a mutated kinase domain.
5. Method according to any of embodiments 1 to 4, wherein said CDKB-RKA polypeptide is a truncated CDKB polypeptide.
35
6. Method according to any of embodiments 1 to 5, wherein said CDKB-RKA polypeptide is a truncated CDKB polypeptide comprising a deletion in the C-terminal half of said polypeptide that reduces, and preferably substantially inactivates, that kinase activity of said CDKB.
- 40 7. Method according to any of embodiments 1 to 6 wherein said CDKB-RKA polypeptide is a truncated CDKB polypeptide comprising a deletion of the kinase domain of said CDKB polypeptide.

8. Method according to any of embodiments 1 to 6 wherein said CDKB-RKA polypeptide is a truncated CDKB polypeptide comprising a deletion of a part of the kinase domain of said CDKB polypeptide, wherein said part preferably comprises at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 or 12 amino acids.
- 5 9. Method according to any of embodiments 1 to 8, wherein said kinase domain of said CDKB polypeptide is represented by a conserved domain from amino acid 140 to 152 in SEQ ID NO: 320.
10. Method according to any of embodiments 1 to 9 wherein said CDKB polypeptide has cyclin-binding activity.
- 10 11. Method according to any of embodiments 1 to 10, wherein said CDKB-RKA polypeptide comprises a cyclin binding domain, preferably represented by a PPTALRE motif with no mismatches or with a mismatch at position 2 and/or 4 from left to right, and more preferably represented by SEQ ID NO: 403, and more preferably represented by SEQ ID NO: 404.
- 15 12. Method according to any of embodiments 1 to 11, wherein said CDKB-RKA polypeptide comprises a protein kinase inhibitory domain.
13. Method according to any of embodiments 1 to 12, wherein said CDKB-RKA polypeptide lacks a T-loop activation kinase domain.
14. Method according to any one of embodiments 1 to 13, wherein said nucleic acid
20 encoding an CDKB-RKA is selected from any one of
 - a nucleic acid derived from a nucleic acid encoding any one of the polypeptides listed in Table A2;
 - a nucleic acid derived from an orthologue or paralogue of any of the polypeptides given in Table A2;
 - 25 - a portion of a nucleic acid encoding any one of the polypeptides listed in Table A2 or of an orthologue or paralogue of any of the polypeptides given in Table A2;
 - a nucleic acid capable of hybridising with a nucleic acid encoding any one of the polypeptides listed in Table A2 or of an orthologue or paralogue of any of the
30 polypeptides given in Table A2.
15. Method according to any one of embodiments 1 to 14, wherein said nucleic acid encodes a polypeptide represented by SEQ ID NO: 318 or a homologue, orthologue or paralogue thereof, or a portion of said nucleic acid, or a nucleic acid capable of hybridising with said nucleic acid.
- 35 16. Method according to any one of embodiments 1 to 15, wherein said nucleic acid encoding said CDKB 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*.
17. Method according to any one of embodiments 1 to 16, wherein said nucleic acid
40 encodes a polypeptide having has at least about 30-40%, preferably at least about 41-50%, more preferably at least about 51-60%, 61-70%, 71-80%, 81-85%, 86-90%, 91-95%, 96-99% or more than 99% overall sequence identity to SEQ ID NO: 318, and

preferably corresponding to SEQ ID NO: 318.

18. Method according to any of embodiments 1 to 17, wherein said increased yield relative to control plants comprises (i) increased seed yield and/or (ii) increased vegetative biomass, such as increased leaf biomass and/or increased root biomass.
- 5 19. Method according to any one of embodiments 1 to 18, wherein said enhanced yield-related traits are obtained under non-stress conditions.
20. Method according to any one of embodiments 1 to 19 wherein said nucleic acid is operably linked to a constitutive promoter, preferably to a constitutive plant promoter, more preferably to a medium strength constitutive plant promoter.
- 10 21. Transgenic plant, transgenic plant part thereof, including seeds, or transgenic plant cell, obtainable by a method according to any one of embodiments 1 to 20, wherein said plant, plant part or plant cell comprises a recombinant nucleic acid encoding a CDKB-RKA polypeptide as defined in any of embodiments 1 and 17.
22. Construct comprising:
 - 15 (i) nucleic acid encoding a CDKB-RKA polypeptide as defined in any of embodiments 1 to 17
 - (ii) one or more control sequences capable of driving expression of the nucleic acid sequence of (i); and optionally
 - (iii) a transcription termination sequence.
- 20 23. Construct according to embodiment 22, wherein one of said control sequences is a constitutive promoter, preferably to a constitutive plant promoter, more preferably to a medium strength constitutive plant promoter.
24. Use of a construct according to embodiment 22 or 23 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.
- 25 25. Transgenic plant, transgenic plant part or transgenic plant cell transformed with a construct according to embodiment 22 or 23.
26. 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:
 - 30 (a) introducing and expressing in a plant cell or plant a nucleic acid encoding a CDKB-RKA polypeptide as defined in any of embodiments 1 to 17; and
 - 35 (b) cultivating said plant cell or plant under conditions promoting plant growth and development.
27. 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 the introduction and expression therein of a nucleic acid encoding a CDKB-RKA polypeptide as defined in any of embodiments 1 to 17, or a transgenic plant cell derived from said transgenic plant.
- 40 28. Transgenic plant according to embodiment 21, 25 or 27, 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.

- 5 29. Harvestable parts of a plant according to any of embodiments 21, 25, 27 or 28 wherein said harvestable parts preferably are seeds.
30. Products derived from or produced from a plant according to embodiment 29 and/or from harvestable parts of a plant according to any of embodiments 21, 25, 27 or 28.
- 10 31. Use of a nucleic acid encoding a CDKB-RKA polypeptide as defined in any of embodiments 1 to 17 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.
- 15 32. A method for the production of a product comprising the steps of growing the plants according to any of embodiments 21, 25, 27 or 28 and producing said product from or by
 - (a) said plants; or
 - (b) parts, including seeds, of said plants.
33. Construct according to embodiment 22 or 23 comprised in a plant cell.
- 20 34. Recombinant chromosomal DNA comprising the construct according to embodiment 22 or 23.
- 25 35. Use of a nucleic acid encoding a CDKB-RKA polypeptide as defined in any of embodiments 1 to 17 as molecular marker for selecting for enhanced yield-related traits in plants relative to control plants, preferably for selecting for increased yield, and more preferably selecting for increasing seed yield and/or selecting for increasing biomass in plants relative to control plants.

In other embodiments, the invention is characterized by one or more of the following embodiments I to XV:

- 30 I. A method for increasing yield in a plant relative to a control plant, comprising introducing and expressing in said plant a nucleic acid molecule encoding a CDKB-RKA polypeptide, wherein said CDKB-RKA polypeptide is a modified CDKB polypeptide.
- II. Method according to embodiment I, wherein said CDKB-RKA polypeptide lacks catalytic kinase activity.
- 35 III. Method according to embodiment I or II, wherein said CDKB-RKA polypeptide is a mutated and/or a truncated CDKB polypeptide.
- IV. Method according to any of embodiments I to III, wherein said CDKB-RKA polypeptide is a truncated CDKB polypeptide comprising a deletion in the C-terminal half of said polypeptide that reduces, and preferably substantially inactivates, that kinase activity of said CDKB polypeptide.
- 40 V. Method according to any of embodiments I to IV, wherein said CDKB-RKA polypeptide is a truncated CDKB polypeptide comprising a deletion of the kinase

domain or a part thereof of said CDKB polypeptide.

5 VI. Method according to any of embodiments I to V, wherein said CDKB-RKA polypeptide comprises a cyclin binding domain represented by a PPTALRE motif with no mismatches or with a mismatch at position 2 and/or 4 from left to right, preferably represented by SEQ ID NO: 403, and more preferably represented by SEQ ID NO: 404.

10 VII. Method according to any one of embodiments I to VI, wherein said nucleic acid encodes a polypeptide represented by SEQ ID NO: 318 or a homologue, orthologue or paralogue thereof, or a portion of said nucleic acid, or a nucleic acid capable of hybridising with said nucleic acid.

VIII. Method according to any of embodiments I to VII, wherein said increased yield relative to control plants comprises (i) increased seed yield and/or (ii) increased vegetative biomass, such as increased leaf biomass and/or increased root biomass.

15 IX. Method according to any one of embodiments I to VIII, wherein said nucleic acid is operably linked to a constitutive promoter, preferably to a constitutive plant promoter, more preferably to a medium strength constitutive plant promoter.

20 X. Plant, plant part thereof, including seeds, or plant cell, obtainable by a method according to any one of embodiments I to IX, wherein said plant, plant part or plant cell comprises a recombinant nucleic acid encoding a CDKB-RKA polypeptide as defined in any of embodiments I to VII.

XI. Construct comprising:

- 25 (i) nucleic acid encoding a CDKB-RKA polypeptide as defined in any of embodiments I to VII;
- (ii) one or more control sequences capable of driving expression of the nucleic acid sequence of (i); wherein one of said control sequences is a constitutive promoter, preferably to a constitutive plant promoter, more preferably to a medium strength constitutive plant promoter and optionally
- (iii) a transcription termination sequence.

30 XII. Plant, plant part or plant cell transformed with a construct according to embodiment XI.

XIII. Method for the production of a transgenic plant having increased yield relative to control plants, and preferably increased seed yield and/or increased biomass such as increased vegetative biomass relative to control plants, comprising:

- 35 (i) introducing and expressing in a plant cell or plant a nucleic acid encoding a CDKB-RKA polypeptide as defined in any of embodiments I to VII; and
- (ii) cultivating said plant cell or plant under conditions promoting plant growth and development.

40 XIV. Transgenic plant having increased yield relative to control plants, and preferably increased seed yield and/or increased biomass such as increased vegetative biomass, resulting from the introduction and expression therein of a nucleic acid encoding a CDKB-RKA polypeptide as defined in any of embodiments I to VII or a transgenic plant cell derived from said transgenic plant.

XV. Use of a nucleic acid encoding a CDKB-RKA polypeptide as defined in any of embodiments I to VII 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 and/or use of a construct according to embodiment XI 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.

10 3. UPA20-like polypeptides

The present invention is further characterized by one or more of the following embodiments.

1. A method for enhancing yield-related traits in plants relative to control plants, comprising modulating expression in a plant of a nucleic acid encoding an UPA20-like polypeptide, wherein said UPA20-like polypeptide comprises a bHLH domain, and preferably comprises one or more domains selected from the group SM00353, PTHR12565:SF7, PTHR12565, PF00010, PS50888, SSF47459; and G3DSA:4.10.280.10.
2. Method according to embodiment 1, wherein said modulated expression is effected by introducing and expressing in a plant said nucleic acid encoding said UPA20-like polypeptide.
3. Method according to embodiment 1 or 2, wherein said enhanced yield-related traits comprise increased yield relative to control plants, and preferably comprise increased seed yield and/increase biomass relative to control plants.
4. Method according to any one of embodiments 1 to 3, wherein said increased seed yield comprises any one or more of increased total seed weight, increased number of seeds, increased number of flowers per panicle, increased thousand kernel weight, and increased fill rate.
5. Method according to any one of embodiments 1 to 4, wherein said enhanced yield-related traits are obtained under non-stress conditions.
6. Method according to any one of embodiments 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 embodiments 1 to 6, wherein said bHLH domain comprises an amino acid sequence having at least 50% overall sequence identity to the amino acid represented by SEQ ID NO: 544.
8. Method according to any of embodiments 1 to 7, wherein said UPA20-like polypeptide comprises one or more of the following motifs:
 - (i) Motif 2: YIHVRARRGQATDSHSLAERVRR[KR]ISERM[KR][LIF]LQ[DL]LVPGC [ND]K[IV]TGKA[LV]ML (SEQ ID NO: 538),
 - (ii) Motif 3: DEIINYVQSLQ[RN]QVEFLSMKLA[ST]V[NS]P[RLV]L[DY] (SEQ ID NO: 539),
 - (iii) Motif 4: [MN][AS][RK]KRK[SA][RA]x[KG][FGS] (SEQ ID NO: 540)

9. Method according to any of embodiments 1 to 8, wherein said UPA20-like polypeptide further comprises one or more of the following motifs:
 - (i) Motif 5: RVRRE[KR]ISERM[RK][MIV]LQ[ARLS]LVP GCDK[IV]TGKAL[MI]LDEIIN YVQSLQNQVEFLSM (SEQ ID NO: 541),
 - 5 (ii) Motif 6: [KRS]KV[HE][EGD]E[PAE]P[KAT]GYIHVRARRGQATDSHSLAE (SEQ ID NO: 542),
 - (iii) Motif 7: [KR][LI]AS[LMV][SN]P[VLM][LF]Y[DG]FGMD[LSR] (SEQ ID NO: 543)
10. Method according to any of embodiments 1 to 9, wherein said UPA20-like polypeptide comprises a conserved domain (or motif) with at least 50% sequence identity to a conserved domain selected from any one of:
 - 10 (i) amino acid 186 to 236 in SEQ ID NO:2 (motif 8 - SEQ ID NO: 556);
 - (ii) amino acid 174 to 231 in SEQ ID NO:2 (motif 9 - SEQ ID NO: 557);
 - (iii) amino acid 183 to 230 in SEQ ID NO:2 (motif 10 - SEQ ID NO: 558);
 - (iv) amino acid 181 to 231 in SEQ ID NO:2 (motif 11 - SEQ ID NO: 559);
 - 15 (v) amino acid 176 to 255 in SEQ ID NO:2 (motif 12 - SEQ ID NO: 560) ; and
 - (vi) amino acid 176 to 257 in SEQ ID NO:2 (motif 13 - SEQ ID NO: 561).
11. Method according to any one of embodiments 1 to 10, wherein said nucleic acid encoding a UPA20-like polypeptide is of plant origin, preferably from a dicotyledonous plant, further preferably from the family *Salicaceae*, more preferably from the genus *Populus*, most preferably the nucleic acid is from *Populus trichocarpa*.
- 20 12. Method according to any one of embodiments 1 to 11, wherein said nucleic acid encoding a UPA20-like polypeptide 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 such a nucleic acid.
- 25 13. Method according to any one of embodiments 1 to 12, wherein said nucleic acid sequence encodes an orthologue or paralogue of any of the polypeptides given in Table A3.
14. Method according to any one of embodiments 1 to 13, wherein said nucleic acid encodes the polypeptide represented by SEQ ID NO: 407 or a homologue thereof.
- 30 15. Method according to any one of embodiments 1 to 14, 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.
16. Plant, plant part thereof, including seeds, or plant cell, obtainable by a method according to any one of embodiments 1 to 15, wherein said plant, plant part or plant cell comprises a recombinant nucleic acid encoding a UPA20-like polypeptide as defined in any of embodiments 1 and 7 to 14.
- 35 17. Plant according to embodiment 16, or a 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, secale, einkorn, teff, milo or oats.
- 40 18. Construct comprising:

- (i) nucleic acid encoding a UPA20-like polypeptide as defined in any of embodiments 1 and 7 to 14;
 - (ii) one or more control sequences capable of driving expression of the nucleic acid sequence of (i); and optionally
 - 5 (iii) a transcription termination sequence.
19. Construct according to embodiment 18, 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.
- 10 20. Use of a construct according to embodiment 18 or 19 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.
- 15 21. Plant, plant part or plant cell transformed with a construct according to embodiment 18 or 19.
22. Plant according to embodiment 21, or a 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, secale, einkorn, teff, milo or oats.
- 20 23. 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:
- 25 (i) introducing and expressing in a plant cell or plant a nucleic acid encoding a UPA20-like polypeptide as defined in any of embodiments 1 and 7 to 14; and
 - (ii) cultivating said plant cell or plant under conditions promoting plant growth and development.
24. Transgenic plant having enhanced yield-related traits relative to control plants, preferably increased yield relative to control plants, and more preferably increased seed yield, resulting from modulated expression of a nucleic acid encoding a UPA20-like polypeptide as defined in any of embodiments 1 and 7 to 14 or a transgenic plant cell derived from said transgenic plant.
- 30 25. Transgenic plant according to embodiment 24, 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, secale, einkorn, teff, milo or oats.
- 35 26. Harvestable parts of a plant according to any of embodiments 16-17, 21-22, and 24-25, wherein said harvestable parts are preferably shoot biomass and/or seeds.
27. Products derived from a plant according to any of embodiments 16-17, 21-22, and 24-40 25, and/or from harvestable parts of a plant according to embodiment 26.
28. Isolated nucleic acid molecule selected from:
- (i) a nucleic acid represented by any one of SEQ ID NO: 458, 460 and 496;

- (ii) the complement of a nucleic acid represented by any one of SEQ ID NO: 458, 460 and 496;
- (iii) a nucleic acid encoding a UPA20-like 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: 459, 461 and 497, and additionally or alternatively comprising one or more motifs having in increasing order of preference at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more sequence identity to any one or more of the motifs given in SEQ ID NO: 538, 539, 540, 541, 542, 543, 556, 557, 558, 559, 560, and 561 (motifs 2 to 13), 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.

29. Isolated polypeptide selected from:

- (i) an amino acid sequence represented by any one of SEQ ID NO: 459, 461 and 497;
- (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: 459, 461 and 497, and additionally or alternatively comprising one or more motifs having in increasing order of preference at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more sequence identity to any one or more of the motifs given in SEQ ID NO: 538, 539, 540, 541, 542, 543, 556, 557, 558, 559, 560, and 561 (motifs 2 to 13), 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.

30. Use of a nucleic acid encoding a UPA20-like polypeptide as defined in any of embodiments 1 and 7 to 14 and 28 for enhancing yield-related traits in plants relative to control plants, preferably for increasing yield, and more preferably for increasing seed yield in plants relative to control plants.

31. Use of a nucleic acid as defined in embodiment 28 and encoding a UPA20-like polypeptide for enhancing yield-related traits in plants relative to control plants, preferably for increasing yield, and more preferably for increasing seed yield in plants relative to control plants.

32. Use of a nucleic acid encoding a UPA20-like polypeptide as defined in any of embodiments 1 and 7 to 14 and 29 as molecular marker.
33. Use of a nucleic acid as defined in embodiment 28 and encoding a UPA20-like polypeptide as defined in any of embodiments 1 and 7 to 14 and 29 as molecular marker.

More particularly, the invention is characterized by one or more of the following embodiments 1 to 12.

1. A method for the production of a transgenic plant having enhanced seed yield relative to a control plant, comprising the steps of:
 - (i) introducing and expressing in a plant cell or plant a nucleic acid encoding an UPA20-like polypeptide, wherein said nucleic acid is operably linked to a constitutive plant promoter, and wherein said UPA20-like polypeptide comprises the polypeptide represented by SEQ ID NO: 407 or a homologue thereof which has at least 80% overall sequence identity to SEQ ID NO : 407, and
 - (ii) cultivating said plant cell or plant under conditions promoting plant growth and development

wherein said increased seed yield comprises at least two parameters selected from the group comprising of increased total seed weight, increased number of seeds, increased number of flowers per panicle, and increased fill rate.
2. Method according to embodiment 1, wherein said nucleic acid is operably linked to a GOS2 promoter.
3. Method according to embodiment 1 or 2, wherein said GOS2 promoter is a GOS2 promoter from rice.
4. Method according to any of embodiments 1 to 3, wherein said increased seed yield comprises at least three parameters selected from the group comprising of increased total seed weight, increased number of seeds, increased number of flowers per panicle, and increased fill rate.
5. Method according to any of embodiments 1 to 4, wherein said increased seed yield consist of an increase in all parameters from the group consisting of increased total seed weight, increased number of seeds, increased number of flowers per panicle, and increased fill rate.
6. Method according to any of embodiments 1 to 5, wherein said increase in seed yield comprises an increase of at least 5 % in said plant as compared to control plants of each of said parameters.
7. Method according to any of embodiments 1 to 6, wherein said increase in seed yield further comprises an increased thousand kernel weight of said plant as compared to a control plant.
8. Method according to any of embodiments 1 to 7, wherein said increased yield is obtained under non-stress conditions.
9. Method according to any one for embodiments 1 to 8, wherein said plant is a

monocotyledonous plant.

10. Method according to embodiment 9, wherein said plant is a cereal.
11. Transgenic plant having enhanced seed yield as defined in any of embodiments 4 to 7 relative to control plants, resulting from introduction and expression of a nucleic acid encoding a UPA20-like polypeptide as defined in embodiment 1 in said plant, or a transgenic plant cell derived from said transgenic plant.
12. Use of a nucleic acid encoding a UPA20-like polypeptide as defined in embodiment 1 for enhancing seed yield as defined in any of embodiments 4 to 7 in a transgenic plant relative to a control plant.

Definitions

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

Peptide(s)/Protein(s)

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

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

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

Homologue(s)

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

Orthologues and paralogues are two different forms of homologues and encompass

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

5

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

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

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

Table 1: Examples of conserved amino acid substitutions

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

Amino acid substitutions, deletions and/or insertions may readily be made using peptide synthetic techniques known in the art, such as solid phase peptide synthesis and the like, or

by recombinant DNA manipulation. Methods for the manipulation of DNA sequences to produce substitution, insertion or deletion variants of a protein are well known in the art. For example, techniques for making substitution mutations at predetermined sites in DNA are well known to those skilled in the art and include M13 mutagenesis, T7-Gen in vitro
5 mutagenesis (USB, Cleveland, OH), QuickChange Site Directed mutagenesis (Stratagene, San Diego, CA), PCR-mediated site-directed mutagenesis or other site-directed mutagenesis protocols (see Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989 and yearly updates)).

10 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
15 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
20 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
25 peptides such as FLAG, HIS6 or thioredoxin (for a review of tagging peptides, see Terpe, Appl. Microbiol. Biotechnol. 60, 523-533, 2003).

Domain, Motif/Consensus sequence/Signature

The term “domain” refers to a set of amino acids conserved at specific positions along an alignment of sequences of evolutionarily related proteins. While amino acids at other
30 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
35 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
40 conserved domain (if all of the amino acids of the motif fall outside of a defined domain).

Specialist databases exist for the identification of domains, for example, SMART (Schultz et

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

Methods for the alignment of sequences for comparison are well known in the art, such methods include GAP, BESTFIT, BLAST, FASTA and TFASTA. GAP uses the algorithm of Needleman and Wunsch ((1970) J Mol Biol 48: 443-453) to find the global (i.e. spanning the complete sequences) alignment of two sequences that maximizes the number of matches and minimizes the number of gaps. The BLAST algorithm (Altschul et al. (1990) J Mol Biol 215: 403-10) calculates percent sequence identity and performs a statistical analysis of the similarity between the two sequences. The software for performing BLAST analysis is publicly available through the National Centre for Biotechnology Information (NCBI). Homologues may readily be identified using, for example, the ClustalW multiple sequence alignment algorithm (version 1.83), with the default pairwise alignment parameters, and a scoring method in percentage. Global percentages of similarity and identity may also be determined using one of the methods available in the MatGAT software package (Campanella et al., BMC Bioinformatics. 2003 Jul 10;4:29. MatGAT: an application that generates similarity/identity matrices using protein or DNA sequences.). Minor manual editing may be performed to optimise alignment between conserved motifs, as would be apparent to a person skilled in the art. Furthermore, instead of using full-length sequences for the identification of homologues, specific domains may also be used. The sequence identity values may be determined over the entire nucleic acid or amino acid sequence or over selected domains or conserved motif(s), using the programs mentioned above using the default parameters. For local alignments, the Smith-Waterman algorithm is particularly useful (Smith TF, Waterman MS (1981) J. Mol. Biol 147(1);195-7).

Reciprocal BLAST

Typically, this involves a first BLAST involving BLASTing a query sequence (for example using any of the sequences listed in Tables A of the Examples section) against any sequence database, such as the publicly available NCBI database. BLASTN or TBLASTX (using standard default values) are generally used when starting from a nucleotide sequence, and BLASTP or TBLASTN (using standard default values) when starting from a protein sequence. The BLAST results may optionally be filtered. The full-length sequences

of either the filtered results or non-filtered results are then BLASTed back (second BLAST) against sequences from the organism from which the query sequence is derived. The results of the first and second BLASTs are then compared. A paralogue is identified if a high-ranking hit from the first blast is from the same species as from which the query sequence is derived, a BLAST back then ideally results in the query sequence amongst the highest hits; an orthologue is identified if a high-ranking hit in the first BLAST is not from the same species as from which the query sequence is derived, and preferably results upon BLAST back in the query sequence being among the highest hits.

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

Hybridisation

- The term "hybridisation" as defined herein is a process wherein substantially homologous complementary nucleotide sequences anneal to each other. The hybridisation process can occur entirely in solution, i.e. both complementary nucleic acids are in solution. The hybridisation process can also occur with one of the complementary nucleic acids immobilised to a matrix such as magnetic beads, Sepharose beads or any other resin. The hybridisation process can furthermore occur with one of the complementary nucleic acids immobilised to a solid support such as a nitro-cellulose or nylon membrane or immobilised by e.g. photolithography to, for example, a siliceous glass support (the latter known as nucleic acid arrays or microarrays or as nucleic acid chips). In order to allow hybridisation to occur, the nucleic acid molecules are generally thermally or chemically denatured to melt a double strand into two single strands and/or to remove hairpins or other secondary structures from single stranded nucleic acids.

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

medium stringency hybridisation conditions may sometimes be needed to identify such nucleic acid molecules.

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

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

2) DNA-RNA or RNA-RNA hybrids:

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

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

$$\text{For } <20 \text{ nucleotides: } T_m = 2 (I_n)$$

$$\text{For } 20\text{--}35 \text{ nucleotides: } T_m = 22 + 1.46 (I_n)$$

^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; I_n , = effective length of primer = $2 \times (\text{no. of G/C}) + (\text{no. of A/T})$.

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.

- 5 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.
- 10 For example, typical high stringency hybridisation conditions for DNA hybrids longer than 50 nucleotides encompass hybridisation at 65°C in 1x SSC or at 42°C in 1x SSC and 50% formamide, followed by washing at 65°C in 0.3x SSC. Examples of medium stringency hybridisation conditions for DNA hybrids longer than 50 nucleotides encompass hybridisation at 50°C in 4x SSC or at 40°C in 6x SSC and 50% formamide, followed by
- 15 washing at 50°C in 2x SSC. The length of the hybrid is the anticipated length for the hybridising nucleic acid. When nucleic acids of known sequence are hybridised, the hybrid length may be determined by aligning the sequences and identifying the conserved regions described herein. 1×SSC is 0.15M NaCl and 15mM sodium citrate; the hybridisation solution and wash solutions may additionally include 5x Denhardt's reagent, 0.5-1.0% SDS,
- 20 100 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).

25

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

30 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).

35

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

40 is usually less than 100 bp. SNPs and INDELs form the largest set of sequence variants in naturally occurring polymorphic strains of most organisms.

Endogenous gene

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

Gene shuffling/Directed evolution

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

Construct

Artificial DNA (such as but, not limited to plasmids or viral DNA) capable of replication in a host cell and used for introduction of a DNA sequence of interest into a host cell or host organism. Host cells of the invention may be any cell selected from bacterial cells, such as Escherichia coli or Agrobacterium species cells, yeast cells, fungal, algal or cyanobacterial cells or plant cells. The skilled artisan is well aware of the genetic elements that must be present on the genetic construct in order to successfully transform, select and propagate host cells containing the sequence of interest. The sequence of interest is operably linked to one or more control sequences (at least to a promoter) as described herein. Additional regulatory elements may include transcriptional as well as translational enhancers. Those skilled in the art will be aware of terminator and enhancer sequences that may be suitable for use in performing the invention. An intron sequence may also be added to the 5' untranslated region (UTR) or in the coding sequence to increase the amount of the mature message that accumulates in the cytosol, as described in the definitions section. Other control sequences (besides promoter, enhancer, silencer, intron sequences, 3'UTR and/or 5'UTR regions) may be protein and/or RNA stabilizing elements. Such sequences would be known or may readily be obtained by a person skilled in the art.

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

For the detection of the successful transfer of the nucleic acid sequences as used in the methods of the invention and/or selection of transgenic plants comprising these nucleic

acids, it is advantageous to use marker genes (or reporter genes). Therefore, the genetic construct may optionally comprise a selectable marker gene. Selectable markers are described in more detail in the “definitions” section herein. The marker genes may be removed or excised from the transgenic cell once they are no longer needed. Techniques for marker removal are known in the art, useful techniques are described above in the definitions section.

Regulatory element/Control sequence/Promoter

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

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

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

5 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

- 10 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.

15

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
20 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:

25

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 α -globulin Glb-1	Wu et al, Plant Cell Physiology 39(8) 885-889, 1998
rice OSH1	Sato et al, Proc. Natl. Acad. Sci. USA, 93: 8117-8122, 1996
rice α -globulin REB/OHP-1	Nakase et al. Plant Mol. Biol. 33: 513-522, 1997
rice ADP-glucose pyrophos- phorylase	Trans Res 6:157-68, 1997
maize ESR gene family	Plant J 12:235-46, 1997
sorghum α -kafirin	DeRose et al., Plant Mol. Biol 32:1029-35, 1996
KNOX	Postma-Haarsma et al, Plant Mol. Biol. 39:257-71, 1999
rice oleosin	Wu et al, J. Biochem. 123:386, 1998
sunflower oleosin	Cummins et al., Plant Mol. Biol. 19: 873-876, 1992
PRO0117, putative rice 40S ribosomal protein	WO 2004/070039
PRO0136, rice alanine aminotransferase	unpublished
PRO0147, trypsin inhibitor ITR1 (barley)	unpublished
PRO0151, rice WSI18	WO 2004/070039
PRO0175, rice RAB21	WO 2004/070039
PRO005	WO 2004/070039
PRO0095	WO 2004/070039
α -amylase (Amy32b)	Lanahan et al, Plant Cell 4:203-211, 1992; Skriver et al, Proc Natl Acad Sci USA 88:7266-7270, 1991
cathepsin β -like gene	Cejudo et al, Plant Mol Biol 20:849-856, 1992
Barley Ltp2	Kalla et al., Plant J. 6:849-60, 1994

Chi26	Leah et al., Plant J. 4:579-89, 1994
Maize B-Peru	Selinger et al., Genetics 149:1125-38,1998

Table 2d: examples of endosperm-specific promoters

Gene source	Reference
glutelin (rice)	Takaiwa et al. (1986) Mol Gen Genet 208:15-22; Takaiwa et al. (1987) FEBS Letts. 221:43-47
zein	Matzke et al., (1990) Plant Mol Biol 14(3): 323-32
wheat LMW and HMW glutenin-1	Colot et al. (1989) Mol Gen Genet 216:81-90, Anderson et al. (1989) NAR 17:461-2
wheat SPA	Albani et al. (1997) Plant Cell 9:171-184
wheat gliadins	Rafalski et al. (1984) EMBO 3:1409-15
barley 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 pyrophosphorylase	Russell et al. (1997) Trans Res 6:157-68
maize ESR gene family	Opsahl-Ferstad et al. (1997) Plant J 12:235-46
sorghum kafirin	DeRose et al. (1996) Plant Mol Biol 32:1029-35

Table 2e: Examples of embryo specific promoters:

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

5

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.

5

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	

10

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.

15

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 <i>et al.</i> (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
5 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

10 "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
15 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,
20 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
25 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/luciferase system) or fluorescence (Green Fluorescent Protein, GFP, and derivatives thereof). This list represents only a small number of possible markers. The
30 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,
35 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
40 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)

5 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
10 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
15 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.

20 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
25 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
30 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
35 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.

40 Modulation

The term "modulation" means in relation to expression or gene expression, a process in which the expression level is changed by said gene expression in comparison to the control

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

10 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
 5 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).

10

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
 15 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
 20 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
 25 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%,
 30 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.

35 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
 40 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
5 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
10 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
15 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
20 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
25 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
30 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
35 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
40 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 α -anomeric nucleic acid sequence. An α -anomeric nucleic acid sequence forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual β -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).

5 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
10 (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.,
15 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
20 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.

25 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
30 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
35 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
40 inhibition. Effects of miRNA overexpression are thus often reflected in decreased mRNA levels of target genes.

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

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

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

Transformation

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

The transfer of foreign genes into the genome of a plant is called transformation. Transformation of plant species is now a fairly routine technique. Advantageously, any of several transformation methods may be used to introduce the gene of interest into a suitable ancestor cell. The methods described for the transformation and regeneration of plants from plant tissues or plant cells may be utilized for transient or for stable transformation. Transformation methods include the use of liposomes, electroporation, chemicals that increase free DNA uptake, injection of the DNA directly into the plant, particle gun bombardment, transformation using viruses or pollen and microprojection. Methods may be selected from the calcium/polyethylene glycol method for protoplasts (Krens, F.A. et al., (1982) Nature 296, 72-74; Negrutiu I et al. (1987) Plant Mol Biol 8: 363-373); electroporation of protoplasts (Shillito R.D. et al. (1985) Bio/Technol 3, 1099-1102); microinjection into plant material (Crossway A et al., (1986) Mol. Gen Genet 202: 179-185); DNA or RNA-coated particle bombardment (Klein TM et al., (1987) Nature 327: 70) infection with (non-integrative) viruses and the like. Transgenic plants, including transgenic crop plants, are preferably produced via *Agrobacterium*-mediated transformation. An advantageous transformation method is the transformation *in planta*. To this end, it is possible, for example, to allow the agrobacteria to act on plant seeds or to inoculate the plant meristem with agrobacteria. It has proved particularly expedient in accordance with the invention to allow a suspension of transformed agrobacteria to act on the intact plant or at least on the flower primordia. The plant is subsequently grown on until the seeds of the treated plant are obtained (Clough and Bent, Plant J. (1998) 16, 735-743). Methods for *Agrobacterium*-mediated transformation of rice include well known methods for rice transformation, such as those described in any of the following: European patent application EP 1198985 A1, Aldemita and Hodges (Planta 199: 612-617, 1996); Chan et al. (Plant Mol Biol 22 (3): 491-506, 1993), Hiei et al. (Plant J 6 (2): 271-282, 1994), which disclosures are incorporated by reference herein as if fully set forth. In the case of corn transformation, the preferred method is as described in either Ishida et al. (Nat. Biotechnol 14(6): 745-50, 1996) or Frame et al. (Plant Physiol 129(1): 13-22, 2002), which disclosures are incorporated by reference herein as if fully set forth. Said methods are further described by way of example in B. Jenes et al., Techniques for Gene Transfer, in: Transgenic Plants, Vol. 1, Engineering and Utilization, eds. S.D. Kung and R. Wu, Academic Press (1993) 128-143 and in Potrykus Annu. Rev. Plant Physiol. Plant Molec. Biol. 42 (1991) 205-225). The nucleic acids or the construct to be expressed is preferably cloned into a vector, which is suitable for transforming *Agrobacterium tumefaciens*, for example pBin19 (Bevan et al., Nucl. Acids Res. 12 (1984) 8711). Agrobacteria transformed by such a vector can then be used in known manner for the transformation of plants, such as plants used as a model, like *Arabidopsis* (*Arabidopsis thaliana* is within the scope of the present invention not considered as a crop plant), or crop plants such as, by way of example, tobacco plants, for example by immersing bruised leaves or chopped leaves in an agrobacterial solution and then culturing them in suitable media. The transformation of plants by means of *Agrobacterium tumefaciens* is described, for example, by Höfgen and Willmitzer in Nucl.

Acid Res. (1988) 16, 9877 or is known inter alia from F.F. White, Vectors for Gene Transfer in Higher Plants; in Transgenic Plants, Vol. 1, Engineering and Utilization, eds. S.D. Kung and R. Wu, Academic Press, 1993, pp. 15-38.

- 5 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
- 10 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,
- 15 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
- 20 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
- 25 transformation of plastids is of advantages because plastids are inherited maternally in 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
- 30 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
- 35 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 marker gene (Klaus et al., 2004, Nature Biotechnology 22(2), 225-229).
- 40 The genetically modified plant cells can be regenerated via all methods with which the skilled worker is familiar. Suitable methods can be found in the abovementioned publications by S.D. Kung and R. Wu, Potrykus or Höfgen and Willmitzer. Alternatively, the

genetically modified plant cells are non-regenerable into a whole plant.

Generally after transformation, plant cells or cell groupings are selected for the presence of one or more markers which are encoded by plant-expressible genes co-transferred with the gene of interest, following which the transformed material is regenerated into a whole plant.

To select transformed plants, the plant material obtained in the transformation is, as a rule, subjected to selective conditions so that transformed plants can be distinguished from untransformed plants. For example, the seeds obtained in the above-described manner can be planted and, after an initial growing period, subjected to a suitable selection by spraying.

A further possibility consists in growing the seeds, if appropriate after sterilization, on agar plates using a suitable selection agent so that only the transformed seeds can grow into plants. Alternatively, the transformed plants are screened for the presence of a selectable marker such as the ones described above.

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

The generated transformed plants may be propagated by a variety of means, such as by clonal propagation or classical breeding techniques. For example, a first generation (or T1) transformed plant may be selfed and homozygous second-generation (or T2) transformants selected, and the T2 plants may then further be propagated through classical breeding techniques. The generated transformed organisms may take a variety of forms. For example, they may be chimeras of transformed cells and non-transformed cells; clonal transformants (e.g., all cells transformed to contain the expression cassette); grafts of transformed and untransformed tissues (e.g., in plants, a transformed rootstock grafted to an untransformed scion).

T-DNA activation tagging

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

TILLING

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

Homologous recombination

“Homologous recombination” allows introduction in a genome of a selected nucleic acid at a defined selected position. Homologous recombination is a standard technology used routinely in biological sciences for lower organisms such as yeast or the moss *Physcomitrella*. Methods for performing homologous recombination in plants have been described not only for model plants (Offringa et al. (1990) EMBO J 9(10): 3077-84) but also for crop plants, for example rice (Terada et al. (2002) Nat Biotech 20(10): 1030-4; Iida and Terada (2004) Curr Opin Biotech 15(2): 132-8), and approaches exist that are generally applicable regardless of the target organism (Miller et al, Nature Biotechnol. 25, 778-785, 2007).

Yield related Trait(s)

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

Reference herein to enhanced yield-related traits, relative to of control plants is taken to mean one or more of an increase in early vigour and/or in biomass (weight) of one or more parts of a plant, which may include (i) aboveground parts and preferably aboveground

harvestable parts and/or (ii) parts below ground and preferably harvestable below ground. In particular, such harvestable parts are seeds.

Yield

5 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.

10 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.

15 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
20 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.

25 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
30 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
35 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.
40 Flowering time of plants can be assessed by counting the number of days (“time to flower”) between sowing and the emergence of a first inflorescence. The “flowering time” of a plant can for instance be determined using the method as described in WO 2007/093444.

Early vigour

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

Increased growth rate

The increased growth rate may be specific to one or more parts of a plant (including seeds), or may be throughout substantially the whole plant. Plants having an increased growth rate may have a shorter life cycle. The life cycle of a plant may be taken to mean the time needed to grow from a mature seed up to the stage where the plant has produced mature seeds, similar to the starting material. This life cycle may be influenced by factors such as speed of germination, early vigour, growth rate, greenness index, flowering time and speed of seed maturation. The increase in growth rate may take place at one or more stages in the life cycle of a plant or during substantially the whole plant life cycle. Increased growth rate during the early stages in the life cycle of a plant may reflect enhanced vigour. The increase in growth rate may alter the harvest cycle of a plant allowing plants to be sown later and/or harvested sooner than would otherwise be possible (a similar effect may be obtained with earlier flowering time). If the growth rate is sufficiently increased, it may allow for the further sowing of seeds of the same plant species (for example sowing and harvesting of rice plants followed by sowing and harvesting of further rice plants all within one conventional growing period). Similarly, if the growth rate is sufficiently increased, it may allow for the further sowing of seeds of different plants species (for example the sowing and harvesting of corn plants followed by, for example, the sowing and optional harvesting of soybean, potato or any other suitable plant). Harvesting additional times from the same rootstock in the case of some crop plants may also be possible. Altering the harvest cycle of a plant may lead to an increase in annual biomass production per square meter (due to an increase in the number of times (say in a year) that any particular plant may be grown and harvested). An increase in growth rate may also allow for the cultivation of transgenic plants in a wider geographical area than their wild-type counterparts, since the territorial limitations for growing a crop are often determined by adverse environmental conditions either at the time of planting (early season) or at the time of harvesting (late season). Such adverse conditions may be avoided if the harvest cycle is shortened. The growth rate may be determined by deriving various parameters from growth curves, such parameters may be: T-Mid (the time taken for plants to reach 50% of their maximal size)

and T-90 (time taken for plants to reach 90% of their maximal size), amongst others.

Stress resistance

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

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

The “abiotic stress” may be an osmotic stress caused by a water stress, e.g. due to drought, salt stress, or freezing stress. Abiotic stress may also be an oxidative stress or a cold stress. “Freezing stress” is intended to refer to stress due to freezing temperatures, i.e. temperatures at which available water molecules freeze and turn into ice. “Cold stress”, also called “chilling stress”, is intended to refer to cold temperatures, e.g. temperatures below 10°, or preferably below 5°C, but at which water molecules do not freeze. As reported in Wang et al. (Planta (2003) 218: 1-14), abiotic stress leads to a series of morphological, physiological, biochemical and molecular changes that adversely affect plant growth and productivity. Drought, salinity, extreme temperatures and oxidative stress are known to be interconnected and may induce growth and cellular damage through similar mechanisms. Rabbani et al. (Plant Physiol (2003) 133: 1755-1767) describes a particularly high degree of “cross talk” between drought stress and high-salinity stress. For example, drought and/or salinisation are manifested primarily as osmotic stress, resulting in the disruption of homeostasis and ion distribution in the cell. Oxidative stress, which frequently accompanies high or low temperature, salinity or drought stress, may cause denaturing of functional and structural proteins. As a consequence, these diverse environmental stresses often activate similar cell signalling pathways and cellular responses, such as the production of stress proteins, up-regulation of anti-oxidants, accumulation of compatible solutes and growth arrest. The term “non-stress” conditions as used herein are those environmental conditions that allow optimal growth of plants. Persons skilled in the art are aware of normal soil conditions and climatic conditions for a given location. Plants with optimal growth

conditions, (grown under non-stress conditions) typically yield in increasing order of preference at least 97%, 95%, 92%, 90%, 87%, 85%, 83%, 80%, 77% or 75% of the average production of such plant in a given environment. Average production may be calculated on harvest and/or season basis. Persons skilled in the art are aware of average yield productions of a crop.

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

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

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

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

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

In yet another example, the methods of the present invention may be performed under stress conditions such as salt stress to give plants having increased yield relative to control plants. The term salt stress is not restricted to common salt (NaCl), but may be any one or more of: NaCl, KCl, LiCl, MgCl₂, CaCl₂, amongst others.

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

Increase/Improve/Enhance

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

Seed yield

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

- (a) an increase in seed biomass (total seed weight) which may be on an individual seed basis and/or per plant and/or per square meter;
- (b) increased number of flowers per plant;
- (c) increased number of seeds;
- (d) increased seed filling rate (which is expressed as the ratio between the number

of filled florets divided by the total number of florets);

(e) increased harvest index, which is expressed as a ratio of the yield of harvestable parts, such as seeds, divided by the biomass of aboveground plant parts; and

(f) increased thousand kernel weight (TKW), which is extrapolated from the number of seeds counted and their total weight. An increased TKW may result from an increased seed size and/or seed weight, and may also result from an increase in embryo and/or endosperm size.

The terms “filled florets” and “filled seeds” may be considered synonyms.

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

Greenness Index

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

Biomass

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

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

Marker assisted breeding

Such breeding programmes sometimes require introduction of allelic variation by mutagenic

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

Use as probes in (gene mapping)

Use of nucleic acids encoding the protein of interest for genetically and physically mapping the genes requires only a nucleic acid sequence of at least 15 nucleotides in length. These nucleic acids may be used as restriction fragment length polymorphism (RFLP) markers. Southern blots (Sambrook J, Fritsch EF and Maniatis T (1989) Molecular Cloning, A Laboratory Manual) of restriction-digested plant genomic DNA may be probed with the nucleic acids encoding the protein of interest. The resulting banding patterns may then be subjected to genetic analyses using computer programs such as MapMaker (Lander et al. (1987) Genomics 1: 174-181) in order to construct a genetic map. In addition, the nucleic acids may be used to probe Southern blots containing restriction endonuclease-treated genomic DNAs of a set of individuals representing parent and progeny of a defined genetic cross. Segregation of the DNA polymorphisms is noted and used to calculate the position of the nucleic acid encoding the protein of interest in the genetic map previously obtained using this population (Botstein et al. (1980) Am. J. Hum. Genet. 32:314-331).

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

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

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

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

Plant

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

Plants that are particularly useful in the methods of the invention include all plants which belong to the superfamily Viridiplantae, in particular monocotyledonous and dicotyledonous plants including fodder or forage legumes, ornamental plants, food crops, trees or shrubs selected from the list comprising *Acer* spp., *Actinidia* spp., *Abelmoschus* spp., *Agave sisalana*, *Agropyron* spp., *Agrostis stolonifera*, *Allium* spp., *Amaranthus* spp., *Ammophila arenaria*, *Ananas comosus*, *Annona* spp., *Apium graveolens*, *Arachis* spp., *Artocarpus* spp., *Asparagus officinalis*, *Avena* spp. (e.g. *Avena sativa*, *Avena fatua*, *Avena byzantina*, *Avena fatua* var. *sativa*, *Avena hybrida*), *Averrhoa carambola*, *Bambusa* sp., *Benincasa hispida*, *Bertholletia excelsea*, *Beta vulgaris*, *Brassica* spp. (e.g. *Brassica napus*, *Brassica rapa* ssp. [canola, oilseed rape, turnip rape]), *Cadaba farinosa*, *Camellia sinensis*, *Canna indica*, *Cannabis sativa*, *Capsicum* spp., *Carex elata*, *Carica papaya*, *Carissa macrocarpa*, *Carya* spp., *Carthamus tinctorius*, *Castanea* spp., *Ceiba pentandra*, *Cichorium endivia*, *Cinnamomum* spp., *Citrullus lanatus*, *Citrus* spp., *Cocos* spp., *Coffea* spp., *Colocasia esculenta*, *Cola* spp., *Corchorus* sp., *Coriandrum sativum*, *Corylus* spp., *Crataegus* spp., *Crocus sativus*, *Cucurbita* spp., *Cucumis* spp., *Cynara* spp., *Daucus carota*, *Desmodium* spp., *Dimocarpus longan*, *Dioscorea* spp., *Diospyros* spp., *Echinochloa* spp., *Elaeis* (e.g. *Elaeis guineensis*, *Elaeis oleifera*), *Eleusine coracana*, *Eragrostis tef*, *Erianthus* sp., *Eriobotrya japonica*, *Eucalyptus* sp., *Eugenia uniflora*, *Fagopyrum* spp., *Fagus* spp.,

Festuca arundinacea, *Ficus carica*, *Fortunella* spp., *Fragaria* spp., *Ginkgo biloba*, *Glycine* spp. (e.g. *Glycine max*, *Soja hispida* or *Soja max*), *Gossypium hirsutum*, *Helianthus* spp. (e.g. *Helianthus annuus*), *Hemerocallis fulva*, *Hibiscus* spp., *Hordeum* spp. (e.g. *Hordeum vulgare*), *Ipomoea batatas*, *Juglans* spp., *Lactuca sativa*, *Lathyrus* spp., *Lens culinaris*,
 5 *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.,
 10 *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.,
 15 *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.,
 20 *Tripsacum dactyloides*, *Triticosecale rimpai*, *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.

25

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
 30 plant to be assessed. The control plant may also be a nullizygote of the plant to be assessed. Nullizygotes (or null control plants) are individuals missing the transgene by segregation. Further, control plants are grown under equal growing conditions to the growing conditions of the plants of the invention, i.e. in the vicinity of, and simultaneously with, the plants of the invention. A "control plant" as used herein refers not only to whole
 35 plants, but also to plant parts, including seeds and seed parts.

Description of figures

The present invention will now be described with reference to the following figures in which:
Figure 1 represents the binary vector used for increased expression in *Oryza sativa* of a
 40 WAK-like-encoding nucleic acid under the control of a rice GOS2 promoter (pGOS2).
Figure 2 represents an alignment of a CDKB-RKA polypeptide represented by SEQ ID NO: 318 with a CDKB polypeptide (CDKB1,2) from *Arabidopsis thaliana* (AT2G38620.2; SEQ

ID NO: 320) with indication of conserved domains.

Figure 3 represents a multiple alignment of various full length CDKB 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. A conserved cyclin binding domain is represented with a black box.

Figure 4 represents the binary vector used for increased expression in *Oryza sativa* of a CDKB-RKA-encoding nucleic acid as defined herein under the control of a suitable promoter.

Figure 5 represents the domain structure of SEQ ID NO: 407 with conserved motifs 2 to 7 and bHLH domain (bold and underlined).

Figure 6 represents a multiple alignment of various UPA20-like polypeptides. The aligned sequences have the SEQ ID NOs as described in the legend of Figure 7. These alignments can be used for defining further motifs or signature sequences, when using conserved amino acids. The bHLH domain is indicated with a box.

Figure 7 shows the MATGAT table of Example 3. The represented polypeptides correspond to the following legend:

1. *P.trichocarpa* 575437; 2. *A.lyrata* 472791; 3. *A.lyrata* 476021; 4. *A.lyrata* 491205; 5. *A.lyrata* 494918; 6. *A.lyrata* 495154; 7. *A.thaliana* AT1G26260.1; 8. *A.thaliana* AT1G68920.1; 9. *A.thaliana* AT3G07340.1; 10. *A.thaliana* AT5G48560.1; 11. *A.thaliana* AT5G50915.1; 12. *Aquilegia* sp TC24120; 13. *Aquilegia* sp TC24680; 14. *Aquilegia* sp TC26724; 15. *B.napus* TC73386; 16. *C.annuum* NP13055023; 17. *C.annuum* TC15396; 18. *C.sinensis* TC686; 19. *G.hirsutum* TC144761; 20. *G.max* Glyma01g09400.1; 21. *G.max* Glyma02g13860.1; 22. *G.max* Glyma03g21770.1; 23. *G.max* Glyma04g37690.1; 24. *G.max* Glyma05g38450.1; 25. *G.max* Glyma06g17420.1; 26. *G.max* Glyma08g46040.1; 27. *G.max* Glyma18g32560.1; 28. *G.max* GM06MC18997 59691073@18643; 29. *G.raimondii* TC6315; 30. *G.raimondii* TC9358; 31. *H.tuberosus* TA2737 4233; 32. *M.domestica* TC28668; 33. *M.truncatula* AC150842 12.5; 34. *M.truncatula* AC171266 17.4; 35. *N.tabacum* FG645212; 36. *O.sativa* LOC Os04g28280.1; 37. *O.sativa* LOC Os05g01256.1; 38. *O.sativa* LOC Os06g16400.1; 39. *O.sativa* LOC Os08g41320.1; 40. *O.sativa* LOC Os08g42470.1; 41. *O.sativa* LOC Os09g32510.4; 42. *O.sativa* LOC Os09g33580.1; 43. *P.hybrida* TC3261; 44. *P.persica* TC11169; 45. *P.trichocarpa* 553223; 46. *P.trichocarpa* 572918; 47. *P.trichocarpa* 773249; 48. *P.virgatum* TC28308; 49. *P.vulgaris* TC12782 ; 50. *R.communis* TA2848 3988 ; 51. *S.bicolor* Sb02g029530.1 ; 52. *S.bicolor* Sb07g025920.1 ; 53. *S.lycopersicum* TC196509 ; 54. *S.lycopersicum* TC215759 ; 55. *S.officinarum* TC78398 ; 56. *S.tuberosum* TC168229 ; 57. *T.aestivum* TC301717 ; 58. *Poptr* TA1 ; 59. *Triphysaria* sp TC7138 ; 60. *V.vinifera* GSVIVT00014998001 ; 61. *Z.mays* TC460846 ; 62. *Z.mays* TC476562; 63. *Z.mays* TC511518; 64. *Z.mays* ZM07MC24807 BFb0298G14@24734; 65. *Z.mays* ZM07MC32776 BFb0337M22@32678; 66. *Z.officinale* TA5989 94328

Figure 8 represents the binary vector used for increased expression in *Oryza sativa* of a

UPA20-like-encoding nucleic acid under the control of a rice GOS2 promoter (pGOS2).

Figure 9 shows a phylogenetic tree of a number of UPA20-like polypeptides.

Examples

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

10 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
15 Protocols. Standard materials and methods for plant molecular work are described in Plant Molecular Biology Labfax (1993) by R.D.D. Croy, published by BIOS Scientific Publications Ltd (UK) and Blackwell Scientific Publications (UK).

Example 1: Identification of sequences related to the nucleic acid sequence used in the
20 methods of intervention

1. WAK-like polypeptides

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
25 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
30 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
35 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.

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

Table A1: Examples of WAK-like nucleic acids and polypeptides:

Plant Source	Nucleic acid SEQ ID NO:	Protein SEQ ID NO:
<i>Oryza sativa</i>	1	2
<i>Oryza sativa</i>	3	4
<i>Oryza sativa</i>	5	6
<i>Oryza sativa</i>	7	8
<i>Oryza sativa</i>	9	10
<i>Oryza sativa</i>	11	12
<i>Oryza sativa</i>	13	14
<i>Oryza sativa</i>	15	16
<i>Oryza sativa</i>	17	18
<i>Oryza sativa</i>	19	20
<i>Oryza sativa</i>	21	22
<i>Oryza sativa</i>	23	24
<i>Oryza sativa</i>	25	26
<i>Oryza sativa</i>	27	28
<i>Oryza sativa</i>	29	30
<i>Oryza sativa</i>	31	32
<i>Oryza sativa</i>	33	34
<i>Oryza sativa</i>	35	36
<i>Oryza sativa</i>	37	38
<i>Oryza sativa</i>	39	40
<i>Oryza sativa</i>	41	42
<i>Oryza sativa</i>	43	44
<i>Oryza sativa</i>	45	46
<i>Oryza sativa</i>	47	48
<i>Oryza sativa</i>	49	50
<i>Oryza sativa</i>	51	52
<i>Oryza sativa</i>	53	54
<i>Oryza sativa</i>	55	56
<i>Oryza sativa</i>	57	58
<i>Oryza sativa</i>	59	60
<i>Oryza sativa</i>	61	62
<i>Oryza sativa</i>	63	64
<i>Oryza sativa</i>	65	66
<i>Oryza sativa</i>	67	68
<i>Oryza sativa</i>	69	70
<i>Oryza sativa</i>	71	72
<i>Oryza sativa</i>	73	74

<i>Oryza sativa</i>	75	76
<i>Oryza sativa</i>	77	78
<i>Oryza sativa</i>	79	80
<i>Oryza sativa</i>	81	82
<i>Oryza sativa</i>	83	84
<i>Oryza sativa</i>	85	86
<i>Oryza sativa</i>	87	88
<i>Oryza sativa</i>	89	90
<i>Oryza sativa</i>	91	92
<i>Oryza sativa</i>	93	94
<i>Oryza sativa</i>	95	96
<i>Oryza sativa</i>	97	98
<i>Oryza sativa</i>	99	100
<i>Oryza sativa</i>	101	102
<i>Oryza sativa</i>	103	104
<i>Oryza sativa</i>	105	106
<i>Oryza sativa</i>	107	108
<i>Oryza sativa</i>	109	110
<i>Oryza sativa</i>	111	112
<i>Oryza sativa</i>	113	114
<i>Oryza sativa</i>	115	116
<i>Oryza sativa</i>	117	118
<i>Oryza sativa</i>	119	120
<i>Oryza sativa</i>	121	122
<i>Oryza sativa</i>	123	124
<i>Oryza sativa</i>	125	126
<i>Oryza sativa</i>	127	128
<i>Oryza sativa</i>	129	130
<i>Oryza sativa</i>	131	132
<i>Oryza sativa</i>	133	134
<i>Oryza sativa</i>	135	136
<i>Oryza sativa</i>	137	138
<i>Oryza sativa</i>	139	140
<i>Oryza sativa</i>	141	142
<i>Oryza sativa</i>	143	144
<i>Oryza sativa</i>	145	146
<i>Oryza sativa</i>	147	148
<i>Oryza sativa</i>	149	150
<i>Oryza sativa</i>	151	152
<i>Oryza sativa</i>	153	154
<i>Oryza sativa</i>	155	156

<i>Oryza sativa</i>	157	158
<i>Oryza sativa</i>	159	160
<i>Oryza sativa</i>	161	162
<i>Oryza sativa</i>	163	164
<i>Oryza sativa</i>	165	166
<i>Oryza sativa</i>	167	168
<i>Oryza sativa</i>	169	170
<i>Oryza sativa</i>	171	172
<i>Oryza sativa</i>	173	174
<i>Oryza sativa</i>	175	176
<i>Oryza sativa</i>	177	178
<i>Oryza sativa</i>	179	180
<i>Oryza sativa</i>	181	182
<i>Oryza sativa</i>	183	184
<i>Oryza sativa</i>	185	186
<i>Oryza sativa</i>	187	188
<i>Oryza sativa</i>	189	190
<i>Oryza sativa</i>	191	192
<i>Oryza sativa</i>	193	194
<i>Oryza sativa</i>	195	196
<i>Oryza sativa</i>	197	198
<i>Oryza sativa</i>	199	200
<i>Oryza sativa</i>	201	202
<i>Oryza sativa</i>	203	204
<i>Oryza sativa</i>	205	206
<i>Oryza sativa</i>	207	208
<i>Oryza sativa</i>	209	210
<i>Oryza sativa</i>	211	212
<i>Oryza sativa</i>	213	214
<i>Oryza sativa</i>	215	216
<i>Oryza sativa</i>	217	218
<i>Oryza sativa</i>	219	220
<i>Oryza sativa</i>	221	222
<i>Oryza sativa</i>	223	224
<i>Oryza sativa</i>	225	226
<i>Oryza sativa</i>	227	228
<i>Oryza sativa</i>	229	230
<i>Oryza sativa</i>	231	232
<i>Oryza sativa</i>	233	234
<i>Oryza sativa</i>	235	236
<i>Oryza sativa</i>	237	238

<i>Oryza sativa</i>	239	240
<i>Oryza sativa</i>	241	242
<i>Oryza sativa</i>	243	244
<i>Oryza sativa</i>	245	246
<i>Oryza sativa</i>	247	248
<i>Oryza sativa</i>	249	250
<i>Oryza sativa</i>	251	252
<i>Oryza sativa</i>	253	254
<i>Oryza sativa</i>	255	256
<i>Oryza sativa</i>	257	258
<i>Oryza sativa</i>	259	260
<i>Oryza sativa</i>	261	262
<i>Oryza sativa</i>	263	264
<i>Oryza sativa</i>	265	266
<i>Oryza sativa</i>	267	268
<i>Oryza sativa</i>	269	270
<i>Arabidopsis thaliana</i>	271	272
<i>Arabidopsis thaliana</i>	273	274
<i>Arabidopsis thaliana</i>	275	276
<i>Arabidopsis thaliana</i>	277	278
<i>Arabidopsis thaliana</i>	279	280
<i>Arabidopsis thaliana</i>	281	282
<i>Arabidopsis thaliana</i>	283	284
<i>Arabidopsis thaliana</i>	285	286
<i>Arabidopsis thaliana</i>	287	288
<i>Arabidopsis thaliana</i>	289	290
<i>Arabidopsis thaliana</i>	291	292
<i>Arabidopsis thaliana</i>	293	294
<i>Arabidopsis thaliana</i>	295	296
<i>Arabidopsis thaliana</i>	297	298
<i>Arabidopsis thaliana</i>	299	300
<i>Arabidopsis thaliana</i>	301	302

2. CDKB-RKA polypeptides

The nucleic acid molecule represented by SEQ ID NO: 319 and encoding a B-type CDK (CDKB) polypeptide represented by SEQ ID NO: 320 is an example of a full length CDKB polypeptide that has catalytic kinase activity.

Starting from this sequence, other sequences (full length cDNA, ESTs or genomic) related to SEQ ID NO: 319 and SEQ ID NO: 320 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: 319 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 A2 provides a list of nucleic acid sequences related to and including SEQ ID NO: 319 and SEQ ID NO: 320.

Table A2: Examples of B-type nucleic acids and polypeptides:

plant source	Nucleic acid sequence	amino acid sequence
A.thaliana_AT2G38620.2#1	319	320
A.thaliana_AT3G54180.1#1	321	322
A.thaliana_AT2G38620.1#1	323	324
A.thaliana_AT1G76540.1#1	325	326
A.thaliana_AT1G20930.1#1	327	328
G.max_Glyma14g39760.1#1	329	330
G.max_Glyma07g02400.1#1	331	332
G.max_Glyma07g07640.1#1	333	334
G.max_Glyma09g08250.1#1	335	336
G.max_Glyma17g38210.1#1	337	338
H.annuus_CD855729#1	339	340
H.vulgare_TC160457#1	341	342
H.vulgare_TC172640#1	343	344
M.truncatula_TC115190#1	345	346
M.truncatula_TC117212#1	347	348
O.sativa_LOC_Os01g67160.1#1	349	350

O.sativa_LOC_Os08g40170.1#1	351	352
P.patens_TC41028#1	353	354
P.patens_NP13120379#1	355	356
P.patens_NP13127560#1	357	358
P.patens_TC43373#1	359	360
P.patens_TC35664#1	361	362
P.patens_TC47562#1	363	364
P.trichocarpa_scaff_28.331#1	365	366
P.trichocarpa_scaff_XVI.1376#1	367	368
P.trichocarpa_scaff_II.31#1	369	370
P.trichocarpa_scaff_3712.1#1	371	372
S.lycopersicum_TC198540#1	373	374
S.lycopersicum_TC205060#1	375	376
T.aestivum_TC283074#1	377	378
T.aestivum_TC290551#1	379	380
T.aestivum_TC282781#1	381	382
Z.mays_c62195783gm030403@12303#1	383	384
M.sativa_P93321	385	386
R.communis_XP_002519974.1	387	388
N.tobacco_AAG01532.1	389	390
S.esculentum_CAC15503.1	391	392
Lotus japonicus	393	394
V.vitis_XP002266623.1	395	396
Anthirinum_CAA66235.1	397	398
C.rubrum_CAC17703.1	399	400
S.bicolor_XP_002441781.1	401	402

Polypeptides related to SEQ ID NO: 320, such as those represented in Table A2 above, may be used in the present invention, provided that the polypeptides sequences have reduced or preferably lack catalytic kinase activity. Hence, starting from sequences such as those listed in Table A2, CDKB-RKA polypeptides according to the invention, can be designed, for instance by introducing one or more mutations in said polypeptides, and preferably in the catalytic kinase domain of said polypeptides, or by partially or substantially completely truncating the catalytic kinase domain from the CDKB polypeptide.

10 3. UPA20-like polypeptides

Sequences (full length cDNA, ESTs or genomic) related to SEQ ID NO: 406 and SEQ ID

NO: 407 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: 406 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 A3 provides SEQ ID NO: 406 and 407 and a list of nucleic acid sequences related to SEQ ID NO: 406 and SEQ ID NO: 407.

Table A3: Examples of UPA20-like nucleic acids and polypeptides:

Plant Source	Nucleic acid SEQ ID NO	Amino acid SEQ ID NO	Clade
P.trichocarpa_575437#1	406	407	A
A.lyrata_495154#1	408	409	A
A.thaliana_AT5G50915.1#1	410	411	A
Aquilegia_sp_TC24680#1	412	413	A
B.napus_TC73386#1	414	415	A
C.annuum_NP13055023#1	416	417	A
C.sinensis_TC686#1	418	419	A
G.max_Glyma04g37690.1#1	420	421	A
G.max_Glyma05g38450.1#1	422	423	A
G.max_Glyma06g17420.1#1	424	425	A
G.raimondii_TC6315#1	426	427	A
M.domestica_TC28668#1	428	429	A
N.tabacum_FG645212#1	430	431	A
O.sativa_LOC_Os05g01256.1#1	432	433	A
O.sativa_LOC_Os08g42470.1#1	434	435	A

WO 2012/117330

101

PCT/IB2012/050888

O.sativa_LOC_Os09g33580.1#1	436	437	A
P.hybrida_TC3261#1	438	439	A
P.trichocarpa_773249#1	440	441	A
P.virgatum_TC28308#1	442	443	A
S.bicolor_Sb02g029530.1#1	444	445	A
S.bicolor_Sb07g025920.1#1	446	447	A
S.officinarum_TC78398#1	448	449	A
T.aestivum_TC301717#1	450	451	A
Z.mays_TC460846#1	452	453	A
Z.mays_TC476562#1	454	455	A
Z.mays_TC511518#1	456	457	A
Z.mays_ZM07MC24807_BFb0298G14@24734#1	458	459	A
Z.mays_ZM07MC32776_BFb0337M22@32678#1	460	461	A
A.lyrata_472791#1	462	463	B
A.lyrata_476021#1	464	465	B
A.lyrata_491205#1	466	467	B
A.lyrata_494918#1	468	469	B
A.thaliana_AT1G26260.1#1	470	471	B
A.thaliana_AT1G68920.1#1	472	473	B
A.thaliana_AT3G07340.1#1	474	475	B
A.thaliana_AT5G48560.1#1	476	477	B
Aquilegia_sp_TC24120#1	478	479	B
Aquilegia_sp_TC26724#1	480	481	B
C.annuum_TC15396#1	482	483	B
G.hirsutum_TC144761#1	484	485	B
G.max_Glyma01g09400.1#1	486	487	B
G.max_Glyma02g13860.1#1	488	489	B
G.max_Glyma03g21770.1#1	490	491	B
G.max_Glyma08g46040.1#1	492	493	B
G.max_Glyma18g32560.1#1	494	495	B
G.max_GM06MC18997_59691073@18643#1	496	497	B
G.raimondii_TC9358#1	498	499	B
H.tuberosus_TA2737_4233#1	500	501	B
M.truncatula_AC150842_12.5#1	502	503	B
M.truncatula_AC171266_17.4#1	504	505	B
O.sativa_LOC_Os04g28280.1#1	506	507	B
O.sativa_LOC_Os06g16400.1#1	508	509	B

WO 2012/117330

102

PCT/IB2012/050888

O.sativa_LOC_Os08g41320.1#1	510	511	B
O.sativa_LOC_Os09g32510.4#1	512	513	B
P.persica_TC11169#1	514	515	B
P.trichocarpa_553223#1	516	517	B
P.trichocarpa_572918#1	518	519	B
P.vulgaris_TC12782#1	520	521	B
Poptr_TA1	522	523	B
R.communis_TA2848_3988#1	524	525	B
S.lycopersicum_TC196509#1	526	527	B
S.lycopersicum_TC215759#1	528	529	B
S.tuberosum_TC168229#1	530	531	B
Triphysaria_sp_TC7138#1	532	533	B
V.vinifera_GSVIVT00014998001#1	534	535	B
Z.officinale_TA5989_94328#1	536	537	B

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

Example 2: Alignment of sequences to the polypeptide sequences used in the methods of the invention

1. WAK-like polypeptides

Alignment of polypeptide sequences is performed using the ClustalW 2.0 algorithm of progressive alignment (Thompson et al. (1997) Nucleic Acids Res 25:4876-4882; Chenna et al. (2003). Nucleic Acids Res 31:3497-3500) with standard setting (slow alignment, similarity matrix: Gonnet, gap opening penalty 10, gap extension penalty: 0.2). Minor manual editing is done to further optimise the alignment.

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

2. CDKB-RKA polypeptides

Alignment of CDKB polypeptide sequences as represented in Table A2 was performed using the ClustalW 1.81 algorithm of progressive alignment (Thompson et al. (1997) Nucleic Acids Res 25:4876-4882; Chenna et al. (2003). Nucleic Acids Res 31:3497-3500) with standard setting (slow alignment, similarity matrix: Gonnet, gap opening penalty 10, gap extension penalty: 0.2). Minor manual editing was done to further optimise the alignment. B-type CDK (CDKB) polypeptides are aligned in Figure 3.

3. UPA20-like polypeptides

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). 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. A representative number of UPA20-like polypeptides are aligned in Figure 6.

A phylogenetic tree of a number of UP20 like polypeptides (Figure 5) can be constructed by aligning UP20 like sequences using MAFFT (Kato 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.

It is further noted that according to the Arabidopsis bHLH tree given in Toledo-Ortiz et al. (2003) both the proteins from Clade A and from Clade B belong to the Subclade 18 shown in figure 2 of this paper.

Example 3: Calculation of global percentage identity between polypeptide sequences

Global percentages of similarity and identity between full length polypeptide sequences useful in performing the methods of the invention were determined using 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 Blosom 62 (for polypeptides), and then places the results in a distance matrix.

Results of the analysis show 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 are: Scoring matrix: Blosum62, First Gap: 12, Extending Gap: 2. A MATGAT table for local alignment of a specific domain, or data on % identity/similarity between specific domains can also be made.

5

1. UPA20-like polypeptides

Results of the analysis are shown in Figure 7 for the global similarity and identity over the full length of a representative number of UP20 like 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 UPA20-like polypeptide sequences useful in performing the methods of the invention is generally higher than 25% compared to SEQ ID NO: 407.

Results of the analysis for the global similarity and identity over the full length of a number of polypeptide sequences, which, when used in the construction of a phylogenetic tree, such as the one depicted in Figure 9, clusters with the group of polypeptides comprising the amino acid sequence represented by SEQ ID NO: 407 rather than with any other group, are shown in Table B1. In this table, the following legend is used: 1. *P.trichocarpa* 575437; 2. *A.lyrata* 495154; 3. *A.thaliana* AT5G50915.1; 4. *Aquilegia* sp TC24680; 5. *B.napus* TC73386; 6. *C.annuum* NP13055023; 7. *C.sinensis* TC686; 8. *G.max* Glyma04g37690.1; 9. *G.max* Glyma05g38450.1; 10. *G.max* Glyma06g17420.1; 11. *G.raimondii* TC6315; 12. *M.domestica* TC28668; 13. *N.tabacum* FG645212; 14. *O.sativa* LOC Os05g01256.1; 15. *O.sativa* LOC Os08g42470.1; 16. *O.sativa* LOC Os09g33580.1; 17. *P.hybrida* TC3261; 18. *P.trichocarpa* 773249; 19. *P.virgatum* TC28308; 20. *S.bicolor* Sb02g029530.1; 21. *S.bicolor* Sb07g025920.1; 22. *S.officinatum* TC78398; 23. *T.aestivum* TC301717; 24. *Z.mays* TC460846. 25. *Z.mays* TC476562; 26. *Z.mays* TC511518; 27. *z.mays* ZM07MC24807 BFb0298G14@24734; 28. *Z.mays* ZM07MC32776 BFb0337M22@32678

30 **Table B1**

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1		38	39	43	39	50	51	54	50	54	53	54	40	34	41	38	41
2	53		87	38	60	43	38	42	40	41	41	38	37	36	43	39	39
3	53	90		39	61	41	39	42	41	41	40	40	37	35	41	39	38
4	66	56	56		39	45	45	47	42	46	45	45	35	36	40	39	38
5	55	70	72	60		41	37	41	38	41	41	38	34	36	38	36	35
6	66	58	57	68	59		46	47	46	50	51	48	56	36	42	39	55
7	67	52	51	62	52	62		52	46	51	47	53	37	36	39	37	40
8	74	54	56	69	56	67	67		63	94	54	56	42	35	39	38	41
9	67	54	54	64	55	64	61	78		64	50	52	40	34	39	36	39

WO 2012/117330

105

PCT/IB2012/050888

10	74	55	56	69	55	68	67	96	78		55	57	42	36	38	38	42
11	71	57	57	66	58	70	64	73	66	73		51	42	38	40	40	43
12	72	53	55	64	54	65	67	72	65	72	67		38	36	38	35	40
13	49	55	53	50	48	61	47	52	50	52	53	50		33	39	39	83
14	50	50	50	54	53	57	49	51	51	54	52	51	43		39	35	32
15	56	60	60	57	53	57	53	56	55	56	57	50	53	55		63	39
16	54	58	56	57	49	57	51	55	53	54	55	51	55	51	78		41
17	50	55	53	51	47	59	50	51	48	50	54	50	91	44	53	56	
18	86	55	56	67	56	66	69	76	69	76	72	74	50	54	54	53	49
19	54	55	54	56	51	60	55	57	55	57	57	54	52	54	86	78	52
20	54	58	56	57	48	54	52	53	50	52	51	52	56	50	75	86	53
21	52	59	58	58	52	58	54	56	53	54	59	54	54	53	87	79	54
22	38	41	42	41	37	39	38	38	37	37	39	39	46	38	47	53	45
23	45	54	53	49	50	49	46	46	46	45	47	45	55	43	64	70	60
24	54	51	52	53	49	51	48	51	50	51	49	51	50	53	70	75	49
25	53	50	53	57	54	57	50	57	53	56	54	50	47	50	79	69	48
26	54	57	56	57	50	52	52	55	51	54	53	53	53	51	74	82	52
27	55	52	55	59	52	57	54	54	53	52	58	54	51	52	82	72	51
28	51	55	57	53	51	51	51	52	49	52	50	51	50	49	67	73	50

	18	19	20	21	22	23	24	25	26	27	28
1	77	39	38	38	31	35	38	35	37	38	37
2	39	41	40	41	32	39	36	36	39	37	37
3	42	41	39	41	33	39	36	36	37	40	37
4	46	39	38	41	30	35	37	36	37	38	35
5	39	36	37	40	30	36	35	35	37	36	38
6	47	42	39	43	31	38	36	36	37	40	35
7	53	41	37	39	30	33	36	34	34	37	34
8	59	40	38	41	31	34	37	39	37	38	36
9	53	38	36	38	30	33	36	35	35	38	33
10	59	39	38	39	30	32	36	36	37	37	36
11	53	39	37	40	31	34	38	36	35	37	34
12	57	38	34	37	29	34	34	36	35	35	34
13	38	39	39	39	29	36	37	36	37	35	36
14	34	39	36	37	30	34	37	36	37	38	35

15	39	76	60	75	41	51	55	65	57	69	50
16	37	61	67	62	44	58	59	52	62	56	54
17	41	38	39	39	31	39	38	36	38	34	38
18		40	36	39	31	37	35	35	36	37	36
19	55		59	87	43	48	53	72	56	79	50
20	54	74		60	61	58	79	51	86	56	76
21	55	92	76		40	51	54	75	55	82	50
22	40	49	64	48		53	52	35	56	41	56
23	44	63	70	65	63		52	43	55	48	54
24	51	72	84	70	55	63		49	75	51	67
25	55	84	66	85	44	56	65		48	70	43
26	54	71	92	75	60	67	83	65		53	90
27	54	87	71	87	48	61	68	83	69		48
28	51	65	84	69	61	65	77	59	93	64	

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

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

1. CDKB-RKA polypeptides

The results of the InterPro scan (InterPro database, release 31.0) of the polypeptide sequence as represented by SEQ ID NO: 320 are presented in Table C1.

Table C1: InterPro scan results (major accession numbers) of full length polypeptide sequence as represented by SEQ ID NO: 320.

Database/ Method	Accession number	Accession name	Name	Amino acid coordinates on SEQ ID NO 320
PROFILE	PS50011	IPR000719	Protein kinase catalytic domain	4-303

SMART	SM00220	IPR002290	Serine/threonine-protein kinase domain	4-303
PROSITE	PS00108	IPR008271	Serine/threonine-protein kinase active site	140-152
PROSITE	PS00107	IPR017441	Protein kinase, ATP binding site	10-33
PFAM	PF00069	IPR017442	Serine/threonine-protein kinase domain	4-303
SMART	SM00219	IPR020635	Tyrosine-protein kinase, catalytic domain	4-303

2. UPA20-like polypeptides

The results of the InterPro scan (InterPro database, release 30.0) of the polypeptide sequence as represented by SEQ ID NO: 407 are presented in Table C2.

5

Table C2: InterPro scan results (major accession numbers) of the polypeptide sequence as represented by SEQ ID NO: 407.

Database	Accession Number	Accession name	Amino acid coordinates on SEQ ID NO: 407	E -value	accession
HMMSmart	SM00353	HLH	[186-236]	6,60E+05	IPR001092
HMMPanther	PTHR12565:SF7	PTHR12565:SF7	[183-230]	1,10E+09	NULL
ProfileScan	PS50888	HLH	[174-231]	0,00E+00	IPR001092
HMMPanther	PTHR12565	PTHR12565	[183-230]	1,10E+09	NULL
HMMPfam	PF00010	HLH	[181-231]	5,60E+07	IPR001092
superfamily	SSF47459	HLH_basic	[176-255]	2,10E+00	IPR011598
Gene3D	G3DSA:4.10.280.10	HLH_DNA_bd	[176-257]	8,00E+07	IPR011598

Example 5: Topology prediction of the polypeptide sequences useful in performing the methods of invention

10

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.

15

20

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

5 A number of parameters are 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).

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

- ChloroP 1.1 hosted on the server of the Technical University of Denmark;
- 10 • 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;
- 15 • TMHMM, hosted on the server of the Technical University of Denmark
- PSORT (URL: psort.org)
- PLOC (Park and Kanehisa, Bioinformatics, 19, 1656-1663, 2003).

20 Example 6: Functional assay related to the polypeptide sequences useful in performing the methods of the invention

1. CDKB-RKA polypeptides

In accordance with the present invention, a CDKB-RKA polypeptide is provided for use in methods according to the invention, wherein said polypeptide has reduced, or preferably lacks catalytic kinase activity as compared to a (unmodified or full length) CDKB polypeptide. There are many ways to measure kinase activity, well known to a person skilled in the art.

30 In one example, catalytic kinase activity of a mutated or truncated CDKB polypeptide as defined herein can be determined by means of a kinase assay and compared to the kinase activity of a corresponding unmodified CDKB polypeptide.

A kinase assay may, for example, be performed as described in Cockcroft *et al.*, 2000 (Nature, 405 (6786), 575-579). The assay involves grinding plant material in liquid nitrogen and resuspending in an appropriate buffer (such as 1ml of 50mM Tris-HCl pH 7.5, 75mM NaCl, 15mM EGTA, 15mM MgCl₂, 1 mM dithiothreitol, 0.1% Tween 20, 1X complete Tm protease inhibitor, 1 mM NaF, 0.2mM NaV, 2mM Na-pyrophosphate, 60mM beta-glycerophosphate). The suspension is then homogenized for 4 x 30sec with 30 sec on ice between homogeneizations. The supernatant is then incubated with 20 microliters of protein A Sepharose (50% suspension) for 30 min at 4°C. The supernatant is incubated with 1 microliter of antiserum (directed against the carboxy terminal peptide of a CDC2b protein, as described in Setiady *et al.*, 1996 (Plant Cell Physiology 37 (3), 369-376) for 2h on ice, then 20 microliters protein A-Sepharose added and the sample rotated at 4°C for an hour.

Samples are washed 2 times with kinase buffer (50mM Tris-HCl pH7.5, 100mM NaCl, 5mM EGTA, 1mM DTT), resuspended in 15 microliters assay buffer (50mM Tris-HCl pH7.5, 100mM NaCl, 5mM EGTA, 10mM MgCl₂, 1mM DTT, 1mM NaF, 0.2 mM Na-Vanadate, 2 mM Na pyrophosphate, 25mM beta glycerophosphate, 0.5 mM histone H1 as a substrate, 0.5mM PMSF, and 74 kBq [gamma ³²P ATP (>185 TBq mmol⁻¹) per 15 microliters of reaction) and incubated at room temperature for 30 min. The reaction may be stopped by adding gel loading buffer and samples may be analysed using SDS-PAGE and quantified using a phosphoimager (Molecular Dynamics).

10 Example 7: Cloning of the nucleic acid sequence used in methods of the invention

1. WAK-like polypeptides

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 prm16798 (SEQ ID NO: 307; sense): 5'-ggggacaagttt gtacaaaaaagcaggcttaaacaatgaccaagcttctccac-3' and prm16761 (SEQ ID NO: 308; reverse, complementary): 5'-ggggaccactttgtacaagaaagctgggttcattatttgatgagctgtagctttc-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", pWAK-like. Plasmid pDONR201 was purchased from Invitrogen, as part of the Gateway® technology.

The entry clone comprising SEQ ID NO: 1 was then used in an LR reaction with a destination vector used for *Oryza sativa* transformation. This vector contained as functional elements within the T-DNA borders: a plant selectable marker; a screenable marker expression cassette; and a Gateway cassette intended for LR *in vivo* recombination with the nucleic acid sequence of interest already cloned in the entry clone. A rice GOS2 promoter (SEQ ID NO: 306) for constitutive expression was located upstream of this Gateway cassette.

After the LR recombination step, the resulting expression vector pGOS2::WAK-like (Figure 1) was transformed into *Agrobacterium* strain LBA4044 according to methods well known in the art.

In the same way an expression vector comprising SEQ ID NO: 3 was prepared using primers prm16656 (SEQ ID NO: 309, sense): 5'-atggcaaagatgatgcttta-3' and prm16657 (SEQ ID NO: 310; reverse, complementary): 5'-ggggaccactttgtacaagaaagctgggttcattactg tcttatcttcg caagtttt-3'.

Furthermore, in the same way also an expression vector comprising SEQ ID NO: 5 was

prepared using primers prm16710 (SEQ ID NO: 311, sense): 5'-ggggacaagtttgtaaaaaa
gcaggcttaaacaatgagcctcttcgatggattac-3' and prm16711 (SEQ ID NO: 312; reverse,
complementary): 5'-ggggaccactttgtacaagaaagctgggttcattattgcttttgatgtctttttcca-3'; an
expression vector comprising SEQ ID NO: 7 was prepared using primers prm16603 (SEQ
5 ID NO: 313, sense): 5'-ggggacaagtttgtaaaaaaagcaggcttaaacaatgggacctgtgctgctgt-3' and
prm16604 (SEQ ID NO: 314; reverse, complementary): 5'-ggggaccactttgtacaagaaagc
tgggttcatta tttctttgcatatgtctcctcc-3'; and an expression vector comprising SEQ ID NO: 9 was
prepared using primers prm16118 (SEQ ID NO: 315, sense): 5'-ggggacaagtttgtaaaaaa
agcaggcttaaacaatgtctccccaccgag-3' and prm16119 (SEQ ID NO: 316; reverse,
10 complementary): 5'-ggggaccactttgtacaagaaagctgggtcatcgatcatctcgcaa-3'.

2. CDKB-RKA polypeptides

An entry clone comprising SEQ ID NO: 317 was used in an LR reaction with a destination
vector used for *Oryza sativa* transformation. This vector contained as functional elements
15 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 High Mobility Group
Protein (HMGP) promoter (SEQ ID NO: 405) for constitutive expression was located
upstream of this Gateway cassette. After the LR recombination step, the resulting
20 expression vector promoter::CDKB-RKA (Figure 4) was transformed into *Agrobacterium*
strain LBA4044 according to methods well known in the art.

The nucleic acid sequence represented by SEQ ID NO: 317 can be readily obtained by
applying common PCR technology, using as template a custom-made *Arabidopsis thaliana*
25 seedlings cDNA library, e.g. *Arabidopsis* ecotype An-1. PCR can be performed using a
commercially available proofreading Taq DNA polymerase in standard conditions, using
200 ng of template in a 50 µl PCR mix.

The CDKB-RKA polypeptide represented by SEQ ID NO: 318, and encoded by SEQ ID NO:
30 317, is obtained by a nucleotide substitution at position 388 of SEQ ID NO: 320, where an A
was replaced by a T (and thus a Lys by a stop codon). Based on this information, a skilled
person can design a forward and a reverse primer for use in the PCR mix. The above-
indicated point mutation can be included anywhere in the reverse primer, which is designed
to have a similar melting temperature as the forward primer. More than one substitution can
35 be made, for example to create another stop codon than TAA (TAG, TGA). Both the forward and
the reverse primer include the AttB sites for Gateway recombination. The amplified PCR
fragment can be purified also using standard methods. The first step of the Gateway
procedure, the BP reaction, is then performed, during which the PCR fragment recombined
in vitro with the pDONR201 plasmid to produce, according to the Gateway terminology, an
40 "entry clone", pCDKB-RKA. Plasmid pDONR201 can be purchased from Invitrogen, as part
of the Gateway® technology.

3. UPA20-like polypeptides

3.1. SEQ ID NO 406

The nucleic acid sequence was amplified by PCR using as template a custom-made *Populus trichocarpa* 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 prm15499 (SEQ ID NO: 548; sense, start codon in bold): 5'-ggggacaagtttgtacaaaaaagcaggcttaaaca**atgg**caggctttcatatcaa-3' and prm15500 (SEQ ID NO: 549; reverse, complementary): 5'-ggggaccactttgtacaagaaagc tgggtggcagcgcttggtatatttatt-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", pUP20A. Plasmid pDONR201 was purchased from Invitrogen, as part of the Gateway® technology.

The entry clone comprising SEQ ID NO: 406 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: 545) for constitutive expression was located upstream of this Gateway cassette.

After the LR recombination step, the resulting expression vector pGOS2::UP20A like (Figure 8) was transformed into *Agrobacterium* strain LBA4044 according to methods well known in the art.

3.2. SEQ ID NO 410

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 proofreading Taq DNA polymerase in standard conditions, using 200 ng of template in a 50 µl PCR mix. The primers used were prm15497 (SEQ ID NO: 550; sense, start codon in bold): 5'-ggggacaagtttgtacaaaaaagcaggcttaaaca**atgg**caactttctcttatttc-3' and prm15498 (SEQ ID NO: 551; reverse, complementary): 5'-ggggaccactttgtacaagaaagctgggt tttgtgtcgacatgtaatttt-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", pUP20A. Plasmid pDONR201 was purchased from Invitrogen, as part of the Gateway® technology.

The entry clone comprising SEQ ID NO: 410 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: 545) for constitutive expression was located upstream of this Gateway cassette.

After the LR recombination step, the resulting expression vector pGOS2::UP20A like (similar as represented in Figure 8) was transformed into *Agrobacterium* strain LBA4044 according to methods well known in the art.

3.3. SEQ ID NO 440

The nucleic acid sequence was amplified by PCR using as template a custom-made *Populus trichocarpa* 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 prm15501 (SEQ ID NO: 552; sense, start codon in bold): 5'-ggggacaagtttgtacaaaaaagcaggcttaaaca**at**ggcagccttttcgtatc-3' and prm15502 (SEQ ID NO: 553; reverse, complementary): 5'-ggggaccactttgtacaagaaagctgggtagggcagctcttggtatctat-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", pUP20A. Plasmid pDONR201 was purchased from Invitrogen, as part of the Gateway® technology.

The entry clone comprising SEQ ID NO: 440 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: 545) for constitutive expression was located upstream of this Gateway cassette.

After the LR recombination step, the resulting expression vector pGOS2::UP20A like (similar as represented in Figure 8) was transformed into *Agrobacterium* strain LBA4044 according to methods well known in the art.

3.4. SEQ ID NO 508

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 prm14919 (SEQ ID NO: 554; sense, start codon in

bold): 5'-ggggacaagtttgtacaaaaaagcaggcttaaacaatgaatgagaaggacgcca-3' and prm14920 (SEQ ID NO: 555; reverse, complementary): 5'-ggggaccactttgtacaagaaagctgggtcttgcttcagttgtggaatca-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", pUP20A. Plasmid pDONR201 was purchased from Invitrogen, as part of the Gateway® technology.

10 The entry clone comprising SEQ ID NO: 508 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: 545) for constitutive expression was located upstream of this Gateway cassette.

After the LR recombination step, the resulting expression vector pGOS2::UP20A like (similar as represented in Figure 8) was transformed into *Agrobacterium* strain LBA4044 according to methods well known in the art.

Example 8: Plant transformation

Rice transformation

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

Agrobacterium strain LBA4404 containing the expression vector was used for co-cultivation. *Agrobacterium* was inoculated on AB medium with the appropriate antibiotics and cultured for 3 days at 28°C. The bacteria were then collected and suspended in liquid co-cultivation medium to a density (OD₆₀₀) of about 1. The calli were immersed in the suspension for 1 to 15 minutes. The callus tissues were then blotted dry on a filter paper and transferred to solidified, co-cultivation medium and incubated for 3 days in the dark at 25°C. After washing away the *Agrobacterium*, the calli were grown on 2,4-D-containing medium for 10 to 14 days (growth time for indica: 3 weeks) under light at 28°C - 32°C in the presence of a selection agent. During this period, rapidly growing resistant callus developed. After transfer of this material to regeneration media, the embryogenic potential was released and shoots

developed in the next four to six weeks. Shoots were excised from the calli and incubated for 2 to 3 weeks on an auxin-containing medium from which they were transferred to soil. Hardened shoots were grown under high humidity and short days in a greenhouse.

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

35 to 90 independent T0 rice transformants were generated for one construct. The primary transformants were transferred from a tissue culture chamber to a greenhouse. After a
10 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).

15

Example 9: 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-
20 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
25 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
30 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
40 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
5 obtain regenerating plants have been described. For example, these can be selected from the cultivar Rangeland (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
10 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 acetosyringone. The explants are washed in half-strength Murashige-Skoog medium (Murashige and Skoog, 1962) and
15 plated on the same SH induction medium without acetosyringone 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
20 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
25 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
30 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-
35 500 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
40 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

5 Seeds of sugarbeet (*Beta vulgaris* L.) are sterilized in 70% ethanol for one minute followed by 20 min. shaking in 20% Hypochlorite bleach e.g. Clorox® regular bleach (commercially available from Clorox, 1221 Broadway, Oakland, CA 94612, USA). Seeds are rinsed with sterile water and air dried followed by plating onto germinating medium (Murashige and Skoog (MS) based medium (Murashige, T., and Skoog, ., 1962. *Physiol. Plant*, vol. 15, 473-
10 497) including B5 vitamins (Gamborg et al.; *Exp. Cell Res.*, vol. 50, 151-8.) supplemented with 10 g/l sucrose and 0,8% agar). Hypocotyl tissue is used essentially for the initiation of shoot cultures according to Hussey and Hepher (Hussey, G., and Hepher, A., 1978. *Annals of Botany*, 42, 477-9) and are maintained on MS based medium supplemented with 30g/l sucrose plus 0,25mg/l benzylamino purine and 0,75% agar, pH 5,8 at 23-25°C with a 16-
15 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)
20 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
25 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.
30 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
35 Gallois, P., 1990. *Journal of Experimental Botany*; vol. 41, No. 226; 529-36) or the methods published in the international application published as WO9623891A.

Sugarcane transformation

Spindles are isolated from 6-month-old field grown sugarcane plants (Arencibia et al., 1998. *Transgenic Research*, vol. 7, 213-22; Enriquez-Obregon et al., 1998. *Planta*, vol. 206, 20-
40 27). Material is sterilized by immersion in a 20% Hypochlorite bleach e.g. Clorox® regular bleach (commercially available from Clorox, 1221 Broadway, Oakland, CA 94612, USA) for

20 minutes. Transverse sections around 0,5cm are placed on the medium in the top-up direction. Plant material is cultivated for 4 weeks on MS (Murashige, T., and Skoog, ., 1962. *Physiol. Plant*, vol. 15, 473-497) based medium incl. B5 vitamins (Gamborg, O., et al., 1968. *Exp. Cell Res.*, vol. 50, 151-8) supplemented with 20g/l sucrose, 500 mg/l casein hydrolysate, 0,8% agar and 5mg/l 2,4-D at 23°C in the dark. Cultures are transferred after 4 weeks onto identical fresh medium. *Agrobacterium tumefaciens* strain carrying a binary plasmid harbouring a selectable marker gene, for example hpt, is used in transformation experiments. One day before transformation, a liquid LB culture including antibiotics is grown on a shaker (28°C, 150rpm) until an optical density (O.D.) at 600 nm of ~0,6 is reached. Overnight-grown bacterial cultures are centrifuged and resuspended in MS based inoculation medium (O.D. ~0,4) including acetosyringone, pH 5,5. Sugarcane embryogenic callus pieces (2-4 mm) are isolated based on morphological characteristics as compact structure and yellow colour and dried for 20 min. in the flow hood followed by immersion in a liquid bacterial inoculation medium for 10-20 minutes. Excess liquid is removed by filter paper blotting. Co-cultivation occurred for 3-5 days in the dark on filter paper which is placed on top of MS based medium incl. B5 vitamins containing 1 mg/l 2,4-D. After co-cultivation calli are washed with sterile water followed by a non-selective cultivation period on similar medium containing 500 mg/l cefotaxime for eliminating remaining *Agrobacterium* cells. After 3-10 days explants are transferred to MS based selective medium incl. B5 vitamins containing 1 mg/l 2,4-D for another 3 weeks harbouring 25 mg/l of hygromycin (genotype dependent). All treatments are made at 23°C under dark conditions. Resistant calli are further cultivated on medium lacking 2,4-D including 1 mg/l BA and 25 mg/l hygromycin under 16 h light photoperiod resulting in the development of shoot structures. Shoots are isolated and cultivated on selective rooting medium (MS based including, 20g/l sucrose, 20 mg/l hygromycin and 500 mg/l cefotaxime). Tissue samples from regenerated shoots are used for DNA analysis. Other transformation methods for sugarcane are known in the art, for example from the in-ternational application published as WO2010/151634A and the granted European patent EP1831378.

30 Example 10: Phenotypic evaluation procedure

10.1 Evaluation setup

35 to 90 independent T0 rice transformants were generated. The primary transformants were transferred from a tissue culture chamber to a greenhouse for growing and harvest of T1 seed. Six events, of which the T1 progeny segregated 3:1 for presence/absence of the transgene, were retained. For each of these events, approximately 10 T1 seedlings containing the transgene (hetero- and homo-zygotes) and approximately 10 T1 seedlings lacking the transgene (nullizygotes) were selected by monitoring visual marker expression. The transgenic plants and the corresponding nullizygotes were grown side-by-side at random positions. Greenhouse conditions were of 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 were grown in potting soil under normal conditions until they approached the heading stage. They were then transferred to a "dry" section where irrigation was withheld. Soil moisture probes were inserted in randomly chosen pots to monitor the soil water content (SWC). When SWC went below certain thresholds, the plants were automatically re-watered continuously until a normal level was reached again. The plants were then re-transferred again to normal conditions. The rest of the cultivation (plant maturation, seed harvest) was the same as for plants not grown under abiotic stress conditions. Growth and yield parameters were recorded as detailed for growth under normal conditions.

Nitrogen use efficiency screen

T1 or T2 plants were grown in potting soil under normal conditions except for the nutrient solution. The pots were 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) was the same as for plants not grown under abiotic stress. Growth and yield parameters were 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.

5

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 WO 2010/031780. These measurements were used to determine different parameters.

10

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.

15

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.

20

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.

25

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.

30

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

35

40

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

5 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

10 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.

15 The Harvest Index (HI) in the present invention is defined as the ratio between the total seed weight and the above ground area (mm²), multiplied by a factor 10⁶.

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

20 (expressed as a %) of the number of filled seeds (i.e. florets containing seeds) over the total number of seeds (i.e. total number of florets). In other words, the seed filling rate is the percentage of florets that are filled with seed.

Example 10: Results of the phenotypic evaluation of the transgenic plants25 1. WAK-like polypeptides

The results of the evaluation of transgenic rice plants expressing a WAK-like polypeptide under non-stress conditions are presented below. See previous Examples for details on the generations of the transgenic plants.

30 The results of the evaluation of transgenic rice plants in the T1 generation and expressing a nucleic acid comprising SEQ ID NO: 1 under non-stress conditions are presented below.

An increase of at least 5 % was observed for seed fill rate, and positive effects on thousand kernel weight (TKW).

35 **Table D1:** Data summary for transgenic rice plants expressing a nucleic acid comprising SEQ ID NO: 1 under non-stress conditions; for each parameter, the overall percent increase is shown, for each parameter the p-value is <0.05.

Parameter	Overall increase
fillrate	5.6

The results of the evaluation of transgenic rice plants in the T1 generation and expressing a nucleic acid comprising SEQ ID NO: 3 in the nitrogen use efficiency screen as described above, are presented below.

- 5 An increase in root biomass was observed; as well as positive effects on seed fill rate, the amount of flowers per panicle and the greenness.

The results of the evaluation of transgenic rice plants in the T1 generation and expressing a nucleic acid comprising SEQ ID NO: 5 under non-stress conditions are presented below.

- 10 An increase of at least 5 % was observed for seed fill rate, and positive effects on seed yield by positive effects on total weight of seeds and number of filled seeds, as well as positive effect on harvest index.

- 15 **Table D2:** Data summary for transgenic rice plants expressing a nucleic acid comprising SEQ ID NO: 5 under non-stress conditions; for each parameter, the overall percent increase is shown, for each parameter the p-value is <0.05.

Parameter	Overall increase
fillrate	6.6

- 20 Evaluation of transgenic rice plants in the T1 generation and expressing a nucleic acid comprising SEQ ID NO: 7 under the drought screen as described above, showed an improved seed yield because of better fill rate.

- 25 Evaluation of transgenic rice plants in the T1 generation and expressing a nucleic acid comprising SEQ ID NO: 9 under non stress conditions, showed an improved seed yield because of better fill rate and larger amount of flowers per panicle.

2. CDKB-RKA polypeptides

- 30 The results of the evaluation of transgenic rice plants in the T1 generation and expressing a nucleic acid encoding the CDKB-RKA polypeptide of SEQ ID NO: 318 under non-stress conditions are presented below in Table D3.

- 35 When grown under non-stress conditions, an increase of at least 5 % was observed for aboveground biomass (AreaMax), root biomass (RootMax and RootThinMax), and for seed yield including total seed weight (totalwgseeds), number of seeds (nrfilledseed). In addition, plants expressing a nucleic acid encoding a CDK polypeptide of SEQ ID NO: 318 showed early vigor (EmerVigor).

- 40 **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	10.5
EmerVigor	24.8
RootMax	7.2
totalwgseeds	14.1
nrfilledseed	13.2
RootThinMax	5.0

3. UPA20-like polypeptides

3.1. SEQ ID NO: 406 under non-stress conditions

- 5 The results of the evaluation of transgenic rice plants of the T1 generation expressing the nucleic acid encoding of the UPA20 like polypeptide of SEQ ID NO: 407 under non-stress conditions are presented below in Table D4.

10 When grown under non-stress conditions, an increase of at least 3% was observed for the parameter thousand kernel weight (TKW) and of at least 5 % was observed for parameters such as: biomass, including above ground biomass (area max) and root biomass (root max and root thick max); and seed yield, including total weight of seeds, harvest index, number of filled seeds, and number of flowers per panicle. In addition, it was noted that plants of at least one event also showed an increase of at least 5% in fill rate as compared to control plants.

15 **Table D4:** Data summary for transgenic rice plants; for each parameter, the overall percent increase is shown for the plants of the T1 generation as compared to control plants, for each parameter the p-value is < 0.05.

20

Parameter	Overall increase
AreaMax	9.1
RootMax	7.9
totalwgseeds	19.2
harvestindex	8.8
Nrfilledseed	13.7
RootThickMax	8.8

3.2. SEQ ID NO: 410 under non stress-conditions

- 25 The results of the evaluation of transgenic rice plants of the T1 generation expressing the nucleic acid encoding of the UPA20 like polypeptide of SEQ ID NO: 411 under non-stress conditions indicated that when grown under non-stress conditions, transgenic plants showed an increase of at least 5%, and in particular of 9.4% in number of flowers per panicle as compared to control plants.

3.3. SEQ ID NO: 440 under non-stress conditions

The results of the evaluation of transgenic rice plants of the T1 generation expressing the nucleic acid encoding of the UPA20 like polypeptide of SEQ ID NO: 441 under non-stress conditions are presented below in Table D5.

5

When grown under non-stress conditions, an increase of at least 5% was observed for parameters such as: time to flower, and seed yield, including TKW and fill rate. In addition, it was noted that plants of at least two event also showed a significant increase of at least 5% in parameters such as harvest index and number of filled seeds as compared to control plants.

10

Table D5: Data summary for transgenic rice plants; for each parameter, the overall percent increase is shown for the plants of the T1 generation as compared to control plants, for each parameter the p-value is <0.05.

Parameter	Overall increase
TimetoFlower	5.6
fillrate	17.7
TKW	5.8

15

3.3. SEQ ID NO: 508 under-stress conditions

Transgenic rice plants of the T1 generation expressing the nucleic acid encoding of the UPA20 like polypeptide of SEQ ID NO: 509 under stress conditions, more particularly under conditions of nitrogen deficiency, showed the following results as compared to control plants:

20

1. plants of at least two events showed an increase of at least 5% in biomass, more particular above ground biomass;
2. plants of at least one event showed an increase of at least 5% in parameters such as above ground biomass, total seed weight, fill rate, harvest index and number of filled seeds as compared to control plants, and an increase of at least 3% in TKW as compared to control plants;
3. plants of at least two events showed an increase of at least 5% in TKW as compared to control plants.

25

30

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 WAK-like polypeptide, wherein said WAK-like polypeptide comprises an EGF domain and a transmembrane domain.
2. Method according to claim 1, wherein said modulated expression is effected by introducing and expressing in a plant said nucleic acid encoding said WAK-like polypeptide.
3. Method according to claim 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 claims 1 to 3, wherein said enhanced yield-related traits are obtained under non-stress conditions.
5. Method according to any one of claims 1 to 3, wherein said enhanced yield-related traits are obtained under conditions of drought stress or nitrogen deficiency.
6. Method according to any one of claims 1 to 5, wherein said nucleic acid encoding a WAK-like polypeptide 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*.
7. Method according to any one of claims 1 to 6, wherein said nucleic acid encoding a WAK-like 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, provided these nucleic acids encode polypeptides comprising an EGF domain and a transmembrane domain.
8. Method according to any one of claims 1 to 7, wherein said nucleic acid sequence encodes an orthologue or paralogue of any of the polypeptides given in Table A1, provided these nucleic acids encode polypeptides comprising an EGF domain and a transmembrane domain.
9. Method according to any one of claims 1 to 8, wherein said nucleic acid encodes the polypeptide represented by SEQ ID NO: 2 or 4.
10. Method according to any one of claims 1 to 9, 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.
11. Plant, plant part thereof, including seeds, or plant cell, obtainable by a method according to any one of claims 1 to 10, wherein said plant, plant part or plant cell comprises a recombinant nucleic acid encoding a WAK-like polypeptide as defined in any of claims 1 and 4 to 9.
12. Construct comprising:
 - (i) nucleic acid encoding a WAK-like as defined in any of claims 1 and 4 to 9;
 - (ii) one or more control sequences capable of driving expression of the nucleic acid sequence of (i); and optionally

(iii) a transcription termination sequence.

13. Construct according to claim 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.
14. Use of a construct according to claim 12 or 13 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.
15. Plant, plant part or plant cell transformed with a construct according to claim 12 or 13.
16. 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 WAK-like polypeptide as defined in any of claims 1 and 4 to 9; and
 - (ii) cultivating said plant cell or plant under conditions promoting plant growth and development.
17. 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 WAK-like polypeptide as defined in any of claims 1 and 4 to 9 or a transgenic plant cell derived from said transgenic plant.
18. Transgenic plant according to claim 11, 15 or 17, 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.
19. Harvestable parts of a plant according to claim 18, wherein said harvestable parts are preferably shoot biomass and/or seeds.
20. Products derived from a plant according to claim 18 and/or from harvestable parts of a plant according to claim 19.
21. Use of a nucleic acid encoding a WAK-like polypeptide as defined in any of claims 1 and 4 to 9 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.
22. A method for the production of a product comprising the steps of growing the plants according to claim 11, 14, 15, 17 or 18 and producing said product from or by
 - (i) said plants; or
 - (ii) parts, including seeds, of said plants.
23. Construct according to claim 12 or 13 comprised in a plant cell.
24. Recombinant chromosomal DNA comprising the construct according to claim 12 or 13.

25. A method for increasing yield in a plant relative to a control plant, comprising introducing and expressing in said plant a nucleic acid molecule encoding a CDKB-RKA polypeptide, wherein said CDKB-RKA polypeptide is a modified CDKB polypeptide, and wherein said CDKB-RKA polypeptide has reduced catalytic kinase activity, as compared to said CDKB polypeptide.
26. Method according to claim 25, wherein said CDKB-RKA polypeptide lacks catalytic kinase activity.
27. Method according to claim 25 or 26, wherein said CDKB-RKA polypeptide is a mutated CDKB polypeptide.
28. Method according to claim 25, 26 or 27, wherein said CDKB-RKA polypeptide is a CDKB polypeptide comprising a mutated kinase domain.
29. Method according to any of claims 25 to 28, wherein said CDKB-RKA polypeptide is a truncated CDKB polypeptide.
30. Method according to any of claims 25 to 29, wherein said CDKB-RKA polypeptide is a truncated CDKB polypeptide comprising a deletion in the C-terminal half of said polypeptide that reduces, and preferably substantially inactivates, that kinase activity of said CDKB.
31. Method according to any of claims 25 to 30 wherein said CDKB-RKA polypeptide is a truncated CDKB polypeptide comprising a deletion of the kinase domain of said CDKB polypeptide.
32. Method according to any of claims 25 to 30 wherein said CDKB-RKA polypeptide is a truncated CDKB polypeptide comprising a deletion of a part of the kinase domain of said CDKB polypeptide, wherein said part preferably comprises at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 or 12 amino acids.
33. Method according to any of claims 25 to 32, wherein said kinase domain of said CDKB polypeptide is represented by a conserved domain from amino acid 140 to 152 in SEQ ID NO: 320.
34. Method according to any of claims 25 to 33 wherein said CDKB polypeptide has cyclin-binding activity.
35. Method according to any of claims 25 to 34, wherein said CDKB-RKA polypeptide comprises a cyclin binding domain, preferably represented by a PPTALRE motif with no mismatches or with a mismatch at position 2 and/or 4 from left to right, and more preferably represented by SEQ ID NO: 403, and more preferably represented by SEQ ID NO: 404.
36. Method according to any of claims 25 to 35, wherein said CDKB-RKA polypeptide comprises a protein kinase inhibitory domain.
37. Method according to any of claims 25 to 36, wherein said CDKB-RKA polypeptide lacks a T-loop activation kinase domain.
38. Method according to any one of claims 25 to 37, wherein said nucleic acid encoding an CDKB-RKA is selected from any one of
 - a nucleic acid derived from a nucleic acid encoding any one of the polypeptides listed in Table A2;

- a nucleic acid derived from an orthologue or paralogue of any of the polypeptides given in Table A2;
 - a portion of a nucleic acid encoding any one of the polypeptides listed in Table A2 or of an orthologue or paralogue of any of the polypeptides given in Table A2;
 - a nucleic acid capable of hybridising with a nucleic acid encoding any one of the polypeptides listed in Table A2 or of an orthologue or paralogue of any of the polypeptides given in Table A2.
39. Method according to any one of claims 25 to 38, wherein said nucleic acid encodes a polypeptide represented by SEQ ID NO: 318 or a homologue, orthologue or paralogue thereof, or a portion of said nucleic acid, or a nucleic acid capable of hybridising with said nucleic acid.
40. Method according to any one of claims 25 to 39 wherein said nucleic acid encoding said CDKB 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*.
41. Method according to any one of claims 25 to 40, wherein said nucleic acid encodes a polypeptide having has at least about 30-40%, preferably at least about 41-50%, more preferably at least about 51-60%, 61-70%, 71-80%, 81-85%, 86-90%, 91-95%, 96-99% or more than 99% overall sequence identity to SEQ ID NO: 318, and preferably corresponding to SEQ ID NO: 318.
42. Method according to any of claims 25 to 41, wherein said increased yield relative to control plants comprises (i) increased seed yield and/or (ii) increased vegetative biomass, such as increased leaf biomass and/or increased root biomass.
43. Method according to any one of claims 25 to 42, wherein said enhanced yield-related traits are obtained under non-stress conditions.
44. Method according to any one of claims 25 to 43 wherein said nucleic acid is operably linked to a constitutive promoter, preferably to a constitutive plant promoter, more preferably to a medium strength constitutive plant promoter.
45. Transgenic plant, transgenic plant part thereof, including seeds, or transgenic plant cell, obtainable by a method according to any one of claims 25 to 44, wherein said plant, plant part or plant cell comprises a recombinant nucleic acid encoding a CDKB-RKA polypeptide as defined in any of claims 25 and 41.
46. Construct comprising:
- (i) nucleic acid encoding a CDKB-RKA polypeptide as defined in any of claims 25 to 41;
 - (ii) one or more control sequences capable of driving expression of the nucleic acid sequence of (i); and optionally
 - (iii) a transcription termination sequence.
47. Construct according to claim 46, wherein one of said control sequences is a constitutive promoter, preferably to a constitutive plant promoter, more preferably to a medium strength constitutive plant promoter.

48. Use of a construct according to claim 46 or 47 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.
- 5 49. Transgenic plant, transgenic plant part or transgenic plant cell transformed with a construct according to claim 46 or 47.
50. 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:
- 10 (a) introducing and expressing in a plant cell or plant a nucleic acid encoding a CDKB-RKA polypeptide as defined in any of claims 25 to 41; and
- (b) cultivating said plant cell or plant under conditions promoting plant growth and development.
- 15 51. 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 the introduction and expression therein of a nucleic acid encoding a CDKB-RKA polypeptide as defined in any of claims 25 to 41, or a transgenic plant cell derived from said transgenic plant.
- 20 52. Transgenic plant according to claim 45, 49 or 51, 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.
53. Harvestable parts of a plant according to any of claims 45, 49, 51 or 52 wherein said
- 25 harvestable parts preferably are seeds.
54. Products derived from or produced from a plant according to claim 53 and/or from harvestable parts of a plant according to any of claims 45, 49, 51 or 52.
55. Use of a nucleic acid encoding a CDKB-RKA polypeptide as defined in any of claims 25 to 41 for enhancing yield-related traits in plants relative to control plants, preferably
- 30 for increasing yield, and more preferably for increasing seed yield and/or for increasing biomass in plants relative to control plants.
56. A method for the production of a product comprising the steps of growing the plants according to any of claims 45, 49, 51 or 52 and producing said product from or by
- (a) said plants; or
- 35 (b) parts, including seeds, of said plants.
57. Construct according to claim 46 or 47 comprised in a plant cell.
58. Recombinant chromosomal DNA comprising the construct according to claim 46 or 47.
59. Use of a nucleic acid encoding a CDKB-RKA polypeptide as defined in any of claims
- 40 25 to 41 as molecular marker for selecting for enhanced yield-related traits in plants relative to control plants, preferably for selecting for increased yield, and more preferably selecting for increasing seed yield and/or selecting for increasing biomass

in plants relative to control plants.

60. A method for enhancing yield-related traits in plants relative to control plants, comprising modulating expression in a plant of a nucleic acid encoding an UPA20-like polypeptide, wherein said UPA20-like polypeptide comprises a bHLH domain, and preferably comprises one or more domains selected from the group SM00353, PTHR12565:SF7, PTHR12565, PF00010, PS50888, SSF47459; and G3DSA:4.10.280.10.
61. Method according to claim 60, wherein said modulated expression is effected by introducing and expressing in a plant said nucleic acid encoding said UPA20-like polypeptide.
62. Method according to claim 60 or 61, wherein said enhanced yield-related traits comprise increased yield relative to control plants, and preferably comprise increased seed yield and/increase biomass relative to control plants.
63. Method according to any one of claims 60 to 62, wherein said increased seed yield comprises any one or more of increased total seed weight, increased number of seeds, increased number of flowers per panicle, increased thousand kernel weight, and increased fill rate.
64. Method according to any one of claims 60 to 63, wherein said enhanced yield-related traits are obtained under non-stress conditions.
65. Method according to any one of claims 60 to 63, wherein said enhanced yield-related traits are obtained under conditions of drought stress, salt stress or nitrogen deficiency.
66. Method according to any one of claims 60 to 65, wherein said bHLH domain comprises an amino acid sequence having at least 50% overall sequence identity to the amino acid represented by SEQ ID NO: 544.
67. Method according to any of claims 60 to 66, wherein said UPA20-like polypeptide comprises one or more of the following motifs:
- (i) Motif 2: YIHVRARRGQATDSHSLAERVRR[KR]ISERM[KR][LIF]LQ[DL]LVPGC [ND]K[IV]TGKA[LV]ML (SEQ ID NO: 538),
 - (ii) Motif 3: DEIINYVQSLQ[RN]QVEFLSMKLA[ST]V[NS]P[RLV]L[DY] (SEQ ID NO: 539),
 - (iii) Motif 4: [MN][AS][RK]KRK[SA][RA]x[KG][FGS] (SEQ ID NO: 540)
68. Method according to any of claims 60 to 67, wherein said UPA20-like polypeptide further comprises one or more of the following motifs:
- (i) Motif 5: RVRRE[KR]ISERM[RK][MIV]LQ[ARLS]LVPGCDK[IV]TGKAL[MI]LDEIIN YVQSLQNQVEFLSM (SEQ ID NO: 541),
 - (ii) Motif 6: [KRS]KV[HE][EGD]E[PAE]P[KAT]GYIHVRARRGQATDSHSLAE (SEQ ID NO: 542),
 - (iii) Motif 7: [KR][LI]AS[LMV][SN]P[VLM][LF]Y[DG]FGMD[LSR] (SEQ ID NO: 543)
69. Method according to any of claims 60 to 68, wherein said UPA20-like polypeptide comprises a conserved domain (or motif) with at least 50% sequence identity to a conserved domain selected from any one of:

- (i) amino acid 186 to 236 in SEQ ID NO:2 (motif 8 - SEQ ID NO: 556);
 - (ii) amino acid 174 to 231 in SEQ ID NO:2 (motif 9 - SEQ ID NO: 557);
 - (iii) amino acid 183 to 230 in SEQ ID NO:2 (motif 10 - SEQ ID NO: 558);
 - (iv) amino acid 181 to 231 in SEQ ID NO:2 (motif 11 - SEQ ID NO: 559);
 - 5 (v) amino acid 176 to 255 in SEQ ID NO:2 (motif 12 - SEQ ID NO: 560) ; and
 - (vi) amino acid 176 to 257 in SEQ ID NO:2 (motif 13 - SEQ ID NO: 561).
70. Method according to any one of claims 60 to 69, wherein said nucleic acid encoding a UPA20-like polypeptide is of plant origin, preferably from a dicotyledonous plant, further preferably from the family *Salicaceae*, more preferably from the genus
- 10 *Populus*, most preferably the nucleic acid is from *Populus trichocarpa*.
71. Method according to any one of claims 60 to 70, wherein said nucleic acid encoding a UPA20-like polypeptide 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 such a nucleic acid.
- 15 72. Method according to any one of claims 60 to 71, wherein said nucleic acid sequence encodes an orthologue or paralogue of any of the polypeptides given in Table A3.
73. Method according to any one of claims 60 to 72, wherein said nucleic acid encodes the polypeptide represented by SEQ ID NO: 407 or a homologue thereof.
- 20 74. Method according to any one of claims 60 to 73, 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.
- 25 75. Plant, plant part thereof, including seeds, or plant cell, obtainable by a method according to any one of claims 60 to 74, wherein said plant, plant part or plant cell comprises a recombinant nucleic acid encoding a UPA20-like polypeptide as defined in any of claims 60 and 66 to 73.
- 30 76. Plant according to claim 75, or a 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, secale, einkorn, teff, milo or oats.
- 35 77. Construct comprising:
- (i) nucleic acid encoding a UPA20-like polypeptide as defined in any of claims 60 and 66 to 73;
 - (ii) one or more control sequences capable of driving expression of the nucleic acid sequence of (i); and optionally
 - (iii) a transcription termination sequence.
- 40 78. Construct according to claim 77, 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.
79. Use of a construct according to claim 77 or 78 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.

80. Plant, plant part or plant cell transformed with a construct according to claim 77 or 78.

81. Plant according to claim 80, or a 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, secale, einkorn, teff, milo or oats.

82. 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 UPA20-like polypeptide as defined in any of claims 60 and 66 to 73; and
- (ii) cultivating said plant cell or plant under conditions promoting plant growth and development.

83. Transgenic plant having enhanced yield-related traits relative to control plants, preferably increased yield relative to control plants, and more preferably increased seed yield, resulting from modulated expression of a nucleic acid encoding a UPA20-like polypeptide as defined in any of claims 60 and 66 to 73 or a transgenic plant cell derived from said transgenic plant.

84. Transgenic plant according to claim 83, 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, secale, einkorn, teff, milo or oats.

85. Harvestable parts of a plant according to any of claims 75-76, 80-81, and 83-84, wherein said harvestable parts are preferably shoot biomass and/or seeds.

86. Products derived from a plant according to any of claims 75-76, 80-81, and 83-84, and/or from harvestable parts of a plant according to claim 85.

87. Isolated nucleic acid molecule selected from:

- (i) a nucleic acid represented by any one of SEQ ID NO: 458, 460 and 496;
- (ii) the complement of a nucleic acid represented by any one of SEQ ID NO: 458, 460 and 496;
- (iii) a nucleic acid encoding a UPA20-like 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: 459, 461 and 497, and additionally or alternatively comprising one or more motifs having in increasing order of preference at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more sequence identity to any one or more of the motifs given in SEQ ID NO: 538,

539, 540, 541, 542, 543, 556, 557, 558, 559, 560, and 561 (motifs 2 to 13), 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.

88. Isolated polypeptide selected from:

- (i) an amino acid sequence represented by any one of SEQ ID NO: 459, 461 and 497;

- (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: 459, 461 and 497, and additionally or alternatively comprising one or more motifs having in increasing order of preference at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more sequence identity to any one or more of the motifs given in SEQ ID NO: 538, 539, 540, 541, 542, 543, 556, 557, 558, 559, 560, and 561 (motifs 2 to 13), 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.

89. Use of a nucleic acid encoding a UPA20-like polypeptide as defined in any of claims 60 and 66 to 73 and 87 for enhancing yield-related traits in plants relative to control plants, preferably for increasing yield, and more preferably for increasing seed yield in plants relative to control plants.

90. Use of a nucleic acid as defined in claim 87 and encoding a UPA20-like polypeptide for enhancing yield-related traits in plants relative to control plants, preferably for increasing yield, and more preferably for increasing seed yield in plants relative to control plants.

91. Use of a nucleic acid encoding a UPA20-like polypeptide as defined in any of claims 60 and 66 to 83 and 88 as molecular marker.

92. Use of a nucleic acid as defined in claim 87 and encoding a UPA20-like polypeptide as defined in any of claims 60 and 66 to 83 and 88 as molecular marker.

1/42

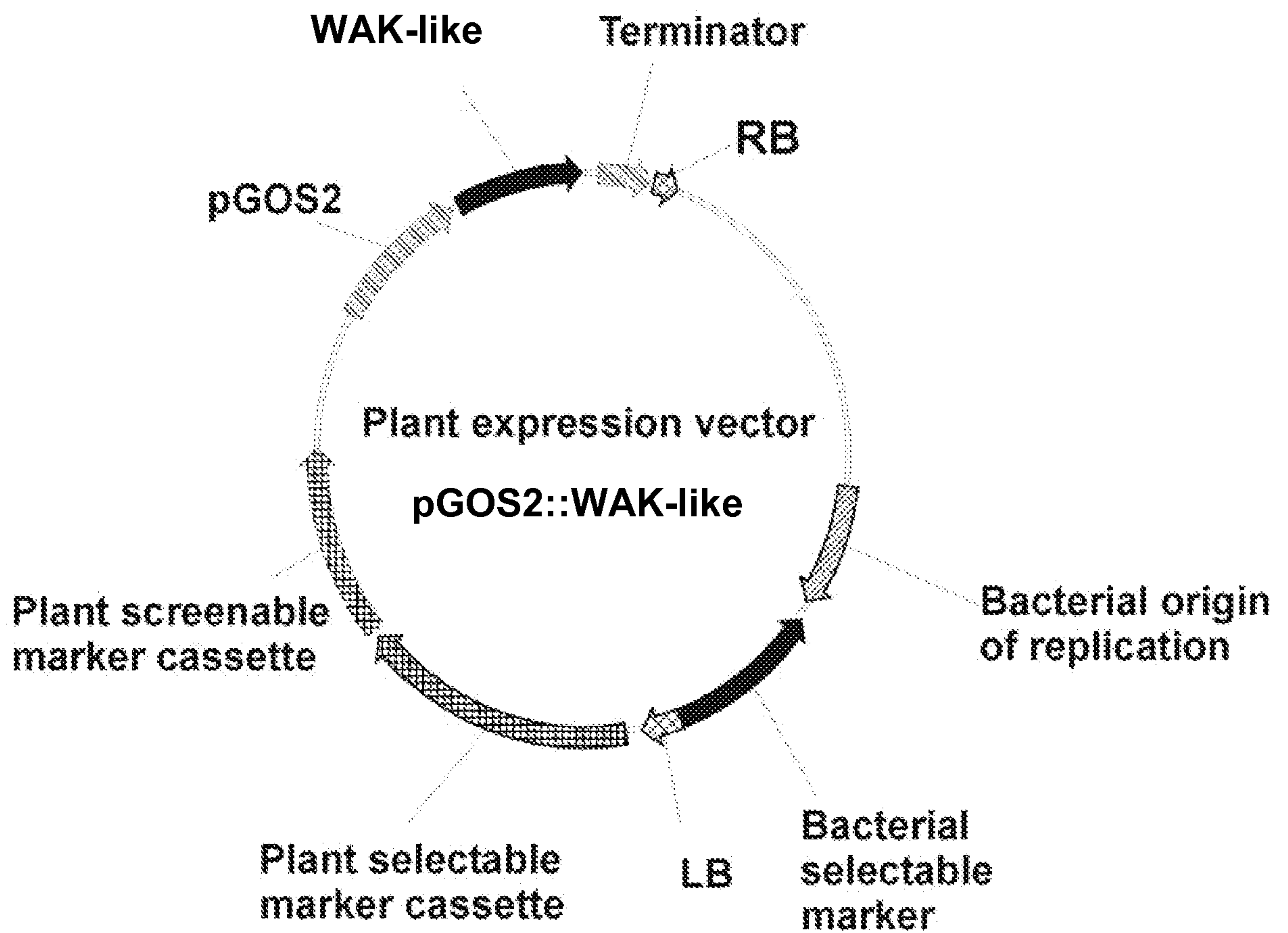


FIGURE 1

2/42

		1		50
SEQ ID NO: 320	(1)	MEKYEKLEKV	GEGTYGKVKAMEKTTGKLVALKKTRLEMDEEGIPPTALR	
SEQ ID NO: 318	(1)	MEKYEKLEKV	GEGTYGKVKAMEKTTGKLVALKKTRLEMDEEGIPPTALR	
Consensus	(1)	MEKYEKLEKV	GEGTYGKVKAMEKTTGKLVALKKTRLEMDEEGIPPTALR	
			* *TY*	
		51		100
SEQ ID NO: 320	(51)	EISLLQMLSQSIYIVRLLC	VEHV IQSKDSTVSHSPKSNLYLVFEYLDTDL	
SEQ ID NO: 318	(51)	EISLLQMLSQSIYIVRLLC	VEHV IQSKDSTVSHSPKSNLYLVFEYLDTDL	
Consensus	(51)	EISLLQMLSQSIYIVRLLC	VEHV IQSKDSTVSHSPKSNLYLVFEYLDTDL	
		E		
		101		150
SEQ ID NO: 320	(101)	KKFIDSHRKGSNPRPLEASLVQR	FMFQLFKGV AHCHSHGV LHRDLKPQNL	
SEQ ID NO: 318	(101)	KKFIDSHRKGSNPRPLEASLVQR	FMFQLF-----	
Consensus	(101)	KKFIDSHRKGSNPRPLEASLVQR	FMFQLF	
		151		200
SEQ ID NO: 320	(151)	LLDKDKGILKIADLGLSRAFTVPLKAYT	HEIVTLWYRAPEVLLGSTHYST	
SEQ ID NO: 318	(130)	-----	-----	
Consensus	(151)	-----	-----	
			T	
		201		250
SEQ ID NO: 320	(201)	AVDIWSVGCIFAEMIRRALFP	GDSEFQQLLHIFRLLGTPTEQQWPGVMA	
SEQ ID NO: 318	(130)	-----	-----	
Consensus	(201)	-----	-----	
			GDSEID	
		251		300
SEQ ID NO: 320	(251)	LRDWHVYPKWEPQDLSRAVPSLS	PEGIDLLTQMLKYNPAERISAKAALDH	
SEQ ID NO: 318	(130)	-----	-----	
Consensus	(251)	-----	-----	
		301	311	
SEQ ID NO: 320	(301)	PYFDSL	DKSQF	
SEQ ID NO: 318	(130)	-----	-----	
Consensus	(301)	-----	-----	

FIGURE 2

A.thaliana_AT1G76540.1#1	-----MDEG-----VIAVSAMD	36
A.thaliana_AT1G20930.1#1	-----MDNNGV-----KPAVSAME	38
H.annuus_CD855729#1	-----MEST-----KTAMEAF	33
G.max_Glyma14g39760.1#1	-----MEKTG-----VLSAKEAF	35
G.max_Glyma17g38210.1#1	-----MEKPGG-----GGVLSAKE	38
G.max_Glyma07g07640.1#1	-----MEKTGA-----GVVLSAKE	39
G.max_Glyma09g08250.1#1	-----MEMEKTGA-----GVAVLS	41
M.truncatula_TC117212#1	-----MEKPGET-----IAVVL	40
P.trichocarpa_scaff_II.31#1	-----MEKAGGVVAAAPTNTVVSAM	45
S.lycopersicum_TC205060#1	-----MEAAKN-----AGKSVS	39
H.vulgare_TC172640#1	MATLQHQNQGAAAPMTTGGGLRAMD	50
T.aestivum_TC282781#1	MATLQHQNQGAAAPMTTGGGLRAMD	50
O.sativa_LOC_Os08g40170.1#1	MAALHHQAAAAPVTTTDDGGELRAM	50
sorghumXP_002441781.1	-----MEINIVDKYEKLEKVGEG	31
Z.mays_c62195783gm030403@12303	-----MEINIVDKYEKLEKVGEG	31
O.sativa_LOC_Os01g67160.1#1	-----MEKYEKLEKVGEGTYGK	26
H.vulgare_TC160457#1	-----MEKYEKLEKVGEGTYGK	26
T.aestivum_TC283074#1	-----MEKYEKLEKVGEGTYGK	26
T.aestivum_TC290551#1	-----MEKYEKLEKVGEGTYGK	26
SEQ ID NO:318	-----MEKYEKLEKVGEGTYGK	26
A.thaliana_AT2G38620.1#1	-----MEKYEKLEKVGEGTYGK	26
A.thaliana_AT2G38620.2#1	-----MEKYEKLEKVGEGTYGK	26

FIGURE 3

A.thaliana_AT3G54180.1#1	-----MEKYEKLEKVGEGTYGKVYKAMEKGT	26
ChenopodiumCAC17703.1	-----MEKYEKLEKVGEGTYGKVYKAKEKTT	26
P.trichocarpa_scaff_XVI.1376#1	-----MEKYEKLEKVGEGTYGKVYKAKDKLT	26
P.trichocarpa_scaff_3712.1#1	-----MEKYEKLEKVGEGTYGKVYKAKDKLT	26
P.trichocarpa_scaff_28.331#1	-----MEKYEKLEKVGEGTYGKVYKAKDKVT	26
ricinusXP_002519974.1	-----MEKYEKLEKVGEGTYGKVYKAKDRVT	26
vitisXP_002266623.1	-----MDKYEKLEKVGEGTYGKVYKAKDKTS	26
S.lycopersicum_TC198540#1	-----MEKYEKLEKVGEGTYGKVYKAKDKAT	26
tomatoCAC15503.1	-----MEKYEKLEKVGEGTYGKVYKAKDKAT	26
tobaccoAAG01532.1	-----MEKYEKLEKVGEGTYGKVYKAKDKAT	26
anthirinumCAA66235.1	-----MEKYEKLEKVGEGTYGKVYKALEKST	26
M.truncatula_TC115190#1	-----MEKYEKLEKVGEGTYGKVYKAKEIST	26
medicagoP93321	-----MEKYEKLEKVGEGTYGKVYKAKELST	26
lotus	-----MENYEKLEKVGEGTYGKVYKAKEKST	26
G.max_Glyma07g02400.1#1	-----MEKYEKLEKVGEGTYGKVYKAREKAS	26
P.patens_NP13120379#1	-----MENYEKLEKIGEGMYGKVYKARDRRT	26
P.patens_TC47562#1	-----MENYEKIEKIGEGMYGKVYKAWDKRT	26
P.patens_NP13127560#1	-----MDEYEKLEKIGEGTYGKVYKARNKST	26
P.patens_TC43373#1	-----MDGYEKLEKVGEGTYGKVYKARDKRT	26
P.patens_TC41028#1	-----MSVSGMDNYEKLEKVGEGTYGKVYKARDKRS	31
P.patens_TC35664#1	-----MDKYEKLEKVGEGTYGKVYKARDKLT	26
	: : ** : ** : * * * * * : *	:

FIGURE 3 (continued

A.thaliana_AT1G76540.1#1	GKIVALKKTRRLHEDEEGV	PSTTLRE	SILRMLARDPHVVR-LMDVKQGLS	85
A.thaliana_AT1G20930.1#1	GMIVALKKTRRLHEDEEGV	PPTTLRE	SILRMLARDPHIVR-LMDVKQGIN	87
H.annuus_CD855729#1	GKIVALKKTRRLHEDEEGV	PPTTLRE	SILRMLSIDPHVVK-LMDVKQGVN	82
G.max_Glyma14g39760.1#1	GKIVALKKTRRLHEDEEGV	PPTTLREV	SILRMLSRDPHVVR-LMDVKQGQN	84
G.max_Glyma17g38210.1#1	GKIVALKKTRRLHEDEEGV	PPTTLREV	SILRMLSRDPHVVR-LMDVKQGQN	87
G.max_Glyma07g07640.1#1	GKIVALKKTRRLHEDQDGV	PPTTLREV	SILRMLSRDPHVVS-LMDVKQGQN	88
G.max_Glyma09g08250.1#1	GKIVALKKTRRLHEDQEGV	PPTTLREV	SILRMLSRDPHVVR-LMDVKQGQN	90
M.truncatula_TC117212#1	GKIVALKKTRRLHEDDEGV	PPTTLREV	SILRMLSRDPHVVR-LLDVKQGQN	89
P.trichocarpa_scaff_II.31#1	GKIVALKKTRRLHEDDEGV	PPTTLREV	SILRMLSRDPHVVR-LMDVKQGQN	94
S.lycopersicum_TC205060#1	GKIVALKKTRRLHEDEEGV	PPTTLRE	SLLKMLSRDPHIVR-LMDVKQGQN	88
H.vulgare_TC172640#1	GRIVALKKTRLPEDDEGV	PPTALREV	SLLRMLSQDPHVVR-LLDLKQGQN	99
T.aestivum_TC282781#1	GRIVALKKTRLPEDDEGV	PPTALREV	SLLRMLSQDPHVVR-LLDLKQGQN	99
O.sativa_LOC_Os08g40170.1#1	GRIVALKKTRLPEDDEGV	PPTALREV	SLLRMLSQDSHVVR-LLDLKQGQN	99
sorghumXP_002441781.1	GQIVALKKTRLEMDEEGI	PPTALRE	SLLNLLSHSIYVVR-LLAVEQAAK	80
Z.mays_c62195783gm030403@12303	GQIVALKKTRLEMDEEGI	PPTALRE	SLLNLLSHSIYIVR-LLAVEQAAK	80
O.sativa_LOC_Os01g67160.1#1	GQIVALKKTRLEMDEEGI	PPTALRE	SILRLLSQSLYVVR-LLSVEQATK	75
H.vulgare_TC160457#1	GQVVALKKTRLEMDDDEGI	PPTALRE	SLLRLLSSSLYVVR-LLAVEQTTK	75
T.aestivum_TC283074#1	GQVVALKKTRLEMDDDEGI	PPTALRE	SLLRLLSSSLYVVR-LLAVEQATK	75
T.aestivum_TC290551#1	GQVVALKKTRLEMDEEGI	PPTALRE	SLLRLLSASLYVVR-LLSVEQVTK	75
SEQ ID NO: 318	GKLVALKKTRLEMDEEGI	PPTALRE	SLLQMLSQSIYIVR-LLCVEHVIQ	75
A.thaliana_AT2G38620.1#1	GKLVALKKTRLEMDEEGI	PPTALRE	SLLQMLSQSIYIVR-LLCVEHVIQ	75
A.thaliana_AT2G38620.2#1	GKLVALKKTRLEMDEEGI	PPTALRE	SLLQMLSQSIYIVR-LLCVEHVIQ	75

FIGURE 3 (continued

A.thaliana_AT3G54180.1#1	GKLVALKKTRLEMDDEEGIPPTALREISLLQMLSTSIYVVR-LLCVEHVHQ	75
ChenopodiumCAC17703.1	GQLVALKKTRLEMDDEEGVPPTALREVSLLQMLSQSLYVVRLLLCVEHVDK	76
P.trichocarpa_scaff_XVI.1376#1	GQLVALKKTRLQMDDEEGVPPTALREVSLLQMLSQSLYVVR-LLSVEHLD-	74
P.trichocarpa_scaff_3712.1#1	GQLVALKKTRLQMDDEEGVPPTALREVSLLQMLSQSLYVVR-LLSVEHID-	74
P.trichocarpa_scaff_28.331#1	GQIVALKKTRLEMDDEEGVPPTALREVSLLQMLSQSLYVVR-LLSVEHVD-	74
ricinusXP_002519974.1	GQIVALKKTRLEMDDEEGIPPTALREVSLLQMLSQSLYVVR-LLSVEHVD-	74
vitisXP_002266623.1	GQVVALKKTRLEMDDEEGVPPTALREVSLLQMLSQSLYVVR-LLCVEHLD-	74
S.lycopersicum_TC198540#1	GQLVALKKTRLEMDDEEGIPPTALREISLLQMLSHSLYIVR-LLCVEHID-	74
tomatoCAC15503.1	GQLVALKKTRLEMDDEEGIPPTALREISLLQMLSHSLYIVR-LLCVEHID-	74
tobaccoAAG01532.1	GQLVALKKTRLEMDDEEGIPPTALREISLLQMLSNSLYIVR-LLCVEQID-	74
anthirinumCAA66235.1	GQVVALKKTRLEMDDEEGVPPTALREVSLLQMLSQSLYVVR-LLSVEHVDC	75
M.truncatula_TC115190#1	GQIVALKKTRLEMDDEEGVPPTALREVSLLQMLSQSLYIVR-LLNVEHIDK	75
medicagoP93321	GQIVALKKTRLEMDDEEGVPPTALREVSLLQMLSQSLYIVR-LLNVEHIDK	75
lotus	GQVVALKKTRLEMDDEEGVPPTALREVSLLQMLSQSIYIVR-LLKVEHVDK	75
G.max_Glyma07g02400.1#1	GSLVALKKTRLEMDDEEGVPPTALREVSLLQLLSQSIYIVR-LLSVEHVDK	75
P.patens_NP13120379#1	GALVALKKIKIENEEEGIPATTLREVSLLQMLSKCIYFVR-LLGVEHLHK	75
P.patens_TC47562#1	GAMVALKKNKIENEKEGIPATTLREVSLLQVLSTCIYFVR-LLSVENSHK	75
P.patens_NP13127560#1	GRLVALKKTRLEMEDEGV PSTALREVSLLQMLSHSIYIVR-LLSVESVDE	75
P.patens_TC43373#1	GRFVALKKTRLENEDEGV PSTTLREVSLLQMLSHSIYIVR-LLCVESVDE	75
P.patens_TC41028#1	GQLVALKKTRLEMEEEGV PSTALREVSLLQMLSHSMYIVR-LLCVEHVEK	80
P.patens_TC35664#1	GQLVALKKTRLEMEEEGV PSTALREVSLLQMLSHSIYIVR-LLCVEHVEK	75
	* .***** :: :.:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*	

FIGURE 3 (continued

A.thaliana_AT1G76540.1#1	KEG-----KTVLYLVFFEYMDTDVKKFIRSFRS--T	113
A.thaliana_AT1G20930.1#1	KEG-----KTVLYLVFFEYVDTDLKKFIRSFRQ--A	115
H.annuus_CD855729#1	KNG-----NTVLYLVFFEYMETDLKKFIRSFRQ--T	110
G.max_Glyma14g39760.1#1	KEG-----KTVLYLVFFEYMDTDLKKFIRSFRQ--S	112
G.max_Glyma17g38210.1#1	KEG-----KTVLYLVFFEYMDTDLKKFIRSFRQ--T	115
G.max_Glyma07g07640.1#1	KEG-----KTVLYLVFFEYMDTDLKKFIRSFDQ--P	116
G.max_Glyma09g08250.1#1	KEG-----KTVLYLVFFEYMDTDLKKFIRSFRQ--T	118
M.truncatula_TC117212#1	KEG-----KTVLYLVFFEYMDTDLKKFIRSFRQ--T	117
P.trichocarpa_scaff_II.31#1	KQG-----KTVLYLVFFEYMDTDLKKFIRSFLQ--T	122
S.lycopersicum_TC205060#1	KEG-----KTVLYLVFFEYMDTDVKKFIRTFRA--K	116
H.vulgare_TC172640#1	KEG-----QTILYLVFFEYMDTDLKKFIRGHRQ--N	127
T.aestivum_TC282781#1	KEG-----QTILYLVFFEYMDTDLKKFIRGHRQ--N	127
O.sativa_LOC_Os08g40170.1#1	KEG-----QTILYLVFFEYMDTDLKKFIRAHQR--N	127
sorghumXP_002441781.1	N-----GKPVLYLVFEFLDTDLKKYLDVYRRGPA	109
Z.mays_c62195783gm030403@12303	N-----GKPVLYLVFEFLDTDLKKYLDVYRRGPS	109
O.sativa_LOC_Os01g67160.1#1	N-----GKPVLYLVFEFLDTDLKKFVDAYRKGN	104
H.vulgare_TC160457#1	GGGAGG-----GGKPVLYLVFEFLDTDLKKFVDAYRRGPA	110
T.aestivum_TC283074#1	GGGAGS-----GGKPVLYLVFEFLDTDLKKFVDAYRRGPA	110
T.aestivum_TC290551#1	P-----CGKTVLYLVFEFLDTDLKKFVDGFRKGPS	105
SED ID NO: 318	SKDSTVSHSP-----KSNLYLVFFEYLDTDLKKFIDSHRKGSN	112
A.thaliana_AT2G38620.1#1	SKDSTVSHSP-----KSNLYLVFFEYLDTDLKKFIDSHRKGSN	112
A.thaliana_AT2G38620.2#1	SKDSTVSHSP-----KSNLYLVFFEYLDTDLKKFIDSHRKGSN	112

FIGURE 3 (continued

A.thaliana_AT3G54180.1#1	P--STKSQST-----KSNLYLVFFEYLDTDLKKFIDSYRKGP	110
ChenopodiumCAC17703.1	K--NTTGIS-----KPILYLVFFEYLDTDLKKFIDSHRKGS	111
P.trichocarpa_scaff_XVI.1376#1	-ANNSD--DDSD-----KSNLYLVFFEFLDADLKKFIDSHRKGP	109
P.trichocarpa_scaff_3712.1#1	-ANNSD--DDSD-----KSNLYLVFFEFLDADLKKFIDSHRKGP	109
P.trichocarpa_scaff_28.331#1	-ANSTNSNLDDSNPSKRNHNNNKSNLYLVFFEYLDTDLKKFIDSHRKGA	123
ricinusXP_002519974.1	-VNSAD--DNN---INKENAIKSNLYLVFFEYLDTDLKKFIDSHRKGP	117
vitisXP_002266623.1	-KNGKP-----FLYLVFFEYLDTDLKKFIDSHRKPP	104
S.lycopersicum_TC198540#1	-KNGKP-----ILYLVFFEYLDTDLKKFIDSHRKGP	104
tomatoCAC15503.1	-KNGKP-----ILYLVFFEYLDTDLKKFIDSHRKGP	104
tobaccoAAG01532.1	-KNGKP-----LLYLVFFEYLDTDLKKFVDSHRKGP	104
anthirinumCAA66235.1	AKNGKP-----LLYLVFFEYLDTDLKKFIDSHRKGP	106
M.truncatula_TC115190#1	PPKNATHTP-----AKPLLYLVFFEYLDTDLKKFIDTFRKGV	112
medicagoP93321	PPKNATHTP-----SKPLLYLVFFEYLDTDLKKFIDTFRKGT	112
lotus	TPKNNHHPSAA---ADHPPPPSKPLLYLVFFEYLDTDLKKYIDTYRKGP	122
G.max_Glyma07g02400.1#1	VPKSQKSSN-----PLTKPILYLVFFEYLDTDLKKFIDSHRKGP	115
P.patens_NP13120379#1	NG-----KPALYLVFFEYIDTDLRRFLDLSWPG-L	103
P.patens_TC47562#1	NG-----KTSLFMVFFEYIDTDLRRFMDLSWPG-P	103
P.patens_NP13127560#1	NG-----KPQLYLVFFEYLDSDLKKFIDSYGRGSG	104
P.patens_TC43373#1	NG-----KPQLYLVFFEYLDSDLKKYIDLHGRGPG	104
P.patens_TC41028#1	GG-----KPMLYLVFFEYMDTDLKKYIDLHGRGPS	109
P.patens_TC35664#1	GG-----KPMLYLVFFEYMDTDLKKYIDLHGRGPS	104
	*::**::*::*::*::*::*::*::*::	

FIGURE 3 (continued

A.thaliana_AT1G76540.1#1	GKNIP	TQTI	KS	LM	YQ	LC	KG	MA	FCH	GH	GIL	HR	DL	KP	HN	LL	MD	PK	TM	RL	KIA	163												
A.thaliana_AT1G20930.1#1	GQNI	PQNT	VK	CL	MY	Q	LC	KG	MA	FCH	GH	GVL	HR	DL	KP	HN	LL	MD	RK	TM	TL	KIA	165											
H.annuus_CD855729#1	GDNIP	PQI	VK	SL	MY	Q	LC	KG	IA	FCH	AH	GIL	HR	DL	KP	HN	LL	MD	RK	TI	ML	KIA	160											
G.max_Glyma14g39760.1#1	GETI	PPHI	IK	SL	MY	Q	LC	KG	VA	FCH	GH	GIL	HR	DL	KP	HN	LL	MD	RK	TM	ML	KIA	162											
G.max_Glyma17g38210.1#1	GQTV	PPQ	TI	KS	LM	Y	Q	LC	KG	VA	FCH	GH	GIL	HR	DL	KP	HN	LL	MD	PK	TM	ML	KIA	165										
G.max_Glyma07g07640.1#1	GQNI	PPET	IK	SL	MY	Q	LC	KG	IA	FCH	GH	GIL	HR	DL	KP	HN	LL	MD	RK	TM	ML	KIA	166											
G.max_Glyma09g08250.1#1	GQSI	PPQ	TI	KS	LM	Y	Q	LC	KG	IA	FCH	GH	GIL	HR	DL	KP	HN	LL	MD	RK	TM	ML	KIA	168										
M.truncatula_TC117212#1	GQNI	PPP	TI	K	L	MY	Q	LC	KG	VA	FCH	GH	GIL	HR	DL	KP	HN	LL	MD	RK	TM	ML	KIA	167										
P.trichocarpa_scaff_II.31#1	GENI	PV	KN	V	K	SL	MY	Q	LC	KG	VA	FCH	GH	GVL	HR	DL	KP	HN	LL	MD	RK	TM	ML	KIA	172									
S.lycopersicum_TC205060#1	GETM	PL	KI	V	K	SL	MY	Q	LC	KG	VA	FCH	GH	GVL	HR	DL	KP	HN	LL	MD	RK	TN	VL	KLA	166									
H.vulgare_TC172640#1	HQKI	PA	HT	V	KI	L	MY	Q	LC	KG	VA	FCH	GR	GVL	HR	DL	KP	HN	LL	MD	RK	TM	AL	KIA	177									
T.aestivum_TC282781#1	HEKI	PA	HT	V	KI	L	MY	Q	LC	KG	VA	FCH	GR	GVL	HR	DL	KP	HN	LL	MD	RK	TM	AL	KIA	177									
O.sativa_LOC_Os08g40170.1#1	LQKI	PV	PT	V	KI	L	MY	Q	LC	KG	VA	FCH	GR	GVL	HR	DL	KP	HN	LL	MD	RK	TM	AL	KIA	177									
sorghumXP_002441781.1	ARPL	PA	TL	I	K	N	F	L	Y	Q	LC	KG	VA	HCH	GH	GVL	HR	DL	KP	Q	N	L	L	V	D	K	D	K	G	I	L	KIA	159	
Z.mays_c62195783gm030403@12303	ARPL	PA	TL	I	K	N	F	L	Y	Q	LC	KG	VA	HCH	SH	GVL	HR	DL	KP	Q	N	L	L	V	D	K	E	K	G	I	L	KIA	159	
O.sativa_LOC_Os01g67160.1#1	PRPL	PT	N	V	I	K	S	F	L	Y	Q	LC	KG	VA	HCH	GH	GVL	HR	DL	KP	Q	N	L	L	V	D	K	E	K	G	I	L	KIA	154
H.vulgare_TC160457#1	PKPL	PT	H	V	V	K	S	F	L	Y	Q	LC	KG	IA	HCH	GH	GVL	HR	DL	KP	Q	N	L	L	V	D	K	E	K	M	I	L	KIA	160
T.aestivum_TC283074#1	PKPL	PT	Q	V	V	K	S	F	L	Y	Q	LC	KG	IA	HCH	GH	GVL	HR	DL	KP	Q	N	L	L	V	D	K	E	K	M	I	L	KIA	160
T.aestivum_TC290551#1	PRPV	PT	Q	V	V	K	S	F	L	Y	Q	LC	KG	TA	HCH	GH	GVL	HR	DL	KP	Q	N	L	L	I	D	K	E	K	G	V	L	KIA	155
SEQ ID NO: 318	PRPLE	AS	L	V	Q	R	F	M	F	Q	L	F	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	129	
A.thaliana_AT2G38620.1#1	PRPLE	AS	L	V	Q	R	F	M	F	Q	L	F	K	G	VA	HCH	SH	GVL	HR	DL	KP	Q	N	L	L	D	K	D	K	G	I	L	KIA	162
A.thaliana_AT2G38620.2#1	PRPLE	AS	L	V	Q	R	F	M	F	Q	L	F	K	G	VA	HCH	SH	GVL	HR	DL	KP	Q	N	L	L	D	K	D	K	G	I	L	KIA	162

FIGURE 3 (continued

A.thaliana_AT3G54180.1#1	PKPLEPFLIQKLMFQLCKGVAHCHSHGVLHRDLKQPQNLLLVKDKELLKIA	160
ChenopodiumCAC17703.1	PRPLPPCQIQSFLFQLLKGVSHCHSHGVLHRDLKQPQNLLLVKDKGILKIA	161
P.trichocarpa_scaff_XVI.1376#1	PRPLSPSLIQSFLFQLCKGVAHCHSHGVLHRDLKQPQNLLLDQERGILKIA	159
P.trichocarpa_scaff_3712.1#1	PRPLSPSLIQSFLFQLCKGVAHCHSHGVLHRDLKQPQNLLLDQERGILKIA	159
P.trichocarpa_scaff_28.331#1	PRPLSPSLIQSFLFQLCKGVAHCHSHGVLHRDLKQPQNLLLDQEKGILKIA	173
ricinusXP_002519974.1	PTPLAPGLIQSFLFQLCKGVAHCHSHGVLHRDLKQPQNLLLDQEKGILKIA	167
vitisXP_002266623.1	PRPMPPALIQSFLYQLCKGVAHCHSHGVLHRDLKQPQNLLVLDKDKGILKIA	154
S.lycopersicum_TC198540#1	ARALPTALIQSFLFQLCKGVAHCHSHGVLHRDLKQPQNLLVLDKEKGILKIA	154
tomatoCAC15503.1	ARALPTALIQSFLFQLCKGVAHCHSHGVLHRDLKQPQNLLVLDKEKGILKIA	154
tobaccoAAG01532.1	PRPLPPSLIQSFLYQLCKGVAHCHSHGVLHRDLKQPQNLLVLDKEKGILKIA	154
anthirinumCAA66235.1	PRPLPPQQIQSFLFQLCKGVSHCHAHGVLHRDLKQPQNLLLDKDKGVLKIA	156
M.truncatula_TC115190#1	PRPLPNTLVQSFLFQLCKGVAHCHSHGVLHRDLKQPQNLLLDQAKGILKIA	162
medicagoP93321	PRPLPNTLVQSFLFQLCKGVAHCHSHGVLHRDLKQPQNLLLDQAKGILKIA	162
lotus	PRPLPPTLIQSFLFQLCMGVAHCHSHGVLHRDLKQPQNLLLDQQKGILKIA	172
G.max_Glyma07g02400.1#1	PRPLPPPLIQSFLFQLCKGVAHCHSHGVLHRDLKQPQNLLLDQHKGILKIA	165
P.patens_NP13120379#1	DNPLPQNTIKSFVYQLLKGVAHCHSHGVMHRDLKQPQNLLVDWSKGLLKIA	153
P.patens_TC47562#1	DNPLPQNTIKSFMFQQLLKGVAHCHSHGIMHRDLKQPQNLLVDWKKGLLKIA	153
P.patens_NP13127560#1	-NPIPAKTIQSFMYQLLKGVAHCHGHGVMHRDLKQPQNLLVDKEKGILKIA	153
P.patens_TC43373#1	TNPIPAKTVQSFMYQLLKGVAHCHSHGVMHRDLKQPQNLLVDKEKGILKIA	154
P.patens_TC41028#1	GKPLPPKVQVQSFMYQLCTGLAHCHGHGVMHRDLKQPQNLLVDKQTRRLKIA	159
P.patens_TC35664#1	GKPLPPKVQVQSFMYQLCTGLAHCHGHGVMHRDLKQPQNLLVDKQTRRLKIA	154
	: :: : :: *	

FIGURE 3 (continued

A. thaliana_AT1G76540.1#1	DLGLARAF	TLPMKKY	THEIL	TLWYRAPEV	LLGATHYST	AVDMWSV	GCIFA	213	
A. thaliana_AT1G20930.1#1	DLGLARAF	TLPMKKY	THEIL	TLWYRAPEV	LLGATHYST	GVDMWSV	GCIFA	215	
H. annuus_CD855729#1	DLGLARAF	TLPIKKY	THEIL	TLWYRAPEV	LLGSTHYST	AVDMWSV	GCIFA	210	
G.max_Glyma14g39760.1#1	DLGLARAF	TVPIKKY	THEIL	TLWYRAPEV	LLGATHYSMA	AVDMWSV	GCIFA	212	
G.max_Glyma17g38210.1#1	DLGLARAF	TVPIKKY	THEIL	TLWYRAPEV	LLGATHYSMA	VDIWSV	GCIFA	215	
G.max_Glyma07g07640.1#1	DLGLARAF	TVPIKKY	THEIL	TLWYRAPEV	LLGATHYSMA	VDIWSV	GCIFA	216	
G.max_Glyma09g08250.1#1	DLGLARAF	TVPIKKY	THEIL	TLWYRAPEV	LLGATHYSMA	VDIWSV	GCIFA	218	
M.truncatula_TC117212#1	DLGLARAF	TVPLKKY	THEIL	TLWYRAPEV	LLGATHYSMA	VDMWSV	ACIFA	217	
P.trichocarpa_scaff_II.31#1	DLGLARAF	TLPIKKY	THEIL	TLWYRAPEV	LLGATHYST	AVDVWSV	GCIFA	222	
S.lycopersicum_TC205060#1	DFGLGRAY	TLPIKKY	THEIL	TLWYRAPEV	LLGATHYST	AVDMWSV	GCIFA	216	
H.vulgare_TC172640#1	DLGLSRAF	TVPLKKY	THEIL	TLWYRAPEV	LLGATHYST	VPDMWSV	GCIFA	227	
T.aestivum_TC282781#1	DLGLSRAF	TVPLKKY	THEIL	TLWYRAPEV	LLGATHYST	VPDMWSV	GCIFA	227	
O.sativa_LOC_Os08g40170.1#1	DLGLSRAF	TVPLKKY	THEIL	TLWYRAPEV	LLGA	AHYST	VPDIWSV	GCIFA	227
sorghumXP_002441781.1	DLGLGRAFT	VPMKSY	THEI	VT	TLWYRAPEV	LLGATHYST	GVDMWSV	GCIFA	209
Z.mays_c62195783gm030403@12303	DLGLGRAFT	VPMKSY	THEI	VT	TLWYRAPEV	LLGATHYST	GVDMWSV	GCIFA	209
O.sativa_LOC_Os01g67160.1#1	DLGLGRAFT	VPMKSY	THEI	VT	TLWYRAPEV	LLGSTHYST	GVDIWSV	GCIFA	204
H.vulgare_TC160457#1	DLGLSRAF	TVPMKSY	THEI	VT	TLWYRAPEV	LLGATHYST	GVDIWSI	GCIFA	210
T.aestivum_TC283074#1	DLGLSRAF	TVPMKSY	THEI	VT	TLWYRAPEV	LLGATHYST	SVDIWSI	GCIFA	210
T.aestivum_TC290551#1	DLGLSRAF	TVPMKSY	THEI	VT	TLWYRAPEV	LLGATHYST	GVDIWSI	GCIFA	205
SEQ ID NO: 318	-----								
A. thaliana_AT2G38620.1#1	DLGLSRAF	TVPLKAY	THEI	VT	TLWYRAPEV	LLGSTHYST	AVDIWSV	GCIFA	212
A. thaliana_AT2G38620.2#1	DLGLSRAF	TVPLKAY	THEI	VT	TLWYRAPEV	LLGSTHYST	AVDIWSV	GCIFA	212

FIGURE 3 (continued

A.thaliana_AT3G54180.1#1	DLGLGRAFTVPLKSYTHEIVTLWYRAPEVLLGSTHYSTGVDMWSVGCIFA	210
ChenopodiumCAC17703.1	DLGLGRAFTVPMKSYTHEIVTLWYRAPEVLLGSTHYSTGVDIWSVGCIFA	211
P.trichocarpa_scaff_XVI.1376#1	DLGLGRAFTVPLKSYTHEIVTLWYRAPEVLLGSTHYSI AIDMWSVGCIFA	209
P.trichocarpa_scaff_3712.1#1	DLGLGRAFTVPLKSYTHEIVTLWYRAPEVLLGSTHYSI AIDMWSVGCIFA	209
P.trichocarpa_scaff_28.331#1	DLGLGRAFTVPLKSYTHEIVTLWYRAPEVLLGSTHYSTGVDMWSVGCIFA	223
ricinusXP_002519974.1	DLGLGRAFTVPLKSYTHEIVTLWYRAPEVLLGSTHYSTAVDMWSVGCIFA	217
vitisXP_002266623.1	DLGLGRAFTVPLKSYTHEIVTLWYRAPEVLLGSTHYSTGVDMWSVGCIFA	204
S.lycopersicum_TC198540#1	DLGLGRAFTVPMKSYTHEIVTLWYRAPEVLLGSTHYSTAVDMWSVGCIFA	204
tomatoCAC15503.1	DLGLGRAFTVPMKSYTHEIVTLWYRAPEVLLGSTHYSTAVDMWSVGCIFA	204
tobaccoAAG01532.1	DLGLGRAFTVPIKSYTHEIVTLWYRAPEVLLGSTHYSTAVDMWSVGCIFA	204
anthirinumCAA66235.1	DLGLARAFTVPLKSYTHEIVTLSYRAPEVLLGSSHYSTAVDMSSVGCIFA	206
M.truncatula_TC115190#1	DLGLGRAFTVPLKSYTHEIVTLWYRAPEVLLGSSYSTYSTSVDIWSVGCIFA	212
medicagoP93321	DLGLGRAFTVPLKSYTHEIVTLWYRAPEVLLGSSYSTYSTGVDIWSVGCIFA	212
lotus	DLGLGRAFTVPLKSYTHEIVTLWYRAPEVLLGSTHYSTGVDMWSVGCIFA	222
G.max_Glyma07g02400.1#1	DLGLGRAFTVPLKSYTHEIVTLWYRAPEVLLGSTHYSTGVDIWSVGCIFA	215
P.patens_NP13120379#1	DLGLGRAFTVPVKSYTHEVVTLWYRAPEILLGASHYSTPVDMWSVGCIFA	203
P.patens_TC47562#1	DLGLGRAFTVPMKSYTHEVVTLWYRAPEILLGASHYSTPVDMWSVGCVFA	203
P.patens_NP13127560#1	DLGLGRAFTIPLKSYTHEIVTLWYRAPEVLLGSSQYSTSVDMWSVGCIFA	203
P.patens_TC43373#1	DLGLGRAFTVPLKSYTHEIVTLWYRAPEVLLGASHYSTSVDVWSVGCIFA	204
P.patens_TC41028#1	DLGLGRAFTVPMKSYTHEIVTLWYRAPEVLLGATHYSLPVDIWSVGCIFA	209
P.patens_TC35664#1	DLGFGRAFTVPMKSYTHEIVTLWYRAPXVFS-----E	186

FIGURE 3 (continued

13/42

A.thaliana_AT1G76540.1#1	ELVTNQAI	FQGDSE	LQQLLH	IFKLF	GTPNEE	MWPGV	STLKN	WHEY	PPQW	KP	263						
A.thaliana_AT1G20930.1#1	ELVTKQAI	FAGDSE	LQQLRI	FRLLG	TPNEE	VWPGV	SKLKD	WHEY	PPQW	KP	265						
H.annuus_CD855729#1	ELVTNTAH	FC-----	WRF-----								223						
G.max_Glyma14g39760.1#1	ELVTKQAL	FPGDSE	LQQLLH	IFRLLG	TPNED	VWPGV	SKLMN	WHEY	PPQW	NP	262						
G.max_Glyma17g38210.1#1	ELVTKQAL	FPGDSE	LQQLLH	IFRLLG	TPNED	VWPGV	SKLMN	WHEY	PPQW	NP	265						
G.max_Glyma07g07640.1#1	ELVTRRAL	FPGDSE	LQQLLH	IFRLLG	TPNEE	VWPGV	SKLKD	WHEY	PPQW	NS	266						
G.max_Glyma09g08250.1#1	ELVTKQAL	FAGDSE	LQQLLH	IFRLLG	TPNEE	VWPGV	SKLKD	WHEY	PPQW	NP	268						
M.truncatula_TC117212#1	ELVTKTAL	FPGDSE	LQQLLH	IFRLLG	TPNED	VWPGV	SKLMN	WHEY	PPQW	GP	267						
P.trichocarpa_scaff_II.31#1	ELATKQPL	FPGDSE	LQQLLH	IFRLLG	TPNEE	MWPGV	SNLMN	WHEY	PPQW	KP	272						
S.lycopersicum_TC205060#1	ELVTNQAL	FPGDSE	LQQLLH	IFRLLG	TPNEE	LWPGV	SKLVN	WHEY	PPQW	KP	266						
H.vulgare_TC172640#1	ELITTTAL	FPGDSE	VQQLLH	IFQLLG	TPNEE	VWPGV	GKLPN	WHEY	PPQW	NV	277						
T.aestivum_TC282781#1	ELITTTAL	FPGDSE	VQQLLH	IFKLLG	TPNED	VWPGV	GKLPN	WHEY	PPQW	NV	277						
O.sativa_LOC_Os08g40170.1#1	ELATNQPL	FAGDSE	VQQLLH	IFKLLG	TPNEQ	VWPGV	SKLPN	WHEY	PPQW	NP	277						
sorghumXP_002441781.1	EMARRQAL	FPGDSE	LQQLLH	IFRLLG	TPSEE	QWPGV	SELRD	WHE	FFPQ	WKP	259						
Z.mays_c62195783gm030403@12303	EMARRQAL	FPGDSE	LQQLLH	IFRLLG	TPTEE	QWPGV	SDLRD	WHE	FFPQ	WKP	259						
O.sativa_LOC_Os01g67160.1#1	EMVRRQAL	FPGDSE	LQQLLH	IFRLLG	TPTEE	QWPGV	TDLRD	WHE	FFPQ	WKP	254						
H.vulgare_TC160457#1	EMVRKQAL	FPGDSE	LQQLLH	IFRLMG	TPTEED	WPGVTAL	LRD	WHE	FFPQ	WKA	260						
T.aestivum_TC283074#1	EMVXKQAL	FPGDSE	LQQLLH	IFKLLGN---	SYRGG	LAW	SN	C	SE	G---L	252						
T.aestivum_TC290551#1	EMVRRQAL	FPGDSE	LQQLLH	IFRLLG	TPNEE	KWPGV	CALRD	WHD	Y	PPQW	KP	255					
SEQ ID NO: 318	-----																
A.thaliana_AT2G38620.1#1	EMIRRRQAL	FPGDSE	FQQLLH	IFRYL-----	LLV	F	D	A	N	V	Y	VRN	Q	251			
A.thaliana_AT2G38620.2#1	EMIRRRQAL	FPGDSE	FQQLLH	IFRLLG	TPTEQ	QWPGV	MALRD	W	H	V	Y	P	K	W	E	P	262

FIGURE 3 (continued

A.thaliana_AT3G54180.1#1	EMVRRQALFPGDSEFQQLLHIFRLLGTPTEQQWPGVSTLRDWHVYPKWEP	260
ChenopodiumCAC17703.1	EMVRRQALFPGDSEFQQLLHIFRLLGTPTEQQWPGVTSLRGMCIRNGSL	261
P.trichocarpa_scaff_XVI.1376#1	EMSRRQALFPGDSEIQQLL-----GTPTEEQWPGVTSLRDWHVYPKWEP	253
P.trichocarpa_scaff_3712.1#1	EMSRRQALFPGDSEIQQLLHIFRLLGTPTEEQWPGVTSLRDWHVYPKWEP	259
P.trichocarpa_scaff_28.331#1	EMSRRQALFPGDSEFQQLLHIFRLLGTPTEEQWPGVTSLRDWHVYPKWEP	273
ricinusXP_002519974.1	EMARRQALFPGDSEFQQLLHIFRLLGTPTEKQWPGVTSFRDWHVYPQWEP	267
vitisXP_002266623.1	EMVRRQALFPGDSEFQQLLHIFRLLGTPTEKQWPGVSSLRDWHVYPQWEP	254
S.lycopersicum_TC198540#1	EMVRRQALFPGDSEFQQLLHIFRLLGTPTDKQWPGVSSLRDWHVYPQWEP	254
tomatoCAC15503.1	EMVRRQALFPGDSEFQQLLHIFRLLGTPTDKQWPGVSSLRDWHVYPQWEP	254
tobaccoAAG01532.1	EMVRRQALFPGDSEFQQLLHIFRLLGTPTEKQWPGVSSLRDWHVYPKWEP	254
anthirinumCAA66235.1	EMVRRQALFPGDSEFQQLLHIFRLLGTPSDEQWPGVSSLRDWHVYPQWEP	256
M.truncatula_TC115190#1	EMVRRQALFPGDSEFQQLLNIFKLLGTPTDQQWPGVSSLRDWHVYPRWEP	262
medicagoP93321	EMVRRQALFPGDSEFQQLLNIFKLLGTPTEQQWPGVSSLRDWHVYPRWEP	262
lotus	EMARRQALFPGDSEFQQLLNIFKVLGTPTEEQWPGVTSLRDWHVYPRWEP	272
G.max_Glyma07g02400.1#1	EMVRRQALFPGDSEFQQLIHIFKMLGTPTEENWPGVTSLRDWHVYPRWEP	265
P.patens_NP13120379#1	ELCRKTPLPFGNSELQQLLYIFRLLGTPNEQVWPGVTTLRDWHAYPQWRA	253
P.patens_TC47562#1	ELCRKSPLPFGNSELQQLLFIFRLLGTPNEQIWPGVTTLRNWHSYPPQWKP	253
P.patens_NP13127560#1	ELSRKAPLPFGDSEIQQLLHIFRLLGTPTEESWPGVKKLRDWHYPPQWQP	253
P.patens_TC43373#1	ELSRKAPLPFGDSEIQQLLHIFRMLGTPKEECWPGVNKLRDWHYPPQWPA	254
P.patens_TC41028#1	ELVRKMPLFTGDSEIQQLLHIFRLLGTPNETIWPGVSQHRDWHYPPQWRP	259
P.patens_TC35664#1	QPTTPFLLIFG-----LLDVFSLSFFAKCRSSPGIQNFNSSSTSSGYLA	230

FIGURE 3 (continued

A.thaliana_AT1G76540.1#1	STLSSAVP-----NLDEAGVDLLSKMLQYEP	AKRISAKMAMEHPYFDDL	307
A.thaliana_AT1G20930.1#1	LSLSTAVP-----NLDEAGLDLLSKMLEYEP	AKRISAKKAMEHPYFDDL	309
H.annuus_CD855729#1	-----	-----	
G.max_Glyma14g39760.1#1	QSLSTAVP-----SLDELGLDLLSQMLKYEP	SKRISAKKAMEHVYFDDL	306
G.max_Glyma17g38210.1#1	QSLSTAVP-----SLDELGLDLLSQMLKYEP	SKRISAKKAMEHAYFDDL	309
G.max_Glyma07g07640.1#1	QSLSTAVP-----GLEELGLDLLSQMLEYEP	SKRISAKKAMEHAYFDDL	310
G.max_Glyma09g08250.1#1	KSLSTAVP-----GLDELGLDLLSQMLEYEP	SKRISAKKAMEHAYFNDL	312
M.truncatula_TC117212#1	QSLSKAVP-----GLEEAGVDLLSQMLQYEP	SKRRLSAKKAMEHPYFDDL	311
P.trichocarpa_scaff_II.31#1	QSLSSSVT-----NLDKDGLDLLSQMLQYD	PSKRISAKKAMEHPYFDEL	316
S.lycopersicum_TC205060#1	QPLSTVVP-----GLDEdGIHLLSEMLHYEP	SRRIISAKKAMEHPYFDDL	310
H.vulgare_TC172640#1	S-----	-----KAVFCNSQSRCRWY---	292
T.aestivum_TC282781#1	SKLSSVIP-----SLDADGIDLLEKMLQYEP	AKRISAKKAMEHPYFDDV	321
O.sativa_LOC_Os08g40170.1#1	SKVSDLVH-----GLDADALDLEKMLQYEP	SKRISAKKAMEHPYFNdv	321
sorghumXP_002441781.1	QSLARVVP-----TLEPEGVDLLSKMLQLD	PSNRI SAI AAMEHPYFDSL	303
Z.mays_c62195783gm030403@12303	QGLARVVP-----TLEPEGVDLLSKMLQLD	PSNRI SAL AAMEHPYFNsl	303
O.sativa_LOC_Os01g67160.1#1	QILERQVP-----SLEPEGVDLLSKMLQYN	PANRI SAK AAMEHPYFDSL	298
H.vulgare_TC160457#1	QRMTRAVP-----TLETegIDLLSKMLQFD	PANRI SAK AALEHPYFNsl	304
T.aestivum_TC283074#1	ARVS-SVE-----STEPRRNDLLSEMFQFD	PGKRISAKVGLKIPFSTAS	295
T.aestivum_TC290551#1	QHLsRAVP-----TLEPEGLDLLAKMLQFD	PANRI SAK AAMDHPYFDSL	299
SEQ ID NO: 318	-----	-----	
A.thaliana_AT2G38620.1#1	FRTL AH-----	-----	257
A.thaliana_AT2G38620.2#1	QDLsRAVP-----SLSPEGIDLLTQMLKYN	PAERI SAK AALDHPYFDSL	306

FIGURE 3 (continued

A.thaliana_AT3G54180.1#1	QDLTLAVP-----SLSPQGVDDLTKMLKYNPAERISAKTALDHPYFD	304
ChenopodiumCAC17703.1	RILLVQFHRWDLMVLI SYLKCSNMIQLKGSRRQLWIIPSLIILISLSK	311
P.trichocarpa_scaff_XVI.1376#1	QNLARAVP-----SLGPQGVDDL SKMLKYDPAERISAKAAMDHPYFD	297
P.trichocarpa_scaff_3712.1#1	QNLARAVP-----SLGPQGVDDL SKMLKYDPAERISAKAAMDHPYFD	303
P.trichocarpa_scaff_28.331#1	QNLARVVQ-----SLGPEGVDLLSKMLKYDPAERISAKAAMDHPYFD	317
ricinusXP_002519974.1	QNLARAVS-----SLGPDGVDLLSEMLKYDPAERISAKAAMDHPYFD	311
vitisXP_002266623.1	QNLARAVP-----SLGPDGVDLLSKMLKYDPSEIRISAKAALDHPYFD	298
S.lycopersicum_TC198540#1	QNLASAVP-----ALGPDGVDLLTKMLKFDPSPDRISAKAALDHPYFD	298
tomatoCAC15503.1	QNLASAVP-----ALGPDGVDLLTKMLKFDPSPDRISAKAALDHPYFD	298
tobaccoAAG01532.1	QNLASAVP-----ALGPDGVDLLTKMLQYDPADRIISAKAALDHPYFD	298
anthirinumCAA66235.1	QNSAPAVP-----SLGPDGLDLLTKTLKYDPADRIISAKAALDHPYFD	300
M.truncatula_TC115190#1	QNLARAVP-----SLSPDGVDLLTKMLKYNPAERISAKAALDHPYFD	306
medicagoP93321	QNLARAVP-----SLSPDGVDLLTKMLKYNPAERISAKAALDHPYFD	306
lotus	QTLARAVP-----SLGPEGVDLLSKMLKYNPADRIISAKAALDHPYFD	316
G.max_Glyma07g02400.1#1	QSLAKNVP-----SLGPDGVDLLSKMLKYNPSEIRISAKAALDHPYFD	309
P.patens_NP13120379#1	HDIAQAVP-----GIEPSGVDLLDRMLQYNPANRISAKEALIHYPYFD	297
P.patens_TC47562#1	HEIAQAVP-----RVERSGVDLLDRLLQYNPANRISAKEALVHPYFD	297
P.patens_NP13127560#1	QNLGRAVP-----DMGPEALDLLTRMLMFDPAKRVSAKAALNHPFFD	297
P.patens_TC43373#1	KDLSLAVP-----DMSPDALDLSRMLVFDPAKRISAKAALHHPFFD	298
P.patens_TC41028#1	QDLSLAVP-----GLSAVGLDLLAKMLVFEPKRIISAKAALSHTYFAD	303
P.patens_TC35664#1	HRMKRFGP-----GLASI-----VIGTSSRSGGPKSYLWRFLVYA	265

FIGURE 3 (continued

A.thaliana_AT1G76540.1#1	PEKSSL-----	313
A.thaliana_AT1G20930.1#1	PKSSL-----	315
H.annuus_CD855729#1	-----	
G.max_Glyma14g39760.1#1	DKRHL-----	311
G.max_Glyma17g38210.1#1	DKRHL-----	314
G.max_Glyma07g07640.1#1	DKRNL-----	315
G.max_Glyma09g08250.1#1	DKGHL-----	317
M.truncatula_TC117212#1	DKTHL-----	316
P.trichocarpa_scaff_II.31#1	EKECL-----	321
S.lycopersicum_TC205060#1	DKTPL-----	315
H.vulgare_TC172640#1	-----	
T.aestivum_TC282781#1	DKALY-----	326
O.sativa_LOC_Os08g40170.1#1	NKELY-----	326
sorghumXP_002441781.1	DKSQF-----	308
Z.mays_c62195783gm030403@12303	DKSQF-----	308
O.sativa_LOC_Os01g67160.1#1	DKSQF-----	303
H.vulgare_TC160457#1	DKSQF-----	309
T.aestivum_TC283074#1	TVQSWRSTSGGFSHMGWLGPHFSAWEYV	324
T.aestivum_TC290551#1	DKSQF-----	304
SEQ ID NO: 318	-----	
A.thaliana_AT2G38620.1#1	-----	
A.thaliana_AT2G38620.2#1	DKSQF-----	311

FIGURE 3 (continued

A.thaliana_AT3G54180.1#1	DKSQF-----	309
ChenopodiumCAC17703.1	NAIVSY-----	317
P.trichocarpa_scaff_XVI.1376#1	DKSQF-----	302
P.trichocarpa_scaff_3712.1#1	DKSQF-----	308
P.trichocarpa_scaff_28.331#1	DKSQF-----	322
ricinusXP_002519974.1	DKSQF-----	316
vitisXP_002266623.1	DKSQF-----	303
S.lycopersicum_TC198540#1	DKSQF-----	303
tomatoCAC15503.1	DKSQF-----	303
tobaccoAAG01532.1	DKSQF-----	303
anthirinumCAA66235.1	DKSQF-----	305
M.truncatula_TC115190#1	DKSQY-----	311
medicagoP93321	DKSQY-----	311
lotus	DKSQY-----	321
G.max_Glyma07g02400.1#1	DKSQF-----	314
P.patens_NP13120379#1	DKSQY-----	302
P.patens_TC47562#1	DKSQF-----	302
P.patens_NP13127560#1	DKSQF-----	302
P.patens_TC43373#1	DKSQF-----	303
P.patens_TC41028#1	DKTAT-----	308
P.patens_TC35664#1	Q-----	266

FIGURE 3 (continued

19/42

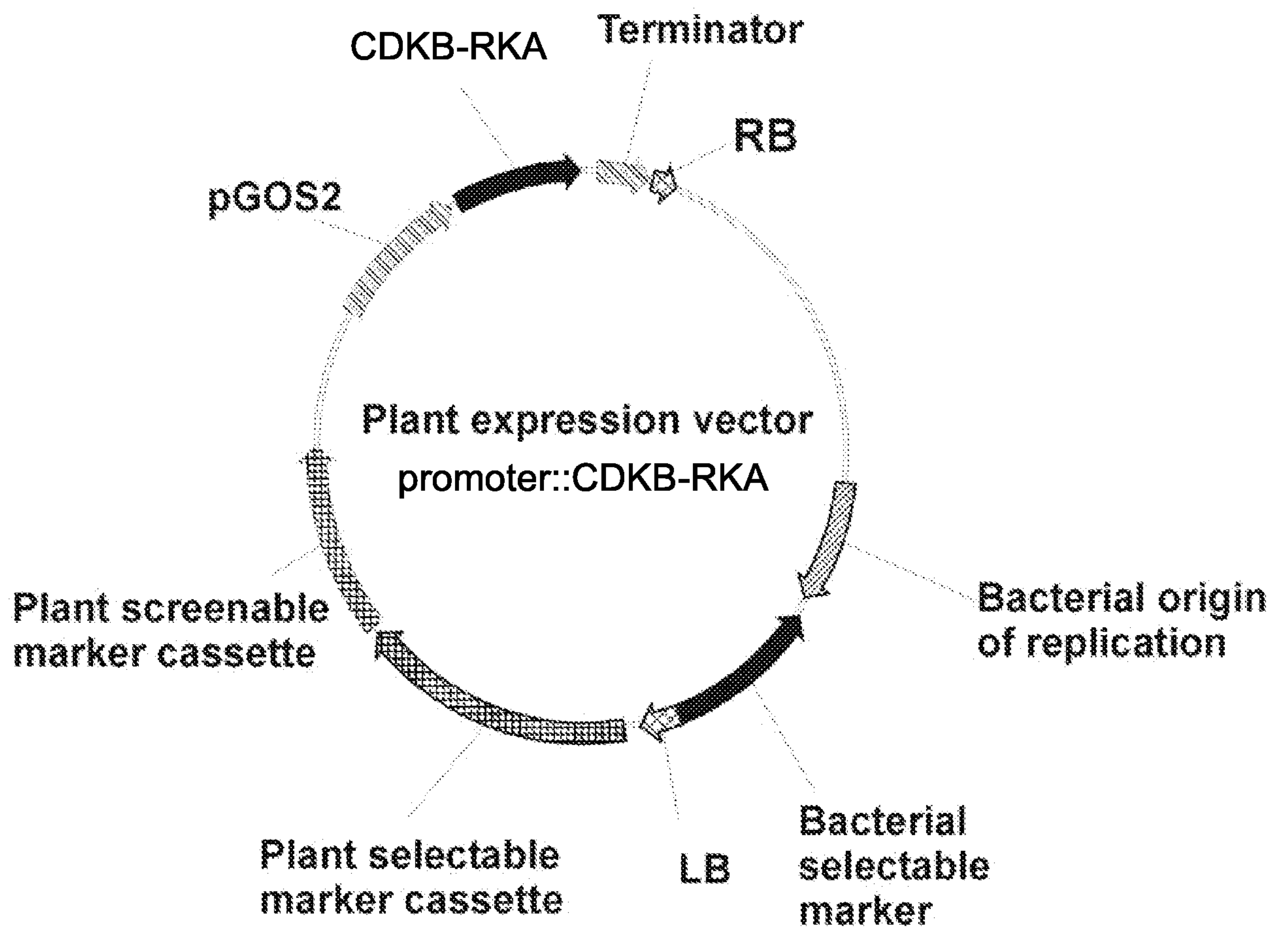


FIGURE 4

20/42

MAGFSYQHQPFLFPDSAFLPSIATPTKNMNNNMYGSFEEAGNMMNTNGFSQIYSPET

FHETPSLDVRFHQSSHPDDHSYKVSLSDNETSLTKKQSTNSSTVVDKLESEHVTQE

VTPMARKRKSANGFLNSAQSKDARKVKSKRQNKCSGDMKHEEKKPKVEKKVHGEP

-----3-----

-----5-----

AGYIHVRA**RRGQATDSHSLAERVRRERISERMKILQLLVP****GCDKITGKALMLDEII**

-----5-----

-----4-----

-----1-----

--2--

NYVQSLQNQVEFLSMKLASVNPLLYDFGMDRDAFMVRPERLSSMSPPPLPSLQHNSP

-----6-----

-----4-----

-----2-----

IQPTAFADTASATTATFATEENNYPLIDNSATLFLQGMRPSTFTTHQDSGYLMWDV

DEQRQKFLNPSGLATNLCSFH

FIGURE 5

21/42

	1	50
A.lyrata 472791	(1)	-----MDEFAAKKK-----DLVNTP-----VDLYPPENPMSG--
A.thaliana AT1G26260.1	(1)	-----MSDKDEFAAKKK-----DLVNTP-----VDLYPPENPMLG--
A.lyrata 476021	(1)	-----MDLSEKDEFSAEKRNPdNYDSVNNPSGDWRVDSYPSENLI SAG-
A.thaliana AT1G68920.1	(1)	-----MDLSAKDEFSAEKRNPdNYDSVNNPSGDWRVDSYPSENLI SAG-
G.max Glyma01g09400.1	(1)	-----MGDKKEKFELDRNE----DPMSYSSGMHSDWRFGG-----
G.max Glyma02g13860.1	(1)	-----MSDKKEKFEVDRNE----DHVSYS SGMHSDWRFGGSNLANSSV
M.truncatula AC150842 12.5	(1)	-----MSESENF EVDRKE----GTLDWRFVSGNLANSSVG--LVSME
O.sativa LOC Os06g16400.1	(1)	-----
A.lyrata 491205	(1)	-----
Aquilegia sp TC24120	(1)	-----
R.communis TA2848 3988	(1)	-----
S.lycopersicum TC196509	(1)	-----
H.tuberosus TA2737 4233	(1)	-----
V.vinifera GSVIVT00014998001	(1)	-----
Z.officinale TA5989 94328	(1)	-----
A.lyrata 494918	(1)	MDNELFMNTEFPPPEMATHFEP-PPSSSSAMMLNWALMDPNPPQDSSFI
A.thaliana AT5G48560.1	(1)	MDNELFMNTEFPPPEMATHFEH-QQSSSSAMMLNWALMDPNPHQDSSFL
A.thaliana AT3G07340.1	(1)	MENELFMNAGVSHPPVMTSPS-----SSSAMLKWVSMETQP-VDPSLS
Aquilegia sp TC26724	(1)	-----
G.hirsutum TC144761	(1)	-----
G.raimondii TC9358	(1)	-----
P.persica TC11169	(1)	-----MGKDR TATPSWNPSSFGMEMQSNE LNCGSQQLPNSL
P.trichocarpa 572918	(1)	-----MERDKLFVSEGANTAATIWN SC SFGMEMQANELSCGPEKLANCF
Poptr TA1	(1)	-----MERDKLFVSEGANTAATIWN SC SFGMEMQANELSCGPEKLANCF
G.max Glyma03g21770.1	(1)	MENQFFLN SGVSQTQH HHPLHLESSPPSNPSSVPSWQSLSPPNMGIQPTV
G.max GM06MC18997 59691073018643	(1)	-----
P.vulgaris TC12782	(1)	-----
G.max Glyma08g46040.1	(1)	MENQFFLNAGVPPQLHFEPSPH--HHHHHLCPPPSWQSSLSSAMDVQVTV
G.max Glyma18g32560.1	(1)	MENQFFLNAGVPPQLHFEP S----PHGHQPCPPPSWQSSLSSAMDVQVTV
M.truncatula AC171266 17.4	(1)	MENQFFLNAGFEPYQTQDCFFN-----PNWDKSTDQNLN-----
P.trichocarpa 553223	(1)	MESEYFFNAGVPPQALLFEPTS-----TSSSMPVWQSLSSPMEMQDNCS
Triphysaria sp TC7138	(1)	-----
S.lycopersicum TC215759	(1)	MDNDCF SNGGIQPPFHFD RKIP-----LNSLHSPQEQNSDYF
S.tuberosum TC168229	(1)	-----
C.annuum TC15396	(1)	-----
O.sativa LOC Os08g41320.1	(1)	-----
O.sativa LOC Os09g32510.4	(1)	-----
A.lyrata 495154	(1)	-----
A.thaliana AT5G50915.1	(1)	-----
B.napus TC73386	(1)	-----
C.annuum NP13055023	(1)	-----
N.tabacum FG645212	(1)	-----
P.hybrida TC3261	(1)	-----
G.max Glyma04g37690.1	(1)	-----
G.max Glyma06g17420.1	(1)	-----
G.max Glyma05g38450.1	(1)	-----
G.raimondii TC6315	(1)	-----
P.trichocarpa 575437	(1)	-----
P.trichocarpa 773249	(1)	-----
M.domestica TC28668	(1)	-----
C.sinensis TC686	(1)	-----
O.sativa LOC Os05g01256.1	(1)	-----
O.sativa LOC Os08g42470.1	(1)	-----
P.virgatum TC28308	(1)	-----
S.bicolor Sb07g025920.1	(1)	-----
Z.mays_ZM07MC24807 BFb0298G14024734	(1)	-----
Z.mays TC476562	(1)	-----
O.sativa LOC Os09g33580.1	(1)	-----
T.aestivum TC301717	(1)	-----
S.bicolor Sb02g029530.1	(1)	-----
Z.mays TC511518	(1)	-----
Z.mays_ZM07MC32776 BFb0337M22032678	(1)	-----
S.officinarum TC78398	(1)	-----
Z.mays TC460846	(1)	-----
Aquilegia sp TC24680	(1)	-----
O.sativa LOC Os04g28280.1	(1)	-----
Consensus	(1)	

FIGURE 6

22/42

	51	100
A.lyrata 472791	(28)	-----PSQMMDSFRET LWND-----
A.thaliana AT1G26260.1	(31)	-----PSPMDSFRET LWHD-----
A.lyrata 476021	(44)	-----PAS-----CSPSQMMDSFGQTLWYDPTSVQ-AMGY
A.thaliana AT1G68920.1	(44)	-----PAS-----CSPSQMMDSFGQTLWYDPTSVQ-AVGY
G.max Glyma01g09400.1	(31)	-DLIGSSS-----CSSASMVDSFGPNYWENPTCSQ-NLGF
G.max Glyma02g13860.1	(39)	GDLIGSSS-----CSSASMVDSLSPNYWENPTSSQ-KLGF
M.truncatula AC150842 12.5	(37)	NSMMGTSS-----CP-NSMVDSYGPNFLDLLTNSENLGGF
O.sativa LOC Os06g16400.1	(1)	-----
A.lyrata 491205	(1)	-----
Aquilegia sp TC24120	(1)	-----
R.communis TA2848 3988	(1)	-----
S.lycopersicum TC196509	(1)	-----
H.tuberosus TA2737 4233	(1)	-----
V.vinifera GSVIVT00014998001	(1)	-----
Z.officinale TA5989 94328	(1)	-----
A.lyrata 494918	(50)	WEKSTEQQ-----Q-QQSIFDSALSSLVSSPTPS----NS
A.thaliana AT5G48560.1	(50)	WEKSTEQQ-----Q-QQSIFDSALSSLVSSPTPS----NS
A.thaliana AT3G07340.1	(43)	RNLFWEKS-----T-EQSIFDSALSSLVSSPTPS----NS
Aquilegia sp TC26724	(1)	-----
G.hirsutum TC144761	(1)	-----
G.raimondii TC9358	(1)	-----
P.persica TC11169	(37)	FNANWDNS-----MDQSDPFESALSSIVSSPAASNAAIAA
P.trichocarpa 572918	(45)	LNP NWDNS-----LDQSDPFESALSSIVSSPVASGANANA
Poptr TA1	(45)	LNP NWDNS-----LDQSDPFESALSSIVSSPVASGANANA
G.max Glyma03g21770.1	(51)	MNYQGLHH-----FDSSALSSMVSSPAASSNP-NN
G.max GM06MC18997 59691073018643	(1)	-----
P.vulgaris TC12782	(1)	-----
G.max Glyma08g46040.1	(49)	LNCSTEQNPDCFYNNPTWDKSTDH HGLQFDSALSSMVSSPAAS----NS
G.max Glyma18g32560.1	(47)	LNCSTEQTQDCFYNTTPTWDKSTDHG-LQFDSALSSMVSSPAAS----NS
M.truncatula AC171266 17.4	(35)	-----YDSTLSSIVSSPAATSSPSNP
P.trichocarpa 553223	(45)	ARRLFNSN-----WEKSTDHSPHFESSLSSMVSSPGVS----NC
Triphysaria sp TC7138	(1)	-----
S.lycopersicum TC215759	(38)	LNTHWDNN-----STDNQYTRFDSALSSIVSSPVPS-----
S.tuberosum TC168229	(1)	-----
C.annuum TC15396	(1)	-----
O.sativa LOC Os08g41320.1	(1)	-----
O.sativa LOC Os09g32510.4	(1)	-----
A.lyrata 495154	(1)	-----
A.thaliana AT5G50915.1	(1)	-----
B.napus TC73386	(1)	-----
C.annuum NP13055023	(1)	-----
N.tabacum FG645212	(1)	-----
P.hybrida TC3261	(1)	-----
G.max Glyma04g37690.1	(1)	-----
G.max Glyma06g17420.1	(1)	-----
G.max Glyma05g38450.1	(1)	-----
G.raimondii TC6315	(1)	-----
P.trichocarpa 575437	(1)	-----
P.trichocarpa 773249	(1)	-----
M.domestica TC28668	(1)	-----
C.sinensis TC686	(1)	-----
O.sativa LOC Os05g01256.1	(1)	-----
O.sativa LOC Os08g42470.1	(1)	-----
P.virgatum TC28308	(1)	-----
S.bicolor Sb07g025920.1	(1)	-----
Z.mays_ZM07MC24807 BFb0298G14024734	(1)	-----
Z.mays TC476562	(1)	-----
O.sativa LOC Os09g33580.1	(1)	-----
T.aestivum TC301717	(1)	-----
S.bicolor Sb02g029530.1	(1)	-----
Z.mays TC511518	(1)	-----
Z.mays_ZM07MC32776 BFb0337M22032678	(1)	-----
S.officinarum TC78398	(1)	-----
Z.mays TC460846	(1)	-----
Aquilegia sp TC24680	(1)	-----
O.sativa LOC Os04g28280.1	(1)	-----
Consensus	(51)	-----

FIGURE 6 (continued)

	101	150
A.lyrata 472791	(43) GGFNVHTDAGTSFR-----	
A.thaliana AT1G26260.1	(46) GGFNVHTDADTSFRG-----	
A.lyrata 476021	(73) GGFNGGNASSSSFRG-----	
A.thaliana AT1G68920.1	(73) AGFNGGNASSSSFRG-----	
G.max Glyma01g09400.1	(64) CDINVHNNGGSSSTVAIRKDGFGFGRAGQ-----	
G.max Glyma02g13860.1	(73) CDIN-----N-----	
M.truncatula AC150842 12.5	(71) CDVN----GQSNE---VRKDGFGFARVG-----	
O.sativa LOC Os06g16400.1	(1) -----	
A.lyrata 491205	(1) -----	
Aquilegia sp TC24120	(1) -----	
R.communis TA2848 3988	(1) -----	-----MN
S.lycopersicum TC196509	(1) -----	-----M
H.tuberosus TA2737 4233	(1) -----	
V.vinifera GSVIVT00014998001	(1) -----	
Z.officinale TA5989 94328	(1) -----	
A.lyrata 494918	(80) NFSGG---GGDGFLIRELIGKLG NIGNNNNN-----SG-----EIY	
A.thaliana AT5G48560.1	(80) NFSGG---GGDGFLIRELIGKLG NIGNNNNN-----SG-----EIY	
A.thaliana AT3G07340.1	(73) NFSVGG-VGGENVIMRELIGKLG NIG-----DIY	
Aquilegia sp TC26724	(1) -----	
G.hirsutum TC144761	(1) -----	
G.raimondii TC9358	(1) -----	
P.persica TC11169	(72) GKG-----GGDGEMIRELIGRLGSICNSGE-----ISSHSYMC	
P.trichocarpa 572918	(80) NAIPNAGVGGDSL MIRELIGRLGNICNSGD-----ISLQSFVN	
Poptr TA1	(80) NAIPNAGVGGDSL MIRELIGRLGNICNSGD-----ISLQSFVN	
G.max Glyma03g21770.1	(80) NMSNE-----NFIIRELMGKLGAIG--NSDEIPQHSPHPLVVASSYMNT	
G.max GM06MC18997 59691073018643	(1) -----	-----MNS
P.vulgaris TC12782	(1) -----	
G.max Glyma08g46040.1	(95) NMSSE-----NFVIRELIGKLG NIGAGSDEIEISPHSQPLVGASSYINC	
G.max Glyma18g32560.1	(92) NISSE-----NFVIRELIGKLG NIGGGGSDEISPHSQPLVGASFYINC	
M.truncatula AC171266 17.4	(56) NMSNE-----NFVMRELIGKLGTFG-----	
P.trichocarpa 553223	(80) NVSSE-----SFMVRELIGNLGNIDNSNNPGEISPHSQPMLAASYITAA	
Triphysaria sp TC7138	(1) -----	
S.lycopersicum TC215759	(69) NSVNS-----NSSLCELIGKLG SICTSPATPFTSN-----	
S.tuberosum TC168229	(1) -----	
C.annuum TC15396	(1) -----	
O.sativa LOC Os08g41320.1	(1) -----	
O.sativa LOC Os09g32510.4	(1) -----	
A.lyrata 495154	(1) -----	
A.thaliana AT5G50915.1	(1) -----	
B.napus TC73386	(1) -----	
C.annuum NP13055023	(1) -----	
N.tabacum FG645212	(1) -----	
P.hybrida TC3261	(1) -----	
G.max Glyma04g37690.1	(1) -----	
G.max Glyma06g17420.1	(1) -----	
G.max Glyma05g38450.1	(1) -----	
G.raimondii TC6315	(1) -----	
P.trichocarpa 575437	(1) -----	
P.trichocarpa 773249	(1) -----	
M.domestica TC28668	(1) -----	
C.sinensis TC686	(1) -----	
O.sativa LOC Os05g01256.1	(1) -----	
O.sativa LOC Os08g42470.1	(1) -----	
P.virgatum TC28308	(1) -----	
S.bicolor Sb07g025920.1	(1) -----	
Z.mays_ZM07MC24807 BFb0298G14024734	(1) -----	
Z.mays TC476562	(1) -----	
O.sativa LOC Os09g33580.1	(1) -----	
T.aestivum TC301717	(1) -----	
S.bicolor Sb02g029530.1	(1) -----	
Z.mays TC511518	(1) -----	
Z.mays_ZM07MC32776 BFb0337M22032678	(1) -----	
S.officinarum TC78398	(1) -----	
Z.mays TC460846	(1) -----	
Aquilegia sp TC24680	(1) -----	
O.sativa LOC Os04g28280.1	(1) -----	
Consensus	(101)	

FIGURE 6 (continued)

	151	200
A.lyrata 472791	(57)	-----
A.thaliana AT1G26260.1	(61)	-NNNIDIPL-----EMGW
A.lyrata 476021	(88)	-NIDRSLEMGWNLPLNLLPPKGN-----GLFLPNASTFLPP
A.thaliana AT1G68920.1	(88)	-SIDRSLEMGWNLPLNLLPPKGN-----GLFLPNASSFLPP
G.max Glyma01g09400.1	(93)	-DHHETLEMGWNPANS-----ILSNGPGMFPH
G.max Glyma02g13860.1	(78)	-DHHGTLEMGWNHANS-----MLPNGPVMFPH
M.truncatula AC150842.12.5	(92)	-CDDRTVGMGWDLGSSSMIKRD-----G-VLPNGAEIFPQ
O.sativa LOC_Os06g16400.1	(1)	--MNEKDATDLEERSEASEHGQALSFGGAMFLQEAQIASPAAANNALTS
A.lyrata 491205	(1)	-----
Aquilegia sp TC24120	(1)	-----MLHCAADTMTVLERQQARQKWQQVPLLQQQQQQQQHQQSYFSDT
R.communis TA2848.3988	(3)	RALLPEMLHCTD--LTVLERQRLKQQE-QQHLQLEQHQQHHQGENY
S.lycopersicum TC196509	(2)	NIALPEMLHNITSNGSSELSVLDRTKWQVQ-----QQEMSYPNGQNDQ
H.tuberosus TA2737.4233	(1)	-----MNPNG
V.vinifera GSVIVT00014998001	(1)	MGIDDNNGNMGFPNTSQSILNCPSSGMNTHPISEKVTGMTMSSASMYKSSN
Z.officinale TA5989.94328	(1)	-----
A.lyrata 494918	(113)	GTPMSRSASCYATPMSSPPPP-----T-----NTNSQMMMNRRTT
A.thaliana AT5G48560.1	(113)	GTPMSRSASCYATPMSSPPPP-----T-----NSNSQMMMNRRTT
A.thaliana AT3G07340.1	(101)	GITASNGNSCYATPMSSPPPP-----GSMMETKTTT
Aquilegia sp TC26724	(1)	-----
G.hirsutum TC144761	(1)	-----
G.raimondii TC9358	(1)	-----
P.persica TC11169	(105)	GNN-STNTSCYSTPLNSSPKLN-----LSMIDPQMRGNLPIPGNHLPSHP
P.trichocarpa 572918	(118)	NNNNSTNTSCYSTPMNSPPKLN-----LSMDSQMRGNLPIPGNSVVKHP
Poptr TA1	(118)	NNNNSTNTSCYSTPMNSPPKLN-----LSMDSQMRGNLPIPGNSVVKHP
G.max Glyma03g21770.1	(122)	KGNNSTNTSCYSTPLSSPPKVNIVNSLVNRLANLGGGKS---TTTVLNS
G.max GM06MC18997.59691073018643	(4)	NGNNSTYTSCYSTPLSSPPKVNIVHSLVNERLANLGGGK---STMALNS
P.vulgaris TC12782	(1)	-----
G.max Glyma08g46040.1	(139)	N--NNSTNTSCYSTPLSSSPKVNMNKIPTMVKHLVKEGMPLSLGTSTSLNS
G.max Glyma18g32560.1	(136)	N--NNSTNTSCYSTPLSSPPKVNTIKIPTMVNHLVKEGMPPSLETPLMSLNS
M.truncatula AC171266.17.4	(76)	-----EISQHSNP-----LVLPMMS
P.trichocarpa 553223	(124)	N--NSANTSCYTTPLNSPPKLN--MPVMDQFSKEHLNIPSLGKPMGLNS
Triphysaria sp TC7138	(1)	-----
S.lycopersicum TC215759	(99)	--CDSTRTSCYTTPMSSPPK-----LHIPIMNQIGKDKVPNLG
S.tuberosum TC168229	(1)	-----
C.annuum TC15396	(1)	-----
O.sativa LOC_Os08g41320.1	(1)	MDGGRVGGDYISSLLSSSPRLD-----FGVPVLDAIVAPGGGGGGGDC
O.sativa LOC_Os09g32510.4	(1)	-----MDSYVAGLLMS-----AAAAGLDLGVLDDGGG---A
A.lyrata 495154	(1)	-----MATFSYFQNYPHS
A.thaliana AT5G50915.1	(1)	-----MATFSYFQNYPHS
B.napus TC73386	(1)	-----MASFSYFQHYPHS
C.annuum NP13055023	(1)	-----MSTFSSYQLQQHNTPS
N.tabacum FG645212	(1)	-----
P.hybrida TC3261	(1)	-----
G.max Glyma04g37690.1	(1)	-----MAAFSYQYHPF
G.max Glyma06g17420.1	(1)	-----MAAFSYQYHPF
G.max Glyma05g38450.1	(1)	-----MAAFSSQYHSHF
G.raimondii TC6315	(1)	-----MATYSFQSHPF
P.trichocarpa 575437	(1)	-----MAGFSYQHQPPL
P.trichocarpa 773249	(1)	-----MAAFSYQHPPPL
M.domestica TC28668	(1)	-----MAAFSYQNLHHLS
C.sinensis TC686	(1)	-----MAAFSYQQHHHPF
O.sativa LOC_Os05g01256.1	(1)	-----MESLCGQLEHE
O.sativa LOC_Os08g42470.1	(1)	-----MADFS--PHHS
P.virgatum TC28308	(1)	-----MADFS--SHHS
S.bicolor Sb07g025920.1	(1)	-----MADFS--QNDS
Z.mays_ZM07MC24807_BF50298G14024734	(1)	-----MAGFSR-PNNS
Z.mays TC476562	(1)	-----MADADLTPNHS
O.sativa LOC_Os09g33580.1	(1)	-----
T.aestivum TC301717	(1)	-----
S.bicolor Sb02g029530.1	(1)	-----
Z.mays TC511518	(1)	-----
Z.mays_ZM07MC32776_BFb0337M22032678	(1)	-----
S.officinarum TC78398	(1)	-----
Z.mays TC460846	(1)	-----
Aquilegia sp TC24680	(1)	-----MAAFSYQ
O.sativa LOC_Os04g28280.1	(1)	-----
Consensus	(151)	

FIGURE 6 (continued)

25/42

	201	250
A.lyrata 472791	(57)	-----ESAGLFLQDTMI
A.thaliana AT1G26260.1	(73)	NMAQFPADSGFIERAAKFSFFGCGEMMMNQSSSLGVPDSTGLFLQDTQI
A.lyrata 476021	(122)	SMAQFPADSGFIERAARFSLFSGGN-FSDMVNQPLGNPESIGLFLQGGGT
A.thaliana AT1G68920.1	(122)	SMAQFPADSGFIERAARFSLFSGGN-FSDMVNQPLGNSEAIGLFLQGGGT
G.max Glyma01g09400.1	(119)	TLSQFPDTSDFIERAARFSCFSGGN-FSDMVN-SYGIAQSMGLYGARDAI
G.max Glyma02g13860.1	(104)	SLSQFPDTSDFIERAARFSCFSGGN-FGDMVN-SYGIAQSMGLYGARDAI
M.truncatula AC150842.12.5	(125)	SLSQFSTDSGFVD-AARMSCFSAGS-FVDMMN-SCGIPQSMALP-----
O.sativa LOC_Os06g16400.1	(49)	MANPFPIPPGLWNPPSHNMGLGETSFSSLLGMLSAGAPPPFVATPGFVDS
A.lyrata 491205	(1)	----MNGAIGDILLNCTDMSVLERQRAHLKYFNPSFDSPLAGFFADSSM
Aquilegia sp TC24120	(45)	HFNLLPNSSSSSQFHIMNDDSVI--RDFVNRSIKLDP-SLDNNGWQDF
R.communis TA2848.3988	(50)	FLGDLSGVFQLOQQQQQQQGFQGD--GEVVIRSVKPDPGFLDNGCWSNT
S.lycopersicum TC196509	(45)	LMNSFHQTAEAAQFHGLINVNDQSL--NELVTRAIKPD--CMENSWGGE
H.tuberosus TA2737.4233	(5)	FMQQQEAFFSVNDHHHLQDMFTLGMVDKDSGLNIVKPDPEMEECWVGFDA
V.vinifera GSVIVT00014998001	(51)	GGDPFFGSGWDPIVSLSQNENFGGSSMVSHSEFANSAYPVVLENQGIGST
Z.officinale TA5989.94328	(1)	-----
A.lyrata 494918	(147)	PLTEFSADPGFAERASRFSCFGSRS--FNGR-----TNTNLPINNGNNIV
A.thaliana AT5G48560.1	(147)	PLTEFSADPGFAERAARFSCFGSRS--FNGR-----TNTNLPINNGNNMV
A.thaliana AT3G07340.1	(131)	PMAELSGDPGFAERAARFSCFGSRS--FNSR-----TNSPPFINN-EPPI
Aquilegia sp TC26724	(1)	-----
G.hirsutum TC144761	(1)	-----
G.raimondii TC9358	(1)	-----
P.persica TC11169	(149)	SLAPFQADPGFVERAARFSCFGGNN--FGGLNGQVNLNEAELAYRSMPKI
P.trichocarpa 572918	(163)	GLAPFPAD--FVERAARYSCFGSNN--PGGINKQFGLNESELINRLMPRV
Poptr TA1	(163)	GLAPFPAD--FVERAARYSCFGSNN--PGGINKQFGLNESELINRLMPRV
G.max Glyma03g21770.1	(169)	SVAEFSADPGFAERAARFSCFGSRS--FNDRSVQLRVNNAELAQRSAPAM
G.max GM06MC18997.59691073018643	(50)	SVAEFSADPGFAERAARFSCFGSRS--FNDRSVQLGVNNAELAQRSAPAM
P.vulgaris TC12782	(1)	-----
G.max Glyma08g46040.1	(188)	TVAEFSADPGFAERAARLSCFGSRS--FNGRTTQLCLNIAELAQRSSPLV
G.max Glyma18g32560.1	(185)	TVAKFSADPGFAERAARFSCFGSRS--LNGRTTQLGLNNAELAQRSS-LV
M.truncatula AC171266.17.4	(92)	RVAEFSADPGFVQRAAKFSCFGSKS--FNERGNQMVMNNVELAQRSHNLM
P.trichocarpa 553223	(169)	SVAEFTADPGFAERAARFSCFGSRS--FNGRISQLGLNNAEMANGCNPLM
Triphysaria sp TC7138	(1)	-----MGNGKLSRVSSSP--VLKQDGLSKKQMEIPRPGNSNYN
S.lycopersicum TC215759	(135)	NSAVMNSPPFPSLSAARFSCFGSRS--FNGRTSQFELNNEDSRYGSGTGV
S.tuberosum TC168229	(1)	-----
C.annuum TC15396	(1)	-----MDPQTSLINHAGGFQSPPPPPF
O.sativa LOC_Os08g41320.1	(44)	GLDKLCGDPGFAERAARLSSFNNGG--GGVGQRYGGAGAGLFGMPPPAPG
O.sativa LOC_Os09g32510.4	(30)	FLETLCGGPGFAERAARLCGG-----GAGLFGLPVAVGNA
A.lyrata 495154	(14)	LLDPLLFP-----T-PHPSINLTGFIDQNHLHPLP
A.thaliana AT5G50915.1	(14)	LLDPLLFP-----T-PHSSINLTGFIDQNHLHPLP
B.napus TC73386	(14)	LLDPLLFS-----S-PNSSTKLSGFIDQNPLYPPP
C.annuum NP13055023	(17)	LLDSVFLPSSPLKMSGFFEE---PNNSCVVQNCFSQFYQPEFPTNNVNVN
N.tabacum FG645212	(1)	-----
P.hybrida TC3261	(1)	-----
G.max Glyma04g37690.1	(12)	LVDPAFF---PNNLNTN-----SPPP---HQFQQLPHETSSFNVT
G.max Glyma06g17420.1	(12)	LVDPAFF---PNNLNTN-----SPPPPLPSQFQQLPHETSSFNVT
G.max Glyma05g38450.1	(13)	VVDSACFSITPPIINTSS-----NFLPSPHYLHNNNIQETSYSVTD
G.raimondii TC6315	(12)	LLDSIFLPNTSTTKVSG-----LVEEGNIINPNCFPQFYQDSPVGSML
P.trichocarpa 575437	(12)	FPDSAFLPSIATPTKNMNNMYGSFEEAGNMMNTNGFSQIYSPETFHETP
P.trichocarpa 773249	(12)	FLDSVILPNIITPIMNMNNSMYWFYDEAG-GINTNSFYQVYPPETFHEAP
M.domestica TC28668	(14)	LLDSIFLPNAPTNTSINTDS---NFSQSFP--SEPLTQEIQADNNSRAV
C.sinensis TC686	(14)	LLDSSEFIEEGSFQLYYP---PKFPTHDDHQEMIVVNNNSNNLLDAAA
O.sativa LOC_Os05g01256.1	(12)	HEQQQLHQSSSSSLLPYCFLSAPEPLVPVPGFPAGCSNLEEKAAAAAMAAYEY
O.sativa LOC_Os08g42470.1	(10)	LLLKATAAG-----A-AIATTNDPNISSFFLYNHS-HGSQ
P.virgatum TC28308	(10)	LLLKMPA-----ALTNGHSHNLSSLLFYGQN-HGQG
S.bicolor Sb07g025920.1	(10)	LLLKMP-----S---HSPNFSSLLFYGHN-HGEA
Z.mays_ZM07MC24807_BF50298G14024734	(11)	LLLKMPA-----TLASSHTPNLSSLLFYDQN-HGQA
Z.mays TC476562	(12)	LLLKMPA-----ASAHASPSLSLLLYGQESHGEA
O.sativa LOC_Os09g33580.1	(1)	-----MAGFSHLPQQMEHGLITNNGFLFC
T.aestivum TC301717	(1)	-----MAGFSHLPQQMEYGLANAGSFLLC
S.bicolor Sb02g029530.1	(1)	-----MASFPHF-TQMEHGLAGASAHGTP
Z.mays TC511518	(1)	-----MASFPHL-TQMEHGLASGASHGTP
Z.mays_ZM07MC32776_BFb0337M22032678	(1)	-----MASFPHL-TQMEHGLASGASHGTP
S.officinarum TC78398	(1)	-----MASFPHL-TQMEHGLAG-ASHGTP
Z.mays TC460846	(1)	-----M-ASLFPQL-AQMEHGVAG-ASHGAP
Aquilegia sp TC24680	(8)	HLSPFLDLVLLPNSSPIKMSGLLLEEGYIDNISTCFPYSSLSEHFQEFPI
O.sativa LOC_Os04g28280.1	(1)	-----
Consensus	(201)	L

FIGURE 6 (continued)

	251	300
A.lyrata 472791	(69) P-----SGSKLDN-----GSLKDASKLVKD	
A.thaliana AT1G26260.1	(123) P-----SGSKLDN-----GPLTDASKLVKE	
A.lyrata 476021	(171) M-----QGQCQCDELNVGEPHNDVSTAVKD	
A.thaliana AT1G68920.1	(171) M-----QGQCQSNEELNVGEPHNDVSVAVKE	
G.max Glyma01g09400.1	(167) AGHGMKSVTGG-QSQGGDMNVVEATKDVSPSVEHLVAAKGSPLKSDRRSE	
G.max Glyma02g13860.1	(152) AGHGLKSVIAGGQSQGGDMNVVE---DVPPSVEHLVAAKGSPLKSDRRSE	
M.truncatula AC150842.12.5	(166) -----LHVSRSVEHLGSD-GSPIQNDRRS	
O.sativa LOC_Os06g16400.1	(99) TAGFPCYNGGN-----LGAMINHPFPGIHQPLGDFQNGV	
A.lyrata 491205	(47) INGGE-----MDGYLATAGLNLPMIYGETTVEG-----	
Aquilegia sp TC24120	(92) GKYPVSGN-----GFGVNGVDLVDYSLRTSSCPP-----	
R.communis TA2848.3988	(98) SSTDLVG-----YGPCGFGNMNFAISRTSSCPPT----	
S.lycopersicum TC196509	(91) GTTGTNGFD-----YVPVGVGHHGGMSPSEMNYAISRTSSCPP----	
H.tuberosus TA2737.4233	(55) GAYGG-----QLGYGYVNHNTVLKVSNIHG-----	
V.vinifera GSVIVT00014998001	(101) PHLVLYPSNSS-----LVEMVPKLPFCGSGSFSEMVASFGLPECGQTANS	
Z.officinale TA5989.94328	(1) -----	
A.lyrata 494918	(190) NNSGK-----LTRVSSTPALKALVSPEVAP-----	
A.thaliana AT5G48560.1	(190) NNSGK-----LTRVSSTPALKALVSPEVTP-----	
A.thaliana AT3G07340.1	(173) TTNEK-----MPRVSSSPVFKPLASHVPAGE-----	
Aquilegia sp TC26724	(1) -----	
G.hirsutum TC144761	(1) -----MVTGSQVHVPDSNKNTLQDGNSCSDKRNRLS	
G.raimondii TC9358	(1) -----MVTGSQVHVPDSNKNTLQDGNSCSDKRNRLS	
P.persica TC11169	(197) DSGKLSRASSN---QSLKVAAGSQLGVQESNKSSPQGGNSAPDKKFGFRFS	
P.trichocarpa 572918	(209) EPGKLSRVSSN---NSMKVTVS-QANVQESNKSSPQDGLNSEKKFSRQS	
Poptr TA1	(209) EPGKLSRVSSN---NSMKVTVS-QANVQESNKSSPQDGLNSEKKFSRQS	
G.max Glyma03g21770.1	(217) EHGGK-----LPRVSSSPLLKTLG-----	
G.max GM06MC18997.59691073018643	(98) EHGGK-----LPRVSSSPLLKTLGSQMGAEQENKNSA	
P.vulgaris TC12782	(1) -----	
G.max Glyma08g46040.1	(236) ENGKKQ-----LPRVSSSPSLKVLGSQMGTQENKNSP	
G.max Glyma18g32560.1	(232) ENGKR-----LPRVSSSPSLKVLESQMGTQENKNSP	
M.truncatula AC171266.17.4	(140) ENGMK-----LSRVSSSPSLKTFGSQMVNHNENKNSS	
P.trichocarpa 553223	(217) GNGKL-----ARVASSPLLKAVGSQKGNKSTPLQD	
Triphysaria sp TC7138	(37) NNNNG-----VVLGSEMFKFSKLSGS-----NGN--	
S.lycopersicum TC215759	(183) MGIGN-----LTRISSSPCVVQNK-----NSSLM	
S.tuberosum TC168229	(1) -----	
C.annuum TC15396	(23) NLADIWQFPGN-----AGGNATAYSFPVSAS-----	
O.sativa LOC_Os08g41320.1	(92) DFAGG-----GSREASSVSDPASSAMKDAA-----	
O.sativa LOC_Os09g32510.4	(64) ERGG-----CSREGSSVSDPAWAHATGGG-----	
A.lyrata 495154	(43) NISTVEDISFN-----PFLDEYNVDRPEN-----	
A.thaliana AT5G50915.1	(43) NISTVEDISF-----LEYNVDKTEN-----	
B.napus TC73386	(43) NISTIVDTSLN-----PFLDSFNVEKTES-----	
C.annuum NP13055023	(64) ENSYCLDQNTN-----VTNKSSSSISLDMSSS-----	
N.tabacum FG645212	(1) -----	
P.hybrida TC3261	(1) -----	
G.max Glyma04g37690.1	(46) NQETSCVDQSS-----KITISDNESPSVTENLS-----	
G.max Glyma06g17420.1	(50) NQETSCVDQSS-----KITISDNESPSVTKNLS-----	
G.max Glyma05g38450.1	(54) NKETSCFEQSS-----KVTISDIELSVVKNHS-----	
G.raimondii TC6315	(55) HQSSCLDHSSN-----EPSTVTKKQSTE-----	
P.trichocarpa 575437	(62) SLDVRFHQSSHP-----DDHSYKVSLSDNETSLTKKQ--S-----	
P.trichocarpa 773249	(61) -LDVRFHEFSH-----HDHSSKVSLSDNETSLTKKQ--S-----	
M.domestica TC28668	(59) ESSCTMDHSST-----KVAFSDNENDPSVTKNKSTE-----	
C.sinensis TC686	(61) VFETSCGDNHDH-----HSDDNCNEPSNNNSLATKNQS-T-----	
O.sativa LOC_Os05g01256.1	(62) ESSS-----CSSLDPTSMMPVYSPIVLQ-----	
O.sativa LOC_Os08g42470.1	(43) APQPAN---AA-----AAAIVEDASLESSVS-----	
P.virgatum TC28308	(40) ATANANAASGT-----AAAMAEDASLESSS-----	
S.bicolor Sb07g025920.1	(35) AAP-----AN-----ANAMVEDASLESSS-----	
Z.mays_ZM07MC24807_BF50298G14024734	(41) VNAN-----AA-----AMMVEDDASLQSSS-----	
Z.mays TC476562	(43) APANAN-ANAA-----AATVEDAASLETSSA-----	
O.sativa LOC_Os09g33580.1	(25) HGSHGGAATTT-----APAIPEDASMETSS-----	
T.aestivum TC301717	(25) HGNP-----ED-----AAASLEGSSGVLDT-----	
S.bicolor Sb02g029530.1	(24) SFLFCH-----GAATADSASLETSS-----	
Z.mays TC511518	(24) SFFFCH-----GTATADSASLETSS-----	
Z.mays_ZM07MC32776_BFb0337M22032678	(24) SFFFCH-----GTATADSASLETSS-----	
S.officinarum TC78398	(23) SFLFCH-----GAAMADSASLETSS-----	
Z.mays TC460846	(24) SFLFCNGA-----GAATADSASLETSS-----	
Aquilegia sp TC24680	(58) EKQR-----TDSSSIVVDLEGSTS-----	
O.sativa LOC_Os04g28280.1	(1) -----	
Consensus	(251)	L

FIGURE 6 (continued)

		301		350
A.lyrata 472791	(89)	RSIN-----VSEDSQSSGCNG-----HDDAKCGV--		
A.thaliana AT1G26260.1	(143)	RSINN-----VSEDSQSSGGNG-----HDDAKCGQ--		
A.lyrata 476021	(196)	PTVRSCEQAKPNVPGSGNVSEDTQSSGGNGRKGRETSNTKKRKRNGQ--		
A.thaliana AT1G68920.1	(196)	STVRSSEQAKPNVPGSGNVSEDTQSSGGNGQKGRETSNTKKRKRNGQ--		
G.max Glyma01g09400.1	(216)	GHVISQDEGKQSLVRPANESDRAESSDDGGGQ---DDSPMLEGTSGE--		
G.max Glyma02g13860.1	(199)	GHVIFQDEGKQSLVRNANESDRAESSDDGGGQ---DDSPMLEGTSGE--		
M.truncatula AC150842.12.5	(190)	CPVMSQDEGKQVLGRSCNEADGDESSGDDG-----SQMLDCASGE--		
O.sativa LOC_Os06g16400.1	(133)	EPCREIEDIEIEGSKNVSTGEKQQGDGETTHAVDSSSKELSMGRNGGA		
A.lyrata 491205	(75)	-----DPRMSISPETTLGTGNFKKRKFDTETKDRNEKKKK--		
Aquilegia sp TC24120	(122)	-----ADVVEAKGGEVVVAEKLGGKVGKESFKKR---KAQN--		
R.communis TA2848.3988	(127)	-----VADAGPVLVKGRESVSEKLTGCVGSESTKKRKKVDKVQN--		
S.lycopersicum TC196509	(129)	-----TMADNVVKPKDTRLSSNRGRESFKKRK---ADKN--		
H.tuberosus TA2737.4233	(80)	-----NASSVLVSPKKRKADKQTLQVVSEKIKKIG--		
V.vinifera GSVIVT00014998001	(146)	GCPPNFPPNKEGLTEKSLNGAQSQEGHQISEGDAVDASPSGKRKSSFD		
Z.officinale TA5989.94328	(1)	-----MDPP-EEENSDDSKR--		
A.lyrata 494918	(215)	-----GGEFPRKRKSVPGKGSKENPIST--		
A.thaliana AT5G48560.1	(215)	-----GGEFSRKRKSVPGKGSKENPIST--		
A.thaliana AT3G07340.1	(199)	-----S-----SGELSRKRK----TKSKQNSPSA--		
Aquilegia sp TC26724	(1)	-----MKSVNDVN-ARKRKTAPK--		
G.hirsutum TC144761	(33)	RCS---SPENGDSKDESSISEQIP-GVGSSIKAQNDAN-ARKRKSMPR--		
G.raimondii TC9358	(33)	RCS---SPENGDSKDESSISEQIP-GVGSSIKAQNDAN-ARKRKSMPR--		
P.persica TC11169	(244)	RSSTPENAEGLDSREGSSVSEQIP-GGMSVKAENVTN-SRKRKPVAR--		
P.trichocarpa 572918	(255)	R---PTTSENGDSREESSLSEQVP-GGKLSMKSQNDAN-SRKRKSIPR--		
Poptr TA1	(255)	R---PTTSENGDSREESSLSEQVP-GGKLSMKSQNDAN-SRKRKSIPR--		
G.max Glyma03g21770.1	(236)	-----SQMEESTISEQ-TPNGEIGVKTSQDIMNSRKRKASSK--		
G.max GM06MC18997.59691073018643	(129)	IHQEQEKMEGANSQEESTISEQ-APNGEIGVKTSQDMNSRKRKASSK--		
P.vulgaris TC12782	(1)	-----MEVTNSQEESTISEQ-TPNGEIGMKIHQDMTNSRKRKASSK--		
G.max Glyma08g46040.1	(268)	---LQDLMEVANSQEESAISEQ-TPNGDTGEKPSPYVN-SRKRKGPSK--		
G.max Glyma18g32560.1	(263)	---LQDLMELANSQEESTISEQ-TPNGDTGVKPSPYVN-SRKRKGPSK--		
M.truncatula AC171266.17.4	(171)	LQQENEKMEVANSQEESTISEQNTTPNGEIGVKASPD MN-SRKRKASS---		
P.trichocarpa 553223	(247)	R-----SELTNSQESSVSEQ-IPSGEAGLKASNELN-SRKRKALS K--		
Triphysaria sp TC7138	(61)	-----SSNEESSVSEQ-----TGSKASNEMN-SRKRKVVS R--		
S.lycopersicum TC215759	(207)	MCERLNLGKISGRNEECVSEQDP-NGEMGSKTRNVLN-SKKRKAVKS--		
S.tuberosum TC168229	(1)	-----		
C.annuum TC15396	(49)	-----AVAAVEQEHNMSDVPENDPMRRARKRREDDECA		
O.sativa LOC_Os08g41320.1	(117)	-----ANAKKRKSTAAAAAAKKGKKEP-PVGEEKESD--		
O.sativa LOC_Os09g32510.4	(88)	-----GDNARKRKAPASAAAGKDKDAVV-GGGSSPCEV--		
A.lyrata 495154	(67)	-----SGLEKQANTTKTATTRSSSCD---HGS--		
A.thaliana AT5G50915.1	(63)	-----SGSEKLANTTKTATTGSSSSCDQLSHGP--		
B.napus TC73386	(67)	-----SDVKKQINTNATAALTGDQLS----PG--		
C.annuum NP13055023	(92)	-----VTDKMESGNNKPNVTSPMDKKRKSRE-GS--		
N.tabacum FG645212	(1)	-----		
P.hybrida TC3261	(1)	-----		
G.max Glyma04g37690.1	(73)	-----PQSSMVVDKLETGEQVTQKVNTPVEKKRKRTRNGS--		
G.max Glyma06g17420.1	(77)	-----PQSSMVVDKLETGEQVTQKVNI PVDKKRRTNGS--		
G.max Glyma05g38450.1	(81)	-----PETSMVVDKLEKGEQVTQKVVT RMEKKRRARNGS--		
G.raimondii TC6315	(78)	-----DDSSTAVDKLESGEQVTQKKRTRNGTTS----N---		
P.trichocarpa 575437	(95)	-----TNSSTVVDKLES-EHVTQEVTPMARKRKS A--NG--		
P.trichocarpa 773249	(92)	-----TGSSTVVDKLETGEQVTQEVTPVDRKRKTT--NG--		
M.domestica TC28668	(90)	-----SSTVVDVDRIEIGDQVTQKVTPMGKKRKS R--TG--		
C.sinensis TC686	(95)	-----AVSSTEQQVTQSTVTTTATTTTRMDKKRKNRTNAS--		
O.sativa LOC_Os05g01256.1	(85)	-----PQECPLSFVFDNAAAAAGDNKWVPGIQGSCP		
O.sativa LOC_Os08g42470.1	(66)	-----AVLDTSPS-----VDRKRK-----		
P.virgatum TC28308	(65)	-----AVVDTSPQ-----GSASPMDRKRKATEDS---		
S.bicolor Sb07g025920.1	(54)	-----AVVDTSPQ-----DSASPMERKRKATEDS---		
Z.mays_ZM07MC24807.BF50298G14024734	(61)	-----AVVDTSPQPQGSAAAASPMERKRKPAEDS---		
Z.mays TC476562	(68)	-----VVDSASPME-----RKRK-ATAEDS---		
O.sativa LOC_Os09g33580.1	(50)	-----VVLDTSPQD-----K-----KRKPREE---		
T.aestivum TC301717	(45)	-----PLVATASVE-----K-----KRKPKE---		
S.bicolor Sb02g029530.1	(44)	-----GVLDTS PQG-----TASDDKKP-RKPRED S---		
Z.mays TC511518	(44)	-----GVLDTS PRG-----TSSDDNKACRKHKEDSG---		
Z.mays_ZM07MC32776.BFb0337M22032678	(44)	-----GVLDTS PRG-----TSSDDNKACRKHKEDSG---		
S.officinarum TC78398	(43)	-----GVLDTS PQG-----TASDDKKP-RKPRED S---		
Z.mays TC460846	(46)	-----GVLDTS PRGG--ADSGGRKAPT KPREDSASA		
Aquilegia sp TC24680	(77)	-----DHQVTDQMDMTPVEEKKRKNRDPGSCLS--		
O.sativa LOC_Os04g28280.1	(1)	-----MNPTTAAAADQPSKPSAAAAARKRKSSAKPKASSS--		
Consensus	(301)			RK

FIGURE 6 (continued)

		351	400
A.lyrata 472791	(113)	----TSSKGFNSKKRKRIG-----	KDCEGEEDKMQKDEQ
A.thaliana AT1G26260.1	(168)	----TSSKGFSSKKRKRIG-----	KDCEEEEDKKQKDEQ
A.lyrata 476021	(244)	----VNSEAAQSHRSQQSE-----	EEPNNNGDKKRNSEQ
A.thaliana AT1G68920.1	(244)	----KNSEAAQSHRSQQSE-----	EEPNNNGDEKRNDEQ
G.max Glyma01g09400.1	(260)	----PSSKGLNTKKRKRSGQDGDNDKANGAQELPSEGAEDNYENQQKGDH	
G.max Glyma02g13860.1	(243)	----PSSKGLNSKKRKRSGRDGDNDKANGAQELPSEGAAGNSENQQKGDQ	
M.truncatula AC150842.12.5	(230)	----PSIKGLNPKKRKRNGQDGDSDKATGTLELPSETAKDNCESTRKKGKQ	
O.sativa LOC_Os06g16400.1	(183)	G-HDEGTRVSCSKKRKRSGQDGGVKHAEGGEQLATVGSAQKNEDEKGEF	
A.lyrata 491205	(110)	----MNREEVTEEEEEKSK-----	ITELNNGSPKSIKKM
Aquilegia sp TC24120	(156)	-----SKVVAEEVIRDKRMKGDSEE---ES-KVTEVNNSN-SSTSNNNN	
R.communis TA2848.3988	(166)	----NTKVVAEDDNCRDKRIKVCAGEG--ESKMITEKNNNNKSSSSKNSN	
S.lycopersicum TC196509	(160)	----QHLKEVAEEETKDKKLKECIEEED-DSSKVTEKSKNSKRSATNSSN	
H.tuberosus TA2737.4233	(112)	-----CCDGD SKLTRQNS-----	NGNNNNNNKE--AFT
V.vinifera GSVIVT00014998001	(196)	RPPLNTSKSADGEQPKGLPWE-----N---SEFSKEQEEKKLKIDQ	
Z.officinale TA5989.94328	(14)	----CKAAEAGGADKDAP-----	VKPKTEQNSSDDGNAK
A.lyrata 494918	(238)	----ASPSPSFSKTAEEKKEEWGGKG-----	SKSSEEEKGGKRRREDE
A.thaliana AT5G48560.1	(238)	----ASPSPSFSKTAEN---GGKGG-----	SKSSEEEKGGKRRREDE
A.thaliana AT3G07340.1	(219)	----VS-SS---KEIEEK-----	EDSDPKRCCKSE
Aquilegia sp TC26724	(18)	----SKAKDITSLSSSSKDAKEATDDDK-SNGKRSKSNDTNGNEKASAKA	
G.hirsutum TC144761	(76)	----GKAKETPSP--AAADAKIAEENDE-STAKRSK---QQ-NDSAKAE-	
G.raimondii TC9358	(76)	----GKAKETPSP--AAAEAKIAEENNK-STAKRSK---QQ-NDSAKAE-	
P.persica TC11169	(290)	----GKAKETSSS-PSVKDGVVAAEKEE-PNSKRSKTDEASGNEKAAAKA	
P.trichocarpa 572918	(298)	----GKAKETPSSSPSASDVKVAAEENDE-SKAKRSKSDETNGSDKDTAKE	
Poptr TA1	(298)	----GKAKETPSSSPSASDVKVAAEENDE-SKAKRSKSDETNGSDKDTAKE	
G.max Glyma03g21770.1	(272)	----GKAKETSNTTNPT-----	KVE
G.max GM06MC18997.59691073@18643	(176)	----GKAKETSNSFNPTKGVEGS-ED---SNSKRSKPNEGDNENGQVKV	
P.vulgaris TC12782	(41)	----GKTRETSNSINPSKGVEVN-ED---SNSMRTKPNDEGEGNEKGPVKV	
G.max Glyma08g46040.1	(311)	----GKAKETSTSTNPPMAAEAS-DD---SNAKRSKPNEGEGNENGQVKA	
G.max Glyma18g32560.1	(306)	----GKAKETSASINPPMAAEAS-ED---WNAKRSKPNAEGEGNENGQVKA	
M.truncatula AC171266.17.4	(217)	-----SKGKAPNSTNPTKGVEGSGED---FNAKKIKANEGERNENGVRNM	
P.trichocarpa 553223	(286)	----GKAKQSASNPASATKDAETDDN--SNTKRIKPNEGEENENSPVKA	
Triphysaria sp TC7138	(91)	----GKSKEDGSNLAKRVEGVN-----ENAKRSK---TTKSVKIETKT	
S.lycopersicum TC215759	(253)	----KDFVPIVDETGKK-----	RAKSTQGNGSNNGTVKM
S.tuberosum TC168229	(1)	-----	
C.annuum TC15396	(82)	K-GVSTSGNGLTESDSKR-----	LKATRLNENCEARCDG
O.sativa LOC_Os08g41320.1	(149)	----GKRCKTGNGEKESS-----	VKPKAEQAGSDSSVED
O.sativa LOC_Os09g32510.4	(120)	----GEAKAPDSKKCKAE-----	VNPKVEEAASDGSVGD
A.lyrata 495154	(91)	----SAITTIKGRGRKAR-----N-----	VSNSKEGVEGTRKSK
A.thaliana AT5G50915.1	(90)	----SAITNTGKTRGRKAR-----N-----	SNNKEGVEGRKSK
B.napus TC73386	(90)	----PSTTSAGKKQRRKTR-----N-----	GSKSKEGVEGRKCK
C.annuum NP13055023	(120)	----SSMTSANSKNVTQGD-----N-----	GKKNKSNKSLVAKD
N.tabacum FG645212	(1)	-----MSSAHSKNVKQVD-----N-----	GKKKKSNQSVAKE
P.hybrida TC3261	(1)	-----MSSAHSKNVREGD-----H-----	GKKKKNNNSQ-VSKD
G.max Glyma04g37690.1	(107)	----SLTSNPQSKDAATER-----K-----	NKKQRKNGGGLKEE
G.max Glyma06g17420.1	(111)	----SFTSNPQSKDTATEG-----K-----	NKKQRKNNGGLKEE
G.max Glyma05g38450.1	(115)	----SSS-DPLSKESTE-----	GGKKKQ
G.raimondii TC6315	(107)	-----SPQFKGAKE-G-----K-----	TKKQKKCNDGLKNE
P.trichocarpa 575437	(126)	----FLN-SAQSKDARK-V-----K-----	SKRQNKCSGDMKHE
P.trichocarpa 773249	(124)	----SLN-SAQSKDVKE-V-----K-----	SKRQKKCRGDMKQE
M.domestica TC28668	(122)	----SSFNSAQSKGTKEQK-----	GKKQKKIDGGTKKD
C.sinensis TC686	(129)	----SSFNSAQSKDPREVK-----S-----	KKQKKGNVNDKKE
O.sativa LOC_Os05g01256.1	(116)	C-SLGSTQDMDASWGKSRK-----	HKRSNVGLKGLE
O.sativa LOC_Os08g42470.1	(80)	----AAEDSAHSKDSCK-----	DGKSRRGKKAS---
P.virgatum TC28308	(89)	----ATLSSAHSKDCNQ-----	EGKSKRGKRSN---
S.bicolor Sb07g025920.1	(78)	----ATLSSAQSKDCKQQ-----	ESKSKRGKRPN---
Z.mays_ZM07MC24807_BF50298G14@24734	(90)	----ATLSSAQSKEDCKQQ-----Q-----	ESKSRGKRPHNKE
Z.mays TC476562	(87)	----AAISSPQPKDCKQ-----	ESKSRGKRPR---
O.sativa LOC_Os09g33580.1	(67)	--DTASLNSAHSKEAK-----	ENGRKRGKK---
T.aestivum TC301717	(61)	--DTASLNSAHSKETK-----	ESTRKRGGKKQ---
S.bicolor Sb02g029530.1	(68)	----ASFSSAHSKDSNSK-----	ESTKKKGKGRD---
Z.mays TC511518	(70)	----ASFSSARSKDSNSK-----	ESSKRSGGKRD---
Z.mays_ZM07MC32776_BFb0337M22@32678	(70)	----ASFSSARSKDSNSK-----	ESSKRSGGKRD---
S.officinarum TC78398	(67)	----TSFSSAHSKDSNSK-----	ESTKKRGKGRD---
Z.mays TC460846	(76)	S-ASFSSAHSKDSNSK-----	GSTKNRGKGR---
Aquilegia sp TC24680	(104)	----SSEKDSKISTRK-----	KQRNCSGNINDAE
O.sativa LOC_Os04g28280.1	(36)	----SLPTATATTNAPKR-----	SKVAAGAGDDGDADA
Consensus	(351)	SK	

FIGURE 6 (continued)

	401	450
A.lyrata 472791	(143)	SSNSNVN----KTNSEK-----QASDSLKDGYIHMRAARRGQATNSHSLAE
A.thaliana AT1G26260.1	(198)	SPTSNAN----KTNSEK-----QPSDSLKDGYIHMRAARRGQATNSHSLAE
A.lyrata 476021	(274)	SPNSPGN----KTNSGKRQ-GKQSSDLPKDGYIHVRAARRGQATNSHSLAE
A.thaliana AT1G68920.1	(274)	SPNSPGK---KNSNGKQQ-GKQSSDPPKDGYIHVRAARRGQATNSHSLAE
G.max Glyma01g09400.1	(306)	QPTSTA----KASGKNAKLGSQASDPPKEEYIHVRAARRGQATNSHSLAE
G.max Glyma02g13860.1	(289)	QPISTAN----KACGKNAKLGSQASDPPKEEYIHVRAARRGQATNSHSLAE
M.truncatula AC150842.12.5	(276)	QTSSTA-----KASGKNAKQGSQASDPPNEGYVHVRAARRGQATNSHSLAE
O.sativa LOC_Os06g16400.1	(232)	KRSSVAS-----GKSSGKQIKDNAGSPKEDYIHVRAARRGQATNSHSLAE
A.lyrata 491205	(140)	KNKAKKE-----ENNCSDSSKVTKELEKTDYIHVRAARRGQATDSHSLAE
Aquilegia sp TC24120	(195)	KETS-----GDSVKENSKASGAQKQDYIHVRAARRGQATDSHSLAE
R.communis TA2848.3988	(210)	KENS-----AETSKDNSKVTEVQKPDYIHVRAARRGQATDSHSLAE
S.lycopersicum TC196509	(205)	SKEN-----SDTSKEKSKITDDKKLDYIHVRAARRGQATDSHSLAE
H.tuberosus TA2737.4233	(138)	KAS-----SKDKQKVSEDQKKDYIHVRAARRGQATDSHSLAE
V.vinifera GSVIVT00014998001	(234)	NMSPNLRG---KQPNKHAKDNSSNGEAPKENYIHVRAARRGQATNSHSLAE
Z.officinale TA5989.94328	(44)	P-----PEPPKDYIHVRAARRGQATDSHSLAE
A.lyrata 494918	(275)	DEEE-----GEGEGN--NNTKPPEPPKDYIHVRAARRGQATDSHSLAE
A.thaliana AT5G48560.1	(273)	DDEEE-----EGEGEGNKSNNTPPEPPKDYIHVRAARRGQATDSHSLAE
A.thaliana AT3G07340.1	(241)	ENGD-----KTKSIDPYKDYIHVRAARRGQATDSHSLAE
Aquilegia sp TC26724	(63)	KGEPNSGDGAK---KQNKDTPKPSPEPPKDYIHVRAARRGQATDSHSLAE
G.hirsutum TC144761	(114)	-----NKENPK-PPEPPKDYIHVRAARRGQATDSHSLAE
G.raimondii TC9358	(114)	-----NKENPK-PPEPPKDYIHVRAARRGQATDSHSLAE
P.persica TC11169	(334)	KIEPSGGSKATGDGAQKQTKDNSEPPEAPKDYIHVRAARRGQATDSHSLAE
P.trichocarpa 572918	(343)	KEEENGN-----QKQNKNSKPPPEPPKDYIHVRAARRGQATDSHSLAE
Poptr TA1	(343)	KEEENGN-----QKQNKNSKPPPEPPKDYIHVRAARRGQATDSHSLAE
G.max Glyma03g21770.1	(288)	EE-----SKAEEEKQSKSNSKPPPEPPKDYIHVRAARRGQATDSHSLAE
G.max GM06MC18997.59691073@18643	(218)	EEE-----SKAEEEKQNKSNSKPPPEPPKDYIHVRAARRGQATDSHSLAE
P.vulgaris TC12782	(83)	EEE-----SKAVEEKQNKSNSKPPPEPPKDYIHVRAARRGQATDSHSLAE
G.max Glyma08g46040.1	(353)	EEESKGGN--NSNANDEKQNKSNSKPPPEPPKDYIHVRAARRGQATDSHSLAE
G.max Glyma18g32560.1	(348)	EEESKGGN--SNANDEKQNKSNSKPPPEPPKDYIHVRAARRGQATDSHSLAE
M.truncatula AC171266.17.4	(259)	EEEIKEG--TPNAGEEKQNKSDSKPPEPQKDYIHVRAARRGQATDSHSLAE
P.trichocarpa 553223	(330)	EEEP-----KSGGDDIQNKANSRPPEPPKDYIHVRAARRGQATDSHSLAE
Triphysaria sp TC7138	(128)	EDETKGG---SNDESEKQ----KAAEPPKDYIHVRAARRGQATDSHSLAE
S.lycopersicum TC215759	(283)	EEQKGN----EDDGAEKETKENRKIAEPPKDYIHVRAARRGQATDSHSLAE
S.tuberosum TC168229	(1)	MEES-----KDDEIQKVMEHQKPPPEPPKDYIHVRAARRGQATDSHSLAE
C.annuum TC15396	(115)	EGNS-----RKSAEQPAKPAEPPKDYIHVRAARRGQATDSHSLAE
O.sativa LOC_Os08g41320.1	(179)	GGGG-----GQKQKGKGN-AKPVE-PPKDYVHVRAARRGQATDSHSLAE
O.sativa LOC_Os09g32510.4	(150)	R-----VQKQKGKGNSSKPAAEPPKDYVHVRAARRGQATDSHSLAE
A.lyrata 495154	(121)	KQ-----KRGSKKEPPKDYIHVRAARRGQATDSHSLAE
A.thaliana AT5G50915.1	(120)	KQ-----KRGSKKEPPTDYIHVRAARRGQATDSHSLAE
B.napus TC73386	(120)	KPRNGSI-----STDVKVKRASEQEPPKDYIHVRAARRGQATDSHSLAE
C.annuum NP13055023	(150)	EKK-----ANEEAPTGYIHVRAARRGQATDSHSLAE
N.tabacum FG645212	(29)	EKKKGK-----EEKKANEEAPTGYIHVRAARRGQATDSHSLAE
P.hybrida TC3261	(28)	AKEKKGS-----DDNKKEENKSPEEAPTGYIHVRAARRGQATDSHSLAE
G.max Glyma04g37690.1	(137)	DKA-----KEEKKEQRKCPEEPPTGYIHVRAARRGQATDSHSLAE
G.max Glyma06g17420.1	(141)	DKA-----KEEKKDQRKCPEEPPTGYIHVRAARRGQATDSHSLAE
G.max Glyma05g38450.1	(133)	KKP-----KEVTKDKKIGAEDPPTGYIHVRAARRGQATDSHSLAE
G.raimondii TC6315	(132)	KKGCN-S-----KADKKDQEKSTEQPPTGYIHVRAARRGQATDSHSLAE
P.trichocarpa 575437	(154)	EK--K-P-----KVEKKVHG---EPPAGYIHVRAARRGQATDSHSLAE
P.trichocarpa 773249	(152)	EK--R-P-----KAVKKVPE---EPPTGYVHVRAARRGQATDSHSLAE
M.domestica TC28668	(151)	EEKK--I-----KSDEEDQVKVREGPPTGYIHVRAARRGQATDSHSLAE
C.sinensis TC686	(159)	EKE--N-----SPKADSEKKVAKEPPKDYIHVRAARRGQATDSHSLAE
O.sativa LOC_Os05g01256.1	(146)	EKKARRVVLHQHDDDVKKKAKEAAGGEPPAGYIHVRAARRGQATDSHSLAE
O.sativa LOC_Os08g42470.1	(104)	KEV-----EEKS--TTEDEPPKGYIHVRAARRGQATDSHSLAE
P.virgatum TC28308	(113)	KEA-----EEKS--TTDEAPKGYIHVRAARRGQATDSHSLAE
S.bicolor Sb07g025920.1	(103)	KET-----EEKS--TTEDEATKGYIHVRAARRGQATDSHSLAE
Z.mays_ZM07MC24807_BF50298G14@24734	(120)	ANT-----EEKS--ATEDEATRGYIHVRAARRGQATDSHSLAE
Z.mays TC476562	(111)	KET-----EEKSSTDEDEASKGYIHVRAARRGQATDSHSLAE
O.sativa LOC_Os09g33580.1	(91)	-----HSRDQMEEEA-PQGFIHVRAARRGQATDSHSLAE
T.aestivum TC301717	(86)	-----GKEMDEEEEEE-PKGYIHVRAARRGQATDSHSLAE
S.bicolor Sb02g029530.1	(93)	-----RSSKEVDEEEE-PKGYIHVRAARRGQATDSHSLAE
Z.mays TC511518	(95)	-----RSSKKVDEEEE-PKGYIHVRAARRGQATDNHSLAE
Z.mays_ZM07MC32776_BFb0337M22@32678	(95)	-----RSSKKVDEEEE-PKGYIHVRAARRGQATDNHSLAE
S.officinarum TC78398	(92)	-----RSSKKVDEE--PKGYIHVRAARRGQATDSHSLAK
Z.mays TC460846	(104)	-----RSSKEVDEEEEPKGYIHVRAARRGQATDSHSLAE
Aquilegia sp TC24680	(131)	KKMSKDE-----KKNQRKDSVAQSGYVHVRAARRGEATDSHSLAE
O.sativa LOC_Os04g28280.1	(66)	DAAE-----EKPEPAKDYIHVRAARRGQATDSHSLAE
Consensus	(401)	EEPPKDYIHVRAARRGQATDSHSLAE

FIGURE 6 (continued)

		451	500
A.lyrata 472791	(184)	RVRREKISERMKFLQDLVPGCDKVTGKAVMLDEI	IN YVQSLQCQIEFLSM
A.thaliana AT1G26260.1	(239)	RVRREKISERMKFLQDLVPGCDKVTGKAVMLDEI	IN YVQSLQCQIEFLSM
A.lyrata 476021	(319)	RVRREKISERMKFLQDLVPGCNKVTGKAVMLDEI	IN YVQSLQRQVEFLSM
A.thaliana AT1G68920.1	(319)	RVRREKISERMKFLQDLVPGCNKVTGKAVMLDEI	IN YVQSLQRQVEFLSM
G.max Glyma01g09400.1	(351)	RVRREKISERMKFLQDLVPGCSKVTGKAVMLDEI	IN YVQSLQRQVEFLSM
G.max Glyma02g13860.1	(335)	RVRREKISERMKFLQDLVPGCSKVTGKAVMLDEI	IN YVQSLQRQVEFLSM
M.truncatula AC150842.12.5	(321)	RVRREKISERMKFLQDLVPGCNKVTGKAVMLDEI	IN YVQSLQQQVEFLSM
O.sativa LOC_Os06g16400.1	(276)	RVRREKISERMKYLQDLVPGCSKVTGKAVMLDEI	IN YVQSLQRQVEFLSM
A.lyrata 491205	(185)	RVRREKISERMKFLQDLVPGCDKITGKAGMLDEI	IN YVQSLQRQIEFLSM
Aquilegia sp TC24120	(235)	RVRREKISERMKYLQDLVPGCNKITGKAGMLDEI	IN YVQSLQRQVEFLSM
R.communis TA2848.3988	(250)	RVRREKISERMKYLQDLVPGCNKITGKAGMLDEI	IN YVQSLQRQVEFLSM
S.lycopersicum TC196509	(245)	RVRREKISERMFLQDLVPGCNKITGKAGMLDEI	IN YVQSLQRQVEFLSM
H.tuberosus TA2737.4233	(174)	RVRREKISERMKYLQDLVPGCNKIRGKAGMLDEI	IN YVQSLQRQVEFLSM
V.vinifera GSVIVT00014998001	(281)	RVRREKISERMRLQLQELVPGCNKITGKAVMLDEI	IN YVQSLQQQVEFLSM
Z.officinale TA5989.94328	(70)	RVRREKISERMKLLQDLVPGCNKVTGKAVMLDEI	IN YVQSLQRQVEFLSM
A.lyrata 494918	(315)	RVRREKIGERMKLLQDLVPGCNKVTGKALMLDEI	IN YVQSLQRQVEFLSM
A.thaliana AT5G48560.1	(317)	RVRREKIGERMKLLQDLVPGCNKVTGKALMLDEI	IN YVQSLQRQVEFLSM
A.thaliana AT3G07340.1	(274)	RVRREKISERMKLLQDLVPGCNKVTGKALMLDEI	IN YVQSLQRQVEFLSM
Aquilegia sp TC26724	(109)	RVRREKISERMKFLQDLVPGCNKVTGKALMLDEI	IN YVQSLQRQVEFLSM
G.hirsutum TC144761	(146)	RVRREKISERMKFLQDLVPGCNKVTGKALMLDEI	IN YVQSLQRQVEFLSM
G.raimondii TC9358	(146)	RVRREKISERMKFLQDLVPGCNKVTGKALMLDEI	IN YVQSLQRQVEFLSM
P.persica TC11169	(384)	RVRREKISERMKFLQDLVPGCNKVTGKAVMLDEI	IN YVQSLQRQVEFLSM
P.trichocarpa 572918	(385)	RVRREKISERMKFLQDLVPGCNKVTGKAVMLDEI	IN YVQSLQRQVEFLSM
Poptr TA1	(385)	RVRREKISERMKLLQDLVPGCNKVTGKALMLDEI	IN YVQSLQRQVEFLSM
G.max Glyma03g21770.1	(330)	RVRREKISERMKLLQDLVPGCNKVTGKALMLDEI	IN YVQSLQRQVEFLSM
G.max GM06MC18997.59691073018643	(261)	RVRREKISERMKLLQDLVPGCNKVTGKALMLDEI	IN YVQSLQRQVEFLSM
P.vulgaris TC12782	(126)	RVRREKISERMKLLQDLVPGCNKVTGKALMLDEI	IN YVQSLQRQVEFLSM
G.max Glyma08g46040.1	(402)	RVRREKISERMKLLQDLVPGCNKVTGKALMLDEI	IN YVQSLQRQVEFLSM
G.max Glyma18g32560.1	(396)	RVRREKISERMKLLQDLVPGCNKVTGKALMLDEI	IN YVQSLQRQVEFLSM
M.truncatula AC171266.17.4	(307)	RVRREKISERMKLLQDLVPGCNKVTGKALMLDEI	IN YVQSLQRQVEFLSM
P.trichocarpa 553223	(374)	RVRREKISERMKLLQDLVPGCNKVTGKALMLDEI	IN YVQSLQRQVEFLSM
Triphysaria sp TC7138	(170)	RVRREKISERMKLLQDIVPGCNKVTGKALMLDEI	IN YVQLLRQVEFLSM
S.lycopersicum TC215759	(329)	RVRREKISQRMKLLQDLVPGCNKVTGKALMLDEI	IN YVRSLQHQVEFLSM
S.tuberosum TC168229	(45)	RVRREKISERMKLLQDLVPGCNKVTGKALMLDEI	IN YVQSLQRQVEFLSM
C.annuum TC15396	(154)	RARREKISERMKVLQDIVPGCNKVIGKALVLDEI	IN YIQSLQHQVEFLSM
O.sativa LOC_Os08g41320.1	(220)	RVRREKISQRMKVLQDLVPGCNKVIGKALMLDEI	IN YVQSLQRQVEFLSM
O.sativa LOC_Os09g32510.4	(190)	RVRREKISQRMKVLQDLVPGCNKVVGKALMLDEI	IN YVQSLQQQVEFLSM
A.lyrata 495154	(153)	RVRREKISERMRTLQNLVPGCDKVTGKALMLDEI	IN YVQTLQNQVEFLSM
A.thaliana AT5G50915.1	(152)	RVRREKISERMRTLQNLVPGCDKVTGKALMLDEI	IN YVQTLQTQVEFLSM
B.napus TC73386	(163)	RARREKISERMRTLQNLVPGCDKVTGKALMLDEI	IN YVQSLQNQVEFLSM
C.annuum NP13055023	(180)	RVRREKISERMKMLQSLVPGCDKVTGKALMLDEI	IN YVQSLQNQVEFLSM
N.tabacum FG645212	(65)	RVRREKISERMKILQSIIVPGCDKVTGKALMLDEI	IS YVQSLQNQVEFLFM
P.hybrida TC3261	(71)	RVRREKISERMKILQSLVPGCDKVTGKALMLDEI	IN YVQSLQNQVEFLSM
G.max Glyma04g37690.1	(176)	RVRREKISERMKVLQRLVPGCDKVTGKALMLDEI	IN YVQSLQNQVEFLSM
G.max Glyma06g17420.1	(180)	RVRREKISERMKVLQRLVPGCDKVTGKALMLDEI	IN YVQSLQNQVEFLSM
G.max Glyma05g38450.1	(172)	RVRREKISKRMTTLQRLVPGCDKVTGKALVLDEI	IN YVQSLQNQVEFLSM
G.raimondii TC6315	(174)	RVRREKISERMKTLQRLVPGCDKVTGKALMLDEI	IN YVQSLQNQVEFLSM
P.trichocarpa 575437	(190)	RVRREKISERMKILQLLVPGCDKITGKALMLDEI	IN YVQSLQNQVEFLSM
P.trichocarpa 773249	(188)	RVRREKISERMKMLQRLVPGCDKVTGKALMLDEI	IN YVQSLQNQVEFLSM
M.domestica TC28668	(192)	RVRREQISERMKMLQRLVPGCDKITGRAVMLDEI	IN YVQSLQNQVEFLSV
C.sinensis TC686	(199)	RVRREKISERMKILQKLVPGCDKVTGKAFMLDEI	IN YVQFLQNQVEFLSM
O.sativa LOC_Os05g01256.1	(196)	RVRREKISERMKMLQSLVPGCDKVTGKALMLDEI	IS YVQSLQNQVEFLSM
O.sativa LOC_Os08g42470.1	(139)	RVRREKISERMRLQALVPGCDKVTGKALILDEI	IN YVQSLQNQVEFLSM
P.virgatum TC28308	(148)	RVRREKISERMRLQALVPGCDKVTGKALILDEI	IN YVQSLQNQVEFLSM
S.bicolor Sb07g025920.1	(138)	RVRREKISERMRLQALVPGCDKVTGKALILDEI	IN YVQSLQNQVEFLSM
Z.mays_ZM07MC24807_BF50298G14024734	(155)	RVRREKISERMRLQALVPGCDKVTGKALILDEI	IN YVQSLQNQVEFLSM
Z.mays TC476562	(147)	RVRREKISERMRLQALVPGCDKVTGKAMVLDEI	IN YVQSLQNQVEFLSM
O.sativa LOC_Os09g33580.1	(123)	RVRREKISERMRLQALVPGCDKVTGKALILDEI	IN YVQSLQNQVEFLSM
T.aestivum TC301717	(118)	RVRREKISERMRLQSLVPGCDKVTGKALILDEI	IN YVQSLQNQVEFLSM
S.bicolor Sb02g029530.1	(125)	RVRREKISERMRLQALVPGCDKVTGKALILDEI	IN YVQSLQNQVEFLSM
Z.mays TC511518	(127)	RVRREKISERMRLQALVPGCDKVTGKALILDEI	IN YVQSLQNQVEFLSM
Z.mays_ZM07MC32776_BFb0337M22032678	(127)	RVRREKISERMRLQALVPGCDKVTGKALILDEI	IN YVQSLQNQVEFLSM
S.officinarum TC78398	(123)	RVRREKISERMRLVXLVPGCDKVTGKALILDEI	IN YVQSLQNQVEFLSM
Z.mays TC460846	(137)	RVRREKISERMRLVXLVPGCDKVTGKALVLDEI	IN YVQSLQNQVEFLSM
Aquilegia sp TC24680	(171)	RVRREKISERLKLQSLVPGCDKVTSKALMLDEI	IN YVQSLQHQVEFLSM
O.sativa LOC_Os04g28280.1	(97)	RVRREKISERMKLLQSLVPGCNKITGKALMLDEI	IN YVQSLQRQVEFLSM
Consensus	(451)	RVRREKISERMKLLQDLVPGC KVTGKALMLDEI	IN YVQSLQ QVEFLSM

FIGURE 6 (continued)

31/42

		501		550
A.lyrata 472791	(234)	KLAVN-PVLDNFNLESLLAKDALQSSAPT-----FPQNMSMLYPPL		
A.thaliana AT1G26260.1	(289)	KLAVN-PVLDNFNLESLLAKDALQSSAPT-----FPHNMSMLYPPL		
A.lyrata 476021	(369)	KLATVN-PQMDFNLEGLLAKDALQLRAGSSS---TTTPFTPNMAMAYPPL		
A.thaliana AT1G68920.1	(369)	KLATVN-PQMDFNLEGLLAKDALQLRAGSS---STTPFPNMSMAYPPL		
G.max Glyma01g09400.1	(401)	KLATVN-PRLDNFNIEGLLAKDILQQRPGPS---SALGFPLDMSMAFPPL		
G.max Glyma02g13860.1	(385)	KLATVN-PRLDNFNIEGLLAKDILQQRPDPS---TALGFPLDMSMAFPPL		
M.truncatula AC150842.12.5	(371)	KLATVN-PHVDFNMERLLPKDILQHQPGPS---SALGFPLREMPMAFPPL		
O.sativa LOC_Os06g16400.1	(326)	KLAVN-PTLDNFNIEERILSKDIFQCRG-----TTASSAFGFPPDIVH		
A.lyrata 491205	(235)	KLAVN-PRPDEFMDDDIFDKEVIVS-----T		
Aquilegia sp TC24120	(285)	KLAVN-PRLDNFNIEENLFA-KEMFSANTAGFP-IGTSSEMANPS-----		
R.communis TA2848 3988	(300)	KLAVN-PRLDNFNIDNLIA-KETFPFCPTNFPAIGLSSDMTNP-----		
S.lycopersicum TC196509	(295)	KLAVN-PRLDIDADNFFN-KDIFATSTSTFSAVGAGTSSEMLSM-----		
H.tuberosus TA2737 4233	(224)	KLAVN-PELDNFNMDNAFTKEMFDSSTSEHPALACSSANSAYFDLNSX-G		
V.vinifera GSVIVT00014998001	(331)	KLATVN-PELNIDIERLLSKDILN-----SRG		
Z.officinale TA5989 94328	(120)	KLATLDPQ-LEFDTETLLSKEMHQVHVP---LPQQVLFPLDMAT-----		
A.lyrata 494918	(365)	KLSSVNDTRLDFNVDALVSKDVMIPSSNNR--LHEEGLQSKSSSR-HHQ		
A.thaliana AT5G48560.1	(367)	KLSSVNDTRLDFNVDALVSKDVMIPSSNNR--LHEEGLQSKSSS---HHHQ		
A.thaliana AT3G07340.1	(324)	KLSSVN-TRLDFNMDALLSKDIFP-SSNNL--MHHQQVLQLDSS---AE		
Aquilegia sp TC26724	(159)	KLATVN-PRLDNFNMEATLT-KDLLQSRG---SLSQAMYQLNSSVS---A		
G.hirsutum TC144761	(196)	KLATVN-PRMNANTEALLS-KDMFRSRG---SLPHSHYPMDSSAL---A		
G.raimondii TC9358	(196)	KLATVN-PRMNANTEALLS-KDMFRSRG---SLPHSHYPMDSSAL---A		
P.persica TC11169	(434)	KLSTVN-PRMDLNMEALLS-KEILQSRG---SLQNALYQLDSAIP---S		
P.trichocarpa 572918	(435)	KLSSVN-PRMEINMETLLS-KDIFQSRG---SMPHSLYPLDASTP---V		
Poptr TA1	(435)	KLSSVN-PRMEINMETLLS-KDIFQSRG---SMPHSLYPLDASTP---V		
G.max Glyma03g21770.1	(380)	KLAVN-TRMDLSIESLVT-KDVFQSNNSLATHPNAIIFPLGSSA---Q		
G.max GM06MC18997 59691073018643	(311)	KLAVN-TRMDLSIENLIS-KDVFQSNNSLATLPN-AVFPLDSSA---Q		
P.vulgaris TC12782	(176)	KLAVN-TRMDLGIENLVS-KDVFQSNNSLATHQN-PIFPLDSSA---Q		
G.max Glyma08g46040.1	(452)	KLAVN-TRLDIFSIESLIS-KDIFQSNN---SLA-QPIFPIDSSA---P		
G.max Glyma18g32560.1	(446)	KLAVN-TRLDIFSIESLIS-KDIFQSNN---SLA-HPIFLIDSSA---P		
M.truncatula AC171266 17.4	(357)	KLSSVN-TKMDLSIESLVVSKDIFQSNN---SLQN-SIFQLDSSV---P		
P.trichocarpa 553223	(424)	KLAVN-TRLDFNMDTLIS-KDIFQSSQ---PLPHPIFPLDSSA---P		
Triphysaria sp TC7138	(220)	KLAVN-TGLDVNMENLIS-KDMFQNA---ITLPQQIYPFDSSSPAAAA		
S.lycopersicum TC215759	(379)	KLAVN-PRTDIHIDSLH-----TEISQ		
S.tuberosum TC168229	(95)	KLATVN-P---NLDYPLS-KDICQPNG---SLPHVPFVVDKTS---S		
C.annuum TC15396	(204)	KLEAVNSRAASIEGFRSKDFGHQTFDPNAMAFAFGSQATREYARGTSPEWLH		
O.sativa LOC_Os08g41320.1	(270)	KLATVNP--LDFSNLPTLLQKDMFQACG---PSASSVFSLESSN-----		
O.sativa LOC_Os09g32510.4	(240)	KLATVNPQ-LDFGNLSTLLQKDMFQSCG---PSVNSVFPLESAG-----		
A.lyrata 495154	(203)	KLTSISPVVYDFGSDLDGLVLR-----EMGSPEVG-----		
A.thaliana AT5G50915.1	(202)	KLTSISPVVYDFGSDLDGLILQS-----EMGSPEVG-----		
B.napus TC73386	(213)	KLASISPVVYDFGSDLDGIIVRP-----EMGSADV-----		
C.annuum NP13055023	(230)	KLASLNPMYYDFGMDLDALMVRP--D-----QNLSGLETQ---LPN		
N.tabacum FG645212	(115)	KLASLNPMYYDFGMDLDALMVRP--D-----QSLSGLGTP---LPN		
P.hybrida TC3261	(121)	KLASSNPMYYDFGMDLDALMVRP--N-----QNLGGLETQ---LPN		
G.max Glyma04g37690.1	(226)	KLAVNPMFYDLATDLTLVLR-----EKLNSMASPSPPLPS		
G.max Glyma06g17420.1	(230)	KLAVNPMFYDLATDLTLVLR-----EKLNSMASPSP-LPS		
G.max Glyma05g38450.1	(222)	KLAVNPMFFDSAMDLDLTMVRP-----QKLSNIASPSP-LPC		
G.raimondii TC6315	(224)	KLASLSPTFYDFGVLETLMIRP-----ERVSTSTETSPPLMPC		
P.trichocarpa 575437	(240)	KLAVNPLLYDFGMDRDAFMVRP-----ERLSSMSP---PLPS		
P.trichocarpa 773249	(238)	KLAVNPMFYDFGMELDAFMVRP-----ERLSSMSP---PLPS		
M.domestica TC28668	(242)	KLDSENPMFYDFGIDIGALVVKP-----ERLCSTESP---YQS		
C.sinensis TC686	(249)	KLAVNPMFYDFGMDLDNTFMVRPDDQ-----QNLNSIIAASAPQ		
O.sativa LOC_Os05g01256.1	(246)	KLASLSPLMYEFPGIDMHPDVLR-----QLAKM		
O.sativa LOC_Os08g42470.1	(189)	RIASMSPVLYGFGMDS-----DGLHD-----QKIGGMFQEALAMPN		
P.virgatum TC28308	(198)	RIASMSPVLYGFGLD-----DGLHDH-----AQKMGMFQEALAMPN		
S.bicolor Sb07g025920.1	(188)	RIASMSPVLYGFGLD-----DGLHDH-----AQKMGMFQEALAMPN		
Z.mays_ZM07MC24807 BF50298G14024734	(205)	RIASMSPVLYGFGLD-----DGLHDH-----AQKMGMFQEALAMAG		
Z.mays TC476562	(197)	RIASMSPVLYGFGTNS-----EGLHDH-----APKMGGVSFQEAAIAV		
O.sativa LOC_Os09g33580.1	(173)	RIASLSPVLYGFGIDS-DAFSDHSQK-----MEGMFHEAVAIPA		
T.aestivum TC301717	(168)	RIASLSPVLYGFGMESGAADFSDQSQK-----IEGMFHDAALPT		
S.bicolor Sb02g029530.1	(175)	RIASLSPVLYGFGMDT-DAFSDHTQK-----IEGMLHHEALGMP-		
Z.mays TC511518	(177)	RIASLSPVLYGFGMDSSDAFSDLITQK-----IEGMAHHEAFAMPS		
Z.mays_ZM07MC32776 BFb0337M22032678	(177)	RIASLSPVLYGFGMDSSDAFSDLITQK-----IEGMAHHEAFAMPS		
S.officinarum TC78398	(173)	KDCFPEPCAVWFHGWQ-----		
Z.mays TC460846	(187)	KIASLSPVLYGFGMDTTDAFSTDHTTQK-----IEGMLQLHHEALAVP-		
Aquilegia sp TC24680	(221)	KLASMSPIAYDMGGLDINSLVKQ-----EKLSSLSALLPPVQ		
O.sativa LOC_Os04g28280.1	(147)	KLATVN-PQLDFDGHYMPMSKDMHMPVP-----AYPSSDPTTTTAF		
Consensus	(501)	KLAVN LDF MD LS L L L		

FIGURE 6 (continued)

		551	600
A.lyrata 472791	(274)	SYLSQTGFMEPN-----RSSMSLLN	
A.thaliana AT1G26260.1	(329)	SYLSQTGFMQPN-----ISSMSLLS	
A.lyrata 476021	(415)	PHGFMQQTLSSIGR-----TISSPLSSMN	
A.thaliana AT1G68920.1	(414)	PHGFMQQTLSSIGR-----TITSPLSPMN	
G.max Glyma01g09400.1	(446)	HP-PQPGLIHPVIPNMA-----NS---SDILQRTIHPQLAPLN	
G.max Glyma02g13860.1	(430)	HP-PQPGLIHPVIPNMT-----NS---SDILQRTIHPQLAPLN	
M.truncatula AC150842.12.5	(416)	SHSYQPGLIHSTLPNMA-----NL---SDMLRQTVQPQLEPLA	
O.sativa LOC_Os06g16400.1	(367)	PRLHPPKYTQVG-----MPSIVNPTDAFGRVIHAPLGT	
A.lyrata 491205	(260)	P-----MTVVVSPPEMVHSGYSHEMVNS	
Aquilegia sp TC24120	(326)	AYLPLNSLQQ-----MISNCGLDMGMNSS-DISLQRTSAPVS--D	
R.communis TA2848 3988	(341)	AYLQFNFPVQQQQQQQQQLVTCCLDMGINNP-DMGIRRTISAPVSIPE	
S.lycopersicum TC196509	(338)	AQRQFNSLQQ-----IMSSSGLEMGIVNLNEMALRRRTSAPVPIPE	
H.tuberosus TA2737 4233	(272)	SVVSYPGLDMG-----INSAEMELRRSISAPTFIPE	
V.vinifera GSVIVT00014998001	(357)	G-----STSVLGFPGPMSSSHPHYHGISQ	
Z.officinale TA5989 94328	(160)	--SAFSLAQLPLKLAEN--PPDSMFSSLQLPPIDHLANLWEDDLRYVVQT	
A.lyrata 494918	(412)	QQLNMYNSNNAQLLPNI--SSNNMMLQSPMNSLEAS-TLARSFTHLPTLT	
A.thaliana AT5G48560.1	(413)	QQLNIYN-NNSQLLPNI--SSNNMMLQSPMNSLETS-TLARSFTHLPTLT	
A.thaliana AT3G07340.1	(366)	TLLGDHNNKLNQLNPDI--SSN-----NVINPLETS-ETRSFISHLPTLA	
Aquilegia sp TC26724	(200)	FHYGHQGE-GTPVQRE--ISNRTETQSSMNPLGPTLQR-NTSLELPPLD	
G.hirsutum TC144761	(237)	FGLRYQPQQGLPMNNG--ISNDTQIQFSMNPNANAALHK-TQGLQLPPVD	
G.raimondii TC9358	(237)	FGLRYHPQQGLPMTNG--ISNAPQIKFSMNPNANAPLHK-TQRLQLPPVE	
P.persica TC11169	(475)	FPPGYQPQQLPPLHSSS--ISSGTETQFPESPLNAAMRQ-SQGMQLPTFD	
P.trichocarpa 572918	(476)	FPYGYQSQQGLALQNG--MPSNAETQFSMNPLNAALRR-NPSMHLPHLD	
Poptr TA1	(476)	FPYGYQSQQGLALQNG--MPSNAETQFSMNPLNAALRR-NPSMHLPHLD	
G.max Glyma03g21770.1	(424)	AFYGHQPQQNNPVFHNN--IPNRTVTHCSVDPLDTS-LCQNLAMQLSPLD	
G.max GM06MC18997 59691073018643	(354)	TFYGHQPQQNNPVAHNN--IPNRSVTHCSVDPLDTS-LCQNLAMQLPPLN	
P.vulgaris TC12782	(219)	AFYGHQPQQN-PVIHNN--IPNRTVTHCSVDPLDTS-LCQNLALQLPPLN	
G.max Glyma08g46040.1	(492)	PFYGQQTQPN-PAIHNN--IPNGTMTHNSVDPLDTG-MCQNLGMHLPHLN	
G.max Glyma18g32560.1	(486)	PFYGHQPQPN-PAVHNN--IPNGTMTHNSVDPLDTG-LCQNLGMHLPHLN	
M.truncatula AC171266 17.4	(398)	SFYGQQPQQN-QAIHNN--IPNTNGSHCSVEPLDSS-LCHNIGMHLPFLN	
P.trichocarpa 553223	(464)	AAIFSHQQQNNPPLHSN--ISNGAVTHCSVDPLDTTGLCQTLNAQLPPLD	
Triphysaria sp TC7138	(265)	AFYHQQNVQQLQHNSTN--ISGGPLSQCPGDPLSLD--LRNLGIHLFPVD	
S.lycopersicum TC215759	(402)	-----PSG-----S--LHQHVFPVD	
S.tuberosum TC168229	(131)	SYQQGR-----SD--IPNGALSRCSVDTLDNS-LCRSLGMQLPPLD	
C.annuum TC15396	(254)	MQLGGGFERTT-----	
O.sativa LOC_Os08g41320.1	(309)	--SAFRFAEQGDVFQQF--AQNSMESQCTLNQDLALSQATNAAQYAFQD	
O.sativa LOC_Os09g32510.4	(280)	--TAFPFCDQADFFQSF--GLGAMENQCSLDLANTALPHTG-STQYAFQK	
A.lyrata 495154	(234)	-----TS-----FTNTMPTTT	
A.thaliana AT5G50915.1	(233)	-----TS-----FTNAMPTTT	
B.napus TC73386	(244)	-----TS-----FANAMPTTT	
C.annuum NP13055023	(267)	IQQASTT-----TSQAAEVIPNTN--SGYPFLD	
N.tabacum FG645212	(151)	MQQASPTN-----ITSQAAEVIPNINN-SGYPFLLD	
P.hybrida TC3261	(157)	IQQTSITD-----TTPQAAAVIPNIN--NGYPFLD	
G.max Glyma04g37690.1	(264)	VSHCNSPNNQATTF-----ADTTMTPTNIFHIAS----DYLLD	
G.max Glyma06g17420.1	(267)	MSHCNSPNNQATTF-----ADTTMTPTNIFHTTS----DYLLD	
G.max Glyma05g38450.1	(260)	VAQFRPHQAI AFAD-----TITPTTTTNSFPTANN--DGYLLD	
G.raimondii TC6315	(262)	PQQCNPIQTT-----IAPS-----HINYPLLD	
P.trichocarpa 575437	(275)	LQHNSPIQPTAFAD-----TASATTATFATEE----NNYPLID	
P.trichocarpa 773249	(273)	LQQCSPIQPTAFAD-----AAAAATTTATPTTSFATANNYPLID	
M.domestica TC28668	(277)	VPQCNPTQPTPFADTST-----NITATTAVAAAAATTN--NNCPFLD	
C.sinensis TC686	(291)	VQCSPTAQLTAFAD-----TTTAFAP---ADHRNNYPVLD	
O.sativa LOC_Os05g01256.1	(275)	P-----HEMVQCMGQM	
O.sativa LOC_Os08g42470.1	(225)	P-----VLNQSSP---A-----PSQAIMDTTSTT---SYSLQSQHGA	
P.virgatum TC28308	(236)	P-----VLSQASP---A-----PSQAI IETTSTT---PFSLQGQ-GA	
S.bicolor Sb07g025920.1	(226)	P-----VLSQASP---A-----PSQAIMDTTSTT---PYSLQGQ-GS	
Z.mays_ZM07MC24807 BF50298G14024734	(243)	P-----VLSQASPA---A-----PSEAIMDTTTTPY---SLHGHGQ-GS	
Z.mays TC476562	(236)	PAGPAAAVSQASSRP-A-----PSEATMNTACTTTTPYSVSLQAQ-AS	
O.sativa LOC_Os09g33580.1	(211)	S-----VLNRGSSP--A-----QSHAIMDTSNTSPTPYT---LQVQG	
T.aestivum TC301717	(207)	GP----PMNRSSSP--A-----PSQAMMDKGTSSSFQP-----	
S.bicolor Sb02g029530.1	(213)	SS----VLNRPPS-----ETHMDTHNSTSSPS----YEVHG	
Z.mays TC511518	(218)	SS----VLNGTPS-----EALVDTHTSSSSPS----YEVHG	
Z.mays_ZM07MC32776 BFb0337M22032678	(217)	SS----VLNGTPS-----EALVDTHTSSSSPS----YEVHG	
S.officinarum TC78398	(189)	-----	
Z.mays TC460846	(230)	SS----ALSRPPF-----EAIADSTHATASASSPRSYELHG	
Aquilegia sp TC24680	(259)	LSNTVADITASFTTANS-----S----NLLMNSPVSY	
O.sativa LOC_Os04g28280.1	(187)	SYTGSPATADPFTVYN-----CWELDLHTAMQMGATTGLSQDGPIAT	
Consensus	(551)	M	L L

FIGURE 6 (continued)

33/42

		601	650
A.lyrata 472791	(294)	GGLKR-----QETHGYESDHHNVVHMNHETGTAPDHEDATGESKSPIL	
A.thaliana AT1G26260.1	(349)	GGLKR-----QETHGYESDHHNLVHMNHETGTAPDHEDTTADMKVEP-	
A.lyrata 476021	(439)	GGFKR-----QETNGWEGDLQNVIHINYGAGDVTPDPQAAATASLPAA	
A.thaliana AT1G68920.1	(438)	GGFKR-----QETNGWEGDLQNVIHINYGAGDVTPDPQAAATASLPAA	
G.max Glyma01g09400.1	(480)	GGLKEP-----NQLPDVWEDELHNVVQMSFATTAPLTSQDFDGTGPASQM	
G.max Glyma02g13860.1	(464)	GGFKEP-----NQLPDVWEDELHNVVQMSFATTAPPTSQDVDGTGPASQM	
M.truncatula AC150842_12.5	(451)	GGFKEP-----NQVNPLQATKD-----	
O.sativa LOC_Os06g16400.1	(400)	NSAFKEPK---HQMPNNLNGEFQDVIEMPFTHDHGSNDQP-----	
A.lyrata 491205	(282)	GYSNEMVNSGYHHVNPMQQAETSSDPLSCFNNGQAPSMWDSYVQNLGSL	
Aquilegia sp TC24120	(364)	AFLDTSCFTQAQPSSSAWETDLQSLYNVEFNQTRP--VTFPPQSFTGILE	
R.communis TA2848_3988	(390)	SYIDSSCFNQIQ--SSSTWDADLQNLYNVAFDQGRS--TSFPTQFPTGAID	
S.lycopersicum TC196509	(379)	MFLDSSSINQVQ-SFQTWNTDLDNMYAMELQQGRS--AQFLPHPCTGFAE	
H.tuberosus TA2737_4233	(303)	TFMDSSTFX-----	
V.vinifera GSVIVT00014998001	(381)	GTLPGIPTPQFHSTQAVWDGELQSLLMGFDSNPS---SNNLGTNGKCQK	
Z.officinale TA5989_94328	(206)	GFP-----QSQEA--PFSSGALRAE--	
A.lyrata 494918	(459)	QFTDSISQ--YQ---MFSEEDLQSIVGMGVALNPNNESQ-----	
A.thaliana AT5G48560.1	(459)	QFTDSISQ--YQ---MFSEEDLQSIVGMGVAENPNNESQ-----	
A.thaliana AT3G07340.1	(408)	HFTDSISQ--YS---TFSEDDLHSIIHMGFAQNRLQELNQGS----SNQV	
Aquilegia sp TC26724	(245)	GFGE--AAAHFA---NFWEDDLQSVVQMGFNQNSQESGFETHSFQGSLS	
G.hirsutum TC144761	(283)	GISD--ANPQIG---SFLG-----	
G.raimondii TC9358	(283)	GIFE--ATPQIG---SFWEDELQTIQVQMFGGQNHQP-----SCQSKHS	
P.persica TC11169	(522)	RFGG--AAPQAP---QFFEDDLQSVVQMGFGQIQQE-----SLHGSMA	
P.trichocarpa 572918	(522)	GFQDP-AALQAS---AMWEDDLQSVVQMGYGQNHQE-----SFQGSVP	
Poptr TA1	(522)	GFQDP-AALQAS---AMWEDDLQSVVQMGYGQNHQE-----SFQGSVP	
G.max Glyma03g21770.1	(471)	VFNEGGSQ-FPL---AFLEDDLHTIVQMFGGQAANRKTPIQSSSFNGSNN	
G.max GM06MC18997_59691073@18643	(401)	VFNDGGSQ-FPL---AFLEDDLHTIVQMFGGQAANRKTPIQSPSFNGSNN	
P.vulgaris TC12782	(265)	GFSEVGSQ-FPL---TFCEDDLHTIVQMFGGQTANWKTPIQSPSFNGSNN	
G.max Glyma08g46040.1	(538)	GFNEGGFQ-YPI---TFSEDDLHTIVQMFGGQTANRKTPIQSQSFNGSNQ	
G.max Glyma18g32560.1	(532)	DFNEGGSQ-YAK---PFSEDDLHTIIQMFGGQNANRITPIQSQSFNGSNQ	
M.truncatula AC171266_17.4	(444)	GFNEGGSQQYPL---TFSEEDLNTVVQMFGGQTSNRNTPIHSPSFNGSNQ	
P.trichocarpa 553223	(512)	GFTQNAHQ--YP---TFCEDDLQTIQVQMFGGQNP-LETFLPQNFHGSNQ	
Triphysaria sp TC7138	(311)	GFNETLHQ-QFP---SISEDMMQSMIHMGFVQNP-----IG---FPVQNQ	
S.lycopersicum TC215759	(415)	GYAENLAQ--LP---TICEDDLQSIQVQMGFNQKL-----	
S.tuberosum TC168229	(169)	GFAEYLNQ--FP---AICEDDLQSFVQMGFSQNP---TKDMTLQSQGPHQ	
C.annuum TC15396	(265)	-----	
O.sativa LOC_Os08g41320.1	(355)	GTAGANLQ--QR---NFWEDDLQSVFHIENGQSQEN--GVSA PN FHGQQQ	
O.sativa LOC_Os09g32510.4	(325)	-----Q--QR---DLWEDNTFQYNDE---QSQED--AVSAPNFDGQLQ	
A.lyrata 495154	(245)	SIFPSLLD-N-----AIVQEEGEGRE--KFVDRSGFNNNNF	
A.thaliana AT5G50915.1	(244)	PIFPSLLD-NSVVPHT-----AQVQEEGEERE--NFVDRSGFNNNNF	
B.napus TC73386	(255)	TNFSSLLD-DSILPAE-----AHLQEGWRGKG--EDCRQKWVQQQQL	
C.annuum NP13055023	(293)	NSASLMFQ-QAHFPNSIHQGIG--QLLWGAEDQRQ--KIINQSGFSN-NF	
N.tabacum FG645212	(180)	NSASLMFQ-QAHFPNSISQNG--QLLWGADDQRQ--KFINQSGLSN-NF	
P.hybrida TC3261	(185)	NSASLLFQ-QAHFPNSNSQNG--QLLWGADDQRQK-ITIDQSGFSN-NF	
G.max Glyma04g37690.1	(299)	NSVSFFLQ-G---QRSNVFSEEDTGSFWD AEDQRQ--KFLHPCGFSN-NL	
G.max Glyma06g17420.1	(302)	NSVSFFLQ-G---QRPNVLSEEDTGSFWD AEDQRQ--KFLHPYGFNS-NL	
G.max Glyma05g38450.1	(297)	YSSSLFLQ-G---HRPNIFEDN--GGQFWDVEDQRQ--KFLHPYRFGN-NS	
G.raimondii TC6315	(284)	VSTALFLQ-QGLRPNLFSSHQDNGSVLWDVEDQRQ--KIFNSFQLND-NL	
P.trichocarpa 575437	(309)	NSATLFLQ-G-MRPSDFTTHQDSGYLMWDVDEQRQ--KFLNPSGLAT-NL	
P.trichocarpa 773249	(312)	NSTSLLLQ-G-MRPSAFTT-EDSCNLMWDVDERRQ--KFLSPSGLTS-NL	
M.domestica TC28668	(317)	SSASLLLQ-R---PNTFFQDGE--SLLWDVEDQRQ--SFLNPSGFNSNL	
C.sinensis TC686	(323)	TAFLLLQQG-Q--RPNALIYQDNGNSLLWDVEDQKQ--RFLNPSGFTS-NL	
O.sativa LOC_Os05g01256.1	(286)	GSTGISLQGLGGGPTGFAQDGSSHINMVVMQVGEQGGQQGSLHQVEMSSH	
O.sativa LOC_Os08g42470.1	(256)	ISFS-----QDNGS---YLM-QAVGEPRQ-QE--MLNQLVFNNM	
P.virgatum TC28308	(266)	VSFS-----QDNGS---TYLMQQAVGEPT-R-QE--LLNQLVFNNM	
S.bicolor Sb07g025920.1	(256)	ISFS-----QENGSS---TYLMQQAVGEPT-R-QE--LLNQLVFNNM	
Z.mays_ZM07MC24807_BF50298G14@24734	(274)	ISFS-----QDNGSSSTYLMQHQQAAAGEPT-RQ-QE--LLSOLVFNNM	
Z.mays TC476562	(276)	ISFSP-----S-QENG--TYLVQQAVGEQATRQEELLSQLVFSNM	
O.sativa LOC_Os09g33580.1	(243)	GSNNN-----SLSQDNGSY-----IMQTVGEPRQ--E-LFNQVVLNNYM	
T.aestivum TC301717	(234)	-----	
S.bicolor Sb02g029530.1	(241)	DGGTS-----ISFPQDNGSY-----MVQTVGEPRQ--E-LFNQVVFNSHM	
Z.mays TC511518	(246)	DGGTA----SVSFPQDSGSY-----MVQTVGEPRQ--E-LFSQVVFNSHM	
Z.mays_ZM07MC32776_BFb0337M22@32678	(245)	DGGTA----SVSFPQVCLQS-----VRFHLRTSTT--R-DHSGLFVSSKV	
S.officinarum TC78398	(189)	-----	
Z.mays TC460846	(262)	DGGGGGAS--ISFPQDGGGY-----MAQTVGGPRP--E-LFNQVVFNSHM	
Aquilegia sp TC24680	(287)	LLQQAAGQ---R-PNNVLCQQDNGNLLWSVDQQRP---TRFFNQSDFSSL	
O.sativa LOC_Os04g28280.1	(229)	MAPSPSPLPHHPPLHGFYGGQQQQGTTVNHMKAEP-----	
Consensus	(601)	D IV	

FIGURE 6 (continued)

34/42

		651	680
A.lyrata 472791	(337)	SYCCNQILKNR-----	
A.thaliana AT1G26260.1	(391)	-----	
A.lyrata 476021	(482)	NMKVEP-----	
A.thaliana AT1G68920.1	(481)	NMKVEP-----	
G.max Glyma01g09400.1	(525)	KVEL-----	
G.max Glyma02g13860.1	(509)	KVEL-----	
M.truncatula AC150842.12.5	(468)	-----	
O.sativa LOC_Os06g16400.1	(438)	-----	
A.lyrata 491205	(332)	GV-----	
Aquilegia sp TC24120	(412)	ANN-LKMEM-----	
R.communis TA2848.3988	(437)	AGN-LKMEM-----	
S.lycopersicum TC196509	(426)	AGHDLKMEM-----	
H.tuberosus TA2737.4233	(311)	-----	
V.vinifera GSVIVT00014998001	(428)	FI-----	
Z.officinale TA5989.94328	(224)	-----	
A.lyrata 494918	(493)	---HMKIEL-----	
A.thaliana AT5G48560.1	(493)	---HMKIEL-----	
A.thaliana AT3G07340.1	(449)	-PSHMKAE-----	
Aquilegia sp TC26724	(290)	-TANMKVEL-----	
G.hirsutum TC144761	(297)	-----	
G.raimondii TC9358	(321)	-----	
P.persica TC11169	(560)	-SAQMKVEL-----	
P.trichocarpa 572918	(561)	-STHMKIEL-----	
Poptr TA1	(561)	-STQMKIEL-----	
G.max Glyma03g21770.1	(517)	-VPQMKVEL-----	
G.max GM06MC18997.59691073018643	(447)	-VPQMKVEL-----	
P.vulgaris TC12782	(311)	-VPQMKVEL-----	
G.max Glyma08g46040.1	(584)	-EPL-----	
G.max Glyma18g32560.1	(578)	-VPL-----	
M.truncatula AC171266.17.4	(491)	-STQMKIEL-----	
P.trichocarpa 553223	(556)	-VSHMKIEL-----	
Triphysaria sp TC7138	(349)	-TANMKVER-----	
S.lycopersicum TC215759	(444)	-----	
S.tuberosum TC168229	(211)	-TSHMKIEM-----	
C.annuum TC15396	(265)	-----	
O.sativa LOC_Os08g41320.1	(398)	-AGHMKMEF-----	
O.sativa LOC_Os09g32510.4	(358)	AADHTEIEF-----	
A.lyrata 495154	(278)	CSFP-----	
A.thaliana AT5G50915.1	(283)	CSFP-----	
B.napus TC73386	(294)	LFFLLINHMSTQTQKPAFVDSISSNTQGRG	
C.annuum NP13055023	(337)	CSFH-----	
N.tabacum FG645212	(224)	CSFH-----	
P.hybrida TC3261	(230)	CSFH-----	
G.max Glyma04g37690.1	(343)	CSFH-----	
G.max Glyma06g17420.1	(346)	CSFH-----	
G.max Glyma05g38450.1	(339)	GSFN-----	
G.raimondii TC6315	(330)	CFFH-----	
P.trichocarpa 575437	(354)	CSFH-----	
P.trichocarpa 773249	(356)	CSFH-----	
M.domestica TC28668	(359)	CSFN-----	
C.sinensis TC686	(367)	CSFN-----	
O.sativa LOC_Os05g01256.1	(336)	CFFH-----	
O.sativa LOC_Os08g42470.1	(288)	CSFQ-----	
P.virgatum TC28308	(300)	CSFQ-----	
S.bicolor Sb07g025920.1	(290)	CSFQ-----	
Z.mays_ZM07MC24807.BF50298G14024734	(312)	CSFQ-----	
Z.mays TC476562	(316)	CSFQ-----	
O.sativa LOC_Os09g33580.1	(279)	CSFQ-----	
T.aestivum TC301717	(234)	-----	
S.bicolor Sb02g029530.1	(278)	CSFQ-----	
Z.mays TC511518	(284)	CPFQ-----	
Z.mays_ZM07MC32776.BFb0337M22032678	(283)	SWSV-----	
S.officinarum TC78398	(189)	-----	
Z.mays TC460846	(302)	CSFQ-----	
Aquilegia sp TC24680	(330)	CSFQ-----	
O.sativa LOC_Os04g28280.1	(264)	-----	
Consensus	(651)		

FIGURE 6 (continued)

35/42

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1		34	29	34	28	38	31	29	31	29	39	33	43	32	39	50	30	51	34	26	25	28
2	50		42	34	27	33	75	42	30	28	33	35	32	35	33	35	32	31	35	36	34	28
3	42	52		28	32	29	50	95	34	32	28	32	25	30	29	30	27	30	29	44	43	32
4	48	53	41		27	33	32	28	30	27	34	34	33	34	33	34	33	35	35	25	26	27
5	41	41	51	39		27	29	32	58	95	27	30	25	30	26	30	26	30	30	31	30	42
6	53	49	38	48	36		30	27	28	27	87	29	38	34	60	43	35	38	36	24	25	25
7	49	81	61	50	46	45		50	32	31	31	32	30	32	29	30	31	32	32	38	38	31
8	42	52	97	41	49	37	62		34	33	28	32	26	29	28	31	27	30	30	45	43	33
9	45	46	53	45	69	39	49	52		58	28	32	28	33	26	32	27	29	31	33	33	45
10	41	41	50	39	97	37	46	51	69		27	31	26	29	28	29	27	30	30	31	30	42
11	53	49	39	49	36	90	47	38	39	37		28	39	36	61	41	33	39	35	25	25	27
12	47	50	48	50	46	41	49	48	50	46	42		30	34	31	32	28	32	31	31	32	31
13	66	55	41	54	40	56	51	39	44	41	56	47		32	39	45	30	45	34	25	25	25
14	47	50	39	50	39	52	45	39	46	40	52	45	50		34	34	33	31	46	29	30	32
15	55	53	41	50	36	70	45	40	40	37	72	43	60	49		41	34	37	36	24	26	27
16	66	52	40	52	42	58	48	41	43	42	57	44	68	46	59		31	46	37	26	27	31
17	44	45	36	45	34	51	43	37	37	33	49	36	46	49	48	43		31	36	25	26	27
18	67	51	42	45	43	52	49	44	43	44	51	49	62	46	52	62	41		34	27	29	31
19	47	49	38	50	40	51	47	40	43	40	51	43	47	64	52	49	50	46		29	29	32
20	36	45	60	38	51	35	51	62	48	52	37	47	38	38	37	39	33	38	39		87	35
21	38	45	60	38	51	37	50	62	47	52	37	49	39	39	39	38	34	40	40	91		34
22	40	40	49	40	60	34	45	52	63	59	35	47	39	42	37	41	34	41	40	51	53	
23	74	54	40	47	39	54	53	41	43	39	56	46	69	46	56	67	46	67	49	37	39	43
24	67	53	39	49	38	54	49	40	42	38	54	47	64	47	55	64	44	61	47	37	39	42
25	74	54	41	48	40	55	49	41	41	40	56	46	69	45	55	68	43	67	48	39	39	42
26	37	37	47	34	58	32	42	47	55	58	31	41	37	38	33	37	31	36	37	49	49	72
27	36	35	47	35	59	33	43	47	57	60	33	42	35	38	33	37	32	35	37	50	49	73
28	43	47	51	43	55	37	49	53	61	55	40	51	45	49	41	46	38	45	49	51	52	74
29	71	53	40	50	39	57	48	42	42	39	57	45	66	50	58	70	44	64	53	39	40	40
30	47	50	40	51	42	49	46	43	44	44	49	44	51	64	52	54	48	48	90	40	41	42
31	48	50	40	53	38	51	45	40	41	38	49	52	55	50	49	52	46	49	53	37	39	37
32	72	48	41	47	41	53	49	42	45	42	55	45	64	47	54	65	41	67	46	37	38	38
33	41	47	58	43	47	37	53	60	52	48	38	47	41	37	41	43	37	42	40	70	69	48

FIGURE 7

36/42

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
34	42	43	51	41	62	38	49	52	63	63	38	47	43	43	40	42	35	41	43	54	55	72
35	49	42	29	41	30	55	38	29	32	30	53	37	50	49	48	61	44	47	45	28	29	29
36	44	41	35	41	33	50	38	34	37	32	49	39	45	54	46	44	51	42	52	31	32	34
37	50	51	39	49	36	50	45	40	40	36	50	42	54	45	53	57	44	49	46	35	38	38
38	43	45	51	43	45	40	49	53	47	47	39	49	46	40	43	43	38	45	41	54	53	45
39	43	49	46	47	45	42	53	46	47	45	41	50	44	50	45	43	41	46	47	45	45	45
40	56	49	37	46	37	60	45	35	41	37	60	39	57	51	53	57	50	53	53	35	35	37
41	45	50	43	50	42	45	50	45	46	41	47	48	47	51	48	48	43	49	50	44	45	43
42	54	48	34	45	36	58	43	33	39	36	56	40	57	52	49	57	50	51	54	31	33	33
43	50	42	30	41	31	55	40	30	34	31	53	37	51	50	47	59	49	50	45	27	29	28
44	37	38	49	37	55	35	44	51	54	54	35	44	35	41	35	36	33	39	41	53	51	63
45	40	39	48	37	61	36	42	50	62	61	34	44	36	39	35	37	31	36	39	51	50	71
46	36	39	49	38	55	34	43	51	55	55	34	44	37	41	36	36	31	39	43	52	49	63
47	86	47	42	49	40	55	51	41	46	41	56	46	67	44	56	66	43	69	47	35	39	38
48	54	52	36	49	36	55	44	37	40	35	54	41	56	49	51	60	49	55	52	35	34	36
49	46	54	39	52	43	48	47	41	48	43	48	47	51	68	49	51	47	45	63	40	40	52
50	40	44	49	50	47	40	48	46	51	47	38	70	41	42	40	41	35	43	41	47	46	45
51	54	48	34	45	35	58	42	35	40	35	56	40	57	52	48	54	48	52	55	31	32	35
52	52	53	36	50	37	59	43	36	42	36	58	40	58	51	52	58	51	54	54	34	34	36
53	45	47	48	50	47	40	48	47	51	46	38	68	44	43	40	44	36	44	45	45	48	49
54	44	43	48	45	53	40	46	49	56	55	39	51	42	39	41	42	36	43	40	46	47	55
55	38	38	27	38	28	41	34	28	28	27	42	29	41	37	37	39	46	38	43	24	25	25
56	37	39	29	39	32	43	33	31	35	32	44	35	39	58	42	39	44	36	50	31	31	33
57	45	41	31	44	35	54	42	31	35	34	53	36	49	45	50	49	49	46	51	28	29	31
58	36	39	49	38	55	34	43	51	55	55	34	44	38	41	36	36	31	38	43	52	49	63
59	50	53	42	52	47	46	50	43	50	47	45	50	50	56	47	54	46	53	55	42	43	45
60	47	47	50	46	50	40	51	50	51	48	41	49	44	41	43	45	39	46	43	49	48	47
61	54	49	36	46	37	51	43	36	39	38	52	40	53	49	49	51	46	48	53	32	33	33
62	53	51	37	48	37	50	46	37	40	37	53	43	57	51	54	57	46	50	48	32	33	36
63	54	48	35	45	35	57	42	36	39	35	56	40	57	56	50	52	50	52	53	33	34	35
64	55	49	39	50	36	52	45	38	41	36	55	43	59	51	52	57	46	54	51	34	35	37
65	51	47	35	48	36	55	42	36	38	37	57	40	53	56	51	51	50	51	53	34	34	35
66	40	42	31	44	32	47	37	32	38	33	48	37	42	58	44	41	47	37	52	32	34	32

FIGURE 7 (continued)

37/42

	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44
1	54	50	54	25	27	29	53	35	34	54	26	28	40	33	34	32	33	41	32	38	41	26
2	32	31	34	27	27	32	34	35	32	31	35	30	29	30	32	33	31	32	35	31	31	29
3	29	28	29	30	31	32	29	28	28	25	40	32	21	25	29	36	33	26	31	24	22	35
4	34	31	35	25	25	31	31	34	38	33	28	28	31	30	32	31	33	31	33	32	31	26
5	27	25	27	43	44	38	29	32	28	29	29	41	22	24	26	28	32	27	30	25	25	39
6	42	40	41	23	24	28	41	35	37	38	27	28	37	33	36	28	29	43	32	39	39	25
7	33	28	31	29	30	32	31	31	32	28	38	31	26	28	31	33	32	30	33	28	27	33
8	29	28	29	31	32	34	29	30	28	27	40	32	21	25	29	35	32	26	32	25	22	36
9	31	28	30	43	44	39	31	33	31	30	32	43	26	29	30	31	32	27	31	26	27	38
10	28	25	28	43	44	38	29	32	28	29	30	42	22	25	26	31	33	27	31	25	24	40
11	42	41	41	23	25	29	40	34	35	40	27	30	37	32	35	29	30	41	33	39	38	25
12	34	30	33	29	31	33	30	34	39	31	30	32	27	28	29	32	35	29	34	27	27	30
13	47	42	46	24	24	28	45	32	35	45	26	28	35	30	36	28	28	40	32	39	38	26
14	33	33	34	29	29	37	34	49	33	32	26	31	39	39	31	27	36	33	36	33	39	32
15	41	38	41	25	25	31	41	35	33	38	27	29	34	33	36	30	33	38	34	36	35	25
16	47	46	50	26	27	33	51	38	36	48	26	31	56	31	36	29	30	42	32	39	55	27
17	33	30	30	24	26	30	31	36	34	29	27	26	28	34	31	28	32	33	34	33	29	24
18	52	46	51	28	26	32	47	35	34	53	28	30	37	31	36	31	32	39	31	37	40	29
19	35	34	36	30	29	37	35	88	36	30	30	33	36	38	34	28	37	34	39	36	37	33
20	26	25	26	34	32	35	27	29	26	26	56	34	21	23	25	37	33	24	32	23	22	35
21	27	26	27	33	33	37	30	30	27	26	55	34	22	24	26	37	34	26	32	24	23	35
22	30	28	29	63	63	66	29	34	28	26	30	55	23	25	27	29	31	25	28	23	22	44
23		63	94	26	27	32	54	36	34	56	28	31	42	33	35	28	32	39	33	38	41	26
24	78		64	27	27	30	50	34	33	52	27	28	40	29	34	30	30	39	33	36	39	25
25	96	78		26	27	32	55	38	33	57	27	31	42	30	36	30	31	38	33	38	42	26
26	36	38	37		90	58	26	31	25	25	30	60	20	24	26	30	29	22	28	22	21	47
27	36	38	37	93		58	27	31	26	25	30	59	20	24	25	31	30	23	28	22	23	47
28	43	45	44	65	65		34	39	31	29	30	64	26	30	28	32	37	29	33	28	29	44
29	73	66	73	35	36	47		36	35	51	29	29	42	33	38	30	30	40	32	40	43	26
30	48	48	50	38	38	49	54		35	32	29	35	36	38	33	29	38	35	39	35	36	33
31	49	51	46	34	33	43	51	52		33	28	29	31	32	31	32	33	35	34	31	31	26
32	72	65	72	35	35	46	67	45	48		27	29	38	30	36	30	30	38	31	35	40	26
33	41	41	41	44	46	51	43	41	40	42		31	23	25	28	34	30	25	32	24	22	33

FIGURE 7 (continued)

38/42

	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44
34	41	40	42	69	70	76	39	45	41	44	50		23	26	27	32	32	25	31	24	23	45
35	52	50	52	26	27	34	53	46	44	50	32	28		38	33	25	26	39	28	39	83	19
36	46	46	45	30	29	39	44	48	46	41	35	34	54		31	27	32	36	34	34	37	23
37	51	51	54	35	34	41	52	46	46	51	40	40	43	44		30	31	39	32	35	32	26
38	41	42	44	44	44	50	43	42	44	43	53	47	34	37	45		31	27	31	27	25	33
39	44	44	43	42	42	56	44	50	44	45	47	49	34	39	45	46		29	57	30	28	33
40	56	55	56	33	33	44	57	49	53	50	38	38	53	51	55	41	42		33	63	39	24
41	45	48	46	38	39	49	47	53	50	46	43	46	39	45	46	45	69	44		32	31	30
42	55	53	54	31	32	38	55	50	49	51	37	35	55	47	51	38	42	78	45		41	22
43	51	48	50	28	28	35	54	45	46	50	31	30	91	55	44	33	36	53	40	56		21
44	36	36	37	66	67	58	38	43	35	37	44	61	27	30	36	46	46	34	40	31	27	
45	37	38	38	77	77	63	37	42	36	36	48	69	26	32	35	45	44	35	42	33	28	67
46	37	37	36	65	67	58	37	46	36	38	45	62	25	30	36	46	45	34	41	31	27	79
47	76	69	76	36	37	46	72	47	47	74	41	42	50	44	54	43	45	54	46	53	49	37
48	57	55	57	33	33	41	57	50	55	54	39	40	52	47	54	41	44	86	46	78	52	33
49	50	49	50	45	46	65	51	64	52	48	39	54	48	54	47	45	50	53	51	48	48	42
50	41	41	41	40	41	49	41	41	50	43	49	46	34	34	41	50	46	37	45	37	32	42
51	53	50	52	31	31	39	51	50	49	52	35	36	56	52	50	40	40	75	43	86	53	33
52	56	53	54	32	33	39	59	53	51	54	36	39	54	46	53	39	43	87	45	79	54	31
53	44	42	45	40	42	50	45	47	52	48	47	46	34	38	41	47	48	40	42	38	33	42
54	42	43	43	54	55	59	42	42	40	45	47	56	31	33	44	48	47	40	43	38	31	53
55	38	37	37	24	24	30	39	41	41	39	28	27	46	42	38	31	31	47	34	53	45	25
56	40	39	40	30	30	39	40	49	40	34	30	35	57	48	38	30	38	44	40	46	55	28
57	46	46	45	29	26	35	47	47	47	45	33	31	55	52	43	36	37	64	43	70	60	29
58	37	37	37	65	67	58	37	46	36	38	45	62	25	30	36	46	45	34	41	31	27	79
59	51	50	53	43	44	56	51	58	50	50	44	52	39	44	47	43	52	49	55	48	40	43
60	45	44	46	43	43	53	44	46	43	47	50	49	35	36	47	55	48	40	47	38	35	47
61	51	50	51	31	32	40	49	53	49	51	37	37	50	51	53	40	42	70	45	75	49	34
62	57	53	56	32	32	39	54	50	51	50	37	39	47	48	50	37	44	79	47	69	48	34
63	55	51	54	33	33	40	53	51	51	53	37	36	53	54	51	38	41	74	45	82	52	32
64	54	53	52	33	34	39	58	53	53	54	38	37	51	49	52	40	44	82	45	72	51	32
65	52	49	52	32	32	39	50	50	52	51	36	37	50	54	49	39	41	67	44	73	50	32
66	42	44	41	29	28	38	42	51	44	38	33	34	60	57	42	33	41	50	43	50	57	30

FIGURE 7 (continued)

39/42

	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66
1	29	26	77	39	32	30	38	38	31	29	31	27	35	26	33	31	38	35	37	38	37	30
2	27	28	33	34	34	30	34	33	32	29	30	28	30	28	35	30	33	30	33	33	32	33
3	33	33	29	26	28	30	25	27	31	29	22	22	25	33	29	34	25	25	24	27	25	26
4	25	26	32	31	33	39	32	32	35	29	29	28	33	26	35	32	30	33	30	29	32	31
5	46	41	28	27	31	30	26	27	32	38	20	24	26	41	34	31	27	26	25	24	24	24
6	24	23	39	41	34	29	40	41	31	28	32	28	39	23	31	30	36	36	39	37	37	34
7	31	31	34	32	30	32	31	31	32	29	26	24	29	31	32	32	30	31	29	29	30	27
8	35	34	30	25	29	30	24	26	30	29	21	23	25	34	31	36	25	25	24	28	24	24
9	47	42	31	29	36	34	28	28	33	39	22	28	24	42	37	31	28	26	28	29	27	28
10	46	41	27	26	31	32	26	26	30	39	21	25	24	41	35	32	27	25	25	26	25	25
11	25	24	42	41	32	26	39	41	29	29	33	29	39	24	32	30	36	36	37	40	37	33
12	30	30	31	28	33	56	27	30	49	29	22	28	26	30	35	31	28	30	27	29	26	28
13	24	24	46	39	32	29	38	41	28	27	30	29	35	24	33	29	37	36	37	38	35	30
14	30	33	32	33	52	29	34	34	29	26	29	45	34	33	40	28	35	32	37	32	37	45
15	25	25	39	36	33	28	37	40	27	26	30	29	36	25	34	28	35	35	37	36	38	33
16	27	27	47	42	35	31	39	43	30	29	31	29	38	27	35	29	36	36	37	40	35	33
17	24	23	29	33	32	28	32	33	29	28	35	30	35	23	33	29	33	31	32	33	32	33
18	28	28	53	41	34	31	37	39	31	28	30	28	33	28	34	29	36	34	34	37	34	30
19	31	36	35	34	44	32	36	35	34	29	32	36	38	36	43	32	34	32	35	32	35	39
20	35	33	26	26	29	29	23	23	30	30	20	24	21	33	29	35	23	23	24	24	24	26
21	33	33	27	25	31	30	24	25	30	30	20	25	21	33	30	33	24	24	25	25	25	27
22	55	43	28	23	47	31	23	24	29	40	20	28	22	43	33	30	25	26	25	24	25	24
23	29	26	59	40	34	29	38	41	31	29	31	31	34	26	34	31	37	39	37	38	36	32
24	26	25	53	38	35	28	36	38	30	28	30	28	33	25	34	31	36	35	35	38	33	31
25	28	27	59	39	34	29	38	39	31	28	30	30	32	27	35	31	36	36	37	37	36	31
26	62	46	26	24	41	28	22	23	28	42	20	25	22	46	35	32	22	23	23	24	23	23
27	62	47	27	24	41	29	23	23	28	43	19	24	20	47	35	31	23	22	24	24	23	22
28	51	41	32	29	62	32	28	29	31	37	24	32	27	41	43	32	29	28	28	27	28	29
29	27	27	53	39	33	30	37	40	30	29	31	29	34	27	34	32	38	36	35	37	34	33
30	33	37	35	33	47	31	35	34	33	29	31	39	34	37	44	33	34	31	35	32	35	40
31	28	26	33	32	32	39	34	32	39	28	31	29	34	26	35	31	35	31	34	32	34	33
32	27	27	57	38	32	30	34	37	30	28	29	26	34	27	33	29	34	36	35	35	34	30
33	32	30	28	26	26	29	25	26	28	32	22	23	25	30	28	34	25	23	25	26	24	27

FIGURE 7 (continued)

40/42

	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66
34	52	45	30	27	48	30	25	27	30	39	22	28	23	45	40	32	27	26	24	26	25	25
35	21	20	38	39	37	25	39	39	26	22	29	41	36	20	31	25	37	36	37	35	36	42
36	25	23	32	34	40	27	36	34	28	22	31	39	36	23	35	27	37	34	36	34	35	42
37	25	25	34	39	28	28	36	37	27	30	30	27	34	25	32	31	37	36	37	38	35	30
38	31	31	32	28	29	31	30	28	31	30	25	24	27	31	31	33	30	26	28	29	28	25
39	33	33	32	32	35	32	28	31	34	29	26	30	29	33	40	30	29	29	29	29	28	34
40	26	23	39	76	35	27	60	75	28	25	41	31	51	23	33	27	55	65	57	69	50	35
41	31	30	32	34	34	33	32	33	31	29	29	31	31	30	39	31	33	33	32	31	32	36
42	22	22	37	61	33	25	67	62	29	24	44	30	58	22	33	28	59	52	62	56	54	35
43	23	20	41	38	36	27	39	39	24	23	31	41	39	21	31	25	38	36	38	34	38	42
44	49	65	27	24	31	29	24	23	28	39	21	23	22	65	33	33	25	24	22	24	22	23
45		49	26	25	38	30	24	24	29	43	20	27	23	48	37	30	25	25	25	24	24	25
46	69		26	24	31	29	23	24	29	41	19	25	22	100	33	31	23	24	23	24	23	24
47	36	36		40	33	28	36	39	30	29	31	28	37	25	33	29	35	35	36	37	36	31
48	33	33	55		33	28	59	87	28	25	43	28	48	24	34	30	53	72	56	79	50	33
49	47	42	46	52		31	34	33	31	28	29	45	32	31	45	28	32	32	34	30	34	40
50	44	43	40	40	43		26	27	50	28	23	25	25	29	33	31	27	25	28	27	27	27
51	33	32	54	74	48	36		60	27	25	61	32	58	23	33	30	79	51	86	56	76	34
52	34	33	55	92	52	38	76		27	24	40	29	51	24	32	29	54	75	55	82	50	33
53	44	43	43	41	45	70	37	41		29	23	26	26	29	33	30	30	27	28	26	27	26
54	56	54	45	38	41	47	37	38	47		23	24	24	41	30	31	25	24	24	25	24	25
55	25	24	40	49	38	29	64	48	29	31		26	53	19	27	24	52	35	56	41	56	30
56	32	30	36	41	55	32	46	44	33	31	42		29	24	38	25	30	30	32	29	33	48
57	30	28	44	63	42	32	70	65	34	33	63	48		22	31	28	52	43	55	48	54	34
58	69	100	36	33	42	43	32	33	42	54	24	30	28		33	31	23	24	23	24	23	24
59	46	43	52	48	62	47	47	48	49	47	36	48	42	43		30	32	31	33	33	32	35
60	44	44	47	44	42	48	41	41	50	50	30	33	36	44	48		26	29	28	30	27	27
61	33	32	51	72	50	39	84	70	39	38	55	41	63	32	49	40		49	75	51	67	34
62	35	34	55	84	50	37	66	85	41	39	44	42	56	34	52	42	65		48	70	43	32
63	33	31	54	71	50	39	92	75	40	37	60	44	67	31	48	40	83	65		53	90	36
64	35	33	54	87	50	39	71	87	42	39	48	41	61	32	48	45	68	83	69		48	30
65	33	31	51	65	52	39	84	69	41	38	61	46	65	31	47	41	77	59	93	64		33
66	32	31	39	45	54	35	49	44	34	33	43	68	53	31	47	34	46	42	50	44	49	

FIGURE 7 (continued)

41/42

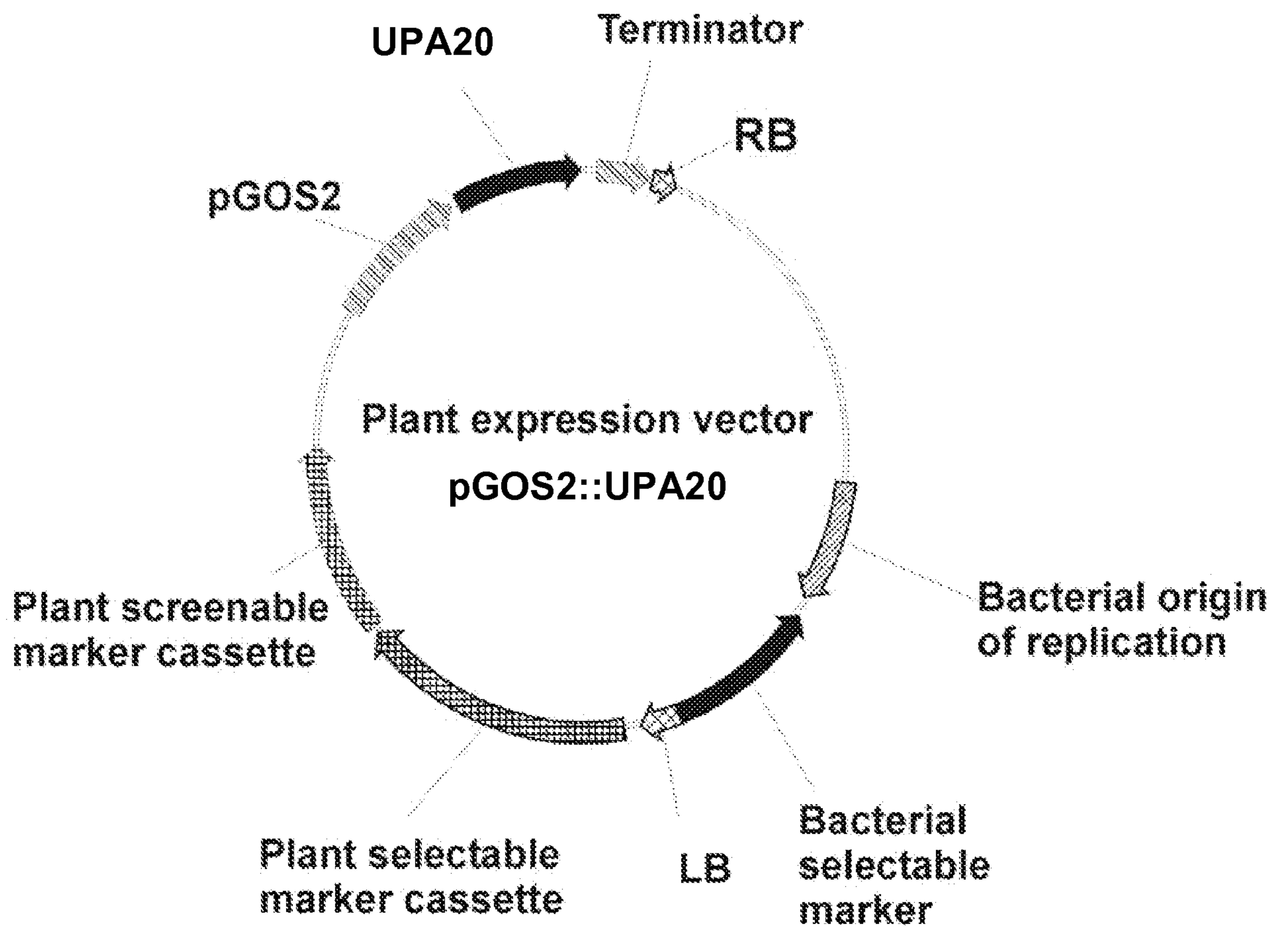


FIGURE 8

