



(51) International Patent Classification:

C12N 15/82 (2006.01) *C12N 9/18* (2006.01)
A01H 5/00 (2006.01) *C12N 9/90* (2006.01)
C07K 14/415 (2006.01) *C12N 15/61* (2006.01)
C12N 9/16 (2006.01)

(21) International Application Number:

PCT/IB2012/050878

(22) International Filing Date:

27 February 2012 (27.02.2012)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

61/447,119	28 February 2011 (28.02.2011)	US
61/447,127	28 February 2011 (28.02.2011)	US
11156244.3	28 February 2011 (28.02.2011)	EP
11156187.4	28 February 2011 (28.02.2011)	EP
61/478,068	22 April 2011 (22.04.2011)	US
11163577.7	22 April 2011 (22.04.2011)	EP

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

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

(72) Inventors; and

(75) Inventors/Applicants (for US only): **LOUWERS, Marieke** [BE/BE]; Prinses Clementinalaan 9, B-9000 Gent (BE). **REUZEAU, Christophe** [FR/FR]; rue de Cimetière, F-24350 La Chapelle Gonaguet (FR). **SANZ MOLINERO, Ana Isabel** [ES/ES]; Isla Malaita 1, 6B,

28035, Madrid (ES). **HATZFELD, Yves** [FR/FR]; 6B rue du Chemin de Fer, F-59000 Lille (FR).

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

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, 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))

(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 DUF642 (Protein containing a Domain of Unknown Function) polypeptide, or an epimerase-related like polypeptide, or a phospholipase/carboxylesterase (PLPCase) polypeptide, and plants having modulated expression of a nucleic acid encoding a DUF642 polypeptide, or an epimerase-related like polypeptide, or a PLPCase polypeptide, which plants have enhanced yield-related traits relative to control plants. Also provided are DUF642-encoding nucleic acids, or epimerase-related like polypeptides, or PLPCase-encoding nucleic acids, and constructs comprising the same, useful in performing the methods for enhancing yield-related traits in plants.



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 DUF642 (Protein containing a Domain of Unknown Function) polypeptide. The present invention also concerns plants having modulated expression of a nucleic acid encoding a DUF642 polypeptide, which plants have enhanced yield-related traits relative to corresponding wild type plants or other control plants. The invention also provides constructs useful in the methods of the invention.

The present invention furthermore concerns a method for enhancing yield-related traits in plants by modulating expression in a plant of a nucleic acid encoding an epimerase-related like polypeptide. The present invention also concerns plants having modulated expression of a nucleic acid encoding an epimerase-related like polypeptide, which plants have enhanced yield-related traits relative to corresponding wild type plants or other control plants. The invention also provides constructs useful in the methods of the invention.

The present invention also concerns a method for enhancing yield-related traits in plants by modulating expression in a plant of a nucleic acid encoding a Phospholipase/carboxylesterase (PLPCase) polypeptide. The present invention also concerns plants having modulated expression of a nucleic acid encoding a PLPCase polypeptide, which plants have enhanced yield-related traits relative to corresponding wild type plants or other control plants. The invention also provides constructs useful in the methods of the invention.

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

A trait of particular economic interest is increased yield. Yield is normally defined as the measurable produce of economic value from a crop. This may be defined in terms of quantity and/or quality. Yield is directly dependent on several factors, for example, the

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

5

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.

10 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
15 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.

20 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
25 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
30 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.

35

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

Concerning DUF642 polypeptides, the family of DUF642 proteins represents a conserved region found in a number of uncharacterised plant proteins. DUF642 proteins were
40 identified e.g. by Irshad et al., *BMC Plant Biology* 2008, 8:94.

Concerning epimerase-related like polypeptides, the present invention relates to the use of epimerase-related like polynucleotides and polypeptides in the field of plant biotechnology. Epimerases can be defined as isomerase enzymes that catalyze the inversion of stereochemistry in biological molecules.

5

Ascencio-Ibáñez et al. (2008, Plant Physiology 148(1):436-454) have reported amongst genes that are significantly changed during CaLCuV injection an Arabidopsis epimerase gene, which is listed in a supplemental table (Table 2) of this paper. In this paper, the Arabidopsis gene is called a "Putative protein Isopenicillin-N-epimerase". The authors described microarray analyses of the Arabidopsis (*Arabidopsis thaliana*) transcriptome in response to cabbage leaf curl virus (CaLCuV) infection, and indicated that they uncovered 5,365 genes (false discovery rate <0.005) differentially expressed in infected rosette leaves at 12 days postinoculation, including an Arabidopsis isomerase gene (At3g62130).

10

15 Mehta and Christen (1994 Biochem. Biophys. Res. Commun. 14;198(1):138-43) described the homology of isopenicillin-N-epimerases to aminotransferases.

Concerning PLPCase polypeptides, the PLPCase polypeptide family contains both phospholipases and carboxylesterases with broad substrate specificity, and is structurally related to alpha/beta hydrolases. Members of the PLPCase superfamily are present in plants, but also in other organisms such as yeast and bacteria. In Arabidopsis, SOBER1 gene encoding a PLPCase is involved in resistance to plant pathogen *P. syringae* pv tomato DC3000 (Kirik and Mudgett (2009), PNAS 106 (48): 20532-20537). Recent studies showed that PLPCase might have arisen from rare horizontal gene transfer from fungi to plant (Richards et al. (2009), The Plant Cell, 21(7): 1897-1911).

20

25

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.

30

35 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 DUF642 polypeptide, or an epimerase-related like polypeptide, or a PLPCase polypeptide in a plant.

Detailed description of the invention

40

The present invention shows that modulating expression in a plant of a nucleic acid encoding a DUF642 polypeptide, or an epimerase-related like polypeptide, or a PLPCase 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 DUF642 polypeptide, or an epimerase-related like polypeptide, or a PLPCase polypeptide and optionally selecting for plants having enhanced yield-related traits. According to another embodiment, the present invention provides a method for producing plants having enhanced yield-related traits relative to control plants, wherein said method comprises the steps of modulating expression in said plant of a nucleic acid encoding a DUF642 polypeptide, or an epimerase-related like polypeptide, or a PLPCase polypeptide as described herein and optionally selecting for plants having enhanced yield-related traits.

A preferred method for modulating (preferably, increasing) expression of a nucleic acid encoding a DUF642 polypeptide, or an epimerase-related like polypeptide, or a PLPCase polypeptide is by introducing and expressing in a plant a nucleic acid encoding a DUF642 polypeptide, or an epimerase-related like polypeptide, or a PLPCase polypeptide.

Any reference hereinafter to a "protein useful in the methods of the invention" is taken to mean a DUF642 polypeptide, or an epimerase-related like polypeptide, or a PLPCase 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 DUF642 polypeptide, or an epimerase-related like polypeptide, or a PLPCase 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 "*DUF642* nucleic acid", or "*epimerase-related like* nucleic acid", or "*PLPCase* nucleic acid" or "*DUF642* gene", or "*epimerase-related like* gene", or "*PLPCase* gene".

A "DUF642 polypeptide" as defined herein refers to any polypeptide comprising an InterPro accession IPR006946 DUF642 domain, preferably said DUF642 polypeptide comprises the signature sequence FSAARTCAQ (SEQ ID NO: 194).

The term "DUF642" or "DUF642 polypeptide" as used herein also intends to include homologues as defined hereunder of "DUF642 polypeptide".

Preferably, the DUF642 polypeptide comprises one or more of the following motifs:

- (i) Motif 1: N[LAM][VL]KNG[DG]FEEGP[WYH][MI][FLI]P[NG][TS][STR][WL]GVL[LVI]P[PTS][NKM][LQDV][EDV][DE][ED][THY]S[PS]L[PS][GP]W[IMT][IV] (SEQ ID NO: 181),
- (ii) Motif 2: [VLA][EKAT][KPR]G[SAM][LRH]Y[SA][LIV][TI]F[SG]A[AS]RTCAQ[DLAS]E[SVRL]L[NR][VI] (SEQ ID NO: 182),
- (iii) Motif 3: D[PE][AT]CGP[LI][IL]D[AS][VIF]AI[KR] (SEQ ID NO: 183)

More preferably, the DUF642 polypeptide comprises one or more of the following motifs:

- (i) Motif 4: N[ML][LV][KR]NG[DG]FEEGPY[IF][FLI]P[GND][TA][SPR]WGV[LVI]P[PS][MN][DIV][EV] (SEQ ID NO: 184),
- 5 (ii) Motif 5: N[VI][ST][VAI][SAT][PG][EQ][SW][GA][VE][LI]P[IM]QT[VIM]Y[TGS]S[SN]GWDSY[SA]WA[FW] (SEQ ID NO: 185),
- (iii) Motif 6: Y[SA][IL][TI]FSAARTCAQ[AS]E (SEQ ID NO: 186)

Even more preferably, the DUF642 polypeptide comprises in increasing order of
10 preference, at least 2, at least 3, at least 4, at least 5, or all 6 motifs.

Motifs 1 to 6 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
15 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.

Additionally or alternatively, the homologue of a DUF642 protein has in increasing order of
preference at least 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%,
20 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
25 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).

30 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 SEQ ID
NO: 4.

Compared to overall sequence identity, the sequence identity will generally be higher when
35 only conserved domains or motifs are considered. Preferably the motifs in a DUF642
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: 181 to SEQ ID NO: 186 (Motifs 1 to 6).

40 In other words, in another embodiment a method is provided wherein said DUF642
polypeptide comprises a conserved domain (or motif) with at least 70%, 71%, 72%, 73%,

74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the conserved motifs starting with amino acid 202 up to amino acid 241 (motif 1), amino acid 108 up to amino acid 132 (motif 2), amino acid 178 up to amino acid 191 (motif 3) in SEQ ID NO:2; the conserved motifs starting with amino acid 202 up to amino acid 230 (motif 4), amino acid 130 up to amino acid 159 (motif 5), amino acid 114 up to amino acid 128 (motif 6) in SEQ ID NO:2; the conserved motifs starting with amino acid 121 up to amino acid 160 (motif 1), amino acid 27 up to amino acid 51 (motif 2), amino acid 97 up to amino acid 110 (motif 3) in SEQ ID NO:4; the conserved motifs starting with amino acid 121 up to amino acid 149 (motif 4), amino acid 50 up to amino acid 78 (motif 5), amino acid 33 up to amino acid 47 (motif 6) in SEQ ID NO:4.

An "epimerase-related like polypeptide" or "epimerase-related like protein" as defined herein refers to any polypeptide having an amino acid sequence which comprises one or more of the following motifs:

- (i) Motif 7: CREEGV[DE][QK]VFVD[AG]AH[AS]IG[QSC]V[PDE][VI][DN][VM][KR][ED]IGADFY[TV]SNLHKWFFCAPP[SA]VAFL[YH] (SEQ ID NO: 273),
- (ii) Motif 8: EF[SA]HH[DN]P[GAN]VAR[IV]NNGSFG[CS]CP[AG]S[VI][LI]AAQ[ARK][RN]WQ[LR][LRQ]FL[RQA]QPD[AD]FYF[ND]xL[QRK][PK]G (SEQ ID NO: 274),
- (iii) Motif 9: S[LI]VDNATTAAAIVLQ[HQ][VAI][AG][WR][AS]FAEG[RKN]FA[KR][GN]D[AVT]V[LV]MLH[YC]AY[GQ][AS]VKKSI[EQH]AYV (SEQ ID NO: 275)

Motifs 7 to 9 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 epimerase-related like polypeptide comprises in increasing order of preference, at least 1, at least 2, or all 3 motifs.

In another embodiment, an epimerase-related like polypeptide as used herein comprises a polypeptide having one or more of the domains as listed in Table C, and in particular having one or more of the following domains:

- a domain as determined with the superfamily database and having accession number SSF53383;
- a domain as determined with the HMMPanther database and having accession number as PTHR11601
- a domain as determined with the HMMPfam database and having accession number PF00266; and
- a domain as determined with the Gene3D database and having accession number G3DSA:3.40.640.10

In another preferred embodiment, an epimerase-related like polypeptide as used herein 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 consists of a conserved domain as represented by any one of:

- amino acid coordinates 32 to 449 in SEQ ID NO: 198 (motif 10 – SEQ ID NO: 280);
- amino acid coordinates 47 to 276 in SEQ ID NO: 198 (motif 11 – SEQ ID NO: 281);
- amino acid coordinates 92 to 364 in SEQ ID NO: 198 (motif 12 – SEQ ID NO: 282);
- and
- amino acid coordinates 59 to 323 in SEQ ID NO: 198 (motif 13 – SEQ ID NO: 283).

The term “epimerase-related like polypeptide” as used herein also intends to include homologues as defined hereunder of an “epimerase-related like polypeptide”.

Additionally or alternatively, the homologue of an epimerase-related like protein has in increasing order of preference at least 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: 198, provided that the homologous protein comprises any one or more of the conserved motifs or domains 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: 198.

Compared to overall sequence identity, the sequence identity will generally be higher when only conserved domains or motifs are considered. Preferably the motifs in an epimerase-related 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 1 to 7 (as represented by SEQ ID NO 273, 274, 275, 280, 281, 282, 283, respectively) as described above.

A “PLPCase polypeptide” as defined herein refers to any polypeptide comprising an InterPro accession IPR003140 corresponding to PFAM accession number PF022030

Phospholipase/carboxylesterase (PLPCase) domain.

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

5

In a preferred embodiment, the PLPCase polypeptide comprises one or more of the following motifs:

- (i) Motif 14: E[FY]G[KR]T[HY]VVRPKG[KR]HQATIVWLHGLGDNG[LSA]S[SW]SQLL[ED][ST]LPLPNIKWICPTA (SEQ ID NO: 348),
- 10 (ii) Motif 15: PDD[WIVL]EGLDASAAH[IV]ANLLS[TS]EP[AS]D[VI]K[VL]G[IV]G (SEQ ID NO: 349),
- (iii) Motif 16: FSMGAA[IT]ALYSA[TA]C[YF]A[MHL] (SEQ ID NO: 350)

In a more preferred embodiment, the PLPCase polypeptide comprises one or more of the following motifs:

- 15 (i) Motif 17: E[FY]G[KR]T[HY]VVRPKGKH[QK]ATIVWLHGLGDNG[LS]S[SW]SQLLE[ST]LPLPNIKWICPTA (SEQ ID NO: 351)
- (ii) Motif 18: ED[GA]PDD[WLV]EGLDA[SA]AAH[IV]ANLLSTEPA[DN][VI]K (SEQ ID NO: 352)
- (iii) Motif 19: [LF]GGFP[CS]TAWFD (SEQ ID NO: 353)

20

In an alternative preferred embodiment, the PLPCase polypeptide comprises one or more of the following motifs:

- (i) Motif 20: T[HY]VVRPKG[KR]HQATIVWLHG[LI]GDNG[LAG]S[SW]SQLL[ED]SLPLPN[IV]KWICPTAP[TS]RP (SEQ ID NO: 354)
- 25 (ii) Motif 21: FPCTAWFDV[GE][ED][LT]S[ELV]DG[PHR]DD[WI]EG[LM]DASA[AS]H[IV]ANLLS[TS]EP[AS]DV[KS][VL]GIGGFSM (SEQ ID NO: 355)
- (iii) Motif 22: FK[SP]Y[END]G[IL]GHYT[VI]P[RE]EM[GD][ED]V[SC][TKN] (SEQ ID NO: 356)

30 In a most preferred embodiment, the PLPCase polypeptide comprises one or more of the following motifs:

- (i) Motif 23: EFG[KR]T[HY]VVRPKGKHQATIVWLHGLGDNG[LS]SS[SY]QLLES LPLPNIKWICPTA (SEQ ID NO: 357)
- (ii) Motif 24: SEDG[PH]DDWEGLDASA[AS]HIANLLSTEPADV (SEQ ID NO: 358)
- 35 (iii) Motif 25: G[IT]SDDVV[LP]Y[KR][YF]GEKSAQSLS[SM]AGFRY[VI][AMT]FK[SP]Y (SEQ ID NO: 259)

Motifs 14 to 25 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.

40

More preferably, the PLPCase polypeptide comprises in increasing order of preference, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, or all 12 motifs.

5

Additionally or alternatively, the homologue of a PLPCase 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: 285, 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 PLPCase 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: 348 to SEQ ID NO: 359 (Motifs 14 to 25).

In another preferred embodiment a method is provided wherein said PLPCase polypeptide comprises a conserved domain (or motif) with at least 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the conserved domain starting with amino acid 20 up to amino acid 69 and/or amino acid 92 to amino acid 118 and/or amino acid 194 to amino acid 225 in SEQ ID NO:285).

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

Preferably, the polypeptide sequence which when used in the construction of a phylogenetic tree, such as the one depicted in Figure 5, clusters with the group of DUF642 polypeptides comprising the amino acid sequence represented by SEQ ID NO: 2 and/or 4 rather than with any other group.

In addition, DUF642 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 when grown under non-stress conditions, total seed yield (Totalwgseeds), number of seeds (nrfilledseed), seed fill rate (fillrate), harvest index,

greenness before flowering, total number of seeds (nrtotalseed) and for aboveground biomass (AreaMax).

5 Preferably, the polypeptide sequence which when used in the construction of a phylogenetic tree, such as the one depicted in Figure 9, clusters with the group of epimerase-related like polypeptides comprising the amino acid sequence represented by SEQ ID NO: 198 rather than with any other group.

10 In addition, epimerase-related 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 and/or increase seed yield, and more particularly any one or more of increased total weight of seeds, harvest index, fill rate, thousand kernel weight (TKW), above ground biomass, and number of filled seeds.

15 Preferably, the polypeptide sequence which when used in the construction of a phylogenetic tree, such as the one depicted in Figure 14, clusters with the group of PLPCase polypeptides comprising the amino acid sequence represented by SEQ ID NO: 285 rather than with any other group.

20 Furthermore, PLPCase polypeptides (at least in their native form) typically have phospholipase and/or carboxylesterase activity. Tools and techniques for measuring phospholipase and/or carboxylesterase activity are well known in the art, as e.g. Kirik and Mudgett (2009), PNAS 106 (48): 20532-20537.

25 In addition, PLPCase 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 root biomass and greenness index.

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

35 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 PLPCase polypeptides 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.

40 The present invention is illustrated by transforming plants with the nucleic acid sequence represented by SEQ ID NO: 1, encoding the polypeptide sequence of SEQ ID NO: 2. However, performance of the invention is not restricted to these sequences; the methods of the invention may advantageously be performed using any DUF642-encoding nucleic acid

or DUF642 polypeptide as defined herein.

Examples of nucleic acids encoding DUF642 polypeptides are given in Table A1 of the Examples section herein. Such nucleic acids are useful in performing the methods of the invention. The amino acid sequences given in Table A1 of the Examples section are example sequences of orthologues and paralogues of the DUF642 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; or SEQ ID NO: 3 or SEQ ID NO: 4, the second BLAST (back-BLAST) would be against rice sequences.

The invention also provides hitherto unknown DUF642-encoding nucleic acids and DUF642 polypeptides useful for conferring enhanced yield-related traits in plants relative to control plants.

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

- (i) a nucleic acid represented by SEQ ID NO: 33, 39, 91, 169, 171, 175;
- (ii) the complement of a nucleic acid represented by SEQ ID NO: 33, 39, 91, 169, 171, 175;
- (iii) a nucleic acid encoding a DUF642 polypeptide having in increasing order of preference at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence represented by SEQ ID NO: 34, 40, 92, 170, 172, 176, 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: 181 to SEQ ID NO: 186, and further preferably conferring enhanced yield-related traits relative to control plants.
- (iv) a nucleic acid molecule which hybridizes with a nucleic acid molecule of (i) to (iii) under high stringency hybridization conditions and preferably confers enhanced yield-related traits relative to control plants.

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

- (i) an amino acid sequence represented by SEQ ID NO: 34, 40, 92, 170, 172, 176;
- (ii) an amino acid sequence having, in increasing order of preference, at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%,

79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence represented by SEQ ID NO: 34, 40, 92, 170, 172, 176, 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: 181 to SEQ ID NO: 186, 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.

The present invention is illustrated by transforming plants with the nucleic acid sequence represented by SEQ ID NO: 197, encoding the polypeptide sequence of SEQ ID NO: 198. However, performance of the invention is not restricted to these sequences; the methods of the invention may advantageously be performed using any nucleic acid encoding epimerase-related like polypeptides or any epimerase-related like polypeptide as defined herein.

Examples of nucleic acids encoding epimerase-related like polypeptides are given in Table A2 of the Examples section herein. Such nucleic acids are useful in performing the methods of the invention. The amino acid sequences given in Table A2 of the Examples section are example sequences of orthologues and paralogues of the epimerase-related like polypeptide represented by SEQ ID NO: 198, 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: 197 or SEQ ID NO: 198, the second BLAST (back-BLAST) would be against poplar sequences.

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

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

- (i) a nucleic acid represented by SEQ ID NO: 199;
- (ii) the complement of a nucleic acid represented by SEQ ID NO: 199;
- (iii) a nucleic acid encoding epimerase-related 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 SEQ ID NO: 200, 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 7 to 13 as given herein, and further preferably conferring enhanced yield-related traits relative to control plants,

- 5 (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.

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

- (i) an amino acid sequence represented by SEQ ID NO: 200;
- (ii) an amino acid sequence having, in increasing order of preference, at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence represented by SEQ ID NO: 200, 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 7 to 13 as given herein, 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.

25 The present invention is illustrated by transforming plants with the nucleic acid sequence represented by SEQ ID NO: 284, encoding the polypeptide sequence of SEQ ID NO: 285. However, performance of the invention is not restricted to these sequences; the methods of the invention may advantageously be performed using any PLPCase-encoding nucleic acid or PLPCase polypeptide as defined herein.

30 Examples of nucleic acids encoding PLPCase 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 PLPCase polypeptide represented by SEQ ID NO: 285, 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: 284 or SEQ ID NO: 285, the second BLAST (back-BLAST) would be against *Medicago truncatula* sequences.

40 The invention also provides hitherto unknown PLPCase-encoding nucleic acids and PLPCase 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:

- 5 (i) a nucleic acid represented by any one of SEQ ID NO: 298, 302, 304, 312, 316, 336 or 346;
- (ii) the complement of a nucleic acid represented by any one of SEQ ID NO: 298, 302, 304, 312, 316, 336 or 346;
- 10 (iii) a nucleic acid encoding a PLPCase 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: 299, 303, 305, 313, 317, 337 or 347, 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: 348 to SEQ ID NO: 359, and further preferably conferring enhanced yield-related traits relative to control plants.
- 15 (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

25 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: 299, 303, 305, 313, 317, 337 or 347;
- 30 (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: 299, 303, 305, 313, 317, 337 or 347, 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: 348 to SEQ ID NO: 359, and further preferably conferring enhanced yield-related traits relative to control plants;
- 35
- 40 (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 to 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 to 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 DUF642 polypeptides, or epimerase-related like polypeptides, or PLPCase polypeptides, nucleic acids hybridising to nucleic acids encoding POI polypeptides, splice variants of nucleic acids encoding DUF642 polypeptides, or epimerase-related like polypeptides, or PLPCase polypeptides, allelic variants of nucleic acids encoding DUF642 polypeptides, or epimerase-related like polypeptides, or PLPCase polypeptides and variants of nucleic acids encoding DUF642 polypeptides, or epimerase-related like polypeptides, or PLPCase polypeptides obtained by gene shuffling. The terms hybridising sequence, splice variant, allelic variant and gene shuffling are as described herein.

Nucleic acids encoding DUF642 polypeptides, or epimerase-related like polypeptides, or PLPCase 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 to 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 to 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 DUF642 polypeptides, portions useful in the methods of the invention, encode a DUF642 polypeptide as defined herein, and have substantially the same biological activity as the amino acid sequences given in Table A1 of the Examples section. Preferably, the portion is a portion of any one of the nucleic acids given in Table A1 of the Examples section, or is a portion of a nucleic acid encoding an orthologue or paralogue of any one of

the amino acid sequences given in Table A1 of the Examples section. Preferably the portion is at least 250, 300, 350, 400, 450 consecutive nucleotides in length, the consecutive nucleotides being of any one of the nucleic acid sequences given in Table A1 of the Examples section, or of a nucleic acid encoding an orthologue or paralogue of any one of the amino acid sequences given in Table A1 of the Examples section. Most preferably the portion is a portion of the nucleic acid of SEQ ID NO: 1. Preferably, the portion encodes a fragment of an amino acid sequence which, when used in the construction of a phylogenetic tree, such as the one depicted in Figure 5, clusters with the group of DUF642 polypeptides comprising the amino acid sequence represented by SEQ ID NO: 2 or 4 rather than with any other group, and/or comprises at least one of the motifs 1 to 6 (SEQ ID NO: 181 to 186), and/or has at least 50% sequence identity to SEQ ID NO: 2 or 4.

Concerning epimerase-related like polypeptides, portions useful in the methods of the invention, encode an epimerase-related like polypeptide as defined herein, and have substantially the same biological activity as the amino acid sequences given in Table A2 of the Examples section. Preferably, the portion is a portion of any one of the nucleic acids given in Table A2 of the Examples section, or is a portion of a nucleic acid encoding an orthologue or paralogue of any one of the amino acid sequences given in Table A2 of the Examples section. Preferably the portion is at least 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, 1450, 1500, 1550, consecutive nucleotides in length, the consecutive nucleotides being of any one of the nucleic acid sequences given in Table A2 of the Examples section, or of a nucleic acid encoding an orthologue or paralogue of any one of the amino acid sequences given in Table A2 of the Examples section. Most preferably the portion is a portion of the nucleic acid of SEQ ID NO: 197.

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 10, clusters with the group of polypeptides comprising the amino acid sequence represented by SEQ ID NO: 198 rather than with any other group;
- comprises one or more of the following domains SSF53383, PTHR11601; PF00266, and G3DSA:3.40.640.10
- comprises any one or more of the motifs 7 to 13 (as represented by SEQ ID NO 273, 274, 275, 280, 281, 282, 283, respectively) as provided herein, and
- has at least 40% sequence identity to SEQ ID NO: 198.

Concerning PLPCase polypeptides, portions useful in the methods of the invention, encode a PLPCase 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 350, 400, 450, 500, 550, 600, 650, 700, 750, 780

consecutive nucleotides in length, the consecutive nucleotides being of any one of the nucleic acid sequences given in Table A3 of the Examples section, or of a nucleic acid encoding an orthologue or paralogue of any one of the amino acid sequences given in Table A3 of the Examples section. Most preferably the portion is a portion of the nucleic acid of SEQ ID NO: 284. Preferably, the portion encodes a fragment of an amino acid sequence which, when used in the construction of a phylogenetic tree, such as the one depicted in Figure 14, clusters with the group of PLPCase polypeptides comprising the amino acid sequence represented by SEQ ID NO: 285 rather than with any other group, and/or comprises motifs any of the motifs 14 to 16, preferably any of the motifs 17 to 19, or preferably any of the motifs 20 to 22, more preferably comprises any of the motifs 23 to 25, and/or has phospholipase and/or carboxylesterase biological activity, and/or has at least 60% sequence identity to SEQ ID NO: 285.

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 DUF642 polypeptide, or an epimerase-related like polypeptide, or a PLPCase polypeptide as defined herein, or with a portion as defined herein.

According to the present invention, there is provided a method for enhancing yield-related traits in plants, comprising introducing and expressing in a plant a nucleic acid capable of hybridizing to the complement of a nucleic acid encoding any one of the proteins given in Table A of the Examples section, or to the complement of a nucleic acid encoding an orthologue, paralogue or homologue of any one of the proteins given in Table A of the Examples section.

Hybridising sequences useful in the methods of the invention encode a DUF642 polypeptide, or an epimerase-related like polypeptide, or a PLPCase polypeptide as defined herein, having substantially the same biological activity as the DUF642, epimerase-related like or PLPCase amino acid sequences given in Table A of the Examples section. Preferably, the hybridising sequence is capable of hybridising to the complement of a nucleic acid encoding any one of the proteins given in Table A of the Examples section, or to a portion of any of these sequences, a portion being as defined herein, or the hybridising sequence is capable of hybridising to the complement of a nucleic acid encoding an orthologue or paralogue of any one of the amino acid sequences given in Table A of the Examples section.

Concerning DUF642 polypeptides, the hybridising sequence is most preferably capable of hybridising to the complement of a nucleic acid encoding the polypeptide as represented by SEQ ID NO: 1 or to a portion thereof.

Preferably, the hybridising sequence encodes a polypeptide with an amino acid sequence which, when full-length and used in the construction of a phylogenetic tree, clusters with the group of DUF642 polypeptides comprising the amino acid sequence represented by SEQ ID NO: 2 and/or 4, rather than with any other group, comprises at least one of the motifs 1 to 6 (SEQ ID NO: 181 to 186), and/or has at least 50% sequence identity to SEQ ID NO: 2 or 4.

In one embodiment the hybridising sequence is capable of hybridising to the complement of a nucleic acid as represented by SEQ ID NO: 1 or SEQ ID NO: 3 or to a portion thereof under conditions of medium or high stringency, preferably high stringency as defined above. In another embodiment the hybridising sequence is capable of hybridising to the complement of a nucleic acid as represented by SEQ ID NO: 1 or SEQ ID NO: 3 under stringent conditions.

Concerning epimerase-related like polypeptides, the hybridising sequence is most preferably capable of hybridising to the complement of a nucleic acid encoding the polypeptide as represented by SEQ ID NO: 197 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 10, clusters with the group of polypeptides comprising the amino acid sequence represented by SEQ ID NO: 198 rather than with any other group;
- comprises one or more of the following domains SSF53383, PTHR11601; PF00266, and G3DSA:3.40.640.10
- comprises any one or more of the motifs 7 to 13 (as represented by SEQ ID NO 273, 274, 275, 280, 281, 282, 283, respectively) as provided herein, and
- has at least 40% sequence identity to SEQ ID NO: 198.

In one embodiment the hybridising sequence is capable of hybridising to the complement of a nucleic acid as represented by SEQ ID NO: 197 or to a portion thereof under conditions of medium or high stringency, preferably high stringency as defined above. In another embodiment the hybridising sequence is capable of hybridising to the complement of a nucleic acid as represented by SEQ ID NO: 197 under stringent conditions.

Concerning PLPCase polypeptides, the hybridising sequence is most preferably capable of hybridising to the complement of a nucleic acid encoding the polypeptide as represented by SEQ ID NO: 284 or to a portion thereof. In one embodiment, the hybridization conditions are of medium stringency, preferably of high stringency, as defined herein.

Preferably, the hybridising sequence encodes a polypeptide with an amino acid sequence which, when full-length and used in the construction of a phylogenetic tree, such as the one

depicted in Figure 14, clusters with the group of PLPCase polypeptides comprising the amino acid sequence represented by SEQ ID NO: 285 rather than with any other group, and/or comprises motifs any of the motifs 14 to 16, preferably any of the motifs 17 to 19, or preferably any of the motifs 20 to 22, more preferably comprises any of the motifs 23 to 25, and/or has phospholipase and/ or carboxylesterase biological activity, and/or has at least 60% sequence identity to SEQ ID NO: 285.

Another nucleic acid variant useful in the methods of the invention is a splice variant encoding a DUF642 polypeptide, or an epimerase-related like polypeptide, or a PLPCase 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 A of the Examples section, or a splice variant of a nucleic acid encoding an orthologue, paralogue or homologue of any of the amino acid sequences given in Table A of the Examples section.

Concerning DUF642 polypeptides, preferred splice variants are splice variants of a nucleic acid represented by SEQ ID NO: 1, or a splice variant of a nucleic acid encoding an orthologue or paralogue of SEQ ID NO: 2. Preferably, the amino acid sequence encoded by the splice variant, when used in the construction of a phylogenetic tree, clusters with the group of DUF642 polypeptides comprising the amino acid sequence represented by SEQ ID NO: 2 and/or 4, rather than with any other group, comprises at least one of the motifs 1 to 6 (SEQ ID NO: 181 to 186), and/or has at least 50% sequence identity to SEQ ID NO: 2 or 4.

Concerning epimerase-related like polypeptides, preferred splice variants are splice variants of a nucleic acid represented by SEQ ID NO: 197, or a splice variant of a nucleic acid encoding an orthologue or paralogue of SEQ ID NO: 198. 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 10, clusters with the group of polypeptides comprising the amino acid sequence represented by SEQ ID NO: 198 rather than with any other group;
- comprises one or more of the following domains SSF53383, PTHR11601; PF00266, and G3DSA:3.40.640.10
- comprises any one or more of the motifs 7 to 13 (as represented by SEQ ID NO 273, 274, 275, 280, 281, 282, 283, respectively) as provided herein, and
- has at least 40% sequence identity to SEQ ID NO: 198.

Concerning PLPCase polypeptides, preferred splice variants are splice variants of a nucleic acid represented by SEQ ID NO: 284, or a splice variant of a nucleic acid encoding an orthologue or paralogue of SEQ ID NO: 285. Preferably, the amino acid sequence encoded

by the splice variant, when used in the construction of a phylogenetic tree, such as the one depicted in Figure 14, clusters with the group of PLPCase polypeptides comprising the amino acid sequence represented by SEQ ID NO: 285 rather than with any other group, and/or comprises motifs any of the motifs 14 to 16, preferably any of the motifs 17 to 19, or preferably any of the motifs 20 to 22, more preferably comprises any of the motifs 23 to 25, and/or has phospholipase and/or carboxylesterase biological activity, and/or has at least 60% sequence identity to SEQ ID NO: 285.

Another nucleic acid variant useful in performing the methods of the invention is an allelic variant of a nucleic acid encoding a DUF642 polypeptide, or an epimerase-related like polypeptide, or a PLPCase polypeptide as defined herein, 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 A of the Examples section, or comprising introducing and expressing in a plant an allelic variant of a nucleic acid encoding an orthologue, paralogue or homologue of any of the amino acid sequences given in Table A of the Examples section.

Concerning DUF642 polypeptides, the polypeptides encoded by allelic variants useful in the methods of the present invention have substantially the same biological activity as the DUF642 polypeptide of SEQ ID NO: 2 and any of the amino acids depicted in Table A1 of the Examples section. Allelic variants exist in nature, and encompassed within the methods of the present invention is the use of these natural alleles. Preferably, the allelic variant is an allelic variant of SEQ ID NO: 1 or an allelic variant of a nucleic acid encoding an orthologue or paralogue of SEQ ID NO: 2. Preferably, the amino acid sequence encoded by the allelic variant, when used in the construction of a phylogenetic tree, clusters with the DUF642 polypeptides comprising the amino acid sequence represented by SEQ ID NO: 2 and/or 4, rather than with any other group, comprises at least one of the motifs 1 to 6 (SEQ ID NO: 181 to 186), and/or has at least 50% sequence identity to SEQ ID NO: 2 or 4.

Concerning epimerase-related like polypeptides, the polypeptides encoded by allelic variants useful in the methods of the present invention have substantially the same biological activity as the epimerase-related like polypeptide of SEQ ID NO: 198 and any of the amino acids depicted in Table A2 of the Examples section. Allelic variants exist in nature, and encompassed within the methods of the present invention is the use of these natural alleles. Preferably, the allelic variant is an allelic variant of SEQ ID NO: 197 or an allelic variant of a nucleic acid encoding an orthologue or paralogue of SEQ ID NO: 198. 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 10, clusters with the group of polypeptides comprising the amino acid sequence represented by SEQ ID NO: 198 rather than with any other group;

- comprises one or more of the following domains SSF53383, PTHR11601; PF00266, and G3DSA:3.40.640.10
- comprises any one or more of the motifs 7 to 13 (as represented by SED ID NO 273, 274, 275, 280, 281, 282, 283, respectively) as provided herein, and
- has at least 40% sequence identity to SEQ ID NO: 198.

Concerning PLPCase polypeptides, the polypeptides encoded by allelic variants useful in the methods of the present invention have substantially the same biological activity as the PLPCase polypeptide of SEQ ID NO: 285 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: 284 or an allelic variant of a nucleic acid encoding an orthologue or paralogue of SEQ ID NO: 285. Preferably, the amino acid sequence encoded by the allelic variant, when used in the construction of a phylogenetic tree, such as the one depicted in Figure 14, clusters with the PLPCase polypeptides comprising the amino acid sequence represented by SEQ ID NO: 285 rather than with any other group, and/or comprises motifs any of the motifs 14 to 16, preferably any of the motifs 17 to 19, or preferably any of the motifs 20 to 22, more preferably comprises any of the motifs 23 to 25, and/or has phospholipase and/or carboxylesterase biological activity, and/or has at least 60% sequence identity to SEQ ID NO: 285.

Gene shuffling or directed evolution may also be used to generate variants of nucleic acids encoding DUF642 polypeptides, or epimerase-related like polypeptides, or PLPCase polypeptides as defined herein; 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 A of the Examples section, or comprising introducing and expressing in a plant a variant of a nucleic acid encoding an orthologue, paralogue or homologue of any of the amino acid sequences given in Table A of the Examples section, which variant nucleic acid is obtained by gene shuffling.

Concerning DUF642 polypeptides, the amino acid sequence encoded by the variant nucleic acid obtained by gene shuffling, when used in the construction of a phylogenetic tree, preferably clusters with the group of DUF642 polypeptides comprising the amino acid sequence represented by SEQ ID NO: 2 and/or 4, rather than with any other group, comprises at least one of the motifs 1 to 6 (SEQ ID NO: 181 to 186), and/or has at least 50% sequence identity to SEQ ID NO: 2 or 4.

Concerning epimerase-related like polypeptides, the amino acid sequence encoded by the variant nucleic acid obtained by gene shuffling, preferably is 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 10, clusters with the group of polypeptides comprising the amino acid sequence represented by SEQ ID NO: 198 rather than with any other group;
- comprises one or more of the following domains SSF53383, PTHR11601; PF00266, and G3DSA:3.40.640.10
- comprises any one or more of the motifs 7 to 13 (as represented by SEQ ID NO 273, 274, 275, 280, 281, 282, 283, respectively) as provided herein, and
- has at least 40% sequence identity to SEQ ID NO: 198.

Concerning PLPCase polypeptides, the amino acid sequence encoded by the variant nucleic acid obtained by gene shuffling, when used in the construction of a phylogenetic tree such as the one depicted in Figure 14, preferably clusters with the group of PLPCase polypeptides comprising the amino acid sequence represented by SEQ ID NO: 285 rather than with any other group, and/or comprises motifs any of the motifs 14 to 16, preferably any of the motifs 17 to 19, or preferably any of the motifs 20 to 22, more preferably comprises any of the motifs 23 to 25, and/or has phospholipase and/or carboxylesterase biological activity, and/or has at least 60% sequence identity to SEQ ID NO: 285.

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.).

Concerning DUF642 polypeptides, nucleic acids encoding DUF642 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 DUF642 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*.

For example, the nucleic acid encoding the DUF642 polypeptide of SEQ ID NO:2 can be generated from the nucleic acid encoding the DUF642 polypeptide of SEQ ID NO:2 by alteration of several nucleotides, e.g. by site-directed mutagenesis using PCR based methods (see Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989 and yearly updates)). DUF642 polypeptides differing from the sequence of SEQ ID NO: 2 or SEQ ID NO: 4 by one or several amino acids may be used to increase the yield of plants in the methods and constructs and plants of the invention.

Concerning epimerase-related like polypeptides, nucleic acids encoding epimerase-related 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 epimerase-related like polypeptide-encoding nucleic acid is from a plant, further preferably from a dicotyledonous plant, further preferably from the family *Salicaceae*, more preferably the nucleic acid is from the genus *Populus*, and
5 most preferably the nucleic acid is from *Populus trichocarpa*.

Concerning PLPCase polypeptides, nucleic acids encoding PLPCase 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
10 manipulation. Preferably the PLPCase polypeptide-encoding nucleic acid is from a plant, further preferably from a dicotyledonous plant, more preferably from the family Fabaceae, more preferably from the genus *Medicago*, most preferably the nucleic acid is from *Medicago truncatula*.

15 PLPCase polypeptides differing from the sequence of SEQ ID NO: 285 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.

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

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

Reference herein to enhanced yield-related traits is taken to mean an increase of early vigour and/or in biomass (weight) of one or more parts of a plant, which may include (i)
40 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.

The present invention provides a method for increasing yield, especially seed yield of plants, relative to control plants, which method comprises modulating expression in a plant
5 of a nucleic acid encoding a DUF642 polypeptide as defined herein.

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

The present invention also provides a method for increasing yield, especially enhancing root biomass of plants, relative to control plants, which method comprises modulating expression in a plant of a nucleic acid encoding a PLPCase 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
20 encoding a DUF642 polypeptide, or an epimerase-related like polypeptide, or a PLPCase 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
25 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 DUF642 polypeptide, or an epimerase-related like polypeptide, or a PLPCase polypeptide.
30

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
35 expression in a plant of a nucleic acid encoding a DUF642 polypeptide, or an epimerase-related like polypeptide, or a PLPCase polypeptide.

Performance of the methods of the invention gives plants grown under conditions of nutrient deficiency, particularly under conditions of nitrogen deficiency, increased yield relative to
40 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 DUF642 polypeptide, or an epimerase-related like polypeptide, or a PLPCase polypeptide.

Performance of the methods of the invention gives plants grown under conditions of salt stress, increased yield relative to control plants grown under comparable conditions. Therefore, according to the present invention, there is provided a method for increasing yield in plants grown under conditions of salt stress, which method comprises modulating expression in a plant of a nucleic acid encoding a DUF642 polypeptide, or an epimerase-related like polypeptide, or a PLPCase polypeptide.

The invention also provides genetic constructs and vectors to facilitate introduction and/or expression in plants of nucleic acids encoding DUF642 polypeptides, or epimerase-related like polypeptides, or PLPCase 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 DUF642 polypeptide, or an epimerase-related like polypeptide, or a PLPCase polypeptide as defined herein;
- (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 DUF642 polypeptide, or an epimerase-related like polypeptide, or a PLPCase polypeptide is as defined herein. The term "control sequence" and "termination sequence" are as defined herein.

In a preferred embodiment, the present invention provides a construct comprising:

- (a) a nucleic acid encoding a PLPCase polypeptide as represented by any of SEQ ID NO: 299, 303, 305, 313, 317, 337 or 347;
- (b) one or more control sequences capable of driving expression of the nucleic acid sequence of (a); and optionally
- (c) a transcription termination sequence.

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 herein. Thus the invention furthermore provides plants or host cells transformed with a construct as described above. In particular, the invention provides plants transformed with a construct as described above, which plants have increased yield-related traits as described herein.

In one embodiment the genetic construct of the invention confers increased yield or yield related traits(s) to a living plant cell when it has been introduced into said plant cell and express the nucleic acid encoding the DUF642 polypeptide, or the epimerase-related like polypeptide, or the PLPCase 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).

The promoter in such a genetic construct may be a non-native promoter to the nucleic acid described above, i.e. a promoter which does not regulate the expression of the herein disclosed polynucleotides in its native surrounding.

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

The constitutive promoter is preferably a ubiquitous constitutive promoter of medium strength. More preferably it is a plant derived promoter, e.g. a promoter of plant chromosomal origin, such as a GOS2 promoter or a promoter of substantially the same strength and having substantially the same expression pattern (a functionally equivalent promoter), more preferably the promoter is the promoter GOS2 promoter from rice. Further preferably the constitutive promoter is represented by a nucleic acid sequence substantially similar to SEQ ID NO: 189, or SEQ ID NO: 278, or SEQ ID NO: 361, most preferably the constitutive promoter is as represented by SEQ ID NO: 189, or SEQ ID NO: 278, or SEQ ID NO: 361. See the "Definitions" section herein for further examples of constitutive promoters.

Concerning DUF642 polypeptides, it should be clear that the applicability of the present invention is not restricted to the DUF642 polypeptide-encoding nucleic acid represented by SEQ ID NO: 1 or 3, nor is the applicability of the invention restricted to expression of a DUF642 polypeptide-encoding nucleic acid when driven by a constitutive promoter.

Concerning epimerase-related like polypeptides, it should be clear that the applicability of the present invention is not restricted to the epimerase-related like polypeptide-encoding nucleic acid represented by SEQ ID NO: 197, nor is the applicability of the invention restricted to expression of an epimerase-related like polypeptide-encoding nucleic acid when driven by a constitutive promoter, or when driven by a root-specific promoter.

Concerning PLPCase polypeptides, it should be clear that the applicability of the present invention is not restricted to the PLPCase polypeptide-encoding nucleic acid represented by SEQ ID NO: 284, nor is the applicability of the invention restricted to expression of a PLPCase polypeptide-encoding nucleic acid when driven by a constitutive promoter.

5

Concerning DUF642 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: 189, operably linked to the nucleic acid encoding the DUF642 polypeptide. More preferably, the construct comprises a zein terminator (t-zein) linked to the 3' end of the DUF642 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: 187 or 188 (pPRO::DUF642::t-zein sequence). Furthermore, one or more sequences encoding selectable markers may be present on the construct introduced into a plant.

Concerning epimerase-related 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: 278, operably linked to the nucleic acid encoding the epimerase-related like polypeptide. More preferably, the construct comprises a zein terminator (t-zein) linked to the 3' end of the coding sequence of an epimerase-related like polypeptide. Most preferably, the expression cassette comprises a sequence having in increasing order of preference at least 95%, at least 96%, at least 97%, at least 98%, at least 99% identity to the sequence represented by SEQ ID NO: 279 (pPRO::ERL::t-zein sequence). Furthermore, one or more sequences encoding selectable markers may be present on the construct introduced into a plant.

Concerning PLPCase 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: 361, operably linked to the nucleic acid encoding the PLPCase polypeptide. More preferably, the expression cassette comprises the sequence represented by SEQ ID NO: 360 (pGOS2:: PLPCase::t-zein sequence). Furthermore, one or more sequences encoding selectable markers may be present on the construct introduced into a plant.

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

40

As mentioned above, a preferred method for modulating expression of a nucleic acid encoding a DUF642 polypeptide, or a epimerase-related like polypeptide, or a PLPCase

polypeptide is by introducing and expressing in a plant a nucleic acid encoding a DUF642 polypeptide, or a epimerase-related like polypeptide, or a PLPCase 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 DUF642 polypeptide, or a epimerase-related like polypeptide, or a PLPCase 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 yield, more particularly increased seed yield, which method comprises:

- (i) introducing and expressing in a plant or plant cell a DUF642 polypeptide-encoding nucleic acid or a genetic construct comprising a DUF642 polypeptide-encoding nucleic acid; and
- (ii) cultivating the plant cell under conditions promoting plant growth and development.

The nucleic acid of (i) may be any of the nucleic acids capable of encoding a DUF642 polypeptide as defined herein.

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

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

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

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

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

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

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

10 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.

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

25 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.

30 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 DUF642 polypeptide, or a epimerase-related like polypeptide, or a PLPCase polypeptide as defined herein, preferably in a genetic construct such as an expression cassette. The present invention extends further to encompass the progeny of a primary transformed or transfected cell, tissue, organ or whole plant that has been produced by any of the aforementioned methods, the only requirement being that progeny exhibit the same genotypic and/or phenotypic characteristic(s) as those produced by the parent in the methods according to the invention.

40 In a further embodiment the invention extends to seeds comprising the expression cassettes of the invention, the genetic constructs of the invention, or the nucleic acids encoding the DUF642 polypeptides, or the epimerase-related like polypeptides, or the PLPCase polypeptides and/or the DUF642 polypeptides, or the epimerase-related like polypeptides, or the PLPCase polypeptides as described above.

The invention also includes host cells containing an isolated nucleic acid encoding a DUF642 polypeptide, or a epimerase-related like polypeptide, or a PLPCase polypeptide as defined herein. In one embodiment host cells according to the invention are plant cells, yeasts, bacteria or fungi. Host plants for the nucleic acids, construct, expression cassette or the vector used in the method according to the invention are, in principle, advantageously all plants which are capable of synthesizing the polypeptides used in the inventive method. In a particular embodiment the plant cells of the invention overexpress the nucleic acid molecule of the invention.

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

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 DUF642 polypeptide, or a epimerase-related like polypeptide, or a PLPCase polypeptide. The invention furthermore relates to products derived or produced, preferably directly derived or produced, from a harvestable part of such a plant, such as dry pellets, meal or powders, oil, fat and fatty acids, starch or proteins.

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

Examples of such methods would be growing corn plants of the invention, harvesting the corn cobs and remove the kernels. These may be used as feedstuff or processed to starch
10 and oil as agricultural products. The product may be produced at the site where the plant has been grown, or the plants or parts thereof may be removed from the site where the plants have been grown to produce the product. Typically, the plant is grown, the desired harvestable parts are removed from the plant, if feasible in repeated cycles, and the product made from the harvestable parts of the plant. The step of growing the plant may be
15 performed only once each time the methods of the invention is performed, while allowing repeated times the steps of product production e.g. by repeated removal of harvestable parts of the plants of the invention and if necessary further processing of these parts to arrive at the product. It is also possible that the step of growing the plants of the invention is repeated and plants or harvestable parts are stored until the production of the product is
20 then performed once for the accumulated plants or plant parts. Also, the steps of growing the plants and producing the product may be performed with an overlap in time, even simultaneously to a large extent, or sequentially. Generally the plants are grown for some time before the product is produced.

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

30 In yet another embodiment the polynucleotides or the polypeptides of the invention are comprised in an agricultural product. In a particular embodiment the nucleic acid sequences and protein sequences of the invention may be used as product markers, for example where an agricultural product was produced by the methods of the invention. Such a marker
35 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
40 methods for protein detection.

The present invention also encompasses use of nucleic acids encoding DUF642 polypeptides, epimerase-related like polypeptides, or PLPCase polypeptides as described herein and use of these DUF642 polypeptides, or epimerase-related like polypeptides, or PLPCase polypeptides in enhancing any of the aforementioned yield-related traits in plants. For example, nucleic acids encoding DUF642 polypeptides, or epimerase-related like polypeptides, or PLPCase polypeptides described herein, or the DUF642 polypeptides, or the epimerase-related like polypeptides, or the PLPCase polypeptides themselves, may find use in breeding programmes in which a DNA marker is identified which may be genetically linked to a DUF642 polypeptide-encoding gene, or an epimerase-related like polypeptide-encoding gene, or a PLPCase polypeptide-encoding gene. The nucleic acids/genes, or the DUF642 polypeptides, or the epimerase-related like polypeptides, or the PLPCase 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 DUF642 polypeptide-encoding nucleic acid/gene, or an epimerase-related like polypeptide-encoding nucleic acid/gene, or a PLPCase polypeptide-encoding nucleic acid/gene may find use in marker-assisted breeding programmes. Nucleic acids encoding DUF642 polypeptides, or epimerase-related like polypeptides, or PLPCase 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.

Concerning DUF642 polypeptides, in one embodiment any comparison to determine sequence identity percentages is performed

- in the case of a comparison of nucleic acids over the entire coding region of SEQ ID NO: 1 or SEQ ID NO: 3, or
- in the case of a comparison of polypeptide sequences over the entire length of SEQ ID NO: 2 or SEQ ID NO: 4.

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

Concerning epimerase-related like polypeptides, in one embodiment any comparison to determine sequence identity percentages is performed

- in the case of a comparison of nucleic acids over the entire coding region of SEQ ID NO: 197, or
- in the case of a comparison of polypeptide sequences over the entire length of

SEQ ID NO: 198.

For example, a sequence identity of 50% sequence identity in this embodiment means that over the entire coding region of SEQ ID NO: 197, 50 percent of all bases are identical between the sequence of SEQ ID NO: 197 and the related sequence. Similarly, in this embodiment a polypeptide sequence is 50 % identical to the polypeptide sequence of SEQ ID NO: 198, when 50 percent of the amino acids residues of the sequence as represented in SEQ ID NO: 198, are found in the polypeptide tested when comparing from the starting methionine to the end of the sequence of SEQ ID NO: 198.

Moreover concerning DUF642 polypeptides, the present invention relates to the following specific embodiments:

1. A method for enhancing yield-related traits in plants relative to control plants, comprising modulating expression in a plant of a nucleic acid encoding a DUF642 polypeptide, wherein said DUF642 polypeptide comprises an InterPro accession IPR006946 DUF642 domain, preferably said DUF642 polypeptide comprises the signature sequence FSAARTCAQ (SEQ ID NO: 194).
2. Method according to embodiment 1, wherein said modulated expression is effected by introducing and expressing in a plant said nucleic acid encoding said DUF642 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 of embodiments 1 to 4, wherein said DUF642 polypeptide comprises one or more of the following motifs:
 - (i) Motif 1: N[LAM][VL]KNG[DG]FEEGP[WYH][MI][FLI]P[NG][TS][STR][WL]GVL[LVI]P[PTS][NKM][LQDV][EDV][DE][ED][THY]S[PS]L[PS][GP]W[IMT][IV] (SEQ ID NO: 181),
 - (ii) Motif 2: [VLA][EKAT][KPR]G[SAM][LRH]Y[SA][LIV][TI]F[SG]A[AS]RTCAQ[DLAS]E[SVRL]L[NR][VI] (SEQ ID NO: 182),
 - (iii) Motif 3: D[PE][AT]CGP[LI][IL]D[AS][VIF]AI[KR] (SEQ ID NO: 183)
6. Method according to any one of embodiments 1 to 5, wherein said nucleic acid encoding a DUF642 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 DUF642 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.
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.

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 SEQ ID NO: 4 or a nucleic acid encoding a polypeptide having at least 90% sequence identity to SEQ ID NO: 2 and/or SEQ ID NO: 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 DUF642 polypeptide as defined in any of embodiments 1 and 5 to 9.
12. Construct comprising:
- 15 (i) nucleic acid encoding a DUF642 as defined in any of embodiments 1 and 5 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.
- 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.
- 30 15. Plant, plant part or plant cell transformed with a construct according to embodiment 12 or 13.
- 35 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 DUF642 polypeptide as defined in any of embodiments 1 and 5 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 DUF642 polypeptide as defined in any of embodiments 1 and 5 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.

19. Harvestable parts of a plant according to embodiment 18, wherein said harvestable parts are preferably shoot biomass and/or seeds.

5 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 DUF642 polypeptide as defined in any of embodiments 1 and 5 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
10 increasing biomass in plants relative to control plants.

22. Plant having increased yield, particularly increased biomass and/or increased seed yield, relative to control plants, resulting from modulated expression of a nucleic acid encoding a DUF642 polypeptide, or a transgenic plant cell originating from or being part of said transgenic plant.

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

20 24. Plant according to embodiment 11, 15, or 21, or a transgenic plant cell originating thereof, or a method according to embodiment 22, wherein said plant is a crop plant, preferably a dicot such as sugar beet, alfalfa, trefoil, chicory, carrot, cassava, cotton, soybean, canola or a monocot, such as sugarcane, or a cereal, such as rice, maize, wheat, barley, millet, rye, triticale, sorghum emmer, spelt, secale, einkorn, teff, milo and oats.

25 25. Construct according to embodiment 12 or 13 comprised in a plant cell.

26. Recombinant chromosomal DNA comprising the construct according to embodiment 12 or 13.

30 Moreover concerning epimerase-related like polypeptides, the present invention relates to the following specific items:

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 epimerase-related like polypeptide, wherein said epimerase-related like polypeptide comprises one or more of the following motifs:

35 (i) Motif 7: CREEGV[DE][QK]VFVD[AG]AH[AS]IG[QSC]V[PDE][VI][DN][VM][KR][ED]IGADFY[TV]SNLHKWFFCPP[SA]VAFL[YH] (SEQ ID NO: 273),

(ii) Motif 8: EF[SA]HH[DN]P[GAN]VAR[IV]NNGSFG[CS]CP[AG]S[VI][LI]AAQ[ARK][RN]WQ[LR][LRQ]FL[RQA]QPD[AD]FYF[ND]xL[QRK][PK]G (SEQ ID NO: 274),

40 (iii) Motif 9: S[LI]VDNATTA AIVLQ[HQ][VAI][AG][WR][AS]FAEG[RKN]FA[KR][GN]D[AVT]V[LV]MLH[YC]AY[GQ][AS]VKKSI[EQH]AYV (SEQ ID NO: 275)

2. Method according to item 1, wherein said modulated expression is effected by introducing and expressing in a plant said nucleic acid encoding said epimerase-

related like polypeptide.

3. Method according to item 1 or 2, wherein said enhanced yield-related traits comprise increased yield relative to control plants, and preferably comprise increased biomass and/or increased seed yield relative to control plants.
- 5 4. Method according to any of items 1 to 3, wherein said epimerase-related like polypeptide as used herein comprises a conserved domain with at least 50% sequence identity to a conserved domain as represented by any one of:
 - amino acid coordinates 32 to 449 in SEQ ID NO:198 (motif 10);
 - amino acid coordinates 47 to 276 in SEQ ID NO:198 (motif 11);
 - 10 - amino acid coordinates 92 to 364 in SEQ ID NO:198 (motif 12); and
 - amino acid coordinates 59 to 323 in SEQ ID NO:198 (motif 13).
5. Method according to any of items 1 or 4, wherein said nucleic acid encodes a polypeptide having at least 40 % overall sequence identity to the amino acid represented by SEQ ID NO: 198.
- 15 6. Method according to any of items 1 to 5, wherein said nucleic acid encoding said epimerase-related 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 from *Populus trichocarpa*.
7. Method according to any one of items 1 to 6, wherein said nucleic acid encoding said
 20 epimerase-related like polypeptide encodes any one of the polypeptides listed in Table A2 or is a portion of such a nucleic acid, or a nucleic acid capable of hybridising with such a nucleic acid.
8. Method according to any one of items 1 to 7, wherein said nucleic acid sequence encodes an orthologue or paralogue of any of the polypeptides given in Table A2.
- 25 9. Method according to any one of items 1 to 8, wherein said enhanced yield-related traits are obtained under non-stress conditions.
10. Method according to any one of items 1 to 8, wherein said enhanced yield-related traits are obtained under conditions of drought stress, salt stress or nitrogen deficiency.
11. Method according to any one of items 1 to 10, wherein said nucleic acid is operably
 30 linked to a constitutive promoter, preferably to a medium strength constitutive promoter, preferably to a plant promoter, more preferably to a GOS2 promoter, most preferably to a GOS2 promoter from rice.
12. Plant, plant part thereof, including seeds, or plant cell, obtainable by a method according to any one of items 1 to 11, wherein said plant, plant part or plant cell
 35 comprises a recombinant nucleic acid encoding an epimerase-related like polypeptide as defined in any of items 1 and 4 to 8.
13. Plant according to item 12 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,
 40 sorghum, emmer, spelt, secale, einkorn, teff, milo or oats.
14. Construct comprising:
 - (i) nucleic acid encoding an epimerase-related like polypeptide as defined in any of

items 1 and 4 to 8;

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

(iii) a transcription termination sequence.

- 5 15. Construct according to item 14, 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 16. Use of a construct according to item 14 or 15 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.
17. Plant, plant part or plant cell transformed with a construct according to item 14 or 15.
- 15 18. Plant according to item 17 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 19. 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 an epimerase-related like polypeptide as defined in any of items 1 and 4 to 8; and
- (ii) cultivating said plant cell or plant under conditions promoting plant growth and development.
- 30 20. Transgenic plant having enhanced yield-related traits relative to control plants, preferably increased yield relative to control plants, and more preferably increased seed yield and/or increased biomass, resulting from modulated expression of a nucleic acid encoding an epimerase-related like polypeptide as defined in any of items 1 and 4 to 8; or a transgenic plant cell derived from said transgenic plant.
- 35 21. Transgenic plant according to item 20 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.
22. Harvestable parts of a plant according to any of items 12-13, 17-18 or 20-21, wherein said harvestable parts are preferably shoot biomass and/or seeds.
23. Products derived from a plant according to any of items 12-13, 17-18 or 20-21 and/or from harvestable parts of a plant according to item 22.
- 40 24. Isolated nucleic acid molecule selected from:
- (i) a nucleic acid represented by SEQ ID NO: 199 ;
- (ii) the complement of a nucleic acid represented by SEQ ID NO: 199;
- (iii) a nucleic acid encoding epimerase-related 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 SEQ ID NO: 200, 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 7 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.
25. Isolated polypeptide selected from:
- (i) an amino acid sequence represented by SEQ ID NO: 200;
- (ii) an amino acid sequence having, in increasing order of preference, at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid sequence represented by SEQ ID NO: 200, 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 7 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.
26. Use of a nucleic acid encoding an epimerase-related like polypeptide as defined in any of items 1 and 4 to 8 and 25; 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.
27. Use of a nucleic acid as defined in item 24 and encoding an epimerase-related 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.
28. Use of a nucleic acid encoding an epimerase-related like polypeptide as defined in any of items 1 and 4 to 8 and 25; as molecular marker.
29. Use of a nucleic acid as defined in item 24 and encoding an epimerase-related like polypeptide as defined in any of items 1 and 4 to 8 and 25 as molecular marker.
30. A method for the production of a product comprising the steps of growing the plants according to item 12, 18 or 20 and producing said product from or by
- (i) said plants; or
- (ii) parts, including seeds, of said plants.

31. Construct according to item 14 or 15 comprised in a plant cell.
32. Recombinant chromosomal DNA comprising the construct according to item 14 or 15 comprised.

5 Moreover concerning epimerase-related like polypeptides, the invention is in particular characterized by any one or more of the following embodiments:

I. 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 epimerase-related like polypeptide by introducing and expressing in a plant said nucleic acid
10 encoding said epimerase-related like polypeptide, wherein said epimerase-related like polypeptide has at least 80 % overall sequence identity to the amino acid represented by SEQ ID NO: 198.

II. Method according to embodiment I, wherein said epimerase-related like polypeptide comprises one or more of the following motifs:

15 (i) Motif 7: CREEGV[DE][QK]VFVD[AG]AH[AS]IG[QSC]V[PDE][VI][DN][VM][KR][ED]IGADFY[TV]SNLHKWFFCPP[SA]VAFL[YH] (SEQ ID NO: 273),

(ii) Motif 8: EF[SA]HH[DN]P[GAN]VAR[IV]NNGSFG[CS]CP[AG]S[VI][LI]AAQ[ARK][RN]WQ[LR][LRQ]FL[RQA]QPD[AD]FYF[ND]xL[QRK][PK]G (SEQ ID NO: 274),

20 (iii) Motif 9: S[LI]VDNATTAAIVLQ[HQ][VAI][AG][WR][AS]FAEG[RKN]FA[KR][GN]D[AVT]V[LV]MLH[YC]AY[GQ][AS]VKKSI[EQH]AYV (SEQ ID NO: 275)

III. Method according to embodiment I or II, wherein said enhanced yield-related traits comprises increased biomass.

IV. Method according to any one of embodiments I to III, wherein said enhanced yield-related traits comprises increased seed yield relative to control plants.

25 V. Method according to any of embodiments I to IV, wherein said epimerase-related like polypeptide as used herein comprises at least two conserved domains selected from the group comprising PyrdxIP-dep_Trfase_major (SSF53383), PTHR11601, Aminotran_5 (PF00266) and PyrdxIP-dep_Trfase_major_sub1 (G3DSA:3.40.640.10), and preferably comprises i) an Aminotran_5 conserved domain and ii) at least one
30 other conserved domain selected from the group comprising PyrdxIP-dep_Trfase_major (SSF53383), PTHR11601, and PyrdxIP-dep_Trfase_major_sub1 (G3DSA:3.40.640.10),

VI. Method according to any of embodiments I to IV, wherein said epimerase-related like polypeptide comprises a conserved domain with at least 50% sequence identity to a conserved domain as represented by any one of:

- amino acid coordinates 32 to 449 in SEQ ID NO:198 (motif 10);
- amino acid coordinates 47 to 276 in SEQ ID NO:198 (motif 11);
- amino acid coordinates 92 to 364 in SEQ ID NO:198 (motif 12); and
- amino acid coordinates 59 to 323 in SEQ ID NO:198 (motif 13).

40 VII. Method according to any of embodiments I to VI, wherein said nucleic acid encoding said epimerase-related 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 from *Populus trichocarpa*.

VIII. Method according to any one of embodiments I to VII, wherein said nucleic acid encoding said epimerase-related like polypeptide encodes any one of the polypeptides listed in Table A2 or is a portion of such a nucleic acid, or a nucleic acid capable of hybridising with such a nucleic acid.

IX. Method according to any one of embodiments I to VIII, wherein said nucleic acid sequence encodes an orthologue or paralogue of any of the polypeptides given in Table A2.

X. Method according to any one of embodiments I to IX, wherein said nucleic acid is operably linked to a GOS2 promoter, preferably to a GOS2 promoter from rice.

XI. Construct comprising:

(i) nucleic acid encoding an epimerase-related like polypeptide as defined in any of embodiments I and V to IX;

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

(iii) a transcription termination sequence.

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 enhanced yield-related traits relative to control plants, preferably increased yield relative to control plants, and more preferably increased seed yield and/or increased biomass relative to control plants, comprising:

(i) introducing and expressing in a plant cell or plant a nucleic acid encoding an epimerase-related like polypeptide as defined in any of embodiments I and V to IX; and

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

XIV. 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 an epimerase-related like polypeptide as defined in any of embodiments I and V to IX; or a transgenic plant cell derived from said transgenic plant.

XV. Use of a nucleic acid encoding an epimerase-related like polypeptide as defined in any of embodiments I and V to IX; 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.

Moreover concerning PLPCase polypeptides, the present invention relates to the following specific embodiments:

1. A method for enhancing yield-related traits in plants relative to control plants, comprising

modulating expression in a plant of a nucleic acid encoding a Phospholipase/carboxylesterase (PLPCase) polypeptide, wherein said PLPCase polypeptide comprises an InterPro accession IPR003140 corresponding to PFAM accession number PF022030 Phospholipase/carboxylesterase (PLPCase) domain.

- 5 2. Method according to embodiment 1, wherein said modulated expression is effected by introducing and expressing in a plant said nucleic acid encoding said PLPCase polypeptide.
3. Method according to embodiments 1 or 2, wherein said enhanced yield-related traits comprise increased yield relative to control plants, and preferably comprise increased
- 10 root biomass and/or increased seed yield relative to control plants.
4. Method according to any one of embodiments 1 to 3, wherein said enhanced yield-related traits are obtained under non-stress conditions.
5. Method according to any one of embodiments 1 to 3, wherein said enhanced yield-related traits are obtained under conditions of drought stress, salt stress or nitrogen
- 15 deficiency.
6. Method according to any of embodiments 1 to 5, wherein said PLPCase polypeptide comprises one or more of the following motifs:
 - (i) Motif 14: E[FY]G[KR]T[HY]VVRPKG[KR]HQATIVWLHGLGDNG[LSA]S[SW]SQLL [ED][ST]LPLPNIKWICPTA (SEQ ID NO: 348),
 - 20 (ii) Motif 15: PDD[WIVL]EGLDASAAH[IV]ANLLS[TS]EP[AS]D[VI]K[VL]G[IV]G (SEQ ID NO: 349),
 - (iii) Motif 16: FSMGAA[IT]ALYSA[TA]C[YF]A[MHL] (SEQ ID NO: 350)
7. Method according to any one of embodiments 1 to 6, wherein said nucleic acid encoding a PLPCase is of plant origin, preferably from a dicotyledonous plant, further
- 25 preferably from the family Fabaceae, more preferably from the genus Medicago, most preferably from Medicago truncatula.
8. Method according to any one of embodiments 1 to 7, wherein said nucleic acid encoding a PLPCase 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
- 30 nucleic acid.
9. Method according to any one of embodiments 1 to 8, wherein said nucleic acid sequence encodes an orthologue or paralogue of any of the polypeptides given in Table A3.
10. Method according to any one of embodiments 1 to 9, wherein said nucleic acid
- 35 encodes the polypeptide represented by SEQ ID NO: 285.
11. Method according to any one of embodiments 1 to 10, wherein said nucleic acid is operably linked to a constitutive promoter, preferably to a medium strength constitutive promoter, preferably to a plant promoter, more preferably to a GOS2 promoter, most preferably to a GOS2 promoter from rice.
- 40 12. Plant, plant part thereof, including seeds, or plant cell, obtainable by a method according to any one of embodiments 1 to 11, wherein said plant, plant part or plant cell comprises a recombinant nucleic acid encoding a PLPCase polypeptide as defined in

any of embodiments 1 and 6 to 10.

13. Construct comprising:

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

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

15. Use of a construct according to embodiment 13 or 14 in a method for making plants having enhanced yield-related traits, preferably increased yield relative to control plants, and more preferably increased root biomass, increased greenness index and/or increased seed yield relative to control plants.

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

17. Method for the production of a transgenic plant having enhanced yield-related traits relative to control plants, preferably increased yield relative to control plants, and more preferably increased root biomass, increased greenness index and/or increased seed yield relative to control plants, comprising:

- (i) introducing and expressing in a plant cell or plant a nucleic acid encoding a PLPCase polypeptide as defined in any of embodiments 1 and 6 to 10; and
- (ii) cultivating said plant cell or plant under conditions promoting plant growth and development.

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

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

20. Harvestable parts of a plant according to embodiment 18 or 19, wherein said harvestable parts are preferably root biomass and/or seeds.

21. Products derived from a plant according to embodiment 18 or 19 and/or from harvestable parts of a plant according to embodiment 20.

22. Use of a nucleic acid encoding a PLPCase polypeptide as defined in any of embodiments 1 and 6 to 10 for enhancing yield-related traits in plants relative to control plants, preferably for increasing yield, and more preferably for root biomass, increased

greenness index and/or increased seed yield in plants relative to control plants.

23. A method for the production of a product comprising the steps of growing the plants according to embodiment 16, 18 or 19 and producing said product from or by

(i) said plants; or

5 (ii) parts, including seeds, of said plants.

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

The present invention also relates to the following embodiments:

10 1. A method for enhancing yield-related traits in plants relative to control plants, comprising introducing and expressing in a plant of a nucleic acid encoding a DUF642 polypeptide, wherein said DUF642 polypeptide comprises the signature sequence FSAARTCAQ (SEQ ID NO: 194).

15 2. Method according to embodiment 1, wherein said nucleic acid encodes the polypeptide represented by SEQ ID NO: 2 or SEQ ID NO: 4 or a nucleic acid encoding a polypeptide having at least 90% sequence identity to SEQ ID NO: 2 and/or SEQ ID NO: 4.

20 3. 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.

25 4. Plant, plant part thereof, including seeds, or plant cell, obtainable by a method according to any one of embodiments 1 to 3, wherein said plant, plant part or plant cell comprises a recombinant nucleic acid encoding a DUF642 polypeptide as defined in any of embodiments 1 to 3.

30 5. Construct comprising:
(i) nucleic acid encoding a DUF642 as defined in any of embodiments 1 to 3;
(ii) one or more control sequences capable of driving expression of the nucleic acid sequence of (i); and optionally
(iii) a transcription termination sequence.

35 6. Construct according to embodiment 5, 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.

40 7. Use of a construct according to embodiment 5 or 6 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.

8. Plant, plant part or plant cell transformed with a construct according to embodiment 5 or 6.

5

9. 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 (i) introducing and expressing in a plant cell or plant a nucleic acid encoding a DUF642 polypeptide as defined in any of embodiments 1 to 3; and
(ii) cultivating said plant cell or plant under conditions promoting plant growth and development.

- 15 10. 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 DUF642 polypeptide as defined in any of embodiments 1 to 3 or a transgenic plant cell derived from said transgenic plant.

20

11. Transgenic plant according to embodiment 4, 8 or 10, 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.

25

12. Harvestable parts of a plant according to embodiment 11, wherein said harvestable parts are preferably shoot biomass and/or seeds.

30

13. Products derived from a plant according to embodiment 11 and/or from harvestable parts of a plant according to embodiment 12.

35

14. Use of a nucleic acid encoding a DUF642 polypeptide as defined in any of embodiments 1 to 3 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.

40

15. 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 epimerase-related like polypeptide by introducing and expressing in a plant said nucleic acid encoding said epimerase-related like polypeptide, wherein said epimerase-related like polypeptide has at least 80 % overall sequence identity to the amino acid represented by SEQ ID NO: 198.

16. Method according to embodiment 15, wherein said epimerase-related like polypeptide comprises one or more of the following motifs:

- (i) Motif 7: CREEGV[DE][QK]VFVD[AG]AH[AS]IG[QSC]V[PDE][VI][DN][VM][KR]
[ED]IGADFY[TV]SNLHKWFFCPP[SA]VAFL[YH] (SEQ ID NO: 273),
- (ii) Motif 8: EF[SA]HH[DN]P[GAN]VAR[IV]NNGSFG[CS]CP[AG]S[VI][LI]AAQ[ARK]
[RN]WQ[LR][LRQ]FL[RQA]QPD[AD]FYF[ND]xL[QRK][PK]G (SEQ ID NO: 274),
- (iii) Motif 9: S[LI]VDNATTAIAIVLQ[HQ][VAI][AG][WR][AS]FAEG[RKN]FA[KR][GN]
D[AVT]V[LV]MLH[YC]AY[GQ][AS]VKKSI[EQH]AYV (SEQ ID NO: 275)

17. Method according to embodiment 15 or 16, wherein said enhanced yield-related traits comprises increased biomass.

18. Method according to any one of embodiments 15 to 17, wherein said enhanced yield-related traits comprises increased seed yield relative to control plants.

19. Method according to any of embodiments 15 to 18, wherein said epimerase-related like polypeptide as used herein comprises at least two conserved domains selected from the group comprising PyrDXIP-dep_Trfase_major (SSF53383), PTHR11601, AminoTran_5 (PF00266) and PyrDXIP-dep_Trfase_major_sub1 (G3DSA:3.40.640.10), and preferably comprises an AminoTran_5 conserved domain and at least one other conserved domains selected from the group comprising PyrDXIP-dep_Trfase_major (SSF53383), PTHR11601, and PyrDXIP-dep_Trfase_major_sub1 (G3DSA:3.40.640.10),

20. Method according to any of embodiments 15 to 18, wherein said epimerase-related like polypeptide comprises a conserved domain with at least 50% sequence identity to a conserved domain as represented by any one of:

- amino acid coordinates 32 to 449 in SEQ ID NO:198 (motif 10);
- amino acid coordinates 47 to 276 in SEQ ID NO: 198 (motif 11);
- amino acid coordinates 92 to 364 in SEQ ID NO: 198 (motif 12); and
- amino acid coordinates 59 to 323 in SEQ ID NO: 198 (motif 13).

21. Method according to any of embodiments 15 to 20, wherein said nucleic acid encoding said epimerase-related 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 from *Populus trichocarpa*.

22. Method according to any one of embodiments 15 to 21, wherein said nucleic acid encoding said epimerase-related like polypeptide encodes any one of the polypeptides listed in Table A or is a portion of such a nucleic acid, or a nucleic acid capable of hybridising with such a nucleic acid.

23. Method according to any one of embodiments 15 to 22, wherein said nucleic acid sequence encodes an orthologue or paralogue of any of the polypeptides given in Table A2.
- 5
24. Method according to any one of embodiments 15 to 23, wherein said nucleic acid is operably linked to a GOS2 promoter, preferably to a GOS2 promoter from rice.
25. Construct comprising:
- 10 (i) nucleic acid encoding an epimerase-related like polypeptide as defined in any of embodiments 15 and 19 to 23;
- (ii) one or more control sequences capable of driving expression of the nucleic acid sequence of (i), and preferably a constitutive promoter, preferably to a medium strength constitutive promoter, preferably to a plant promoter, more preferably to
- 15 a GOS2 promoter, most preferably to a GOS2 promoter from rice, and optionally (iii) a transcription termination sequence.
26. Plant, plant part or plant cell transformed with a construct according to embodiment 25.
- 20 27. Method for the production of a transgenic plant having enhanced yield-related traits relative to control plants, preferably increased yield relative to control plants, and more preferably increased seed yield and/or increased biomass relative to control plants, comprising:
- (i) introducing and expressing in a plant cell or plant a nucleic acid encoding an
- 25 epimerase-related like polypeptide as defined in any of embodiments 15 and 19 to 23; and
- (ii) cultivating said plant cell or plant under conditions promoting plant growth and development.
- 30 28. 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 an epimerase-related like polypeptide as defined in any of embodiments 15 and 19 to 23; or a transgenic plant cell derived from said
- 35 transgenic plant.
29. Use of a nucleic acid encoding an epimerase-related like polypeptide as defined in any of embodiments 15 and 19 to 23; for enhancing yield-related traits in plants relative to control plants, preferably for increasing yield, and more preferably for increasing seed
- 40 yield and/or for increasing biomass in plants relative to control plants.
30. A method for enhancing yield-related traits in plants relative to control plants,

comprising introducing and expressing in a plant of a nucleic acid encoding SEQ ID NO: 285 or homologues having 80% sequence identity to SEQ ID NO: 285.

- 5 31. Method according to embodiment 30, wherein said enhanced yield-related traits comprise increased yield relative to control plants, and preferably comprise increased root biomass and/or increased seed yield relative to control plants.
- 10 32. Method according to any one of embodiments 30 to 31, wherein said enhanced yield-related traits are obtained under non-stress conditions.
- 15 33. Method according to any one of embodiments 30 to 31, wherein said enhanced yield-related traits are obtained under conditions of drought stress, salt stress or nitrogen deficiency.
- 20 34. Method according to any one of embodiments 30 to 33, wherein said nucleic acid encoding a PLPCase is of plant origin, preferably from a dicotyledonous plant, further preferably from the family Fabaceae, more preferably from the genus *Medicago*, most preferably from *Medicago truncatula*.
- 25 35. Method according to any one of embodiments 30 to 34, wherein said nucleic acid encoding a PLPCase 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, provided said nucleic acid has 80% sequence identity to SEQ ID NO: 285.
- 30 36. Method according to any one of embodiments 30 to 35, 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.
- 35 37. Plant, plant part thereof, including seeds, or plant cell, obtainable by a method according to any one of embodiments 30 to 36, wherein said plant, plant part or plant cell comprises a recombinant nucleic acid encoding a PLPCase polypeptide as defined in any of embodiments 30 and 35 to 39.
- 40 38. Construct comprising:
(i) nucleic acid encoding a PLPCase as defined in any of embodiments 30 and 34 to 36;
(ii) one or more control sequences capable of driving expression of the nucleic acid sequence of (i); and optionally
(iii) a transcription termination sequence.

39. Construct according to embodiment 38, 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.
- 5
40. Use of a construct according to embodiment 38 or 39 in a method for making plants having enhanced yield-related traits, preferably increased yield relative to control plants, and more preferably increased root biomass, increased greenness index and/or increased seed yield relative to control plants.
- 10
41. Plant, plant part or plant cell transformed with a construct according to embodiment 38 or 39.
42. 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 root biomass, increased greenness index and/or increased seed yield relative to control plants, comprising:
- 15
- a. introducing and expressing in a plant cell or plant a nucleic acid encoding a PLPCase polypeptide as defined in any of embodiments 30 and 34 to 36; and
- 20
- b. cultivating said plant cell or plant under conditions promoting plant growth and development.
43. 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 PLPCase polypeptide as defined in any of embodiments 30 and 34 to 36 or a transgenic plant cell derived from said transgenic plant.
- 25
44. Transgenic plant according to embodiment 37, 41 or 43, 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.
- 30
45. Harvestable parts of a plant according to embodiment 43 or 44, wherein said harvestable parts are preferably root biomass and/or seeds.
- 35
46. Products derived from a plant according to embodiment 43 or 44 and/or from harvestable parts of a plant according to embodiment 45.
- 40
47. Use of a nucleic acid encoding a PLPCase polypeptide as defined in any of embodiments 30 and 34 to 36 for enhancing yield-related traits in plants relative to

control plants, preferably for increasing yield, and more preferably for root biomass, increased greenness index and/or increased seed yield in plants relative to control plants.

5 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.

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

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

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

Table 1: Examples of conserved amino acid substitutions

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

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

mutagenesis protocols (see Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989 and yearly updates)).

Derivatives

5 "Derivatives" include peptides, oligopeptides, polypeptides which may, compared to the amino acid sequence of the naturally-occurring form of the protein, such as the protein of interest, comprise substitutions of amino acids with non-naturally occurring amino acid residues, or additions of non-naturally occurring amino acid residues. "Derivatives" of a protein also encompass peptides, oligopeptides, polypeptides which comprise naturally
10 occurring altered (glycosylated, acylated, prenylated, phosphorylated, myristoylated, sulphated etc.) or non-naturally altered amino acid residues compared to the amino acid sequence of a naturally-occurring form of the polypeptide. A derivative may also comprise one or more non-amino acid substituents or additions compared to the amino acid sequence from which it is derived, for example a reporter molecule or other ligand,
15 covalently or non-covalently bound to the amino acid sequence, such as a reporter molecule which is bound to facilitate its detection, and non-naturally occurring amino acid residues relative to the amino acid sequence of a naturally-occurring protein. Furthermore, "derivatives" also include fusions of the naturally-occurring form of the protein with tagging peptides such as FLAG, HIS6 or thioredoxin (for a review of tagging peptides, see Terpe,
20 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
25 positions can vary between homologues, amino acids that are highly conserved at specific positions indicate amino acids that are likely essential in the structure, stability or function of a protein. Identified by their high degree of conservation in aligned sequences of a family of protein homologues, they can be used as identifiers to determine if any polypeptide in question belongs to a previously identified polypeptide family.

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

35 Specialist databases exist for the identification of domains, for example, SMART (Schultz et al. (1998) Proc. Natl. Acad. Sci. USA 95, 5857-5864; Letunic et al. (2002) Nucleic Acids Res 30, 242-244), InterPro (Mulder et al., (2003) Nucl. Acids. Res. 31, 315-318), Prosite (Bucher and Bairoch (1994), A generalized profile syntax for biomolecular sequences motifs and its function in automatic sequence interpretation. (In) ISMB-94; Proceedings 2nd International Conference on Intelligent Systems for Molecular Biology. Altman R., Brutlag D., Karp P., Lathrop R., Searls D., Eds., pp53-61, AAAI Press, Menlo Park; Hulo et al.,

Nucl. Acids. Res. 32:D134-D137, (2004)), or Pfam (Bateman et al., Nucleic Acids Research 30(1): 276-280 (2002)). A set of tools for *in silico* analysis of protein sequences is available on the ExPASy proteomics server (Swiss Institute of Bioinformatics (Gasteiger et al., ExPASy: the proteomics server for in-depth protein knowledge and analysis, Nucleic Acids Res. 31:3784-3788(2003)). Domains or motifs may also be identified using routine techniques, such as by sequence alignment.

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

Reciprocal BLAST

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

same species as from which the query sequence is derived, and preferably results upon BLAST back in the query sequence being among the highest hits.

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

Hybridisation

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

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

The T_m is the temperature under defined ionic strength and pH, at which 50% of the target sequence hybridises to a perfectly matched probe. The T_m is dependent upon the solution conditions and the base composition and length of the probe. For example, longer

sequences hybridise specifically at higher temperatures. The maximum rate of hybridisation is obtained from about 16°C up to 32°C below T_m . The presence of monovalent cations in the hybridisation solution reduce the electrostatic repulsion between the two nucleic acid strands thereby promoting hybrid formation; this effect is visible for sodium concentrations of up to 0.4M (for higher concentrations, this effect may be ignored). Formamide reduces the melting temperature of DNA-DNA and DNA-RNA duplexes with 0.6 to 0.7°C for each percent formamide, and addition of 50% formamide allows hybridisation to be performed at 30 to 45°C, though the rate of hybridisation will be lowered. Base pair mismatches reduce the hybridisation rate and the thermal stability of the duplexes. On average and for large probes, the 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. Generally, suitable stringent conditions for nucleic acid hybridisation assays or gene amplification detection procedures are as set forth above. More or less stringent conditions

may also be selected. The skilled artisan is aware of various parameters which may be altered during washing and which will either maintain or change the stringency conditions.

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

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

Splice variant

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

Allelic variant

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

Endogenous gene

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

expression and/or substantial reduction of expression of the endogenous gene. The isolated gene may be isolated from an organism or may be manmade, for example by chemical synthesis.

5 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).

10

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
15 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
20 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
25 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.

30

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.

35

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
40 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	Van der Zaal et al., Plant Mol. Biol. 16, 983, 1991.

gene	
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	Mol Gen Genet 216:81-90, 1989; NAR 17:461-2, 1989

glutenin-1	
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 pyrophosphorylase	Trans Res 6:157-68, 1997
maize ESR gene family	Plant J 12:235-46, 1997
sorghum α -kafrin	DeRose et al., Plant Mol. Biol 32:1029-35, 1996
KNOX	Postma-Haarsma et al, Plant Mol. Biol. 39:257-71, 1999
rice oleosin	Wu et al, J. Biochem. 123:386, 1998
sunflower oleosin	Cummins et al., Plant Mol. Biol. 19: 873-876, 1992
PRO0117, putative rice 40S ribosomal protein	WO 2004/070039
PRO0136, rice alanine aminotransferase	unpublished
PRO0147, trypsin inhibitor ITR1 (barley)	unpublished
PRO0151, rice WSI18	WO 2004/070039
PRO0175, rice RAB21	WO 2004/070039
PRO005	WO 2004/070039
PRO0095	WO 2004/070039
α -amylase (Amy32b)	Lanahan et al, Plant Cell 4:203-211, 1992; Skriver et al, Proc Natl Acad Sci USA 88:7266-7270, 1991
cathepsin β -like gene	Cejudo et al, Plant Mol Biol 20:849-856, 1992
Barley Ltp2	Kalla et al., Plant J. 6:849-60, 1994
Chi26	Leah et al., Plant J. 4:579-89, 1994
Maize B-Peru	Selinger et al., Genetics 149:1125-38, 1998

Table 2d: examples of endosperm-specific promoters

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

Table 2e: Examples of embryo specific promoters:

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

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

- 20 The term “terminator” encompasses a control sequence which is a DNA sequence at the end of a transcriptional unit which signals 3' processing and polyadenylation of a primary transcript and termination of transcription. The terminator can be derived from the natural gene, from a variety of other plant genes, or from T-DNA. The terminator to be added may

be derived from, for example, the nopaline synthase or octopine synthase genes, or alternatively from another plant gene, or less preferably from any other eukaryotic gene.

Selectable marker (gene)/Reporter gene

5 "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
10 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,
15 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
20 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
25 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,
30 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
35 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).

40

Since the marker genes, particularly genes for resistance to antibiotics and herbicides, are no longer required or are undesired in the transgenic host cell once the nucleic acids have

been introduced successfully, the process according to the invention for introducing the nucleic acids advantageously employs techniques which enable the removal or excision of these marker genes. One such a method is what is known as co-transformation. The co-transformation method employs two vectors simultaneously for the transformation, one vector bearing the nucleic acid according to the invention and a second bearing the marker gene(s). A large proportion of transformants receives or, in the case of plants, comprises (up to 40% or more of the transformants), both vectors. In case of transformation with *Agrobacteria*, the transformants usually receive only a part of the vector, i.e. the sequence flanked by the T-DNA, which usually represents the expression cassette. The marker genes can subsequently be removed from the transformed plant by performing crosses. In another method, marker genes integrated into a transposon are used for the transformation together with desired nucleic acid (known as the Ac/Ds technology). The transformants can be crossed with a transposase source or the transformants are transformed with a nucleic acid construct conferring expression of a transposase, transiently or stable. In some cases (approx. 10%), the transposon jumps out of the genome of the host cell once transformation has taken place successfully and is lost. In a further number of cases, the transposon jumps to a different location. In these cases the marker gene must be eliminated by performing crosses. In microbiology, techniques were developed which make possible, or facilitate, the detection of such events. A further advantageous method relies on what is known as recombination systems; whose advantage is that elimination by crossing can be dispensed with. The best-known system of this type is what is known as the Cre/lox system. Cre1 is a recombinase that removes the sequences located between the loxP sequences. If the marker gene is integrated between the loxP sequences, it is removed once transformation has taken place successfully, by expression of the recombinase. Further recombination systems are the HIN/HIX, FLP/FRT and REP/STB system (Tribble et al., J. Biol. Chem., 275, 2000: 22255-22267; Velmurugan et al., J. Cell Biol., 149, 2000: 553-566). A site-specific integration into the plant genome of the nucleic acid sequences according to the invention is possible. Naturally, these methods can also be applied to microorganisms such as yeast, fungi or bacteria.

Transgenic/Transgene/Recombinant

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

- (a) the nucleic acid sequences encoding proteins useful in the methods of the invention, or
- (b) genetic control sequence(s) which is operably linked with the nucleic acid sequence according to the invention, for example a promoter, or
- (c) a) and b)

are not located in their natural genetic environment or have been modified by recombinant

methods, it being possible for the modification to take the form of, for example, a substitution, addition, deletion, inversion or insertion of one or more nucleotide residues. The natural genetic environment is understood as meaning the natural genomic or chromosomal locus in the original plant or the presence in a genomic library. In the case of a genomic library, the natural genetic environment of the nucleic acid sequence is preferably retained, at least in part. The environment flanks the nucleic acid sequence at least on one side and has a sequence length of at least 50 bp, preferably at least 500 bp, especially preferably at least 1000 bp, most preferably at least 5000 bp. A naturally occurring expression cassette – for example the naturally occurring combination of the natural promoter of the nucleic acid sequences with the corresponding nucleic acid sequence encoding a polypeptide useful in the methods of the present invention, as defined above – becomes a transgenic expression cassette when this expression cassette is modified by non-natural, synthetic ("artificial") methods such as, for example, mutagenic treatment. Suitable methods are described, for example, in US 5,565,350 or WO 00/15815.

A transgenic plant for the purposes of the invention is thus understood as meaning, as above, that the nucleic acids used in the method of the invention are not present in, or originating from, the genome of said plant, or are present in the genome of said plant but not at their natural locus in the genome of said plant, it being possible for the nucleic acids to be expressed homologously or heterologously. However, as mentioned, transgenic also means that, while the nucleic acids according to the invention or used in the inventive method are at their natural position in the genome of a plant, the sequence has been modified with regard to the natural sequence, and/or that the regulatory sequences of the natural sequences have been modified. Transgenic is preferably understood as meaning the expression of the nucleic acids according to the invention at an unnatural locus in the genome, i.e. homologous or, preferably, heterologous expression of the nucleic acids takes place. Preferred transgenic plants are mentioned herein.

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

Modulation

The term "modulation" means in relation to expression or gene expression, a process in which the expression level is changed by said gene expression in comparison to the control plant, the expression level may be increased or decreased. The original, unmodulated expression may be of any kind of expression of a structural RNA (rRNA, tRNA) or mRNA with subsequent translation. For the purposes of this invention, the original unmodulated expression may also be absence of any expression. The term "modulating the activity" shall mean any change of the expression of the inventive nucleic acid sequences or encoded

proteins, which leads to increased yield and/or increased growth of the plants. The expression can increase from zero (absence of, or immeasurable expression) to a certain amount, or can decrease from a certain amount to immeasurable small amounts or zero.

5 Expression

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

Increased expression/overexpression

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

Methods for increasing expression of genes or gene products are well documented in the art and include, for example, overexpression driven by appropriate promoters, the use of transcription enhancers or translation enhancers. Isolated nucleic acids which serve as promoter or enhancer elements may be introduced in an appropriate position (typically upstream) of a non-heterologous form of a polynucleotide so as to upregulate expression of a nucleic acid encoding the polypeptide of interest. For example, endogenous promoters may be altered in vivo by mutation, deletion, and/or substitution (see, Kmiec, US 5,565,350; Zarling et al., WO9322443), or isolated promoters may be introduced into a plant cell in the proper orientation and distance from a gene of the present invention so as to control the expression of the gene.

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

An intron sequence may also be added to the 5' untranslated region (UTR) or the coding sequence of the partial coding sequence to increase the amount of the mature message that accumulates in the cytosol. Inclusion of a spliceable intron in the transcription unit in both plant and animal expression constructs has been shown to increase gene expression at both the mRNA and protein levels up to 1000-fold (Buchman and Berg (1988) Mol. Cell biol. 8: 4395-4405; Callis et al. (1987) Genes Dev 1:1183-1200). Such intron enhancement

of gene expression is typically greatest when placed near the 5' end of the transcription unit. Use of the maize introns Adh1-S intron 1, 2, and 6, the Bronze-1 intron are known in the art. For general information see: The Maize Handbook, Chapter 116, Freeling and Walbot, Eds., Springer, N.Y. (1994).

5

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
10 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
15 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
20 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%,
25 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.

30 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
35 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
40 part thereof (in this case a stretch of substantially contiguous nucleotides derived from the gene of interest, or from any nucleic acid capable of encoding an orthologue, paralogue or homologue of the protein of interest), preferably capable of forming a hairpin structure. The

inverted repeat is cloned in an expression vector comprising control sequences. A non-coding DNA nucleic acid sequence (a spacer, for example a matrix attachment region fragment (MAR), an intron, a polylinker, etc.) is located between the two inverted nucleic acids forming the inverted repeat. After transcription of the inverted repeat, a chimeric RNA with a self-complementary structure is formed (partial or complete). This double-stranded RNA structure is referred to as the hairpin RNA (hpRNA). The hpRNA is processed by the plant into siRNAs that are incorporated into an RNA-induced silencing complex (RISC). The RISC further cleaves the mRNA transcripts, thereby substantially reducing the number of mRNA transcripts to be translated into polypeptides. For further general details see for example, Grierson et al. (1998) WO 98/53083; Waterhouse et al. (1999) WO 99/53050).

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

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

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

Another example of an RNA silencing method involves the use of antisense nucleic acid sequences. An "antisense" nucleic acid sequence comprises a nucleotide sequence that is complementary to a "sense" nucleic acid sequence encoding a protein, i.e. complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA

transcript sequence. The antisense nucleic acid sequence is preferably complementary to the endogenous gene to be silenced. The complementarity may be located in the "coding region" and/or in the "non-coding region" of a gene. The term "coding region" refers to a region of the nucleotide sequence comprising codons that are translated into amino acid residues. The term "non-coding region" refers to 5' and 3' sequences that flank the coding region that are transcribed but not translated into amino acids (also referred to as 5' and 3' untranslated regions).

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

The antisense nucleic acid sequence can be produced biologically using an expression vector into which a nucleic acid sequence has been subcloned in an antisense orientation (i.e., RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest). Preferably, production of antisense nucleic acid sequences in plants occurs by means of a stably integrated nucleic acid construct comprising a promoter, an operably linked antisense oligonucleotide, and a terminator.

The nucleic acid molecules used for silencing in the methods of the invention (whether introduced into a plant or generated in situ) hybridize with or bind to mRNA transcripts and/or genomic DNA encoding a polypeptide to thereby inhibit expression of the protein, e.g., by inhibiting transcription and/or translation. The hybridization can be by conventional

nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid sequence which binds to DNA duplexes, through specific interactions in the major groove of the double helix. Antisense nucleic acid sequences may be introduced into a plant by transformation or direct injection at a specific tissue site.

- 5 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
10 nucleic acid sequences can also be delivered to cells using the vectors described herein.

According to a further aspect, the antisense nucleic acid sequence is an a-anomeric nucleic acid sequence. An a-anomeric nucleic acid sequence forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual b-units, the strands run
15 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).

- 20 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
25 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
30 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).

35 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).

- 40 Gene silencing may also occur if there is a mutation on an endogenous gene and/or a mutation on an isolated gene/nucleic acid subsequently introduced into a plant. The

reduction or substantial elimination may be caused by a non-functional polypeptide. For example, the polypeptide may bind to various interacting proteins; one or more mutation(s) and/or truncation(s) may therefore provide for a polypeptide that is still able to bind interacting proteins (such as receptor proteins) but that cannot exhibit its normal function (such as signalling ligand).

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

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

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

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

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

Convenient tools for design and generation of amiRNAs and their precursors are also available to the public (Schwab et al., Plant Cell 18, 1121-1133, 2006).

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

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

Transformation

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

The transfer of foreign genes into the genome of a plant is called transformation. Transformation of plant species is now a fairly routine technique. Advantageously, any of several transformation methods may be used to introduce the gene of interest into a suitable ancestor cell. The methods described for the transformation and regeneration of

plants from plant tissues or plant cells may be utilized for transient or for stable transformation. Transformation methods include the use of liposomes, electroporation, chemicals that increase free DNA uptake, injection of the DNA directly into the plant, particle gun bombardment, transformation using viruses or pollen and microprojection.

5 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
10 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
15 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
20 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
25 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
30 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
35 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
40 and R. Wu, Academic Press, 1993, pp. 15-38.

In addition to the transformation of somatic cells, which then have to be regenerated into

intact plants, it is also possible to transform the cells of plant meristems and in particular those cells which develop into gametes. In this case, the transformed gametes follow the natural plant development, giving rise to transgenic plants. Thus, for example, seeds of *Arabidopsis* are treated with agrobacteria and seeds are obtained from the developing plants of which a certain proportion is transformed and thus transgenic [Feldman, KA and Marks MD (1987). *Mol Gen Genet* 208:1-9; Feldmann K (1992). In: C Koncz, N-H Chua and J Shell, eds, *Methods in Arabidopsis Research*. Word Scientific, Singapore, pp. 274-289]. Alternative methods are based on the repeated removal of the inflorescences and incubation of the excision site in the center of the rosette with transformed agrobacteria, whereby transformed seeds can likewise be obtained at a later point in time (Chang (1994). *Plant J.* 5: 551-558; Katavic (1994). *Mol Gen Genet*, 245: 363-370). However, an especially effective method is the vacuum infiltration method with its modifications such as the "floral dip" method. In the case of vacuum infiltration of *Arabidopsis*, intact plants under reduced pressure are treated with an agrobacterial suspension [Bechthold, N (1993). *C R Acad Sci Paris Life Sci*, 316: 1194-1199], while in the case of the "floral dip" method the developing floral tissue is incubated briefly with a surfactant-treated agrobacterial suspension [Clough, SJ and Bent AF (1998) *The Plant J.* 16, 735-743]. A certain proportion of transgenic seeds are harvested in both cases, and these seeds can be distinguished from non-transgenic seeds by growing under the above-described selective conditions. In addition the stable transformation of plastids is of advantages because plastids are inherited maternally is most crops reducing or eliminating the risk of transgene flow through pollen. The transformation of the chloroplast genome is generally achieved by a process which has been schematically displayed in Klaus et al., 2004 [*Nature Biotechnology* 22 (2), 225-229]. Briefly the sequences to be transformed are cloned together with a selectable marker gene between flanking sequences homologous to the chloroplast genome. These homologous flanking sequences direct site specific integration into the plastome. Plastidal transformation has been described for many different plant species and an overview is given in Bock (2001) *Transgenic plastids in basic research and plant biotechnology. J Mol Biol.* 2001 Sep 21; 312 (3):425-38 or Maliga, P (2003) *Progress towards commercialization of plastid transformation technology. Trends Biotechnol.* 21, 20-28. Further biotechnological progress has recently been reported in form of marker free plastid transformants, which can be produced by a transient co-integrated marker gene (Klaus et al., 2004, *Nature Biotechnology* 22(2), 225-229).

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

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

subjected to selective conditions so that transformed plants can be distinguished from untransformed plants. For example, the seeds obtained in the above-described manner can be planted and, after an initial growing period, subjected to a suitable selection by spraying. A further possibility consists in growing the seeds, if appropriate after sterilization, on agar plates using a suitable selection agent so that only the transformed seeds can grow into plants. Alternatively, the transformed plants are screened for the presence of a selectable marker such as the ones described above.

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

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

T-DNA activation tagging

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

TILLING

The term "TILLING" is an abbreviation of "Targeted Induced Local Lesions In Genomes" and refers to a mutagenesis technology useful to generate and/or identify nucleic acids encoding proteins with modified expression and/or activity. TILLING also allows selection of plants carrying such mutant variants. These mutant variants may exhibit modified expression, either in strength or in location or in timing (if the mutations affect the promoter

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

Homologous recombination

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

Yield related Trait(s)

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

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

Yield

The term “yield” in general means a measurable produce of economic value, typically related to a specified crop, to an area, and to a period of time. Individual plant parts directly

contribute to yield based on their number, size and/or weight, or the actual yield is the yield per square meter for a crop and year, which is determined by dividing total production (includes both harvested and appraised production) by planted square meters.

- 5 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.

10 Flowers in maize are unisexual; male inflorescences (tassels) originate from the apical stem and female inflorescences (ears) arise from axillary bud apices. The female inflorescence produces pairs of spikelets on the surface of a central axis (cob). Each of the female spikelets encloses two fertile florets, one of them will usually mature into a maize kernel once fertilized. Hence a yield increase in maize may be manifested as one or more of the following: increase in the number of plants established per square meter, an increase in the
15 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.

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

30

Early flowering time

Plants having an “early flowering time” as used herein are plants which start to flower earlier than control plants. Hence this term refers to plants that show an earlier start of flowering. Flowering time of plants can be assessed by counting the number of days (“time to flower”) between sowing and the emergence of a first inflorescence. The “flowering time” of a plant
35 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
40 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.

5

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.

10 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,
15 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.

20 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

25 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;
- 35 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
- 40 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
5 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
10 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
15 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
20 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
25 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
30 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
40 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.

5

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.

10 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
15 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).

20

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
25 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
30 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
35 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
40 (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.

10 Plant

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

Plants that are particularly useful in the methods of the invention include all plants which belong to the superfamily Viridiplantae, in particular monocotyledonous and dicotyledonous plants including fodder or forage legumes, ornamental plants, food crops, trees or shrubs selected from the list comprising *Acer* spp., *Actinidia* spp., *Abelmoschus* spp., *Agave sisalana*, *Agropyron* spp., *Agrostis stolonifera*, *Allium* spp., *Amaranthus* spp., *Ammophila arenaria*, *Ananas comosus*, *Annona* spp., *Apium graveolens*, *Arachis* spp., *Artocarpus* spp., *Asparagus officinalis*, *Avena* spp. (e.g. *Avena sativa*, *Avena fatua*, *Avena byzantina*, *Avena fatua* var. *sativa*, *Avena hybrida*), *Averrhoa carambola*, *Bambusa* sp., *Benincasa hispida*, *Bertholletia excelsea*, *Beta vulgaris*, *Brassica* spp. (e.g. *Brassica napus*, *Brassica rapa* ssp. [canola, oilseed rape, turnip rape]), *Cadaba farinosa*, *Camellia sinensis*, *Canna indica*, *Cannabis sativa*, *Capsicum* spp., *Carex elata*, *Carica papaya*, *Carissa macrocarpa*, *Carya* spp., *Carthamus tinctorius*, *Castanea* spp., *Ceiba pentandra*, *Cichorium endivia*, *Cinnamomum* spp., *Citrullus lanatus*, *Citrus* spp., *Cocos* spp., *Coffea* spp., *Colocasia esculenta*, *Cola* spp., *Corchorus* sp., *Coriandrum sativum*, *Corylus* spp., *Crataegus* spp., *Crocus sativus*, *Cucurbita* spp., *Cucumis* spp., *Cynara* spp., *Daucus carota*, *Desmodium* spp., *Dimocarpus longan*, *Dioscorea* spp., *Diospyros* spp., *Echinochloa* spp., *Elaeis* (e.g. *Elaeis guineensis*, *Elaeis oleifera*), *Eleusine coracana*, *Eragrostis tef*, *Erianthus* sp., *Eriobotrya japonica*, *Eucalyptus* sp., *Eugenia uniflora*, *Fagopyrum* spp., *Fagus* spp., *Festuca arundinacea*, *Ficus carica*, *Fortunella* spp., *Fragaria* spp., *Ginkgo biloba*, *Glycine* spp. (e.g. *Glycine max*, *Soja hispida* or *Soja max*), *Gossypium hirsutum*, *Helianthus* spp. (e.g. *Helianthus annuus*), *Hemerocallis fulva*, *Hibiscus* spp., *Hordeum* spp. (e.g. *Hordeum vulgare*), *Ipomoea batatas*, *Juglans* spp., *Lactuca sativa*, *Lathyrus* spp., *Lens culinaris*, *Linum usitatissimum*, *Litchi chinensis*, *Lotus* spp., *Luffa acutangula*, *Lupinus* spp., *Luzula sylvatica*, *Lycopersicon* spp. (e.g. *Lycopersicon esculentum*, *Lycopersicon lycopersicum*,

Lycopersicon pyriforme), *Macrotyloma* spp., *Malus* spp., *Malpighia emarginata*, *Mammea americana*, *Mangifera indica*, *Manihot* spp., *Manilkara zapota*, *Medicago sativa*, *Melilotus* spp., *Mentha* spp., *Miscanthus sinensis*, *Momordica* spp., *Morus nigra*, *Musa* spp., *Nicotiana* spp., *Olea* spp., *Opuntia* spp., *Ornithopus* spp., *Oryza* spp. (e.g. *Oryza sativa*,
5 *Oryza latifolia*), *Panicum miliaceum*, *Panicum virgatum*, *Passiflora edulis*, *Pastinaca sativa*, *Pennisetum* sp., *Persea* spp., *Petroselinum crispum*, *Phalaris arundinacea*, *Phaseolus* spp., *Phleum pratense*, *Phoenix* spp., *Phragmites australis*, *Physalis* spp., *Pinus* spp., *Pistacia vera*, *Pisum* spp., *Poa* spp., *Populus* spp., *Prosopis* spp., *Prunus* spp., *Psidium* spp., *Punica granatum*, *Pyrus communis*, *Quercus* spp., *Raphanus sativus*, *Rheum rhabarbarum*,
10 *Ribes* spp., *Ricinus communis*, *Rubus* spp., *Saccharum* spp., *Salix* sp., *Sambucus* spp., *Secale cereale*, *Sesamum* spp., *Sinapis* sp., *Solanum* spp. (e.g. *Solanum tuberosum*, *Solanum integrifolium* or *Solanum lycopersicum*), *Sorghum bicolor*, *Spinacia* spp., *Syzygium* spp., *Tagetes* spp., *Tamarindus indica*, *Theobroma cacao*, *Trifolium* spp., *Tripsacum dactyloides*, *Triticosecale rimpaui*, *Triticum* spp. (e.g. *Triticum aestivum*, *Triticum durum*, *Triticum turgidum*, *Triticum hybernum*, *Triticum macha*, *Triticum sativum*, *Triticum monococcum* or *Triticum vulgare*), *Tropaeolum minus*, *Tropaeolum majus*, *Vaccinium* spp.,
15 *Vicia* spp., *Vigna* spp., *Viola odorata*, *Vitis* spp., *Zea mays*, *Zizania palustris*, *Ziziphus* spp., amongst others.

20 Control plant(s)

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

30

Description of figures

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

Figure 1 represents the domain structure of **SEQ ID NO: 2** with conserved motifs.

Figure 2 represents the domain structure of **SEQ ID NO: 4** with conserved motifs.

35 **Figure 3** represents a multiple alignment of various DUF642 polypeptides. These alignments can be used for defining further motifs or signature sequences, when using conserved amino acids. The corresponding SEQ ID NO's for the aligned polypeptide sequences shown in **Figure 3** are:

Annotation	SEQ ID NO:
A.lyrata_340195	58
A.thaliana_AT1G80240.1	68

B.napus_TC65362	72
B.napus_TC77368	74
O.sativa_LOC_Os01g42520.1	6
P.virgatum_TC46275	18
S.bicolor_Sb03g027650.1	20
Z.mays_TC458850	36
Zea_mays_GRMZM2G034985_T02	42
T.aestivum_c60008104@11410	34
O.sativa_LOC_Os04g41710.1	8
O.sativa_LOC_Os04g41740.1	10
Os_DUF642.3	2
O.sativa_LOC_Os04g41750.1	12
S.bicolor_Sb06g021280.1	24
Z.mays_ZM07MC22756_BFb0088A01@22693	40
Zea_mays_GRMZM2G051898_T01	44
S.bicolor_Sb06g021290.1	26
O.sativa_LOC_Os04g41759.1	14
S.bicolor_Sb06g021320.1	32
S.bicolor_Sb06g021300.1	28
Z.mays_TC468877	38
S.bicolor_Sb06g021310.1	30
O.sativa_LOC_Os04g41770.1	16
S.bicolor_Sb06g021270.1	22
Zea_mays_GRMZM2G051912_T01	46
Zea_mays_GRMZM2G405071_T01	48
O.sativa_LOC_Os02g11040.1	50
Os_DUF642.2	4
P.virgatum_TC392	52
S.bicolor_Sb04g007160.1	54
Z.mays_TC467870	56
P.sitchensis_TA14284_3332	122
P.taeda_TA10429_3352	124
A.lyrata_912808	62
A.thaliana_AT1G29980.1	66
A.lyrata_934049	64

C.annuum_TC19321	78
S.tuberosum_TC189725	150
J.hindsii_x_regia_TA1156_432290	102
G.hirsutum_TC135941	84
G.raimondii_TC1124	96
G.max_GM06MC25068_saa05g06@24507	92
G.max_TC286277	94
L.japonicus_TC41787	104
M.domestica_TC32901	108
P.trifoliata_TA6477_37690	134
P. trichocarpa_DUF	158
P. trichocarpa_MUB2	160
T.officinale_TA3338_50225	156
H.vulgare_TC163732	98
T.aestivum_TC292522	152
O.sativa_LOC_Os01g55190.1	116
P.virgatum_TC7090	136
S.bicolor_Sb03g034960.1	142
officinarum_TC76182	148
Z.mays_c62082358gm030403@7546	172
Z.mays_ZM07MC26525_BFb0211C02@26447	176
Zea_mays_GRMZM2G324705_T01	180
Z.officinale_TA720_94328	178
P.sitchensis_TA10533_3332	120
P.taeda_TA4773_3352	126
O.sativa_LOC_Os03g59300.1	118
S.bicolor_Sb01g004270.1	140
Z.mays_c57213778gm030403@389	170
Z.mays_TC506958	174
P.trichocarpa_735032	130
V.vinifera_GSVIVT00025589001	166
S.lycopersicum_TC198542	144
Triphysaria_sp_TC11756	162
G.max_Glyma09g36220.1	88
M.truncatula_AC174305_19.4	110

G.max_Glyma15g43180.1	90
P.vulgaris_TC10932	138
A.lyrata_478145	60
A.thaliana_AT3G08030.1	70
B.napus_TC89135	76
G.hirsutum_TC161608	86
T.cacao_TC4474	154
L.serriola_TC5542	106
C.sinensis_TC3550	80
P.trichocarpa_549292	128
P.trichocarpa_833200	132
V.vinifera_GSVIVT00028751001	168
F.vesca_TA10399_57918	82
N.benthamiana_TC11106	112
N.tabacum_TC43533	114
S.lycopersicum_TC199778	146
Triphysaria_sp_TC1599	164
I.nil_TC4208	100

Figure 4 represents a multiple alignment of a selection (clade A) of DUF642 polypeptides from Figure 3. The corresponding SEQ ID NO's for the aligned polypeptide sequences shown in Figure 4 are:

Annotation	SEQ ID NO:
O.sativa_LOC_Os01g42520.1	6
P.virgatum_TC46275	18
S.bicolor_Sb03g027650.1	20
Z.mays_TC458850	36
Zea_mays_GRMZM2G034985_T02	42
T.aestivum_c60008104@11410	34
O.sativa_LOC_Os04g41710.1	8
O.sativa_LOC_Os04g41740.1	10
Os_DUF642.3	2
O.sativa_LOC_Os04g41750.1	12
SEQ ID NO: 24 for S.bicolor_Sb06g021280.1	24
Z.mays_ZM07MC22756_BFb0088A01@22693	40
Zea_mays_GRMZM2G051898_T01	44

S.bicolor_Sb06g021290.1	26
O.sativa_LOC_Os04g41759.1	14
S.bicolor_Sb06g021320.1	32
S.bicolor_Sb06g021300.1	28
Z.mays_TC468877	38
S.bicolor_Sb06g021310.1	30
O.sativa_LOC_Os04g41770.1	16
S.bicolor_Sb06g021270.1	22
Zea_mays_GRMZM2G051912_T01	46
Zea_mays_GRMZM2G405071_T01	48
O.sativa_LOC_Os02g11040.1	50
Os_DUF642.2	4
P.virgatum_TC392	52
S.bicolor_Sb04g007160.1	54
Z.mays_TC467870	56

Figure 5 shows a phylogenetic tree of DUF642 polypeptides, as described in Example 2. Clade A is encircled, the remainder of the branches of the tree are grouped in clade B.

Figure 6 shows the MATGAT table of Example 3.

- 5 **Figure 7** represents the binary vector used for increased expression in *Oryza sativa* of a DUF642-encoding nucleic acid under the control of a rice GOS2 promoter (pGOS2).

Figure 8 represents the domain structure of SEQ ID NO: 198 with conserved motifs.

- 10 **Figure 9** represents a multiple alignment of the epimerase-related like polypeptides as listed in table A. These alignments can be used for defining further motifs or signature sequences, when using conserved amino acids.

Figure 10 shows phylogenetic tree of epimerase-related like polypeptides. To obtain the tree, protein sequences were aligned with MAFFT, a bootstrapped Neighbour Joining (NJ) tree was calculated with QuickTree (100 repeats, uncorrected), which was visualized with Dendroscope (for further details on the used techniques see example 2)

- 15 **Figure 11** shows the MATGAT table of Example 3.

Figure 12 represents the binary vector used for increased expression in *Oryza sativa* of a nucleic acid encoding an epimerase-related like polypeptide under the control of a rice GOS2 promoter (pGOS2).

Figure 13 represents the domain structure of SEQ ID NO: 285 with conserved motifs.

- 20 **Figure 14** shows phylogenetic tree of PLPCase polypeptides, showing three clusters.

Figure 15 shows the MatGAT table of Example 3.

Figure 16 represents the binary vector used for increased expression in *Oryza sativa* of a PLPCase-encoding nucleic acid under the control of a rice GOS2 promoter (pGOS2).

Examples

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

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

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

1. DUF642 polypeptides

Sequences (full length cDNA, ESTs or genomic) related to SEQ ID NO: 1 and SEQ ID NO: 2; and SEQ ID NO: 3 and SEQ ID NO: 4 were identified amongst those maintained in the Entrez Nucleotides database at the National Center for Biotechnology Information (NCBI) using database sequence search tools, such as the Basic Local Alignment Tool (BLAST) (Altschul et al. (1990) J. Mol. Biol. 215:403-410; and Altschul et al. (1997) Nucleic Acids Res. 25:3389-3402). The program is used to find regions of local similarity between sequences by comparing nucleic acid or polypeptide sequences to sequence databases and by calculating the statistical significance of matches. For example, the polypeptide encoded by the nucleic acid of SEQ ID NO: 1 was used for the TBLASTN algorithm, with default settings and the filter to ignore low complexity sequences set off. The output of the analysis was viewed by pairwise comparison, and ranked according to the probability score (E-value), where the score reflect the probability that a particular alignment occurs by chance (the lower the E-value, the more significant the hit). In addition to E-values, comparisons were also scored by percentage identity. Percentage identity refers to the number of identical nucleotides (or amino acids) between the two compared nucleic acid (or polypeptide) sequences over a particular length. In some instances, the default parameters may be adjusted to modify the stringency of the search. For example the E-value may be increased to show less stringent matches. This way, short nearly exact matches may be identified.

Table A1 provides a list of nucleic acid and protein sequences related to SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3 and SEQ ID NO: 4.

Table A1: Examples of DUF642 nucleic acids and polypeptides:

Plant Source_name	Nucleic acid SEQ ID NO:	Protein SEQ ID NO:
O. sativa _DUF642.3	1	2
O. sativa _DUF642.2	3	4
O.sativa_LOC_Os01g42520.1	5	6
O.sativa_LOC_Os04g41710.1	7	8
O.sativa_LOC_Os04g41740.1	9	10
O.sativa_LOC_Os04g41750.1	11	12
O.sativa_LOC_Os04g41759.1	13	14
O.sativa_LOC_Os04g41770.1	15	16
P.virgatum_TC46275	17	18
S.bicolor_Sb03g027650.1	19	20
S.bicolor_Sb06g021270.1	21	22
S.bicolor_Sb06g021280.1	23	24
S.bicolor_Sb06g021290.1	25	26
S.bicolor_Sb06g021300.1	27	28
S.bicolor_Sb06g021310.1	29	30
S.bicolor_Sb06g021320.1	31	32
T.aestivum_c60008104@11410	33	34
Z.mays_TC458850	35	36
Z.mays_TC468877	37	38
Z.mays_ZM07MC22756_BFb0088A01@22693	39	40
Z. mays_GRMZM2G034985_T02	41	42
Z. mays_GRMZM2G051898_T01	43	44
Z. mays_GRMZM2G051912_T01	45	46
Z. mays_GRMZM2G405071_T01	47	48
O.sativa_LOC_Os02g11040.1	49	50
P.virgatum_TC392	51	52
S.bicolor_Sb04g007160.1	53	54
Z.mays_TC467870	55	56
A.lyrata_340195	57	58
A.lyrata_478145	59	60
A.lyrata_912808	61	62
A.lyrata_934049	63	64
A.thaliana_AT1G29980.1	65	66
A.thaliana_AT1G80240.1	67	68
A.thaliana_AT3G08030.1	69	70
B.napus_TC65362	71	72
B.napus_TC77368	73	74
B.napus_TC89135	75	76

C.annuum_TC19321	77	78
C.sinensis_TC3550	79	80
F.vesca_TA10399_57918	81	82
G.hirsutum_TC135941	83	84
G.hirsutum_TC161608	85	86
G.max_Glyma09g36220.1	87	88
G.max_Glyma15g43180.1	89	90
G.max_GM06MC25068_saa05g06@24507	91	92
G.max_TC286277	93	94
G.raimondii_TC1124	95	96
H.vulgare_TC163732	97	98
I.nil_TC4208	99	100
J.hindsii_x_regia_TA1156_432290	101	102
L.japonicus_TC41787	103	104
L.serriola_TC5542	105	106
M.domestica_TC32901	107	108
M.truncatula_AC174305_19.4	109	110
N.benthamiana_TC11106	111	112
N.tabacum_TC43533	113	114
O.sativa_LOC_Os01g55190.1	115	116
O.sativa_LOC_Os03g59300.1	117	118
P.sitchensis_TA10533_3332	119	120
P.sitchensis_TA14284_3332	121	122
P.taeda_TA10429_3352	123	124
P.taeda_TA4773_3352	125	126
P.trichocarpa_549292	127	128
P.trichocarpa_735032	129	130
P.trichocarpa_833200	131	132
P.trifoliata_TA6477_37690	133	134
P.virgatum_TC7090	135	136
P.vulgaris_TC10932	137	138
S.bicolor_Sb01g004270.1	139	140
S.bicolor_Sb03g034960.1	141	142
S.lycopersicum_TC198542	143	144
S.lycopersicum_TC199778	145	146
S.officinatum_TC76182	147	148
S.tuberosum_TC189725	149	150
T.aestivum_TC292522	151	152
T.cacao_TC4474	153	154
T.officinale_TA3338_50225	155	156
P.trichocarpa_DUF	157	158

P. trichocarpa PP_MUB2	159	160
Triphysaria_sp_TC11756	161	162
Triphysaria_sp_TC1599	163	164
V.vinifera_GSVIVT00025589001	165	166
V.vinifera_GSVIVT00028751001	167	168
Z.mays_c57213778gm030403@389	169	170
Z.mays_c62082358gm030403@7546	171	172
Z.mays_TC506958	173	174
Z.mays_ZM07MC26525_BFb0211C02@26447	175	176
Z.officinale_TA720_94328	177	178
Z. mays_GRMZM2G324705_T01	179	180

2. epimerase-related like polypeptides

Sequences (full length cDNA, ESTs or genomic) related to SEQ ID NO: 197 and SEQ ID NO: 198 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: 197 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 SEQ ID NO: 197 and SEQ ID NO: 198.and a list of nucleic acid and protein sequences related to SEQ ID NO: 197 and SEQ ID NO: 198.

Table A2: Examples of an epimerase-related like nucleic acids and polypeptides:

Plant Source	SEQ ID NO: protein	SEQ ID NO: nucleic acid
Poptr_epimerase-related	198	197
Z.mays_ZM07MC24472_BFb0052A17@24402	200	199
A.lyrata_486720	202	201

A.lyrata_489502	204	203
A.thaliana_AT3G62130.1	206	205
A.thaliana_AT5G26600.1	208	207
Aquilegia_sp_TC26581	210	209
B.napus_TC63646	212	211
C.clementina_TC23968	214	213
C.sinensis_TC2568	216	215
G.hirsutum_TC132670	218	217
G.max_Glyma02g01070.1	220	219
G.max_Glyma03g38740.1	222	221
G.max_Glyma19g41330.1	224	223
G.max_Glyma20g22430.1	226	225
H.vulgare_TC158694	228	227
M.truncatula_AC151948_16.5	230	229
M.truncatula_AC166315_9.4	232	231
O.sativa_LOC_Os01g18640.1	234	233
O.sativa_LOC_Os01g18660.1	236	235
P.patens_175806	238	237
P.patens_TC39863	240	239
P.persica_TC12771	242	241
P.trichocarpa_570656	244	243
P.trichocarpa_572528	246	245
S.bicolor_Sb03g011840.1	248	247
S.bicolor_Sb03g011840.2	250	249
S.bicolor_Sb03g011840.3	252	251
S.bicolor_Sb03g011850.1	254	253
S.lycopersicum_TC191918	256	255
S.moellendorffii_232064	258	257
S.moellendorffii_413706	260	259
S.tuberosum_TC167698	262	261
T.aestivum_TC278728	264	263
V.vinifera_GSVIVT00028240001	266	265
Z.mays_TC513613	268	267
Zea_mays_GRMZM2G077212_T01	270	269
Zea_mays_GRMZM2G322186_T01	272	271

3. PLPCase polypeptides

Sequences (full length cDNA, ESTs or genomic) related to SEQ ID NO: 284 and SEQ ID NO: 285 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: 284 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 a list of nucleic acid and protein sequences related to SEQ ID NO: 284 and SEQ ID NO: 285.

Table A3: Examples of PLPCase nucleic acids and polypeptides:

Plant Source	Nucleic acid SEQ ID NO:	Protein SEQ ID NO:
M.truncatula_PLPCase	284	285
A.thaliana_AT3G15650.1#1	286	287
A.thaliana_AT1G52470.1#1	288	289
G.max_Glyma13g43990.1#1	290	291
G.max_Glyma08g22420.1#1	292	293
G.max_Glyma07g03670.1#1	294	295
A.thaliana_AT1G52700.1#1	296	297
G.max_GM06MC00001_45093526@1#1	298	299
P.trichocarpa_TC89207#1	300	301
T.aestivum_c55164328@10990#1	302	303
Z.mays_c58236581gm030403@14151#1	304	305
H.vulgare_BI948365#1	306	307
H.annuus_TC40709#1	308	309
B.napus_TC92944#1	310	311
B.napus_BN06MC12702_43481925@12666#1	312	313
A.thaliana_AT5G20060.1#1	314	315

B.napus_BN06MC00524_46010623@523#1	316	317
H.annuus_TC45565#1	318	319
M.truncatula_AC148176_15.5#1	320	321
O.sativa_LOC_Os01g07960.1#1	322	323
Z.mays_TC480822#1	324	325
P.trichocarpa_scaff_VI.1824#1	326	327
P.trichocarpa_809106#1	328	329
G.max_TC317555#1	330	331
G.max_Glyma04g07280.1#1	332	333
Z.mays_TC563901#1	334	335
Z.mays_ZM07MC31087_BFb0267E17@30994#1	336	337
T.aestivum_TC296583#1	338	339
Zea_mays_GRMZM2G481362_T01#1	340	341
O.sativa_LOC_Os01g42690.1#1	342	343
O.sativa_LOC_Os05g51050.1#1	344	345
Z.mays_ZM07MC23122_BFb0221H24@23057#1	346	347

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. DUF642 polypeptides

Alignment of polypeptide sequences was performed using the AlignX alignment of Vector NTI, Invitrogen, 10.3 (2006) with gap opening penalty of 10.0, gap extension penalty of 0.05 and gap separation penalty range 8. The DUF642 polypeptides are aligned in Figure 2, showing DUF642 polypeptides of clades A and B. Figure 3 shows an alignment of the DUF642 polypeptides of clade A (as shown in Figure 5).

A phylogenetic tree of DUF642 polypeptides (Figure 5) was constructed by aligning DUF642 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.

2. epimerase-related like polypeptides

Alignment of polypeptide sequences was performed using MAFFT (version 6.624, L-INS-I method - Katoh and Toh (2008) - Briefings in Bioinformatics 9:286-298). Minor manual editing was done to further optimize the alignment. A representative number of epimerase-related like polypeptides are aligned in Figure 9.

A phylogenetic tree of epimerase-related like polypeptides (Figure 10) was constructed by aligning epimerase-related like sequences using MAFFT (Katoh and Toh (2008) - Briefings in Bioinformatics 9:286-298). A neighbour-joining tree was calculated using Quick-Tree (Howe et al. (2002), Bioinformatics 18(11): 1546-7), 100 bootstrap repetitions. The dendrogram was drawn using Dendroscope (Huson et al. (2007), BMC Bioinformatics 8(1):460). Confidence levels for 100 bootstrap repetitions are indicated for major branchings.

3. PLPCase 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 was done to further optimise the alignment.

A phylogenetic tree of PLPCase polypeptides (Figure 14) was constructed by aligning PLPCase sequences using MAFFT (Katoh and Toh (2008) - Briefings in Bioinformatics 9:286-298). A neighbour-joining tree was calculated using Quick-Tree (Howe et al. (2002), Bioinformatics 18(11): 1546-7), 100 bootstrap repetitions. A circular dendrogram was drawn using Dendroscope 2.0.1 (Huson et al. (2007), BMC Bioinformatics 8(1):460). Confidence levels for 100 bootstrap repetitions are indicated for major branchings.

Example 3: Calculation of global percentage identity between polypeptide sequences

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

1. DUF642 polypeptides

Results of an analysis for clade A are shown in Figure 6 for the global similarity and identity over the full length of the polypeptide sequences. Sequence similarity is shown in the

bottom half of the dividing line and sequence identity is shown in the top half of the diagonal dividing line. Parameters used in the comparison were: Scoring matrix: Blosum62, First Gap: 12, Extending Gap: 2. The sequence identity (in %) between the DUF642 polypeptide sequences of clade A+B useful in performing the methods of the invention can be as low as 29,7 % (is generally higher than 29,7%) compared to SEQ ID NO: 2 or 4. The sequence identity (in %) between the DUF642 polypeptide sequences of clade A useful in performing the methods of the invention can be as low as 38,4 % (is generally higher than 38,4%) compared to SEQ ID NO: 2 or 4.

10 **Table B1:** Description of proteins in figure 6:

1. Os_DUF642.3
2. Os_DUF642.2
3. O.sativa_LOC_Os01g42520.1
4. O.sativa_LOC_Os04g41710.1
5. O.sativa_LOC_Os04g41740.1
6. O.sativa_LOC_Os04g41750.1
7. O.sativa_LOC_Os04g41759.1
8. O.sativa_LOC_Os04g41770.1
9. P.virgatum_TC46275
10. S.bicolor_Sb03g027650.1
11. S.bicolor_Sb06g021270.1
12. S.bicolor_Sb06g021280.1
13. S.bicolor_Sb06g021290.1
14. S.bicolor_Sb06g021300.1
15. S.bicolor_Sb06g021310.1
16. S.bicolor_Sb06g021320.1
17. T.aestivum_c60008104@11410
18. Z.mays_TC458850
19. Z.mays_TC468877
20. Z.mays_ZM07MC22756_BFb0088A01@22693
21. Zea_mays_GRMZM2G034985_T02
22. Zea_mays_GRMZM2G051898_T01
23. Zea_mays_GRMZM2G051912_T01
24. Zea_mays_GRMZM2G405071_T01
25. O.sativa_LOC_Os02g11040.1
26. P.virgatum_TC392
27. S.bicolor_Sb04g007160.1
28. Z.mays_TC467870

2. epimerase-related like polypeptides

Results of the analysis are shown in Figure 11 for the global similarity and identity over the full length of the polypeptide sequences. Sequence similarity is shown in the bottom half of the dividing line and sequence identity is shown in the top half of the diagonal dividing line.

- 5 Parameters used in the comparison were: Scoring matrix: Blosum62, First Gap: 12, Extending Gap: 2. The sequence identity (in %) between the epimerase-related like polypeptide sequences useful in performing the methods of the invention is generally higher than 40% compared to SEQ ID NO: 198.
- 10 A MATGAT table for local alignment of a specific domain, or data on % identity/similarity between specific domains can also be made.

3. PLPCase polypeptides

- 15 Results of the analysis are shown in Figure 15 for the global similarity and identity over the full length of the polypeptide sequences. Sequence similarity is shown in the bottom half of the dividing line and sequence identity is shown in the top half of the diagonal dividing line. Parameters used in the comparison were: Scoring matrix: Blosum62, First Gap: 12, Extending Gap: 2. The sequence identity (in %) between the PLPCase polypeptide sequences useful in performing the methods of the invention can be as low as 24%, but is
- 20 generally higher than 24%, compared to SEQ ID NO: 285.

Table B2: Description of proteins in figure 15:

1. M.truncatula_PLPCase
2. A.thaliana_AT3G15650.1#1
3. A.thaliana_AT1G52470.1#1
4. G.max_Glyma13g43990.1#1
5. G.max_Glyma08g22420.1#1
6. G.max_Glyma07g03670.1#1
7. A.thaliana_AT1G52700.1#1
8. G.max_GM06MC00001_45093526@1#1
9. P.trichocarpa_TC89207#1
10. T.aestivum_c55164328@10990#1
11. Z.mays_c58236581gm030403@14151#1
12. H.vulgare_BI948365#1
13. H.annuus_TC40709#1
14. B.napus_TC92944#1
15. B.napus_BN06MC12702_43481925@12666#1
16. A.thaliana_AT5G20060.1#1
17. B.napus_BN06MC00524_46010623@523#1
18. H.annuus_TC45565#1
19. M.truncatula_AC148176_15.5#1
20. O.sativa_LOC_Os01g07960.1#1

21. Z.mays_TC480822#1
22. P.trichocarpa_scaff_VI.1824#1
23. P.trichocarpa_809106#1
24. G.max_TC317555#1
25. G.max_Glyma04g07280.1#1
26. Z.mays_TC563901#1
27. Z.mays_ZM07MC31087_BFb0267E17@30994#1
28. T.aestivum_TC296583#1
29. Zea_mays_GRMZM2G481362_T01#1
30. O.sativa_LOC_Os01g42690.1#1
31. O.sativa_LOC_Os05g51050.1#1
32. Z.mays_ZM07MC23122_BFb0221H24@23057#1

As can be seen from the MatGAT table in figure 15, PLPCase can be separated into three clusters: Cluster A: containing PLPCase of dicot origin, Cluster B: containing PLPCase of dicot and monocot origin and Cluster C: containing PLPCase of monocot origin.

5

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.

10

1. DUF642 polypeptides

The results of the InterPro scan (InterPro database, release 4.7) of the polypeptide sequence as represented by SEQ ID NO: 2 and 4, respectively, are presented in Table C1.

20

Table C1: InterPro scan results (major accession numbers) of the polypeptide sequence as represented by SEQ ID NO: 2 and 4.

Interpro ID	Domain ID	Domain name	Amino acid coordinates on SEQ ID NO 2	Amino acid coordinates on SEQ ID NO 4
IPR006946	PF04862	DUF642	51-371	1-292

25

In an embodiment a DUF642 polypeptide comprises a conserved domain (or motif) with at least 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%,

85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to a conserved domain from amino acid 51 to 371 in SEQ ID NO:2 or from amino acid 28 to 194 of SEQ ID NO: 4.

5 2. epimerase-related like polypeptides

The results of the InterPro scan (InterPro database, release 30.0) of the polypeptide sequence as represented by SEQ ID NO: 198 are presented in Table C2.

Table C2: InterPro scan results (major accession numbers) of the polypeptide sequence as represented by SEQ ID NO: 198.

Database	Accession number	Accession name	Amino acid coordinates on SEQ ID NO:198	E -value	accession
superfamily	SSF53383	PyrdxIP-dep_Trfase_major (Pyridoxal phosphate-dependent transferase, major domain)	[32-449]	1,2 E-55	IPR015424
HMMPanther	PTHR11601	PTHR11601	[47-276]	7,89 E-15	NULL
HMMPfam	PF00266	Aminotran_5 (Aminotransferase, class V/Cysteine desulfurase Biological Process: metabolic process (GO:0008152))	[92-364]	6,7 E-18	IPR000192
Gene3D	G3DSA:3.40.640.10	PyrdxIP-dep_Trfase_major_sub1 (Pyridoxal phosphate-dependent transferase, major region, subdomain 1)	[59-323]	1,09 E-42	IPR015421

An epimerase-related 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 consists of a conserved domain selected from any one of:

- amino acid coordinates 32 to 449 in SEQ ID NO: 198;
- amino acid coordinates 47 to 276 in SEQ ID NO: 198;
- amino acid coordinates 92 to 364 in SEQ ID NO: 198; and
- amino acid coordinates 59 to 323 in SEQ ID NO: 198.

3. PLPCase polypeptides

The results of the InterPro scan (InterPro database, version 4.6) of the polypeptide sequence as represented by SEQ ID NO: 285 are presented in Table C3.

Table C3: InterPro scan results (major accession numbers) of the polypeptide sequence as represented by SEQ ID NO: 285.

Database	Accession number	Domain name Short name	e-value [Amino acid coordinates on SEQ ID NO 285]
IPR003140 InterPro	PF02230 PFAM	Phospholipase/carboxylesterase Abhydrolase_2	6.2e-41 [24-248]T
Unintegrated	G3DSA:3.40.50.1820 GENE3D	Unintegrated NA	5.3e-56 [12-250]T
	PTHR10655 PANTHER	Unintegrated <i>LYSOPHOSPHOLIPASE-RELATED</i>	1.3e-94 [60-254]T
	PTHR10655:SF5 PANTHER	Unintegrated <i>LYSOPHOSPHOLIPASE-RELATED</i>	1.3e-94 [60-254]T
	SSF53474 SUPERFAMILY	Unintegrated <i>alpha/beta-Hydrolases</i>	2e-42 [21-251]T

In an embodiment a PLPCase polypeptide comprises a conserved domain or motif with at least 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to a conserved domain from amino acid 24 to 248 in SEQ ID NO: 285.

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

TargetP 1.1 predicts the subcellular location of eukaryotic proteins. The location assignment is based on the predicted presence of any of the N-terminal pre-sequences: chloroplast transit peptide (cTP), mitochondrial targeting peptide (mTP) or secretory pathway signal peptide (SP). Scores on which the final prediction is based are not really probabilities, and they do not necessarily add to one. However, the location with the highest score is the most

likely according to TargetP, and the relationship between the scores (the reliability class) may be an indication of how certain the prediction is. The reliability class (RC) ranges from 1 to 5, where 1 indicates the strongest prediction. For the sequences predicted to contain an N-terminal presequence a potential cleavage site can also be predicted. TargetP is maintained at the server of the Technical University of Denmark.

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

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

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

Example 6: Cloning of the nucleic acid sequence used in methods of the invention

1. DUF642 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.

In a first experiment A, the primers used were prm16414 (SEQ ID NO: 190; sense): 5'-ggggacaagttgtacaaaaaagcaggctaaacaatgagaagaaaaacagagatg-3' and prm16415 (SEQ ID NO: 191; reverse, complementary): 5'-ggggaccactttgtacaagaaagctgggtataatattacagaaggaaccg-3', which include the AttB sites for Gateway recombination.

In a second experiment B, the primers used were prm16412 (SEQ ID NO: 192; sense): 5'-ggggacaagttgtacaaaaaagcaggctaaacaatggtgctggtgtccc-3' and prm16413 (SEQ ID NO: 193; reverse, complementary): 5'-ggggaccactttgtacaagaaagctgggttagtacgacatcgcgga-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", pDUF642. Plasmid pDONR201 was purchased from Invitrogen, as part of the Gateway® technology.

In the first experiment A, 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: 189) for constitutive expression was located upstream of this Gateway cassette.

In the second experiment B, the entry clone comprising SEQ ID NO: 3 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: 189) for constitutive expression was located upstream of this Gateway cassette.

In a further experiment, all steps as described above were executed for the cloning of SEQ ID NO: 159. Primers used were prm13182 (SEQ ID NO: 195; sense): 5'-ggggacaagttgtacaaaaagcaggcttaacaatgtttcaagct caacaaga-3' and prm13183 (SEQ ID NO: 196; reverse, complementary): 5'-ggggaccactttgtta caagaaagctgggtctcttctcccatcacaac 3', which include the AttB sites for Gateway recombination.

The entry clone comprising SEQ ID NO: 159 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: 189) for constitutive expression was located upstream of this Gateway cassette.

After the LR recombination step, the resulting expression vectors pGOS2::DUF642 (Figure 7) were transformed into *Agrobacterium* strain LBA4044 according to methods well known in the art.

2. epimerase-related like polypeptides

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 prm18733 (SEQ ID NO: 276; sense, start codon in bold): 5' ggggacaagttgtacaaaaagcaggcttaacaatgcaagaaaactccaaagat 3' and prm18734 (SEQ ID NO: 277; reverse, complementary): 5' ggggaccactttgtacaagaaagctg ggtgcataccttctcactctggat 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”, pERL (“ERL” stands for “epimerase-related like”). Plasmid pDONR201 was purchased from Invitrogen, as part of the Gateway® technology.

The entry clone comprising SEQ ID NO: 197 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: 278) for constitutive expression was located upstream of this Gateway cassette.

After the LR recombination step, the resulting expression vector pGOS2::ERL (Figure 12) was transformed into *Agrobacterium* strain LBA4044 according to methods well known in the art.

3. PLPCase polypeptides

The nucleic acid sequence was amplified by PCR using as template a custom-made *Medicago truncatula* 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 prm15409 (SEQ ID NO: 362; sense, start codon in bold): 5'-ggggacaagttgtacaaaaagcaggcttaaca**atg**agtcacgcgcattctc-3' and prm15410 (SEQ ID NO: 363; reverse, complementary): 5'-ggggaccacttgtacaagaaa gctgggtaa aatagtttgcatttacgacg-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”, pPLPCase. Plasmid pDONR201 was purchased from Invitrogen, as part of the Gateway® technology.

The entry clone comprising SEQ ID NO: 284 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: 361) for constitutive expression was located upstream of this Gateway cassette.

After the LR recombination step, the resulting expression vector pGOS2:: PLPCase (Figure 16) 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: 286 was prepared using primers prm15411 (SEQ ID NO: 364, sense): 5'- ggggacaagttgtacaaaaagcaggcttaacaa
5 tgagctattctcgtaaagc-3' and prm15412 (SEQ ID NO: 365; reverse, complementary): 5'-
ggggaccactttgtacaagaaagctgggtgctgagaagatgcttgtaaa-3'.

Example 7: Plant transformation

Rice transformation

10 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
15 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. Alternatively, the incubation step in ethanol
was followed by a 30 minutes incubation in 0.2% HgCl₂, followed by a 6 times 15 minutes
wash with sterile distilled water. The sterile seeds were then germinated on a medium
20 containing 2,4-D (callus induction medium). After incubation in the dark for four weeks,
embryogenic, scutellum-derived calli were excised and propagated on the same medium.
After two weeks, the calli were multiplied or propagated by subculture on the same medium
for another 2 weeks. Embryogenic callus pieces were sub-cultured on fresh medium 3 days
before co-cultivation (to boost cell division activity).

25 *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
30 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. Alternatively co-cultivated calli were grown on 2,4-D-containing medium for
35 4 weeks in the dark at 28°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
40 humidity and short days in a greenhouse.

Transformation of rice cultivar indica can also be done in a similar way as give above

according to techniques well known to a skilled person.

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

Example 8: Transformation of other crops

Corn transformation

Transformation of maize (*Zea mays*) is performed with a modification of the method described by Ishida et al. (1996) Nature Biotech 14(6): 745-50. Transformation is genotype-dependent in corn and only specific genotypes are amenable to transformation and regeneration. The inbred line A188 (University of Minnesota) or hybrids with A188 as a parent are good sources of donor material for transformation, but other genotypes can be used successfully as well. Ears are harvested from corn plant approximately 11 days after pollination (DAP) when the length of the immature embryo is about 1 to 1.2 mm. Immature embryos are cocultivated with *Agrobacterium tumefaciens* containing the expression vector, and transgenic plants are recovered through organogenesis. Excised embryos are grown on callus induction medium, then maize regeneration medium, containing the selection agent (for example imidazolinone but various selection markers can be used). The Petri plates are incubated in the light at 25 °C for 2-3 weeks, or until shoots develop. The green shoots are transferred from each embryo to maize rooting medium and incubated at 25 °C for 2-3 weeks, until roots develop. The rooted shoots are transplanted to soil in the greenhouse. T1 seeds are produced from plants that exhibit tolerance to the selection agent and that contain a single copy of the T-DNA insert.

Wheat transformation

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

Soybean transformation

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

Rapeseed/canola transformation

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

Alfalfa transformation

A regenerating clone of alfalfa (*Medicago sativa*) is transformed using the method of (McKersie et al., 1999 Plant Physiol 119: 839–847). Regeneration and transformation of alfalfa is genotype dependent and therefore a regenerating plant is required. Methods to obtain regenerating plants have been described. For example, these can be selected from the cultivar Rangelander (Agriculture Canada) or any other commercial alfalfa variety as described by Brown DCW and A Atanassov (1985. Plant Cell Tissue Organ Culture 4: 111-

112). Alternatively, the RA3 variety (University of Wisconsin) has been selected for use in tissue culture (Walker et al., 1978 Am J Bot 65:654-659). Petiole explants are cocultivated with an overnight culture of *Agrobacterium tumefaciens* C58C1 pMP90 (McKersie et al., 1999 Plant Physiol 119: 839–847) or LBA4404 containing the expression vector. The explants are cocultivated for 3 d in the dark on SH induction medium containing 288 mg/ L Pro, 53 mg/ L thioproline, 4.35 g/ L K₂SO₄, and 100 µm acetosyringinone. The explants are washed in half-strength Murashige-Skoog medium (Murashige and Skoog, 1962) and plated on the same SH induction medium without acetosyringinone but with a suitable selection agent and suitable antibiotic to inhibit *Agrobacterium* growth. After several weeks, somatic embryos are transferred to BOi2Y development medium containing no growth regulators, no antibiotics, and 50 g/ L sucrose. Somatic embryos are subsequently germinated on half-strength Murashige-Skoog medium. Rooted seedlings were transplanted into pots and grown in a greenhouse. T1 seeds are produced from plants that exhibit tolerance to the selection agent and that contain a single copy of the T-DNA insert.

Cotton transformation

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

Sugarbeet transformation

Seeds of sugarbeet (*Beta vulgaris* L.) are sterilized in 70% ethanol for one minute followed by 20 min. shaking in 20% Hypochlorite bleach e.g. Clorox® regular bleach (commercially available from Clorox, 1221 Broadway, Oakland, CA 94612, USA). Seeds are rinsed with

sterile water and air dried followed by plating onto germinating medium (Murashige and Skoog (MS) based medium (Murashige, T., and Skoog, ., 1962. *Physiol. Plant*, vol. 15, 473-497) including B5 vitamins (Gamborg et al.; *Exp. Cell Res.*, vol. 50, 151-8.) supplemented with 10 g/l sucrose and 0,8% agar). Hypocotyl tissue is used essentially for the initiation of shoot cultures according to Hussey and Hepher (Hussey, G., and Hepher, A., 1978. *Annals of Botany*, 42, 477-9) and are maintained on MS based medium supplemented with 30g/l sucrose plus 0,25mg/l benzylamino purine and 0,75% agar, pH 5,8 at 23-25°C with a 16-hour photoperiod. *Agrobacterium tumefaciens* strain carrying a binary plasmid harbouring a selectable marker gene, for example nptII, is used in transformation experiments. One day before transformation, a liquid LB culture including antibiotics is grown on a shaker (28°C, 150rpm) until an optical density (O.D.) at 600 nm of ~1 is reached. Overnight-grown bacterial cultures are centrifuged and resuspended in inoculation medium (O.D. ~1) including Acetosyringone, pH 5,5. Shoot base tissue is cut into slices (1.0 cm x 1.0 cm x 2.0 mm approximately). Tissue is immersed for 30s in liquid bacterial inoculation medium. Excess liquid is removed by filter paper blotting. Co-cultivation occurred for 24-72 hours on MS based medium incl. 30g/l sucrose followed by a non-selective period including MS based medium, 30g/l sucrose with 1 mg/l BAP to induce shoot development and cefotaxim for eliminating the *Agrobacterium*. After 3-10 days explants are transferred to similar selective medium harbouring for example kanamycin or G418 (50-100 mg/l genotype dependent). Tissues are transferred to fresh medium every 2-3 weeks to maintain selection pressure. The very rapid initiation of shoots (after 3-4 days) indicates regeneration of existing meristems rather than organogenesis of newly developed transgenic meristems. Small shoots are transferred after several rounds of subculture to root induction medium containing 5 mg/l NAA and kanamycin or G418. Additional steps are taken to reduce the potential of generating transformed plants that are chimeric (partially transgenic). Tissue samples from regenerated shoots are used for DNA analysis. Other transformation methods for sugarbeet are known in the art, for example those by Linsey & Gallois (Linsey, K., and Gallois, P., 1990. *Journal of Experimental Botany*; vol. 41, No. 226; 529-36) or the methods published in the international application published as WO9623891A.

Sugarcane transformation

Spindles are isolated from 6-month-old field grown sugarcane plants (Arencibia et al., 1998. *Transgenic Research*, vol. 7, 213-22; Enriquez-Obregon et al., 1998. *Planta*, vol. 206, 20-27). Material is sterilized by immersion in a 20% Hypochlorite bleach e.g. Clorox® regular bleach (commercially available from Clorox, 1221 Broadway, Oakland, CA 94612, USA) for 20 minutes. Transverse sections around 0,5cm are placed on the medium in the top-up direction. Plant material is cultivated for 4 weeks on MS (Murashige, T., and Skoog, ., 1962. *Physiol. Plant*, vol. 15, 473-497) based medium incl. B5 vitamins (Gamborg, O., et al., 1968. *Exp. Cell Res.*, vol. 50, 151-8) supplemented with 20g/l sucrose, 500 mg/l casein hydrolysate, 0,8% agar and 5mg/l 2,4-D at 23°C in the dark. Cultures are transferred after 4 weeks onto identical fresh medium. *Agrobacterium tumefaciens* strain carrying a binary plasmid harbouring a selectable marker gene, for example hpt, is used in transformation

experiments. One day before transformation, a liquid LB culture including antibiotics is grown on a shaker (28°C, 150rpm) until an optical density (O.D.) at 600 nm of ~0,6 is reached. Overnight-grown bacterial cultures are centrifuged and resuspended in MS based inoculation medium (O.D. ~0,4) including acetosyringone, pH 5,5. Sugarcane embryogenic callus pieces (2-4 mm) are isolated based on morphological characteristics as compact structure and yellow colour and dried for 20 min. in the flow hood followed by immersion in a liquid bacterial inoculation medium for 10-20 minutes. Excess liquid is removed by filter paper blotting. Co-cultivation occurred for 3-5 days in the dark on filter paper which is placed on top of MS based medium incl. B5 vitamins containing 1 mg/l 2,4-D. After co-cultivation calli are washed with sterile water followed by a non-selective cultivation period on similar medium containing 500 mg/l cefotaxime for eliminating remaining *Agrobacterium* cells. After 3-10 days explants are transferred to MS based selective medium incl. B5 vitamins containing 1 mg/l 2,4-D for another 3 weeks harbouring 25 mg/l of hygromycin (genotype dependent). All treatments are made at 23°C under dark conditions. Resistant calli are further cultivated on medium lacking 2,4-D including 1 mg/l BA and 25 mg/l hygromycin under 16 h light photoperiod resulting in the development of shoot structures. Shoots are isolated and cultivated on selective rooting medium (MS based including, 20g/l sucrose, 20 mg/l hygromycin and 500 mg/l cefotaxime). Tissue samples from regenerated shoots are used for DNA analysis. Other transformation methods for sugarcane are known in the art, for example from the in-ternational application published as WO2010/151634A and the granted European patent EP1831378.

Example 9: Phenotypic evaluation procedure

9.1 Evaluation setup

35 to 90 independent T0 rice transformants were generated. The primary transformants were transferred from a tissue culture chamber to a greenhouse for growing and harvest of T1 seed. Six events, of which the T1 progeny segregated 3:1 for presence/absence of the transgene, were retained. For each of these events, approximately 10 T1 seedlings containing the transgene (hetero- and homo-zygotes) and approximately 10 T1 seedlings lacking the transgene (nullizygotes) were selected by monitoring visual marker expression. The transgenic plants and the corresponding nullizygotes were grown side-by-side at random positions. Greenhouse conditions were of 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.

5 *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.

15

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.

20

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.

25

9.2 Statistical analysis: F test

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

30

When two experiments with overlapping events are carried out, a combined analysis is

performed. This is useful to check consistency of the effects over the two experiments, and if this is the case, to accumulate evidence from both experiments in order to increase confidence in the conclusion. The method used is a mixed-model approach that takes into account the multilevel structure of the data (i.e. experiment - event - segregants). P values are obtained by comparing likelihood ratio test to chi square distributions.

9.3 Parameters measured

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

Biomass-related parameter measurement

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

Increase in root biomass is expressed as an increase in total root biomass ("rootMax") (measured as maximum biomass of roots observed during the lifespan of a plant); or as an increase in the root/shoot index ("RootShInd"), measured as the ratio between root mass and shoot mass in the period of active growth of root and shoot. In other words, the root/shoot index is defined as the ratio of the rapidity of root growth to the rapidity of shoot growth in the period of active growth of root and shoot. Root biomass can be determined using a method as described in WO 2006/029987.

Parameters related to development time

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

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

The "time to flower" or "flowering time" ("TimetoFlower") of the plant can be determined

using the method as described in WO 2007/093444.

Seed-related parameter measurements

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

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

The total seed weight ("totalwgseeds") was measured by weighing all filled husks harvested from a plant.

The total number of seeds (or florets = "nrtotalseeds") 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 ("totalwgseeds/nrfilledseeds * 1000").

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

The number of flowers per panicle ("flowerperpan") as defined in the present invention is the ratio between the total number of seeds over the number of mature primary panicles (= "nrtotalseeds/firstpan" and "firstpan" = number of panicles in the first flush).

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

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

1. DUF642 polypeptides

The results of the evaluation of transgenic rice plants expressing a DUF642 nucleic acid under non-stress conditions are presented below.

In the first experiment A using SEQ ID NO: 1, an increase of at least 5 % was observed for total seed yield (Totalwgseeds), number of seeds (nrfilledseed), seed fill rate (fillrate) and harvest index.

The results of the evaluation of transgenic rice plants in the T1 generation and expressing a nucleic acid encoding the DUF642 polypeptide of SEQ ID NO: 2 under non-stress conditions are presented below in Table D1. When grown under non-stress conditions, an increase of at least 5 % was observed for total seed yield (Totalwgseeds), number of seeds

(nrfilledseed), seed fill rate (fillrate) and harvest index. Moreover, two lines showed an increase in aboveground biomass (AreaMax), and one line showed an increase of Maximum height of the gravity center of the leafy biomass (GravityYMax).

Table D1: Data summary for transgenic rice plants; for each parameter, the overall percent increase is shown for the T1 generation), for each parameter the p-value is <0.05.

Parameter	Overall increase
totalwgseeds	16.5
nrfilledseed	15.8
fillrate	14.1
harvestindex	15.0

In the second experiment B using SEQ ID NO: 3, an increase of at least 5 % was observed for total seed yield (Totalwgseeds), seed fill rate (fillrate) and greenness before flowering.

The results of the evaluation of transgenic rice plants in the T1 generation and expressing a nucleic acid encoding the DUF642 polypeptide of SEQ ID NO: 4 under non-stress conditions are presented below in Table D2. When grown under non-stress conditions, an increase of at least 5 % was observed for total seed yield (Totalwgseeds), seed fill rate (fillrate) and greenness before flowering. Moreover, two lines showed an increase in aboveground biomass (AreaMax), Height of the highest tip of the plant (HeightMax) and Maximum height of the gravity center of the leafy biomass (GravityYMax).

Table D2: Data summary for transgenic rice plants; for each parameter, the overall percent increase is shown for the T1 generation, for each parameter the p-value is <0.05.

Parameter	Overall increase
totalwgseeds	7.2
fillrate	8.8
Greenness before flowering	3.8

In a further experiment, SEQ ID NO: 159 (in clade B) was used instead of SEQ ID NO: 1 in a drought screen as described above. The results of the evaluation of transgenic rice plants in the T1 generation and expressing a nucleic acid encoding the DUF642 polypeptide of SEQ ID NO: 160 under drought conditions showed an increase of at least 5 % total number of seeds (nrtotalseed) and for aboveground biomass (AreaMax).

In yet another experiment, rice plants transformed with SEQ ID NO: 173 under control of the rice GOS2 promoter were used in a nitrogen use efficiency screen described above.

The transgenic rice plants showed an increased fillrate compared to the control plants (overall increase for the 6 tested T1 lines was 10%, p<0.05).

2. epimerase-related like polypeptides

The results of the evaluation of transgenic rice plants of the T1 generation expressing the nucleic acid encoding an epimerase-related like polypeptide of SEQ ID NO: 198 under non-stress conditions are presented below in **Table D3**.

5

When grown under non-stress conditions, an increase of at least 5 % was observed for parameters such as seed yield, including total weight of seeds (totalwgseeds), harvest index, and fill rate.

- 10 **Table D3:** 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
totalwgseeds	8.6
fillrate	5.6
harvestindex	6.9

- 15 In addition, it could be noted that transgenic rice plants of at least one event showed an increase of at least 3% for the parameter thousand kernel weight (TKW) as compared to control plants.

Transgenic rice plants of at least two events showed an increase of at least 5% (and in particular on average 8 and 11% increase, respectively) in biomass, particularly above ground biomass(AreaMax), as compared to control plants.

20 Furthermore, there was a clear tendency for transgenic plants of at least four events to have an increase of at least 5% in number of filled seeds (nrfilledseed) as compared to control plants.

3. PLPCase polypeptides

25 The results of the evaluation of transgenic rice plants in the T1 generation and expressing a nucleic acid comprising the longest Open Reading Frame in SEQ ID NO: 284 under nitrogen deficiency-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 encoding the PLPCase polypeptide of SEQ ID NO: 285 under stress conditions are presented below in Table D4. When grown under nitrogen deficiency-stress conditions, an increase of at least 5 % was observed for root biomass (RootMax, which is the maximal rootmass measured during the life of the plant and RootThickMax, which is the maximal mass of thick roots measured during the life of the plant).

35

Table D4: Data summary for transgenic rice plants; for each parameter, the overall percent increase is shown for the T1 generation, for each parameter the p-value is <0.05.

Parameter	Overall increase
RootMax	10.4
RootThickMax	6.0

5 In addition, plants expressing the above PLPCase showed two lines which were clearly positive for greenness index; three lines were clearly positive for roothinmax, which is the maximal mass of thin roots measured during the life of the plant; and one line was clearly positive for number of filled seeds.

10 The evaluation of transgenic rice plants expressing a nucleic acid encoding the PLPCase polypeptide of SEQ ID NO: 287 under non-stress conditions are presented below in Table D5. When grown under non-stress conditions, an increase of at least 5 % was observed for root biomass (Root/Shoot Index), and for seed yield (fill rate). In addition, two plant lines expressing the *FeS* nucleic acid of SEQ ID NO: 286 showed increased total root biomass (resp increase of 14% and 11%, $p \leq 0.1$) and a third line showed increased amount of
15 biomass of roots below a certain thickness threshold (RootThin, +7%, $p \leq 0.1$).

Table D5: Data summary for transgenic rice plants; for each parameter, the overall percent increase is shown, for each parameter the p-value is <0.05.

Parameter	Overall increase
Root/Shoot Index	11.9
fillrate	8.5

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 DUF642 polypeptide, wherein said DUF642 polypeptide comprises an InterPro accession IPR006946 DUF642 domain, preferably said DUF642 polypeptide comprises the signature sequence FSAARTCAQ (SEQ ID NO: 194).
2. Method according to claim 1, wherein said modulated expression is effected by introducing and expressing in a plant said nucleic acid encoding said DUF642 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 or under abiotic stress conditions.
5. Method according to any of claims 1 to 4, wherein said DUF642 polypeptide comprises one or more of the following motifs:
 - (i) Motif 1: N[LAM][VL]KNG[DG]FEEGP[WYH][MI][FLI]P[NG][TS][STR][WL]GVL[LVI]P[PTS][NKM][LQDV][EDV][DE][ED][THY]S[PS]L[PS][GP]W[IMT][IV] (SEQ ID NO: 181),
 - (ii) Motif 2: [VLA][EKAT][KPR]G[SAM][LRH]Y[SA][LIV][TI]F[SG]A[AS]RTCAQ[DLAS]E[SVRL]L[NR][VI] (SEQ ID NO: 182),
 - (iii) Motif 3: D[PE][AT]CGP[LI][IL]D[AS][VIF]AI[KR] (SEQ ID NO: 183)
6. Method according to any one of claims 1 to 5, wherein said nucleic acid encoding a DUF642 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 DUF642 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.
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.
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 SEQ ID NO: 4 or a nucleic acid encoding a polypeptide having at least 90% sequence identity to SEQ ID NO: 2 and/or SEQ ID NO: 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 DUF642 polypeptide as defined in any of claims 1 and 5 to 9.

12. Construct comprising:

- (i) nucleic acid encoding a DUF642 as defined in any of claims 1 and 5 to 9;
- 5 (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 DUF642 polypeptide as defined in any of claims 1 and 5 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 DUF642 polypeptide as defined in any of claims 1 and 5 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 DUF642 polypeptide as defined in any of claims 1 and 5 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. Plant having increased yield, particularly increased biomass and/or increased seed yield, relative to control plants, resulting from modulated expression of a nucleic acid encoding a DUF642 polypeptide, or a transgenic plant cell originating from or being part of said transgenic plant.
- 5 23. A method for the production of a product comprising the steps of growing the plants of the invention and producing said product from or by
- a. the plants of the invention; or
 - b. parts, including seeds, of these plants.
- 10 24. Plant according to claim 11, 15, or 21, or a transgenic plant cell originating thereof, or a method according to claim 22, wherein said plant is a crop plant, preferably a dicot such as sugar beet, alfalfa, trefoil, chicory, carrot, cassava, cotton, soybean, canola or a monocot, such as sugarcane, or a cereal, such as rice, maize, wheat, barley, millet, rye, triticale, sorghum emmer, spelt, secale, einkorn, teff, milo and oats.
25. Construct according to claim 12 or 13 comprised in a plant cell.
- 15 26. Recombinant chromosomal DNA comprising the construct according to claim 12 or 13.
27. 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 epimerase-related like polypeptide, wherein said epimerase-related like polypeptide comprises one or more of the following motifs:
- 20 (i) Motif 7: CREEGV[DE][QK]VFVD[AG]AH[AS]IG[QSC]V[PDE][VI][DN][VM][KR][ED]IGADFY[TV]SNLHKWFFCPP[SA]VAFL[YH] (SEQ ID NO: 273),
 - (ii) Motif 8: EF[SA]HH[DN]P[GAN]VAR[IV]NNGSFG[CS]CP[AG]S[VI][LI]AAQ[ARK][RN]WQ[LR][LRQ]FL[RQA]QPD[AD]FYF[ND]xL[QRK][PK]G (SEQ ID NO: 274),
 - (iii) Motif 9: S[LI]VDNATTAAAIVLQ[HQ][VAI][AG][WR][AS]FAEG[RKN]FA[KR][GN]25 D[AVT]V[LV]MLH[YC]AY[GQ][AS]VKKSI[EQH]AYV (SEQ ID NO: 275)
28. Method according to claim 27, wherein said modulated expression is effected by introducing and expressing in a plant said nucleic acid encoding said epimerase-related like polypeptide.
29. Method according to claim 27 or 28, wherein said enhanced yield-related traits comprise30 increased yield relative to control plants, and preferably comprise increased biomass and/or increased seed yield relative to control plants.
30. Method according to any of claims 27 to 29, wherein said epimerase-related like polypeptide as used herein comprises a conserved domain with at least 50% sequence identity to a conserved domain as represented by any one of:
- 35 - amino acid coordinates 32 to 449 in SEQ ID NO:198 (motif 10);
 - amino acid coordinates 47 to 276 in SEQ ID NO:198 (motif 11);
 - amino acid coordinates 92 to 364 in SEQ ID NO:198 (motif 12); and
 - amino acid coordinates 59 to 323 in SEQ ID NO:198 (motif 13).
31. Method according to any of claims 27 or 30, wherein said nucleic acid encodes a40 polypeptide having at least 40 % overall sequence identity to the amino acid represented by SEQ ID NO: 198.

32. Method according to any of claims 27 to 31, wherein said nucleic acid encoding said epimerase-related 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 from *Populus trichocarpa*.
- 5 33. Method according to any one of claims 27 to 32, wherein said nucleic acid encoding said epimerase-related like polypeptide encodes any one of the polypeptides listed in Table A2 or is a portion of such a nucleic acid, or a nucleic acid capable of hybridising with such a nucleic acid.
34. Method according to any one of claims 27 to 33, wherein said nucleic acid sequence
10 encodes an orthologue or paralogue of any of the polypeptides given in Table A2.
35. Method according to any one of claims 27 to 34, wherein said enhanced yield-related traits are obtained under non-stress conditions.
36. Method according to any one of claims 27 to 34, wherein said enhanced yield-related traits are obtained under conditions of drought stress, salt stress or nitrogen deficiency.
- 15 37. Method according to any one of claims 27 to 36, 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.
38. Plant, plant part thereof, including seeds, or plant cell, obtainable by a method according
20 to any one of claims 27 to 37, wherein said plant, plant part or plant cell comprises a recombinant nucleic acid encoding an epimerase-related like polypeptide as defined in any of claims 27 and 30 to 34.
39. Construct comprising:
- 25 (i) nucleic acid encoding an epimerase-related like polypeptide as defined in any of claims 27 and 30 to 34;
- (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. Construct according to claim 39, wherein one of said control sequences is a constitutive
30 promoter, preferably a medium strength constitutive promoter, preferably to a plant promoter, more preferably a GOS2 promoter, most preferably a GOS2 promoter from rice.
41. Use of a construct according to claim 39 or 40 in a method for making plants having enhanced yield-related traits, preferably increased yield relative to control plants, and
35 more preferably increased seed yield and/or increased biomass relative to control plants.
42. Plant, plant part or plant cell transformed with a construct according to claim 39 or 40.
43. 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
40 preferably increased seed yield and/or increased biomass relative to control plants, comprising:

- (i) introducing and expressing in a plant cell or plant a nucleic acid encoding an epimerase-related like polypeptide as defined in any of claims 27 and 30 to 34; and
- (ii) cultivating said plant cell or plant under conditions promoting plant growth and development.

44. Transgenic plant having enhanced yield-related traits relative to control plants, preferably increased yield relative to control plants, and more preferably increased seed yield and/or increased biomass, resulting from modulated expression of a nucleic acid encoding an epimerase-related like polypeptide as defined in any of claims 27 and 30 to 34; or a transgenic plant cell derived from said transgenic plant.

45. Transgenic plant according to claim 38, 42 or 44, 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.

46. Harvestable parts of a plant according to any of claims 38, 42, 44 or 45, wherein said harvestable parts are preferably shoot biomass and/or seeds.

47. Products derived from a plant according to any of claims 38, 42, 44 or 45 and/or from harvestable parts of a plant according to claim 46.

48. Isolated nucleic acid molecule selected from:

- (i) a nucleic acid represented by SEQ ID NO: 199 ;
- (ii) the complement of a nucleic acid represented by SEQ ID NO: 199;
- (iii) a nucleic acid encoding epimerase-related 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 SEQ ID NO: 200, 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 7 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.

49. Isolated polypeptide selected from:

- (i) an amino acid sequence represented by SEQ ID NO: 200;
- (ii) an amino acid sequence having, in increasing order of preference, at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the amino acid

- sequence represented by SEQ ID NO: 200, 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 7 to 13, and further preferably conferring enhanced yield-related traits relative to control plants;
- 5 (iii) derivatives of any of the amino acid sequences given in (i) or (ii) above.
50. Use of a nucleic acid encoding an epimerase-related like polypeptide as defined in any of claims 27 and 30 to 34 and 49, or a nucleic acid as defined in claim 48 for enhancing yield-related traits in plants relative to control plants, preferably for increasing yield, and
- 10 more preferably for increasing seed yield and/or for increasing biomass in plants relative to control plants.
51. Use of a nucleic acid encoding an epimerase-related like polypeptide as defined in any of claims 27 and 30 to 34 and 49 or a nucleic acid as defined in claim 48 as molecular marker.
- 15 52. A method for the production of a product comprising the steps of growing the plants according to claim 38, 42 or 44 and producing said product from or by
- (i) said plants; or
- (ii) parts, including seeds, of said plants.
53. Construct according to claim 39 or 40 comprised in a plant cell.
- 20 54. Recombinant chromosomal DNA comprising the construct according to claim 39 or 40 comprised.
55. 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 Phospholipase/carboxylesterase (PLPCase) polypeptide, wherein said PLPCase
- 25 polypeptide comprises an InterPro accession IPR003140 corresponding to PFAM accession number PF022030 Phospholipase/carboxylesterase (PLPCase) domain.
56. Method according to claim 55, wherein said modulated expression is effected by introducing and expressing in a plant said nucleic acid encoding said PLPCase polypeptide.
- 30 57. Method according to claims 55 or 56, wherein said enhanced yield-related traits comprise increased yield relative to control plants, and preferably comprise increased root biomass and/or increased seed yield relative to control plants.
58. Method according to any one of claims 55 to 57, wherein said enhanced yield-related traits are obtained under non-stress conditions.
- 35 59. Method according to any one of claims 55 to 57, wherein said enhanced yield-related traits are obtained under conditions of drought stress, salt stress or nitrogen deficiency.
60. Method according to any of claims 55 to 59, wherein said PLPCase polypeptide comprises one or more of the following motifs:
- (i) Motif 14: E[FY]G[KR]T[HY]VVRPKG[KR]HQATIVWLHGLGDNG[LSA]S[SW]SQLL [ED][ST]LPLPNIKWICPTA (SEQ ID NO: 348),
- 40 (ii) Motif 15: PDD[WIVL]EGLDASAAH[IV]ANLLS[TS]EP[AS]D[VI]K[VL]G[IV]G (SEQ ID NO: 349),

(iii) Motif 16: FSMGAA[IT]ALYSA[TA]C[YF]A[MHL] (SEQ ID NO: 350)

61. Method according to any one of claims 55 to 60, wherein said nucleic acid encoding a PLPCase is of plant origin, preferably from a dicotyledonous plant, further preferably from the family Fabaceae, more preferably from the genus *Medicago*, most preferably from *Medicago truncatula*.

62. Method according to any one of claims 55 to 61, wherein said nucleic acid encoding a PLPCase 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.

63. Method according to any one of claims 55 to 62, wherein said nucleic acid sequence encodes an orthologue or paralogue of any of the polypeptides given in Table A3.

64. Method according to any one of claims 55 to 63, wherein said nucleic acid encodes the polypeptide represented by SEQ ID NO: 285.

65. Method according to any one of claims 55 to 64, 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.

66. Plant, plant part thereof, including seeds, or plant cell, obtainable by a method according to any one of claims 55 to 65, wherein said plant, plant part or plant cell comprises a recombinant nucleic acid encoding a PLPCase polypeptide as defined in any of claims 55 and 60 to 64.

67. Construct comprising:

- (i) nucleic acid encoding a PLPCase as defined in any of claims 55 and 60 to 64;
- (ii) one or more control sequences capable of driving expression of the nucleic acid sequence of (i); and optionally
- (iii) a transcription termination sequence.

68. Construct according to claim 67, 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.

69. Use of a construct according to claim 67 or 68 in a method for making plants having enhanced yield-related traits, preferably increased yield relative to control plants, and more preferably increased root biomass, increased greenness index and/or increased seed yield relative to control plants.

70. Plant, plant part or plant cell transformed with a construct according to claim 67 or 68.

71. 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 root biomass, increased greenness index and/or increased seed yield relative to control plants, comprising:

- (i) introducing and expressing in a plant cell or plant a nucleic acid encoding a PLPCase polypeptide as defined in any of claims 55 and 60 to 64; and
- (ii) cultivating said plant cell or plant under conditions promoting plant growth and development.

72. 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 PLPCase polypeptide as defined in any of claims 55 and 60 to 64 or a transgenic plant cell derived from said transgenic plant.
73. Transgenic plant according to claim 66, 70 or 72, 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.
74. Harvestable parts of a plant according to claim 72 or 73, wherein said harvestable parts are preferably root biomass and/or seeds.
75. Products derived from a plant according to claim 72 or 73 and/or from harvestable parts of a plant according to claim 74.
76. Use of a nucleic acid encoding a PLPCase polypeptide as defined in any of claims 55 and 60 to 64 for enhancing yield-related traits in plants relative to control plants, preferably for increasing yield, and more preferably for root biomass, increased greenness index and/or increased seed yield in plants relative to control plants.
77. A method for the production of a product comprising the steps of growing the plants according to claim 70, 72 or 73 and producing said product from or by
- (i) said plants; or
 - (ii) parts, including seeds, of said plants.
78. Construct according to claim 67 or 68 comprised in a plant cell.

1/77

MRRKTEMIQTTPCVALLLLVGVAFAARSASAITDGLLPNGNFEEAPDESQMNGTRV

TGRYAIPQWEISGFVEYIGSGQKQGDMLLPVPEGAYAVRLGNEASIQQRLT**LTRGM**
motif

2

HYSIIFSAARTCAQSELLNITITPESGEIPIQTVYTSSGWDSYSWAFKAKHSVVL
motif 2/motif 6 motif 1

IVHNPGVSDDEACGPLIDSFAIKTLNPPQRTKG**NMLKNGGFEEGPYIFPNTSWGVL**
motif 3 **motif 4**

VPPMEDDDYSPLSPWTILSSTTKSVKYIDAAHYAVPGGARAVELVSGMETAMLQEV
motif 4/motif 1

TVPGRSYRLEFSVG DAGDGC SGSLTVQAYASRGSVKV TYQS QGTGGYKRGLLEFTA

TEKRTRVVFVSMAYTTKSDGTL CGPVIDDASLVSVHSHRRFLL

FIGURE 1

MVLVVPEGSHAVRLGNEASIRQRLAGAARGARYALT**FSAARTCAQAERL**NVSASGQ
motif 2 / motif 6 motif 5

WAVLPMQTMYS SNGWDSYAWAWDAAADAFDVVIHNPGVTEDPACGPLIDSVAIRTL
motif 5 motif 3

NPPRRTNK**NLVKNGDFEEGPYIIPGTRWGV**LIPSMV**DEHSPLPGWMV**ESLKAVKY
motif 4 motif 1

IDSDHFAVPRGRRAVELLAGRESAIAQVIRTVPGRQYALSFTVGDASNGCEGSLVV

EAYAGRESTRVAHESAGRGGA AKRAVL PFRAAAARTRVVFSSFYSTRSDDMSSLC

GPVIDDVAVVSVRARRPAAKRG

FIGURE 2

A.lyrata_340195	(1)	1	50
A.thaliana_AT1G80240.1	(1)	---	---
B.napus_TC65362	(1)	---	---
B.napus_TC77368	(1)	---	---
O.sativa_IOC_Os01g42520.1	(1)	---	---
P.virgatum_TC46275	(1)	---	---
S.bicolor_Sb03g027650.1	(1)	---	---
Z.mays_TC458850	(1)	---	---
Zea_mays_GRMZM2G034985_T02	(1)	---	---
T.aestivum_c60008104l1410	(1)	---	---
O.sativa_IOC_Os04g41710.1	(1)	---	---
O.sativa_IOC_Os04g41740.1	(1)	---	---
O.sativa_IOC_Os_DUF642.3	(1)	---	---
O.sativa_IOC_Os04g41750.1	(1)	---	---
S.bicolor_Sb06g021280.1	(1)	---	---
Z.mays_ZM07MC22756_BF50088A01022693	(1)	---	---
Zea_mays_GRMZM2G051898_T01	(1)	---	---
S.bicolor_Sb06g021290.1	(1)	---	---
O.sativa_IOC_Os04g41759.1	(1)	---	---
S.bicolor_Sb06g021320.1	(1)	---	---
S.bicolor_Sb06g021300.1	(1)	---	---
Z.mays_TC468877	(1)	---	---
S.bicolor_Sb06g021310.1	(1)	---	---
O.sativa_IOC_Os04g41770.1	(1)	MAQPQCPRPQARTPNPLCAATPWPRAGRANWRGIFCSRWSPSCRDART	---
S.bicolor_Sb06g021270.1	(1)	---	---
Zea_mays_GRMZM2G051912_T01	(1)	---	---
Zea_mays_GRMZM2G405071_T01	(1)	---	---
O.sativa_IOC_Os02g11040.1	(1)	---	---
O.sativa_IOC_Os_DUF642.2	(1)	---	---
P.virgatum_TC392	(1)	---	---

FIGURE 3

S.bicolor Sb04g007160.1	(1)	----
Z_mays TC467870	(1)	----
P.sitchensis TA14284 3332	(1)	----
P.taeda TA10429 3352	(1)	----
A.lyrata 9I2808	(1)	----
A.thaliana AT1G29980.1	(1)	----
A.lyrata 934049	(1)	----
C.annuum TC19321	(1)	----
S.tuberosum TC189725	(1)	----
J.hindsii_x_regia TA1156 432290	(1)	----
G.hirsutum TC135941	(1)	----
G.raimondii TC1124	(1)	----
G.max_GM06MC25068_saa05g06@24507	(1)	----
G.max TC286277	(1)	----
L.japonicus TC41787	(1)	----
M.domestica TC32901	(1)	----
P.trifoliata TA6477 37690	(1)	----
P.trichocarpa DUF	(1)	----
P.trichocarpa MUB2	(1)	----
T.officinale TA3338 50225	(1)	----
H.vulgare TC163732	(1)	----
T.aestivum TC292522	(1)	----
O.sativa LOC_Os01g55190.1	(1)	----
P.virgatum TC7090	(1)	----
S.bicolor Sb03g034960.1	(1)	----
S.officinale TC76182	(1)	----
Z_mays c62082358gm030403@7546	(1)	----
Z_mays_ZM07MC26525_Bfb0211C02@26447	(1)	----
Zea_mays GRMZM2G324705 T01	(1)	----
Z.officinale TA720 94328	(1)	----
P.sitchensis TA10533 3332	(1)	----
P.taeda TA4773 3352	(1)	----
O.sativa LOC_Os03g59300.1	(1)	----
S.bicolor Sb01g004270.1	(1)	----

FIGURE 3 (continued)

4/77

Z.mays_c57213778gm030403@389	(1)
Z.mays_TC506958	(1)
P.trichocarpa_735032	(1)
V.vinifera_GSVIVT00025589001	(1)
S.lycopersicum_TC198542	(1)
Triphysaria_sp_TC11756	(1)
G.max_Glyma09g36220.1	(1)
M.truncatula_AC174305_19.4	(1)
G.max_Glyma15g43I80.1	(1)
P.vulgaris_TC10932	(1)
A.lyrata_478145	(1)
A.thaliana_AT3G08030.1	(1)
B.napus_TC89135	(1)
G.hirsutum_TC161608	(1)
T.cacao_TC4474	(1)
L.serriola_TC5542	(1)
C.sinensis_TC3550	(1)
P.trichocarpa_549292	(1)
P.trichocarpa_833200	(1)
V.vinifera_GSVIVT00028751001	(1)
F.vesca_TA10399_57918	(1)
N.benthamiana_TC11106	(1)
N.tabacum_TC43533	(1)
S.lycopersicum_TC199778	(1)
Triphysaria_sp_TC1599	(1)
I.nil_TC4208	(1)
Consensus	(1)

FIGURE 3 (continued)

5/77

	51	100
A.lyrata 340195	(1)	-----MYHYQEAALIVALL
A.thaliana AT1G80240.1	(1)	-----MMYQEAALIVALL
B.napus TC65362	(1)	-----MYQEAALIVALL
B.napus TC77368	(1)	-----MYQEVVLLTLL
O.sativa LOC Os01g42520.1	(1)	-----MRRVTIVILLAC
P.virgatum TC46275	(1)	-----MRRVALLLV
S.bicolor Sb03g027650.1	(1)	-----MRRVALLLLV
Z.mays TC458850	(1)	-----MRRVALLLVAVC
Zea mays GRMZM2G034985 T02	(1)	-----MRLTISARPVSLSPPTQQNQQTATSGAALP
T.aestivum c60008104@11410	(1)	-----MRCVLLLV
O.sativa LOC Os04g41710.1	(1)	-----MAFYNGASVVALLLV
O.sativa LOC Os04g41740.1	(1)	-----MRRKTEMIQTTPCVALLLV
O.sativa LOC Os04g41750.1	(1)	-----MRRKTEMIQTTPCVALLLV
S.bicolor Sb06g021280.1	(1)	-----MRRKTEMVSTCCVLLLV
Z.mays GRMZM2G051898 T01	(1)	-----MGRGSETVPNARRAALFLIV
S.bicolor Sb06g021290.1	(1)	-----MGRGSDAVPDARRAALLLV
O.sativa LOC Os04g41759.1	(1)	-----MGRGSDAVPDARRAALLLV
S.bicolor Sb06g021320.1	(1)	-----MVQNAPHAALFLV
S.bicolor Sb06g021300.1	(1)	-----MTGSARSVALLFL
Z.mays TC468877	(1)	-----MVARRSVAFVIL
S.bicolor Sb06g021310.1	(1)	-----MAESGSSRRGVLLFLL
O.sativa LOC Os04g41770.1	(1)	-----MEGRASSRRGVLLFLL
S.bicolor Sb06g021270.1	(1)	-----MAGRTRSRCPVLLFLL
Zea mays GRMZM2G051912 T01	(51)	IQPARLAASRTHPPCTDHTSCNKQQLLYISSEVRVLDRTVTRAHDHGLGEV
Zea mays GRMZM2G405071 T01	(1)	-----MAIGKRSVHG--AALVLL
O.sativa LOC Os02g11040.1	(1)	-----MAIDERSVHGGAALVLRLL
P.virgatum TC392	(1)	-----MVIDRRSVQR-AALVVLPL
S.bicolor Sb04g007160.1	(1)	-----MAAGSASLILLMVVL
	(1)	-----MGRSSMAGPAMLL
	(1)	-----MAVPVMLL

FIGURE 3 (continued)

6/77

Z.mays TC467870	(1)	-----MGRALALAVPVMLL
P.sitchensis TA14284 3332	(1)	-----MWSSGVSI CRFLCCL
P.taeda TA10429 3352	(1)	-----MWCSRVSICRFLCCL
A.lyrata 9I2808	(1)	-----
A.thaliana AT1G29980.1	(1)	-----MVPSNNRCKWSSIFLFLLSVSAVL
A.lyrata 934049	(1)	-----MVLNYSTWRILVSILILLGLSIVA
C.annuum TC19321	(1)	-----MEIHWSSTRCKWVT---TFL
S.tuberosum TC189725	(1)	-----MEIASSTKFKWVT---AFL
J.hindsii x_regia TA1156 432290	(1)	-----MAQSSSKSWVS--LLML
G.hirsutum TC135941	(1)	-----MAESSAINRCIF--LFML
G.raimondii TC1124	(1)	-----MDAAESWSRTKWVF--HFML
G.max GM06MC25068 saa05g06024507	(1)	-----MARSTRMKWVS-IFTLL
G.max TC286277	(1)	-----MARSTRMKWVS-IFTLL
L.japonicus TC41787	(1)	-----MARSTRMKWVSMFVPL
M.domestica TC32901	(1)	-----MGRSSTLHKWVL--PTLL
P.trifoliata TA6477 37690	(1)	-----MVKSSSTTRWVS--LFMF
P.trichocarpa DUF	(1)	-----MFQSSSTRSKWFT--LFML
P.trichocarpa MUB2	(1)	-----MFQSSSTRSKWFT--LFML
T.officinale TA3338 50225	(1)	-----MASHSATTKSLA--LFCL
H.vulgare TC163732	(1)	-----MERGLALLGLGLALLLLA
T.aestivum TC292522	(1)	-----MERGLALLGLGLALLLLA
O.sativa IOC Os01g55190.1	(1)	-----MEGDSARRIALGLCIGLILLLA
P.virgatum TC7090	(1)	-----MERGSPRRRALLALCSAIIIIA
S.bicolor Sb03g034960.1	(1)	-----MEPDSARKRLALLAIIIIIIA
S.officinale TC76182	(1)	-----MEPDSARKRLALLA--LIIIIA
Z.mays c62082358gm03040307546	(1)	-----MEADSARKRLALLA--LIIIIA
Z.mays_ZM07MC26525 Bfb0211C02026447	(1)	-----MEPDSATTRIALIA--LIIIIA
Zea mays GRMZM2G324705 T01	(1)	-----
Z.officinale TA720 94328	(1)	-----MNNPLARLPHLIVLL
P.sitchensis TA10533 3332	(1)	-----MAKSELPFSSVFRLLQVLI
P.taeda TA4773 3352	(1)	-----MAKTSEL SLLFFRFL ETLV
O.sativa IOC Os03g59300.1	(1)	-----MELPRGALLIVLLAVAAAAS
S.bicolor Sb01g004270.1	(1)	-----MELPWGMLL--LLVLVS VSSS

FIGURE 3 (continued)

7/77

Z.mays_c57213778gm0304030389	(1)	-----MLLTLLLVLSVSS
Z.mays_TC506958	(1)	-----MELPWGMLLPLLLVLSVSSP
P.trichocarpa_735032	(1)	-----MAFLSICVALLFS
V.vinifera_GSVIVT00025589001	(1)	-----MAFLSLCLVFFSI
S.lycopersicum_TC198542	(1)	-----MASFFSLSFILFS
Triphysaria_sp_TC11756	(1)	-----MARFF--LFILLIS
G.max_Glyma09g36220.1	(1)	-----MERGALLATCGLLLQHKTEWLVVRMSETIFTLSLVM
G.max_Glyma15g43I80.1	(1)	-----MAVTICTLFMTLL
P.vulgaris_TC10932	(1)	-----MAVTLTSTLFMALL
A.lyrata_478145	(1)	-----MAPTLSLILLALL
A.thaliana_AT3G08030.1	(1)	-----MKKIQSLS
B.napus_TC89135	(1)	-----MAVPKAIILPILL
G.hirsutum_TC161608	(1)	-----MAVSKAIVIPLLLV
T.cacao_TC4474	(1)	-----MTISASFLLASTLL
L.serriola_TC5542	(1)	-----MTMATSFLIVCTLL
C.sinensis_TC3550	(1)	-----MSKLISMLLVLFV
P.trichocarpa_549292	(1)	-----MV'AHSLSIVLALGL-LL
P.trichocarpa_833200	(1)	-----MAVGISLILLAS-FL
V.vinifera_GSVIVT00028751001	(1)	-----MAVGFSLMIAL-FL
F.vesca_TA10399_57918	(1)	-----MAVLISLILLALSSV
N.benthamiana_TC11106	(1)	-----MALLVSFTLALS
N.tabacum_TC43533	(1)	-----MVVKSIFLLFLFL
S.lycopersicum_TC199778	(1)	-----MVVKSILLILLSL
Triphysaria_sp_TC1599	(1)	-----MVFKVSVIILFSFL
I.nil_TC4208	(1)	-----MSLKPTLLLVLCFL
Consensus	(51)	-----MAAKICFALFFCLV
		L LLL LL

FIGURE 3 (continued)

9/77

Z.mays TC467870	(14)	LLCTTCRLTLGIT-----DGLLPNGNFER
P.sitchensis TA14284 3332	(16)	QFLQLWVSVTAS-----DGLLPNGDFEL
P.taeda TA10429 3352	(16)	QLLQLWGSVTAS-----DGLLPNGDFEL
A.lyrata 9I2808	(1)	-----MG-----RLVINGDFET
A.thaliana AT1G29980.1	(26)	VAVADKSPAVED-----GLVINGDFET
A.lyrata 934049	(26)	AADSAGKTLFVED-----GLVINGDFET
C.annuum TC19321	(18)	LLACSAAGQVED-----GPLVNGNFEF
S.tuberosum TC189725	(18)	LLITCSAAGQVQD-----GSLVNGNFEF
J.hindsii_x_regia TA1156 432290	(17)	LFARWATAIVIED-----GPLPNGDFET
G.hirsutum TC135941	(17)	IFPHFSTLALAE-----GLVANGDFEA
G.raimondii TC1124	(19)	IFAHFSLLSLAED-----GLVANGDFET
G.max GM06MC25068_saa05g06@24507	(18)	FLSHSPLTTFAED-----GLVANGDFEV
G.max TC286277	(18)	FLSHSPLTTFAED-----GLVANGDFEV
L.japonicus TC41787	(19)	ILLHLVLAPEED-----GLVPNGDFEA
M.domestica TC32901	(17)	LLGHLAAEISADD-----GPVPNGDFEL
P.trifoliata TA6477 37690	(17)	VVAHLASSILAED-----GLVANGDFET
P.trichocarpa DUF	(17)	VFALFASAILGED-----GLVTNGDFET
P.trichocarpa MUB2	(17)	VFVLFASAILGED-----GLVTNGDFET
T.officinale TA3338 50225	(17)	LFSHFVAVFTAED-----GLLINGDFET
H.vulgare TC163732	(19)	PSRRLASAAPVED-----GLLINGDFET
T.aestivum TC292522	(19)	PSRRLASAAPVED-----GLLMNGDFET
O.sativa LOC Os01g55190.1	(23)	PS-HLASAAVED-----GLLSNGDFET
P.virgatum TC7090	(24)	APARPVWAAPVED-----GLLSNGDFET
S.bicolor Sb03g034960.1	(21)	APARPASAAPVED-----GLLSNGDFET
S.officinatum TC76182	(20)	APACPASAAPVED-----GLLSNGDFET
Z.mays c62082358gm03040307546	(20)	APVRPASAAPVED-----GLISNGDFET
Z.mays_ZM07MC26525_Bfb0211C02026447	(20)	APARPASAAPVED-----GLLSNGDFET
Zea mays GRMZM2G324705 T01	(1)	--MISDLLASCTT-----GLLSNGDFET
Z.officinale TA720 94328	(16)	LLLTHCVSAAVED-----GLINNGDFET
P.sitchensis TA10533 3332	(21)	CLSLSTFVTAQDGVTAED-----GLLVNGNFEK
P.taeda TA4773 3352	(21)	CLSVLSTFVAED-----GLLVNGNFEK
O.sativa LOC Os03g59300.1	(21)	---SAGATLAVKAPVPVPAPAP-----AHAPPQPKDAEGLLINGNFEF
S.bicolor Sb01g004270.1	(19)	SSSSAAATLAVSAPPPVPAP-----RHAQDAEGLLINGNFEF

FIGURE 3 (continued)

Z.mays_c57213778gm030403@389	(15)	TSTSAATLAVSAPPPPLPVPP-----PRHAQDAEGMLINGNFET
Z.mays_TC506958	(21)	SSSSAAATLAVTAPPPLPVPP-----PRHAQDAEGMLINGNFET
P.trichocarpa_735032	(15)	FCHGLANAIPPP-----LDGLVENGDFEE
V.vinifera_GSVIT00025589001	(15)	VCS--SASATLPY-----LDGLLPNGDFEK
S.lycopersicum_TC198542	(15)	LYN-SFAYANTVPY-----LDDLLENGNFEQ
Triphysaria_sp_TC11756	(13)	CNN-LASAIIVPPP-----LDGLLENGNFER
G.max_Glyma09g36220.1	(38)	CAASAFAAVPHR-----LPEVYLKNGNFEE
M.truncatula_AC174305_19.4	(15)	FAASAFKVVHR-----SPEVFFRNGNFEE
G.max_Glyma15g43180.1	(15)	FAASAFALQPR-----VPEAYLQNGNFEE
P.vulgaris_TC10932	(15)	CAASAFALPR-----VPEAYLQNGNFEE
A.lyrata_478145	(9)	LDLRFIS-----LFTSSGYLRNGNFEE
A.thaliana_AT3G08030.1	(15)	ICGAALG-----APASEGYLRNGNFEE
B.napus_TC89135	(15)	LCAALG-----APAYEGFLRNGNFEE
G.hirsutum_TC161608	(15)	LTGCVLGPISLSAKHGVPY-----AHKKIASKPFEGYLENGNFEE
T.cacao_TC4474	(15)	VTGCALGPKPFVSAKRGVRPY-----GSAKIAAKPFEGYLENGNFEE
L.serriola_TC5542	(15)	STSVLAARPP-----ATPMEGLLPNGNFEE
C.sinensis_TC3550	(17)	LLSGPAFAKY-----FEGYLPNGDFEE
P.trichocarpa_549292	(14)	LFTTSALAAATY-----MEGFLKNGNFEE
P.trichocarpa_833200	(14)	LFIGSASAATY-----MEGFLKNGNFEE
V.vinifera_GSVIT00028751001	(15)	FFTGPAPAGTY-----LEGLLPNGNFEE
F.vesca_TA10399_57918	(15)	ISTSPAFAPK-----PFEGLLANGNFEE
N.benthamiana_TC11106	(15)	FTVSAVRPAP-----LEGLLSNGNFEE
N.tabacum_TC43533	(15)	STVSAVRPAP-----LEGLLPNGNFEE
S.lycopersicum_TC199778	(15)	FTASAVRPAP-----LEGLLPNGNFEE
Triphysaria_sp_TC1599	(15)	YTAALVPPLP-----FEGLLPNGNFEE
I.nil_TC4208	(15)	LAVSALNAP-----LEGFLPNGNFEE
Consensus	(101)	A A DGLL NGNFE

FIGURE 3 (continued)

151		DUF 642	DOMAIN	200
(39)	GPKP-SQ_LKGSVVK-ERTAVPNWDITGFVEYIKSGQK-QDDMVLVVPQGS			
(38)	GPKP-SQ_MKGSVVK-ERTAVPNWNIIGFVEFIKSGQK-QDDMVLVVPQGS			
(37)	GPKP-SQ_LKGSVVK-ERNAVAHWDMTGFVEYIKAGQK-QDDMVLVVPQGS			
(37)	GPKP-SQ_LKGSVIK-GRTAVPNWEITGFVEYIKAGQK-QDDMVLVVPQGS			
(36)	GPKSDLVNGTVVR-GANAI PRWETS GFVEYIESGHK-QGDMVLVVPQGA			
(33)	GPKSAIVNGTVVK-GANAI PRWETS GFVEYIESGHK-QGDMVLVVPQGA			
(34)	GPKSAIVNGTVVS-GANAI PRWETS GFVEYIESGHK-QGDMVLVVPQGA			
(36)	GPKSAIVNGTVVS-GAHAI PRWETS GFVEYIESGHK-QGDMVLVVPQGA			
(83)	GPKSAIVNGTVVS-GAHAI PRWETS GFVEYIESGHK-QGDMVLVVPQGA			
(33)	GPAKSEIVNGTVVK-GGKAI PKWETS GFVEYIESGHK-QGDMVLVVPQGA			
(36)	QPSK-SQ_MNGTRVM-AEYAI PYWKITGFVEYISSGQK-QGDMVLVVPQGA			
(45)	APDK-SQ_MNGTRVT-GRYAI PQWEISGFVEYIGSGQK-QGDMVLVVPQGA			
(45)	APDE-SQ_MNGTRVT-GRYAI PQWEISGFVEYIGSGQK-QGDMVLVVPQGA			
(43)	APDK-SQ_MNGTRVI-GRYAI PQWEISGFVEYIGSGQM-QGDMVLVVPQGA			
(43)	APPK-SQ_LKGSRVK-GRNAI PHWEISGFVEYIGSGQK-QGDMVLVVPQGA			
(43)	APPK-SQ_MNGTRVM-GRNAI PHWEISGFVEYIGSGQK-QGDMVLVVPQGA			
(43)	APPK-SQ_MNGTRVM-GRNAI PHWEISGFVEYIGSGQK-QGDMVLVVPQGA			
(37)	APAK-SELNGTRVK-GRDAI PHWEISGFVEYIGPGQK-QGDMVLVVPQGA			
(39)	GPDK-SQ_MNGTVVT-GRYAI LNWEISGFVEYIESGHR-EQDMILAVPEGA			
(37)	GPDK-SEMNGTRVL-GRYSI PHWEISGFVEYIESGQK-QDDMLIQVPEGE			
(41)	GPDK-SQ_MNGSRVI-GQNSI PCWETS GFVEYIEPGQK-QNDMVLVVPQGA			
(36)	GPDK-YQ_MNGTWVI-GRKSI PCWETS GFVEYIVSGQQQDNGMVAVPEGA			
(41)	GPDK-SQ_MNGTRVM-QQNSI PNWEISGFVEYIGSGQK-QDDMVLVVPQGA			
(125)	WPST-KRQLRARAG-QVPAEPNWEISGFVEYIGSGHK-EQDMILAVPEGA			
(42)	SPNP-SQ_MSGSVVT-GEHAISYWRKTSGHIEFICSGQT-QGDMVLVVPDGA			
(44)	PPNQ-SQ_MSGSVVT-GEHAVSYWKTSGHIEFICSGQK-QGDMVLVVPDGS			
(43)	SPDQ-SQ_MSGSVVA-GEHAISGWKTSGHIELVCSGQK-QGDMVLVVPDGA			
(40)	GPAP-SQ_LRGTRVV-GASAI PSWRKTS GFVEYIPSGRK-QGDMVLVVPQGS			
(1)	-----MVLVVPQGS			
(38)	GPVP-SQ_LRGTRVL-GSSAI PSWQTS GFVEYIPSGKK-QGDMVLVVPQGA			
(33)	GPAP-SQ_LRGTRVM-GSSAI PAWQTS GFVEYIPSGRK-QGDMVLVVPQGA			

FIGURE 3 (continued)

Z.mays TC467870	(38)	SPLP-QQLRGTRVV-GASAIPSWQTSQFVEYISSGQK-QQDMVLVVPEGA
P.sitchensis TA14284 3332	(40)	APKP-KDLKGTVLL-GRNSLPQWRTKGFVEYITSGHH-QRDMVLVVPEGA
P.taeda TA10429 3352	(40)	APKP-KNLKGTVLV-GRNSLPKWRTRGFVEYITSGHR-QRDMVLVVPEGA
A.lyrata 9I2808	(13)	SPS-HQFPD-DGVTDPSPDIPSWKSNGTVELINSQ-QQGMILIVPQGR
A.thaliana AT1G29980.1	(49)	SPS-SQFPD-DGVTDPSPDIPSWKSNGTVELINSQ-QQGMILIVPQGR
A.lyrata 934049	(49)	PPS-NQFPD-DAIIESSEIPEWSRSDGTVELIKSQ-KQGMILIVPEGR
C.annuum TC19321	(41)	PPS-GQFSSGDGFSQTEIPEWSKSNGTVEVESGQ-KQGMILIVPQGR
S.tuberosum TC189725	(41)	PPS-GQFSSGDGFSQTEIPEWSKSNGTVEVESGQ-KQGMILIVPQGR
J.hindsii x regia TA1156 432290	(40)	PPS-NQFPN-GAIAEGPTLVPSWRSNGTVELVEAGQ-KQGMILIVPQGR
G.hirsutum TC135941	(40)	RPS-NQFPS-EAIAADGPTTEIPSWRTKGRVELVSSGEKVSQGMILIVPGGS
G.raimondii TC1124	(42)	PPS-GQFPS-EAISDGAPEIPSWSTNGTVKLVSAGEKVSQGMILIVPRGR
G.max GM06MC25068_saa05g06@24507	(41)	SPS-SQFPNEAIVE-GPSEVPNWSNGNVELVESGQ-KQGMILIVPQGR
G.max TC286277	(41)	SPS-SQFPNEAIVE-GPSEVPNWSNGNVELVESGQ-KQGMILIVPQGR
L.japonicus TC41787	(42)	SPS-NQFPSEATIEGPSEVPNWSNGTVELVESGQ-KQGMILIVPQGR
M.domestica TC32901	(40)	TPS-GQFSN-DGAVDGPAGIPSWKANGTVELVESGQ-KQGMILIVPQGR
P.trifoliata TA6477 37690	(40)	SPS-HQFPS-ESRADGPTTEIPSWKNGTVELVSSGQ-KQGMILIVPQGS
P.trichocarpa DUF	(40)	PPI-TQFPKDEALVEGPTTEIPGWKTNGTVELVSSGQ-KQGMILIVPRGA
P.trichocarpa MUB2	(40)	PPI-TQFPKDEALVEGPTTEIPGWKTNGTVELVSSGQ-KQGMILIVPRGA
T.officinale TA3338 50225	(40)	PPS-GQFAD-EGLDNPNNAIPSWKVNGTVELVASGQ-KQGMILIVPEGR
H.vulgare TC163732	(42)	APP-GQFVKSSSVSEGA-IPSWTINGTVELISAGQ-HQGMILIVPQGD
T.aestivum TC292522	(42)	APP-GQFVKASVSEGA-IPSWTINGTVELISAGQ-HQGMILIVPQGD
O.sativa LOC Os01g55190.1	(45)	APA-GQFVKASVSEGA-IPSWTINGTVELISAGQ-HQGMILIVPQGD
P.virgatum TC7090	(47)	APA-GQFVKASVSEGA-IPSWTINGTVELISAGQ-HQGMILIVPQGD
S.bicolor Sb03g034960.1	(44)	APA-GQFVKASVSEGA-IPSWTINGTVELISAGQ-HQGMILIVPQGD
S.officinale TC76182	(43)	APV-GQFVKASVSEGA-IPSWTINGTVELISAGQ-HQGMILIVPQGD
Z.mays c62082358gm030403@7546	(43)	APT-GQFVKASVSEGA-IPSWTINGTVELISAGQ-HQGMILIVPQGD
Z.mays_ZM07MC26525 Bfb0211C02@26447	(43)	APA-GQFVRSASVSEGA-IPSWTINGTVELISAGQ-HQGMILIVPQGD
Zea mays GRMZM2G324705 T01	(22)	APS-GQFVKASVSEGA-IPSWTINGTVELISAGQ-HQGMILIVPQGD
Z.officinale TA720 94328	(39)	TPAPS-GTAGNDLGEITSLPGWTINGTVELVESGQ-KQGMILIVPEGS
P.sitchensis TA10533 3332	(50)	EGIENKTSR---AGDIDSLPGWNIDGMVEYINEGQ-KQGSMLVVPQEK
P.taeda TA4773 3352	(44)	AGIENKTSR---AGVMDSLPGWTIEGTVEYVNDGE-RQGSMLVVPQEK
O.sativa LOC Os03g59300.1	(63)	APRK---VNKTLIV-GRHSLPGWTLRGHVEYVSGGPQ-PGGMFFAVPHGV
S.bicolor Sb01g004270.1	(56)	APRK---LNKTLIV-GRHSLPGWTVQGHVEYVSGGPQ-PGGMFFAVPHGV

FIGURE 3 (continued)

Z.mays_c57213778gm030403@389	(54)	APRR-T-LNKTLIV-GRHSLPGWTVQGHVEYVSGGPQ-PGGMFFAVPHGV
Z.mays_TC506958	(60)	APRR-T-LNKTLIV-GRHSLPGWTVQGHVEYVSGGPQ-PGGMFFAVPHGV
P.trichocarpa_735032	(41)	APAK-SNLKKTVII-GKHSLPKWEINGLVEYVSGGPQ-PGGFFLAIPRGV
V.vinifera_GSVIVT00025589001	(39)	GPKA-SNLKKTVIK-GKHSLPDWEIKGLVEYVSGGPQ-PGGFFLAIPRGV
S.lycopersicum_TC198542	(40)	GPKP-SNMNKTVII-GKHSLPKWEINSIVEVSGGPQ-EGGFYFPIPRGA
Triphysaria_sp_TC11756	(38)	GNP-RYLKKTVII-GKHSLPGWEISGLVEYVSGGPQ-PGGFYFVAPRGV
G.max_Glyma09g36220.1	(63)	NPNP-RYLKKTTLI-GKYALPKWEISGLVEYVSGGPQ-PGGMYFPVSHGV
M.truncatula_AC174305_19.4	(40)	QPNP-GYIKQTRLM-GKHALPNWETNGLVEYITGGPQ-PGGMFFPVS HGV
G.max_Glyma15g43180.1	(40)	QPNP-KNLQKTKLM-GKYSLPKWEVNGLVEYVSGGPQ-PGGMFFPVT HGI
P.vulgaris_TC10932	(40)	QPNP-RYLKKTTLV-GKFALPKWEISGLIEYVSGGPQ-PGGMFFPVT HGI
A.lyrata_478145	(31)	SPKK-PLDMKKTVII-GKTALPEWEITGFVEYIAGGPQ-PGGMFFPVA HGV
A.thaliana_AT3G08030.1	(37)	SPKK-PLDMKKTVII-GKNALPEWEITGFVEYIAGGPQ-PGGMYFPVA HGV
B.napus_TC89135	(37)	SPKK-PLDMKKTVII-GKTALPEWVITGFVEYIAGGPQ-PGGMYFPVA HGV
G.hirsutum_TC161608	(57)	QSPS-PLALKKTVLK-GAHALPKWTITGLVEYISGGPQ-PGGMYFPVA HGV
T.cacao_TC4474	(57)	QPKP-PLDLKKTVII-GKYALPKWTISGLVEYITGGPQ-PGGMFFPVA HGV
L.serriola_TC5542	(40)	PPKA-PLDIKKTVII-GKTALPKWEISGLVEYIHGGPQ-PGGMYFPVA HGV
C.sinensis_TC3550	(40)	LPKP-PLNLKKTVII-GKHALPKWEINGFVEYIAGGPQ-PGGMFFPVS HGV
P.trichocarpa_549292	(37)	KPKP-GAIKKTIVLK-GKYALPSWEINGFVEYISAGPQ-PGGMYFNVA HGV
P.trichocarpa_833200	(37)	KPKP-GAIRKTVLK-GKNALPSWEINGFVEYISAGPQ-PGGMYFNVA HGV
V.vinifera_GSVIVT00028751001	(38)	SPKP-PLDLKKTVIK-GKYSLPKWEINGFVEYIAGGPQ-PGGMFFAVA HGV
F.vesca_TA10399_57918	(40)	PPKP-SLKKTVII-GKNALPKWEINGFVEYISGGPQ-PGGMYFAVA HGV
N.benthamiana_TC11106	(38)	QPKP-PLKKTIVLV-GKYALPKWEINGLVEYISGGPQ-PGGMYFPVA HGV
N.tabacum_TC43533	(37)	QPKP-PLKKTIVLV-GKYALPKWEINGLVEYISGGPQ-PGGMYFPVA HGI
S.lycopersicum_TC199778	(38)	LPKADLKKKTIVLQ-GKYALPKWEINGLVEYISGGPQ-PGGMYFAVA HGV
Triphysaria_sp_TC1599	(38)	KPKP-TSLHKTKLL-GKNALPKWQIKGLVEYISGGPQ-PGGMFFAVA HGI
I.nil_TC4208	(36)	GPKA-SDIKKTIVLQ-GKHALPRWETSGLVEYISGGPQ-PGGMFFPVA HGV
Consensus	(151)	AP I LK T V G AIP W I G VEYI SGQ QGGMILIVP G

FIGURE 3 (continued)

	201				250
		(86)	SAVRLGNEASISQKIS-VLPGRLYSITFSAARTCAQDEK--LNISVTHES		
		(85)	SAVRLGNEASISQKIS-VLPGRLYSITFSAARTCAQDER--LNISVTHES		
		(84)	SAVRLGNEASISQKIS-VRSGRLYSITFSAARTCAQDER--LNISVTHES		
		(84)	SAVRLGNEASISQKIS-VRSGRLYSITFSAARTCAQDER--LNISVTHES		
		(84)	HAVRLGNEASIRQRLA-VTRGAYAVTFSAARTCAQAEQ--LNVSVSPEW		
		(81)	HAVRLGNEASIRQRLA-VTRGAYAITFSAARTCAQAEK--LNISVSPFW		
		(82)	HAVRLGNEASIRQRLA-VARGAYAITFSAARTCAQAEK--LNVSVSPEW		
		(84)	HAVRLGNEASIRQRLA-VARGAYAITFSAARTCAQAEK--LNVSVSPEW		
		(131)	HAVRLGNEASIRQRLA-VARGAYAITFSAARTCAQAEK--LNVSVSPEW		
		(81)	YAVRLGNDASILQRIIP-VARGSYAITFSAARTCAQAEK--LNVSVSPES		
		(83)	HAVRLGNEASIEQQIS-VTRGMYISITFSAARTCAQSEK--LNVSVAPGP		
		(92)	YAVRLGNEASIQRLT-LTRGMHYSITFSAARTCAQSEL--LNITITP--		
		(92)	YAVRLGNEASIQRLT-LTRGMHYSITFSAARTCAQSEL--LNITITP--		
		(90)	YAVRLGNEASIQRLT-LTRGMHYSITFSAARTCAQSEL--LNITITP--		
		(90)	YAVRLGNEASIQRLA-LTQGAHYSITFSAARTCAQAEQ--LNVTVAP--		
		(90)	YAVRLGNEASIQRLA-VTPGARYSITFSAARTCAQAEQ--LNVTVAP--		
		(90)	YAVRLGNEASIQRLA-VTPGARYSITFSAARTCAQAEQ--LNVTVAP--		
		(84)	YAVRLGNEASIQRLA-VTPGARYSVTFSAARTCAQAEQ--LNVTVAT--		
		(86)	YAVRLGNDATIRQRLS-VTRRAYSITFSAARTCAQSEK--LNVSVTP--		
		(84)	YAVRLGNDATIRQRLA-VTRHTYYSITFSAARTCAQAEK--LNVSVTP--		
		(88)	YAVRLGNDATIRQRLPGLTRMSYSITFSAARTCAQAEQ--LNVSVAA--		
		(84)	YAVRLGNDASIRQRLTGMTRGTYSVTFSAARTCAQAEQ--LNVSVAP--		
		(88)	YAVRLGNDATIRQKLS-VNRKTTYITFSAARTCAQAEQ--LNVSVAA--		
		(172)	YAVRLGNDATIRQRLS-VTRHMYSVTFSAARTCAQAEK--LNVSVTL--		
		(89)	HALRLNGASIEQRIS-LTPGSYSLTFSAARTCAQDEV--LNVTAAPVS		
		(91)	HALRLSGASIEQQIS-LTKGLYSLTFSAARTCSQNEK--LNVTAAPVS		
		(90)	QALRLGASIEQQISGLTQGSYSLTFSAARTCSQDET--LSVTAAPAP		
		(87)	HAVRLGNEASIRQRLAGAARGARYALTFAARTCAQAEK--LNVSVASGQ--		
		(10)	HAVRLGNEASIRQRLAGAARGARYALTFAARTCAQAEK--LNVSVASGQ--		
		(85)	YAVRLGNEASIRQRLAGAARGARYSLTFSAARTCAQAEQ--LNVSVASGQ--		
		(80)	YAVRLGNEASIRQRLHGAARGARYSLTFSAARTCAQAEQ--LNVSVASGQ--		
A.lyrata 340195					
A.thaliana AT1G80240.1					
B.napus TC65362					
B.napus TC77368					
O.sativa LOC_Os01g42520.1					
P.virgatum TC46275					
S.bicolor Sb03g027650.1					
Z.mays TC458850					
Zea_mays GRMZM2G034985.T02					
T.aestivum c600081041I410					
O.sativa LOC_Os04g41710.1					
O.sativa LOC_Os04g41740.1					
O.sativa LOC_Os04g41740.1					
O.sativa LOC_Os04g41750.1					
S.bicolor Sb06g021280.1					
Z.mays_ZM07MC22756 BF50088A01022693					
Zea_mays GRMZM2G051898.T01					
S.bicolor Sb06g021290.1					
O.sativa LOC_Os04g41759.1					
S.bicolor Sb06g021320.1					
S.bicolor Sb06g021300.1					
Z.mays TC468877					
S.bicolor Sb06g021310.1					
O.sativa LOC_Os04g41770.1					
S.bicolor Sb06g021270.1					
Zea_mays GRMZM2G051912.T01					
Zea_mays GRMZM2G05071.T01					
O.sativa LOC_Os02g11040.1					
O.sativa LOC_Os02g11040.1					
P.virgatum TC392					
S.bicolor Sb04g007160.1					

FIGURE 3 (continued)

Z.mays TC467870	(85)	YAVRLGNEASIRQRLRGAARGARYSLTFSAAARTCAQAEQ--LNVASGQ-
P.sitchensis TA14284 3332	(87)	HAVRLGNEASISQSIK-VTRGSYSLTFSAAARTCAQSER--LNVSVPP--
P.taeda TA10429 3352	(87)	HAVRLGNEASISQSIK-VTRGSYSLTFSAAARTCAQFER--LNVSVPP--
A.lyrata 9I2808	(60)	HAVRLGNDAEISQDLT-VEKGFVYSVTFSAAARTCAQLESINVSASVNAD
A.thaliana AT1G29980.1	(96)	HAVRLGNDAEISQDLT-VEKGFVYSVTFSAAARTCAQLESINVSASVNAD
A.lyrata 934049	(96)	HAVRLGNDAEISQELP-VEKGSYISVTFSAAARTCAQLESINVSASN---
C.annuum TC19321	(89)	HAVRLGNDAEISQELK-VEKGSYISITFSAAARTCAQLESINVSPPA---
S.tuberosum TC189725	(89)	HAVRLGNDAEISQELK-VEKGSYISITFSAAARTCAQLESINVSPPA---
J.hindsii x regia TA1156 432290	(87)	HAVRLGNDAEISQDLK-VEKGSYISVTFSAAARTCAQLESINVSPPA---
G.hirsutum TC135941	(88)	KAVRLGNDAEISQEVV-VEKGSYAVTFSAAARTCAQLESINVSPPA---
G.raimondii TC1124	(90)	HAVRLGNDAEISQEVV-VEKGSYAVTFSAAARTCAQLESINVSPPA---
G.max GM06MC25068_saa05g06@24507	(88)	HAVRLGNDAEISQELP-VEKGSYISLTFSAAARTCAQLESINSVAPA---
G.max TC286277	(88)	HAVRLGNDAEISQELP-VEKGSYISLTFSAAARTCAQLESINSVAPA---
L.japonicus TC41787	(90)	HAVRLGNDAEISQELT-VEKGSYISLTFSAAARTCAQLESINSVPPA---
M.domestica TC32901	(87)	HAVRLGNDAEISQVK-VEKGSYISVTFSAAARTCAQLESINSVPPA---
P.trifoliata TA6477 37690	(87)	HAVRLGNDAEISQVK-VEKGSYISVTFSAAARTCAQLESINSVPPA---
P.trichocarpa DUF	(88)	HAVRLGNDAEISQELT-VEKGSYISVTFSAAARTCAQLESINSVLPA---
P.trichocarpa MUB2	(88)	HAVRLGNDAEISQELT-VEKGSYISVTFSAAARTCAQLESINSVLPA---
T.officinale TA3338 50225	(87)	HAVRLGNDAEISQELK-LEKGEIYSITFSAAARTCAQLESINSVPPA---
H.vulgare TC163732	(89)	HAVRLGNDAEISQVQ-VEKGEYAITFSAAARTCAQLESINISAG-S---
T.aestivum TC292522	(89)	HAVRLGNDAEISQVQ-VEKGEYAITFSAAARTCAQLESINISAG-S---
O.sativa LOC Os01g55190.1	(93)	HAVRLGNDAEISQVQ-VEKGEYAITFSAAARTCAQLESINISAG-S---
P.virgatum TC7090	(95)	HAVRLGNDAEISQVQ-VEKGEYAITFSAAARTCAQLESINISAG-S---
S.bicolor Sb03g034960.1	(92)	HAVRLGNDAEISQVQ-VEKGEYAITFSAAARTCAQLESINISAG-S---
S.officinale TA720 94328	(91)	HAVRLGNDAEISQVQ-VEKGEYAITFSAAARTCAQLESINISAG-S---
Z.mays c62082358gm030403@7546	(91)	HAVRLGNDAEISQVQ-VEKGEYAITFSAAARTCAQLESINISAG-S---
Z.mays_ZM07MC26525 Bfb0211C02@26447	(91)	HAVRLGNDAEISQVQ-VEKGEYAITFSAAARTCAQLESINISAG-S---
Zea mays GRMZM2G324705 T01	(91)	HAVRLGNDAEISQVQ-VEKGEYAITFSAAARTCAQLESINISAG-S---
Z.officinale TA720 94328	(70)	HAVRLGNDAEISQVQ-VEKGEYAITFSAAARTCAQLESINISAG-S---
P.sitchensis TA10533 3332	(88)	RAIRLGNEAQISQGLQ-LEKGSYAVTFSAAARTCAQLESINISAG-S---
P.taeda TA4773 3352	(95)	HAVRLGNEAQISQIN-IQKGSYISLTFSAAARTCAQLESINISAG-S---
O.sativa LOC Os03g59300.1	(89)	HAVRLGNDAEISQIS-LQKGSYISLTFSAAARTCAQLESINISAG-S---
S.bicolor Sb01g004270.1	(108)	HALRGGASASQNVV-VKPGALYALTFAATRTCAQDEALRVAVAPSLSP
	(101)	HALRGGASASQNVV-VKPGALYALTFAATRTCAQDEALRVAVAPSLSA

FIGURE 3 (continued)

Z.mays_c57213778gm030403@389	(99)	HAIRLGSRASASQNV	VRPGALYALTFAATRTCAQDEALRVAVAPSLSA
Z.mays_TC506958	(105)	HAIRLGSRASASQNV	VRPGALYALTFAATRTCAQDEALRVAVAPSLSA
P.trichocarpa_735032	(88)	HAVRLGNEASISQILT	VKPGSVYALTTFGATRTCAQDEVLRVSIPGQS
V.vinifera_GSVIVT00025589001	(86)	HAVKLNEASISQNV	VRPGFIYSLTFGATRTCAQDEVLRVSIPGQS
S.lycopersicum_TC198542	(87)	HAVRLGNEASISQYVK	VKPNTIYSLTFGATRTCAQDEVLRVSAGMS
Triphysaria_sp_TC11756	(85)	HAVRLGNEASVSQVKNL	KPGAIYTLTFGATRTCAQHEELTVSVPGLL
G.max_Glyma09g36220.1	(110)	HAVRLGNEASISQTIK	VKPGKWIYALILGASRTCAQDEVLRISVPPQS
M.truncatula_AC174305_19.4	(87)	HAVRLGNEASISQTIK	VKPGTWYAIILGATRTCAQDEVLRISVPLQS
G.max_Glyma15g43180.1	(87)	HAVRLGNEASISQNIK	VKPGQLYALILGASRTCAQDEVLRISVPAQT
P.vulgaris_TC10932	(87)	HAVRLGNEASISQTIK	VKQGQLYALILGASRTCAQDEVLRISVPPQT
A.lyrata_478145	(78)	HAVRLGNEATISQKLE	VKPGSLYALTFGASRTCAQDEVLRVSVPQS
A.thaliana_AT3G08030.1	(84)	HAVRLGNEATISQKLE	VKPGSLYALTFGASRTCAQDEVLRVSVPQS
B.napus_TC89135	(84)	HAVRLGNEAKISQKLE	VKPGSLYALTFGASRTCAQDEVLRVSVPQS
G.hirsutum_TC161608	(104)	HAVKLNEATISQTIK	VKPGVLYALTFGASRTCAQDEVLRVSVPQS
T.cacao_TC4474	(104)	HAVKLNEATISQTIK	VKPGVLYALTFGASRTCAQDEVLRVSVPQS
L.serriola_TC5542	(87)	HAVRLGNEATISQNIK	VKAGSLYALTFGASRTCAQDEVLRVSVPQS
C.sinensis_TC3550	(87)	HAVRLGNEATISQTIK	VKPGALYALTFGASRTCAQDEVLRVSVPQS
P.trichocarpa_549292	(84)	HAVRLGNEASISQTIK	VKAGSLYALTFGASRTCAQDEVLRVSVPQS
P.trichocarpa_833200	(84)	HAVRLGNEASISQTIK	VKAGSLYALTFGASRTCAQDEVLRVSVPQS
V.vinifera_GSVIVT00028751001	(85)	HAVRLGNEASISQTIK	VKPGSLYALTFGASRTCAQDEVLRVSVPQS
F.vesca_TA10399_57918	(87)	HAVRLGNEASISQTIK	VKPGSLYALTFGASRTCAQDEVLRVSVPQS
N.benthamiana_TC11106	(86)	HAVRLGNDASISQTIK	VKPGSLYALTFGASRTCAQDEVLRVSVPQS
N.tabacum_TC43533	(85)	HAVRLGNDASISQTIK	VKPGSLYALTFGASRTCAQDEVLRVSVPQS
S.lycopersicum_TC199778	(86)	HAVRLGNEASISQTIK	VKPGSLYALTFGASRTCAQDEVLRVSVPQS
Triphysaria_sp_TC1599	(85)	HAVRLGNEASISQTIK	VKPGSLYALTFGASRTCAQDEVLRVSVPQS
I.nil_TC4208	(83)	HAVRLGNEASISQNIK	VKPGALYALTFAASRTCAQDEVLRVSVPQS
Consensus	(201)	HAVRLGNEASISQ	L V KGS YALTFAASRTCAQ E L VSV

FIGURE 3 (continued)

		251			300																																					
(133)	A.lyrata 340195	----	GVIP	QTM	YSG	DG	WDS	YAW	AFK	A	----	GPE	I	H	R	F	H	N	P	G																						
(132)	A.thaliana AT1G80240.1	----	GVIP	QTM	YSG	DG	WDS	YAW	AFK	A	----	GPE	I	H	R	F	H	N	P	G																						
(131)	B.napus TC65362	----	GVIP	QTM	YSG	DG	WDS	YAW	AFK	A	----	GPE	I	V	I	R	I	H	N	P	G																					
(131)	B.napus TC77368	----	GVIP	QTM	YSG	DG	WDS	YAW	AFK	A	----	GSE	I	E	I	R	I	H	N	P	G																					
(131)	O.sativa LOC_Os01g42520.1	----	GVLP	MQ	T	I	Y	G	S	N	G	W	D	S	YAW	AFK	A	----	M	D	E	V	A	L	V	I	H	N	P	G												
(128)	P.virgatum TC46275	----	GVLP	MQ	T	I	Y	G	S	N	G	W	D	S	YAW	AFK	A	----	L	D	A	V	T	L	V	I	H	N	P	G												
(129)	S.bicolor Sb03g027650.1	----	GVLP	MQ	T	I	Y	G	S	N	G	W	D	S	YAW	AFK	A	----	L	S	T	V	T	L	V	I	H	N	P	G												
(133)	Z.mays TC458850	----	GVLP	MQ	T	I	Y	G	S	N	G	W	D	S	YAW	AFK	A	----	M	D	A	V	T	L	V	I	H	N	P	G												
(180)	Zea_mays GRMZM2G034985.T02	----	GVLP	MQ	T	I	Y	G	S	N	G	W	D	S	YAW	AFK	A	----	M	D	A	V	T	L	V	I	H	N	P	G												
(128)	T.aestivum c600081041I410	----	GVLP	MQ	T	I	Y	G	S	N	G	W	D	S	YAW	AFK	A	----	L	D	V	V	E	L	V	I	H	N	P	G												
(130)	O.sativa LOC_Os04g41710.1	E	----	SGEL	P	I	Q	T	V	T	S	S	G	W	D	S	YAW	AFK	A	----	R	G	L	V	S	L	I	I	H	N	H	G										
(137)	O.sativa LOC_Os04g41740.1	E	----	SGEI	P	I	Q	T	V	T	S	S	G	W	D	S	YAW	AFK	A	----	H	S	V	V	L	F	I	V	H	N	P	G										
(137)	O.sativa LOC_Os04g41740.1	E	----	SGEI	P	I	Q	T	V	T	S	S	G	W	D	S	YAW	AFK	A	----	H	S	V	V	L	F	I	V	H	N	P	G										
(135)	O.sativa LOC_Os04g41750.1	E	----	IGEV	P	I	Q	T	V	T	S	S	G	W	D	S	YAW	AFK	A	----	R	S	D	V	S	L	I	V	H	N	P	G										
(135)	S.bicolor Sb06g021280.1	E	----	SDIL	P	I	Q	T	V	T	S	S	G	W	D	S	YAW	AFK	A	----	S	S	V	S	L	I	I	H	N	P	G											
(135)	Z.mays_ZM07MC22756.BF50088A01022693	E	----	SDVL	P	I	Q	T	V	T	S	S	G	W	D	S	YAW	AFQ	A	----	G	S	V	T	L	I	V	H	N	P	G											
(135)	Zea_mays GRMZM2G051898.T01	E	----	SDVL	P	I	Q	T	V	T	S	S	G	W	D	S	YAW	AFQ	A	----	G	S	V	T	L	I	V	H	N	P	G											
(129)	S.bicolor Sb06g021290.1	D	----	SDVL	P	I	Q	T	V	T	S	S	G	W	D	S	YAW	AFQ	A	----	R	S	A	V	T	L	I	V	H	N	P	G										
(131)	O.sativa LOC_Os04g41759.1	E	----	FGVL	P	I	Q	T	V	T	S	S	G	W	D	S	YAW	AFK	A	----	H	S	V	V	W	L	S	I	H	N	P	G										
(129)	S.bicolor Sb06g021320.1	E	----	SGVL	P	I	Q	T	V	T	S	S	G	W	D	S	YAW	AFK	A	----	H	S	T	V	W	L	S	I	H	N	P	G										
(134)	S.bicolor Sb06g021300.1	E	----	SGVL	P	V	Q	T	V	T	S	S	G	W	D	S	YAW	AFK	A	----	H	T	A	V	W	L	T	I	H	N	P	G										
(130)	Z.mays TC468877	E	----	SGVL	P	I	Q	T	V	T	S	S	G	W	D	S	YAW	AFR	A	----	H	S	A	A	L	V	I	H	N	P	G											
(133)	S.bicolor Sb06g021310.1	E	----	SGVL	P	V	Q	T	V	T	S	S	G	W	D	S	YAW	AFR	A	----	H	S	T	A	W	L	T	I	H	N	P	S										
(217)	O.sativa LOC_Os04g41770.1	E	----	FGVL	P	I	Q	T	V	T	S	T	G	W	D	S	YAW	AFK	A	----	H	S	A	V	W	L	S	I	H	N	P	G										
(136)	S.bicolor Sb06g021270.1	G	----	SGAA	Q	T	G	E	L	P	I	Q	T	V	T	S	S	G	W	D	S	YAW	AFK	A	----	A	G	L	V	S	I	I	T	H	I	C	G					
(138)	Zea_mays GRMZM2G051912.T01	G	----	SPA	Q	T	G	E	L	P	I	Q	T	V	T	S	S	G	W	D	S	YAW	AFR	A	----	A	G	L	M	S	I	T	V	T	H	N	C	G				
(138)	Zea_mays GRMZM2G405071.T01	G	----	SGAG	S	P	A	Q	T	G	E	L	P	M	Q	T	V	T	S	I	G	W	D	S	YAW	AFR	A	E	G	-	A	G	L	M	S	V	I	T	H	N	C	G
(134)	O.sativa LOC_Os02g11040.1	----	----	WAV	L	P	M	Q	T	M	Y	S	S	N	G	W	D	S	YAW	A	W	A	D	A	----	A	D	A	F	D	V	I	H	N	P	G						
(57)		----	WAV	L	P	M	Q	T	M	Y	S	S	N	G	W	D	S	YAW	A	W	A	D	A	----	A	D	A	F	D	V	I	H	N	P	G							
(132)	P.virgatum TC392	----	SGLL	A	M	Q	T	M	Y	S	S	N	G	W	D	S	YAW	A	W	A	V	A	D	----	A	G	E	V	D	V	I	H	N	P	G							
(127)	S.bicolor Sb04g007160.1	----	SGLL	A	M	Q	T	M	Y	S	S	N	G	W	D	S	YAW	A	W	A	V	A	D	----	A	D	E	V	D	V	I	H	N	P	G							

FIGURE 3 (continued)

Z.mays TC467870	(132)	-----SGLLAMQTMYSNGWDSYAWAWAN-----ADEVDIVIHNP
P.sitchensis TA14284 3332	(132)	-----FSGDISIQTLSSNGWDAYSWAFRAP-----SSVVDVILHNP
P.taeda TA10429 3352	(132)	-----FSGDISIQTLSSNGWDAYSWAFRAL-----SSVVDVILHNP
A.lyrata 9I2808	(109)	AD-----DMVASRNVDLQTLSSVQGWDPYAWAFEAEEDH-----VRLVFKNPG
A.thaliana AT1G29980.1	(145)	AD-----DMLASRNVDLQTLSSVQGWDPYAWAFEAEEDH-----VRLVFKNPG
A.lyrata 934049	(142)	-----EPIASQTIDLQTLVSVQGWDPYAWAFEAVVDR-----VRLVFRNPG
C.annuum TC19321	(135)	-----SQTIDLQTLVSVQGWDSYAWAFEAEEDD-----VRVFTNPG
S.tuberosum TC189725	(135)	-----SQTIDLQTLVSVQGWDSYAWAFEAEEDD-----VRVFTNPG
J.hindsii x regia TA1156 432290	(133)	-----SLTIDLQTLVSVQGWDPYAWAFEAEEDD-----VSLVFRNPG
G.hirsutum TC135941	(134)	-----SQSVDLQTLVSVQGWDPYAWAFEAEEDD-----VPLIFRNTG
G.raimondii TC1124	(136)	-----SQTVDLQTLVSVQGWDPYAWAFEAEEDD-----GRLIFRNP
G.max GM06MC25068 saa05g06@24507	(134)	-----SQTVDLQTLVSVQGWDPYAWAFEAEEDD-----GRLIFRNP
G.max TC286277	(134)	-----SQTVDLQTLVSVQGWDPYAWAFEAEEDD-----GRLIFRNP
L.japonicus TC41787	(136)	-----SQTVDLQTLVSVQGWDPYAWAFEAEEDD-----GRLIFRNP
M.domestica TC32901	(133)	-----SQTVDLQTLVSVQGWDPYAWAFEAEEDD-----GRLIFRNP
P.trifoliata TA6477 37690	(133)	-----SQTVDLQTLVSVQGWDPYAWAFEAEEDD-----GRLIFRNP
P.trichocarpa DUF	(134)	-----SQTVDLQTLVSVQGWDPYAWAFEAEEDD-----GRLIFRNP
P.trichocarpa MUB2	(134)	-----SQTVDLQTLVSVQGWDPYAWAFEAEEDD-----GRLIFRNP
T.officinale TA3338 50225	(133)	-----SQTVDLQTLVSVQGWDPYAWAFEAEEDD-----GRLIFRNP
H.vulgare TC163732	(134)	-----ISQTVDLQTLVSVQGWDPYAWAFEAEEDD-----GRLIFRNP
T.aestivum TC292522	(134)	-----VSQTVDLQTLVSVQGWDPYAWAFEAEEDD-----GRLIFRNP
O.sativa LOC Os01g55190.1	(139)	-----ASQTVDLQTLVSVQGWDPYAWAFEAEEDD-----GRLIFRNP
P.virgatum TC7090	(141)	-----VSQTVDLQTLVSVQGWDPYAWAFEAEEDD-----GRLIFRNP
S.bicolor Sb03g034960.1	(138)	-----ISQTVDLQTLVSVQGWDPYAWAFEAEEDD-----GRLIFRNP
S.officinale TC76182	(137)	-----VSQTVDLQTLVSVQGWDPYAWAFEAEEDD-----GRLIFRNP
Z.mays c62082358gm030403@7546	(137)	-----VSQTVDLQTLVSVQGWDPYAWAFEAEEDD-----GRLIFRNP
Z.mays_ZM07MC26525 Bfb0211C02@26447	(137)	-----VSQTVDLQTLVSVQGWDPYAWAFEAEEDD-----GRLIFRNP
Zea mays GRMZM2G324705 T01	(116)	-----VSQTVDLQTLVSVQGWDPYAWAFEAEEDD-----GRLIFRNP
Z.officinale TA720 94328	(133)	-----AASQTVDLQTLVSVQGWDPYAWAFEAEEDD-----GRLIFRNP
P.sitchensis TA10533 3332	(141)	-----TRNVDLQTLVSVQGWDPYAWAFEAEEDD-----GRLIFRNP
P.taeda TA4773 3352	(135)	-----TRNVDLQTLVSVQGWDPYAWAFEAEEDD-----GRLIFRNP
O.sativa LOC Os03g59300.1	(157)	P-----ADVAVRTLQTLVSVQGWDPYAWAFEAEEDD-----GRLIFRNP
S.bicolor Sb01g004270.1	(150)	P-----ADVAVRTLQTLVSVQGWDPYAWAFEAEEDD-----GRLIFRNP

FIGURE 3 (continued)

Z.mays_c57213778gm030403@389	(148)	P-----ADVAVRTLYSGASADTWAWGFRAS----	SPVAQVTFANPG
Z.mays_TC506958	(154)	P-----ADVAVRTLYSGASADTWAWGFRAS----	SPVAQVTFANPG
P.trichocarpa_735032	(135)	-----SDLPLQTLTSSDGGDTYSLAWKAT----	SKAVKVTFHNPG
V.vinifera_GSVIVT00025589001	(133)	-----ADLPIQTLTSTGGDTYAWAFNAT----	TEVNVVTFHNPG
S.lycopersicum_TC198542	(134)	-----SDLNIQTLTSSADGGDTYAWGFKST----	SNLVKVTFHNPG
Triphysaria_sp_TC11756	(133)	-----SRLPIQTLTGTGGDTYAWSFKAS----	SNMVKVTFHNHG
G.max_Glyma09g36220.1	(157)	-----GEVPLQTLTSLNG-DVIAWGRPT----	SSVAKVILHNPG
M.truncatula_AC174305_19.4	(134)	-----GDVPLQTLTSLNG-DVIAWGFAR----	SSFAKVTFHNPG
G.max_Glyma15g43180.1	(134)	-----GDVPLQTLTSLNG-DVIAWGFAR----	SSVVKVTFHNPG
P.vulgaris_TC10932	(134)	-----GDVPLQTLTSLNG-DVIAWGFAR----	SNVVKVTFHNPG
A.lyrata_478145	(125)	-----GDLPLQTLTLYNSFGGDDVYAWAFVAK----	TSQVTVTFHNPG
A.thaliana_AT3G08030.1	(131)	-----GDLPLQTLTLYNSFGGDDVYAWAFVAK----	TSQVTVTFHNPG
B.napus_TC89135	(131)	-----GDLPLQTLTLYNSFGGDDVYAWAFVAK----	TSVTVTFHNPG
G.hirsutum_TC161608	(151)	-----GDLPLQTLTLYSSIGDDVYAWGFIPK----	TKYATVKFHNPG
T.cacao_TC4474	(151)	-----GDLPLQTLTLYSSIGDDVYAWGFIAK----	SKYITVTFHNPG
L.serriola_TC5542	(134)	-----GDLPLQTLTLYSSDGGDVYAYGFKAN----	SSSVRLTFHNPG
C.sinensis_TC3550	(134)	-----GDLPLQTLTLYDING-DTYAWGFRAK----	TNIVTVTFHNPG
P.trichocarpa_549292	(131)	-----GDLPLQTLTLYSSNGGDTYAWGFKIN----	ATVVKVSFRNTG
P.trichocarpa_833200	(131)	-----GDLPLQTLTLYSSNGGDTYAWGFRIN----	ATVVQVTFHNTG
V.vinifera_GSVIVT00028751001	(132)	-----GDLPLQTLTLYDSFGGDDVYAWGFRAT----	SNEAKVTFHNTG
F.vesca_TA10399_57918	(134)	-----GDLPLQTLTLYSSNGGDTYAWGFKAT----	SNTVKVTFHNPG
N.benthamiana_TC11106	(133)	-----GDLPLQTLTLYSSNGGDTYAWGFYAT----	SNVVKVTFHNPG
N.tabacum_TC43533	(132)	-----GDLPLQTLTLYSSNGGDTYAWGFYAT----	SNVVKVSFHNPG
S.lycopersicum_TC199778	(133)	-----GDLPLQTLTLYSSNGGDTYAWGFYAT----	SNVVKITFHNPG
Triphysaria_sp_TC1599	(132)	-----GDLPLQTLTLYSSNGGDTYAWGFRAT----	SKNVTVSFSNPG
I.nil_TC4208	(130)	-----GDIPLQTLTLYSSNGGDDVVSFAFRAA----	SEAAVTFHNPG
Consensus	(251)	SG LPLQTLTLYSS GWDSYAWAFKA V L FHNPG	

FIGURE 3 (continued)

20/77

A.lyrata_340195
 A.thaliana_AT1G80240.1
 B.napus_TC65362
 B.napus_TC77368
 O.sativa_LOC_Os01g42520.1
 P.virgatum_TC46275
 S.bicolor_Sb03g027650.1
 Z.mays_TC458850
 Zea_mays_GRMZM2G034985_T02
 T.aestivum_c6008104@11410
 O.sativa_LOC_Os04g41710.1
 O.sativa_LOC_Os04g41740.1
 O.sativa_LOC_Os_DUF642.3
 O.sativa_LOC_Os04g41750.1
 S.bicolor_Sb06g021280.1
 Z.mays_ZM07MC22756_BF50088A01@22693
 Zea_mays_GRMZM2G051898_T01
 S.bicolor_Sb06g021290.1
 O.sativa_LOC_Os04g41759.1
 S.bicolor_Sb06g021320.1
 S.bicolor_Sb06g021300.1
 Z.mays_TC468877
 S.bicolor_Sb06g021310.1
 O.sativa_LOC_Os04g41770.1
 S.bicolor_Sb06g021270.1
 Zea_mays_GRMZM2G051912_T01
 Zea_mays_GRMZM2G405071_T01
 O.sativa_LOC_Os02g11040.1
 (94)
 P.virgatum_TC392
 S.bicolor_Sb04g007160.1
 (169)
 (164)

FIGURE 3 (continued)

Z.mays TC467870	(169)	VTEDPACGGLIDSVAIKTLNPPR----	RTKNLVKNGDFEEGPIYIPGTK
P.sitchensis TA14284 3332	(170)	VEEDPACGPVLDAVAIKELFPPR----	FTKFNLIKNGDFEEGPIYMFASSN
P.taeda TA10429 3352	(170)	VTEDPACGPVLDAVAIKELFPPR----	TRFDLIKNGDFEEGPIYMFSSN
A.lyrata 9I2808	(152)	MEDDPTCGPIIDDAIAIKKLFPPD----	KPKDNAVINGDFEEGPIYMFSSN
A.thaliana AT1G29980.1	(188)	MEDDPTCGPIIDDAIAIKKLFPPD----	KPKDNAVINGDFEEGPIYMFSSN
A.lyrata 934049	(183)	MEDDPTCGPIIDDAIAIKKLFPPD----	KTKGNAVINGDFEEGPIYMFSSN
C.annuum TC19321	(172)	MEDDPTCGPIIDDAIAIKKLFPPD----	KSKDNAVINGDFEEGPIYMFSSN
S.tuberosum TC189725	(172)	MEDDPTCGPIIDDAIAIKKLFPPD----	KSKNNAVINGDFEEGPIYMFSSN
J.hindsii x regia TA1156 432290	(170)	MEDDPTCGPIIDDAIAIKKLFPPD----	RLKDNAVINGNLEEGPIYMFSSN
G.hirsutum TC135941	(171)	MEDDPTCGPIIDDAIAIKKLFPPD----	QPKDNAVINGSGFEFGPIYMFSSN
G.raimondii TC1124	(173)	MEDDPTCGPIIDDAIAIKKLFPPD----	KPKDNAVINGSGFEFGPIYMFSSN
G.max GM06MC25068 saa05g06024507	(171)	MEDDPTCGPIIDDAIAIKKLFPPD----	KPKDNAVINGDFEEGPIYMFSSN
L.japonicus TC41787	(173)	MEDDPTCGPIIDDAIAIKKLFPPD----	KPKDNAVINGDFEEGPIYMFSSN
M.domestica TC32901	(170)	MEDDPTCGPIIDDAIAIKKLFPPD----	KPKDNAVINGDFEEGPIYMFSSN
P.trifoliata TA6477 37690	(170)	MEDDPTCGPIIDDAIAIKKLFPPD----	KPKDNAVINGDFEEGPIYMFSSN
P. trichocarpa DUF	(171)	MEDDPTCGPIIDDAIAIKKLFPPD----	KPKDNAVINGDFEEGPIYMFSSN
P. trichocarpa MUB2	(171)	MEDDPTCGPIIDDAIAIKKLFPPD----	KPKDNAVINGDFEEGPIYMFSSN
T.officinale TA3338 50225	(170)	MEDDPTCGPIIDDAIAIKKLFPPD----	KPKDNAVINGDFEEGPIYMFSSN
H.vulgare TC163732	(172)	MEDDPTCGPIIDDAIAIKKLFPPD----	KAKENMAVINGDFEEGPIYMFSSN
T.aestivum TC292522	(172)	MEDDPTCGPIIDDAIAIKKLFPPD----	KAKENMAVINGDFEEGPIYMFSSN
O.sativa LOC_Os01g55190.1	(177)	MEDDPTCGPIIDDAIAIKKLFPPD----	KPKDTVVSNGDFEEGPIYMFSSN
P.virgatum TC7090	(179)	MEDDPTCGPIIDDAIAIKKLFPPD----	KPKDNAVINGDFEEGPIYMFSSN
S.bicolor Sb03g034960.1	(176)	MEDDPTCGPIIDDAIAIKKLFPPD----	KPKDNAVINGDFEEGPIYMFSSN
S.officinale TC76182	(175)	MEDDPTCGPIIDDAIAIKKLFPPD----	KPKDNAVINGDFEEGPIYMFSSN
Z.mays c62082358gm03040307546	(175)	MEDDPTCGPIIDDAIAIKKLFPPD----	KPKDNAVINGDFEEGPIYMFSSN
Z.mays_ZM07MC26525 Bfb0211C02026447	(175)	MEDDPTCGPIIDDAIAIKKLFPPD----	KPKDNAVINGDFEEGPIYMFSSN
Zea_mays GRMZM2G324705 T01	(154)	MEDDPTCGPIIDDAIAIKKLFPPD----	KPKDNAVINGDFEEGPIYMFSSN
Z.officinale TA720 94328	(174)	MEDDPTCGPIIDDAIAIKKLFPPD----	KPKDNAVINGDFEEGPIYMFSSN
P.sitchensis TA10533 3332	(178)	MEDDPTCGPIIDDAIAIKKLFPPD----	KLRDNAVINGDFEEGPIYMFSSN
P.taeda TA4773 3352	(172)	LEEDPTCGPIIDDAIAIKKLFPPD----	KLRDNAVINGDFEEGPIYMFSSN
O.sativa LOC_Os03g59300.1	(194)	VQEDASCGPILDAVAIKELTPY----	PTKDNLIKNEGFEIGPQVFNST
S.bicolor Sb01g004270.1	(187)	VQEDASCGPILDAVAIKELTPY----	PTKDNLIKNEGFEIGPQVFNST

FIGURE 3 (continued)

Z.mays_c57213778gm0304030389	(185)	VQEDPACGGLLDVAIAIKELPAPY----	PTKDNLIKNDGFEIGPQVLNRST
Z.mays_TC506958	(191)	VQEDPACGGLLDVAIAIKELPAPY----	PTKDNLIKNDGFEIGPQVLNRST
P.trichocarpa_735032	(171)	VQEDPACGGLLDVAIAIKELPLK----	RTIGNLVKNGGFEVGFHFKNFS
V.vinifera_GSVIT00025589001	(169)	IQEDPTCGPLLDVAIAIKQMPPLK----	LTRGSLAKNGGYETGPHVFKNFS
S.lycopersicum_TC198542	(170)	TQEDPTCGPLLDHVAIAIKEMLMAT----	YTKGNLVKNGGFEIGPHTFKNFS
Triphysaria_sp_TC11756	(169)	TQEDPACGGLLDVAIAIKEMQPLR----	YTRGNLVKNGGFEIGPHTFKNFS
G.max_Glyma09g36220.1	(192)	IQEDPACGGLLDVAIAIEFCPPK----	PTRANLVKNPGFEVGFPIFNST
M.truncatula_AC174305_19.4	(169)	MQEDPTCGPLLDVAIAIREFYPPM----	PTRANLVKNPGFEVGFPIFNST
G.max_Glyma15g43I80.1	(169)	VQEDPACGGLLDVAIAIREFYPPM----	PTRNLVKNPGFEVGFPIFNST
P.vulgaris_TC10932	(169)	VQEDPACGGLLDVAIAIREFYPPM----	PTRNLVKNPGFEVGFPIFNST
A.lyrata_478145	(161)	VQEDPACGGLLDVAIAIKELVHPI----	YTKGNLVKNGGFEVGFPIFNST
A.thaliana_AT3G08030.1	(167)	VQEDPACGGLLDVAIAIKELVHPI----	YTKGNLVKNGGFEVGFPIFNST
B.napus_TC89135	(167)	VQEDPACGGLLDVAIAIKELVHPM----	YTKGNLVKNGGFEVGFPIFNST
G.hirsutum_TC161608	(187)	VQEDPTCGPLLDVAIAIKELVRPR----	ATKYNLVKNGGFEVGFPIFNST
T.cacao_TC4474	(187)	VQEDPTCGPLLDVAIAIKELVRPM----	PTRGNLVKNPGFEVGFPIFNST
L.serriola_TC5542	(170)	VQEDPKCGPLLDVAIAIKELPPR----	PTRLNIVKNGGFEVGFPIFNST
C.sinensis_TC3550	(169)	VQEDPACGPLLDVAIAIKELYPMP----	PTRNLVKNPGFEVGFPIFNST
P.trichocarpa_549292	(167)	VQEDPACGPLLDVAIAIKELFPPM----	PTRNLVKNHGFEVGFPIFNST
P.trichocarpa_833200	(167)	VQEDPACGPLLDVAIAIKELFPPM----	PTRNLVRNHGFEVGFPIFNST
V.vinifera_GSVIT00028751001	(168)	VQEDPACGGLLDVAIAIKELFPPM----	PTRNLVKNHGFEVGFPIFNST
F.vesca_TA10399_57918	(170)	IQEDPACGGLLDVAIAIKELFPVL----	PTRNLVRNPGFEVGFPIFNST
N.benthamiana_TC11106	(169)	VQEDPACGGLLDVAIAIKELFPVQ----	PTRNLVRNPGFEVGFPIFNST
N.tabacum_TC43533	(168)	VQEDPACGGLLDVAIAIKELFPVR----	PTRNLVRNPGFEVGFPIFNST
S.lycopersicum_TC199778	(169)	VQEDPACGGLLDVAIAIKELFPVR----	PTRNLVRNPGFEVGFPIFNST
Triphysaria_sp_TC1599	(168)	KQEDPACGGLLDVAIAIKELFPMP----	PTRNLVRNPGFEVGFPIFNST
I.nil_TC4208	(166)	KQEDPTCGPLLDVAIAIKELLPVL----	PTRNLVRNPGFEVGFPIFNST
Consensus	(301)	VEEDPACGGLLDVAIAIK LF P	TK NLVKNGDFFGPFYMF NT

FIGURE 3 (continued)

		351	
		400	
(215)	A.lyrata 340195	SGVLIPPFIEDDNSPLPGWMIES-LKAVKYVDKAHFAVPEGHRAIELVGG	
(214)	A.thaliana AT1G80240.1	WGVLIPPFIEDDNSPLPGWMIES-LKAVKYVDKAHFAVPEGHRAIELVGG	
(213)	B.napus TC65362	WGVLIPPFIEDDNSPLPGWIES-LKAVKYVDKAHFAVPEGHRAIELVGG	
(213)	B.napus TC77368	WGVLIPPFIEDDNSPLPGWKIES-LKAVKYVDKAHFAVPEGHRAIELVGG	
(213)	O.sativa LOC_Os01g42520.1	WGVLVPPNIEDDHSPLPAWMIMS-SKAVKYVDAAHFAVPPQGARAVELVGG	
(210)	P.virgatum TC46275	WGVLVPPNIEDDHSPLPAWMIAS-SKAVKYVDAAHFAVPPQGARAVELVGG	
(211)	S.bicolor Sb03g027650.1	WGVLVPPNIEDDHSPLPGWMIAS-SKAVKYVDAAHFAVPPQGARAVELVGG	
(219)	Z.mays TC458850	WGVLVPPNIEDDHSPLPAWMIVS-SKAVKYVDAAHFAVPPQGARAVELVGG	
(266)	Zea mays GRMZM2G034985 T02	WGVLVPPNIEDDHSPLPAWMIVS-SKAVKYVDAAHFAVPPQGARAVELVGG	
(210)	T.aestivum c6000810411410	WGVLVPPNIEDDHSPLPAWMIMS-SKAVKYVDAAHFAVPPQGARAVELVGG	
(214)	O.sativa LOC_Os04g41710.1	WGMVPPPISEDDHSPLPGWVMSDTKAVKCVDSAHFTVPHGARAVELVSG	
(221)	O.sativa LOC_Os04g41740.1	WGVLVPPMEDDDYSPLSPWTILSTTKSVKYIDAHYAVPGGARAVELVSG	
(221)	O.sativa LOC_Os04g41740.1	WGVLVPPMEDDDYSPLSPWTILSTTKSVKYIDAHYAVPGGARAVELVSG	
(219)	O.sativa LOC_Os04g41750.1	WGVLVPPMEDDDYSPLSPWTIMGYTKSVKYVDAAHFAVPPQGARAVELVAG	
(219)	S.bicolor Sb06g021280.1	WGVLVPPMEDDDYSPLSPWMLSTTKSVKYVDAHFAVPHGARAVELVSG	
(219)	Z.mays_ZM07MC22756 BF50088A01022693	WGVLVPPMEDDDYSPLAPWMLSTTKSVKYVDAHFAVPHGARAVELVSG	
(219)	Zea mays GRMZM2G051898 T01	WGVLVPPMEDDDYSPLAPWMLSTTKSVKYVDAHFAVPHGARAVELVSG	
(214)	S.bicolor Sb06g021290.1	WGVLVPPMEDDDYSPLSPWMLSTTKSVKYLDAAHFAVPHGARAVELLSG	
(215)	O.sativa LOC_Os04g41759.1	WGVLVPPPIFEDEHSPLPGWIMSDTKVYVDSPHHRVPPQGARAVELVAG	
(213)	S.bicolor Sb06g021320.1	WGVLVPPITEDEHSPLPGWIMSDTKVYVDAHFAVPRGAGAVELVAG	
(218)	S.bicolor Sb06g021300.1	WGVLVPPMEDDDVSPLPGWIMSDTKAVKYLDAAHFAVPHGAYAVELVAG	
(215)	Z.mays TC468877	WGVLVPPMEDDDVSPLPGWTVMSDTKAVKYLDAAHFAVPRGSRVELVAG	
(217)	S.bicolor Sb06g021310.1	WGVLVPPMEDDDVSPLPGWVMSDTKVYVDAAHFAVPRGARAVELVAG	
(302)	O.sativa LOC_Os04g41770.1	WGVLVPPMEDDVHSPLPGWVMSDTKVYVDSARHAPVSGAHAVEMVAG	
(227)	S.bicolor Sb06g021270.1	CGVMLPPMDQDCVSPPLPGWKIMSYKKSVMYVDSAHFAVPRGARAVELVSG	
(227)	Zea mays GRMZM2G051912 T01	WGVLVPPVEDDDVSPLPGWIMAYKKSVMYVDSAHFAVPRGARAVELVSG	
(237)	Zea mays GRMZM2G405071 T01	WGVLVPPVEDDDVSPLPGWKVMYKKSVMYVDSAHFAVPRGARAVELVSG	
(217)	O.sativa LOC_Os02g11040.1	WGVLIPSMVVDHSPLPGWVMS-LKAVKYIDSDHFAVPRGRRAVELLAG	
(140)	P.virgatum TC392	WGVLIPSMVVDHSPLPGWVMS-LKAVKYIDSDHFAVPRGRRAVELLAG	
(215)	S.bicolor Sb04g007160.1	WGVLIPSRVVDHSPLPGWVMS-LKAVKYIDGGSFAVPRGRRAVELLAG	
(207)		WGVLIPSRVVDHSPLPGWVMS-LKAIKYIDSDSFAVPRGRRAVELLAG	

FIGURE 3 (continued)

24/77

Z.mays TC467870	(215) WGVLLIPSRVVEDHSPLPGWVIES-LKAIKYIDGESFAVPRGRRAVELLAG
P.sitchensis TA14284_3332	(216) SGVLLPPNVEDAISPLPGWIIIES-LKAVKYLDGAHFAVPQGRRAVELVAG
P.taeda TA10429_3352	(216) SGVLLPPNVEDAISPLPGWIIIES-LKAVKYLDGAHFAVPQGRRAVELVAG
A.lyrata 9I2808	(198) LGVLLPTNLDEETSSPLPGWTVES-NRAVRFVDSDFHSPVPGKRAVELLSG
A.thaliana AT1G29980.1	(234) LGVLLPTNLDEEISSPLPGWTVES-NRAVRFVDSDFHSPVPGKRAVELLSG
A.lyrata 934049	(229) LGVLLPTNLDEEISSPLPGWTVES-NRAVRFVDSDFHSPVPGKRAVELLSG
C.annuum TC19321	(218) LGVLLPTNLDEETSSPLPGWIVES-NRAVRYIDTYHFTVPEGKRAVELLSG
S.tuberosum TC189725	(218) LGVLLPTNLDEETSSPLPGWIVES-NRAVRYIDTYHFTVPEGKRAVELLSG
J.hindsii x regia TA1156_432290	(216) LGVLLPTNLDEEISSPLPGWVIES-NRAIRYIDSDHFTVPEGKRAVELLSG
G.hirsutum TC135941	(217) LGVLLPTNLDEETSPPLPGWVIES-TRAVRYIDSNHYAVPEGKRAVELVSG
G.raimondii TC1124	(219) LGVLLPTNLDEETSPPLPGWVIES-NRAVRYIDSNHYAVPEGKRAVELVSG
G.max GM06MC25068_saa05g06_24507	(217) LGVLLPTNLDEEASSPLPGWIVES-NRAVRYIDSDHYSVPQGRRAIELLSR
G.max TC286277	(217) LGVLLPTNLDEETSSPLPGWIVES-NRAVRYIDSDHYSVPQGRRAIELLSR
L.japonicus TC41787	(219) MGVLLPTNLDEETSSMPGWIVES-NRAIRYIDSDHYAVPQGRRAIELLSG
M.domestica TC32901	(216) LGVLLPTNLDEETSSPLPGWIVES-NRAVRYIDSYHYAVPQGRRAIELVSG
P.trifoliata TA6477_37690	(216) LGVLLPTNLDEETSSPLPGWIVES-NRAVRYIDSYHYSPVQGRRAIELVSG
P.trichocarpa DUF	(217) LGVLLPTKLDEETSSPLPGWIIIES-YRAVRYIDSYHYSPVQGRRAIELLSG
P.trichocarpa MUB2	(217) LGVLLPTKLDEETSSPLPGWIIIES-YRAVRYIDSYHYSPVQGRRAIELLSG
T.officinale TA3338_50225	(216) LGVLLPTNLDEEISSPLPGWIVES-NRAVRYIDANHFSPVPGKRAIELLSG
H.vulgare TC163732	(218) FGVLLPTNLDEQTSALPGWMIES-NRAVRFVDSQYTPVQGRRAIELLSG
T.aestivum TC292522	(218) FGVLLPTNLDEQTSALPGWMIES-NRAVRFVDSQYTPVQGRRAIELLSG
O.sativa IOC Os01g55190.1	(223) FGVLLPTNLDEQTSALPGWMIES-NRAVRFVDSQYTPVQGRRAIELLSG
P.virgatum TC7090	(225) FGVLLPTNLDEQTSALPGWMIES-NRAVRFVDSQYTPVQGRRAIELLSG
S.bicolor Sb03g034960.1	(222) FGVLLPTNLDEQTSALPGWMIES-NRAVRYIDSEYKVPQGRRAIELLSG
S.officinale TC76182	(221) FGVLLPTNLDEQTSALPGWMIES-NRAVRYIDSEYKVPQGRRAIELLSG
Z.mays c62082358gm030403@7546	(221) FGVLLPTNLDEQTSALPGWMIES-NRAVRYIDSDQYKVPQGRRAIELLSG
Z.mays_ZM07MC26525_Bfb0211C02@26447	(221) FGVLLPTNLDEQTSALPGWMIES-NRAVRYIDSDQYKVPQGRRAIELLSG
Zea mays GRMZM2G324705.T01	(200) FGVLLPTNLDEQTSALPGWMIES-NRAVRYIDSEYKVPQGRRAIELLSG
Z.officinale TA720_94328	(220) LGVLLPTNLEEAMSAALPGWLIVES-NRAVRYIDSNHFVVPQGRRAIELLSG
P.sitchensis TA10533_3332	(224) LGVLLPTNLDEATSSPLPGWIMES-SKAVRYIDSNHYKVPQGRRAIELLSG
P.taeda TA4773_3352	(218) LGVLLPTNLDEATSSPLPGWIIIES-SKAVRYIDSNHYKVPQGRRAIELLSG
O.sativa IOC Os03g59300.1	(240) VGVLLPPKQKDATSPLPGWIIIES-LKAVRFIDAAHFSVPAGQYAVELVAG
S.bicolor Sb01g004270.1	(233) VGVLLPPKQKDATSPLPGWIIIES-LKAVRFIDAAHFSVPAGQYAVELVAG

FIGURE 3 (continued)

Z.mays_c57213778gm0304030389	(231)	GVLLPPKQKDATSPLPGWIIES-LKAVRFIDAAHFSVPAGQYAVELVAG
Z.mays_TC506958	(237)	GVLLPPKQKDATSPLPGWIIES-LKAVRFIDAAHFSVPAGQYAVELVAG
P.trichocarpa_735032	(217)	TGILIPAKQQDLISPLPGWIVES-LKPVKYIDKKHFFVPSGFAAIEMVAG
V.vinifera_GSVIT00025589001	(215)	TGVLLLPKQQDLISPLPGWIIES-LKPVKYIDSKHFSVPTGLAAIELIGG
S.lycopersicum_TC198542	(216)	TGVLLPLKQDKYSPIPGWMVQF-AKPSKYIDSKHFFVPSGNAAVELIGG
Triphysaria_sp_TC11756	(215)	TGVLLPKQDLYSPIPGWIVES-LKPVKYIDSRHFI VPGVGLAAIELVGG
G.max_Glyma09g36220.1	(238)	NGVLLPPEQEDHVSPLPGWMIES-LKAVKFI DAKHFNVPFGQGAVELIAG
M.truncatula_AC174305_19.4	(215)	NGVLLPPKQQDLVSPLPGWIIES-LKAIFIDSNHFQVFPFGKGAVELVAG
G.max_Glyma15g43180.1	(215)	NGVLLPPQQQDGFSPPLPGWIIES-LKAVKFI DSKHFNVPFGLGAVELVAG
P.vulgaris_TC10932	(215)	NGVLLPPQQQDGFSPPLPGWIIES-LKAVKFI DSKHFNVPFGLGAVELVAG
A.lyrata_478145	(207)	QGVLLPPKQEDLTSPLPGWIIES-LKAVKFI DSKYFNVPFGHAAIELVAG
A.thaliana_AT3G08030.1	(213)	QGVLLPPKQEDLTSPLPGWIIES-LKAVKFI DSKYFNVPFGHAAIELVAG
B.napus_TC89135	(213)	QGVLLPPKQEDLTSPLPGWIIES-LKAVKFI DSKYFNVPFGHAAIELVAG
G.hirsutum_TC161608	(233)	NGVLLPPRQEDFTSPLPGWIIES-LKAVKFINAKHFNVPAGPLNRRTGLP
T.cacao_TC4474	(233)	NGVLLPPRQEDFTSPLPGWIIES-LKAVKFI DSKHFNVPAGKAAVELVAG
L.serriola_TC5542	(216)	NGVLLPPRQEDITSPLPGWIIES-LKAVKFI DNKHFYVPSGVAAIELVAG
C.sinensis_TC3550	(215)	NGVLLPPRQEDLTSPLPGWIIES-LKAVKFI DAKHFNVPYGHAAVELVAG
P.trichocarpa_549292	(213)	SGVLLPPRQEDLTSPLPGWMIES-LKAVKFI DKKHFNVPFGLAAVELVAG
P.trichocarpa_833200	(213)	NGVLLPPRQEDLTSPLPGWIIES-LKAVKFI DKKHFNVPFGLAAVELVAG
V.vinifera_GSVIT00028751001	(214)	NGVLLPPKQEDLTSPLPGWIIES-LKAVKFI DKKHFNVPFGLAAVELLAG
F.vesca_TA10399_57918	(216)	HGVLLPPKQDLATSPLPGWIIES-LKAVKFI DAKHFNVPFGKGAVELVAG
N.benthamiana_TC11106	(215)	HGVLLPPKQEDLTSPLPGWIIES-LKAVKFLDSAHFNVPFGRAAVELLAG
N.tabacum_TC43533	(214)	HGVLLPPKQEDLTSPLPGWIIES-LKAVKFLDSAHFNVPFGQAAVELLAG
S.lycopersicum_TC199778	(215)	HGVLLPPKQEDLTSPLPGWIIES-LKAVKFLDSAHFNVPFGQAAVELLAG
Triphysaria_sp_TC1599	(214)	HGVLLPPKQEDLTSPLPGWIIES-LKAIFYIDAKHFNVPFGKAAVELLAG
I.nil_TC4208	(212)	NGVLLPPKQQDMVSPPLPGWIIES-LKAVKYIDAAHFNIPFGRAAVELVAG
Consensus	(351)	GVLLPP LEDD SPLPGWIIES KAVKYIDS HFAVP G RAVELVAG

FIGURE 3 (continued)

26/77

	401	450
A.lyrata_340195	(264)	-KESAIQIVRTS-LNKFYALTFSVGWARD-GCEGPMT-----VE
A.thaliana_AT1G80240.1	(263)	-KESAIQIVRTS-LNKFYALTFNVGDARD-GCEGPMI-----VE
B.napus_TC65362	(262)	-KESAVSQIVRTS-PNKFYALTFNVGDARD-ACEGPMA-----VE
B.napus_TC77368	(262)	-KESAVSQIVRTSSLNKFYALGFNVGDARD-ACEGPMA-----VE
O.sativa_LOC_Os01g42520.1	(262)	-KESALVQEVR-TVPGWTYRLSFAVGWARD-GCAGSMV-----AE
P.virgatum_TC46275	(259)	-RESALVQEVR-TVPGWSYRLAFAVGDAGD-GCAGPMV-----AE
S.bicolor_Sb03g027650.1	(260)	-RESALVQEVR-TVPGWSYRLAFAVGDSGD-GCTGSMV-----AE
Z.mays_TC458850	(268)	-RESALVQEVR-TVPGWTYRLAFAVGDSGD-GCAGSMA-----AE
Zea_mays_GRMZM2G034985_T02	(315)	-RESALVQEVR-TVPGWTYRLAFAVGDSGD-GCAGSMA-----AE
T.aestivum_c60008104@11410	(259)	-KESALVQEVR-TVPGWAYRLSFAVGDSAD-AAR-----A
O.sativa_LOC_Os04g41710.1	(264)	-LETALMQEVR-TVPGRSYRLEFSVGDAKD-GCVGSMQ-----VK
O.sativa_LOC_Os04g41740.1	(271)	-METAMLQEVTS-TVPGRSYRLEFSVGDAKD-GCSGLT-----VQ
O.sativa_LOC_Os04g41740.1	(271)	-METAMLQEVTS-TVPGRSYRLEFSVGDAKD-GCSGLT-----VQ
O.sativa_LOC_Os04g41750.1	(269)	-MEALVQEVC-TVPGRSYRLEFSVGDAKD-GCVGMS-----VQ
S.bicolor_Sb06g021280.1	(269)	-METALVQNAT-TVPGLPYRLQFSAGDAGN-GCVGSMA-----VQ
Z.mays_ZM07MC22756_BF50088A01@22693	(269)	-METALVQNTVA-TVPGQLYRLQFSAGDAGN-GCVGSMT-----VQ
Zea_mays_GRMZM2G051898_T01	(269)	-METALVQNTVA-TVPGQLYRLQFSAGDAGN-GCVGSMT-----VQ
S.bicolor_Sb06g021290.1	(264)	-VETALAQQDVA-TVAGRPRYLEFTSGDAGD-GCSGSL-----LR
O.sativa_LOC_Os04g41759.1	(265)	-RETALVQEVA-TVPGRSYRLSFSVGDAKD-GCKDSL-----VE
S.bicolor_Sb06g021320.1	(263)	-RECALVQEVR-TVPRRSYKLFSAVGDAAN-GCEGNLA-----VD
S.bicolor_Sb06g021300.1	(268)	-TEAALVQEVQ-TKPGRSYKLFSVGDAAN-GCASTLL-----VD
Z.mays_TC468877	(265)	-TEAALVQEVR-TTPGRSYTLFSAGDAAD-GCAGALR-----VD
S.bicolor_Sb06g021310.1	(267)	-REAALVQEVGTVPGRRYRLSFSVGDAKD-GCEGSLA-----VE
O.sativa_LOC_Os04g41770.1	(352)	-RECALVQEVA-TVPGRRYTLFSVGDAKD-GCIGSLA-----VD
S.bicolor_Sb06g021270.1	(277)	-VETALMQEVYTVEGWSYRLEFSVGDAAN-GCASSYDDSSSGMKVK
Zea_mays_GRMZM2G051912_T01	(277)	-VETALIQEVYTVEGWSYRLEFSAGDAAD-GCASSYDDGSSS--GMKVK
Zea_mays_GRMZM2G405071_T01	(287)	-VEAALIQEVDTTVEGWSYRLEFSAGDAAN-GCASSYDGSSS--GMKLK
O.sativa_LOC_Os02g11040.1	(266)	-RESIAQVIR-TVPGRQYALSFTVGDAKN-GCEGSLV-----VE
P.virgatum_TC392	(189)	-RESIAQVIR-TVPGRQYALSFTVGDAKN-GCEGSLV-----VE
S.bicolor_Sb04g007160.1	(264)	-RESAVAQVIR-TVPGRRYALSFTVGDAKN-ACRGSLM-----VE
	(256)	-RESIAQVIR-TVPGRQYALSFTVGDAKN-ACRGSLM-----VE

FIGURE 3 (continued)

27/77

Z.mays TC467870	(264)	RESAIAQVIR-TVPGRQYVLSFTIGDASN-ACRGSIM-----VE
P.sitchensis TA14284 3332	(265)	KESAIQIVR-TVPGRFYRLSFAVGDAARN-TCLGNMI-----VE
P.taeda TA10429 3352	(265)	KESAIQIVR-TIPGRFYRLSFAVGDAARN-TCMGNMI-----VE
A.lyrata 9I2808	(247)	KEGIIQMVETK-ADKPYTLTFSLGAGD-KCKEPLA-----IM
A.thaliana AT1G29980.1	(283)	KEGIIQMVETK-ADKPYTLTFSLGAGD-KCKEPLA-----IM
A.lyrata 934049	(278)	KEGIIQMVETK-ANIPYKLSFSLGHAGD-KCKEPLA-----VM
C.annuum TC19321	(267)	KEGIIQMVETK-PNKPYRLTFLGHAGD-SCKEPLA-----IM
S.tuberosum TC189725	(267)	KEGIIQMVETK-PNKPYRLTFLGHAGD-SCKEPLA-----IM
J.hindsii_x_regia TA1156 432290	(265)	KEGIIQMVETT-ANKPYTLTFSLGQAGD-KCKQPLA-----VM
G.hirsutum TC135941	(266)	KEGIIQMVETK-PDKLYSLTFSLGAGD-KCKEPLA-----VM
G.raimondii TC1124	(268)	KEGIIQMVETT-PNKLKSLTFSLGAGD-KCKEPLA-----VM
G.max_GM06MC25068_saa05g06024507	(266)	KEGIIQMVETK-PDMLYSLTFSLGAGD-KCKEPLA-----VM
G.max_TC286277	(266)	KEGIIQMVETK-PDMLYSLTFSLGAGD-KCKEPLA-----VM
L.japonicus TC41787	(268)	KEGIIQMVETS-PDKLYSLTFSLGAGD-KCKEPLA-----VM
M.domestica TC32901	(265)	KEGIIQMVETS-PNKPYTLTFSLGAGD-KCKQPLG-----IM
P.trifoliata TA6477 37690	(265)	KEGIIQMVETT-ANKDYSLTFSLGAGD-KCKQPLA-----VM
P. trichocarpa DUF	(266)	KEGIIQMVETT-PSKPYTMSFALGHAGD-KCKQPLA-----VM
P. trichocarpa MUB2	(266)	KEGIIQMVETT-PSKPYTMSFALGHAGD-KCKQPLA-----VM
T.officinale TA3338 50225	(265)	KEGIIQMVTTT-PNKPYRMSFSLGHAGD-MRKQPLA-----VM
H.vulgare TC163732	(267)	KEGIIQMVETT-PQKAYILTFTLSAGD-SCQPPMA-----VM
T.aestivum TC292522	(267)	KEGIIQMVETT-PQKAYILTFTLSAGD-SCQPPMA-----VM
O.sativa LOC_Os01g55190.1	(272)	KEGIIQMVETT-PQKEYSLTFTLSAGD-SCQPPMA-----VM
P.virgatum TC7090	(274)	KEGIIQMVETT-PQKEYSLTFTLSAGD-SCQPPMA-----IM
S.bicolor Sb03g034960.1	(271)	KEGIIQMVETT-PQKVYSLTFTLSAGD-SCQPPMA-----IM
S.officinale TC76182	(270)	KEGIIQMVETT-PQKVYSLTFTLSAGD-SCQPPMA-----VM
Z.mays_c62082358gm03040307546	(270)	KEGIIQMVETT-PQKVYSLTFTLSAGD-SCQPPMA-----VM
Z.mays_ZM07MC26525_Bfb0211C02026447	(270)	KEGIIQMVETT-PEKVYSLTAMLGAGD-SCQPPMA-----IM
Zea_mays_GRMZM2G324705_T01	(249)	KEGIIQMVETT-PQKVYSLTFTLSAGD-SCQPPMA-----VM
Z.officinale TA720 94328	(269)	KEGIIQMVETT-PDKQYSLTFTVGAAGD-SCQAPLA-----VM
P.sitchensis TA10533 3332	(273)	KEGIIQMVETD-PGNKYTMSFLLGQAGD-GCQEPMA-----VM
P.taeda TA4773 3352	(267)	KEGIIQMVETE-PGKKYTMSFLMGQAGD-GCQEPMA-----VM
O.sativa LOC_Os03g59300.1	(289)	RESAIAQVIR-TVANRAYNLSFVVGDAKN-GCHGSMLE-----
S.bicolor Sb01g004270.1	(282)	RESAIAQVIR-TVPNRAYNLSFAVGDAKN-GCHGSMLE-----

FIGURE 3 (continued)

Z.mays_c57213778gm0304030389	(280)	-RESAIAQVIR-TVPNRAYNLSFAVGDARN-GCHGSMLE	V
Z.mays_TC506958	(286)	-RESAIAQVIR-TVPNRAYNLSFAVGDARN-GCHGSMLE	
P.trichocarpa_735032	(266)	-RESAIAQVIR-TIPNKFYNLTFTIGDAKN-ACHGSMME	
V.vinifera_GSVIVT00025589001	(264)	-RESAIAQIIR-TVPNKSYNLTFTIGDAKN-GCHGSMTE	
S.lycopersicum_TC198542	(265)	-RETGIAQTIR-TIPKQFYNLTFTIGDAQN-SCHGTMVQ	
Triphysaria_sp_TC11756	(264)	-RETAISQIIR-TQPNKMYSLSFTIGDAKN-GCHGTMTE	
G.max_Glyma09g36220.1	(287)	-RESVIAQIIR-TVPNKIYNMKFTIGDARN-GCHGSMME	
M.truncatula_AC174305_19.4	(264)	-RESAIAQIIR-TVTNKVYNLKTFTVGDGRN-GCHGSMME	
G.max_Glyma15g43180.1	(264)	-RESAIAQIIR-TVTNKVYNITFSVGDKN-GCHGSMME	
P.vulgaris_TC10932	(264)	-RESAIAQVIR-TDTNKVYNITFSVGDKN-GCHGSMME	
A.lyrata_478145	(256)	-KESIAAQVIR-TSPGQTYTSLFVVGDKN-DCHGSMME	
A.thaliana_AT3G08030.1	(262)	-KESIAAQVIR-TSPGQTYTSLFVVGDKN-DCHGSMME	
B.napus_TC89135	(282)	-RESAIAQVIR-TSPGQTYSLSFVGDKN-DCHGSMME	
G.hirsutum_TC161608	(282)	GKKVAIAQNPQEHPTNHNMTFIHRGPFTGCNGEMME	
T.cacao_TC4474	(282)	-RESAIAQIFR-TVPNQLYDLTFTIGDAWN-GCHGEMME	
L.serriola_TC5542	(265)	-RESAIAQIIR-TIPNKLYSLSFAIGDAKN-GCHGDMME	
C.sinensis_TC3550	(264)	-RESAIAQIIR-TVPNKEYNLTFTIGDAKN-GCHGSMRE	
P.trichocarpa_549292	(262)	-RESAIAQIIR-TIPDKVYGLTFTVGDARN-GCHGSMME	
P.trichocarpa_833200	(262)	-RESAIAQIIR-TTPNKVYSLAFTIGDARN-GCHGSMME	
V.vinifera_GSVIVT00028751001	(263)	-RESAIAQVLR-TVPNKLYNLTFSIGDGKN-GCHGDMME	
F.vesca_TA10399_57918	(265)	-RESAIAQVLR-TVPNKLYSLSFTIGDARN-GCHGSMME	
N.benthamiana_TC11106	(264)	-RESAIAQIIR-TVPKKVYAFATVGDKN-GCHGDMME	
N.tabacum_TC43533	(263)	-RESAIAQIIR-TVPKKVYAFATVGDKN-GCHGDMME	
S.lycopersicum_TC199778	(264)	-RESAIAQIIR-TVPKKVYAFATVGDKN-GCHGDMME	
Triphysaria_sp_TC1599	(263)	-RESAIAQIIR-TVPNKSYDLTFTVIGDAKN-GCHGSMME	
I.nil_TC4208	(261)	-RESAIAQIIR-TVPNKVYTTLAFTVGDKN-GCHGDMME	
Consensus	(401)	KESAISQMVR TVP K Y LTFVSGDA GC G M	

FIGURE 3 (continued)

29/77

		451	500
(301)	A. lyrata 340195	AFAGQKVM-VDYASKGKGGR--GRLVFKAVSARTRVTLSTFFYHMK	
(300)	A.thaliana AT1G80240.1	AFAGQKVM-VDYASKGKGGR--GRLVFKAVSARTRVTLSTFFYHMK	
(299)	B.napus TC65362	AFAGRGKL-VEYVSKGKGFKR--GRLVFKAVSARTRVTLSTFFYHMK	
(300)	B.napus TC77368	AFAGRGKL-VEYVSKGKGFKQ--GRLVFKAVSARTRVTLSTFFYHMK	
(299)	O.sativa LOC Os01g42520.1	AYAAASIK-VPYESKGTGGYK--RAVLEFAAIAANTRVVFQSTFYHMT	
(296)	P.virgatum TC46275	AFAARATVK-VPYESKGTGGYK--RAVLDFATAIANTRVVFQSTFYHMA	
(297)	S.bicolor Sb03g027650.1	AYAAARATVK-VPYESKGTGGYK--RAVLDFATAIANTRVVFQSTFYHMA	
(305)	Z.mays TC458850	AYAAARATVK-VPYQSRGTGGYK--RAVLDFATAIANTRVVFQSTFYHMKP	
(352)	Zea mays GRMZM2G034985 T02	AYAAARATVK-VPYQSRGTGGYK--RAVLDFATAIANTRVVFQSTFYHMKP	
(291)	T.aestivum c6000810401I410	PWPRPTPR-APP-S-----RCRTSPTAPAAATSAPCWTSSPSGTAP	
(301)	O.sativa LOC Os04g41710.1	GYAGQGCTT-VTYSSQGTGGHTR--ASLEFAAIVANTTRVVFVSSTYITKW	
(308)	O.sativa LOC Os04g41740.1	AYASRGSVK-VTYQSQGTGGYK--GLLEFTEATEKRTRVVFVSMAYTTKS	
(308)	O.sativa LOC Os04g41750.1	AYASRGSVK-VTYQSQGTGGYK--GLLEFTEATEKRTRVVFVSMAYTTKS	
(306)	S.bicolor Sb06g021280.1	AYVSHGSVK-VPYESQGRGGYK--GVLEFTEATEKRTRVVFVSMAYTMKP	
(306)	Z.mays_ZM07MC22756 Bf50088A01@22693	AYAGKSVK-AQFQSQGKGHKKR--GVLDFTATADQTRVVFVSMGYIMKP	
(306)	Zea mays GRMZM2G051898 T01	AYAARGSVK-VPYQSQGKGHTR--GVLDFAAIVADQTRVVFVSMGYIMKP	
(306)	S.bicolor Sb06g021290.1	AYAARGSVK-VPYQSQGKGHTR--GVLDFAAIVADQTRVVFVSMGYIMKP	
(301)	O.sativa LOC Os04g41759.1	AYAAAGSVT-VPHEQSGRGGHKKR--GVLDFTATADQTRVAFVSMAYTMKS	
(302)	S.bicolor LOC Os04g41759.1	AYAAARATK-VPYESQGTGGHKKR--AQLEFAAVANLTRVVFQSFNYHTKP	
(300)	S.bicolor Sb06g021320.1	VSAGRAAKLVNVPYESHGTGGYK--AELEFVATDNLTRVVFQSANHYTKS	
(305)	S.bicolor Sb06g021300.1	AYAARGRQP-VSYESHGMGCVR--SELEFAAVANLTRVVFQSMGHYMKP	
(302)	Z.mays TC468877	AYAARGRQP-VSYESRGTTGGV--NELEFAAVANLTRVVFQSMGHYMKP	
(305)	S.bicolor Sb06g021310.1	AYAAARATK-ATYESRGTTGGSIKRAAVVEFAAIAANLTRVVFQSYNHHMKP	
(389)	O.sativa LOC Os04g41770.1	AYAAARATK-VSYESRGTTGGER--AELVFAAVANLTRVVFHSSNHHMK	
(325)	S.bicolor Sb06g021270.1	AYAGSETT-VDIDFHGAGGSKR--GKIEFRATANPTRVVFSLGYHTKC	
(322)	Zea mays GRMZM2G051912 T01	AYAGTAETT-VDVDFRDAGGSKR--GKIEFRATANPTRVVFSLGYHTKS	
(333)	Zea mays GRMZM2G405071 T01	AYAGTSETT-VDIDFRGAGGSKR--GKIEFRATANPTRVVFSLGYHTKS	
(303)	O.sativa LOC Os02g11040.1	AYAGRESTR-VAHESAGRG--AAKRAVLPFRAAAARTRVFFSSFYSTRS	
(226)	O.sativa LOC Os02g11040.1	AYAGRESTR-VAHESAGRG--AAKRAVLPFRAAAARTRVFFSSFYSTRS	
(301)	P.virgatum TC392	AYAGRESTK-VAYESAGKG--GVKRAVLPFRAAAARTRVFFSSFYSTRS	
(293)	S.bicolor Sb04g007160.1	AYAGRESTK-VAYESAGKG--GVKRAVLPFRAAAARTRVFFSSFYSTRS	

FIGURE 3 (continued)

Z.mays TC467870	(301)	AYAGRESTK-VAYESAGKG--GVKRAVLFPFRAASARTLVLFFSSFYSTRS
P.sitchensis TA14284 3332	(302)	AFAGRETVK-VPFESKGAGTFKA--ATLAFTAIAPRTRVTFSTFYTMRS
P.taeda TA10429 3352	(302)	AFAGRETVK-VPFESKGAGTFKA--ATLAFTAIAPRTRVTFSTFYTMRS
A.lyrata 9I2808	(284)	AFAGDQAQNFHYMAQ-ANSSFEK--AGLFTAKADRTVAFYSVYYNTRT
A.thaliana AT1G29980.1	(320)	AFAGDQAQNFHYMAQ-ANSSFEK--AGLFTAKADRTVAFYSVYYNTRT
A.lyrata 934049	(315)	AFAGDQAQNFHYMAQ-ANSSFER--SELNFTAKAERTRIAFYSIYYNTRT
C.annuum TC19321	(304)	AFAGDQAQNIHYTPN-SNSSFQN--ANLFTAKADRTRIAFYSIYYNTRS
S.tuberosum TC189725	(304)	AFAGDQAQDIHYTPN-FNSSFQN--ANLFTAKADRTRIAFYSIYYNTRS
J.hindsii x_regia TA1156 432290	(302)	AFARRPGSKHPLHA-----
G.hirsutum TC135941	(303)	AFAGDQAQNFHYTPD-SNSTFQV--ASVNFTAKAERTRIAFYSVYYNTRT
G.raimondii TC1124	(305)	AFAGDQAQNFHYTPD-SNSTFQV--ASVNFTAKAERTRIAFYSVYYNTRT
G.max GM06MC25068 saa05g06@24507	(303)	AFAGDQAQNIHYTPN-SNSTFQT--ANVNFTAKAERTRIAFYSIYYNTRS
L.japonicus TC41787	(303)	AFAGDQAQNIHYTPN-SNSTFQT--ANVNFTAKAERTRIAFYSIYYNTRS
M.domestica TC32901	(305)	AFAGDQAQNIHYTPN-SNSTFQT--ANLFTAKADRTRIAFYSVYYNTRS
P.trifoliata TA6477 37690	(302)	AFAGDQAQNVHYTPN-SNSTFQT--AGVNFTAKAERTRIAFYSVYYNTRT
P.trichocarpa DUF	(302)	AFAGDQAQIVHYTED-ANSTFHD--ANVNFTAKADRTRIAFYSVYYNTRT
P.trichocarpa MUB2	(303)	AFAGDQAQNIHYTPD-SNSTFQV--ANLFTAKADRTRIAFYSVYYNTRS
T.officinale TA3338 50225	(303)	AFAGGQAQNIHYTPD-SNSTFQV--ANLFTAKADRTRIAFYSVYYNTRS
H.vulgare TC163732	(302)	AFAGDQAENIHYAPN-GNSSFQT--VNVNFTAKAERTVAFYSVYYNTRS
T.aestivum TC292522	(304)	AFAGDQAQNFHYSPM-GNATSQA--VNVNFTARAERSRVALYSVYYNTRS
O.sativa LOC Os01g55190.1	(304)	AFAGDQAQNFHYSPM-GNATSQA--VNVNFTARAERTVAFYSVYYNTRS
P.virgatum TC7090	(309)	AFAGDQAQNFHYSPM-GNATSQA--ANVTFTAERTVAFYSVYYNTRS
S.bicolor Sb03g034960.1	(311)	AFAGDQAQNFHYSPM-GNATSQA--ANVTFTAERTVAFYSVYYNTRS
S.officinale TA720 94328	(308)	AFAGDQAQNFHYSPI-GNATSQA--ANVTFTAERTVAFYSVYYNTRS
Z.mays c62082358gm030403@7546	(307)	AFAGDQAQNFHYSPM-GNATSQA--ANVTFTAERTVAFYSVYYNTRS
Z.mays_ZM07MC26525 Bfb0211C02@26447	(307)	AFAGDQAQNFHYSPM-GNATSQA--ANVTFTAERTVAFYSVYYNTRS
Zea mays GRMZM2G324705 T01	(286)	AFAGDQAQNFHYSPM-GNATSQA--ANVTFTAERTVAFYSVYYNTRS
Z.officinale TA720 94328	(306)	AFASDQAQNFHYAPT-GNATSQA--ANVTFTAERTVAFYSVYYNTRS
P.sitchensis TA10533 3332	(310)	AFAGDQAQSFHNPPANGIYEP--NNVTFTAERTVAFYSIYYNTRV
P.taeda TA4773 3352	(304)	AFAGDQSGNVHNNPPANGIYEP--NNVTFTAERTVAFYSIYYNTRV
O.sativa LOC Os03g59300.1	(326)	AFAGNVTQK-VPFESVGNNGGFKP--ASFREFVAAGVTRVTFYSSYYHTKV
S.bicolor Sb01g004270.1	(319)	AFAGNVTQK-VPFESAGKGGFKA--AAFREFVAAGARTLTFYSSYYHTKV

FIGURE 3 (continued)

Z.mays_c57213778gm030403@389	(317)	AFAGNVTQK-VPFDSAGKGGFKT--AAFRFVAGGVRTRLTFYSSYYHTKV
Z.mays_TC506958	(323)	AFAGNVTQK-VPFDSAGKGGFKT--AAFRFVAGGVRTRLTFYSSYYHTKV
P.trichocarpa_735032	(303)	FAAAKETVK-APYVSQKGGSKT--ASLRFQAI SDRTRITFYSSYYHNKM
V.vinifera_GSV1VT00025589001	(301)	FAAAKETVK-VPHASQKGEFKA--ASLKFKALSARTRITFYSAYYHTKL
S.lycopersicum_TC198542	(302)	AFAGKASTQ-VTFVSNKGWSKT--ATFKFQADSIRTTIAFNYPYHTKI
Triphysaria_sp_TC11756	(301)	AFAGRNSLK-VPFTSHKGWFKK--ASLGFQAVSARTRITFYSPFYHTRL
G.max_Glyma09g36220.1	(324)	FAAAKDTLK-VPFKSEKGEFKT--VSFKFRAIENRTRITFYSSFYHTRI
M.truncatula_AC174305_19.4	(301)	FAAAARETLK-VPFKSVKGIFKT--ANFNFKAVSNRTRITFYSSFYHTKI
G.max_Glyma15g43180.1	(301)	FAAAKDTFK-VPFKSEKGTSTKT--VSFKFKAIAPRTRLTFYSSFYHTRI
P.vulgaris_TC10932	(301)	FAAAKDTFK-VPFKSEKGTFT--VSFKFKAI AKRTRLTFYSSFYHTRS
A.lyrata_478145	(293)	FAAAKDTLK-VPHTSVGGGHVKT--ASF KFKAVEARTRITFFSGFYHTKK
A.thaliana_AT3G08030.1	(299)	FAAARDTLK-VPHTSVGGGHVKT--ASF KFKAVEARTRITFFSGFYHTKK
B.napus_TC89135	(299)	FAAARDTLK-VPHTSVGGGHVKT--ASF KFKAVEARTRITFFSGFYHTKK
G.hirsutum_TC161608	(322)	FAAAKHTVK-VPFTSRKGGEFKA--ASLMFKADTARTVTFYSSYYHTRS
T.cacao_TC4474	(319)	AFADKNTVK-VPFTSRKGGEFKT--ASLKFKAVTARTTRITFFSFYHTRI
L.serriola_TC5542	(302)	FAAAKDTLK-APFKSEKGGKSKV--VSMKFKALSARTRVTFYSSYYHTRA
C.sinensis_TC3550	(301)	AFAGKDTVA-VPFESKGGGFKS--ASLKFKAI AARTTRITFFSTYYHTKI
P.trichocarpa_549292	(299)	FAAAKDTFK-VPFESKGGGEFKT--VSFKFKAVTARTTRITFYSSYYHTRI
P.trichocarpa_833200	(299)	FAAAGDTFK-VPFESKGGGESKS--ASF KFKAISARTRITFYSSYYHTRI
V.vinifera_GSV1VT00028751001	(300)	FAAAKDSFK-APFKSRKGGEFKT--VSFKFKAI AARTTRITFFSSFYHTRI
F.vesca_TA10399_57918	(302)	AFAGKDTLK-VPFNSQKGGFKN--ASF KFKATSFRTRITFFSSFYHTRT
N.benthamiana_TC11106	(301)	FAAAKETFK-VPFKSQKGGSKT--VSFKFTAIADRTTRITFYSSYYHTKI
N.tabacum_TC43533	(300)	FAAAKETFK-VPFKSQKGGSKT--VSFKFTAIADRTTRITFYSSYYHTKI
S.lycopersicum_TC199778	(301)	FAAAKETFK-VPFKSQKGGSKT--VSFKFTAIADRTTRITFYSSFYHTKI
Triphysaria_sp_TC1599	(300)	FAAAKEAMK-APFQSRKGGEFKA--YSEKFTAVSLRTRLTFYSSFYHTKI
I.nil_TC4208	(298)	AFAGKSAFK-VPFERSVGKGHSKS--VSFRFTAISIRTRLTFSSFYHTRI
Consensus	(451)	AFAGR T K V F S G G FK A L F A A RTRV FYS YYHTRS

FIGURE 3 (continued)

501

(348) DHSGSLCGPVIDDVRLVAVRKLRG-----

(347) DHSGSLCGPVIDDVRLVAVGKLRG-----

(346) DHSGSLCGPVDDVRLVAVKTLRG-----

(347) DHSGSLCGPVIDDVRLVAVRTLGG-----

(346) D--GSLCGPIDDASLVGLRKK--TAGRRLLL--

(343) D--GTLCGPLVDDASLVGLRKKPAAGSRLLL--

(344) D--GTLCGPLVDDASVVGLRKKPAAAGRRLLQFL--

(352) D--GTLCGPLVDDASLVGLRKK--KPARRLMQLL--

(399) D--GTLCGPLVDDASLVGLRKK--KPARRLMQLL--

(330) G--SSSRAPSTT-----

(348) D--GTLCGPVDDASLVCV6QQQ--PPARRLLRL--

(355) D--GTLCGPIDDASLVSVHS--HRRFLL--

(355) D--GTLCGPIDDASLVSVHS--HRRFLL--

(353) D--GTLCGPVDDASVGVHS--HRRLLL--

(353) D--GTLCGPVDDSVLVCTRN--HPARRLLL--

(353) D--GTLCGPVDDSVLVCSTRN--HPARRLLL--

(353) D--GTLCGPVDDSVLVCSTRN--HPARRLLL--

(348) D--GTLCGPLDDSVLGLHS--RAARRFLL--

(349) D--GTLCGPLVDDISLVSVRKR--AARL--

(348) D--ATLCGPIDDVRLVPVHAR--RRRM--

(352) D--GTLCGPVDDSVLSLPRH--AARRLLM--

(349) D--GTLCGPVDDSVLAVHKH--AARRLFM--

(354) D--GTLCGPVDDSVLGLRKH--AARRLFL--

(436) D--GTLCGPVDDSVLSVDKH--TVRRLLM--

(372) DMSGTLCGPVDDSVLPVLPQP--YARRLLL--

(369) DMSGTLCGPVDDSVLPVLPQP--YARRLLL--

(380) DMSGTLCGPVDDSVLPVLPQP--YARRLLL--

(351) DDMSSLCGPVIDDVAVSVRARRPAAKRG-----

(274) DDMSSLCGPVIDDVAVSVRARRPAAKRG-----

(348) DDLSSLCGPVDDVAVSVPAKRG-----

(340) DDLSSLCGPVDDVAVSVRTTKRG-----

550

A.lyrata 340195

A.thaliana AT1G80240.1

B.napus TC65362

B.napus TC77368

O.sativa LOC Os01g42520.1

P_virgatum TC46275

S.bicolor Sb03g027650.1

Z_mays TC458850

Zea_mays GRMZM2G034985.T02

T.aestivum c60008104@11410

O.sativa LOC Os04g41710.1

O.sativa LOC Os04g41740.1

O.sativa LOC Os04g41750.1

S.bicolor Sb06g021280.1

Z.mays_ZM07MC22756 BF50088A01@22693

Zea_mays GRMZM2G051898.T01

S.bicolor Sb06g021290.1

O.sativa LOC Os04g41759.1

S.bicolor Sb06g021320.1

S.bicolor Sb06g021300.1

Z_mays TC468877

S.bicolor Sb06g021310.1

O.sativa LOC Os04g41770.1

S.bicolor Sb06g021270.1

Zea_mays GRMZM2G051912.T01

Zea_mays GRMZM2G405071.T01

O.sativa LOC Os02g11040.1

Os_DUF642.2

P_virgatum TC392

S.bicolor Sb04g007160.1

FIGURE 3 (continued)

Z.mays TC467870	(348)	DDLSSLCGPVDDVAVVSVRTKRG
P.sitchensis TA14284 3332	(349)	DDFTLCGPVIDQVKLFSLR
P.taeda TA10429 3352	(349)	DDFTLCGPVIDQVKLFVSR
A.lyrata 9I2808	(331)	DDMSSLCGPVIDDVVWFS
A.thaliana AT1G29980.1	(367)	DDMSSLCGPVIDDVVWFS
A.lyrata 934049	(362)	DDMSSLCGPVIDDVVWFS
C.annuum TC19321	(351)	DDMSSLCGPVDDVAVVWFS
S.tuberosum TC189725	(351)	DDMSSLCGPVDDVAVVWFS
J.hindsii x regia TA1156 432290	(316)	DDMSSLCGPVDDVAVVWFS
G.hirsutum TC135941	(350)	DDMSSLCGPVDDVAVVWFS
G.raimondii TC1124	(352)	DDMSSLCGPVDDVAVVWFS
G.max GM06MC25068 saa05g06@24507	(350)	DDMSSLCGPVDDVAVVWFS
L.japonicus TC41787	(352)	DDMSSLCGPVDDVAVVWFS
M.domestica TC32901	(349)	DDMSSLCGPVDDVAVVWFS
P.trifoliata TA6477 37690	(349)	DDMSSLCGPVDDVAVVWFS
P.trichocarpa DUF	(349)	DDMSSLCGPVDDVAVVWFS
P.trichocarpa MUB2	(350)	DDMSSLCGPVDDVAVVWFS
T.officinale TA3338 50225	(349)	DDMSSLCGPVDDVAVVWFS
H.vulgare TC163732	(351)	DDHSSLCGPVIDDVAVVWFS
T.aestivum TC292522	(351)	DDHSSLCGPVIDDVAVVWFS
O.sativa LOC Os01g55190.1	(356)	DDHSSLCGPVIDDVAVVWFS
P.virgatum TC7090	(358)	DDHSSLCGPVIDDVAVVWFS
S.bicolor Sb03g034960.1	(355)	DDHSSLCGPVIDDVAVVWFS
S.officinale TC76182	(354)	DDHSSLCGPVIDDVAVVWFS
Z.mays c62082358gm030403@7546	(354)	DDHSSLCGPVIDDVAVVWFS
Z.mays_ZM07MC26525 Bfb0211C02@26447	(354)	DDHSSLCGPVIDDVAVVWFS
Zea mays GRMZM2G324705 T01	(333)	DDHSSLCGPVIDDVAVVWFS
Z.officinale TA720 94328	(353)	DDHSSLCGPVIDDVAVVWFS
P.sitchensis TA10533 3332	(358)	DDHSSLCGPVIDDVAVVWFS
P.taeda TA4773 3352	(352)	DDHSSLCGPVIDDVAVVWFS
O.sativa LOC Os03g59300.1	(373)	DDHSSLCGPVIDDVAVVWFS
S.bicolor Sb01g004270.1	(366)	DDHSSLCGPVIDDVAVVWFS

FIGURE 3 (continued)

Z.mays_c57213778gm030403@389	(364)	TDGVSILCGPVLDDQVKVPVKA
Z.mays_TC506958	(370)	TDGVSILCGPVLDDQVKVPVKA
P.trichocarpa_735032	(350)	HDYGHMCGPVLDDQVKVPVKA
V.vinifera_GSVIVT00025589001	(348)	HDYGHMCGPVLDDQVKVPVKA
S.lycopersicum_TC198542	(349)	HDYGHMCGPVLDDQVKVPVKA
Triphysaria_sp_TC11756	(348)	YDFGHMCGPVLDDQVKVPVKA
G.max_Glyma09g36220.1	(371)	HDYGSILCGPVLDDQVKVPVKA
M.truncatula_AC174305_19.4	(348)	DDYGSILCGPVLDDQVKVPVKA
G.max_Glyma15g43180.1	(348)	DDYGSILCGPVLDDQVKVPVKA
P.vulgaris_TC10932	(348)	DDYGSILCGPVLDDQVKVPVKA
A.lyrata_478145	(340)	TDIGSILCGPVLDDQVKVPVKA
A.thaliana_AT3G08030.1	(346)	DDYGSILCGPVLDDQVKVPVKA
B.napus_TC89135	(346)	DDYGSILCGPVLDDQVKVPVKA
G.hirsutum_TC161608	(369)	DDYGSILCGPVLDDQVKVPVKA
T.cacao_TC4474	(366)	DDYGSILCGPVLDDQVKVPVKA
L.serriola_TC5542	(349)	DDYGSILCGPVLDDQVKVPVKA
C.sinensis_TC3550	(348)	DDYGSILCGPVLDDQVKVPVKA
P.trichocarpa_549292	(346)	DDYGSILCGPVLDDQVKVPVKA
P.trichocarpa_833200	(346)	DDYGSILCGPVLDDQVKVPVKA
V.vinifera_GSVIVT00028751001	(347)	DDYGSILCGPVLDDQVKVPVKA
F.vesca_TA10399_57918	(349)	DDYGSILCGPVLDDQVKVPVKA
N.benthamiana_TC11106	(347)	DDYGSILCGPVLDDQVKVPVKA
N.tabacum_TC43533	(348)	DDYGSILCGPVLDDQVKVPVKA
S.lycopersicum_TC199778	(348)	DDYGSILCGPVLDDQVKVPVKA
Triphysaria_sp_TC1599	(347)	DDYGSILCGPVLDDQVKVPVKA
I.nil_TC4208	(345)	DDYGSILCGPVLDDQVKVPVKA
Consensus	(501)	DDYGSILCGPVLDDQVKVPVKA

FIGURE 3 (continued)

35/77

551

A.lyrata 340195
 A.thaliana AT1G80240.1
 B.napus TC65362
 B.napus TC77368
 O.sativa LOC_Os01g42520.1
 P.virgatum TC46275
 S.bicolor Sb03g027650.1
 Z.mays TC458850
 Zea_mays GRMZM2G034985.T02
 T.aestivum c600081041I410
 O.sativa LOC_Os04g41710.1
 O.sativa LOC_Os04g41740.1
 OsDUF642.3
 O.sativa LOC_Os04g41750.1
 S.bicolor Sb06g021280.1
 Z.mays_ZM07MC22756 BF50088A01022693
 Zea_mays GRMZM2G051898.T01
 S.bicolor Sb06g021290.1
 O.sativa LOC_Os04g41759.1
 S.bicolor Sb06g021320.1
 S.bicolor Sb06g021300.1
 Z.mays TC468877
 S.bicolor Sb06g021310.1
 O.sativa LOC_Os04g41770.1
 S.bicolor Sb06g021270.1
 Zea_mays GRMZM2G051912.T01
 Zea_mays GRMZM2G405071.T01
 O.sativa LOC_Os02g11040.1
 OsDUF642.2
 P.virgatum TC392
 S.bicolor Sb04g007160.1

(372)
 (371)
 (370)
 (371)
 (374)
 (373)
 (376)
 (382)
 (429)
 (340)
 (378)
 (380)
 (380)
 (375)
 (373)
 (372)
 (379)
 (376)
 (381)
 (463)
 (401)
 (398)
 (409)
 (380)
 (303)
 (372)
 (365)

FIGURE 3 (continued)

Z.mays TC467870	(372)	-
P.sitchensis TA14284 3332	(369)	-
P.taeda TA10429 3352	(369)	-
A.lyrata 9I2808	(374)	-
A.thaliana AT1G29980.1	(408)	-
A.lyrata 934049	(401)	-
C.annuum TC19321	(392)	-
S.tuberosum TC189725	(392)	-
J.hindsii_x_regia TA1156 432290	(316)	-
G.hirsutum TC135941	(394)	-
G.raimondii TC1124	(396)	-
G.max GM06MC25068_saa05g06024507	(394)	-
G.max TC286277	(394)	-
L.japonicus TC41787	(396)	-
M.domestica TC32901	(361)	-
P.trifoliata TA6477 37690	(394)	-
P. trichocarpa DUF	(394)	-
P. trichocarpa MUB2	(394)	-
T.officinale TA3338 50225	(392)	-
H.vulgare TC163732	(395)	-
T.aestivum TC292522	(395)	-
O.sativa LOC Os01g55190.1	(400)	-
P.virgatum TC7090	(402)	-
S.bicolor Sb03g034960.1	(399)	-
S.officinatum TC76182	(398)	-
Z.mays c62082358gm03040307546	(398)	-
Z.mays_ZM07MC26525 BFb0211C02026447	(398)	-
Zea_mays GRMZM2G324705 T01	(377)	-
Z.officinale TA720 94328	(399)	-
P.sitchensis TA10533 3332	(408)	L
P.taeda TA4773 3352	(402)	L
O.sativa LOC Os03g59300.1	(394)	-
S.bicolor Sb01g004270.1	(387)	-

FIGURE 3 (continued)

37/77

Z.mays_c57213778gm030403@389	(385)	-
Z.mays_TC506958	(391)	-
P.trichocarpa_735032	(370)	-
V.vinifera_GSVIVT0025589001	(368)	-
S.lycopersicum_TC198542	(370)	-
Triphysaria_sp_TC11756	(368)	-
G.max_Glyma09g36220.1	(391)	-
M.truncatula_AC174305_19.4	(368)	-
G.max_Glyma15g43180.1	(368)	-
P.vulgaris_TC10932	(368)	-
A.lyrata_478145	(360)	-
A.thaliana_AT3G08030.1	(366)	-
B.napus_TC89135	(366)	-
G.hirsutum_TC161608	(381)	-
T.cacao_TC4474	(386)	-
I.serriola_TC5542	(370)	-
C.sinensis_TC3550	(369)	-
P.trichocarpa_549292	(366)	-
P.trichocarpa_833200	(366)	-
V.vinifera_GSVIVT0028751001	(369)	-
F.vesca_TA10399_57918	(369)	-
N.benthamiana_TC11106	(368)	-
N.tabacum_TC43533	(367)	-
S.lycopersicum_TC199778	(368)	-
Triphysaria_sp_TC1599	(368)	-
I.nil_TC4208	(365)	-
Consensus	(551)	-

FIGURE 3 (continued)

O. sativa LOC_Os01g42520.1	1	-----	50	-----MRRVTVLV
P_virgatum TC46275	(1)	-----		-----MRRVA
S.bicolor Sb03g027650.1	(1)	-----		-----MRRVAL
Z_mays TC458850	(1)	-----		-----MRRVALLL
Zea mays GRMZM2G034985.T02	(1)	-----		-----MRLTISARPVSLSPPTQQQTATS
T.aestivum c60008104@11410	(1)	-----		-----MRCAY
O. sativa LOC_Os02g11040.1	(1)	-----		-----MAAGSASLLLM
Os_DUF642.2	(1)	-----		-----
P_virgatum TC392	(1)	-----		-----MGRSSMAGPA
S.bicolor Sb04g007160.1	(1)	-----		-----MAVPV
Z_mays TC467870	(1)	-----		-----MGRIALAVPV
O. sativa LOC_Os04g41710.1	(1)	-----		-----MAFYNGAS
S.bicolor Sb06g021270.1	(1)	-----		-----MAIGKRSVHG--AALV
Zea mays GRMZM2G051912.T01	(1)	-----		-----MAIDERSVHGGAALVL
Zea mays GRMZM2G405071.T01	(1)	-----		-----MVIDRRSVQR-AALVV
O. sativa LOC_Os04g41740.1	(1)	-----		-----MRRKTEMIQTTPCVALL
Os_DUF642.3	(1)	-----		-----MRRKTEMIQTTPCVALL
O. sativa LOC_Os04g41750.1	(1)	-----		-----MRRKTEMRSTCCVLL
S.bicolor Sb06g021280.1	(1)	-----		-----MGRGETVPNARRAALF
Z.mays_ZM07MC22756.BFB0088A01@22693	(1)	-----		-----MGRGSDAVPDARRAALL
Zea mays GRMZM2G051898.T01	(1)	-----		-----MGRGSDAVPDARRAALL
S.bicolor Sb06g021290.1	(1)	-----		-----MVQNAPHAAALF
S.bicolor Sb06g021300.1	(1)	-----		-----MAESGSSRRGVLL
Z_mays TC468877	(1)	-----		-----MEGRASSRRGVLL
S.bicolor Sb06g021310.1	(1)	-----		-----MAGTRSRCPVLL
O. sativa LOC_Os04g41759.1	(1)	-----		-----MTGSARSVVAL
O. sativa LOC_Os04g41770.1	(1)	-----		-----MAQPCPRPQARTPNPLCAATPWPRAGRANWRGIFCSRWSPSCRDART
S.bicolor Sb06g021320.1	(1)	-----		-----MVARRSVAF
Consensus	(1)	-----		AVLL

FIGURE 4

O.sativa LOC_Os01g42520.1	(9)	LLAACAAARAAAAVVTDG	51	100
P.virgatum TC46275	(6)	LLLVCAARAAAAVVTDG		
S.bicolor Sb03g027650.1	(7)	LLLVCAARAAAAVVTDG		
Z.mays TC458850	(9)	VAAVCALAAPAAAAVVTDG		
Zea mays GRMZM2G034985 T02	(28)	GAALPCNVHPASVSSKDCNICSSHRPTTRPITIHTDTKCNANGAG		
T.aestivum c60008104@11410	(6)	LLLVFAARAAAAVVTDG		
O.sativa LOC_Os02g11040.1	(13)	VVLCAAASGTALAAITDG		
O.sativa LOC_Os_DUF642.2	(1)	-----		
P.virgatum TC392	(11)	MLLLICTVCLALGITDG		
S.bicolor Sb04g007160.1	(6)	MLLLCATCRLALGITDG		
Z.mays TC467870	(11)	MLLLCTTCRLTLGITDG		
O.sativa LOC_Os04g41710.1	(9)	VVALLLVSTAARAAGDGLLN		
S.bicolor Sb06g021270.1	(15)	LLLVCVAAARAVSAVGDGPLN		
Zea mays GRMZM2G051912 T01	(17)	RLLVCVAAARAVSAVADGPLN		
Zea mays GRMZM2G405071 T01	(16)	LPLVLFVAARAVSAVGDGPLQN		
O.sativa LOC_Os04g41740.1	(18)	LLVGVAFAARSASAITDGLLPN		
O.sativa LOC_Os_DUF642.3	(18)	LLVGVAFAARSASAITDGLLPN		
O.sativa LOC_Os04g41750.1	(18)	LLL--SVAARSASAIMDGLLPN		
S.bicolor Sb06g021280.1	(18)	LIV--CLAAARASAIQDGLLPN		
Z.mays_ZM07MC22756 BF50088A01022693	(18)	LIV--CLVARAASAIQDGLLPN		
Zea mays GRMZM2G051898 T01	(18)	LIV--CLAAARASAIQDGLLPN		
S.bicolor Sb06g021290.1	(12)	LLV--ALAAARASAIQDGLLPN		
S.bicolor Sb06g021300.1	(14)	FLLVGAAAGVASAIPDGLLPN		
Z.mays TC468877	(14)	FLLVG----VASANPDGLLPN		
S.bicolor Sb06g021310.1	(14)	LFLVGAAARVASGVVTDGGLLPN		
O.sativa LOC_Os04g41759.1	(12)	LFLLVGSAARADSAVTDGGLLPN		
O.sativa LOC_Os04g41770.1	(51)	IQPARLAASRTHPPCTDHTSCNKQQLLYISSEVRVLDRTVTRAHDHGLGEV		
S.bicolor Sb06g021320.1	(10)	VILLIAAARAVFAITEGGLPN		
Consensus	(51)	LLLL AARAASAITDG L N		

FIGURE 4 (continued)

O.sativa LOC Os01g42520.1	(27)	LLP	GNFEEGPPKSDLV	NGTVVRGAN	-----	DUF642 DOMAIN	150
P_virgatum TC46275	(24)	LLP	GNFEEGPPKSALV	NGTVVKGAN	-----		AIPRWETSG
S.bicolor Sb03g027650.1	(25)	LLP	GNFEDGPPKSALV	NGTVVSGAN	-----		AIPRWETSG
Z_mays TC458850	(27)	LLP	GNFEDGPPKSALV	NGTVVSGAH	-----		AIPRWETSG
Zea_mays GRMZM2G034985.T02	(74)	LLP	GNFEDGPPKSALV	NGTVVSGAH	-----		AIPRWETSG
T.aestivum c600081041I410	(24)	LLP	GNFELGPAKSELV	NGTVVKGGK	-----		AIPKWETSG
O.sativa LOC Os02g11040.1	(31)	LLA	GNFERGP-APSQ	LRGTRVVGAS	-----		AIPSWRTSG
Os_DUF642.2	(1)				-----		
P_virgatum TC392	(29)	LLP	GNFERGP-VPSQ	LRGTRVLGSS	-----		AIPSWQTSQ
S.bicolor Sb04g007160.1	(24)	LLP	GNFERGP-APSQ	LRGTRVMGSS	-----		AIPAWQTSQ
Z_mays TC467870	(29)	LLP	GNFERSP-LPSQ	LRGTRVVGAS	-----		AIPSWQTSQ
O.sativa LOC Os04g41710.1	(31)	---	GNFEYQP-SKSQ	MNGTRVMAEY	-----		AIPYWKITG
S.bicolor Sb06g021270.1	(37)	---	GNFEDSP-NPSQ	MSGSVVTGEH	-----		AISYWRITSG
Zea_mays GRMZM2G051912.T01	(39)	---	GNFEDPP-NQSQ	MSGSVVTGEH	-----		AVSYWKITSG
Zea_mays GRMZM2G405071.T01	(38)	---	GNFEDSP-DQSQ	MSGSVVAGEH	-----		AISGWKTSQ
O.sativa LOC Os04g41740.1	(40)	---	GNFEEAP-DKSQ	MNGTRVTGRY	-----		AIPQWEISG
Os_DUF642.3	(40)	---	GNFEEAP-DESQ	MNGTRVTGRY	-----		AIPQWEISG
O.sativa LOC Os04g41750.1	(38)	---	GNFEEAP-DKSQ	LNTRVIGRY	-----		AIPQWEISG
S.bicolor Sb06g021280.1	(38)	---	GNFEEAP-PKSQ	LKGSRVVMGRN	-----		AIPHWEISG
Z_mays_ZM07MC22756.BFB0088A01@22693	(38)	---	GNFEEAP-PKSQ	LNTRVVMGRN	-----		AIPHWEISG
Zea_mays GRMZM2G051898.T01	(38)	---	GNFEEAP-PKSQ	LNTRVVMGRN	-----		AIPHWEISG
S.bicolor Sb06g021290.1	(32)	---	GNFEEAP-AKSEL	LNTRVKGRD	-----		AIPHWEISG
S.bicolor Sb06g021300.1	(36)	---	GNFEQGP-DKSQ	MNGSRVIGQN	-----		SIPCWETSG
Z_mays TC468877	(31)	---	GNFEQGP-DKYQ	MNGTWVIGRK	-----		SIPCWEISG
S.bicolor Sb06g021310.1	(36)	---	GNFEFGP-AKSD	LNTRVMGQN	-----		SIPNWEISG
O.sativa LOC Os04g41759.1	(34)	---	GNFEDGP-DKSQ	LNTRVVTGRY	-----		AILNWEISG
O.sativa LOC Os04g41770.1	(101)	HAL	PGSCGGA-RWRG	CTGGLLCHRWFPSTKRQLRARAGQVPAEPNWEISG			
S.bicolor Sb06g021320.1	(32)	---	GHFEQGP-AKSEL	MNGTRVLGRY	-----		SIPHWEISG
Consensus	(101)		GNFEEGP	SQLNTRV G	-----		AIP WETSG

FIGURE 4 (continued)

O.sativa LOC_Os01g42520.1	(62)	FVEYIESGKH-QGDMMLVVPQGAHAVRLGNEASIRQRLA-VTRGAYYAVT	151	200
P_virgatum TC46275	(59)	FVEYIESGKH-QGDMMLVVPQGAHAVRLGNEASIRQRLA-VTRGAYYAIT		
S.bicolor Sb03g027650.1	(60)	FVEYIESGKH-QGDMMLVVPQGAHAVRLGNEASIRQRLA-VARGAYYAIT		
Z.mays TC458850	(62)	FVEYIESGKH-QGDMMLVVPQGAHAVRLGNEASIRQRLA-VARGAYYAIT		
Zea_mays GRMZM2G034985.T02	(109)	FVEYIESGKH-QGDMMLVVPQGAHAVRLGNEASIRQRLA-VARGAYYAIT		
T.aestivum c6000810411410	(59)	FVEYIESGKH-QGDMMLVVPQGAHAVRLGNDASILQRIIP-VARGSYAIT		
O.sativa LOC_Os02g11040.1	(65)	FVEYIPSGRK-QGDMVLVPEGSHAVRLGNEASIRQRLAGAARGARYALT		
Os_DUF642.2	(1)	-----MVLVPEGSHAVRLGNEASIRQRLAGAARGARYALT		
P_virgatum TC392	(63)	FVEYIPSGK-KQGDMMLVPEGAYAVRLGNEASIRQRLRGAARGARYSLT		
S.bicolor Sb04g007160.1	(58)	FVEYIPSGRK-QGDMVLVPEGAYAVRLGNEASIRQRLHGAARGARYSLT		
Z.mays TC467870	(63)	FVEYISSGQK-QGDMVLVPEGAYAVRLGNEASIRQRLRGAARGARYSLT		
O.sativa LOC_Os04g41710.1	(61)	FVEYISSGQK-QGDMMLTVPEGAHAVRLGNEASIEQQIS-VTRGMYYSIT		
S.bicolor Sb06g021270.1	(67)	HIEFICSGQT-QGDMVLTVPDGAHALRLNGASIEQRIS-LTPGSYYSLT		
Zea_mays GRMZM2G051912.T01	(69)	HIEFICSGQK-QGDMVLTVPDGSHALRLGSGASIQQQIS-LTKGLYYSLT		
Zea_mays GRMZM2G405071.T01	(68)	HIELVCSGQK-QGDMVLTVPDGAALRLGAGASIEQQISGLTQGSYYSLT		
O.sativa LOC_Os04g41740.1	(70)	FVEYIGSGQK-QGDMMLPVPEGAYAVRLGNEASIQQRILT-LTRGMHYSIT		
Os_DUF642.3	(70)	FVEYIGSGQK-QGDMMLPVPEGAYAVRLGNEASIQQRILT-LTRGMHYSII		
O.sativa LOC_Os04g41750.1	(68)	FVEYIGSGQM-QGDMMLPVPEGAYAVRLGNEASIQQRILT-LTRGMHYSVT		
S.bicolor Sb06g021280.1	(68)	FVEYIGSGQK-QGDMMLPVPEGAYAVRLGNEASIQQRILA-LTQGAHYSIT		
Z.mays_ZM07MC22756_BFB0088A01022693	(68)	FVEYIGSGQK-QGDMMLPVPEGAYAVRLGNEASIQQRILA-VTPGARYSIT		
Zea_mays GRMZM2G051898.T01	(68)	FVEYIGSGQK-QGDMMLPVPEGAYAVRLGNEASIQQRILA-VTPGARYSIT		
S.bicolor Sb06g021290.1	(62)	FVEYIGPGQK-QGDMMLPVPEGAYAVRLGNEASIQQRILA-VTQGARYSVT		
S.bicolor Sb06g021300.1	(66)	FVEYIEPGQQ-QNDMVLVPEGARAVRLGNDATIRQQLPGLTRSMSYSIT		
Z.mays TC468877	(61)	FVEYIVSGQQQDNGMVVAPEGARAVRLGNDASIRQRLTGMTRGTSYSVT		
S.bicolor Sb06g021310.1	(66)	FVEYIGSGQQ-QDDMILPVPEGAYAVRLGNDATIRQKLS-VNRKTTYISIT		
O.sativa LOC_Os04g41759.1	(64)	FVEYIESGHR-EQDMILAVPEGARAVRLGNDATIRQRLS-VTRRAYYSIT		
O.sativa LOC_Os04g41770.1	(150)	FVEYIGSGHK-EQDMILAVPEGAYAVRLGNDATIRQRIS-VTRHMYYSVT		
S.bicolor Sb06g021320.1	(62)	FVEYIESGQK-QDDMMLQVPEGARAVRLGNDATIQQOLA-VTRHTYYSIT		
Consensus	(151)	FVEYI SGQK QGDMML VPEGAYAVRLGNEASIRQRLA VTRGAYYSIT		

FIGURE 4 (continued)

O.sativa LOC Os01g42520.1	(110)	FSAARTCAQAEQ--LNVSVP-----EWGLPMQTIYGSNGWDSYA	250
P_virgatum TC46275	(107)	FSAARTCAQAEQ--LNISVSP-----EWGLPMQTIYGSNGWDSYA	
S.bicolor Sb03g027650.1	(108)	FSAARTCAQAEA--LNVSVP-----EWGLPMQTIYGSNGWDSYA	
Z_mays TC458850	(110)	FSAARTCAQAEQR LNVSVP-----EWGLPMQTIYGSNGWDSYA	
Zea_mays GRMZM2G034985 T02	(157)	FSAARTCAQAEQR LNVSVP-----EWGLPMQTIYGSNGWDSYA	
T.aestivum c6000810411410	(107)	FSAARTCAQAEQ--LNVSVP-----EWGLPMQTIYGSNGWDSYA	
O.sativa LOC Os02g11040.1	(114)	FSAARTCAQAEQ--LNVSASG-----QWAVLPMQTIYSSNGWDSYA	
Os_DUF642.2	(37)	FSAARTCAQAEQ--LNVSASG-----QWAVLPMQTIYSSNGWDSYA	
P_virgatum TC392	(112)	FSAARTCAQAEQ--LNVSASG-----QGLLAMQTIYSSNGWDSYA	
S.bicolor Sb04g007160.1	(107)	FSAARTCAQAEQ--LNVSASG-----QGLLAMQTIYSSNGWDSYA	
Z_mays TC467870	(112)	FSAARTCAQAEQ--LNVSASG-----QGLLAMQTIYSSNGWDSYA	
O.sativa LOC Os04g41710.1	(109)	FSAARTCAQSEK--LNVSAPG-----PESGELPIQTVYTSSGWDYS	
S.bicolor Sb06g021270.1	(115)	FSAARTCAQDEV LNVTAA PVS-----SGAAQTGELPIQTVYTSSGWDYS	
Zea_mays GRMZM2G051912 T01	(117)	FSAARTCSQNEMLNVTAA PVS-----SPAQTGELPTQTVYTSSGWDYS	
Zea_mays GRMZM2G405071 T01	(117)	FAASRTCSQDETL SVTAA PAPS GAGS PAPAQTGELPMQTVYTSSGWDYS	
O.sativa LOC Os04g41740.1	(118)	FSAARTCAQSEL--LNITITP-----ESGEIPIQTVYTSSGWDYS	
Os_DUF642.3	(118)	FSAARTCAQSEL--LNITITP-----ESGEIPIQTVYTSSGWDYS	
O.sativa LOC Os04g41750.1	(116)	FSAARTCAQSEL--LNITVTP-----EIGEVPIQTVYTSSGWDYS	
S.bicolor Sb06g021280.1	(116)	FSAARTCAQAEQ--LNVTAP-----ESDILPIQTVYTSSGWDYS	
Z_mays_ZM07MC22756 BFB0088A01022693	(116)	FSAARTCAQAEQ--LNVTAP-----ESDILPIQTVYTSSGWDYS	
Zea_mays GRMZM2G051898 T01	(116)	FSAARTCAQAEQ--LNVTAP-----ESDILPIQTVYTSSGWDYS	
S.bicolor Sb06g021290.1	(110)	FSAARTCAQAEQ--LNVTAT-----DSDVLPQTVYTSSGWDYS	
S.bicolor Sb06g021300.1	(115)	FSAARTCAQAEQ--LNVSVA-----ESGVL PVQTVYTSSGWDYS	
Z_mays TC468877	(111)	FSAARTCAQAEQ--LNVSVA-----ESGVL PVQTVYTSSGWDYS	
S.bicolor Sb06g021310.1	(114)	FSAARTCAQAEQ--LNVSVA-----ESGVL PVQTVYTSSGWDYS	
O.sativa LOC Os04g41759.1	(112)	FSAARTCAQKEK--LNMSVTP-----EFGVLPQTVYTSSGWDYS	
O.sativa LOC Os04g41770.1	(198)	FSAARTCAQAEK--LNVSVTL-----EFGVLPQTVYTSTGWDYS	
S.bicolor Sb06g021320.1	(110)	FSAARTCAQAEK--LNVSVTP-----ESGVL PVQTVYTSSGWDYS	
Consensus	(201)	FSAARTCAQAE LNVSVP-----ESGVL PVQTVYTSSGWDYS	

FIGURE 4 (continued)

		251			300
O.sativa LOC_Os01g42520.1	(149)	WAFKAK--MDEVALVIHNP	GVVEEDPACG	PLIDGVAIRALYPPTLAK----	
P_virgatum TC46275	(146)	WAFKAK--LDAVTLVIHNP	GVVEEDPACG	PLIDGVAIRALYPPTLAK----	
S.bicolor Sb03g027650.1	(147)	WAFKAK--LSTVTLVIHNP	GVVEEDPACG	PLIDGVAIRALYPPTLAK----	
Z.mays TC458850	(151)	WAFKAK--MDAVTLVIHNP	GVVEEDPACG	PLIDGVAIRALYPPALARRAGG	
Zea mays GRMZM2G034985 T02	(198)	WAFKAK--MDAVTLVIHNP	GVVEEDPACG	PLIDGVAIRALYPPALARRAGG	
T.aestivum c60008104011410	(146)	WAFKAK--LDVVELVIHNP	GVVEEDPACG	PLIDAVAIRTLYPPTLSK----	
O.sativa LOC_Os02g11040.1	(153)	WAWDAA--ADAFDVVIHNP	GVTEDPACG	PLIDSVAIRTLNPPRRTN----	
Os DUF642.2	(76)	WAWDAA--ADAFDVVIHNP	GVTEDPACG	PLIDSVAIRTLNPPRRTN----	
P.virgatum TC392	(151)	WAWVAD--AGEVDVVIHNP	GVTEDPACG	PLIDSVAIKTLNPPRRTN----	
S.bicolor Sb04g007160.1	(146)	WAWVAD--ADEVDVVIHNP	GVTEDPACG	PLIDSVAIKTLNPPRRTN----	
Z.mays TC467870	(151)	WAWVAN--ADEVDVVIHNP	GVTEDPACG	PLIDSVAIKTLNPPRRTN----	
O.sativa LOC_Os04g41710.1	(150)	WAFKAK--RGLVSLIIHNP	GVTEDPACG	PLIDSVAIKTLNPPRRTN----	
S.bicolor Sb06g021270.1	(162)	WAFKAE--AGLVSLIIITHICGEQED	PACG	PVVDFAIKTLTSQPE----ATG	
Zea mays GRMZM2G051912 T01	(162)	WAFRAE--AGLMSITVTHNCGEEED	PACG	PVVDFAIKTLTSQPE----ASQ	
Zea_mays_GRMZM2G405071_T01	(167)	WAFRAEGAGLMSVVI	THNCGEEEDPACG	PVVDFAIKALHDQFPDQAAAD	
O.sativa LOC_Os04g41740.1	(157)	WAFKAK--HSVVLFIHNP	GVSDDEACG	PLIDSFVAIKTLNPP-QRTKG--	
Os DUF642.3	(157)	WAFKAK--HSVVLFIHNP	GVSDDEACG	PLIDSFVAIKTLNPP-QRTKG--	
O.sativa LOC_Os04g41750.1	(155)	WAFKAR--RSDVSLIIVHNP	GVTDAAACG	PLIDSFVAIKTLQSP-PSTKD--	
S.bicolor Sb06g021280.1	(155)	WAFKAT--SSVVSIIHNP	GVAEADAACG	PLIDLFAIKTLPTP-QSSKN--	
Z.mays_ZM07MC22756_BFB0088A01022693	(155)	WAFQAA--GSVVTIIVHNP	GVAEADAACG	PLIDLFAIKTMPTP-RSSKN--	
Zea mays GRMZM2G051898 T01	(155)	WAFQAA--GSVVTIIVHNP	GVAEADAACG	PLIDLFAIKTMPTP-RSSKN--	
S.bicolor Sb06g021290.1	(149)	WAFEAT--RSVAVTLIIVHNP	GVTEDPACG	PLLDFAIKTLQPPPVLAKN--	
S.bicolor Sb06g021300.1	(154)	YAFKAR--HTAVWLTIHNP	GVVEEDPACG	PVVDFAIKALHPPS-RVKD--	
Z.mays TC468877	(150)	YAFRAR--HSAAWLVIHNP	GVVEEDPACG	PVVDFAIKALHPPSRTAKD--	
S.bicolor Sb06g021310.1	(153)	YAFRAR--HSTAWLTIHNP	SHEDDPACG	PLIDSVAIKALNPPH-HTKG--	
O.sativa LOC_Os04g41759.1	(151)	WAFRAK--HSVWLSIIVHNP	GVTEEDPACG	PLIDSVAIKALNPPH-HTKG--	
O.sativa LOC_Os04g41770.1	(237)	WAFKAE--HSAVWLSIIVHNP	GVTEEDPACG	PLIDSVAIKALNPPH-HTKG--	
S.bicolor Sb06g021320.1	(149)	WAFKAR--HSTVWLSIIVHNP	GVTEEDPACG	PLIDSVAVKTLRSPHHVKN----	
Consensus	(251)	WAFKAK S V LVHNP	GVVEEDPACG	PLIDSVAIKTL PP T	

FIGURE 4 (continued)

O. sativa LOC_Os01g42520.1	(193)	GNMLKNGGFEEGPYFLPNASWGVLPVPPNIEDDHSPLPAWMIMS-SKAVKY	301
P_virgatum TC46275	(190)	GNMLKNGGFEEGPYFLPGASWGVLPVPPNIEDDHSPLPAWMIAS-SKAVKY	350
S.bicolor Sb03g027650.1	(191)	GNMLKNGGFEEGPYFLPNASWGVLPVPPNIEDDHSPLPGWMIAS-SKAVKY	
Z.mays TC458850	(199)	NNLLKNGGFEEGPYFLPGASWGVLPVPPNIEDDHSPLPAWMIAS-SKAVKY	
Zea mays GRMZM2G034985.T02	(246)	NNLLKNGGFEEGPYFLPGASWGVLPVPPNIEDDHSPLPAWMIAS-SKAVKY	
T.aestivum c60008104011410	(190)	GNMLKNGGFEEGPYFLPNASWGVLPVPPNIEDDHSPLPAWMIMS-SKAVKY	
O. sativa LOC_Os02g11040.1	(197)	KNLVKNGDFEEGPIIPGTRWGVLPVPPNIEDDHSPLPGWVES-LKAVKY	
Os_DUF642.2	(120)	KNLVKNGDFEEGPIIPGTRWGVLPVPPNIEDDHSPLPGWVES-LKAVKY	
P.virgatum TC392	(195)	RNLVKNMGDFEEGPIIPGTRWGVLPVPPNIEDDHSPLPGWVES-LKAVKY	
S.bicolor Sb04g007160.1	(187)	KNLVKNGDFEEGPIIPGTRWGVLPVPPNIEDDHSPLPGWVES-LKAVKY	
Z.mays TC467870	(195)	KNLVKNGDFEEGPIIPGTRWGVLPVPPNIEDDHSPLPGWVES-LKAVKY	
O. sativa LOC_Os04g41710.1	(195)	-NMLRNGDFEEGPIIPGTRWGVLPVPPNIEDDHSPLPGWVES-LKAVKY	
S.bicolor Sb06g021270.1	(207)	DNLLRNGDFEEGPIIPGTRWGVLPVPPNIEDDHSPLPGWVES-LKAVKY	
Zea mays GRMZM2G051912.T01	(207)	NNMLRNGDFEEGPIIPGTRWGVLPVPPNIEDDHSPLPGWVES-LKAVKY	
Zea_mays_GRMZM2G405071.T01	(217)	NNMLRNGDFEEGPIIPGTRWGVLPVPPNIEDDHSPLPGWVES-LKAVKY	
O. sativa LOC_Os04g41740.1	(202)	-NMLKNGGFEEGPIIPGTRWGVLPVPPNIEDDHSPLPGWVES-LKAVKY	
Os_DUF642.3	(202)	-NMLKNGGFEEGPIIPGTRWGVLPVPPNIEDDHSPLPGWVES-LKAVKY	
O. sativa LOC_Os04g41750.1	(200)	-NMLKNGGFEEGPIIPGTRWGVLPVPPNIEDDHSPLPGWVES-LKAVKY	
S.bicolor Sb06g021280.1	(200)	-NMLKNGGFEEGPIIPGTRWGVLPVPPNIEDDHSPLPGWVES-LKAVKY	
Z.mays_ZM07MC22756 BF00088A01022693	(200)	-NMLKNGGFEEDGPIIPGTRWGVLPVPPNIEDDHSPLPGWVES-LKAVKY	
Zea mays GRMZM2G051898.T01	(200)	-NMLKNGGFEEDGPIIPGTRWGVLPVPPNIEDDHSPLPGWVES-LKAVKY	
S.bicolor Sb06g021290.1	(195)	-NMIQNGDFEEGPIIPGTRWGVLPVPPNIEDDHSPLPGWVES-LKAVKY	
S.bicolor Sb06g021300.1	(199)	-NMLKNGGFEEGPIIPGTRWGVLPVPPNIEDDHSPLPGWVES-LKAVKY	
Z.mays TC468877	(196)	-NMLKNGGFEEGPIIPGTRWGVLPVPPNIEDDHSPLPGWVES-LKAVKY	
S.bicolor Sb06g021310.1	(198)	-NMLKNGGFEEGPIIPGTRWGVLPVPPNIEDDHSPLPGWVES-LKAVKY	
O. sativa LOC_Os04g41759.1	(196)	-NMLRNGDLEEGPIIPGTRWGVLPVPPNIEDDHSPLPGWVES-LKAVKY	
O. sativa LOC_Os04g41770.1	(283)	-TMLRNGDLEEGPIIPGTRWGVLPVPPNIEDDHSPLPGWVES-LKAVKY	
S.bicolor Sb06g021320.1	(194)	-NMLRNGDLEEGPIIPGTRWGVLPVPPNIEDDHSPLPGWVES-LKAVKY	
Consensus	(301)	NMLKNGGFEEGPIIPGTRWGVLPVPPNIEDDHSPLPGWVES-LKAVKY	

FIGURE 4 (continued)

O.sativa LOC_Os01g42520.1	(242)	VDAAHFAVPQGARAVELVGGKESALVQEV	400
P_virgatum TC46275	(239)	VDAAHFRVPQGARAVELVGGRESALVQEV	
S.bicolor Sb03g027650.1	(240)	VDAAHFAVPQGARAVELVGGRESALVQEV	
Z.mays TC458850	(248)	VDAAHFAVPQGARAVELVGGRESALVQEV	
Zea_mays GRMZM2G034985.T02	(295)	VDAAHFAVPQGARAVELVGGRESALVQEV	
T.aestivum c60008104l1410	(239)	VDAAHFKVPEGARAVELVGGKESALVQEV	
O.sativa LOC_Os02g11040.1	(246)	IDSDFAVPRGRRRAVELLAGRESAIAQVIR	
O.sativa LOC_Os_DUF642.2	(169)	IDSDFAVPRGRRRAVELLAGRESAIAQVIR	
P_virgatum TC392	(244)	IDGSFAVPRGRRRAVELLAGRESAIAQVIR	
S.bicolor Sb04g007160.1	(236)	IDSDFAVPRGRRRAVELLAGRESAIAQVIR	
Z.mays TC467870	(244)	IDGESFAVPRGRRRAVELLAGRESAIAQVIR	
O.sativa LOC_Os04g41710.1	(244)	VDSAHFTVPHGARAVELVSGLETAIQQEV	
S.bicolor Sb06g021270.1	(257)	VDSAHFAVPRGARAVELVSGVETAIQQEV	
Zea_mays GRMZM2G051912.T01	(257)	IDSDFAVPRGRRRAVELVSGVETAIQQEV	
Zea_mays GRMZM2G405071.T01	(267)	VDSAHFAVPRGARAVELVSGVEAALLQEV	
O.sativa LOC_Os04g41740.1	(251)	IDAAHYAVPGGARAVELVSGMETAIQQEV	
O.sativa LOC_Os_DUF642.3	(251)	IDAAHYAVPGGARAVELVSGMETAIQQEV	
O.sativa LOC_Os04g41750.1	(249)	VDAAHYAVPGGARAVELVAGMEALVQEV	
S.bicolor Sb06g021280.1	(249)	VDAPHHVPHGARAVELVSGMETAIQQEV	
Z.mays_ZM07MC22756 BF00088A01022693	(249)	VDAPHHVPHGARAVELVSGMETAIQQEV	
Zea_mays GRMZM2G051898.T01	(249)	VDAPHHVPHGARAVELVSGMETAIQQEV	
S.bicolor Sb06g021290.1	(244)	LDAAHYAVPHGARAVELVSGVETAIQQEV	
S.bicolor Sb06g021300.1	(248)	LDAAHFAVPHGAYAVELVAGTEAALVQEV	
Z.mays TC468877	(245)	LDAAHFAVPRGSRRAVELVAGTEAALVQEV	
S.bicolor Sb06g021310.1	(247)	VDAAHFAVPRGARAVELVAGTEAALVQEV	
O.sativa LOC_Os04g41759.1	(245)	VDSPHHRVPQGARAVELVAGRETAIQQEV	
O.sativa LOC_Os04g41770.1	(332)	VDSARHAPVPSGAHVMVAGRECAIQQEV	
S.bicolor Sb06g021320.1	(243)	VDAPHHVPRGAGAVELVAGRECAIQQEV	
Consensus	(351)	VDAAHFAVP GARAVELVAGRESALVQEV	

FIGURE 4 (continued)

	401	450
O. sativa LOC_Os01g42520.1	(291)	CAGSMV-----AEAYAARAS- IKVPYESKGTG-GYK-RAVLFFAA
P_virgatum TC46275	(288)	CAGPMV-----AEAFAAARAT-VKVPYESKGTG-GYK-RAVLDFTA
S. bicolor Sb03g027650.1	(289)	CTGSMV-----AEAYAARAT-VKVPYESKGTG-GYK-RAVLDFTA
Z_mays TC458850	(297)	CAGSMA-----AEAYAARAT-VKVPYQSRGTG-GYK-RAVLDFTA
Zea_mays GRMZM2G034985 T02	(344)	CAGSMA-----AEAYAARAT-VKVPYQSRGTG-GYK-RAVLDFTA
T.aestivum c60008104011410	(288)	AR-----APWPRPT-PRAPP-S-----RCRTSPTA
O. sativa LOC_Os02g11040.1	(295)	CEGSLV-----VEAYAGRES-TRVAHESAGRGAAGK-RAVLFPFA
Os DUF642.2	(218)	CEGSLV-----VEAYAGRES-TRVAHESAGRGAAGK-RAVLFPFA
P_virgatum TC392	(293)	CRGSLM-----VEAYAGRES-TKVAYESAGKG-GVK-RAVLFPFA
S. bicolor Sb04g007160.1	(285)	CRGSLM-----VEAYAGRES-TKVAYESAGKG-GVK-RAVLFPFA
Z_mays TC467870	(293)	CRGSLM-----VEAYAGRES-TKVAYESAGKG-GVK-RAVLFPFA
O. sativa LOC_Os04g41710.1	(293)	CVGSMQ-----VKGYAGQGC-TTVTYSSQGTGGHTR--ASLEFAA
S.bicolor Sb06g021270.1	(307)	CASPSYDDGSSSGMKVKAYAGSSE-TTVDIDFHAGAGSKR--GKIEFRA
Zea_mays GRMZM2G051912 T01	(306)	CASSYDDGSSS--GMKVKAYAGTAE-TTVDVDFRDAGGSKR--GKIEFRA
Zea_mays GRMZM2G405071 T01	(317)	CASSSDGSSSS--GMKLKAYAGTSE-TTVDIDIFRGAGGSKR--GKIEFRA
O. sativa LOC_Os04g41740.1	(300)	CSGSLT-----VQAYASRGS-VKVTYQSQGTGGYKR--GLLEFTA
Os DUF642.3	(300)	CSGSLT-----VQAYASRGS-VKVTYQSQGTGGYKR--GLLEFTA
O. sativa LOC_Os04g41750.1	(298)	CVGSMG-----VQAYVSHGS-VKVPYESQGRGGYKR--GLLEFTA
S.bicolor Sb06g021280.1	(298)	CVGSMA-----VQAYAGGKS-VKAQFQSQGKGHKKR--GVLDFTA
Z.mays_ZM07MC22756 BF50088A01022693	(298)	CVGSMT-----VQAYAAARGS-VKVPYQSQGKGHKKR--GVLDFAA
Zea_mays GRMZM2G051898 T01	(298)	CVGSMT-----VQAYAAARGS-VKVPYQSQGKGHKKR--GVLDFAA
S.bicolor Sb06g021290.1	(293)	CSGSLG-----LRAYAASGS-VTVPHESQGRGGHKKR--GVLDFTA
S.bicolor Sb06g021300.1	(297)	CASLTL-----VDAYAARGR-QPVSYESHGMGCVR--SELEFAA
Z_mays TC468877	(294)	CAGALR-----VDAYAARGR-LPVSYESRGTGGFVR--NELEFAA
S.bicolor Sb06g021310.1	(297)	CEGSLA-----VEAYAARAT-ARATYESRGTGGGSKRAVVEFAA
O. sativa LOC_Os04g41759.1	(294)	CKDSL A-----VEAYAARAT-AKVVPYESQGTGGHKKR--AQLEFAA
O. sativa LOC_Os04g41770.1	(381)	CIGSLA-----VDAYAARAT-LKVSYESRGTGGHKKR--AELVFAA
S.bicolor Sb06g021320.1	(292)	CEGNLA-----VDVSAGRAAKLNVPYESHGTGGYKR--AELEFVA
Consensus	(401)	C GSL VEAYAARAS VKV YES C GG R AVLEF A

FIGURE 4 (continued)

		451			500
O.sativa LOC_Os01g42520.1		(328)	I	I	I
P.virgatum TC46275		(325)	I	I	I
S.bicolor Sb03g027650.1		(326)	I	I	I
Z.mays TC458850		(334)	I	I	I
Zea_mays GRMZM2G034985.T02		(381)	I	I	I
T.aestivum c6000810411410		(312)	I	I	I
O.sativa LOC_Os02g11040.1		(333)	I	I	I
O.sativa LOC_Os02g11040.1		(256)	I	I	I
P.virgatum TC392		(330)	I	I	I
S.bicolor Sb04g007160.1		(322)	I	I	I
Z.mays TC467870		(330)	I	I	I
O.sativa LOC_Os04g41710.1		(330)	I	I	I
S.bicolor Sb06g021270.1		(354)	I	I	I
Zea_mays GRMZM2G051912.T01		(351)	I	I	I
Zea_mays GRMZM2G405071.T01		(362)	I	I	I
O.sativa LOC_Os04g41740.1		(337)	I	I	I
O.sativa LOC_Os04g41750.1		(337)	I	I	I
S.bicolor Sb06g021280.1		(335)	I	I	I
Z.mays_ZM07MC22756 BF50088A01022693		(335)	I	I	I
Zea_mays GRMZM2G051898.T01		(335)	I	I	I
S.bicolor Sb06g021290.1		(330)	I	I	I
S.bicolor Sb06g021300.1		(334)	I	I	I
Z.mays TC468877		(331)	I	I	I
S.bicolor Sb06g021310.1		(336)	I	I	I
O.sativa LOC_Os04g41759.1		(331)	I	I	I
O.sativa LOC_Os04g41770.1		(418)	I	I	I
S.bicolor Sb06g021320.1		(330)	I	I	I
Consensus		(451)	I	I	I

FIGURE 4 (continued)

48/77

O.sativa	LOC_Os01g42520.1	(374)	--	501
P.virgatum	TC46275	(373)	--	
S.bicolor	Sb03g027650.1	(374)	FL	
Z.mays	TC458850	(380)	LL	
Zea_mays	GRMZM2G034985.T02	(427)	LL	
T.aestivum	c60008104l1410	(340)	--	
O.sativa	LOC_Os02g11040.1	(380)	--	
P.virgatum	TC392	(303)	--	
S.bicolor	Sb04g007160.1	(372)	--	
Z.mays	TC467870	(365)	--	
O.sativa	LOC_Os04g41710.1	(372)	--	
S.bicolor	Sb06g021270.1	(377)	L-	
Zea_mays	GRMZM2G051912.T01	(401)	--	
Zea_mays	GRMZM2G405071.T01	(398)	--	
O.sativa	LOC_Os04g41740.1	(409)	--	
P.virgatum	TC392	(380)	--	
S.bicolor	Sb04g007160.1	(380)	--	
Z.mays	TC467870	(378)	--	
O.sativa	LOC_Os04g41750.1	(380)	--	
S.bicolor	Sb06g021280.1	(380)	--	
Z.mays	GRMZM2G051898.T01	(380)	--	
S.bicolor	Sb06g021290.1	(375)	--	
S.bicolor	Sb06g021300.1	(379)	--	
Z.mays	TC468877	(376)	--	
S.bicolor	Sb06g021310.1	(381)	--	
O.sativa	LOC_Os04g41759.1	(373)	--	
O.sativa	LOC_Os04g41770.1	(463)	--	
S.bicolor	Sb06g021320.1	(372)	--	
Consensus		(501)	--	

FIGURE 4 (continued)

49/77



FIGURE 5

50/77

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1.		44.4	63.5	66.8	99.5	85.0	66.5	50.6	63.4	62.8	55.8	73.2	73.3	64.7
2.	58.6		52.1	46.5	44.7	43.6	47.2	38.4	51.5	51.2	41.6	44.5	47.4	46.2
3.	78.9	62.2		64.7	64.1	65.4	68.8	51.0	89.6	88.9	54.1	63.9	64.2	64.7
4.	79.2	59.7	78.0		67.4	66.1	68.0	50.7	64.8	64.8	62.1	65.6	63.5	64.6
5.	99.7	58.8	79.2	79.4		85.5	67.0	50.9	63.9	63.3	56.1	73.8	73.8	65.2
6.	91.8	57.0	78.8	78.5	92.1		65.8	50.8	64.8	65.0	56.9	75.2	73.4	65.1
7.	79.7	60.8	82.8	78.5	79.9	77.2		58.3	68.9	69.3	56.0	64.9	65.2	69.7
8.	61.9	48.1	61.7	60.2	62.1	60.4	64.9		49.7	50.9	43.0	51.4	50.0	54.2
9.	79.7	63.2	94.6	78.2	79.9	77.7	82.0	60.6		93.9	52.2	64.8	64.8	64.4
10.	80.7	61.6	93.9	78.0	81.0	78.5	82.1	61.7	95.5		52.7	64.2	63.9	64.1
11.	71.5	54.5	68.0	74.3	71.8	70.3	69.8	56.1	67.5	67.8		58.6	55.9	56.1
12.	83.6	58.3	76.5	78.1	83.9	84.2	76.8	61.3	77.6	77.3	71.3		76.6	64.6
13.	85.2	61.2	79.4	78.5	85.5	85.4	78.1	60.2	79.9	79.5	71.3	86.3		64.4
14.	78.6	59.8	79.1	78.6	78.9	77.0	82.3	62.8	78.8	78.8	68.8	75.7	78.8	
15.	79.2	60.5	79.5	77.1	79.5	78.2	83.4	64.9	79.2	79.7	68.0	77.6	76.8	83.2
16.	78.1	59.3	77.5	77.2	78.4	76.9	83.6	63.4	76.3	76.5	69.0	75.2	76.7	79.4
17.	66.8	57.2	76.7	67.9	67.0	66.3	70.4	49.4	76.6	75.7	58.0	64.9	67.1	67.7
18.	79.8	61.7	91.6	76.6	80.1	78.2	80.6	61.7	94.0	94.0	68.8	78.0	78.7	78.0
19.	79.4	61.1	78.9	77.5	79.7	78.2	81.3	63.4	78.1	78.4	68.8	76.8	78.1	89.2
20.	84.7	59.4	77.6	79.7	85.0	84.4	78.6	62.3	78.1	78.6	70.5	94.2	87.3	77.8
21.	70.1	54.9	80.4	68.5	70.3	68.9	72.0	65.6	80.8	81.1	64.0	68.0	69.6	69.6
22.	85.0	59.4	77.8	79.9	85.2	84.7	78.9	61.9	78.4	78.4	70.8	94.5	87.6	77.3
23.	73.8	55.9	69.5	75.8	74.1	71.8	71.5	57.4	69.3	70.5	93.5	72.3	72.3	71.3
24.	69.4	54.7	67.9	71.1	69.6	68.4	68.4	56.5	67.4	67.9	89.2	67.9	68.4	68.6
25.	72.0	79.7	74.7	72.6	72.3	69.9	73.1	57.4	75.7	74.7	68.0	72.8	73.9	74.7
26.	72.0	75.2	74.0	70.3	72.3	69.8	72.8	56.7	74.2	74.4	65.5	72.6	72.5	72.0
27.	70.7	76.6	74.3	70.8	71.0	69.0	71.5	56.3	74.2	73.3	64.5	71.0	71.1	72.0
28.	71.2	75.2	73.7	70.3	71.5	69.5	72.0	56.5	73.9	73.6	64.8	71.2	71.1	70.6

FIGURE 6

51/77

	15	16	17	18	19	20	21	22	23	24	25	26	27	28
1.	64.2	62.2	53.2	62.2	64.3	73.0	55.5	73.2	55.9	52.2	55.9	58.6	57.3	57.3
2.	47.3	46.3	42.1	50.9	47.9	45.1	45.4	45.1	40.8	39.9	79.7	69.0	70.4	68.4
3.	67.0	62.9	71.3	86.1	63.7	65.4	74.8	65.7	53.3	51.6	63.9	62.3	62.5	61.5
4.	63.9	62.5	55.4	61.9	62.9	66.6	55.1	66.8	60.6	56.8	56.4	57.1	57.6	56.8
5.	64.8	62.7	53.7	62.8	64.8	73.5	55.9	73.8	56.1	52.4	56.2	58.9	57.6	57.6
6.	66.3	62.4	54.0	62.2	66.0	73.9	55.7	74.1	55.5	52.9	54.3	57.1	55.5	56.1
7.	73.9	71.6	56.8	67.3	69.4	66.9	60.8	67.2	54.5	51.6	57.4	58.1	57.3	56.8
8.	57.2	55.3	42.1	50.0	54.7	51.9	51.2	52.9	42.7	40.7	45.8	46.7	46.2	45.3
9.	66.1	61.6	71.0	90.6	62.8	66.1	77.9	66.3	52.0	50.7	63.0	60.5	60.0	59.5
10.	66.6	62.1	70.7	90.6	62.9	66.1	77.9	66.1	52.5	50.2	62.9	61.0	60.3	59.9
11.	54.1	52.9	45.4	52.9	54.3	56.9	48.0	57.1	85.1	80.2	51.5	50.4	49.1	49.6
12.	65.6	61.3	53.9	64.3	63.0	89.4	57.8	89.7	57.3	53.7	55.7	58.1	55.8	56.0
13.	64.1	61.9	55.5	62.8	64.4	76.8	56.0	77.1	56.1	53.4	58.2	57.4	56.5	56.1
14.	73.5	68.9	55.4	61.9	82.9	65.4	56.0	65.4	55.6	53.9	57.0	56.3	56.3	55.3
15.		69.3	55.6	65.1	73.4	66.1	58.9	66.4	53.5	52.3	57.9	59.5	58.7	57.4
16.	81.1		54.8	59.6	67.6	61.8	53.3	62.0	52.4	49.9	55.9	55.9	54.5	54.3
17.	66.8	66.3		69.5	53.7	54.2	60.0	54.5	45.5	45.2	53.2	53.6	53.8	52.4
18.	79.3	75.3	75.1		61.3	65.1	85.0	65.3	52.5	50.7	62.2	60.3	60.0	59.0
19.	83.7	78.9	66.4	77.4		65.8	55.7	65.5	54.1	53.8	56.3	57.4	56.1	56.6
20.	78.2	75.5	65.4	78.5	79.4		58.3	99.7	56.5	53.2	56.2	59.2	56.8	57.1
21.	70.6	66.8	65.2	85.5	68.9	68.5		58.3	46.9	46.6	56.3	54.3	53.1	54.2
22.	78.4	75.7	65.7	78.7	79.2	99.7	68.7		56.5	53.4	56.5	58.9	56.8	57.1
23.	70.0	70.0	59.4	69.5	70.5	71.3	65.0	71.3		81.7	50.2	49.1	47.9	48.4
24.	67.9	67.2	58.1	67.4	69.4	67.4	64.3	67.6	88.2		48.9	47.8	47.1	47.9
25.	74.5	72.0	64.6	75.1	73.6	73.4	66.6	73.6	69.3	67.2		81.9	83.4	81.0
26.	73.9	72.0	65.2	73.0	72.5	73.4	64.7	73.1	66.5	64.0	89.7		91.9	91.9
27.	72.9	71.7	66.5	72.2	71.5	71.5	64.3	71.5	65.2	63.2	89.4	94.9		92.7
28.	71.8	72.0	65.0	72.4	71.5	71.8	65.4	71.8	64.5	64.0	89.2	96.5	95.4	

FIGURE 6 (continued)

52/77

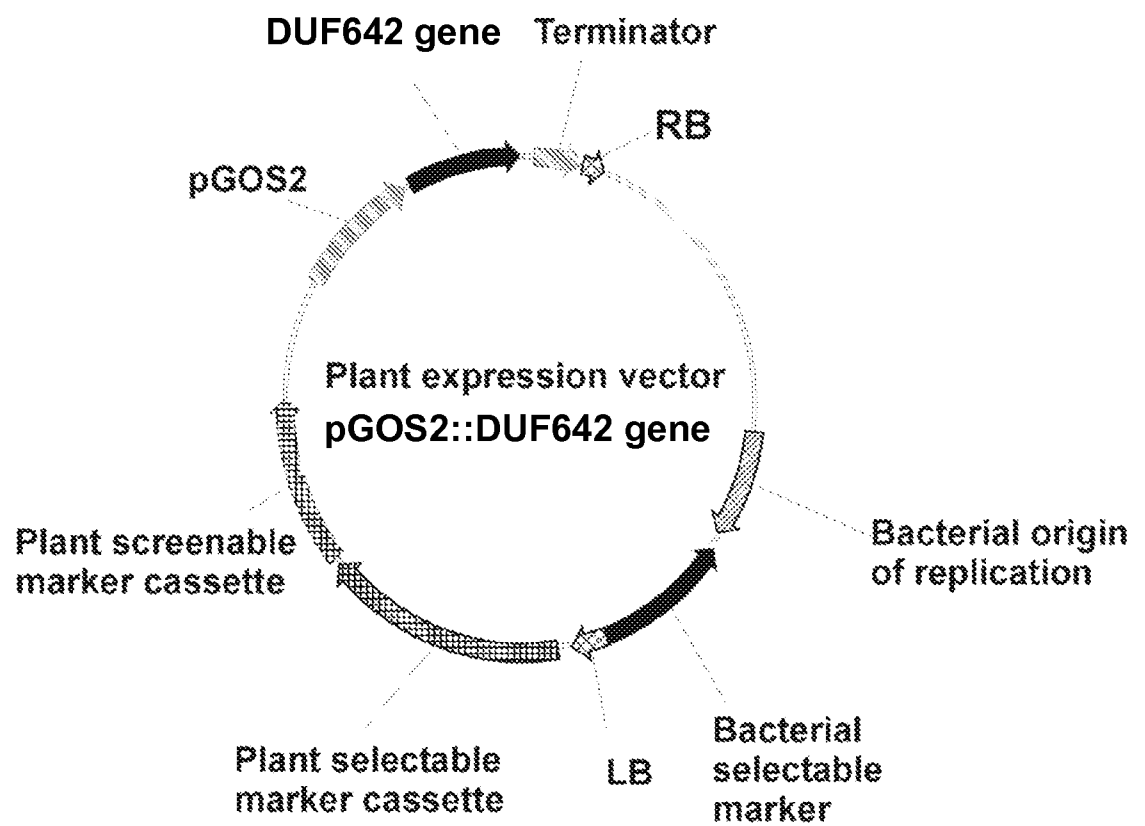


FIGURE 7

53/77

MQENSKDARNGGDSLSTTPTQVSKKPRVSFSITEHEIHHEFSHHNPVVARINNGSFGSCPGS
-----2-----

VLAAQKNWQLQFLQQPDDFYFNTLRKGILHSRTVIKDLINADDVDEISLVDNATTAAIVLQ
-----2-----3-----

QIGRAFAEGNFAKNDTVLMLHCAYQAVKKSIIQAYVTRAGGSVIEIQLPFPVTSNEEIISEFK
-----3-----

RGIEKGKANGKKIRLAIIDHITSMPCVVI PVKELVKICREEGVDQVFVDAAHAIGSVEINVK
-----1-----

EIGADFYVSNLHKWFFCPPSVAFLYCKKAASLEFDVHHPVVSHEYGNGLPIESAWIGTRDYS
-----1-----

SQLVVPAALEFVNRFE DGIQGIMKRNHEEVVKMGKMLAESWGTNLGSSPEMCAGMIMVGLPS

RLRVSSDDALRLRSHLRECHGVEVPIHYQGLKDGEEGVKDKDGVITAYARISHQVYNKSED

YCKLRDAVNRLSENLLIRKTFYPE

FIGURE 8

54/77

	1	50
A.lyrata_486720	(1)	-----
A.thaliana_AT3G62130.1	(1)	-----
C.clementina_TC23968	(1)	-----
P.trichocarpa_572528	(1)	-----
Poptr epimerase-related	(1)	-----
G.hirsutum_TC132670	(1)	-----
V.vinifera_GSVIVT00028240001	(1)	-----
G.max_Glyma02g01070.1	(1)	-----
M.truncatula_AC151948.16.5	(1)	-----
H.vulgare_TC158694	(1)	-----
T.aestivum_TC278728	(1)	-----
O.sativa_LOC_Os01g18640.1	(1)	-----
S.bicolor_Sb03g011840.1	(1)	-----
S.bicolor_Sb03g011840.3	(1)	-----
S.bicolor_Sb03g011840.2	(1)	-----
Zea_mays_GRMZM2G322186.T01	(1)	-----
Z.mays_TC513613	(1)	-----
O.sativa_LOC_Os01g18660.1	(1)	-----
S.bicolor_Sb03g011850.1	(1)	-----
Z.mays_ZM07MC24472.BF50052A17024402	(1)	-----
Zea_mays_GRMZM2G077212.T01	(1)	-----
P.patens_175806	(1)	MQIKRGIIEAVDEGGGSQPLAEPAPHEAGKAEEDDEAGYGEGKRPPGELC
P.patens_TC39863	(1)	MQIKRGIIEAVDEGGGSQPLAEPAPHEAGKAEEDDEAGYGEGKRPPGELC
S.moellendorffii_232064	(1)	-----
S.moellendorffii_413706	(1)	-----
A.lyrata_489502	(1)	-----MASSLSPPEK
A.thaliana_AT5G26600.1	(1)	-----MASSLSPPEE
B.napus_TC63646	(1)	-----MASLSPPEE
C.sinensis_TC2568	(1)	-----
P.trichocarpa_570656	(1)	-----
G.max_Glyma03g38740.1	(1)	-----
G.max_Glyma19g41330.1	(1)	-----
M.truncatula_AC166315.9.4	(1)	-----
G.max_Glyma20g22430.1	(1)	-----
Aquilegia_sp_TC26581	(1)	-----
P.persica_TC12771	(1)	-----MTGNAGSSTQEN
S.lycopersicum_TC191918	(1)	-----
S.tuberosum_TC167698	(1)	-----
Consensus	(1)	-----

FIGURE 9

55/77

	51	100
A.lyrata_486720	(1) -----MESGERRNGDSMSHNN-----	RAP
A.thaliana_AT3G62130.1	(1) -----MEAGERRNGDSMSHNN-----	RAP
C.clementina_TC23968	(1) -----MDQEARNGELTHH-----	VS
P.trichocarpa_572528	(1) -----MEGAKDPRNGGDS-----	PKQVS
Poptr_epimerase-related	(1) -----MQENSKDARNGGDSLSTT-----	PTQVS
G.hirsutum_TC132670	(1) -----MQHDDPRNGESTHR-----	SS
V.vinifera_GSVIVT00028240001	(1) -----MDSDRGENDSSHN-----	HVS
G.max_Glyma02g01070.1	(1) -----MENDNTTHLNGTG-----	KKP
M.truncatula_AC151948_16.5	(1) -----MDNHDDPRNGTVQ-----	PDS
H.vulgare_TC158694	(1) -----MASIPSEVSKNGHGNGNG-----	PSPPRA
T.aestivum_TC278728	(1) -----MASIPSEVSENGHGNGNG-----	PSPPRA
O.sativa_LOC_Os01g18640.1	(1) -----MASIPDDDDAAAAAAGAAENGYNGKNGNGPAPRPPPA	
S.bicolor_Sb03g011840.1	(1) -----MASDPAPDDAAAAAENGHGNDNGHGNGNGNGNGNGPSP---	
S.bicolor_Sb03g011840.3	(1) -----MASDPAPDDAAAAAENGHGNDNGHGNGNGNGNGNGNGPSP---	
S.bicolor_Sb03g011840.2	(1) -----MASDPAPDDAAAAAENGHGNDNGHGNGNGNGNGNGNGPSP---	
Zea_mays_GRMZM2G322186_T01	(1) -----MASDPAPDDAAAAVNGHGNDNGHGNG-----	PSPS
Z.mays_TC513613	(1) -----MASDPAPDDAAAAVNGHGNDNGHGNG-----	PSPS
O.sativa_LOC_Os01g18660.1	(1) -----MASLQSGGDAANGVDADVDGAASPPS-----	A
S.bicolor_Sb03g011850.1	(1) -----MASAPHGDAAEVRNAAPP-----	A
Z.mays_ZM07MC24472_BF50052A17024402	(1) -----MASAPHDDAAKGRNAAPP-----	E
Zea_mays_GRMZM2G077212_T01	(1) -----MASAPHDDAAKGRNAAPP-----	E
P.patens_175806	(51) NGESLSPSRVCQGNSALLIAELLSAGKAGIPGKLDDTA-----	EDDDA
P.patens_TC39863	(51) NGESLSPSRVCQGNSALLIAELLSAGKAGIPGKLDDTA-----	EDDDA
S.moellendorffii_232064	(1) -----	
S.moellendorffii_413706	(1) -----MTVSMQGE-----	AGSIQ
A.lyrata_489502	(11) -----ASYHHRHS-----SKRYSTNGNVESSSVS-----	DFVKR
A.thaliana_AT5G26600.1	(11) -----ASYHHRHTKRYTSSASSASSTTNGTVESS-VS-----	DFVKR
B.napus_TC63646	(10) EAAAAAASFHHRRSKR-----YSSSTTNGLAES-----	VPDKR
C.sinensis_TC2568	(1) -----MDTIGADTKIS-----	NTYLP
P.trichocarpa_570656	(1) -----MASSSLKHHKSHS--L-----	NGFTT
G.max_Glyma03g38740.1	(1) -----MSSTTIHNNHHNGSH-----	IP-KK
G.max_Glyma19g41330.1	(1) -----MVKSLFLISIPMSSTTIHHNGSH-----	IP-KK
M.truncatula_AC166315_9.4	(1) -----MASEKNNHRSIPDINGTH-----	PTPKK
G.max_Glyma20g22430.1	(1) -----MASNGDAVHPHPSKKI-----	KLFSN
Aquilegia_sp_TC26581	(1) -----MNHNNNNNNNN-----	NH
P.persica_TC12771	(13) HFRQRLRVMAEKIVIFCRYVPLIVLLPYRTLQFFYQI-----	KTQVP
S.lycopersicum_TC191918	(1) -----MEPANDDHHRANGSDHSH-----	LAKKP
S.tuberosum_TC167698	(1) -----MEPANDDHHRANGSDHSH-----	LAKKP
Consensus	(51)	

FIGURE 9 (continued)

56/77

A.lyrata_486720	101	150
A.thaliana_AT3G62130.1	(20) KKP---RLAGLLTESDIDSEFAHHQTGVARINNGSFGCCPGSVLEAQREW	
C.clementina_TC23968	(20) KKP---RLAGLLTESDIDSEFAHHQTGVARINNGSFGCCPGSVLEAQREW	
P.trichocarpa_572528	(16) KKP---KLTRCISEAEIRDEFSSHQHGVARINNGSFGSCPKSVLADQQKW	
Poptr_epimerase-related	(20) KKP---RTSGHITEQEIHHEFSHHNLNVARINNGSFGSCPGSVLAAQKNW	
G.hirsutum_TC132670	(24) KKP---RVSF SITEHEIHHEFSHHNPNVARINNGSFGSCPGSVLAAQKNW	
V.vinifera_GSVIVT00028240001	(17) KKP---ISSQFITESDIRDEFSSHHPVARINNGSFGSCPGSVLAAQRHW	
G.max_Glyma02g01070.1	(18) KKP---KLSAFISEEEIRQEFSSHQGIARINNGSFGSCPASIIAAQKEW	
M.truncatula_AC151948_16.5	(17) KLA---HVTHITISEAEIRDEFSSHHRGVARINNGSFGSCPHSVLAAQSTW	
H.vulgare_TC158694	(17) KKP---KLTHPISSHELLHEFSHHQTAVARINNGSFGSCPRSIILTAQTTW	
T.aestivum_TC278728	(26) KRPR----AAFISAAQIRDEF FAHHDPAVARVNNGSFGCCPASVLEAQARW	
O.sativa_LOC_Os01g18640.1	(26) KRPR----AALISAAQIRDEF FAHHDPAVARVNNGSFGCCPASVLEAQARW	
S.bicolor_Sb03g011840.1	(41) KRPR----RSVISAAQIRAEFEHHEAGVARVNNGSFGCCPSSLLDAQARW	
S.bicolor_Sb03g011840.3	(38) KRA-----RSVISADEIRTEFAHHDAAVARVNNGSFGCCPASVLAQAHW	
S.bicolor_Sb03g011840.2	(38) KRA-----RSVISADEIRTEFAHHDAAVARVNNGSFGCCPASVLAQAHW	
Zea_mays_GRMZM2G322186_T01	(31) KRPR----RAVISEAEIRVEFAHHDAAVARVNNGSFGCCPASVLAQAHW	
Z.mays_TC513613	(31) KRPR----RAVISEAEIRVEFAHHDAAVARVNNGSFGCCPASVLAQAHW	
O.sativa_LOC_Os01g18660.1	(29) KRPRAGAGAAAITDAEVRAEFAHHDRAVARLNNGTFGCCPASVLAARARW	
S.bicolor_Sb03g011850.1	(20) KR-----ASAPITEADIRAEFAHHDGVARVNNGTFGCCPASVLAARARW	
Z.mays_ZM07MC24472_BF50052A17024402	(20) KR-----TSATITEADIRAEFAHHDGTVARVNNGTFGCCPASVLAARARW	
Zea_mays_GRMZM2G077212_T01	(20) KR-----TSATITEADIRAEFAHHDGTVARVNNGTFGCCPASVLAARARW	
P.patens_175806	(94) ESKR--VVDRKNLSISIEEEFSHHKPGLARLNNGSFGSCPQSVLAAQENC	
P.patens_TC39863	(94) ESKR--VVDRKNLSISIEEEFSHHKPGLARLNNGSFGSCPQSVLAAQENC	
S.moellendorffii_232064	(1) -MPI--VESGKIWEMLAREFSHHRQGIARINNGSFGSCPRSVLEAQQRW	
S.moellendorffii_413706	(15) RMPI--VGSGKIWEMLAREFSHHRQGIARINNGSFGSCPRSVLEAQQRW	
A.lyrata_489502	(40) PKIS--PQ-NYISCSEIESEFSHHDPDFARINNGSFGCCPSSILALQRDW	
A.thaliana_AT5G26600.1	(47) PKIS--HP-NYISSSEIESEFSHHDPDFARINNGSFGCCPSSILALQRDW	
B.napus_TC63646	(44) HKIS--LPPSFISSEIESEFSHHDPFAFARINNGSFGSCPSSVIAAQRDW	
C.sinensis_TC2568	(17) EKPE--LSPTLISDTELLSEFSHHVQNIARINNGSFGCCPKSVISAQQKW	
P.trichocarpa_570656	(20) PAKR--TKLSFISDSEIQSEFSHHDTIARINNGSFGSCPQSVISAQQNL	
G.max_Glyma03g38740.1	(21) PKLS--ST--FITPSEIESEFGHHDAAVARINNGSFGCCPASVLAQAQRW	
G.max_Glyma19g41330.1	(29) PKLS--S-----LEIESEFGHHDSSVARINNGSFGCCPASVLSAQRRW	
M.truncatula_AC166315_9.4	(25) PKLS--SSSNFITASEIQSEFSHHDSAVSRINNGSFGSCPSSVISAQQQW	
G.max_Glyma20g22430.1	(22) SINS--TNSSEFITHAEIESEFSHHHTVARLNNGAFGCCPASVIAVQREW	
Aquilegia_sp_TC26581	(14) VSKK--LKLSTITNSEIRQEFEEHHDKKTARINNGSFGSCPATILAAQQW	
P.persica_TC12771	(56) IEPK--HSSSFITDSEIQSEFAHHAPGVAMMNGSFGCCPASVISALHQW	
S.lycopersicum_TC191918	(24) KLST--RSSSLITDSEIREEF FAHHTGIARINNGSFGSCPASIIAAQKRW	
S.tuberosum_TC167698	(24) KLST--SSSSLITDSEIREEF FAHHTGIARINNGSFGSCPASIIAAQKRW	
Consensus	(101) KK SS ISESEIR EFSHHD GVARINNGSFGCCPASVLAQA W	

FIGURE 9 (continued)

57/77

A.lyrata_486720	(67)	QLRYLRQPDEFFYFNGLRRLGIVASRTVISDLINADDVDEVSLVDNATTTAAA	151	200
A.thaliana_AT3G62130.1	(67)	QLRYLRQPDEFFYFNGLRRLGIVASRTVISDLINADDVDEVSLVDNATTTAAA		
C.clementina_TC23968	(63)	QLKFLQQPDDFFYFNSLRKGILESRAAVKDLINADDVGEISLVDNATTTAAA		
P.trichocarpa_572528	(67)	QLQFLQQPDDFFYFNTLRKGILHSRTVIKDLINADDVDEISLVDNATTTAAA		
Poptr epimerase-related	(71)	QLQFLQQPDDFFYFNTLRKGILHSRTVIKDLINADDVDEISLVDNATTTAAA		
G.hirsutum_TC132670	(65)	QLQFLRQPDAFYFNSLRDGITASRKIIKDLINADHVDEVSLVDNATTTAAA		
V.vinifera_GSVIVT00028240001	(65)	QLRFLQQPDDFFYFNHLRKGILLESRTVVKGLINADSVDEVSLIDNATTTAAA		
G.max_Glyma02g01070.1	(64)	QLRFLQQPDDFFYFNALRPGILASRAAIQDLINAGHVDDVSLVDNATTTAAA		
M.truncatula_AC151948_16.5	(64)	QLRFLQQPDDFFNTLRNGILDSRKIIKDLINADDVEEISLIDNATTTAAA		
H.vulgare_TC158694	(72)	QRLFLAQPDFAFYFDGLQPLRRSRAAVAALVNAGDVSEISLVDNATTTAAA		
T.aestivum_TC278728	(72)	QRLFLAQPDFAFYFDGLQPLRRSRAAVAALVNAGDVSEISLVDNATTTAAA		
O.sativa_LOC_Os01g18640.1	(86)	QRLFLAQPDFFYFHALQPLRRSRAAVAGLVNAGDVAEVSIVDNATTTAAA		
S.bicolor_Sb03g011840.1	(83)	QGLFLAQPDFAFYFHGLQQGLLRSRAAAVAEAVGAGDVSEVSIVDNATTTAAA		
S.bicolor_Sb03g011840.3	(83)	QGLFLAQPDFAFYFHGLQQGLLRSRAAAVAEAVGAGDVSEVSIVDNATTTAAA		
S.bicolor_Sb03g011840.2	(83)	QGLFLAQPDFAFYFHGLQQGLLRSRAAAVAEAVGAGDVSEVSIVDNATTTAAA		
Zea_mays_GRMZM2G322186_T01	(76)	QRLFLAQPDFAFYFHVLLQGLLRSRAAAVAEAVGAGDVSEVSIVDNATTTAAA		
Z.mays_TC513613	(76)	QRLFLAQPDFAFYFHVLLQGLLRSRAAAVAEAVGAGDVSEVSIVDNATTTAAA		
O.sativa_LOC_Os01g18660.1	(79)	QRLFLSQPDFAFYFHHLLQPLGARSRAAVAAAVGAGDASEVSIVDNVTTAAA		
S.bicolor_Sb03g011850.1	(65)	QRLFLSQPDFAFYFDSLQPLGARSRAAVAAAVGAGDASEVSIVDNATTTAAA		
Z.mays_ZM07MC24472_BF50052A17024402	(65)	QRLFLSQPDFAFYFDSLQPLGARSRAAVAAAVGAGDASEVSIVDNATTTAAA		
Zea_mays_GRMZM2G077212_T01	(65)	QRLFLSQPDFAFYFDSLQPLGARSRAAVAAAVGAGDASEVSIVDNATTTAAA		
P.patens_175806	(142)	SRCWLRQPDKSYFGPLEEGLYRARKEVADLVNAP-VEEVFLLENVTTAAAS		
P.patens_TC39863	(142)	SRCWLRQPDKSYFGPLEEGLYRARKEVADLVNAP-VEEVFLLENVTTAAAS		
S.moellendorffii_232064	(48)	HRQWLEQPDFFYFGLPEKGLRASREAIASIIIRAPDVEEIVLVDNVTTAAA		
S.moellendorffii_413706	(63)	HRQWLEQPDFFYFGLPEKRLRASREAIASIIIRAPNVEEIVLVDNVTTAAA		
A.lyrata_489502	(87)	QLRFLRQPDREFFYFDELKPKISDSRSVIKRLINADHDDEVSLVDNATTTAAA		
A.thaliana_AT5G26600.1	(94)	QLRFLRQPDREFFYFDELKPKISDSRSVIKRLINAEHDDDEVSLVDNATTTAAA		
B.napus_TC63646	(92)	QLRFLRQPDREFFYFDELKPNISASRAAIKRLINAEHDDDEVSLVDNATTTAAA		
C.sinensis_TC2568	(65)	QLEFLKQPDDFFYFNRLKDEILKSRLVIRDLVNADHVDEISIVDNATTATA		
P.trichocarpa_570656	(68)	QLQFLRQPDNFYFNTLRKPSILHSRLIKSLVNAHSVDEISLVDNATTTAAA		
G.max_Glyma03g38740.1	(67)	QLQYLRQPDHFYFNDLQSGILQSRTLIKDLVNADHVNEISIVDNATTTAAA		
G.max_Glyma19g41330.1	(70)	QLQYLRQPDHFYFNDLKSAYLQSRITLIKDLVNADHVDEISIVDNATTTAAA		
M.truncatula_AC166315_9.4	(73)	QLKYLRQPDHFYFNLKPAINNRSRTIQNLVNADKINEISIVDNATTTAAA		
G.max_Glyma20g22430.1	(70)	QMKNLRQPDHFYFNDLKKGLLHSRTIIKDLVNADHVDEISLVDNASTATA		
Aquilegia_sp_TC26581	(62)	QLKFLQQPDDFFYFNQLRKGINESRNLIKDLINAEHVDEVSLVDNATTTAAA		
P.persica_TC12771	(104)	QLKMFRQPSHFLNKLKRNILSRITIKDHINAEHDDEVSLVDNISTAAA		
S.lycopersicum_TC191918	(72)	QLRFLQQPDDFFLNLHLQKRILHSRTIIKDVINAEHVDEVSLVDNATTTAAA		
S.tuberosum_TC167698	(72)	QLRFLQQPDDFFLNLHLQKRILHSRTIIKDVINAEHVDEVSLVDNATTTAAA		
Consensus	(151)	QL FL QPD FYFN LR GLL SRAVI DLVNADDVDEVSLVDNATTTAAA		

FIGURE 9 (continued)

58/77

	201	250
A.lyrata_486720	(117) IVLQKVGRCFSEGGYKKEDTVVMFHCAFQSVKKSIIQAYVSRVGGFTVEVR	
A.thaliana_AT3G62130.1	(117) IVLQKVGRCFSEGGYKKEDTVVMFHCAFQSVKKSIIQAYVSRVGGSTVEVR	
C.clementina_TC23968	(113) IVLQQIGRGFTGRFHRNDTVLMLHCAFQAVKKSIIQAYVTRAGGSVVEVQ	
P.trichocarpa_572528	(117) IVLQQIGRAFAEGKFQKNDTVLMLHCAFQAVKKSIIQAYVTRAGGSVIEIQ	
Poptr_epimerase-related	(121) IVLQQIGRAFAEGNFQKNDTVLMLHCAFQAVKKSIIQAYVTRAGGSVIEIQ	
G.hirsutum_TC132670	(115) IVLQIRGRSFAEGKFQKNDTVLMLHCAFQAVKKSIIQAYVTRAGGSVIEVR	
V.vinifera_GSVIVT00028240001	(115) IVLQQIGRAFAQKGFQKGDVVVLMHCAFQSVKKSIIQAYVTGAGGSVIEVQ	
G.max_Glyma02g01070.1	(114) VVLQQIGRRFVRGYFHRNDVAVVMFHCAFQAVKKSIEAYVSPIGGTIVEVE	
M.truncatula_AC151948_16.5	(114) IVLQQIGHHFFSSGKFCRNDTVIIIFHCAYQAVKKSIEAYVIPVGGSVIEVE	
H.vulgare_TC158694	(122) IVLQHAWSFAEGHFARGDAVLMLHYAYGAVKKSIIQAYVARAGATVVEVP	
T.aestivum_TC278728	(122) IVLQHAWSFAEGHFARGDAVLMLHYAYGAVKKSIIQAYVARAGATVVEVP	
O.sativa_LOC_Os01g18640.1	(136) IVLQHAWSFAEGRFSRGDAVLMLHYAYGAVKKSIIHAYVARAGATVVEVP	
S.bicolor_Sb03g011840.1	(133) IVLQHAWSFAEGRFASGDAVLMLHYAYGAVKKSIIHAYVARAGATVVEVP	
S.bicolor_Sb03g011840.3	(133) IVLQHAWSFAEGRFASGDAVLMLHYAYGAVKKSIIHAYVARAGATVVEVP	
S.bicolor_Sb03g011840.2	(133) IVLQHAWSFAEGRFASGDAVLMLHYAYGAVKKSIIHAYVARAGATVVEVP	
Zea_mays_GRMZM2G322186_T01	(126) IVLQHAWSFAEGHFARGDAVLMLHYAYGAVKKSIIHAYVARAGATVVEVP	
Z.mays_TC513613	(126) IVLQHAWSFAEGHFARGDAVLMLHYAYGAVKKSIIHAYVARAGATVVEVP	
O.sativa_LOC_Os01g18660.1	(129) IIMQHVAWSFAEGDFARGDVVLMFLYTYCSIKNSIIHAYVARAGATVVEVP	
S.bicolor_Sb03g011850.1	(115) IIMQHVAWSFAEGVFARGDVVLMFLHYTYSSVKNSIIHAYVVRAGATVVEVP	
Z.mays_ZM07MC24472_BF50052A17024402	(115) IIMQHVAWSFAEGAFARGDVVLMFLHYTYSSVKNSIIHAYVVRAGATVVEVP	
Zea_mays_GRMZM2G077212_T01	(115) IIMQHVAWSFAEGAFARGDVVLMFLHYTYSSVKNSIIHAYVVRAGATVVEVP	
P.patens_TC39863	(191) MVALDVMWAFAEGRYKKGDSILMLNFTYGALKKSFQAYARAGGRIFQVQ	
S.moellendorffii_232064	(191) MVALDVMWAFAEGRYKKGDSILMLNFTYGALKKSFQAYARAGGRIFQVQ	
S.moellendorffii_413706	(98) MIAQDMAWGFLEKRFHPGDAILMLNCSYAAVKKCFEAYAVRAGARIIEVE	
A.lyrata_489502	(113) MIAQDMAWGFLEKRFHPGDAILMLNCSYAAVKKCFEAYAVRAGARIIEVE	
A.thaliana_AT5G26600.1	(137) IVLQQTAWAFREGRFDRGDAVVMLHYAYGAVKKSVEAYVTRSGGHVIEVQ	
B.napus_TC63646	(144) IVLQQTAWAFREGRFDRGDAVVMLHYAYGAVKKSVEAYVTRSGGHVTEVQ	
C.sinensis_TC2568	(142) IVLQQTAWAFREGRFDRGDAVVMLHYAYGAVKKSVEAYVTRSGGEVIEMQ	
P.trichocarpa_570656	(115) IVLQSVAWFELQGFENKGDVAVMLQYAYGAVKKSLEAYVTRAGGYVIEAW	
G.max_Glyma03g38740.1	(118) IVLQNCWAGFNEGFRFSKGDVAVMLHYAYGAVKKSVOAYVTRAGGEVIEVH	
G.max_Glyma19g41330.1	(117) IVLQHTAWNFRGKFKQGDVVLMMLHYAYGAVKKSMEAYVTRAGGNVVEVP	
M.truncatula_AC166315_9.4	(120) IVLQHTAWNFRGKFKQGDVVLMMLHYAYGAVKKSMEAYVTRAGGNVVEVP	
G.max_Glyma20g22430.1	(123) IVLQHTAWCFREGKFNKGDVVVLMMLHYAYGAVKKSMEAYVTRAGGKVIEWP	
Aquilegia_sp_TC26581	(120) IVLQQAAWAFQEAQFKQGDVVVLMLHYAYGAVKKAIEAYVVRAGGTVIEWP	
P.persica_TC12771	(112) IVLQQIGWAFTEGRFQKGDAAVMLHYAYGAVKKSIIQAYVSRAGGFVIEVQ	
S.lycopersicum_TC191918	(154) IVLQQTAWAFAEKFNKGDVAVIMFSCYTGAVKNSIEAYVTRAGGSVIEVP	
S.tuberosum_TC167698	(122) IVLQHVGVAFAEGRFKKGDVAVMLHCAFQAVKKSIEAYVTRAGGSVIVVH	
Consensus	(122) IVLQHVGVAFAEGRFKKGDVAVMLHCAFQAVKKSIEAYVTRAGGSVIVVH	
	(201) IVLQ VAWAFAEGRF KGDVAVMLHYAYGAVKKSII AYVTRAGGTVEVP	

FIGURE 9 (continued)

59/77

A.lyrata_486720	(167)	LPPFVNSNEEIIISKFREGL	E-KGRANGRT-VRLAIIDHITSMPCVLMPPVR
A.thaliana_AT3G62130.1	(167)	LPPFVNSNEEIIISKFREGL	E-KGRANGRT-VRLAIIDHITSMPCVLMPPVR
C.clementina_TC23968	(163)	LPFPLASEEEIINEFKKGL	E-KGKKDKM-IRLAIIDHITSMPCVVIIPVR
P.trichocarpa_572528	(167)	LPPFVNSNEEIIAEFKRGL	G-KGKANGRK-IRLAIIDHITAMPCVVIIPVK
Poptr_epimerase-related	(171)	LPPFVTSNEEIISEFKRG	E-KGKANGKK-IRLAIIDHITSMPCVVIIPVK
G.hirsutum_TC132670	(165)	LPPFVNSEEEIISELKKS	ID-EGKFNGRR-IRLAIIDHITSMPSVVIPLK
V.vinifera_GSVIVT00028240001	(165)	LPFPLTSKEEIVSEFRK	GLE-KGKSDGRH-VRLAIIDHITSMPCVVPVPE
G.max_Glyma02g01070.1	(164)	LPPFVRSEEEIITEFKK	GLE-KGKLNGGR-VRLAIIDHITSMPSFVLPVR
M.truncatula_AC151948_16.5	(164)	LPPFVNSNEEIIAEFKK	GIE-RGKINGGR-VRLAIIDHITSMPSVVIIPVR
H.vulgare_TC158694	(172)	LPPFVTSPhaIIAEFHA	ALA-VAKAGGRK-VRLAVIDHITSMPSVLIIPVK
T.aestivum_TC278728	(172)	LPPFVTSPhaIIAEFHA	ALA-VAKAGGRK-VRLAVIDHITSMPSVLIIPVK
O.sativa_LOC_Os01g18640.1	(186)	LPPFVASADAIIAEFRA	ALD-VAKAGGRK-VRLAVIDHITSMPSVVIIPVK
S.bicolor_Sb03g011840.1	(183)	LPPFVASADAIIAEFRA	ALA-VAKEGGRR-VRLAVIDHITSMPSVVIIPVK
S.bicolor_Sb03g011840.3	(183)	LPPFVASADAIIAEFRA	ALA-VAKEGGRR-VRLAVIDHITSMPSVVIIPVK
S.bicolor_Sb03g011840.2	(183)	LPPFVASADAIIAEFRA	ALA-VAKEGGRR-VRLAVIDHITSMPSVVIIPVK
Zea_mays_GRMZM2G322186_T01	(176)	LPPFVASADAIIAEFRA	ALA-VAKEGGRR-VRLAVIDHITSMPSVVIIPVK
Z.mays_TC513613	(176)	LPPFVASADAIIAEFRA	ALA-VAKEGGRR-VRLAVIDHITSMPSVVIIPVK
O.sativa_LOC_Os01g18660.1	(179)	LPPFVSSPDAIVAEFRA	ALA-VARDGRRR-VRLAVIDHITAMPTVLIIPVK
S.bicolor_Sb03g011850.1	(165)	LPPFVASPGAVVAEFRT	ALA-LAKAGGRS-VRLAVIDHITSMPSVLLPVK
Z.mays_ZM07MC24472_BF50052A17024402	(165)	LPPFVASAVAVVAEFRA	ALA-LAQAGGRR-VRLAVIDHITSMPSVLLPVK
Zea_mays_GRMZM2G077212_T01	(165)	LPPFVASAVAVVAEFRA	ALA-LAQAGGRR-VRLAVIDHITSMPSVLLPVK
P.patens_175806	(241)	IPFPVSSEEQILQVFEE	ALEEEENPSSIIRMAVFDHIVSMPTMILPIR
P.patens_TC39863	(241)	IPFPVSSEEQILQVFEE	ALEEEENPSSIIRMAVFDHIVSMPTMILPIR
S.moellendorffii_232064	(148)	LALPILSARQIIDSFQ	RSLEQAAS-STTIRLAVIDHITSMPSVVLPAR
S.moellendorffii_413706	(163)	LALPILSARQIIDSFQ	RSLEQAAS-STTIRLAVIDHITSMPSVVLPAR
A.lyrata_489502	(187)	LPPFVNSADEIIDRFRI	GLE-SGKANGRR-VRLALIDHVTSMPSVVIPIK
A.thaliana_AT5G26600.1	(194)	LPPFVISADEIIDRFRI	GLE-SGKANGRR-VRLALIDHVTSMPSVVIPIK
B.napus_TC63646	(192)	LPPFVNSAEEIVNRFRT	GIA-LGKANGRK-VRLALIDHVTSMPSVVIPIK
C.sinensis_TC2568	(165)	FLFPVNSDGEIVSEFR	KALE-RGKANGRK-VRLAMIDHVTIPSVPPIPVK
P.trichocarpa_570656	(168)	LPPFVASKEEIVSEFR	KALE-RGKENGKKVRLAVIDHVTSMPSVVIIPVK
G.max_Glyma03g38740.1	(167)	LPPFVNSNDEIVSEFR	KALE-RGKSNGNR-VRLAVIDHVTSMPCVVIIPVK
G.max_Glyma19g41330.1	(170)	LPPFVNSNDEIVSEFR	KALE-RGKSNGNR-VRLAVIDHVTSMPCVVIIPVK
M.truncatula_AC166315_9.4	(173)	LPPFVSSNDEIVTEFR	KALE-KGKVDGKK-IRLAVIDHVTSMPCVVIIPVK
G.max_Glyma20g22430.1	(170)	LPPFVTSNDDVVNEFR	KALE-RGKSNGNR-IRLAVIDHVTSMPSVVIIPVK
Aquilegia_sp_TC26581	(162)	LPPFVNSNEEIIICEFK	KALQ-QGKLNGRK-VRLAVIDHITSMPSVVIPIK
P.persica_TC12771	(204)	FNFPPLNSNEEIISEF	RKALE-REKNGNR-VRLAVIDHVTCPVTVMVPVK
S.lycopersicum_TC191918	(172)	LPPPLRSEEEIVAEFR	KALA-KGKANGKK-VRLAIIDHITSMPCVVIIPVR
S.tuberosum_TC167698	(172)	LPPPLRSEEEIVAEFR	KALA-KGKANGKK-VRLAIIDHITSMPCVVIIPVR
Consensus	(251)	LPPFV S EEIIAEFR	KALE GKANGRR VRLAVIDHITSMPSVVIIPVK

FIGURE 9 (continued)

60/77

	301		350
A.lyrata_486720	(215)	ELVKICREEGVEQVFVDA	AAHAIGSVKVDVKEIGADYYVSNLHKWLFCCPPS
A.thaliana_AT3G62130.1	(215)	ELVKICREEGVEQVFVDA	AAHAIGSVKVDVKEIGADYYVSNLHKWFFCCPPS
C.clementina_TC23968	(211)	ELVKICRDEGVDQVFVDA	AAHAMGSIKIDVKEIGADFYVSNLHKWFFCCPPS
P.trichocarpa_572528	(215)	ELVKICREEGVEQVFVDA	AAHAIGSVKVDVKEIGADFYVSNLHKWFFCCPPS
Poptr_epimerase-related	(219)	ELVKICREEGVDQVFVDA	AAHAIGSVKVDVKEIGADFYVSNLHKWFFCCPPS
G.hirsutum_TC132670	(213)	ELVRCRAEGIEQVFVDA	AAHAIGSVKVDVKEIGADFYVSNLHKWFFCCPPS
V.vinifera_GSVIVT00028240001	(213)	ELVKICRQEGVDQVFVDA	AAHAIGSVKVDVKEIGADFYVSNLHKWFFCCPPS
G.max_Glyma02g01070.1	(212)	ELIRVCREHGVEQVFVDA	AAHAIGSVKVDVKEIGADFYVSNLHKWFFCCPPS
M.truncatula_AC151948_16.5	(212)	ELIRVCRENEVDQVFVDA	AAHAIGSVKVDVKEIGADFYVSNLHKWFFCCPPS
H.vulgare_TC158694	(220)	ELVAICRQEGVDQVFVDA	AAHAIGSVKVDVKEIGADFYVSNLHKWFFCCPPS
T.aestivum_TC278728	(220)	ELVAICRQEGVDQVFVDA	AAHAIGSVKVDVKEIGADFYVSNLHKWFFCCPPS
O.sativa_LOC_Os01g18640.1	(234)	ELVAICREEGVDQVFVDA	AAHAIGSVKVDVKEIGADFYVSNLHKWFFCCPPS
S.bicolor_Sb03g011840.1	(231)	ELVAICREEGVDQVFVDA	AAHAIGSVKVDVKEIGADFYVSNLHKWFFCCPPS
S.bicolor_Sb03g011840.3	(231)	ELVAICREEGVDQVFVDA	AAHAIGSVKVDVKEIGADFYVSNLHKWFFCCPPS
S.bicolor_Sb03g011840.2	(231)	ELVAICREEGVDQVFVDA	AAHAIGSVKVDVKEIGADFYVSNLHKWFFCCPPS
Zea_mays_GRMZM2G322186_T01	(224)	ELVAICREEGVDQVFVDA	AAHAIGSVKVDVKEIGADFYVSNLHKWFFCCPPS
Z.mays_TC513613	(224)	ELVAICREEGVDQVFVDA	AAHAIGSVKVDVKEIGADFYVSNLHKWFFCCPPS
O.sativa_LOC_Os01g18660.1	(228)	ELVAICREEGVDQVFVDA	AAHAIGSVKVDVKEIGADFYVSNLHKWFFCCPPS
S.bicolor_Sb03g011850.1	(213)	ELVAICREEGVDQVFVDA	AAHAIGSVKVDVKEIGADFYVSNLHKWFFCCPPS
Z.mays_ZM07MC24472_BF50052A17024402	(213)	ELVAICREEGVDQVFVDA	AAHAIGSVKVDVKEIGADFYVSNLHKWFFCCPPS
Zea_mays_GRMZM2G077212_T01	(213)	ELVAICREEGVDQVFVDA	AAHAIGSVKVDVKEIGADFYVSNLHKWFFCCPPS
P.patens_175806	(291)	QLIKLCRSYGVENIFIDGA	HGIGNLELNLTELDADYYTSNLHKWFFCPTT
P.patens_TC39863	(291)	QLIKLCRSYGVENIFIDGA	HGIGNLELNLTELDADYYTSNLHKWFFCPTT
S.moellendorffii_232064	(196)	DLVALCRDAGVEQIFVDA	AAHAIGSVKVDVKEIGADFYVSNLHKWFFCPTT
S.moellendorffii_413706	(211)	DLVALCRDAGVEQIFVDA	AAHAIGSVKVDVKEIGADFYVSNLHKWFFCPTT
A.lyrata_489502	(235)	ELVKICREEGVDQVFVDA	AAHAIGSVKVDVKEIGADFYVSNLHKWFFCPTT
A.thaliana_AT5G26600.1	(242)	ELVKICREEGVDQVFVDA	AAHAIGSVKVDVKEIGADFYVSNLHKWFFCPTT
B.napus_TC63646	(240)	ELVKICREEGVDQVFVDA	AAHAIGSVKVDVKEIGADFYVSNLHKWFFCPTT
C.sinensis_TC2568	(213)	ELIKICREEGVDQVFVDA	AAHAIGSVKVDVKEIGADFYVSNLHKWFFCPTT
P.trichocarpa_570656	(217)	ELVKICREEGVDQVFVDA	AAHAIGSVKVDVKEIGADFYVSNLHKWFFCPTT
G.max_Glyma03g38740.1	(215)	ELIQICREEGVDQVFVDA	AAHAIGSVKVDVKEIGADFYVSNLHKWFFCPTT
G.max_Glyma19g41330.1	(218)	ELIQICREEGVDQVFVDA	AAHAIGSVKVDVKEIGADFYVSNLHKWFFCPTT
M.truncatula_AC166315_9.4	(221)	ELIQICREEGVEQVFVDA	AAHAIGSVKVDVKEIGADFYVSNLHKWFFCPTT
G.max_Glyma20g22430.1	(218)	DLVKICREEGVEQVFVDA	AAHAIGSVKVDVKEIGADFYVSNLHKWFFCPTT
Aquilegia_sp_TC26581	(210)	ELVQICRKEGVDQVFVDA	AAHAIGSVKVDVKEIGADFYVSNLHKWFFCPTT
P.persica_TC12771	(252)	QLVKICRDEGVEQVFVDA	AAHAIGSVKVDVKEIGADFYVSNLHKWFFCPTT
S.lycopersicum_TC191918	(220)	DLVKICREEGVEQVFVDA	AAHAIGSVKVDVKEIGADFYVSNLHKWFFCPTT
S.tuberosum_TC167698	(220)	DLVKICREEGVEQVFVDA	AAHAIGSVKVDVKEIGADFYVSNLHKWFFCPTT
Consensus	(301)	ELVKICREEGVDQVFVDA	AAHAIGSVKVDVKEIGADFYVSNLHKWFFCPTT

FIGURE 9 (continued)

61/77

	351	400
A.lyrata 486720	(265) IAFFYCKKRG----SESDVHHPPVVSHEFGNGLPIESAWIGTRDYSSQLVV	
A.thaliana_AT3G62130.1	(265) IAFFYCKKRG----SESDVHHPPVVSHEFGNGLPIESAWIGTRDYSSQLVV	
C.clementina_TC23968	(261) VVFLYCRKSI----LSSDMHHPGGFT-----	
P.trichocarpa 572528	(265) VSFLYCKKAS----LEFDVHHPPVVSHEYGNGLPIESAWIGTRDYSSQLVV	
Popttr_epimerase-related	(269) VAFLYCKKAAS----LEFDVHHPPVVSHEYGNGLPIESAWIGTRDYSSQLVV	
G.hirsutum TC132670	(263) VAFLYCNKST----TSSDLHHPVIVSSSA-----	
V.vinifera_GSVIVT00028240001	(263) VAFLYCRKSP----LSSEVHHPPVVSHEFGNGLPIESSWIGTRDYSSQLVV	
G.max_Glyma02g01070.1	(262) VAFLYCKE-----KSNDVHHPPVVSQEYKGKLPVESAWVGMRDYSPLVV	
M.truncatula AC151948 16.5	(262) VAFMYCNKNK----KLNDVHHPPVVAHEYGNGLPAESAWVGMRDYSPLVV	
H.vulgare TC158694	(270) VAFLHTRKGG---PIASQLHHPVVSHEYGNGLPMESGWIGTRDYSAQLVV	
T.aestivum TC278728	(270) VAFLHTRKGG---PITSQLHHPVVSHEYGNGLPMESGWIGTRDYSAQLVV	
O.sativa LOC_Os01g18640.1	(284) VAFLHTRKDD---PIASQLHHPVVSHEYGNGLPMESGWIGTRDYSAQLVV	
S.bicolor_Sb03g011840.1	(281) VAFLHTRKD---PIASQLHHPVVSHEYGNGLPMESGWIGTRDYSAQLVV	
S.bicolor_Sb03g011840.3	(281) VAFLHTRKD---PIASQLHHPVVSHEYGNGLPMESGWIGTRDYSAQLVV	
S.bicolor_Sb03g011840.2	(281) VAFLHTRKD---PIASQLHHPVVSHEYGNGLPMESGWIGTRDYSAQLVV	
Zea_mays_GRMZM2G322186 T01	(274) VAFLHTRKDD---PIASQLHHPVVSHEYGNGLPMESGWIGTRDYSAQLVV	
Z.mays TC513613	(274) VAFLHTRKDD---PIASQLHHPVVSHEYGNGLPMESGWIGTRDYSAQLVV	
O.sativa LOC_Os01g18660.1	(278) VAFIHTRKDD---PVSSKLHHPVVSSEYGNGLPMESAWIGVRDYSAQLVV	
S.bicolor_Sb03g011850.1	(263) VAFLHIRKDD---PVAQLHHPVVSSEYGNGLPMESAWIGVRDYSAQLVV	
Z.mays_ZM07MC24472_BF50052A17024402	(263) VAFLHIRKDD---PVAQLHHPVVSSEYGNGLPMESAWIGVRDYSAQLVV	
Zea_mays_GRMZM2G077212 T01	(263) VAFLHIRKDD---PVAQLHHPVVSSEYGNGLPMESAWIGVRDYSAQLVV	
P.patens 175806	(341) AAFVHCKAKH-----LGRLLHHPVVSHEYGNGLPIESSWIGTRDYSSQLAV	
P.patens TC39863	(341) AAFVHCKAKH-----LGRLLHHPVVSHEYGNGLPIESSWIGTRDYSSQLAV	
S.moellendorffii_232064	(246) VAFLLHAKAKH-----LERLHHPVVSHEYNFGLFSECSWVGTRDYSSQLAV	
S.moellendorffii_413706	(261) VAFLLHAKAKH-----LERLHHPVVSHEYNFGLFSECSWVGTRDYSSQLAV	
A.lyrata 489502	(285) VAFLYCRKSSNGGGVADLHHPVVSNEYGNGLAVESSWVGTRDYSAQLVV	
A.thaliana_AT5G26600.1	(292) VAFLYCRKSSNG--GVADLHHPVVSNEYGNGLAVESSWVGTRDYSAQLVV	
B.napus TC63646	(290) VAFLYCRKSGAG----GNLHHPVVSHEYGNGLAVESTWVGTRDYSAQLVV	
C.sinensis TC2568	(263) AAFLYCRKSNE----IDDLHHPVVSHEYGNGLPIESSWIGTRDYSAQLVV	
P.trichocarpa 570656	(267) VAFLYCRKRGEDGKG-GDLHHPVVSHEYGNGLAVESAWIGTRDYSAQLVV	
G.max_Glyma03g38740.1	(265) IAFLYTRRNLKGTDPGSDLHHPVVSHEYGNGLAVESAWIGTRDYSAQLVV	
G.max_Glyma19g41330.1	(268) IAFLYTRRNFKGTSKSGSDLHHPVVSHEYGNGLAVESAWIGTRDYSAQLVV	
M.truncatula AC166315 9.4	(271) IAFLYTKKNPK-TGGGGDLHHPVVSHEYGNGLAVESAWIGTRDYSAQLVV	
G.max_Glyma20g22430.1	(268) VAFLYARASSK---AR-DLHHPVVSHEYGNGLAVESSWTCNRDYSAQLVV	
Aquilegia sp TC26581	(260) VAFLYCKKSD----KLSDLHHPVVSNEYGNGLPIESSWIGTRDYSSQLVV	
P.persica TC12771	(302) VSFLYCRKSATHS--DLELHHPVVSHEYGNGLAVESFWVGTRDYSPYLVL	
S.lycopersicum_TC191918	(270) VAFLYCRKSP----VSPDLHHPVVSHEYGNGLPIESSWIGTRDYSSQLVI	
S.tuberosum TC167698	(270) VAFLYCRKSP----VSPDLHHPVVSHEYGNGLPIESSWIGTRDYSSQLVI	
Consensus	(351) VAFLYCRK SDLHHPVVSHEYGNGL MESAWIGTRDYSAQLVV	

FIGURE 9 (continued)

62/77

A.lyrata_486720	(311)	PSVMEFVNRFEGGIEGIMIRNHD EAVRMGLMLADAWGTNLGSPPEMCGVM	401		450
A.thaliana_AT3G62130.1	(311)	PSVMEFVNRFEGGMEGIMMKNHDEAVRMGLMLADAWGTNLGSPPEMCGVM	(311)		
C.clementina_TC23968	(283)	-----	(283)		
P.trichocarpa_572528	(311)	PAALEFVNRFEDGIGHGIMKRNHEEVVKMGKMLAESWRTNLGSPPEMCGVM	(311)		
Popttr_epimerase-related	(316)	PAALEFVNRFEDGIGIMKRNHEEVVKMGKMLAESWGTNLGSSPEMCGVM	(316)		
G.hirsutum_TC132670	(288)	-----	(288)		
V.vinifera_GSVIVT00028240001	(309)	PSVLEFVNRFEGGIEGIMMRNHEIVVKMGEMLAESWGTNLGAPPEMCGVM	(309)		
G.max_Glyma02g01070.1	(306)	PSILEFVNRFEGGIEGIMKRNHDGVVKMGTM LAESWGTVLGSPPEMCGVM	(306)		
M.truncatula_AC151948_16.5	(308)	PSIMEFVNRFEGGIEGIMKRNHNMVVKMGVMLKEAWGTNLGSPPEMCGVM	(308)		
H.vulgare_TC158694	(317)	PEAIDFVNRFEGGIEGIRSRNHEKVVEMGRMLAEAWGTFLGSPPEMCGVM	(317)		
T.aestivum_TC278728	(317)	PEAIHFVNRFEGGIEGIRSRNHEKVIEMGRMLAEAWGTFLGSPPEMCGVM	(317)		
O.sativa_LOC_Os01g18640.1	(331)	PESIDFVNRFEGGIEGIRSRNHEKVIEMGRMLAEAWGTFLGTPPELCCGM	(331)		
S.bicolor_Sb03g011840.1	(327)	SEAIDFVNRFEGGIEGIRSRNHEKVIEMGMMLAEAWGTFLGTPPELCCGM	(327)		
S.bicolor_Sb03g011840.3	(327)	SEAIDFVNRFEGGIEGIRSRNHEKVIEMGMMLAEAWGTFLGTPPELCCGM	(327)		
S.bicolor_Sb03g011840.2	(327)	SEAIDFVNRFEGGIEGIRSRNHEKVIEMGMMLAEAWGTFLGTPPELCCGM	(327)		
Zea_mays_GRMZM2G322186_T01	(321)	SEAIDFVNRFEGGIEAIRNRNHEKVIEMGRMLAEAWGTFLGSPPELCCGM	(321)		
Z.mays_TC513613	(321)	SEAIDFVNRFEGGIEAIRNRNHEKVIEMGRMLAEAWGTFLGSPPELCCGM	(321)		
O.sativa_LOC_Os01g18660.1	(325)	PDVDFVNRFEGGIEGIRSRNHDGVVKMGTM LAEAWGTFLGTPPEMCGVM	(325)		
S.bicolor_Sb03g011850.1	(310)	PDAVDFMRFEFGGIEAIRSRNHDGVVKMGTM LAEAWGTFLGSPPEMCGVM	(310)		
Z.mays_ZM07MC24472_BF50052A17024402	(310)	PDAVDFMRFEFGGIEAIRSRNHDGVVKMGTM LAEAWGTFLGSPPEMCGVM	(310)		
Zea_mays_GRMZM2G077212_T01	(310)	PDAVDFMRFEFGGIEAIRSRNHDGVVKMGTM LAEAWGTFLGSPPEMCGVM	(310)		
P.patens_175806	(386)	PAAIKFVIDVAGSIENYSKFNHCKVVMAM EMLASSWGTFLGTPPEMCAAM	(386)		
P.patens_TC39863	(291)	TAAVEFIDGNLGLSSLI AFNHRGAMAMAEMLAAAWGTQCGTCAELASSM	(291)		
S.moellendorffii_232064	(306)	TAAVEFIDGNLGLSSLI AFNHRAMAMAEMLAAAWGTQCGTCAELAQGV	(306)		
S.moellendorffii_413706	(335)	PSILEFVNRFEGGIDGIIKRNHESVVMGQMLVKS WGTQLGCPPEMCGVM	(335)		
A.lyrata_489502	(340)	PSILEFVNRFEGGIDGIIKRNHESVVMGQMLVKS WGTQLGCPPEMCGVM	(340)		
A.thaliana_AT5G26600.1	(336)	PSILEFVNRFEGGIDGIIKRNHESVVMGQMLVKS WGTQLGCPPEMCPVM	(336)		
B.napus_TC63646	(309)	PQALEFVNRFEGGIEGIIKRNHKA VVEMGEMLAKAWGTHLGSPPEMCGVM	(309)		
C.sinensis_TC2568	(316)	PAVLEFFNRFEGGIEGIIKRNHKA VVEMGEMLVKA WGTNLGSPPEMCGVM	(316)		
G.max_Glyma03g38740.1	(315)	PAAVEFVNRFEGGIEGIIKRNHKA VVEMGEMLAKAWGTRLGSPPEMCGVM	(315)		
G.max_Glyma19g41330.1	(318)	PAAVEFVNRFEGGIEGIIKRNHKA VVEMGEMLVKA WGTNLGSPPEMCGVM	(318)		
M.truncatula_AC166315_9.4	(320)	PEALEFVNRFEGGIEGIIKRNHKA VVEMGEMLVKA WGTNLGSPPEMCGVM	(320)		
G.max_Glyma20g22430.1	(314)	PAVMEFVKRFEGGIEGIRKRNHD LVVEMGEMLV EAWGTHLGSPSHMSASM	(314)		
Aquilegia_sp_TC26581	(306)	PSVLEFVNRFEGGIEGIIKRNHDA AIE MGKMLAKAWGTNLGSPPELCCSM	(306)		
P.persica_TC12771	(350)	PSVMEFVNRFEGGLKGI I KRNHDA VVEMGKMLAEAWGTNLGCPPEMCGVM	(350)		
S.lycopersicum_TC191918	(316)	PEVLEFINRFEGGIEGIRLRNHKAVIEMGQMLANAWGTSLGCPPEMSPGM	(316)		
S.tuberosum_TC167698	(316)	PEVLEFINRFEGGIEGIRLRNHKAVIEMGQMLANAWGTSLGCPPEMSPGM	(316)		
Consensus	(401)	PAALEFVNRFEGGIEGIRKRNHE VVEMG MLAEAWGT LGSPPEMCGVM	(401)		

FIGURE 9 (continued)

63/77

	451	500
A.lyrata_486720	(361) VMIGLPSKLCVSEDEDATKLRSLRVHYSVEVPVYFLG---LRDGEEGV-	
A.thaliana_AT3G62130.1	(361) VMIGLPSKLCVGSDEDAIKLRSLRVHYSVEVPVYFLG---LRDGEEGV-	
C.clementina_TC23968	(283) -----	
P.trichocarpa_572528	(361) IMVGLPSRLSVSSEDDASRLRSHLRDCHGVEVPPIHYQG---LRDGEEGV-	
Popttr_epimerase-related	(366) IMVGLPSRLRVSSEDDALRLRSHLRCHGVEVPPIHYQG---LKDGEEGV-	
G.hirsutum_TC132670	(288) -----	
V.vinifera_GSVIVT00028240001	(359) IMVGLPSRLFISSEEDAMRLRSLRQHHGIEVPLHYQA---PSD-----	
G.max_Glyma02g01070.1	(356) IMVGLPSRLRVMSVDDALRLRSLRVYHAEVVPVYQA---LRNGDRDP-	
M.truncatula_AC151948_16.5	(358) IMIGLPSKIRVMSDDDALRLRFYLRVYHAEVVPVYQA---LGNGERDA-	
H.vulgare_TC158694	(367) VMVGMPSCLCIESDDDALRVRTMLRKDFKVEVPPIYYNT--RQVEVQE---	
T.aestivum_TC278728	(367) AMVGMPSCLCIESDDDALRVRTMLRKDFKVEVPPIYYNS--RQVKVQE---	
O.sativa_LOC_Os01g18640.1	(381) VMVGLPGCLGVESDDDMVRMTMLRKDFMVEVPPIYYNS--RRVEAQE---	
S.bicolor_Sb03g011840.1	(377) VMVGMPGCLGVESDDDALRVRTMLRKDFQVEVPPIYYNS--RRVEGQE---	
S.bicolor_Sb03g011840.3	(377) VMVGMPGCLGVESDDDALRVRTMLRKDFQVEVPPIYYNS--RRVEGQE---	
S.bicolor_Sb03g011840.2	(377) VMVGMPGCLGVESDDDALRVRTMLRKDFQVEVPPIYYNS--RRVEGQE---	
Zea_mays_GRMZM2G322186_T01	(371) VMVGMPGCLGVESDDAMRVRTMLRKDFQVEVPVYYNS--RRVEGQE---	
Z.mays_TC513613	(371) VMVGMPGCLGVESDDAMRVRTMLRPLVRTSLLSQLMRGGHPVLSWM---	
O.sativa_LOC_Os01g18660.1	(375) LMVGLPGSLGVGSDDAVGLRTMLRKQFKVEVPPIYYNS--KAAAADAPPE	
S.bicolor_Sb03g011850.1	(360) AMVGLPGCLAIESGGDAMRVDRMLRNEFKVEVPPIFHNS--RSVEEGQE---	
Z.mays_ZM07MC24472_BF50052A17024402	(360) AMVGLPGCLGIESDDAMRVDRMLRNQFKVEVPPIFNNS--RSVEEGQ---	
Zea_mays_GRMZM2G077212_T01	(360) AMVGLPGCLGIESDDAMRVDRMLRNEFKVEVPPIFNNS--RSVEEGQ---	
P.patens_175806	(436) AMVALPPALNIHSQVDLLALRRRIREEYQVDVHLYYAGSLAPANIDNAS-	
P.patens_TC39863	(436) AMVALPPALNIHSQVDLLALRRRIREEYQVDVHLYYAGSLAPANIDNAS-	
S.moellendorffii_232064	(341) AMVELPAKLEVASEEAALEMRTLRREFGVEVPICFQAGGSPS--ANS---	
S.moellendorffii_413706	(356) WSGGADLLPGGGIAQREFEEQEIRDGGWSENAMDFCRSSRIP--WRS---	
A.lyrata_489502	(385) IMVGLPVYLGVSDDSVLKLRTFLREKFRIEIPITYFRP---PEDG-----	
A.thaliana_AT5G26600.1	(390) IMVGLPVCLGVSSSDVLKLRTFLREKFRIEIPITYFRP---PGDG-----	
B.napus_TC63646	(386) VMVGLPVCLGVSSSETDAVRLRNLLRERFSIEIPITYFRA---PGE-----	
C.sinensis_TC2568	(359) IMVGLPAILGISSDSIALKLRTHLRNNFSVEVPPIYYRV---PGDG-----	
P.trichocarpa_570656	(366) IMVGLPACLGISNDLDTLKLRLSHLRHEHFQVEVPPIYFRA---PVDG-----	
G.max_Glyma03g38740.1	(365) VMVGLPACLGIGSDSDALKLRTHLRDAFGVEVPPIYYRS---PREG-----	
G.max_Glyma19g41330.1	(368) VMVGLPACLGIESDSALKLRTHLRDFRDFGVEVPPIYYRP---PKEG-----	
M.truncatula_AC166315_9.4	(370) VMVGLPTCFGVRSDSDALTLRTHLRDVFGEVPPIYYRP---PRDG-----	
G.max_Glyma20g22430.1	(364) VMVGLPPSLGIGSDSDAQKLRTLRDEFDVEVPPIYFR-----GG-----	
Aquilegia_sp_TC26581	(356) VMVGLPSSLEISSEADGMLFRTHLRRESFGIEVPPIYYKD---PKSG-----	
P.persica_TC12771	(400) IMVGLPACLGISSDDATKLRPHLCDKFGVEVRIHYQA---PKDG-----	
S.lycopersicum_TC191918	(366) AMVGLPVNLKILSDKDALNLRNHLRDFHFAVEVPPIHYEE---IKELQDG---	
S.tuberosum_TC167698	(366) AMVGLPVNLRLISDKDALTLRNHLRDFHFGVEVPPIHYEE---IKELQDG---	
Consensus	(451) VMVGLP LGV SDDDALRLRT LR F VEVPIYY EG	

FIGURE 9 (continued)

64/77

```

A.lyrata_486720 (407) --KDKDSSLITAYVRISHQIYNKTEDYERLRDAITELVKDQMTCCQNLPSG
A.thaliana_AT3G62130.1 (407) --KDKDSSLITAYVRISHQVYNKTEDYERLRDAITELVKDQMTCCQNLPAL
C.clementina_TC23968 (283) -----
P.trichocarpa_572528 (407) --KDKDG-VITAYARISHQIYNKFEDYCRFDDAVNHLAENRQIRRTISPE
Poptr_epimerase-related (412) --KDKDG-VITAYARISHQVYNKSEDYCKLRDAVNRLSENLLIRKTFYPE
G.hirsutum_TC132670 (288) -----
V.vinifera_GSVIVT00028240001 (400) ---DKDG-LVTGYARISHQVYNFDDYCKFRDAINQLVEQRRSCKMLFME
G.max_Glyma02g01070.1 (402) --RDKDG-FITGYVRISHQVYNNTADDYERLKTAINQLVEDGKVCSGLPKE
M.truncatula_AC151948.16.5 (404) --RDKDG-FITGYVRISHQVYNIVDDYNRLKTAIQLLQDGKICSELPKE
H.vulgare_TC158694 (412) IAKDNNSDPVTGYVRISHQVYNVKEEYERLRDAVNKLVAEGFTSAELRPA
T.aestivum_TC278728 (412) MAKDNSSDPVTGYVRISHQVYNVKEEYERLRDAVNKLVAEGFTSAELRPA
O.sativa_LOC_Os01g18640.1 (426) MAKDKNGDAVTGYVRISHQVYNVTEDYERLRDAVNKLVDGFTSSKLRPS
S.bicolor_Sb03g011840.1 (422) MAKDRNGDPVTGYVRISHQVYNVTKDYERLRDAVNKLVDGFTSSKLRPS
S.bicolor_Sb03g011840.3 (422) MAKDRNGDPVTGYVRISHQVYNVTKDYERLRDAVNKLVDGFTSSKLRPS
S.bicolor_Sb03g011840.2 (422) MAKDRNGDPVTGYVRISHQVYNVTKDYERLRDAVNKLVDGFTSSKLRPS
Zea_mays_GRMZM2G322186.T01 (416) MTKDKNGDPVTGYVRISHQVYNVTEDYERLRDAVNKLVDGFTSSKLRPS
Z.mays_TC513613 (418) LALSGLQQGTVSLVPSRYISYVCYHDIILVSLFVSPSRIIIIISWACCMS
O.sativa_LOC_Os01g18660.1 (423) MVKDGNGDPVTGYVRISHQVYNVREEYEALDAVAKLVADGFTCRKLRPP
S.bicolor_Sb03g011850.1 (406) LAKDAKGDQVSGYVRISHQVYNVREEYETLRDAVHKLVLDGFTSCSKMRPS
Z.mays_ZM07MC24472_BF50052A17024402 (405) --EMAKGDQVTGYVRISHQVYNVTEEYEVLRDAVGKLVLDGFTSCKLRPS
Zea_mays_GRMZM2G077212.T01 (405) --EMAKGDQVTGYVRISHQVYNVTEEYEVLRDAVGKLVLDGFTSCKLRPS
P.patens_175806 (485) ----QGRTRTTAYVRISHQIYNNSEYIKLRDVVNDIARKS-----
P.patens_TC39863 (485) ----QGRTRTTAYVRISHQIYNNSEYIKLRDVVNDIARKS-----
S.moellendorffii_232064 (387) ----KSMRSGTAYARVSHQIYNKLEDYTRLRDAVKSLAR-----
S.moellendorffii_413706 (402) ----AQCCSSTVVNL-----
A.lyrata_489502 (427) ----EIDPITGYVRISHQVYNKPEDYHRLRDAINELVLDGFRGFCASLSS-
A.thaliana_AT5G26600.1 (432) ----EIDPITGYVRISHQVYNKPEDYHRLRDAINELVLDGFRGFCASLSS-
B.napus_TC63646 (428) ----ETDSITGETSYFEGE-----
C.sinensis_TC2568 (401) ----VVNPVTGYARISHHVYNKPDEYYKFRDAVNQLVNDKFTCALLAW-
P.trichocarpa_570656 (408) ----EVDITGYARISHQVYNKAEDYYRFRDAVNQLVSDGFTCASLSN-
G.max_Glyma03g38740.1 (407) ----EVGVVTGYARISHQVYNKVDDYYKFRDAVNQLVQNGFTCVVLSDD
G.max_Glyma19g41330.1 (410) ----EVGVVTGYARISHQVYNKVDDYYKFRDAVNQLVQNGFTCAVLSD-
M.truncatula_AC166315.9.4 (412) ----EVDVPTGYARISHQVYNKVDDYYKFRDAVNQLVQNGFTCAVLSD-
G.max_Glyma20g22430.1 (403) ----EDGSVTAYARISHQVYNKVDDYYKFRDAVNQLVQNGFTCALLSA-
Aquilegia_sp_TC26581 (398) ---NIEDGVTTGYVRISHQIYNTIDDDYNRLRDAINQLIQDGFTCKSLLSR
P.persica_TC12771 (442) ----EVGSTTGYVRISHQIYNKVDDYYKFRDAVNQLVLDGFTCALEINE
S.lycopersicum_TC191918 (411) ----DG-YVTGYARISHQVYNKVDDYIKLKDAILQLVLDGVTCKMLHSE
S.tuberosum_TC167698 (411) ----DG-YVTGYARISHQVYNKVDDYIKLKDAILQLVLDGVTCKMLPSE
Consensus (501) E VTGYVRISHQVYN EDY RLRDAVN LV DGFTC L

```

FIGURE 9 (continued)

65/77

	551	565
A.lyrata_486720	(455)	-----
A.thaliana_AT3G62130.1	(455)	-----
C.clementina_TC23968	(283)	-----
P.trichocarpa_572528	(454)	-----
Poptr_epimerase-related	(459)	-----
G.hirsutum_TC132670	(288)	-----
V.vinifera_GSVIVT00028240001	(446)	-----
G.max_Glyma02g01070.1	(449)	-----
M.truncatula_AC151948_16.5	(451)	-----
H.vulgare_TC158694	(462)	EKQETLA-----
T.aestivum_TC278728	(462)	EKQEALA-----
O.sativa_LOC_Os01g18640.1	(476)	QKQETMA-----
S.bicolor_Sb03g011840.1	(472)	EKPGIIHLKALWPTT
S.bicolor_Sb03g011840.3	(472)	EKQEA-----
S.bicolor_Sb03g011840.2	(472)	EKVSST-----
Zea_mays_GRMZM2G322186_T01	(466)	EKQEA-----
Z.mays_TC513613	(468)	HLLW-----
O.sativa_LOC_Os01g18660.1	(473)	EKEETLA-----
S.bicolor_Sb03g011850.1	(456)	GKALSN-----
Z.mays_ZM07MC24472_BF50052A17024402	(453)	EKVISN-----
Zea_mays_GRMZM2G077212_T01	(453)	EKVISN-----
P.patens_175806	(522)	-----
P.patens_TC39863	(522)	-----
S.moellendorffii_232064	(422)	-----
S.moellendorffii_413706	(413)	-----
A.lyrata_489502	(471)	-----
A.thaliana_AT5G26600.1	(476)	-----
B.napus_TC63646	(443)	-----
C.sinensis_TC2568	(445)	-----
P.trichocarpa_570656	(452)	-----
G.max_Glyma03g38740.1	(452)	-----
G.max_Glyma19g41330.1	(454)	-----
M.truncatula_AC166315_9.4	(456)	-----
G.max_Glyma20g22430.1	(447)	-----
Aquilegia_sp_TC26581	(445)	-----
P.persica_TC12771	(487)	A-----
S.lycopersicum_TC191918	(455)	-----
S.tuberosum_TC167698	(455)	-----
Consensus	(551)	-----

FIGURE 9 (continued)

66/77

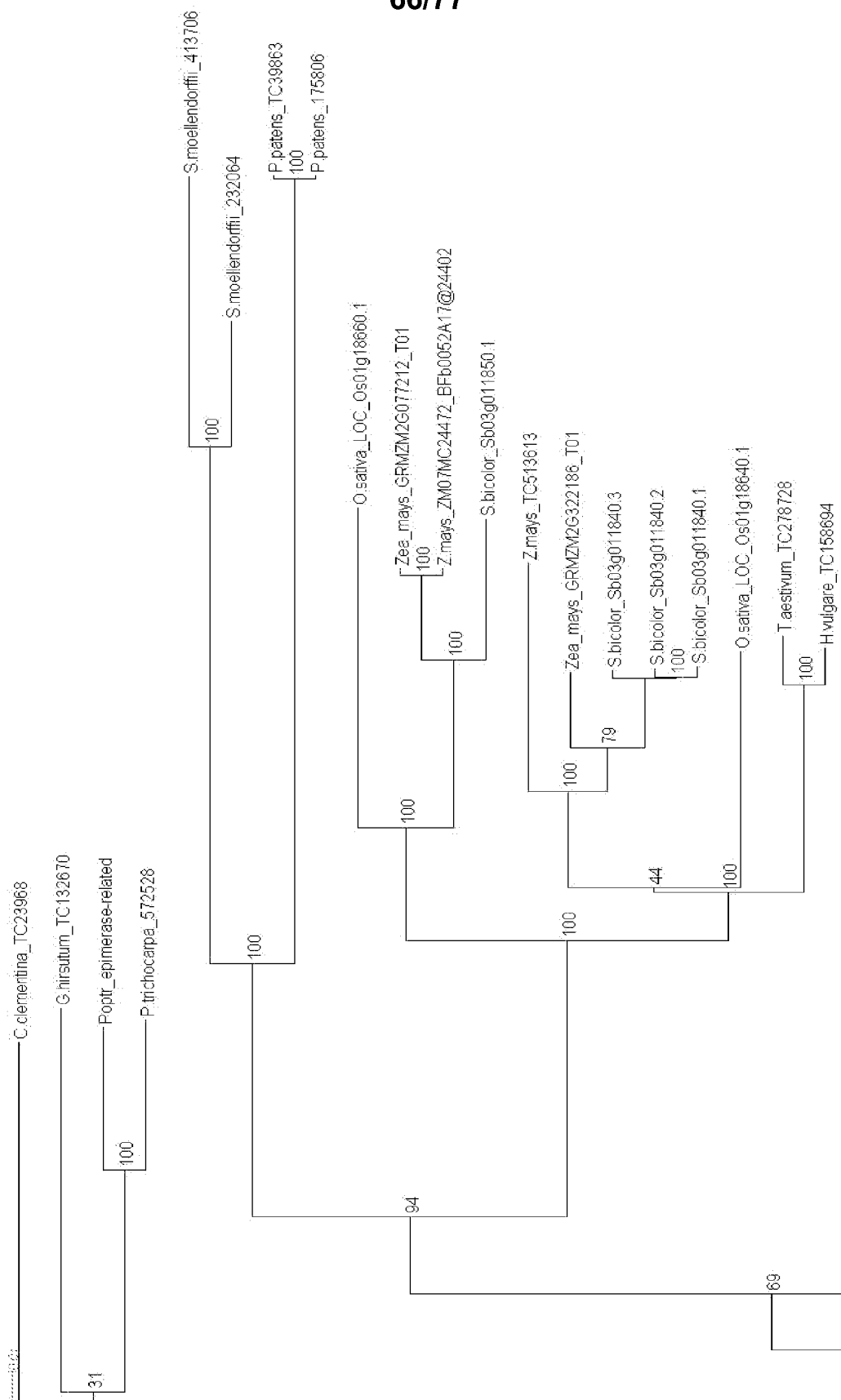


FIGURE 10

67/77

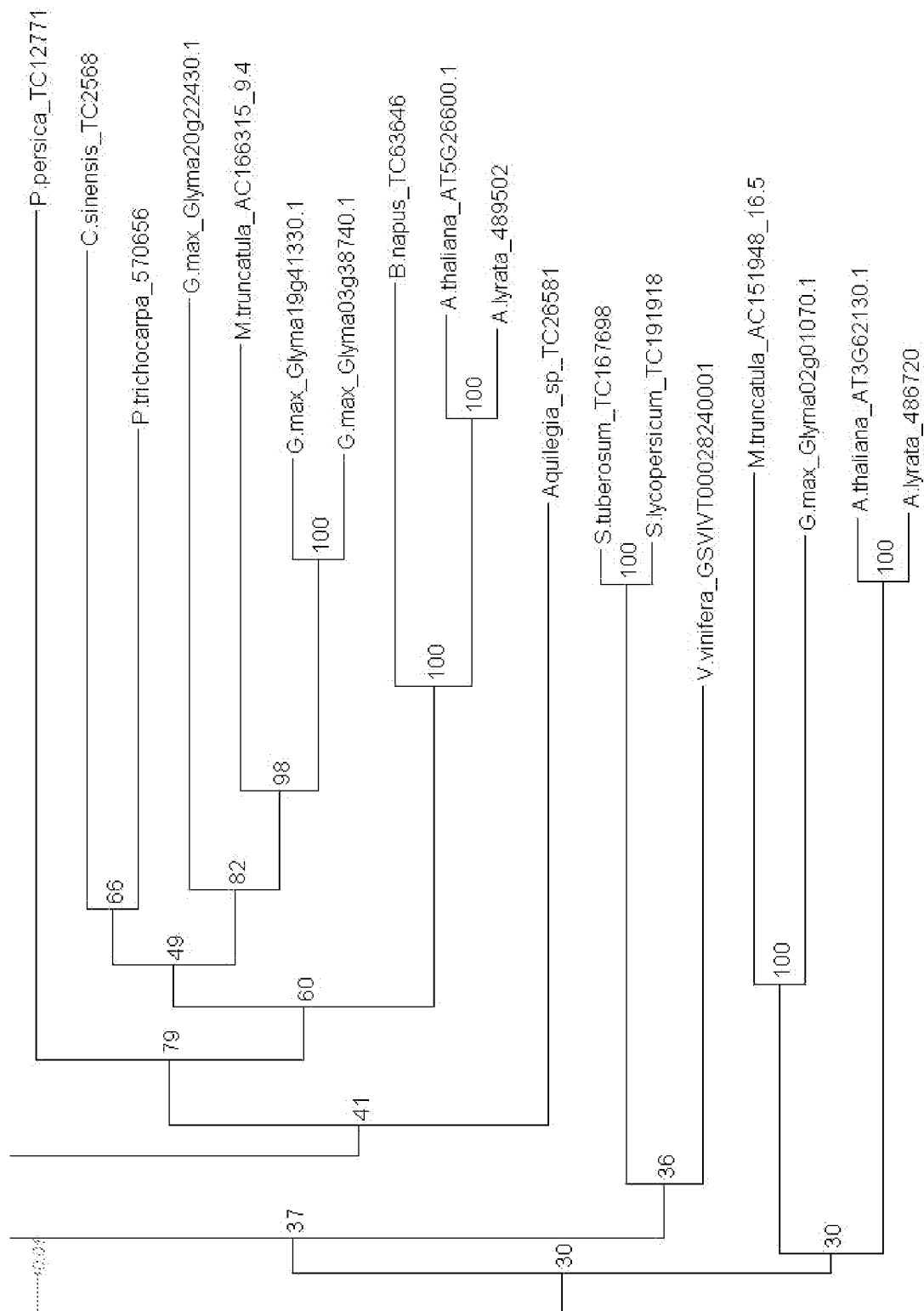


FIGURE 10 (continued)

68/77

	1	2	3	4	5	6	7	8	9	10	11	12	13
1		68.8	60.2	69.3	59.8	66.8	55.7	47.8	60.3	47.5	66.6	63.9	62.5
2	85.8		58.2	96.9	56.9	62.3	53.4	41.4	57.6	43.7	65.0	61.9	60.9
3	75.1	73.8		57.4	93.3	65.5	75.2	37.4	64.0	41.7	58.8	67.7	67.6
4	85.8	99.1	74.5		56.0	61.5	52.6	42.1	57.8	44.0	65.2	61.9	60.6
5	74.9	72.4	95.8	72.8		64.6	75.9	37.3	62.3	39.6	58.2	66.1	66.7
6	80.6	80.4	80.0	80.4	78.3		58.6	43.8	63.8	45.4	63.7	69.3	66.4
7	72.5	70.5	84.3	70.7	84.0	75.7		39.6	57.3	42.5	54.2	60.1	59.4
8	54.1	53.3	47.4	53.5	46.5	52.7	50.2		36.7	72.1	43.6	39.7	38.7
9	77.1	76.7	77.7	76.2	76.6	83.3	75.0	50.5		37.8	57.4	71.8	70.3
10	55.5	54.8	50.2	54.8	47.8	54.7	52.0	86.1	49.8		42.8	42.6	40.8
11	80.8	81.5	72.6	81.7	71.8	79.5	70.8	52.7	75.9	54.5		59.9	58.1
12	80.3	79.7	81.1	80.4	80.0	85.6	78.0	50.6	84.0	51.7	76.9		91.4
13	80.1	78.9	81.1	79.1	80.0	83.9	76.8	49.9	83.0	50.6	75.5	95.4	
14	76.0	76.7	78.1	77.1	76.6	81.2	74.0	50.2	78.5	52.2	74.6	84.9	83.0
15	74.4	72.9	74.7	73.1	72.4	73.1	69.9	44.9	71.2	45.3	70.9	75.9	74.4
16	81.2	81.1	72.8	81.3	71.8	79.8	70.2	52.0	75.6	53.8	87.1	77.8	77.0
17	80.3	77.4	81.9	77.8	81.3	84.2	78.7	49.9	82.4	51.4	76.9	90.1	88.8
18	71.6	71.6	72.6	71.8	72.0	71.6	69.1	45.0	70.5	44.2	70.1	74.3	72.8
19	71.0	69.5	71.0	70.1	70.4	71.4	66.0	43.2	67.0	44.3	69.3	72.0	70.4
20	59.5	56.2	57.6	56.8	58.3	58.0	54.7	38.0	55.7	40.1	56.2	55.9	55.5
21	59.5	56.0	57.6	56.6	58.3	58.0	54.7	38.0	55.7	40.1	56.0	55.9	55.5
22	72.5	70.4	73.7	69.8	72.3	73.1	67.8	44.8	72.3	46.6	69.2	77.6	76.2
23	80.1	76.9	80.6	76.9	79.2	82.9	79.4	50.8	81.2	50.1	76.9	85.6	84.8
24	94.3	85.2	73.4	85.2	73.1	80.8	70.9	54.5	77.5	56.1	79.2	81.0	80.1
25	71.8	70.6	71.4	70.6	70.8	70.2	67.3	43.4	68.5	43.8	67.9	73.0	71.6
26	72.3	71.9	72.7	72.1	72.1	71.5	68.3	44.2	69.8	44.9	69.4	74.6	73.0
27	72.9	72.1	72.7	72.3	72.1	71.4	68.7	44.5	69.7	45.0	69.5	74.4	72.9
28	70.7	70.7	71.3	71.1	69.5	70.9	68.3	44.3	70.3	44.7	70.1	74.0	72.5
29	83.0	81.5	76.0	81.7	74.5	83.3	73.6	54.0	78.9	54.8	79.7	82.6	81.3
30	63.8	63.2	59.6	63.4	58.3	65.8	59.3	46.3	65.1	45.1	64.1	64.1	62.7
31	58.1	57.7	54.5	57.9	54.3	60.6	56.3	48.3	59.0	47.3	58.3	57.9	57.8
32	83.6	81.9	76.0	82.2	74.7	83.3	73.8	53.5	78.4	55.3	79.7	82.6	81.1
33	73.5	72.0	73.8	72.4	72.0	72.6	69.2	44.7	70.5	45.3	70.7	74.8	74.4
34	85.2	84.1	75.1	84.4	74.9	85.4	74.6	56.2	80.2	56.2	83.9	83.1	81.5
35	69.4	66.7	68.8	67.9	68.2	69.6	67.1	46.1	66.5	46.1	65.8	71.3	69.9
36	71.6	71.4	70.9	71.6	70.3	71.8	69.0	44.5	71.0	46.1	70.5	75.3	73.4
37	71.6	71.2	70.9	71.4	70.3	71.8	68.8	44.5	71.0	46.1	70.5	75.3	73.4
38	74.5	72.6	73.8	73.0	72.4	73.0	68.9	46.2	71.1	46.4	71.5	75.1	73.4

FIGURE 11

69/77

	14	15	16	17	18	19	20	21	22	23	24	25	26	27
1	60.5	58.7	67.9	62.9	56.6	53.1	40.4	40.2	55.4	65.2	89.5	56.5	57.3	57.4
2	57.0	57.4	63.5	58.5	55.4	53.5	38.1	38.1	52.5	60.8	69.2	54.3	55.6	55.7
3	63.4	58.2	57.4	67.9	57.7	52.3	38.3	38.3	56.8	70.1	60.2	57.1	58.4	58.3
4	56.8	57.2	64.0	58.4	55.0	54.0	38.1	38.1	52.5	60.1	69.7	54.3	55.3	55.5
5	61.5	56.3	56.8	66.9	57.0	51.9	38.3	38.3	56.6	68.3	59.1	56.3	57.6	57.5
6	62.3	59.5	62.0	65.9	59.3	53.7	37.8	37.8	57.3	68.5	66.7	58.6	60.0	59.9
7	56.7	52.6	52.8	61.0	54.5	48.5	36.3	36.3	50.9	64.2	55.6	53.5	54.3	54.6
8	39.2	37.2	43.2	38.5	35.9	31.9	24.5	24.5	33.3	39.8	46.8	35.6	36.3	36.6
9	63.9	58.1	57.4	68.4	56.2	51.8	37.5	37.5	57.9	71.0	60.3	55.1	56.0	56.1
10	41.1	37.0	43.1	40.0	35.5	32.4	25.6	25.6	35.5	40.8	48.2	36.2	37.1	37.0
11	57.3	56.7	76.7	58.4	55.4	52.5	38.2	38.2	51.1	59.5	67.2	54.9	56.2	56.3
12	73.2	63.1	60.8	80.7	61.1	57.3	38.5	38.5	60.3	73.9	64.1	62.1	63.7	63.4
13	72.2	61.7	59.0	79.0	59.0	55.3	38.1	38.1	60.5	73.1	61.6	59.4	60.7	60.6
14		57.8	56.5	70.1	57.1	53.7	38.4	38.4	58.9	68.7	60.3	57.2	58.5	58.6
15	70.9		53.8	60.7	82.2	73.9	38.7	38.7	50.3	62.8	58.3	80.7	82.0	82.6
16	75.1	69.9		58.7	52.5	49.5	35.3	35.3	51.7	58.9	67.1	52.1	53.2	53.4
17	83.1	75.2	76.7		59.3	55.1	36.7	36.7	58.6	73.9	62.6	58.7	60.0	59.9
18	69.9	90.2	68.7	73.0		71.9	40.0	40.0	50.7	60.3	56.2	84.3	85.7	86.3
19	67.4	84.8	68.1	72.9	83.8		38.4	38.4	49.6	55.9	53.1	71.4	72.5	73.0
20	55.7	55.7	55.9	56.0	57.2	55.9		99.8	37.8	39.3	38.8	38.1	38.7	38.8
21	55.7	55.7	55.9	56.0	57.0	55.9	100.0		37.8	39.3	38.8	38.1	38.7	38.8
22	72.3	67.4	70.2	74.5	68.4	67.8	58.0	58.0		58.3	55.0	49.5	50.0	50.7
23	80.3	75.2	76.7	85.5	73.9	73.3	58.0	57.8	72.5		64.0	59.8	61.1	61.0
24	75.7	73.1	81.0	79.3	71.4	69.1	57.6	57.6	71.3	78.8		55.9	57.0	56.9
25	68.5	87.0	66.5	72.6	89.5	82.7	56.6	56.4	67.4	72.6	69.8		97.5	97.3
26	69.8	88.7	67.9	74.0	90.2	83.9	56.6	56.4	67.6	74.0	71.1	97.7		99.2
27	70.0	89.3	68.1	73.9	90.7	84.3	56.8	56.6	67.4	73.9	70.8	97.3	99.2	
28	69.2	84.0	70.5	73.3	81.1	85.0	54.9	54.9	64.9	74.8	72.0	80.7	81.6	81.7
29	78.0	73.3	79.1	82.4	72.6	71.2	57.6	57.6	71.9	80.6	82.8	72.0	73.0	73.3
30	63.0	59.8	64.4	62.6	59.1	58.0	59.1	59.1	57.7	63.0	64.5	58.2	59.5	59.5
31	58.1	55.6	58.7	57.1	55.8	54.1	55.3	55.1	53.0	59.2	58.5	53.9	54.9	55.0
32	78.6	73.9	78.9	82.6	73.0	71.4	57.8	57.8	72.3	80.8	83.3	72.0	73.4	73.5
33	70.7	98.5	69.4	74.6	89.4	83.9	55.9	55.9	66.1	74.6	71.6	86.8	88.1	88.7
34	80.0	73.9	81.6	81.3	72.8	72.0	57.8	57.8	71.9	80.9	85.9	71.8	72.3	72.5
35	66.7	80.9	64.1	70.9	81.7	76.4	53.2	53.2	62.6	71.5	68.6	84.6	85.1	84.9
36	70.5	84.0	69.9	74.9	82.4	85.0	54.9	54.7	66.1	75.3	71.8	81.5	82.8	82.6
37	70.5	84.2	69.9	74.9	82.6	85.0	54.9	54.7	66.1	75.1	71.8	81.7	83.0	82.8
38	70.6	91.5	69.8	74.0	91.5	84.6	56.2	56.0	66.3	76.0	73.2	92.4	94.1	95.0

FIGURE 11 (continued)

70/77

	28	29	30	31	32	33	34	35	36	37	38
1	53.7	66.2	47.9	41.4	67.2	58.3	72.5	52.9	54.8	54.8	58.5
2	54.0	63.4	46.3	39.6	63.2	56.7	67.8	50.9	55.3	55.3	56.6
3	52.4	60.3	42.9	38.5	60.7	57.4	61.6	52.5	53.4	53.4	57.9
4	54.0	63.8	45.6	40.1	63.6	56.5	68.5	50.2	54.7	54.7	56.4
5	51.6	59.8	42.7	38.1	60.2	56.2	60.7	52.0	53.3	53.3	57.1
6	56.0	67.8	46.8	41.3	68.1	59.3	69.1	55.6	57.3	57.3	60.0
7	50.2	57.4	40.8	39.4	57.8	52.0	57.0	52.8	50.7	50.7	54.3
8	34.0	44.9	32.5	34.3	44.7	37.0	48.5	37.0	33.8	33.8	37.3
9	53.6	63.7	44.1	37.1	63.9	57.9	63.2	51.5	55.4	55.4	56.9
10	34.8	44.5	31.9	34.2	45.2	36.8	45.7	37.8	35.9	35.9	37.9
11	52.9	64.2	45.3	40.0	64.6	56.5	67.5	51.4	54.6	54.6	57.4
12	58.0	67.1	45.5	40.3	67.1	62.4	65.9	57.2	60.3	60.3	62.9
13	55.6	64.7	44.3	39.0	64.7	61.7	63.8	55.8	57.5	57.5	60.9
14	55.6	62.5	46.3	40.6	62.7	57.4	61.9	54.4	56.8	57.0	58.9
15	73.3	57.7	43.8	39.6	57.9	97.0	57.6	72.9	75.3	75.3	84.6
16	49.7	62.0	43.2	38.6	62.2	53.5	65.5	48.8	51.3	51.3	54.0
17	54.6	64.5	44.1	39.3	65.4	60.6	61.6	54.9	56.2	56.2	59.8
18	71.6	58.5	43.8	38.3	58.5	81.4	56.6	73.9	73.9	73.9	86.1
19	76.2	54.8	41.1	37.3	54.8	73.3	54.0	64.7	76.4	76.2	74.4
20	37.0	38.3	43.2	39.0	38.7	38.9	38.7	35.6	36.9	37.1	38.1
21	37.0	38.3	43.2	39.0	38.7	38.9	38.7	35.6	36.9	37.1	38.1
22	45.9	59.0	39.0	34.8	59.4	49.3	55.6	44.2	47.5	47.5	49.3
23	58.0	65.7	45.8	40.5	65.9	62.0	66.0	55.6	60.1	60.1	61.4
24	54.1	65.9	48.0	41.3	66.1	57.7	72.6	52.9	55.6	55.6	58.1
25	71.7	57.8	43.6	38.4	58.0	80.7	56.2	79.5	74.2	74.2	90.6
26	72.7	59.0	44.4	39.1	59.0	81.6	56.8	79.3	75.8	75.8	92.3
27	72.4	59.4	44.5	39.2	59.4	82.2	57.4	79.3	75.4	75.4	93.1
28		56.4	44.3	40.0	56.6	73.4	54.2	66.2	91.5	91.5	75.2
29	72.2		44.3	40.1	98.2	57.7	71.9	54.9	57.9	57.9	59.3
30	59.9	63.7		79.6	44.5	44.7	45.1	41.1	45.2	45.4	44.3
31	56.0	59.3	85.5		39.9	39.3	40.9	40.2	39.7	39.9	39.5
32	72.7	98.7	63.4	59.0		57.9	72.3	54.1	57.9	57.9	59.3
33	83.8	72.9	60.0	56.0	73.5		57.1	72.3	74.8	74.8	84.0
34	71.4	85.9	65.8	60.2	86.1	73.5		53.9	55.2	55.2	58.1
35	75.6	69.9	56.1	55.6	69.4	80.3	69.6		68.6	68.3	84.6
36	94.6	73.8	61.1	56.1	74.2	83.3	71.4	77.5		99.6	78.2
37	94.4	73.6	61.1	56.1	74.0	83.5	71.2	77.3	99.8		78.0
38	83.8	74.9	59.4	56.0	75.1	90.9	73.8	89.6	84.9	84.9	

FIGURE 11 (continued)

71/77

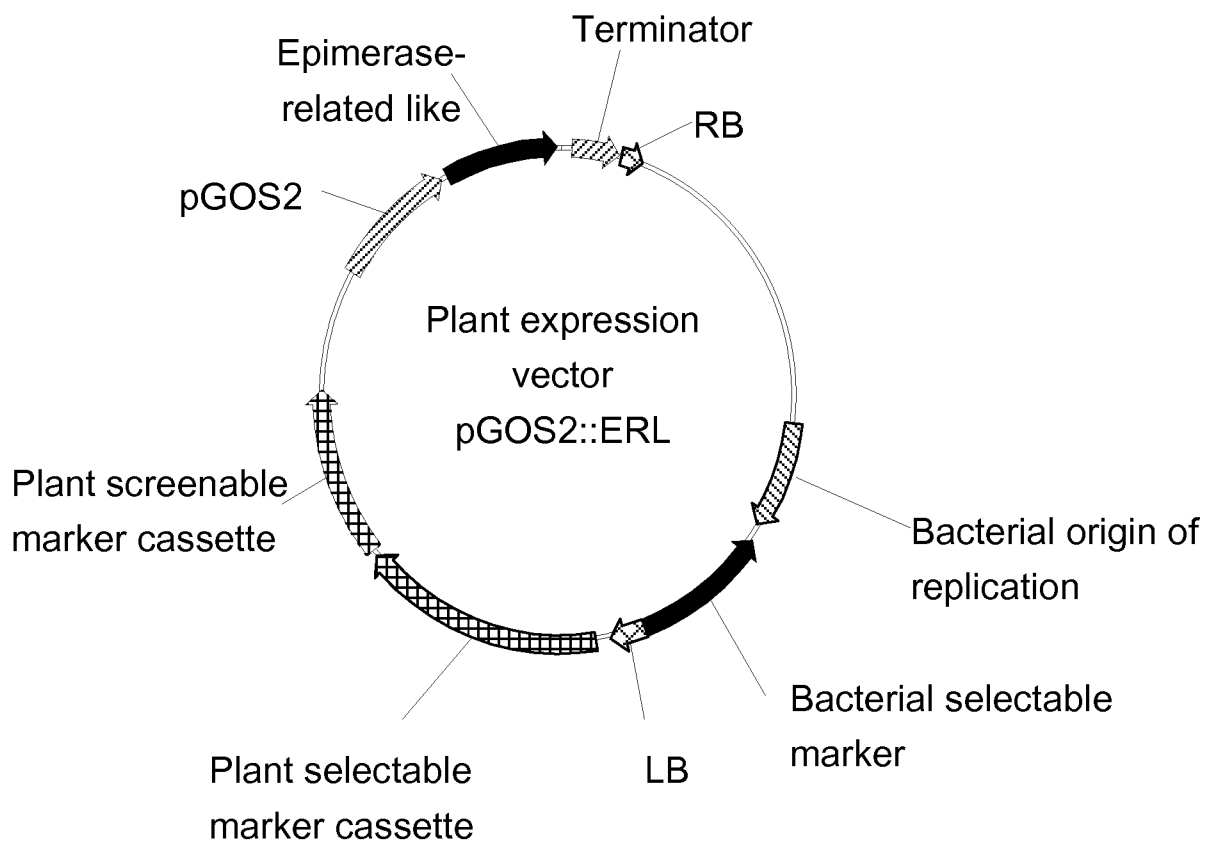


FIGURE 12

72/77

MSHAHSHIGSGSRTTRRTFEFGKTHVVRPKGKHQATIVWLHGLGDNGLS

-----motif 1-----
 -----motif 4-----
 -----motif 7-----
 -----motif 10-----

SSQLLESLPLPNIKWICPTAPTRPVAILGGFPCTAWFDVGELSEDGPDD

 -----motif 6--
 -----motif 8-----

WEGLDASAAHIANLLSTEPADVKGIGGFSMGAAIALYSATCYAMGRYG

-----motif 2----- motif 3-----

NGIPYPVNLRAVVGLSGWLPGRSLGNKIEVSHEARRRAASLPILQCHG

 ISDDVHCKYGEKSAQSLSSAGFRYVAFKSYEGIGHYTVPREMGEVSTW

-----motif 9-----
 -----motif 12-----

LSSRLGLEGFSS

FIGURE 13

73/77

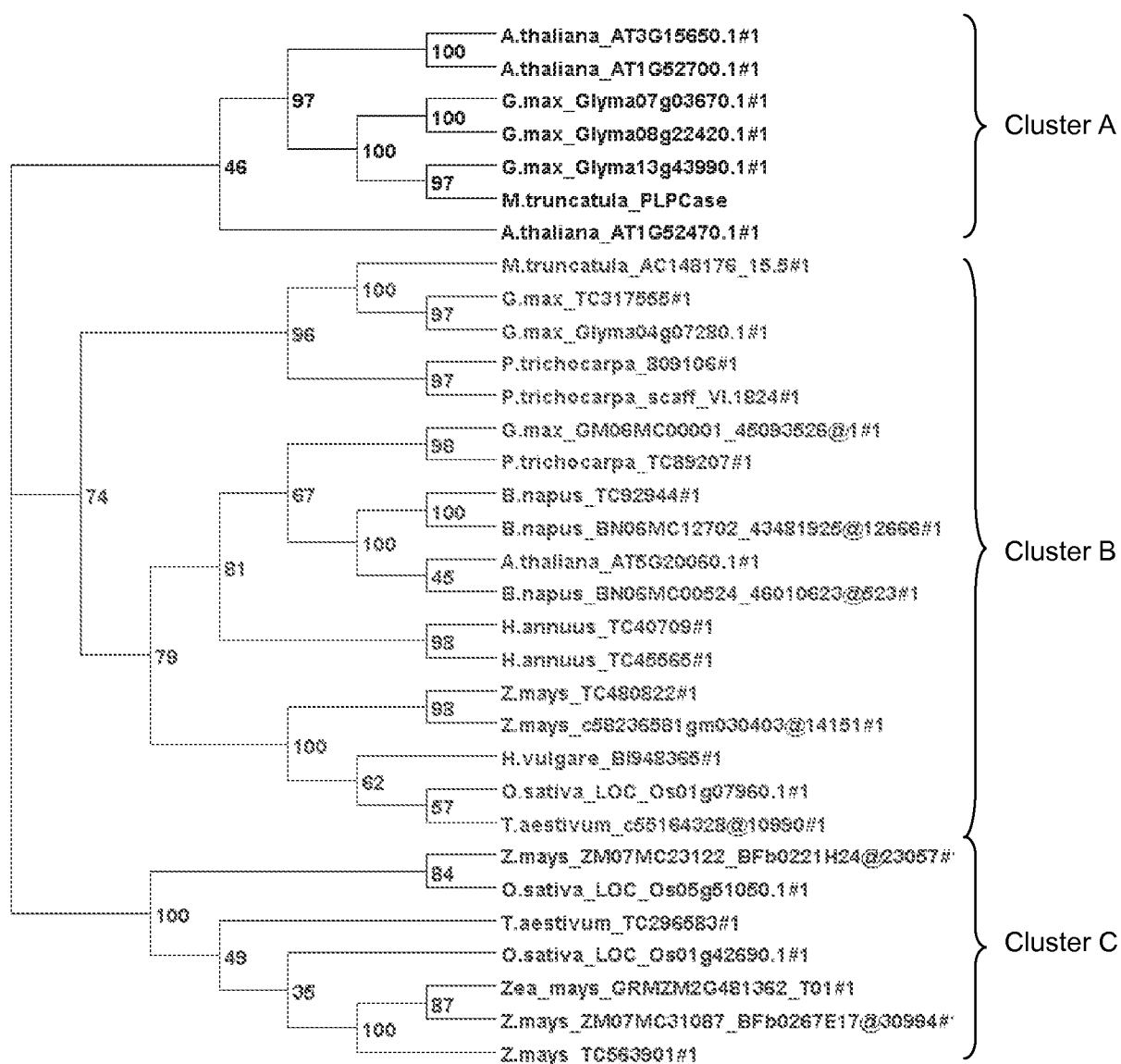


FIGURE 14

74/77

		1	2	3	4	5	6	7	8	9	10	11
Cluster A	1		77.00	34.00	93.00	88.00	85.00	76.00	69.00	67.00	69.00	67.00
	2	85.00		34.00	77.00	75.00	73.00	86.00	67.00	67.00	64.00	65.00
	3	49.00	47.00		34.00	34.00	33.00	36.00	30.00	32.00	32.00	30.00
	4	95.00	85.00	49.00		90.00	86.00	76.00	70.00	69.00	70.00	69.00
	5	93.00	85.00	51.00	93.00		96.00	77.00	69.00	68.00	68.00	66.00
	6	90.00	82.00	51.00	90.00	97.00		74.00	65.00	65.00	65.00	65.00
	7	85.00	92.00	51.00	86.00	87.00	84.00		65.00	66.00	62.00	62.00
Cluster B	8	82.00	80.00	48.00	83.00	82.00	80.00	80.00		81.00	72.00	72.00
	9	79.00	79.00	49.00	81.00	82.00	79.00	79.00	89.00		70.00	71.00
	10	80.00	79.00	50.00	82.00	80.00	77.00	79.00	85.00	83.00		86.00
	11	81.00	80.00	52.00	83.00	82.00	79.00	80.00	85.00	84.00	96.00	
	12	61.00	62.00	38.00	62.00	60.00	58.00	62.00	64.00	61.00	71.00	70.00
	13	79.00	80.00	47.00	80.00	80.00	77.00	80.00	88.00	88.00	82.00	83.00
	14	79.00	77.00	48.00	81.00	79.00	76.00	78.00	82.00	84.00	80.00	82.00
	15	80.00	80.00	48.00	82.00	80.00	77.00	80.00	84.00	85.00	82.00	84.00
	16	82.00	80.00	50.00	84.00	82.00	79.00	81.00	86.00	86.00	83.00	85.00
	17	80.00	79.00	50.00	82.00	80.00	77.00	80.00	85.00	85.00	82.00	84.00
	18	67.00	68.00	48.00	70.00	68.00	66.00	67.00	71.00	72.00	71.00	71.00
	19	80.00	78.00	50.00	80.00	79.00	77.00	78.00	81.00	78.00	80.00	82.00
	20	80.00	78.00	52.00	81.00	80.00	79.00	77.00	85.00	83.00	96.00	95.00
	21	66.00	66.00	42.00	67.00	65.00	62.00	66.00	68.00	67.00	75.00	76.00
	22	68.00	68.00	45.00	67.00	66.00	64.00	66.00	67.00	65.00	68.00	69.00
	23	86.00	85.00	49.00	86.00	86.00	84.00	85.00	87.00	84.00	85.00	86.00
	24	75.00	72.00	51.00	75.00	73.00	71.00	72.00	74.00	73.00	76.00	77.00
	25	83.00	80.00	51.00	83.00	81.00	79.00	80.00	83.00	83.00	83.00	85.00
Cluster C	26	75.00	75.00	48.00	76.00	76.00	74.00	75.00	71.00	74.00	78.00	77.00
	27	78.00	78.00	50.00	79.00	79.00	77.00	78.00	74.00	78.00	81.00	81.00
	28	78.00	77.00	49.00	79.00	79.00	76.00	77.00	78.00	81.00	80.00	81.00
	29	79.00	78.00	51.00	80.00	80.00	77.00	79.00	75.00	78.00	82.00	81.00
	30	79.00	80.00	50.00	81.00	80.00	77.00	80.00	79.00	81.00	83.00	83.00
	31	74.00	74.00	47.00	73.00	74.00	71.00	75.00	73.00	75.00	75.00	77.00
	32	80.00	78.00	49.00	81.00	82.00	78.00	80.00	81.00	82.00	82.00	83.00

FIGURE 15

75/77

		12	13	14	15	16	17	18	19	20	21	22
Cluster A	1	53.00	64.00	63.00	65.00	65.00	63.00	55.00	67.00	69.00	55.00	55.00
	2	49.00	65.00	63.00	65.00	64.00	62.00	54.00	65.00	65.00	54.00	51.00
	3	27.00	29.00	33.00	32.00	33.00	32.00	28.00	31.00	31.00	27.00	24.00
	4	54.00	65.00	66.00	67.00	66.00	64.00	57.00	66.00	70.00	56.00	55.00
	5	52.00	64.00	63.00	64.00	64.00	62.00	55.00	65.00	67.00	54.00	53.00
	6	49.00	61.00	61.00	62.00	61.00	60.00	53.00	63.00	65.00	53.00	51.00
	7	47.00	64.00	62.00	64.00	64.00	62.00	53.00	64.00	62.00	50.00	49.00
Cluster B	8	54.00	76.00	72.00	74.00	74.00	71.00	59.00	70.00	71.00	59.00	54.00
	9	51.00	74.00	70.00	73.00	71.00	70.00	60.00	69.00	70.00	56.00	50.00
	10	69.00	69.00	65.00	67.00	68.00	66.00	58.00	68.00	90.00	67.00	52.00
	11	62.00	69.00	66.00	69.00	68.00	66.00	57.00	70.00	87.00	75.00	51.00
	12		53.00	51.00	53.00	55.00	53.00	58.00	50.00	64.00	74.00	45.00
	13	63.00		68.00	69.00	71.00	69.00	68.00	66.00	67.00	57.00	49.00
	14	63.00	81.00		96.00	88.00	86.00	56.00	62.00	65.00	54.00	48.00
	15	64.00	82.00	96.00		90.00	88.00	58.00	64.00	67.00	56.00	50.00
	16	66.00	86.00	93.00	95.00		89.00	58.00	65.00	67.00	57.00	51.00
	17	64.00	83.00	93.00	95.00	96.00		58.00	61.00	66.00	55.00	49.00
	18	69.00	77.00	70.00	70.00	73.00	72.00		51.00	57.00	61.00	45.00
	19	60.00	77.00	75.00	77.00	79.00	77.00	64.00		69.00	55.00	55.00
	20	69.00	81.00	79.00	82.00	81.00	82.00	71.00	82.00		66.00	52.00
	21	84.00	67.00	66.00	67.00	68.00	68.00	73.00	65.00	74.00		48.00
	22	54.00	66.00	64.00	65.00	65.00	64.00	59.00	68.00	68.00	57.00	
	23	65.00	83.00	82.00	83.00	84.00	83.00	69.00	86.00	86.00	70.00	75.00
	24	68.00	73.00	73.00	73.00	76.00	74.00	74.00	80.00	76.00	74.00	67.00
	25	62.00	82.00	79.00	80.00	82.00	81.00	67.00	92.00	84.00	68.00	71.00
Cluster C	26	59.00	74.00	74.00	75.00	77.00	74.00	66.00	71.00	78.00	64.00	62.00
	27	61.00	77.00	78.00	78.00	80.00	78.00	68.00	74.00	82.00	67.00	63.00
	28	61.00	77.00	78.00	78.00	81.00	80.00	68.00	76.00	81.00	66.00	63.00
	29	62.00	77.00	78.00	78.00	81.00	78.00	69.00	75.00	82.00	67.00	64.00
	30	61.00	79.00	78.00	80.00	82.00	79.00	68.00	78.00	84.00	66.00	64.00
	31	64.00	73.00	74.00	74.00	76.00	75.00	73.00	71.00	77.00	71.00	59.00
	32	59.00	78.00	78.00	79.00	82.00	80.00	69.00	78.00	83.00	65.00	63.00

FIGURE 15 (continued)

76/77

		23	24	25	26	27	28	29	30	31	32
Cluster A	1	76.00	63.00	70.00	62.00	66.00	64.00	66.00	67.00	62.00	65.00
	2	72.00	61.00	69.00	64.00	68.00	64.00	68.00	68.00	63.00	66.00
	3	33.00	29.00	32.00	30.00	33.00	30.00	33.00	30.00	28.00	34.00
	4	76.00	63.00	70.00	63.00	66.00	65.00	67.00	68.00	63.00	66.00
	5	74.00	61.00	68.00	61.00	65.00	63.00	66.00	67.00	63.00	66.00
	6	71.00	58.00	65.00	59.00	63.00	60.00	63.00	64.00	59.00	63.00
	7	72.00	59.00	67.00	61.00	66.00	63.00	66.00	66.00	63.00	65.00
Cluster B	8	75.00	63.00	71.00	57.00	61.00	60.00	61.00	65.00	60.00	64.00
	9	72.00	61.00	70.00	60.00	63.00	65.00	64.00	67.00	62.00	66.00
	10	72.00	63.00	70.00	62.00	66.00	64.00	66.00	67.00	61.00	63.00
	11	73.00	64.00	73.00	60.00	65.00	63.00	65.00	68.00	61.00	64.00
	12	54.00	57.00	52.00	46.00	49.00	49.00	49.00	49.00	52.00	48.00
	13	70.00	60.00	67.00	59.00	63.00	58.00	63.00	63.00	59.00	61.00
	14	68.00	58.00	64.00	56.00	60.00	59.00	61.00	61.00	58.00	60.00
	15	70.00	60.00	66.00	58.00	62.00	61.00	62.00	63.00	60.00	62.00
	16	70.00	61.00	66.00	56.00	60.00	59.00	61.00	62.00	59.00	61.00
	17	68.00	58.00	64.00	54.00	58.00	58.00	58.00	60.00	58.00	58.00
	18	56.00	57.00	53.00	49.00	52.00	49.00	53.00	53.00	57.00	52.00
	19	77.00	75.00	89.00	58.00	62.00	62.00	62.00	63.00	57.00	62.00
	20	74.00	62.00	71.00	62.00	66.00	65.00	67.00	69.00	63.00	64.00
	21	59.00	63.00	59.00	50.00	54.00	52.00	54.00	55.00	57.00	52.00
	22	66.00	56.00	56.00	46.00	47.00	45.00	47.00	49.00	48.00	48.00
	23		71.00	80.00	61.00	65.00	65.00	65.00	69.00	63.00	66.00
	24	81.00		86.00	56.00	59.00	58.00	59.00	58.00	62.00	58.00
	25	90.00	87.00		61.00	64.00	65.00	65.00	67.00	61.00	66.00
Cluster C	26	75.00	68.00	74.00		94.00	75.00	94.00	80.00	68.00	74.00
	27	78.00	72.00	77.00	96.00		79.00	100.00	84.00	72.00	78.00
	28	81.00	72.00	78.00	84.00	88.00		79.00	83.00	71.00	77.00
	29	79.00	72.00	78.00	96.00	100.00	88.00		85.00	72.00	78.00
	30	83.00	72.00	81.00	86.00	90.00	91.00	90.00		80.00	84.00
	31	75.00	76.00	74.00	78.00	82.00	82.00	82.00	86.00		80.00
	32	81.00	73.00	81.00	83.00	87.00	87.00	87.00	92.00	85.00	

FIGURE 15 (continued)

77/77

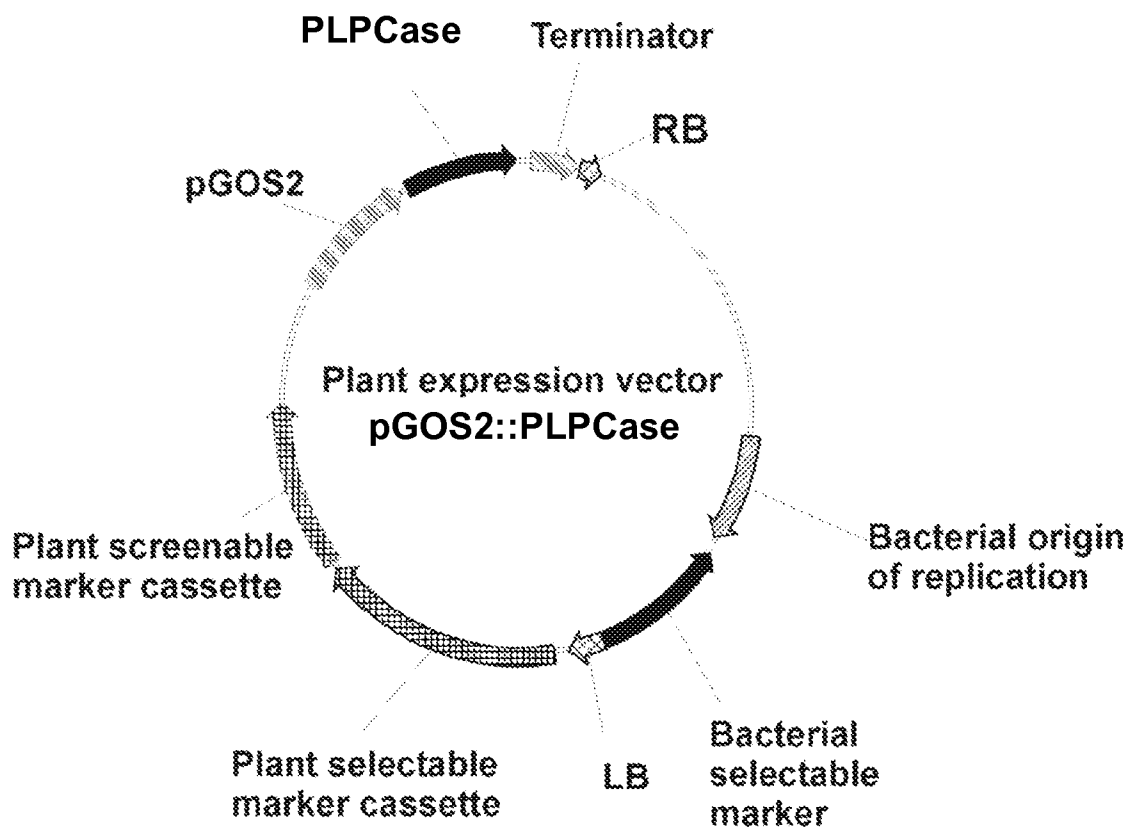


FIGURE 16

INTERNATIONAL SEARCH REPORT

International application No.

PCT/IB2012/050878

A. CLASSIFICATION OF SUBJECT MATTER

See extra sheet

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC: C12N; A01H; C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNABS, CPRSABS, JPABS, CNMED, DWPI, SIPOABS, CPEA, CNTXT, TWTXT, WOTXT, EPTXT, USTXT, CATXT, CNKI, ISI WEB OF KNOWLEDGE, BAIDU, GOOGLE SCHOLAR and keywords: DUF642, yield, trait+, plants, Os04g0494800, AT1g29980 et al. GENBANK, EMBL: SEQ ID NOS: 1-2

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2009141824A2 (EVOGENE LTD ET AL.) 26 Nov. 2009(26.11.2009) see claims 1-34, examples 4-5 and 7, Tables 1 and 23 of the description, SEQ ID NOS: 29 and 79	1-8, 10-26
Y	see claims 1-34, examples 4-5 and 7, Tables 1 and 23 of the description, SEQ ID NOS: 29 and 79	1-26

☒ Further documents are listed in the continuation of Box C.☒ See patent family annex.

* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier application or patent but published on or after the international filing date	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim (S) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&" document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search 13 Jun. 2012(13.06.2012)	Date of mailing of the international search report 26 Jul. 2012 (26.07.2012)
---	--

Name and mailing address of the ISA/CN
The State Intellectual Property Office, the P.R.China
6 Xitucheng Rd., Jimen Bridge, Haidian District, Beijing, China
100088
Facsimile No. 86-10-62019451

Authorized officer

LIN, Junkai

Telephone No. (86-10)62412317

INTERNATIONAL SEARCH REPORT

International application No.

PCT/IB2012/050878

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	GenBank Accession No. BAF15103.1, 07 Aug. 2009(07.08.2009), [retrieved on 13 Jun. 2012(13.06.2012)]. Retrieved from: GenBank database see the amino acid sequence	1-26
A	IRSHAD, M. et al. A new picture of cell wall protein dynamics in elongating cells of <i>Arabidopsis thaliana</i> : Confirmed actors and newcomers. BMC Plant Biology. 16 September 2008(16.09.2008), vol. 8, no. 94, pages 1-16, ISSN 1471-2229 see the whole document	1-26
A	WO 2010086221A1 (BASF PLANT SCI CO GMBH) 05 Aug. 2010(05.08.2010) see the whole document	1-26

INTERNATIONAL SEARCH REPORT

International application No.

PCT/IB2012/050878

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

See extra sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.: 1-26 (all partially)

Remark on protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/IB2012/050878

Continuation of: Box No. III

This Authority considers that there are 92 separate inventions covered by the claims indicated as follows:

Group I:

1: claims 1-26 (all partially)

a DUF642 polypeptide (SEQ ID NO: 2) and subject-matter relating thereto;

2 claims 1-26 (all partially)

a DUF642 polypeptide (SEQ ID NO: 4) and subject-matter relating thereto;

3-90: claims 1-8, 10-26 (all partially)

a DUF642 polypeptide according to SEQ ID NOs:5-180 even numbers, respectively, and subject-matter relating thereto;

Group II:

91: claims 27-54

a epimerase-related like polypeptide and subject-matter relating thereto;

Group III:

92: claims 55-78

a phospholipase/carboxylesterase (PLPCase) polypeptide and subject-matter relating thereto.

Detailed lack of unity reasoning concerning the present application:

The present application concerns proteins potentially useful for enhancing yield-related traits in plants. In the light of the prior art (e.g. WO 2010086221A1), the underlying problem to be solved by the present application is seen in the provision of further proteins potentially useful for enhancing yield-related traits in plants. The problem is solved by providing the structurally completely different polypeptides of Groups I-III.

However, this concept cannot be regarded as the "single general inventive concept" required by Rules 13.1 PCT because it is neither novel nor inventive, since such proteins potentially useful for enhancing yield-related traits in plants are already known in the art (e.g. WO 2010086221A1).

With regard to the prior art (e.g. WO 2010086221A1), the common special technical feature among the inventions 1-90 (common concept) seems to be the fact that all claimed polypeptides represent members of the DUF642 family.

However, as shown by WO 2009141824A2, which discloses a gene BDL141 (AT1G29980) and its encoding protein belonging to DUF642 family (see Table 1), and expressing within the plant the BDL141 gene resulting in the increase of dry weight as compared to control (see example 7 and Table 23), therefore such polypeptides and their use in enhancing plant yield-related traits were already known in the art by the time of the invention.

In conclusion, the 92 groups of claims are not linked by common or corresponding special technical features and define 92 separate inventions not linked by a single general inventive concept. The application, hence does not meet the requirements of unity of invention as defined in Rules 13.1 and 13.2 PCT.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/IB2012/050878

Continuation of: CLASSIFICATION OF SUBJECT MATTER

C12N 15/82 (2006.01) i

A01H 5/00 (2006.01) i

C07K14/415 (2006.01) i

C12N 9/16 (2006.01) i

C12N 9/18 (2006.01) i

C12N 9/90 (2006.01) i

C12N 15/61 (2006.01) i

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/IB2012/050878

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
WO 2009141824A2	26.11.2009	WO 2009141824A3	22.04.2010
		AU 2009250806A1	26.11.2009
		EP 2279258A2	02.02.2011
		CA 2724545A1	26.11.2009
		INMUMNP 201002525E	11.03.2011
		US 2011097771A1	28.04.2011
		MXPA 10012697A	31.03.2011
WO 2010086221A1	05.08.2010	AU 2010209875A1	05.08.2010
		CA 2750007A1	05.08.2010
		MXPA 11007599A	31.08.2011
		AU 2010209875A2	05.08.2010
		US 2011283420A1	17.11.2011
		EP 2391719A1	07.12.2011