



## (51) International Patent Classification:

C12N 15/29 (2006.01) C07K 14/415 (2006.01)  
C12N 15/113 (2010.01)

## (21) International Application Number:

PCT/US2015/053629

## (22) International Filing Date:

2 October 2015 (02.10.2015)

## (25) Filing Language:

English

## (26) Publication Language:

English

## (30) Priority Data:

62/059,069 2 October 2014 (02.10.2014) US

(71) Applicant: **PIONEER HI-BRED INTERNATIONAL, INC.** [US/US]; 7100 N.W. 62nd Avenue, Johnston, Iowa 50131-1014 (US).

(72) Inventors: **NAGAHATENNA, Dilrukshi**; Australian Centre for Plant Functional, Genomics Pty Ltd., University of Adelaide, Waite Campus, Hartley Grove, Urrbrae, South Australia 5064 (AU). **WHITFORD, Ryan**; Australian

Centre for Plant Functional, Genomics Pty Ltd., University of Adelaide, Waite Campus, Hartley Grove, Urrbrae, South Australia 5064 (AU).

(74) Agent: **DEVADAS, Sendil**; PIONEER HI-BRED INTERNATIONAL, INC., 7100 N.W. 62nd Avenue, Johnston, Iowa 50131-1014 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, 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, ST, SZ,

[Continued on next page]

(54) Title: FERROCHELATASE COMPOSITIONS AND METHODS TO INCREASE AGRONOMIC PERFORMANCE OF PLANTS

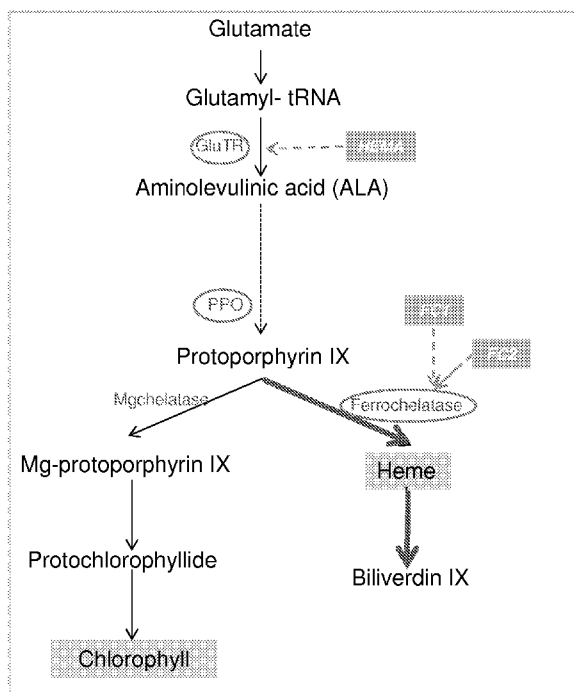


FIG. 1

(57) Abstract: Ferrochelatase expression results in improved agronomic performance including various photosynthetic characteristics and drought tolerance. Methods and compositions that affect yield and other agronomic characteristics in plants are disclosed.



TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, 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))*

## FERROCHELATASE COMPOSITIONS AND METHODS TO INCREASE AGRONOMIC PERFORMANCE OF PLANTS

5

### CROSS REFERENCE

This utility application claims the benefit of priority of U.S. Ser. No. 62/059,069, filed October 2, 2014, which is incorporated herein by reference in its entirety.

10

### REFERENCE TO SEQUENCE LISTING SUBMITTED ELECTRONICALLY

A sequence listing having the file name "20250E\_ST25.txt" created on September 28, 2015, and having a size of 13 kilobytes is filed in computer readable form concurrently with the specification. The sequence listing is part of the specification and is herein incorporated by reference in its entirety.

15

### FIELD

The disclosure relates generally to the field of molecular biology.

### BACKGROUND

20

Improving agronomic traits in crop plants is beneficial to farmers. Several factors influence crop yield. Abiotic stress is the primary cause of crop loss worldwide, causing average yield losses of more than 50% for major crops. Among the various abiotic stresses, drought is a major factor that limits crop productivity worldwide. Exposure of plants to a water-limiting environment during various developmental stages appears to activate various physiological and developmental changes. Molecular mechanisms of abiotic stress responses and the genetic regulatory networks of drought stress tolerance have been studied.

25

Natural responses to abiotic stress vary among plant species and among varieties and cultivars within a plant species. Certain species, varieties or cultivars are more tolerant to abiotic stress such as drought than others. Transgenic approaches are needed for improving drought tolerance in crop plants. Modulation of porphyrin biosynthesis, which is common to all higher plants and responsible for the synthesis of chlorophyll, heme, siroheme and phytychromobilin results in improved agronomic performance of plants.

30

35

Ferrochelatase (FC) is the terminal enzyme of heme biosynthesis. In photosynthetic organisms studied so far, there is evidence for two FC isoforms, which are encoded by two genes (FC1 and FC2).

## SUMMARY

The present disclosure provides polynucleotides, related polypeptides and all conservatively modified variants of barley ferrochelatases that have been shown to affect agronomic parameters in plants. In an embodiment, HvFC1 and HvFCII modulate drought tolerance and one or more other agronomic characteristics of a plant. In an embodiment, transgenic plants overexpressing HvFC1 and HvFCII had increased drought tolerance, improved photosynthetic performance, and altered oxidative stress response.

Methods of improving an agronomic characteristic of a plant, the method includes modulating the expression of (i) a polynucleotide encoding an amino acid sequence comprising SEQ ID NO: 1 or 2 or an amino acid sequence that is at least 95% identical to one of SEQ ID NO: 1 or 2 (ii) a polynucleotide that hybridizes under stringent hybridization conditions to a fragment of polynucleotide comprising SEQ ID NO: 3 or 4, wherein the fragment comprises at least 100 contiguous nucleotides of SEQ ID NO: 3 or 4 (iii) a polynucleotide that encodes an amino acid sequence that is at least 90% identical to SEQ ID NO: 1 or 2, (iv) a polynucleotide encoding a polypeptide comprising one or more deletions or insertions or substitutions of amino acids compared to SEQ ID NO: 1 or 2.

In an embodiment, the expression of the polynucleotide encoding a polypeptide having at least 95% identity to SEQ ID NO: 1 or 2 is increased by transforming the plant with a recombinant polynucleotide operably linked to a heterologous promoter. In an embodiment, the expression of an endogenous polynucleotide encoding a polypeptide having at least 95% identity to SEQ ID NO: 1 or 2 is increased by upregulating a regulatory element operably associated with the endogenous polynucleotide.

In an embodiment, the agronomic characteristic is selected from the group consisting of wilting avoidance, improved photosynthetic performance, increased chlorophyll content, increased photosynthetic rate, improved stomatal conductance, carboxylation efficiency, an increase in grain size, an increase in grain weight, an increase in grain yield, an increase in grain filling rate, and an increase in biomass. The increase in agronomic characteristic is measured with respect to a control plant that does not exhibit elevated levels of HvFCI or HvFCII (or a variant or an ortholog/homolog thereof). In an embodiment, the agronomic performance is an increase in drought tolerance. In an embodiment, the grain weight is increased in relation to a control plant not having an increased expression of the polynucleotide.

In an embodiment, the plant is a monocot. In an embodiment, the plant is wheat, barley, rice or maize. In an embodiment, the plant is a dicot. In an embodiment, the plant is soybean or brassica.

In an embodiment, methods of improving yield of a plant include increasing the expression of a polynucleotide that encodes a polypeptide comprising the amino acid sequence selected from the group consisting of SEQ ID NOS: 1-2 or an allelic variant thereof.

In an embodiment, methods of improving grain yield include the expression of a polynucleotide that encodes a polypeptide comprising the amino acid sequence of SEQ ID NO: 1 or 2 or a variant thereof.

In an embodiment, methods of marker assisted selection of a plant or identifying a native trait associated with increased yield, include:

- a. performing marker-assisted selection of plants that have one or more variations in genomic regions encoding a protein comprising SEQ ID NO: 1 or 2 or a variant thereof or a regulatory sequence thereof; and
- b. identifying the plant that exhibits higher yield.

In an embodiment, methods of identifying one or more alleles in a population of plants that are associated with increased grain yield includes:

- a. evaluating in a population of plants for one or more allelic variations in (i) a genomic region, the genomic region encoding a polypeptide or (ii) the regulatory region controlling the expression of the polypeptide, wherein the polypeptide comprises the amino acid sequence of SEQ ID NO: 1 or 2 or a sequence that is 95% identical to SEQ ID NO: 1 or 2;
- b. obtaining phenotypic values of increased yield for the one or more plants in the population;
- c. associating the allelic variations in the genomic region with the phenotype; and
- d. identifying the one or more alleles that are associated with increased yield.

In an embodiment, a recombinant expression cassette includes the polynucleotide that is operably linked to a regulatory element, wherein the expression cassette is functional in a plant cell. In an embodiment, a host cell includes the expression cassette. In an embodiment, a transgenic plant includes the recombinant expression cassette.

In an embodiment, a transgenic plant part includes a plant regulatory element that operably regulates the expression of a polynucleotide encoding a polypeptide

comprising the amino acid sequence of SEQ ID NO: 1 or 2 or a variant or an ortholog thereof, wherein the regulatory element is heterologous to the polynucleotide.

In an embodiment, the polynucleotide that comprises a fragment of SEQ ID NO: 2, is sufficient to up-regulate the endogenous expression of the polynucleotide that encodes a polypeptide.

In an embodiment, the modulation of the expression is achieved through mutagenesis. In an embodiment, the modulation of the expression is achieved through microRNA mediated gene silencing. In an embodiment, the modulation of the expression is achieved through promoter-mediated gene suppression. In an embodiment, the modulation of the expression is achieved through targeted mutagenesis of an endogenous regulatory element.

In another aspect, the present disclosure relates to a recombinant expression cassette comprising a nucleic acid as described. Additionally, the present disclosure relates to a vector containing the recombinant expression cassette. Further, the vector containing the recombinant expression cassette can facilitate the transcription and translation of the nucleic acid in a host cell. The present disclosure also relates to the host cells able to express the polynucleotide of the present disclosure. A number of host cells could be used, such as but not limited to, microbial, mammalian, plant or insect.

In yet another embodiment, the present disclosure is directed to a transgenic plant or plant cells, containing the nucleic acids of the present disclosure. Preferred plants containing the polynucleotides of the present disclosure include but are not limited to maize, soybean, sunflower, sorghum, canola, wheat, alfalfa, cotton, rice, barley, tomato and millet. In another embodiment, the transgenic plant is a maize plant or plant cells. Another embodiment is the transgenic seeds from the transgenic nitrate uptake-associated polypeptide of the disclosure operably linked to a promoter that drives expression in the plant. The plants of the disclosure can have improved grain quality as compared to a control plant.

#### BRIEF DESCRIPTION OF THE DRAWINGS

**Fig. 1** shows tetrapyrrole biosynthesis in drought stress signaling. The tetrapyrrole biosynthesis pathway with major end products (bold text) and some catalytic enzymes. GluTR; Glutamyl-tRNA reductase, Protoporphyrinogen IX oxidase; Protoporphyrinogen IX oxidase.

**Fig. 2** shows phylogenetic relationship between Ferrochelatase I and II. Phylogenetic relationship of HvFC1 and HvFC2 with other FC from grass and dicot species. At, Arabidopsis (*Arabidopsis thaliana*); Cs, cucumber (*Cucumis sativa*); Hv,

barley (*Hordeum vulgare*); Os, rice (*Oryza sativa*); Sit, foxtail millet (*Setaria italica*); Sbi, Sorghum (*Sorghum bicolor*); Zma, Maize (*Zea mays*). The maximum likelihood tree was constructed and reliability of the tree was estimated using bootstrap method.

**Fig. 3** shows tissue specific and stress responsive expression of HvFC and other markers.

**Fig. 4** shows HvFC1 is transcriptionally responsive to ROS and drought stress; ROS ( $^1O_2$ ) generation in *tigrina d<sup>12</sup>* mutant (A) and drought stress (B)

**Fig. 5** shows Barley FC overexpressing transgenics have higher photosynthetic performance.

**Fig. 6** shows FC overexpressing transgenics maintain higher leaf RWC under drought stress.

**Fig. 7** shows FC overexpressing transgenics showed higher water use efficiency (WUE) under drought.

**Fig. 8** shows FC transgenics improve photosynthetic performance under drought stress.

**Fig. 9** shows (a) Leaf N, and (b) Leaf total Fe concentration of transgenic barley lines over-expressing either HvFC1 or HvFC2 relative to WT and null controls. Data are shown as mean values  $\pm$  standard error from three different plants. Means with the same letter are not significantly different at  $P < 0.05$ , oneway ANOVA.

## DETAILED DESCRIPTION

In certain embodiments, modified ferrochelatase expression improves photosynthesis by about 13 % and the disclosure provides methods and compositions to engineer the tetrapyrrole biosynthetic pathway for improved crop performance. Plants expressing recombinant HvFC1 and HvFC2 exhibited improved photosynthetic rate ( $A_{sat}$ ), stomatal conductance ( $g_s$ ) and carboxylation efficiency (CE), demonstrating that FC1 and FC2 are involved in photosynthesis to improve crop performance.

A method of producing a seed, the method comprising: (a) crossing a first plant with a second plant, wherein at least one of the first plant and the second plant comprises a recombinant DNA construct, wherein the recombinant DNA construct comprises a polynucleotide operably linked to at least one heterologous regulatory element, wherein the polynucleotide encodes a polypeptide having an amino acid sequence of at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity, based on the Clustal V or the Clustal W method of alignment, using the respective default parameters, when compared to SEQ ID NO: 1 or 2; and (b) selecting a seed of the crossing of step (a), wherein the seed

comprises the recombinant DNA construct. A plant grown from the seed may exhibit at least one trait selected from the group consisting of: increased abiotic stress tolerance, increased yield, increased biomass, and altered root architecture, when compared to a control plant not comprising the recombinant DNA construct. The polypeptide may be over-expressed in at least one tissue of the plant, or during at least one condition of abiotic stress, or both. The plant may be selected from the group consisting of: maize, soybean, sunflower, sorghum, canola, wheat, alfalfa, cotton, rice, barley, millet, sugar cane and switchgrass.

A method of producing a plant that exhibits an increase in at least one trait selected from the group consisting of: increased abiotic stress tolerance, increased yield, increased biomass, and altered root architecture, wherein the method comprises growing a plant from a seed comprising a recombinant DNA construct, wherein the recombinant DNA construct comprises a polynucleotide operably linked to at least one heterologous regulatory element, wherein the polynucleotide encodes a polypeptide having an amino acid sequence of at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity, based on the Clustal V or the Clustal W method of alignment, using the respective default parameters, when compared to SEQ ID NO: 1 or 2, wherein the plant exhibits at least one trait selected from the group consisting of: increased nitrogen stress tolerance, increased yield, increased biomass, and altered root architecture, when compared to a control plant not comprising the recombinant DNA construct. The polypeptide may be over-expressed in at least one tissue of the plant, or during at least one condition of abiotic stress, or both. The plant may be selected from the group consisting of: maize, soybean, sunflower, sorghum, canola, wheat, alfalfa, cotton, rice, barley, millet, sugar cane and switchgrass.

Porphyrin biosynthesis is common to all higher plants and is responsible for the synthesis of chlorophyll, heme, siroheme and phytychromobilin which play vital roles in several primary metabolic processes (see FIG. 1). Mg<sup>2+</sup> containing chlorophyll, a cyclic porphyrin, is the most abundant of plant porphyrins. Five distinct chlorophylls, namely a, b, c, d and f have been identified in photosynthetic organisms. As the major light-harvesting compound, chlorophyll plays a key role in photosynthesis which converts light energy into useful chemical energy. Similar to chlorophyll, heme is also a cyclic compound, which contains Fe<sup>2+</sup> instead of Mg<sup>2+</sup>. Although chlorophyll is confined to plastids, heme has a widespread cellular distribution. Heme is capable of binding to a variety of proteins both covalently and non-covalently, acting as a prosthetic group. It is a co-factor for many enzymes involved in respiration and reactive oxygen species (ROS) detoxification within chloroplast, mitochondria and peroxisomes. Siroheme, another

Fe<sup>2+</sup> containing porphyrin, is a prosthetic group to nitrite and sulphite reductases, which are involved in nitrogen and sulphur assimilation, respectively. Phytochromobilin is a linear porphyrin found in plastids and plays a crucial role in photo-perception for photosynthesis.

5 Porphyrin biosynthesis is regulated at two control points, in order to meet greater porphyrin demand. The two major regulatory points are: (1) aminolevulinic acid (ALA) synthesis (2) at the branch point between chlorophyll and heme synthesis (Figure 1). ALA is the universal precursor necessary for the synthesis of all other porphyrins. Therefore, ALA synthesis is regulated at both the transcriptional and post-translational  
10 levels. The main enzyme regulating ALA synthesis is glutamyl-tRNA-reductase (GluTR). ALA synthesis is also regulated post-translationally level by two molecules, fluorescent (FLU) protein and heme. This negative regulation of ALA synthesis via FLU helps to prevent excessive accumulation of the highly photo-oxidative chlorophyll branch intermediate, Pchl<sub>ide</sub>.

15 Fig. 1 shows tetrapyrrole biosynthesis in drought stress signaling and Fig. 2 shows multiple sequence alignment of HvFCI and HvFCII and other ferrochelatases (FC) (Fig. 2A-2B). In addition, Fig. 3 shows phylogenetic relationship between Ferrochelatase I and II.

The practice of the present disclosure will employ, unless otherwise indicated,  
20 conventional techniques of botany, microbiology, tissue culture, molecular biology, chemistry, biochemistry and recombinant DNA technology, which are within the skill of the art. Such techniques are explained fully in the literature. See, e.g., Langenheim and Thimann, (1982) *Botany: Plant Biology and Its Relation to Human Affairs*, John Wiley; Cell Culture and Somatic Cell Genetics of Plants, vol. 1, Vasil, ed. (1984); Stanier, *et al.*,  
25 (1986) *The Microbial World*, 5<sup>th</sup> ed., Prentice-Hall; Dhringra and Sinclair, (1985) *Basic Plant Pathology Methods*, CRC Press; Maniatis, *et al.*, (1982) *Molecular Cloning: A Laboratory Manual*; DNA Cloning, vols. I and II, Glover, ed. (1985); *Oligonucleotide Synthesis*, Gait, ed. (1984); *Nucleic Acid Hybridization*, Hames and Higgins, eds. (1984) and the series *Methods in Enzymology*, Colowick and Kaplan, eds, Academic Press,  
30 Inc., San Diego, CA.

By "amplified" is meant the construction of multiple copies of a nucleic acid sequence or multiple copies complementary to the nucleic acid sequence using at least one of the nucleic acid sequences as a template. Amplification systems include the polymerase chain reaction (PCR) system, ligase chain reaction (LCR) system, nucleic  
35 acid sequence based amplification (NASBA, Cangene, Mississauga, Ontario), Q-Beta Replicase systems, transcription-based amplification system (TAS) and strand

displacement amplification (SDA). See, e.g., *Diagnostic Molecular Microbiology: Principles and Applications*, Persing, *et al.*, eds., American Society for Microbiology, Washington, DC (1993). The product of amplification is termed an amplicon.

It is understood, as those skilled in the art will appreciate, that the disclosure  
5 encompasses more than the specific exemplary sequences. Alterations in a nucleic acid  
fragment which result in the production of a chemically equivalent amino acid at a given  
site, but do not affect the functional properties of the encoded polypeptide, are well  
known in the art. For example, a codon for the amino acid alanine, a hydrophobic amino  
acid, may be substituted by a codon encoding another less hydrophobic residue, such  
10 as glycine, or a more hydrophobic residue, such as valine, leucine, or isoleucine.  
Similarly, changes which result in substitution of one negatively charged residue for  
another, such as aspartic acid for glutamic acid, or one positively charged residue for  
another, such as lysine for arginine, can also be expected to produce a functionally  
equivalent product. Nucleotide changes which result in alteration of the N terminal and  
15 C terminal portions of the polypeptide molecule would also not be expected to alter the  
activity of the polypeptide. Each of the proposed modifications is well within the routine  
skill in the art, as is determination of retention of biological activity of the encoded  
products.

The protein disclosed herein may also be a protein which comprises an amino  
20 acid sequence comprising deletion, substitution, insertion and/or addition of one or more  
amino acids in an amino acid sequence selected from the group consisting of SEQ ID  
NO: 1 or variants thereof. The substitution may be conservative, which means the  
replacement of a certain amino acid residue by another residue having similar physical  
and chemical characteristics. Non-limiting examples of conservative substitution include  
25 replacement between aliphatic group-containing amino acid residues such as Ile, Val,  
Leu or Ala, and replacement between polar residues such as Lys-Arg, Glu-Asp or Gln-  
Asn replacement.

Proteins derived by amino acid deletion, substitution, insertion and/or addition  
can be prepared when DNAs encoding their wild-type proteins are subjected to, for  
30 example, well-known site-directed mutagenesis (see, e.g., *Nucleic Acid Research*, Vol.  
10, No. 20, p.6487-6500, 1982, which is hereby incorporated by reference in its entirety).  
As used herein, the term "one or more amino acids" is intended to mean a possible  
number of amino acids which may be deleted, substituted, inserted and/or added by site-  
directed mutagenesis.

35 Site-directed mutagenesis may be accomplished, for example, as follows using a  
synthetic oligonucleotide primer that is complementary to single-stranded phage DNA to

be mutated, except for having a specific mismatch (i.e., a desired mutation). Namely, the above synthetic oligonucleotide is used as a primer to cause synthesis of a complementary strand by phages, and the resulting duplex DNA is then used to transform host cells. The transformed bacterial culture is plated on agar, whereby  
5     plaques are allowed to form from phage-containing single cells. As a result, in theory, 50% of new colonies contain phages with the mutation as a single strand, while the remaining 50% have the original sequence. At a temperature which allows hybridization with DNA completely identical to one having the above desired mutation, but not with  
10     DNA having the original strand, the resulting plaques are allowed to hybridize with a synthetic probe labeled by kinase treatment. Subsequently, plaques hybridized with the probe are picked up and cultured for collection of their DNA.

Techniques for allowing deletion, substitution, insertion and/or addition of one or more amino acids in the amino acid sequences of biologically active peptides such as enzymes while retaining their activity include site-directed mutagenesis mentioned  
15     above, as well as other techniques such as those for treating a gene with a mutagen, and those in which a gene is selectively cleaved to remove, substitute, insert or add a selected nucleotide or nucleotides, and then ligated.

The protein disclosed herein may also be a protein which is encoded by a nucleic acid comprising a nucleotide sequence comprising deletion, substitution, insertion and/or  
20     addition of one or more nucleotides in a nucleotide sequence selected from the group consisting of sequences encoding SEQ ID NOS: 1, 4-27. Nucleotide deletion, substitution, insertion and/or addition may be accomplished by site-directed mutagenesis or other techniques as mentioned above.

The protein disclosed herein may also be a protein which is encoded by a nucleic acid comprising a nucleotide sequence hybridizable under stringent conditions with the  
25     complementary strand of a nucleotide sequence selected from the group consisting of sequences encoding SEQ ID NOS: 1, 4-27.

The term "under stringent conditions" means that two sequences hybridize under moderately or highly stringent conditions. More specifically, moderately stringent  
30     conditions can be readily determined by those having ordinary skill in the art, e.g., depending on the length of DNA. The basic conditions are set forth by Sambrook et al., Molecular Cloning: A Laboratory Manual, third edition, chapters 6 and 7, Cold Spring Harbor Laboratory Press, 2001 and include the use of a prewashing solution for nitrocellulose filters 5xSSC, 0.5% SDS, 1.0 mM EDTA (pH 8.0), hybridization conditions  
35     of about 50% formamide, 2xSSC to 6xSSC at about 40-50 °C (or other similar hybridization solutions, such as Stark's solution, in about 50% formamide at about 42

°C) and washing conditions of, for example, about 40-60 °C, 0.5-6xSSC, 0.1% SDS. Preferably, moderately stringent conditions include hybridization (and washing) at about 50 °C and 6xSSC. Highly stringent conditions can also be readily determined by those skilled in the art, e.g., depending on the length of DNA.

5           Generally, such conditions include hybridization and/or washing at higher temperature and/or lower salt concentration (such as hybridization at about 65 °C, 6xSSC to 0.2xSSC, preferably 6xSSC, more preferably 2xSSC, most preferably 0.2xSSC), compared to the moderately stringent conditions. For example, highly stringent conditions may include hybridization as defined above, and washing at  
10   approximately 65-68 °C, 0.2xSSC, 0.1% SDS. SSPE (1xSSPE is 0.15 M NaCl, 10 mM NaH<sub>2</sub>PO<sub>4</sub>, and 1.25 mM EDTA, pH 7.4) can be substituted for SSC (1xSSC is 0.15 M NaCl and 15 mM sodium citrate) in the hybridization and washing buffers; washing is performed for 15 minutes after hybridization is completed.

          It is also possible to use a commercially available hybridization kit which uses no  
15   radioactive substance as a probe. Specific examples include hybridization with an ECL direct labeling & detection system. Stringent conditions include, for example, hybridization at 42 °C for 4 hours using the hybridization buffer included in the kit, which is supplemented with 5% (w/v) Blocking reagent and 0.5 M NaCl, and washing twice in 0.4% SDS, 0.5xSSC at 55 °C for 20 minutes and once in 2xSSC at room temperature  
20   for 5 minutes.

          By “encoding” or “encoded,” with respect to a specified nucleic acid, is meant comprising the information for translation into the specified protein. A nucleic acid encoding a protein may comprise non-translated sequences (e.g., introns) within translated regions of the nucleic acid or may lack such intervening non-translated  
25   sequences (e.g., as in cDNA). The information by which a protein is encoded is specified by the use of codons. Typically, the amino acid sequence is encoded by the nucleic acid using the “universal” genetic code. However, variants of the universal code, such as is present in some plant, animal, and fungal mitochondria, the bacterium *Mycoplasma capricolum* (Yamao, *et al.*, (1985) *Proc. Natl. Acad. Sci. USA* 82:2306-9) or  
30   the ciliate *Macronucleus*, may be used when the nucleic acid is expressed using these organisms.

          When the nucleic acid is prepared or altered synthetically, advantage can be taken of known codon preferences of the intended host where the nucleic acid is to be expressed. For example, although nucleic acid sequences of the present disclosure  
35   may be expressed in both monocotyledonous and dicotyledonous plant species, sequences can be modified to account for the specific codon preferences and GC

content preferences of monocotyledonous plants or dicotyledonous plants as these preferences have been shown to differ (Murray, *et al.*, (1989) *Nucleic Acids Res.* 17:477-98 and herein incorporated by reference). Thus, the maize preferred codon for a particular amino acid might be derived from known gene sequences from maize. Maize  
5 codon usage for 28 genes from maize plants is listed in Table 4 of Murray, *et al.*, *supra*.

As used herein, the term "FC" refers to ferrochelataes disclosed herein.

As used herein, "heterologous" in reference to a nucleic acid is a nucleic acid that originates from a foreign species, or, if from the same species, is substantially modified from its native form in composition and/or genomic locus by deliberate human  
10 intervention. Heterologous may also indicate that a particular nucleic acid is foreign to its location in the genome as compared to its native location in the genome. For example, a promoter operably linked to a heterologous structural gene is from a species different from that from which the structural gene was derived or, if from the same species, one or both are substantially modified from their original form. A heterologous  
15 protein may originate from a foreign species or, if from the same species, is substantially modified from its original form by deliberate human intervention.

By "host cell" is meant a cell, which comprises a heterologous nucleic acid sequence of the disclosure, which contains a vector and supports the replication and/or expression of the expression vector. Host cells may be prokaryotic cells such as *E. coli*,  
20 or eukaryotic cells such as yeast, insect, plant, amphibian or mammalian cells. Preferably, host cells are monocotyledonous or dicotyledonous plant cells, including but not limited to maize, sorghum, sunflower, soybean, wheat, alfalfa, rice, cotton, canola, barley, millet and tomato. A particularly preferred monocotyledonous host cell is a maize host cell.

25 The term "hybridization complex" includes reference to a duplex nucleic acid structure formed by two single-stranded nucleic acid sequences selectively hybridized with each other.

The term "introduced" in the context of inserting a nucleic acid into a cell, means "transfection" or "transformation" or "transduction" and includes reference to the  
30 incorporation of a nucleic acid into a eukaryotic or prokaryotic cell where the nucleic acid may be incorporated into the genome of the cell (e.g., chromosome, plasmid, plastid or mitochondrial DNA), converted into an autonomous replicon or transiently expressed (e.g., transfected mRNA).

The terms "isolated" refers to material, such as a nucleic acid or a protein, which  
35 is substantially or essentially free from components which normally accompany or interact with it as found in its naturally occurring environment. The isolated material

optionally comprises material not found with the material in its natural environment. Nucleic acids, which are “isolated”, as defined herein, are also referred to as “heterologous” nucleic acids. Unless otherwise stated, the term “nitrate uptake-associated nucleic acid” means a nucleic acid comprising a polynucleotide (“nitrate uptake-associated polynucleotide”) encoding a full length or partial length nitrate uptake-associated polypeptide.

As used herein, “nucleic acid” includes reference to a deoxyribonucleotide or ribonucleotide polymer in either single- or double-stranded form, and unless otherwise limited, encompasses known analogues having the essential nature of natural nucleotides in that they hybridize to single-stranded nucleic acids in a manner similar to naturally occurring nucleotides (e.g., peptide nucleic acids).

By “nucleic acid library” is meant a collection of isolated DNA or RNA molecules, which comprise and substantially represent the entire transcribed fraction of a genome of a specified organism. Construction of exemplary nucleic acid libraries, such as genomic and cDNA libraries, is taught in standard molecular biology references such as Berger and Kimmel, (1987) *Guide To Molecular Cloning Techniques*, from the series *Methods in Enzymology*, vol. 152, Academic Press, Inc., San Diego, CA; Sambrook, *et al.*, (1989) *Molecular Cloning: A Laboratory Manual*, 2<sup>nd</sup> ed., vols. 1-3; and *Current Protocols in Molecular Biology*, Ausubel, *et al.*, eds, Current Protocols, a joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc. (1994 Supplement).

As used herein “operably linked” includes reference to a functional linkage between a first sequence, such as a promoter, and a second sequence, wherein the promoter sequence initiates and mediates transcription of the DNA corresponding to the second sequence. Generally, operably linked means that the nucleic acid sequences being linked are contiguous and, where necessary to join two protein coding regions, contiguous and in the same reading frame.

As used herein, the term “plant” includes reference to whole plants, plant organs (e.g., leaves, stems, roots, etc.), seeds and plant cells and progeny of same. Plant cell, as used herein includes, without limitation, seeds, suspension cultures, embryos, meristematic regions, callus tissue, leaves, roots, shoots, gametophytes, sporophytes, pollen and microspores. The class of plants, which can be used in the methods of the disclosure, is generally as broad as the class of higher plants amenable to transformation techniques, including both monocotyledonous and dicotyledonous plants including species from the genera: *Cucurbita*, *Rosa*, *Vitis*, *Juglans*, *Fragaria*, *Lotus*, *Medicago*, *Onobrychis*, *Trifolium*, *Trigonella*, *Vigna*, *Citrus*, *Linum*, *Geranium*, *Manihot*,

*Daucus, Arabidopsis, Brassica, Raphanus, Sinapis, Atropa, Capsicum, Datura, Hyoscyamus, Lycopersicon, Nicotiana, Solanum, Petunia, Digitalis, Majorana, Ciahorium, Helianthus, Lactuca, Bromus, Asparagus, Antirrhinum, Heterocallis, Nemesis, Pelargonium, Panieum, Pennisetum, Ranunculus, Senecio, Salpiglossis,*  
 5 *Cucumis, Browaalia, Glycine, Pisum, Phaseolus, Lolium, Oryza, Avena, Hordeum, Secale, Allium and Triticum.* A particularly preferred plant is *Zea mays*.

As used herein, "polynucleotide" includes reference to a deoxyribopolynucleotide, ribopolynucleotide or analogs thereof that have the essential nature of a natural ribonucleotide in that they hybridize, under stringent hybridization conditions, to  
 10 substantially the same nucleotide sequence as naturally occurring nucleotides and/or allow translation into the same amino acid(s) as the naturally occurring nucleotide(s). A polynucleotide can be full-length or a subsequence of a native or heterologous structural or regulatory gene. Unless otherwise indicated, the term includes reference to the specified sequence as well as the complementary sequence thereof. Thus, DNAs or  
 15 RNAs with backbones modified for stability or for other reasons are "polynucleotides" as that term is intended herein. Moreover, DNAs or RNAs comprising unusual bases, such as inosine or modified bases, such as tritylated bases, to name just two examples, are polynucleotides as the term is used herein. It will be appreciated that a great variety of modifications have been made to DNA and RNA that serve many useful purposes known  
 20 to those of skill in the art. The term polynucleotide as it is employed herein embraces such chemically, enzymatically or metabolically modified forms of polynucleotides, as well as the chemical forms of DNA and RNA characteristic of viruses and cells, including *inter alia*, simple and complex cells.

The terms "polypeptide," "peptide" and "protein" are used interchangeably herein  
 25 to refer to a polymer of amino acid residues. The terms apply to amino acid polymers in which one or more amino acid residue is an artificial chemical analogue of a corresponding naturally occurring amino acid, as well as to naturally occurring amino acid polymers.

As used herein "promoter" includes reference to a region of DNA upstream from  
 30 the start of transcription and involved in recognition and binding of RNA polymerase and other proteins to initiate transcription. A "plant promoter" is a promoter capable of initiating transcription in plant cells. Exemplary plant promoters include, but are not limited to, those that are obtained from plants, plant viruses and bacteria which comprise genes expressed in plant cells such *Agrobacterium* or *Rhizobium*. Examples are  
 35 promoters that preferentially initiate transcription in certain tissues, such as leaves, roots, seeds, fibres, xylem vessels, tracheids or sclerenchyma. Such promoters are

referred to as "tissue preferred." A "cell type" specific promoter primarily drives expression in certain cell types in one or more organs, for example, vascular cells in roots or leaves. An "inducible" or "regulatable" promoter is a promoter, which is under environmental control. Examples of environmental conditions that may effect transcription by inducible promoters include anaerobic conditions or the presence of light. Another type of promoter is a developmentally regulated promoter, for example, a promoter that drives expression during pollen development. Tissue preferred, cell type specific, developmentally regulated and inducible promoters constitute the class of "non-constitutive" promoters. A "constitutive" promoter is a promoter, which is active under most environmental conditions. Suitable constitutive promoters include for example, Ubiquitin promoters, actin promoters, and GOS2 promoter (de Pater et al (1992), The Plant Journal, 2: 837–844).

As used herein "recombinant" includes reference to a cell or vector, that has been modified by the introduction of a heterologous nucleic acid or that the cell is derived from a cell so modified. Thus, for example, recombinant cells express genes that are not found in identical form within the native (non-recombinant) form of the cell or express native genes that are otherwise abnormally expressed, under expressed or not expressed at all as a result of deliberate human intervention or may have reduced or eliminated expression of a native gene. The term "recombinant" as used herein does not encompass the alteration of the cell or vector by naturally occurring events (e.g., spontaneous mutation, natural transformation/transduction/transposition) such as those occurring without deliberate human intervention.

As used herein, a "recombinant expression cassette" is a nucleic acid construct, generated recombinantly or synthetically, with a series of specified nucleic acid elements, which permit transcription of a particular nucleic acid in a target cell. The recombinant expression cassette can be incorporated into a plasmid, chromosome, mitochondrial DNA, plastid DNA, virus or nucleic acid fragment. Typically, the recombinant expression cassette portion of an expression vector includes, among other sequences, a nucleic acid to be transcribed and a promoter.

As used herein, "transgenic plant" includes reference to a plant, which comprises within its genome a heterologous polynucleotide. Generally, the heterologous polynucleotide is stably integrated within the genome such that the polynucleotide is passed on to successive generations. The heterologous polynucleotide may be integrated into the genome alone or as part of a recombinant expression cassette. "Transgenic" is used herein to include any cell, cell line, callus, tissue, plant part or plant, the genotype of which has been altered by the presence of heterologous nucleic acid

including those transgenics initially so altered as well as those created by sexual crosses or asexual propagation from the initial transgenic. The term "transgenic" as used herein does not encompass the alteration of the genome (chromosomal or extra-chromosomal) by conventional plant breeding methods or by naturally occurring events such as random cross-fertilization, non-recombinant viral infection, non-recombinant bacterial transformation, non-recombinant transposition or spontaneous mutation.

As used herein, "vector" includes reference to a nucleic acid used in transfection of a host cell and into which can be inserted a polynucleotide. Vectors are often replicons. Expression vectors permit transcription of a nucleic acid inserted therein.

The following terms are used to describe the sequence relationships between two or more nucleic acids or polynucleotides or polypeptides: (a) "reference sequence," (b) "comparison window," (c) "sequence identity," (d) "percentage of sequence identity" and (e) "substantial identity."

As used herein, "reference sequence" is a defined sequence used as a basis for sequence comparison. A reference sequence may be a subset or the entirety of a specified sequence; for example, as a segment of a full-length cDNA or gene sequence or the complete cDNA or gene sequence.

As used herein, "comparison window" means includes reference to a contiguous and specified segment of a polynucleotide sequence, wherein the polynucleotide sequence may be compared to a reference sequence and wherein the portion of the polynucleotide sequence in the comparison window may comprise additions or deletions (i.e., gaps) compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. Generally, the comparison window is at least 20 contiguous nucleotides in length, and optionally can be 30, 40, 50, 100 or longer. Those of skill in the art understand that to avoid a high similarity to a reference sequence due to inclusion of gaps in the polynucleotide sequence a gap penalty is typically introduced and is subtracted from the number of matches.

Methods of alignment of nucleotide and amino acid sequences for comparison are well known in the art. The local homology algorithm (BESTFIT) of Smith and Waterman, (1981) *Adv. Appl. Math* 2:482, may conduct optimal alignment of sequences for comparison; by the homology alignment algorithm (GAP) of Needleman and Wunsch, (1970) *J. Mol. Biol.* 48:443-53; by the search for similarity method (Tfasta and Fasta) of Pearson and Lipman, (1988) *Proc. Natl. Acad. Sci. USA* 85:2444; by computerized implementations of these algorithms, including, but not limited to: CLUSTAL in the PC/Gene program by Intelligenetics, Mountain View, California, GAP, BESTFIT, BLAST, FASTA and TFASTA in the Wisconsin Genetics Software Package,

Version 8 (available from Genetics Computer Group (GCG® programs (Accelrys, Inc., San Diego, CA).). The CLUSTAL program is well described by Higgins and Sharp, (1988) *Gene* 73:237-44; Higgins and Sharp, (1989) *CABIOS* 5:151-3; Corpet, *et al.*, (1988) *Nucleic Acids Res.* 16:10881-90; Huang, *et al.*, (1992) *Computer Applications in the Biosciences* 8:155-65 and Pearson, *et al.*, (1994) *Meth. Mol. Biol.* 24:307-31. The preferred program to use for optimal global alignment of multiple sequences is PileUp (Feng and Doolittle, (1987) *J. Mol. Evol.*, 25:351-60 which is similar to the method described by Higgins and Sharp, (1989) *CABIOS* 5:151-53 and hereby incorporated by reference). The BLAST family of programs which can be used for database similarity searches includes: BLASTN for nucleotide query sequences against nucleotide database sequences; BLASTX for nucleotide query sequences against protein database sequences; BLASTP for protein query sequences against protein database sequences; TBLASTN for protein query sequences against nucleotide database sequences and TBLASTX for nucleotide query sequences against nucleotide database sequences. See, *Current Protocols in Molecular Biology*, Chapter 19, Ausubel *et al.*, eds., Greene Publishing and Wiley-Interscience, New York (1995).

As those of ordinary skill in the art will understand, BLAST searches assume that proteins can be modeled as random sequences. However, many real proteins comprise regions of nonrandom sequences, which may be homopolymeric tracts, short-period repeats, or regions enriched in one or more amino acids. Such low-complexity regions may be aligned between unrelated proteins even though other regions of the protein are entirely dissimilar. A number of low-complexity filter programs can be employed to reduce such low-complexity alignments. For example, the SEG (Wooten and Federhen, (1993) *Comput. Chem.* 17:149-63) and XNU (Claverie and States, (1993) *Comput. Chem.* 17:191-201) low-complexity filters can be employed alone or in combination.

As used herein, "sequence identity" or "identity" in the context of two nucleic acid or polypeptide sequences includes reference to the residues in the two sequences, which are the same when aligned for maximum correspondence over a specified comparison window. When percentage of sequence identity is used in reference to proteins it is recognized that residue positions which are not identical often differ by conservative amino acid substitutions, where amino acid residues are substituted for other amino acid residues with similar chemical properties (e.g., charge or hydrophobicity) and therefore do not change the functional properties of the molecule. Where sequences differ in conservative substitutions, the percent sequence identity may be adjusted upwards to correct for the conservative nature of the substitution. Sequences, which differ by such conservative substitutions, are said to have "sequence

similarity” or “similarity.” Means for making this adjustment are well known to those of skill in the art. Typically this involves scoring a conservative substitution as a partial rather than a full mismatch, thereby increasing the percentage sequence identity. Thus, for example, where an identical amino acid is given a score of 1 and a non-conservative substitution is given a score of zero, a conservative substitution is given a score between zero and 1. The scoring of conservative substitutions is calculated, e.g., according to the algorithm of Meyers and Miller, (1988) *Computer Applic. Biol. Sci.* 4:11-17, e.g., as implemented in the program PC/GENE (Intelligenetics, Mountain View, California, USA).

As used herein, “percentage of sequence identity” means the value determined by comparing two optimally aligned sequences over a comparison window, wherein the portion of the polynucleotide sequence in the comparison window may comprise additions or deletions (i.e., gaps) as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical nucleic acid base or amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison and multiplying the result by 100 to yield the percentage of sequence identity.

The term “substantial identity” of polynucleotide sequences means that a polynucleotide comprises a sequence that has between 50-100% sequence identity, preferably at least 50% sequence identity, preferably at least 60% sequence identity, preferably at least 70%, more preferably at least 80%, more preferably at least 90% and most preferably at least 95%, compared to a reference sequence using one of the alignment programs described using standard parameters. One of skill will recognize that these values can be appropriately adjusted to determine corresponding identity of proteins encoded by two nucleotide sequences by taking into account codon degeneracy, amino acid similarity, reading frame positioning and the like. Substantial identity of amino acid sequences for these purposes normally means sequence identity of between 55-100%, preferably at least 55%, preferably at least 60%, more preferably at least 70%, 80%, 90% and most preferably at least 95%.

#### Orthologs and Paralogs

Homologous sequences as described above can comprise orthologous or paralogous sequences. Several different methods are known by those of skill in the art for identifying and defining these functionally homologous sequences. Three general

methods for defining orthologs and paralogs are described; an ortholog, paralog or homolog may be identified by one or more of the methods described below.

*Variant Nucleotide Sequences in the non-coding regions*

5           The nitrate uptake-associated nucleotide sequences are used to generate variant nucleotide sequences having the nucleotide sequence of the 5'-untranslated region, 3'-untranslated region or promoter region that is approximately 70%, 75%, 80%, 85%, 90% and 95% identical to the original nucleotide sequence of the corresponding SEQ ID NO: 1. These variants are then associated with natural variation in the germplasm for  
10       component traits related to grain quality and/or grain yield. The associated variants are used as marker haplotypes to select for the desirable traits.

*Variant Amino Acid Sequences of FC-associated Polypeptides*

          Variant amino acid sequences of FC-associated polypeptides are generated. In  
15       this example, one amino acid is altered. Specifically, the open reading frames are reviewed to determine the appropriate amino acid alteration. The selection of the amino acid to change is made by consulting the protein alignment (with the other orthologs and other gene family members from various species). An amino acid is selected that is deemed not to be under high selection pressure (not highly conserved) and which is  
20       rather easily substituted by an amino acid with similar chemical characteristics (i.e., similar functional side-chain). Using a protein alignment, an appropriate amino acid can be changed. Once the targeted amino acid is identified, the procedure outlined herein is followed. Variants having about 70%, 75%, 80%, 85%, 90% and 95% nucleic acid sequence identity are generated using this method. These variants are then associated  
25       with natural variation in the germplasm for component traits related to grain quality and/or grain yield. The associated variants are used as marker haplotypes to select for the desirable traits.

*Synthetic Methods for Constructing Nucleic Acids*

30       The isolated nucleic acids of the present disclosure can also be prepared by direct chemical synthesis by methods such as the phosphotriester method of Narang, *et al.*, (1979) *Meth. Enzymol.* 68:90-9; the phosphodiester method of Brown, *et al.*, (1979) *Meth. Enzymol.* 68:109-51; the diethylphosphoramidite method of Beaucage, *et al.*, (1981) *Tetra. Letts.* 22(20):1859-62; the solid phase phosphoramidite triester method  
35       described by Beaucage, *et al.*, *supra*, e.g., using an automated synthesizer, e.g., as described in Needham-VanDevanter, *et al.*, (1984) *Nucleic Acids Res.* 12:6159-68 and

the solid support method of US Patent Number 4,458,066. Chemical synthesis generally produces a single stranded oligonucleotide. This may be converted into double stranded DNA by hybridization with a complementary sequence or by polymerization with a DNA polymerase using the single strand as a template. One of skill will recognize that while  
5 chemical synthesis of DNA is limited to sequences of about 100 bases, longer sequences may be obtained by the ligation of shorter sequences.

#### UTRs and Codon Preference

In general, translational efficiency has been found to be regulated by specific  
10 sequence elements in the 5' non-coding or untranslated region (5' UTR) of the RNA. Positive sequence motifs include translational initiation consensus sequences (Kozak, (1987) *Nucleic Acids Res.*15:8125) and the 5' cap structure (Drummond, *et al.*, (1985) *Nucleic Acids Res.* 13:7375). Negative elements include stable intramolecular 5' UTR stem-loop structures (Muesing, *et al.*, (1987) *Cell* 48:691)  
15 and AUG sequences or short open reading frames preceded by an appropriate AUG in the 5' UTR (Kozak, *supra*, Rao, *et al.*, (1988) *Mol. and Cell. Biol.* 8:284). Accordingly, the present disclosure provides 5' and/or 3' UTR regions for modulation of translation of heterologous coding sequences.

#### 20 Plant Transformation Methods

Numerous methods for introducing foreign genes into plants are known and can be used to insert a nitrate uptake-associated polynucleotide into a plant host, including biological and physical plant transformation protocols. See, e.g., Miki, *et al.*, "Procedure for Introducing Foreign DNA into Plants," in *Methods in Plant Molecular Biology and  
25 Biotechnology*, Glick and Thompson, eds., CRC Press, Inc., Boca Raton, pp. 67-88 (1993). The methods chosen vary with the host plant, and include chemical transfection methods such as calcium phosphate, microorganism-mediated gene transfer such as *Agrobacterium* (Horsch *et al.*, (1985) *Science* 227:1229-31), electroporation, micro-injection and biolistic bombardment.

30 Expression cassettes and vectors and *in vitro* culture methods for plant cell or tissue transformation and regeneration of plants are known and available. See, e.g., Gruber *et al.*, "Vectors for Plant Transformation," in *Methods in Plant Molecular Biology and Biotechnology*, *supra*, pp. 89-119.

The isolated polynucleotides or polypeptides may be introduced into the plant by  
35 one or more techniques typically used for direct delivery into cells. Such protocols may vary depending on the type of organism, cell, plant or plant cell, i.e., monocot or dicot,

targeted for gene modification. Suitable methods of transforming plant cells include microinjection (Crossway, *et al.*, (1986) *Biotechniques* 4:320-334 and US Patent Number 6,300,543), electroporation (Riggs, *et al.*, (1986) *Proc. Natl. Acad. Sci. USA* 83:5602-5606, direct gene transfer (Paszkowski, *et al.*, (1984) *EMBO J.* 3:2717-2722) and

5 ballistic particle acceleration (see, for example, Sanford, *et al.*, US Patent Number 4,945,050; WO 91/10725 and McCabe, *et al.*, (1988) *Biotechnology* 6:923-926). Also see, Tomes, *et al.*, "Direct DNA Transfer into Intact Plant Cells Via Microprojectile Bombardment". pp. 197-213 in *Plant Cell, Tissue and Organ Culture, Fundamental Methods*. eds. Gamborg and Phillips. Springer-Verlag Berlin Heidelberg New York,

10 1995; US Patent Number 5,736,369 (meristem); Weissinger, *et al.*, (1988) *Ann. Rev. Genet.* 22:421-477; Sanford, *et al.*, (1987) *Particulate Science and Technology* 5:27-37 (onion); Christou, *et al.*, (1988) *Plant Physiol.* 87:671-674 (soybean); Datta, *et al.*, (1990) *Biotechnology* 8:736-740 (rice); Klein, *et al.*, (1988) *Proc. Natl. Acad. Sci. USA* 85:4305-4309 (maize); Klein, *et al.*, (1988) *Biotechnology* 6:559-563 (maize); WO 91/10725

15 (maize); Klein, *et al.*, (1988) *Plant Physiol.* 91:440-444 (maize); Fromm, *et al.*, (1990) *Biotechnology* 8:833-839 and Gordon-Kamm, *et al.*, (1990) *Plant Cell* 2:603-618 (maize); Hooydaas-Van Slogteren and Hooykaas (1984) *Nature* (London) 311:763-764; Bytebierm, *et al.*, (1987) *Proc. Natl. Acad. Sci. USA* 84:5345-5349 (Liliaceae); De Wet, *et al.*, (1985) *In The Experimental Manipulation of Ovule Tissues*, ed. Chapman, *et al.*,

20 pp. 197-209. Longman, NY (pollen); Kaeppler, *et al.*, (1990) *Plant Cell Reports* 9:415-418; and Kaeppler, *et al.*, (1992) *Theor. Appl. Genet.* 84:560-566 (whisker-mediated transformation); US Patent Number 5,693,512 (sonication); D'Halluin, *et al.*, (1992) *Plant Cell* 4:1495-1505 (electroporation); Li, *et al.*, (1993) *Plant Cell Reports* 12:250-255 and Christou and Ford, (1995) *Annals of Botany* 75:407-413 (rice); Osjoda, *et al.*, (1996)

25 *Nature Biotech.* 14:745-750; Agrobacterium mediated maize transformation (US Patent Number 5,981,840); silicon carbide whisker methods (Frame, *et al.*, (1994) *Plant J.* 6:941-948); laser methods (Guo, *et al.*, (1995) *Physiologia Plantarum* 93:19-24); sonication methods (Bao, *et al.*, (1997) *Ultrasound in Medicine & Biology* 23:953-959; Finer and Finer, (2000) *Lett Appl Microbiol.* 30:406-10; Amoah, *et al.*, (2001) *J Exp Bot*

30 52:1135-42); polyethylene glycol methods (Krens, *et al.*, (1982) *Nature* 296:72-77); protoplasts of monocot and dicot cells can be transformed using electroporation (Fromm, *et al.*, (1985) *Proc. Natl. Acad. Sci. USA* 82:5824-5828) and microinjection (Crossway, *et al.*, (1986) *Mol. Gen. Genet.* 202:179-185), all of which are herein incorporated by reference.

### *Agrobacterium*-mediated Transformation

The most widely utilized method for introducing an expression vector into plants is based on the natural transformation system of *Agrobacterium*. *A. tumefaciens* and *A. rhizogenes* are plant pathogenic soil bacteria, which genetically transform plant cells.

5 The Ti and Ri plasmids of *A. tumefaciens* and *A. rhizogenes*, respectively, carry genes responsible for genetic transformation of plants. See, e.g., Kado, (1991) *Crit. Rev. Plant Sci.* 10:1. Descriptions of the *Agrobacterium* vector systems and methods for *Agrobacterium*-mediated gene transfer are provided in Gruber, *et al.*, *supra*; Miki, *et al.*, *supra* and Moloney, *et al.*, (1989) *Plant Cell Reports* 8:238.

10 Similarly, the gene can be inserted into the T-DNA region of a Ti or Ri plasmid derived from *A. tumefaciens* or *A. rhizogenes*, respectively. Thus, expression cassettes can be constructed as above, using these plasmids. Many control sequences are known which when coupled to a heterologous coding sequence and transformed into a host organism show fidelity in gene expression with respect to tissue/organ specificity of the original coding sequence. See, e.g., Benfey and Chua, (1989) *Science* 244:174-81. Particularly suitable control sequences for use in these plasmids are promoters for constitutive leaf-specific expression of the gene in the various target plants. Other useful control sequences include a promoter and terminator from the nopaline synthase gene (NOS). The NOS promoter and terminator are present in the plasmid pARC2, available from the American Type Culture Collection and designated ATCC 67238. If such a system is used, the virulence (*vir*) gene from either the Ti or Ri plasmid must also be present, either along with the T-DNA portion or via a binary system where the *vir* gene is present on a separate vector. Such systems, vectors for use therein, and methods of transforming plant cells are described in US Patent Number 4,658,082; US Patent Application Serial Number 913,914, filed October 1, 1986, as referenced in US Patent Number 5,262,306, issued November 16, 1993 and Simpson, *et al.*, (1986) *Plant Mol. Biol.* 6:403-15 (also referenced in the '306 patent), all incorporated by reference in their entirety.

30 Once transformed, these cells can be used to regenerate transgenic plants. For example, whole plants can be infected with these vectors by wounding the plant and then introducing the vector into the wound site. Any part of the plant can be wounded, including leaves, stems and roots. Alternatively, plant tissue, in the form of an explant, such as cotyledonary tissue or leaf disks, can be inoculated with these vectors, and cultured under conditions, which promote plant regeneration. Roots or shoots transformed by inoculation of plant tissue with *A. rhizogenes* or *A. tumefaciens*, containing the gene coding for the fumonisin degradation enzyme, can be used as a

source of plant tissue to regenerate fumonisin-resistant transgenic plants, either via somatic embryogenesis or organogenesis. Examples of such methods for regenerating plant tissue are disclosed in Shahin, (1985) *Theor. Appl. Genet.* 69:235-40; US Patent Number 4,658,082; Simpson, *et al.*, *supra*; and US Patent Application Serial Numbers 913,913 and 913,914, both filed October 1, 1986, as referenced in US Patent Number 5,262,306, issued November 16, 1993, the entire disclosures therein incorporated herein by reference.

#### Direct Gene Transfer

Despite the fact that the host range for *Agrobacterium*-mediated transformation is broad, some major cereal crop species and gymnosperms have generally been recalcitrant to this mode of gene transfer, even though some success has recently been achieved in rice (Hiei, *et al.*, (1994) *The Plant Journal* 6:271-82). Several methods of plant transformation, collectively referred to as direct gene transfer, have been developed as an alternative to *Agrobacterium*-mediated transformation.

A generally applicable method of plant transformation is microprojectile-mediated transformation, where DNA is carried on the surface of microprojectiles measuring about 1 to 4  $\mu\text{m}$ . The expression vector is introduced into plant tissues with a biolistic device that accelerates the microprojectiles to speeds of 300 to 600 m/s which is sufficient to penetrate the plant cell walls and membranes (Sanford, *et al.*, (1987) *Part. Sci. Technol.* 5:27; Sanford, (1988) *Trends Biotech* 6:299; Sanford, (1990) *Physiol. Plant* 79:206 and Klein, *et al.*, (1992) *Biotechnology* 10:268).

Another method for physical delivery of DNA to plants is sonication of target cells as described in Zang, *et al.*, (1991) *BioTechnology* 9:996. Alternatively, liposome or spheroplast fusions have been used to introduce expression vectors into plants. See, e.g., Deshayes, *et al.*, (1985) *EMBO J.* 4:2731 and Christou, *et al.*, (1987) *Proc. Natl. Acad. Sci. USA* 84:3962. Direct uptake of DNA into protoplasts using  $\text{CaCl}_2$  precipitation, polyvinyl alcohol or poly-L-ornithine has also been reported. See, e.g., Hain, *et al.*, (1985) *Mol. Gen. Genet.* 199:161 and Draper, *et al.*, (1982) *Plant Cell Physiol.* 23:451.

Electroporation of protoplasts and whole cells and tissues has also been described. See, e.g., Donn, *et al.*, (1990) *Abstracts of the VIIth Int'l. Congress on Plant Cell and Tissue Culture IAPTC*, A2-38, p. 53; D'Halluin, *et al.*, (1992) *Plant Cell* 4:1495-505 and Spencer, *et al.*, (1994) *Plant Mol. Biol.* 24:51-61.

1. *Polynucleotide-Based Methods:*

In some embodiments of the present disclosure, a plant is transformed with an expression cassette that is capable of expressing a polynucleotide that inhibits the expression of FC of the disclosure. The term "expression" as used herein refers to the biosynthesis of a gene product, including the transcription and/or translation of said gene product. For example, for the purposes of the present disclosure, an expression cassette capable of expressing a polynucleotide that inhibits the expression of at least one nitrate uptake-associated polypeptide is an expression cassette capable of producing an RNA molecule that inhibits the transcription and/or translation of at least one nitrate uptake-associated polypeptide of the disclosure. The "expression" or "production" of a protein or polypeptide from a DNA molecule refers to the transcription and translation of the coding sequence to produce the protein or polypeptide, while the "expression" or "production" of a protein or polypeptide from an RNA molecule refers to the translation of the RNA coding sequence to produce the protein or polypeptide.

Examples of polynucleotides that inhibit the expression of FC are given below.

*i. Sense Suppression/Cosuppression*

In some embodiments of the disclosure, inhibition of the expression of FC may be obtained by sense suppression or cosuppression. For cosuppression, an expression cassette is designed to express an RNA molecule corresponding to all or part of a messenger RNA encoding FC in the "sense" orientation. Over expression of the RNA molecule can result in reduced expression of the native gene. Accordingly, multiple plant lines transformed with the cosuppression expression cassette are screened to identify those that show the greatest inhibition of nitrate uptake-associated polypeptide expression.

The polynucleotide used for cosuppression may correspond to all or part of the sequence encoding the nitrate uptake-associated polypeptide, all or part of the 5' and/or 3' untranslated region of FC transcript or all or part of both the coding sequence and the untranslated regions of a transcript encoding FC. In some embodiments where the polynucleotide comprises all or part of the coding region for the nitrate uptake-associated polypeptide, the expression cassette is designed to eliminate the start codon of the polynucleotide so that no protein product will be translated.

Cosuppression may be used to inhibit the expression of plant genes to produce plants having undetectable protein levels for the proteins encoded by these genes. See, for example, Broin, *et al.*, (2002) *Plant Cell* 14:1417-1432. Cosuppression may also be used to inhibit the expression of multiple proteins in the same plant. See, for example,

US Patent Number 5,942,657. Methods for using cosuppression to inhibit the expression of endogenous genes in plants are described in Flavell, *et al.*, (1994) *Proc. Natl. Acad. Sci. USA* 91:3490-3496; Jorgensen, *et al.*, (1996) *Plant Mol. Biol.* 31:957-973; Johansen and Carrington, (2001) *Plant Physiol.* 126:930-938; Broin, *et al.*, (2002) *Plant Cell* 14:1417-1432; Stoutjesdijk, *et al.*, (2002) *Plant Physiol.* 129:1723-1731; Yu, *et al.*, (2003) *Phytochemistry* 63:753-763 and US Patent Numbers 5,034,323, 5,283,184 and 5,942,657, each of which is herein incorporated by reference. The efficiency of cosuppression may be increased by including a poly-dT region in the expression cassette at a position 3' to the sense sequence and 5' of the polyadenylation signal. See, US Patent Publication Number 2002/0048814, herein incorporated by reference. Typically, such a nucleotide sequence has substantial sequence identity to the sequence of the transcript of the endogenous gene, optimally greater than about 65% sequence identity, more optimally greater than about 85% sequence identity, most optimally greater than about 95% sequence identity. See, US Patent Numbers 5,283,184 and 5,034,323, herein incorporated by reference.

*ii. Antisense Suppression*

In some embodiments of the disclosure, inhibition of the expression of the nitrate uptake-associated polypeptide may be obtained by antisense suppression. For antisense suppression, the expression cassette is designed to express an RNA molecule complementary to all or part of a messenger RNA encoding the nitrate uptake-associated polypeptide. Over expression of the antisense RNA molecule can result in reduced expression of the native gene. Accordingly, multiple plant lines transformed with the antisense suppression expression cassette are screened to identify those that show the greatest inhibition of nitrate uptake-associated polypeptide expression.

*iii. Double-Stranded RNA Interference*

In some embodiments of the disclosure, inhibition of the expression of *FC* may be obtained by double-stranded RNA (dsRNA) interference. For dsRNA interference, a sense RNA molecule like that described above for cosuppression and an antisense RNA molecule that is fully or partially complementary to the sense RNA molecule are expressed in the same cell, resulting in inhibition of the expression of the corresponding endogenous messenger RNA.

Expression of the sense and antisense molecules can be accomplished by designing the expression cassette to comprise both a sense sequence and an antisense sequence. Alternatively, separate expression cassettes may be used for the sense and

antisense sequences. Methods for using dsRNA interference to inhibit the expression of endogenous plant genes are described in Waterhouse, *et al.*, (1998) *Proc. Natl. Acad. Sci. USA* 95:13959-13964, Liu, *et al.*, (2002) *Plant Physiol.* 129:1732-1743 and WO 99/49029, WO 99/53050, WO 99/61631 and WO 00/49035, each of which is herein  
5 incorporated by reference.

*iv. Hairpin RNA Interference and Intron-Containing Hairpin RNA Interference*

In some embodiments of the disclosure, inhibition of the expression of FC may  
10 be obtained by hairpin RNA (hpRNA) interference or intron-containing hairpin RNA (ihpRNA) interference. These methods are highly efficient at inhibiting the expression of endogenous genes. See, Waterhouse and Helliwell, (2003) *Nat. Rev. Genet.* 4:29-38 and the references cited therein.

For hpRNA interference, the expression cassette is designed to express an RNA  
15 molecule that hybridizes with itself to form a hairpin structure that comprises a single-stranded loop region and a base-paired stem. The base-paired stem region comprises a sense sequence corresponding to all or part of the endogenous messenger RNA encoding the gene whose expression is to be inhibited and an antisense sequence that is fully or partially complementary to the sense sequence. Alternatively, the base-paired  
20 stem region may correspond to a portion of a promoter sequence controlling expression of the gene to be inhibited. Thus, the base-paired stem region of the molecule generally determines the specificity of the RNA interference. hpRNA molecules are highly efficient at inhibiting the expression of endogenous genes and the RNA interference they induce is inherited by subsequent generations of plants. See, for example, Chuang and  
25 Meyerowitz, (2000) *Proc. Natl. Acad. Sci. USA* 97:4985-4990; Stoutjesdijk, *et al.*, (2002) *Plant Physiol.* 129:1723-1731 and Waterhouse and Helliwell, (2003) *Nat. Rev. Genet.* 4:29-38. Methods for using hpRNA interference to inhibit or silence the expression of genes are described, for example, in Chuang and Meyerowitz, (2000) *Proc. Natl. Acad. Sci. USA* 97:4985-4990; Stoutjesdijk, *et al.*, (2002) *Plant Physiol.* 129:1723-1731;  
30 Waterhouse and Helliwell, (2003) *Nat. Rev. Genet.* 4:29-38; Pandolfini *et al.*, *BMC Biotechnology* 3:7, and US Patent Application Publication Number 2003/0175965, each of which is herein incorporated by reference. A transient assay for the efficiency of hpRNA constructs to silence gene expression *in vivo* has been described by Panstruga, *et al.*, (2003) *Mol. Biol. Rep.* 30:135-140, herein incorporated by reference.

35 For ihpRNA, the interfering molecules have the same general structure as for hpRNA, but the RNA molecule additionally comprises an intron that is capable of being

spliced in the cell in which the ihpRNA is expressed. The use of an intron minimizes the size of the loop in the hairpin RNA molecule following splicing, and this increases the efficiency of interference. See, for example, Smith, *et al.*, (2000) *Nature* 407:319-320. In fact, Smith, *et al.*, show 100% suppression of endogenous gene expression using  
5 ihpRNA-mediated interference. Methods for using ihpRNA interference to inhibit the expression of endogenous plant genes are described, for example, in Smith, *et al.*, (2000) *Nature* 407:319-320; Wesley, *et al.*, (2001) *Plant J.* 27:581-590; Wang and Waterhouse, (2001) *Curr. Opin. Plant Biol.* 5:146-150; Waterhouse and Helliwell, (2003) *Nat. Rev. Genet.* 4:29-38; Helliwell and Waterhouse, (2003) *Methods* 30:289-295 and  
10 US Patent Application Publication Number 2003/0180945, each of which is herein incorporated by reference.

The expression cassette for hpRNA interference may also be designed such that the sense sequence and the antisense sequence do not correspond to an endogenous RNA. In this embodiment, the sense and antisense sequence flank a loop sequence  
15 that comprises a nucleotide sequence corresponding to all or part of the endogenous messenger RNA of the target gene. Thus, it is the loop region that determines the specificity of the RNA interference. See, for example, WO 02/00904; Mette, *et al.*, (2000) *EMBO J* 19:5194-5201; Matzke, *et al.*, (2001) *Curr. Opin. Genet. Devel.* 11:221-227; Scheid, *et al.*, (2002) *Proc. Natl. Acad. Sci., USA* 99:13659-13662; Aufsatz, *et al.*,  
20 (2002) *Proc. Nat'l. Acad. Sci.* 99(4):16499-16506; Sijen, *et al.*, *Curr. Biol.* (2001) 11:436-440), herein incorporated by reference.

#### v. *Amplicon-Mediated Interference*

Amplicon expression cassettes comprise a plant virus-derived sequence that  
25 contains all or part of the target gene but generally not all of the genes of the native virus. The viral sequences present in the transcription product of the expression cassette allow the transcription product to direct its own replication. The transcripts produced by the amplicon may be either sense or antisense relative to the target sequence (i.e., the messenger RNA for the nitrate uptake-associated polypeptide).  
30 Methods of using amplicons to inhibit the expression of endogenous plant genes are described, for example, in Angell and Baulcombe, (1997) *EMBO J.* 16:3675-3684, Angell and Baulcombe, (1999) *Plant J.* 20:357-362 and US Patent Number 6,646,805, each of which is herein incorporated by reference.

vi. *Ribozymes*

In some embodiments, the polynucleotide expressed by the expression cassette of the disclosure is catalytic RNA or has ribozyme activity specific for the messenger RNA. This method is described, for example, in US Patent Number 4,987,071, herein  
5 incorporated by reference.

vii. *Small Interfering RNA or Micro RNA*

In some embodiments of the disclosure, inhibition of the expression of FC may be obtained by RNA interference by expression of a gene encoding a micro RNA (miRNA). miRNAs are regulatory agents consisting of about 22 ribonucleotides. miRNA  
10 are highly efficient at inhibiting the expression of endogenous genes. See, for example Javier, *et al.*, (2003) *Nature* 425:257-263, herein incorporated by reference.

For miRNA interference, the expression cassette is designed to express an RNA molecule that is modeled on an endogenous miRNA gene. The miRNA gene encodes  
15 an RNA that forms a hairpin structure containing a 22-nucleotide sequence that is complementary to another endogenous gene (target sequence). For suppression of nitrate uptake-associated expression, the 22-nucleotide sequence is selected from a nitrate uptake-associated transcript sequence and contains 22 nucleotides of said nitrate uptake-associated sequence in sense orientation and 21 nucleotides of a corresponding  
20 antisense sequence that is complementary to the sense sequence. miRNA molecules are highly efficient at inhibiting the expression of endogenous genes and the RNA interference they induce is inherited by subsequent generations of plants.

2. *Polypeptide-Based Inhibition of Gene Expression*

25 In one embodiment, the polynucleotide encodes a zinc finger protein that binds to a gene encoding FC, resulting in reduced expression of the gene. In particular embodiments, the zinc finger protein binds to a regulatory region of a nitrate uptake-associated gene. In other embodiments, the zinc finger protein binds to a messenger RNA encoding FC and prevents its translation. Methods of selecting sites for targeting  
30 by zinc finger proteins have been described, for example, in US Patent Number 6,453,242 and methods for using zinc finger proteins to inhibit the expression of genes in plants are described, for example, in US. Patent Application Publication Number 2003/0037355, each of which is herein incorporated by reference.

### 3. *Polypeptide-Based Inhibition of Protein Activity*

In some embodiments of the disclosure, the polynucleotide encodes an antibody that binds to polypeptide of the disclosure. The expression of antibodies in plant cells and the inhibition of molecular pathways by expression and binding of antibodies to proteins in plant cells are well known in the art. See, for example, Conrad and Sonnewald, (2003) *Nature Biotech.* 21:35-36, incorporated herein by reference.

### 4. *Gene Disruption*

In some embodiments of the present disclosure, the activity of FC is reduced or eliminated by disrupting the gene encoding the nitrate uptake-associated polypeptide. The gene encoding the nitrate uptake-associated polypeptide may be disrupted by any method known in the art. For example, in one embodiment, the gene is disrupted by transposon tagging. In another embodiment, the gene is disrupted by mutagenizing plants using random or targeted mutagenesis and selecting for plants that have reduced nitrogen utilization activity.

#### *i. Mutant Plants with Reduced Activity*

Additional methods for decreasing or eliminating the expression of endogenous genes in plants are also known in the art and can be similarly applied to the instant disclosure. These methods include other forms of mutagenesis, such as ethyl methanesulfonate-induced mutagenesis, deletion mutagenesis, and fast neutron deletion mutagenesis used in a reverse genetics sense (with PCR) to identify plant lines in which the endogenous gene has been deleted. For examples of these methods see, Ohshima, *et al.*, (1998) *Virology* 243:472-481; Okubara, *et al.*, (1994) *Genetics* 137:867-874 and Quesada, *et al.*, (2000) *Genetics* 154:421-436, each of which is herein incorporated by reference. In addition, a fast and automatable method for screening for chemically induced mutations, TILLING (Targeting Induced Local Lesions In Genomes), using denaturing HPLC or selective endonuclease digestion of selected PCR products is also applicable to the instant disclosure. See, McCallum, *et al.*, (2000) *Nat. Biotechnol.* 18:455-457, herein incorporated by reference.

Mutations that impact gene expression or that interfere with the function (enhanced nitrogen utilization activity) of the encoded protein are well known in the art. Insertional mutations in gene exons usually result in null-mutants. Mutations in conserved residues are particularly effective in inhibiting the activity of the encoded protein. Conserved residues of plant nitrate uptake-associated polypeptides suitable for mutagenesis with the goal to eliminate nitrate uptake-associated activity have been

described. Such mutants can be isolated according to well-known procedures, and mutations in different nitrate uptake-associated loci can be stacked by genetic crossing. See, for example, Gruis, *et al.*, (2002) *Plant Cell* 14:2863-2882.

In another embodiment of this disclosure, dominant mutants can be used to trigger RNA silencing due to gene inversion and recombination of a duplicated gene locus. See, for example, Kusaba, *et al.*, (2003) *Plant Cell* 15:1455-1467.

The disclosure encompasses additional methods for reducing or eliminating the activity of one or more nitrate uptake-associated polypeptide. Examples of other methods for altering or mutating a genomic nucleotide sequence in a plant are known in the art and include, but are not limited to, the use of RNA:DNA vectors, RNA:DNA mutational vectors, RNA:DNA repair vectors, mixed-duplex oligonucleotides, self-complementary RNA:DNA oligonucleotides and recombinogenic oligonucleobases. Such vectors and methods of use are known in the art. See, for example, US Patent Numbers 5,565,350; 5,731,181; 5,756,325; 5,760,012; 5,795,972 and 5,871,984, each of which are herein incorporated by reference. See also, WO 98/49350, WO 99/07865, WO 99/25821 and Beetham, *et al.*, (1999) *Proc. Natl. Acad. Sci. USA* 96:8774-8778, each of which is herein incorporated by reference.

#### vi. *Modulating Reproductive Tissue Development*

Methods for modulating reproductive tissue development are provided. In one embodiment, methods are provided to modulate floral development in a plant. By "modulating floral development" is intended any alteration in a structure of a plant's reproductive tissue as compared to a control plant in which the activity or level of the nitrate uptake-associated polypeptide has not been modulated. "Modulating floral development" further includes any alteration in the timing of the development of a plant's reproductive tissue (i.e., a delayed or an accelerated timing of floral development) when compared to a control plant in which the activity or level of the nitrate uptake-associated polypeptide has not been modulated. Macroscopic alterations may include changes in size, shape, number, or location of reproductive organs, the developmental time period that these structures form or the ability to maintain or proceed through the flowering process in times of environmental stress. Microscopic alterations may include changes to the types or shapes of cells that make up the reproductive organs.

In general, methods to modify or alter the host endogenous genomic DNA are available. This includes altering the host native DNA sequence or a pre-existing transgenic sequence including regulatory elements, coding and non-coding sequences. These methods are also useful in targeting nucleic acids to pre-engineered target

recognition sequences in the genome. As an example, the genetically modified cell or plant described herein, is generated using "custom" or engineered endonucleases such as meganucleases produced to modify plant genomes (see e.g., WO 2009/114321; Gao et al. (2010) Plant Journal 1:176-187). Another site-directed engineering is through the use of zinc finger domain recognition coupled with the restriction properties of restriction enzyme. See e.g., Urnov, et al., (2010) Nat Rev Genet. 11(9):636-46; Shukla, et al., (2009) Nature 459 (7245):437-41. A transcription activator-like (TAL) effector-DNA modifying enzyme (TALE or TALEN) is also used to engineer changes in plant genome. See e.g., US20110145940, Cermak et al., (2011) Nucleic Acids Res. 39(12) and Boch et al., (2009), Science 326(5959): 1509-12. Site-specific modification of plant genomes can also be performed using the bacterial type II CRISPR (clustered regularly interspaced short palindromic repeats)/Cas (CRISPR-associated) system. See e.g., Belhaj et al., (2013), Plant Methods 9: 39; The CRISPR/Cas system allows targeted cleavage of genomic DNA guided by a customizable small noncoding RNA. Based on the disclosure of the FC coding sequences, polypeptide sequences of the orthologs/homologs and the genomic DNA sequences, site-directed mutagenesis can be readily performed to generate plants expressing a higher level of the endogenous FC polypeptide or an ortholog thereof.

Antibodies to a FC polypeptide disclosed herein or the embodiments or to variants or fragments thereof are also encompassed. The antibodies of the disclosure include polyclonal and monoclonal antibodies as well as fragments thereof which retain their ability to bind to FC polypeptide disclosed herein. An antibody, monoclonal antibody or fragment thereof is said to be capable of binding a molecule if it is capable of specifically reacting with the molecule to thereby bind the molecule to the antibody, monoclonal antibody or fragment thereof. The term "antibody" (Ab) or "monoclonal antibody" (Mab) is meant to include intact molecules as well as fragments or binding regions or domains thereof (such as, for example, Fab and F(ab)<sub>2</sub> fragments) which are capable of binding hapten. Such fragments are typically produced by proteolytic cleavage, such as papain or pepsin. Alternatively, hapten-binding fragments can be produced through the application of recombinant DNA technology or through synthetic chemistry. Methods for the preparation of the antibodies of the present disclosure are generally known in the art. For example, see, Antibodies, A Laboratory Manual, Ed Harlow and David Lane (eds.) Cold Spring Harbor Laboratory, N.Y. (1988), as well as the references cited therein. Standard reference works setting forth the general principles of immunology include: Klein, J. Immunology: The Science of Cell-Noncell Discrimination, John Wiley & Sons, N.Y. (1982); Dennett, *et al.*, Monoclonal Antibodies,

Hybridoma: A New Dimension in Biological Analyses, Plenum Press, N.Y. (1980) and Campbell, "Monoclonal Antibody Technology," In Laboratory Techniques in Biochemistry and Molecular Biology, Vol. 13, Burdon, *et al.*, (eds.), Elsevier, Amsterdam (1984). See also, US Patent Numbers 4,196,265; 4,609,893; 4,713,325; 4,714,681; 4,716,111; 5 4,716,117 and 4,720,459. PtIP-50 polypeptide or PtIP-65 polypeptide antibodies or antigen-binding portions thereof can be produced by a variety of techniques, including conventional monoclonal antibody methodology, for example the standard somatic cell hybridization technique of Kohler and Milstein, (1975) *Nature* 256:495. Other techniques for producing monoclonal antibody can also be employed such as viral or oncogenic 10 transformation of B lymphocytes. An animal system for preparing hybridomas is a murine system. Immunization protocols and techniques for isolation of immunized splenocytes for fusion are known in the art. Fusion partners (e.g., murine myeloma cells) and fusion procedures are also known. The antibody and monoclonal antibodies of the disclosure can be prepared by utilizing a FC polypeptide disclosed herein as 15 antigens.

A kit for detecting the presence of a FC polypeptide disclosed herein or detecting the presence of a nucleotide sequence encoding a FC polypeptide disclosed herein, in a sample is provided. In one embodiment, the kit provides antibody-based reagents for detecting the presence of a FC polypeptide disclosed herein in a tissue sample. In 20 another embodiment, the kit provides labeled nucleic acid probes useful for detecting the presence of one or more polynucleotides encoding FC polypeptide disclosed herein. The kit is provided along with appropriate reagents and controls for carrying out a detection method, as well as instructions for use of the kit.

As discussed above, one of skill will recognize the appropriate promoter to use to 25 modulate floral development of the plant. Exemplary promoters for this embodiment include constitutive promoters, inducible promoters, shoot-preferred promoters and inflorescence-preferred promoters.

Genes of interest are reflective of the commercial markets and interests of those involved in the development of the crop. Crops and markets of interest change, and as 30 developing nations open up world markets, new crops and technologies will emerge also. In addition, as our understanding of agronomic traits and characteristics such as yield and heterosis increase, the choice of genes for transformation will change accordingly. General categories of genes of interest include, for example, those genes involved in information, such as zinc fingers, those involved in communication, such as 35 kinases and those involved in housekeeping, such as heat shock proteins. More specific categories of transgenes, for example, include genes encoding important traits for

agronomics, insect resistance, disease resistance, herbicide resistance, sterility, grain characteristics and commercial products. Genes of interest include, generally, those involved in oil, starch, carbohydrate, or nutrient metabolism as well as those affecting kernel size, sucrose loading, and the like.

5 In certain embodiments the nucleic acid sequences of the present disclosure can be used in combination ("stacked") with other polynucleotide sequences of interest in order to create plants with a desired phenotype.

This disclosure can be better understood by reference to the following non-limiting examples. It will be appreciated by those skilled in the art that other  
10 embodiments of the disclosure may be practiced without departing from the spirit and the scope of the disclosure as herein disclosed and claimed.

Sequences described in this disclosure include: HvFCI amino acid sequence (SEQ ID NO: 1); HvFCII amino acid sequence (SEQ ID NO: 2); HvFCI polynucleotide sequence (SEQ ID NO: 3); and HvFCII polynucleotide sequence (SEQ ID NO: 4).

15

## EXAMPLES

### EXAMPLE 1 – Identification of barley ferrochelatases (HvFC) and their expression patterns

20 The two barley FC isoforms are 55.6% and 11.2% identical to each other at the amino acid and nucleotide levels, respectively. Similarity comparisons revealed that the two HvFCs share a high level of identity with their Arabidopsis orthologues (AtFC1 (62.3%) and AtFC2 (71.2%), respectively). Multiple sequence alignment revealed that the HvFC1 and HvFC2 catalytic domains are highly conserved. Several proline and  
25 glycine residues, which play vital roles in hydrogen bonding, metal binding, and the stability of the protoporphyrin-interacting loop are also highly conserved. FC2 contains an additional chlorophyll a/b binding (CAB) domain which has a light harvesting complex (LHC) motif. This domain is present in many photosynthesis-associated proteins.

High amino acid conservation in the catalytic domains is indicative of shared  
30 catalytic function for HvFC1 and HvFC2. FC catalyses the conversion of Proto IX into heme, a terminal step in the tetrapyrrole biosynthesis pathway.

Despite catalytic domain commonality, plant FC polypeptides form two distinct phylogenetic lineages (Fig 2). These two lineages are unlikely to have arisen from segmental duplication and are separated by the presence of a characteristic C-terminal  
35 CAB domain containing a conserved LHC motif. HvFC2, as with other plant FC2 sequences, contains this domain which is connected to the FC2 catalytic core by a

proline-rich linker sequence. The LHC motif is abundant in proteins associated with light harvesting complex and is involved in anchoring the complex to the chloroplast membrane, binding chlorophyll and carotenoids, and facilitating interactions with other co-localised proteins. Expression of HvFC1 differs compared to HvFC2. HvFC1 and HvFC2 have similar transcript levels within photosynthetic tissues, but HvFC1 is more highly expressed in non-photosynthetic tissues (Fig 3). Together with structural divergence between the two isoforms these differential expression patterns indicate that HvFC1 and HvFC2 may have distinct roles in barley.

The putative evolutionary relationship between HvFCs and those from other grass and dicot species was investigated by constructing a phylogenetic tree (Hall 2013). The resulting dendrogram demonstrated that the two FC isoforms in all plant species studied so far belong to distinct clades (Fig 2).

Expression of HvFCI and HvFCII are shown in Fig. 4. ROS ( $H_2O_2$ ) generation by Paraquat application and subsequent HvFC expression at different times are shown in Fig. 4 (B). Stress responsive HvFC expression and expression of HvFC in different tissues indicate that modulation of HvFC expression result in protection against oxidative stress. In addition, expression of HvCatalase and HvSOD<sub>2</sub> were also measured in response to paraquat application.

#### **EXAMPLE 2 – HvFCI is transcriptionally responsive to ROS and drought stress**

HvFC expression at different times following drought stress (B) and in the *tigrina* mutant background (A) are shown in Fig. 5. In addition, expression of HvCatalase and HvSOD2 were also measured.

#### **EXAMPLE 3- Ferrochelatase overexpressing transgenic plants are phenotypically similar to wild type and null**

Transgenic plants overexpressing ferrochelatases (HvFCI or HvFCII) displayed normal phenotype as measured by the plant height, number of leaves, shoot dry weight, root dry weight, and shoot/root ratio as shown in the table below and as compared to a wild-type control plant (Null).

**Table 1:** Phenotypic characteristics of HvFC1 and HvFC2 expressing plants

| Line | Plant height (cm) | Number of leaves | Shoot dry weight (mg) | Root dry weight (mg) | Shoot/Root |
|------|-------------------|------------------|-----------------------|----------------------|------------|
| GP   | 5.7 bc            | 4.2a             | 45 abc                | 14 a                 | 3.2 a      |

|                  |        |       |          |        |       |
|------------------|--------|-------|----------|--------|-------|
| Null             | 5.4 bc | 4.0 a | 50.2 bc  | 18.6 a | 3.4 a |
| 2X35S::HvFCI-28  | 5.7 bc | 4.0 a | 61.5 c   | 15.3 a | 4.1 a |
| 2X35S::HvFCI-13  | 5.3 bc | 4.0 a | 38.6 ab  | 15.1 a | 3.5 a |
| 2X35S::HvFCI-17  | 4.3 a  | 3.5 a | 29.4 a   | 8.7 a  | 3.2 a |
| 2X35S::HvFCII-29 | 5.6 bc | 4.3 a | 44 abc   | 11.3 a | 1.8 a |
| 2X35S::HvFCI-25  | 6.3 c  | 4.1 a | 48.8 abc | 16.6 a | 2.3 a |
| 2X35S::HvFCI-9   | 4.9 ab | 3.7 a | 53.6 bc  | 22.2 a | 4.7 a |

Means within each column followed by the same letter did not significantly differ at  $P \leq 0.05$  according to Duncan's multiple range test.

#### 5 **EXAMPLE 4 - Barley FC overexpressing transgenics have higher photosynthetic performance**

Photosynthesis is a highly complex and highly-regulated process ultimately determined by three factors ( $A_{\text{sat}}$ ,  $g_s$ , and CE). Overexpression of HvFC1 and HvFC2 each significantly improve  $A_{\text{sat}}$  (+13%),  $g_s$  (+16%) and CE (+11%) (Fig. 5), indicating  
 10 that both barley FC isoforms are directly involved in photosynthesis or the regulation of photosynthetic components.

Barley transgenics (cv. Golden Promise) that ectopically expressed either HvFC1 or HvFC2 were generated. Coding regions of FC were cloned under the control of the 2x35SCaMV promoter. Twenty-nine independent T0 transgenic lines were obtained for  
 15 each FC construct, using Agrobacterium-mediated transformation. Southern blot hybridization showed that most T0 transgenic lines had 2-5 copies of the transgene. Low copy number transgenic lines were selected and confirmed for transgene copy number by qPCR and subsequently analysed for FC expression by quantitative RT-PCR. Three  
 20 single-copy transgenic lines, each ectopically overexpressing either HvFC1 or HvFC2, were selected for further analysis. The highest HvFC1 and HvFC2 overexpressing

transgenic lines demonstrated increased HvFC protein content in leaves relative to wild-type and null controls. T2 transgenic plants were phenotypically evaluated under controlled conditions for growth and development. Untransformed plants and non-transgenic sibs (null segregants) were used as controls.

5 Molecular characterization of these transgenic lines confirmed that HvFC1 and HvFC2 were constitutively overexpressed and showed no obvious negative developmental defects relative to untransformed and null controls (Table 1). Four-week old T2 transgenic plants (with the exception of line 2x35S::FC1-17) did not show a significant difference in plant height, leaf number, tiller number and shoot or root  
10 biomass when compared to controls. Total chlorophyll content and chlorophyll a/b ratios were similar across all transgenic lines and relative to controls (one-way ANOVA,  $P < 0.05$ ) (Fig. 5).  $A_{\text{sat}}$  increased 13% when comparing transgenic lines to controls, however no significant differences (one-way ANOVA,  $P < 0.05$ ) were observed between 2x35S::FC1 and 2x35S::FC2 transgenics (Fig. 5). Stomatal conductance ( $g_s$ ) was  
15 observed to be 16 % higher in two of the three 2x35S::FC1 lines and only one of the 2x35S::FC2 lines relative to controls (Fig. 5). Whereas CE was observed to be 11% higher in all three 2x35S::FC1 lines and two of the three 2x35S::FC2 lines when compared to controls (Fig. 5). These findings show that HvFC1 and FC2 genes, when ectopically overexpressed, are able to significantly increase photosynthetic rate ( $A_{\text{sat}}$ ) in  
20 barley, therefore demonstrating that the FC isoforms are likely to play important roles during photosynthesis.

Leaf nitrogen content, as a surrogate indicator for the amount of Rubisco was measured in transgenic plants relative to untransformed controls and null segregants. Total leaf N concentration was not significantly different between transgenics and  
25 controls (one-way ANOVA,  $P < 0.05$ ), except for one line (2x35S::FC2-29) which showed a lower concentration (Fig. 9a). These results indicate that the improved photosynthetic performance of the transgenic lines is unlikely to be a consequence of increased Rubisco content.

Because FCs catalyse the insertion of ferrous iron ( $\text{Fe}^{2+}$ ) into protoporphyrin IX, it  
30 is possible that the observed photosynthetic differences may be a consequence of altered Fe homeostasis. Total Fe concentration was measure in photosynthesizing leaf tissue. No significant differences were observed between leaf Fe concentration of the transgenic and control lines. These results indicate that the observed phenotypic differences in photosynthetic performance are not likely to be the consequence of  
35 altered Fe acquisition and/or distribution (Fig. 9b).

Transgenic barley plants were grown at 22 /18 0C, 350-400 lux, 55-60% RH, 12/12hr light/dark. Chlorophyll content, photosynthetic rate, stomatal conductance, and carboxylation efficiency were measured (FIG. 5). As shown in Fig. 5, transgenic plants showed improved Chlorophyll content, photosynthetic rate, stomatal conductance, and carboxylation efficiency compared to the control plants.

Without being limited by a particular theory or hypothesis, HvFC1 and HvFC2 ectopic overexpressors may induce an increase in the free heme pool, which may, in turn, trigger nuclear gene expression for enzymes that affect carboxylation rate.

Based on the disclosure and guidance provided herein, tetrapyrrole biosynthesis is a simple target for engineering photosynthetic yield potential, a trait considered as physiologically complex. The molecular identity of these gene sequences now allows beneficial expression alleles to be identified, tracked and ultimately deployed into cereal breeding programs.

**EXAMPLE 5 - FC overexpressing transgenics maintain higher leaf relative water content (RWC) under drought stress and showed higher WUE under drought**

Drought stress profile is shown in Fig. 6. Also shown in Fig. 6 is the relative water content (RWC) of transgenic plant expressing ferrochelatases under drought stress. As shown herein, higher RWC indicates drought tolerance and the ability of the transgenic plants to withstand lower water availability. Fig. 7 demonstrates both instantaneous water use efficiency (WUE) and intrinsic WUE as measured by  $\mu\text{mol CO}_2.\text{mmolH}_2\text{O}.\text{m}^{-2}.\text{s}^{-1}$ . Both HvFC1 and HvFCII expressing plants show improved WUE as compared to control plants

**EXAMPLE 6 - Barley FC1 is targeted to plastids**

A transient expression assay in onion epidermal cells (*Allium cepa* L.) was used. HvFC1-GFP fusion proteins were detected in either irregular or oval shaped structures consistent with the size and morphology of onion cell proplastids and associated stromules. GFP fluorescence was not detected in small punctate structures, as expected if it were localized to mitochondria.

**Example 7: FC transgenics improve photosynthetic performance under drought stress**

Several photosynthetic performance characteristics were measured as shown in Fig. 8. Total chlorophyll content, photosynthesis rate, stomatal conductance and transpiration rates improved when transgenic plants expressed higher levels of

ferrochelatases as compared to the control plants indicating a direct relation to drought tolerance or drought avoidance.

5 **Example 8: Barley FC overexpressing transgenics have higher expression of genes involved in ROS scavenging.**

Higher expression of catalase, superoxide dismutase (SOD) and ascorbate peroxidase (APX) in FC overexpressing transgenic plants demonstrate that increased expression of ferrochelatases result in improved oxidative stress response.

10

**Example 9: HvFC overexpression reduces red fluorescence and photo-bleaching phenotypes of tigrina d12 mutants.**

*tigrina d<sup>12</sup>* mutant was used as a phenotypic reporter for mitigating some of the oxidative stress conditions of the mutants. *tigrina d<sup>12</sup>* mutants overexpressing *HvFC* were developed. Seeds were germinated under dark in petri plates. *tigrina d<sup>12</sup>* mutants were identified under UV light. *tigrina d<sup>12</sup>* mutant phenotype was confirmed by using CAPS markers. Transgenes specific primers were used to detect the presence of transgene. In summary, FC overexpression reduced damage to PSII system by *tigrina* d12 mutants. PSII efficiency by chlorophyll fluorescence was also demonstrated by the overexpression of FC in the *tigrina* mutant background. Red fluorescence was by image analysis; and photo-bleaching was by determination of chlorophyll content.

**EXAMPLE 10 – Crop improvement by increasing ferrochelatase expression**

25 FCI and FCII are differentially responsive to stress conditions. Ectopic overexpression of HvFCI and HvFCII improves photosynthesis performance. Ectopic overexpression of HvFCI and HvFCII improves plant performance under drought by maintaining higher RWC, higher instantaneous and intrinsic WUE, higher chlorophyll content, higher photosynthesis rate, higher expression of genes encoding ROS scavenging enzymes. Further, ectopic overexpression of ferrochelatase prevents excessive accumulation of chlorophyll branch intermediates and therefore potential photo-oxidative damages.

Increased expression of targeted ferrochelatase genes in a crop plant confer tolerance to herbicides targeting components of the plant heme or porphyrin system. 35 The crop plants gain tolerance to, for example, heme or porphyrin pathway targeting herbicides whereas weeds will be susceptible. The herbicide tolerance may be enhanced by ferrochelatase expression, targeted expression within the cell, and also by cooperative expression from other genes, especially those that contribute to mitigation of

side-effects of singlet oxygen and other reactive oxygen species, which can emanate from heme systems challenged with herbicide stress, and even by abiotic stresses such as drought. Examples of such 'companion' genes may be super oxide dismutase, super oxide reductase, catalase, thioredoxin, among others. In addition, engineering resident  
5 ferrochelatase genes, by selective targeted changes to the gene structure, as by mutagenesis and genome editing (such as via the CRSPR-CAS9 system), can similarly result in heme-targeting herbicide tolerance.

The disclosure has been described with reference to various specific and preferred embodiments and techniques. However, it should be understood that many  
10 variations and modifications may be made while remaining within the spirit and scope of the disclosure.

## WHAT IS CLAIMED IS:

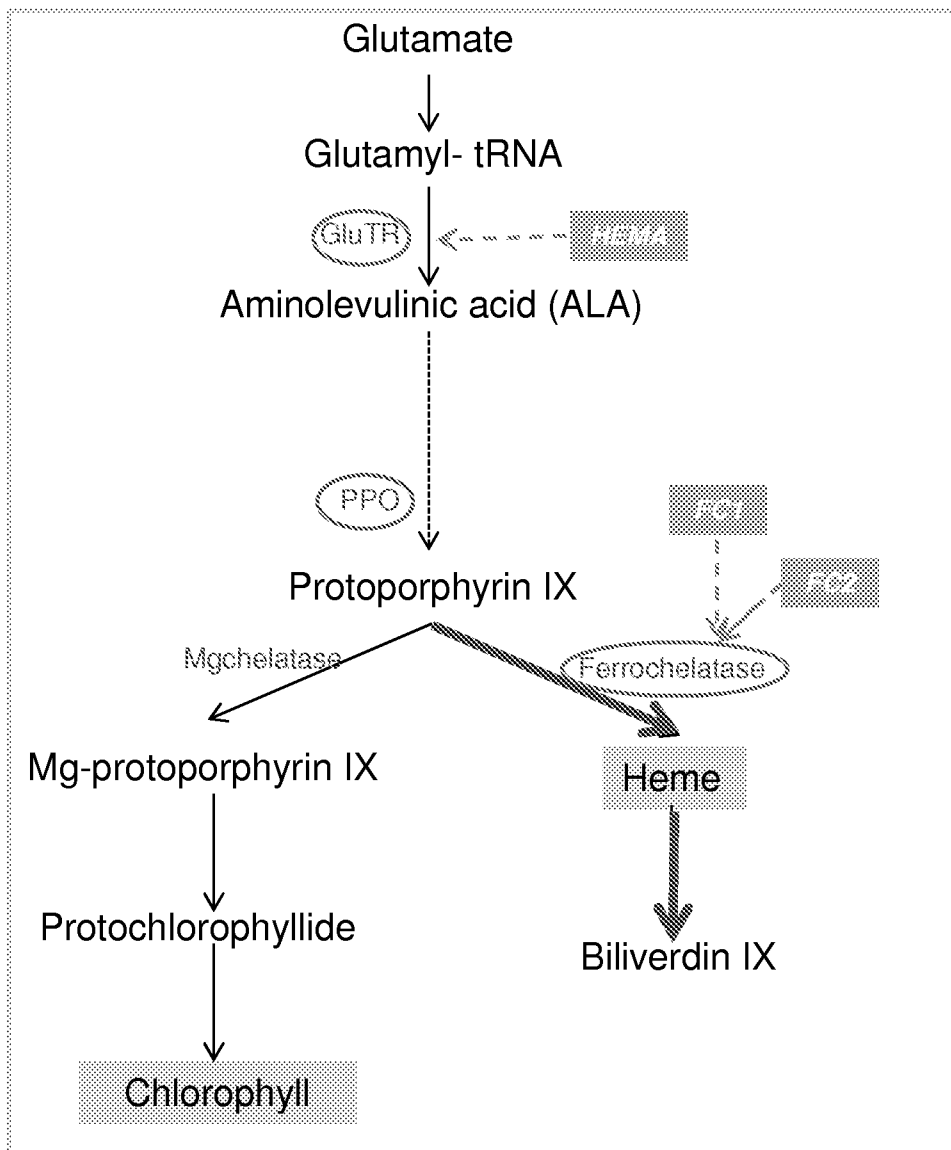
1. A method of improving an agronomic characteristic of a plant, the method comprising modulating the expression of (i) a polynucleotide encoding an amino acid sequence comprising SEQ ID NO: 1 or 2 or an amino acid sequence that is at least 95% identical to one of SEQ ID NO: 1 or 2 (ii) a polynucleotide that hybridizes under stringent hybridization conditions to a fragment of polynucleotide comprising SEQ ID NO: 3 or 4, wherein the fragment comprises at least 100 contiguous nucleotides of SEQ ID NO: 3 or 4 (iii) a polynucleotide that encodes an amino acid sequence that is at least 90% identical to SEQ ID NO: 1 or 2, (iv) a polynucleotide encoding a polypeptide comprising one or more deletions or insertions or substitutions of amino acids compared to SEQ ID NO: 1 or 2.
2. The method of claim 1, wherein the expression of the polynucleotide encoding a polypeptide having at least 95% identity to SEQ ID NO: 1 or 2 is increased by transforming the plant with a recombinant polynucleotide operably linked to a heterologous promoter.
3. The method of claim 1, wherein the expression of an endogenous polynucleotide encoding a polypeptide having at least 95% identity to SEQ ID NO: 1 or 2 is increased by upregulating a regulatory element operably associated with the endogenous polynucleotide.
4. The method of claim 1, wherein the expression of the polynucleotide is increased by expressing the polynucleotide under a tissue preferred regulatory element.
5. The method of claim 1, wherein the agronomic characteristic is selected from the group consisting of wilting avoidance, improved photosynthetic performance, increased chlorophyll content, increased photosynthetic rate, improved stomatal conductance, carboxylation efficiency, an increase in grain size, an increase in grain weight, an increase in grain yield, an increase in grain filling rate, and an increase in biomass.
6. The method of claim 1, wherein the agronomic performance is an increase in drought tolerance during vegetative and/or reproductive stages.
7. The method of claim 1, wherein the grain weight is increased in relation to a control plant not having an increased expression of the polynucleotide.
8. The method of claim 1, wherein the plant is a monocot.
9. The method of claim 1, wherein the plant is selected from the group consisting of maize, rice, wheat, barley, rye, millet, and sorghum.
10. The method of claim 1, wherein the plant is a dicot.

11. The method of claim 1, wherein the plant is soybean or brassica.
12. A method of improving yield of a plant, the method comprising increasing the expression of a polynucleotide that encodes a polypeptide comprising the amino acid sequence selected from the group consisting of SEQ ID NOS: 1, 2 or an allelic variant thereof.
13. A method of improving rice grain yield, the method comprising the expression of a polynucleotide that encodes a polypeptide comprising the amino acid sequence of SEQ ID NO: 1 or a variant thereof.
14. A method of marker assisted selection of a plant for improved yield, the method comprising:
- performing marker-assisted selection of plants that have one or more variations in genomic regions encoding a protein comprising SEQ ID NO: 1 or 2 or a variant thereof or a regulatory sequence thereof; and
  - identifying the plant that exhibits higher yield.
15. A method of identifying one or more alleles in a population of plants that are associated with increased grain yield, the method comprising:
- evaluating in a population of plants for one or more allelic variations in (i) a genomic region, the genomic region encoding a polypeptide or (ii) the regulatory region controlling the expression of the polypeptide, wherein the polypeptide comprises the amino acid sequence of SEQ ID NO: 1 or 2 or a sequence that is 95% identical to SEQ ID NO: 1 or 2;
  - obtaining phenotypic values of increased yield for the one or more plants in the population;
  - associating the allelic variations in the genomic region with the phenotype; and
  - identifying the one or more alleles that are associated with increased yield.
16. A recombinant expression cassette, comprising the polynucleotide of claim 1, wherein the polynucleotide is operably linked to a regulatory element, wherein the expression cassette is functional in a plant cell.
17. A host cell comprising the expression cassette of claim 16.
18. A transgenic plant comprising the recombinant expression cassette of claim 16.
19. A transgenic plant comprising a plant regulatory element that operably regulates the expression of a polynucleotide encoding a polypeptide comprising the amino acid sequence of SEQ ID NO: 1 or 2 or a variant or an ortholog thereof, wherein the regulatory element is heterologous to the polynucleotide and wherein the

transgenic plant exhibits substantially normal phenotype as compared to a control plant.

20. The plant of claim 19, wherein the plant is a wheat, rice, barley, maize, rye, millet, sorghum, triticale, oat, teff, wild rice, spelt, buckwheat, turf grass, rye grass, switchgrass, Miscanthus, or Festuca plant.
- 5

1/9

**FIG. 1**

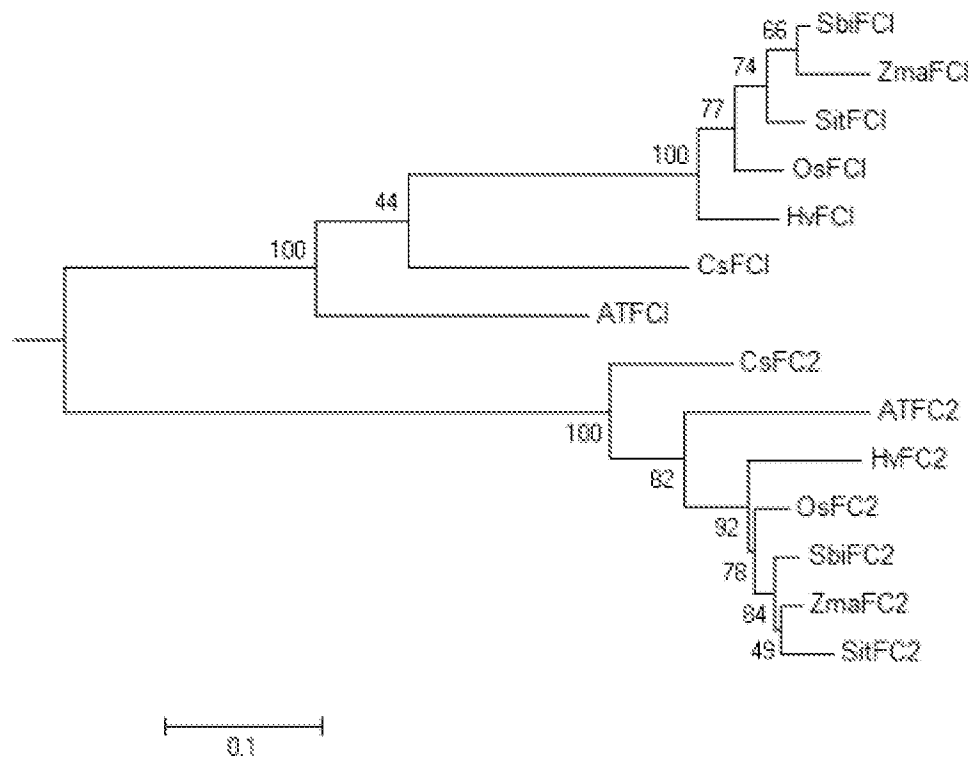


FIG. 2

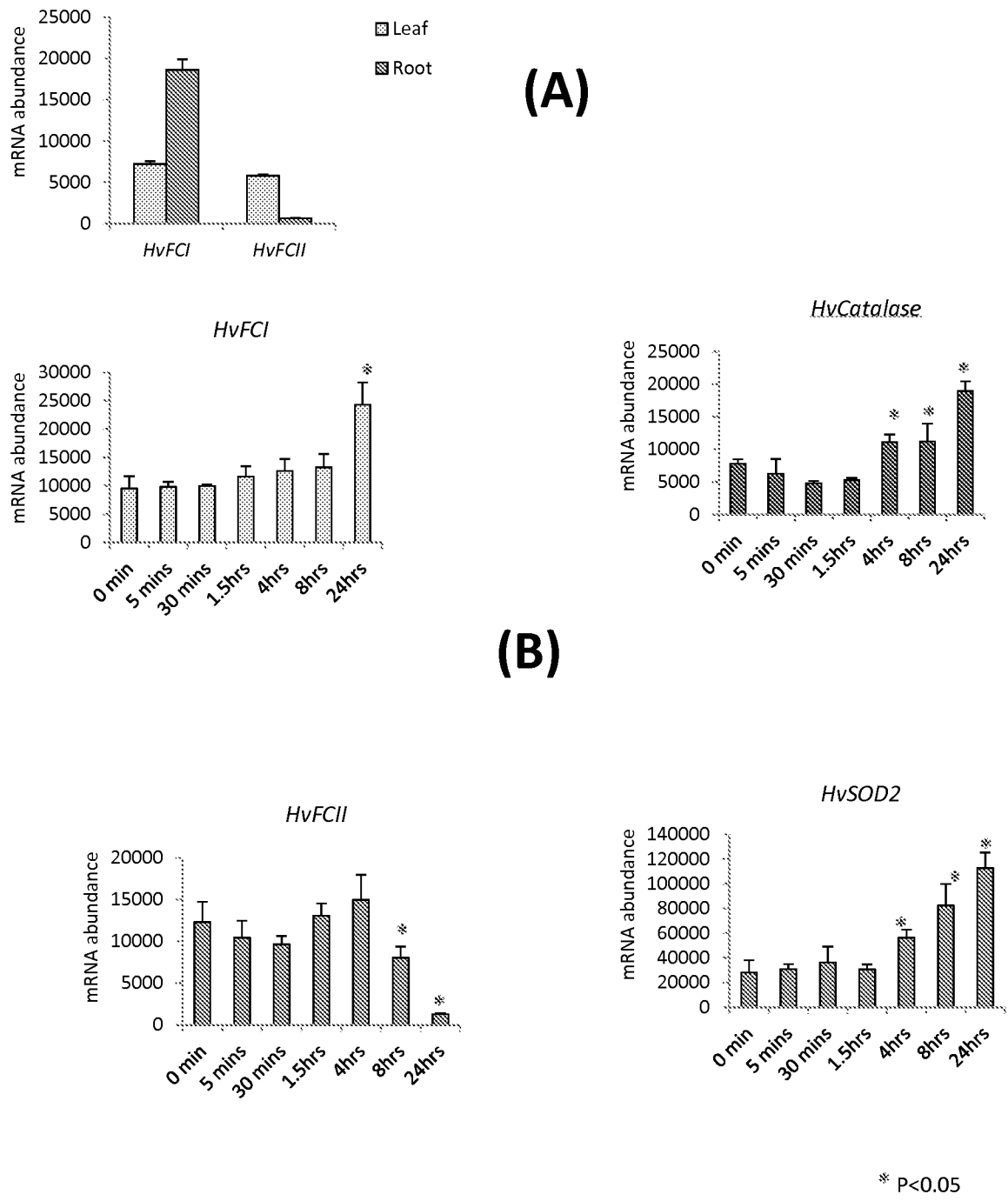
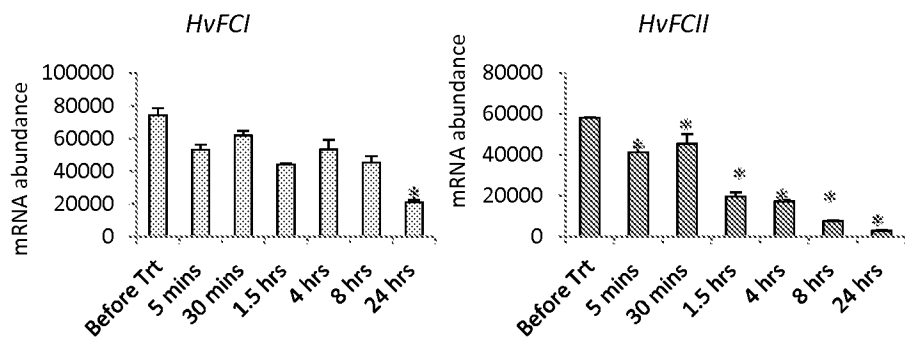


FIG. 3

4/9

(A)



(B)

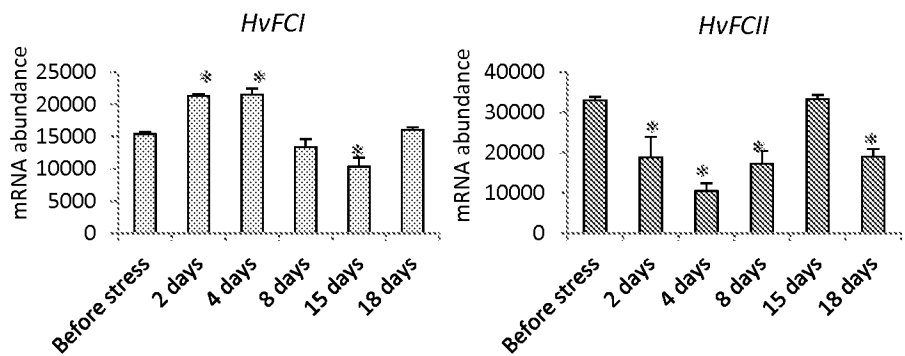


FIG. 4

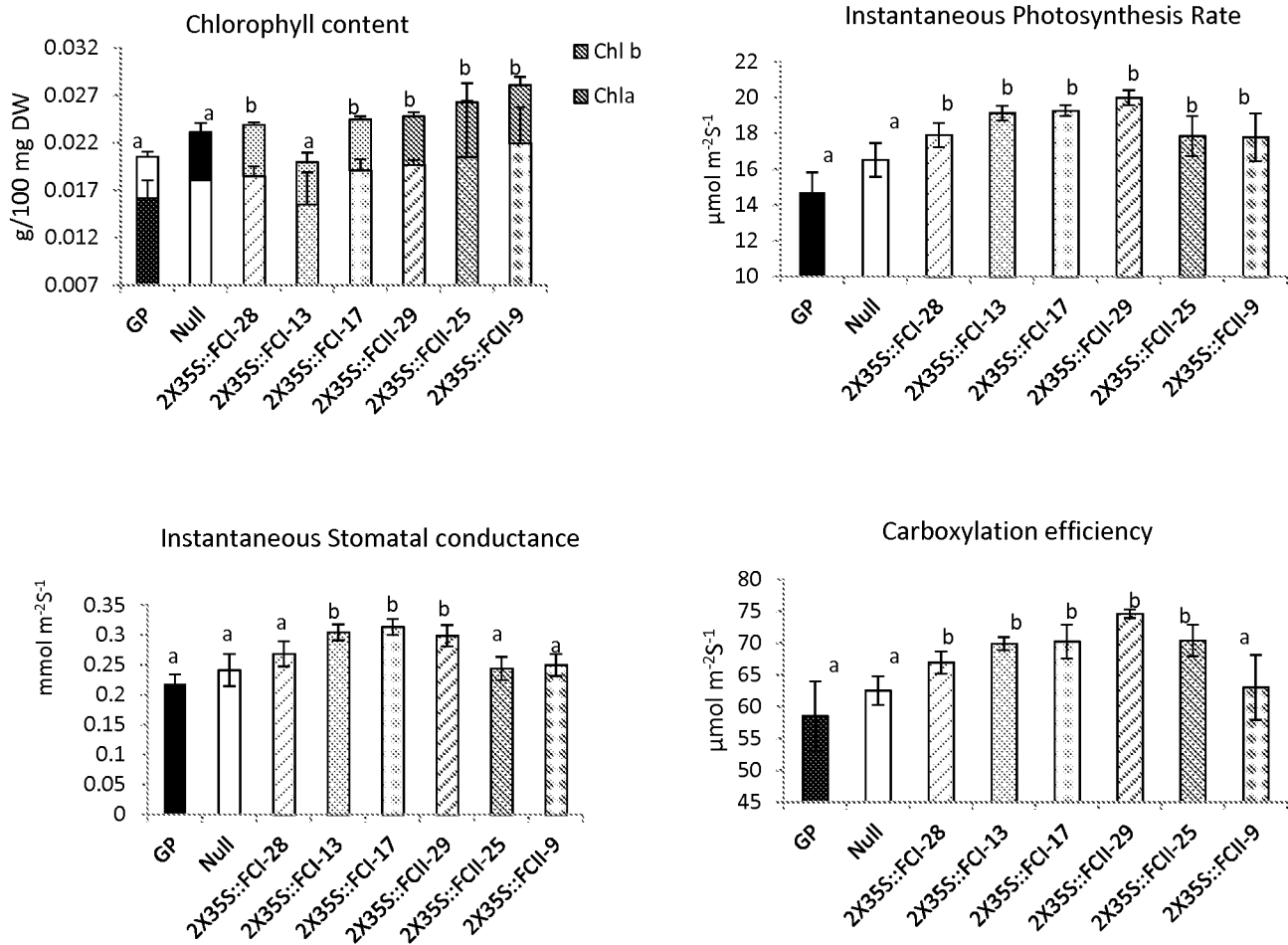


FIG. 5

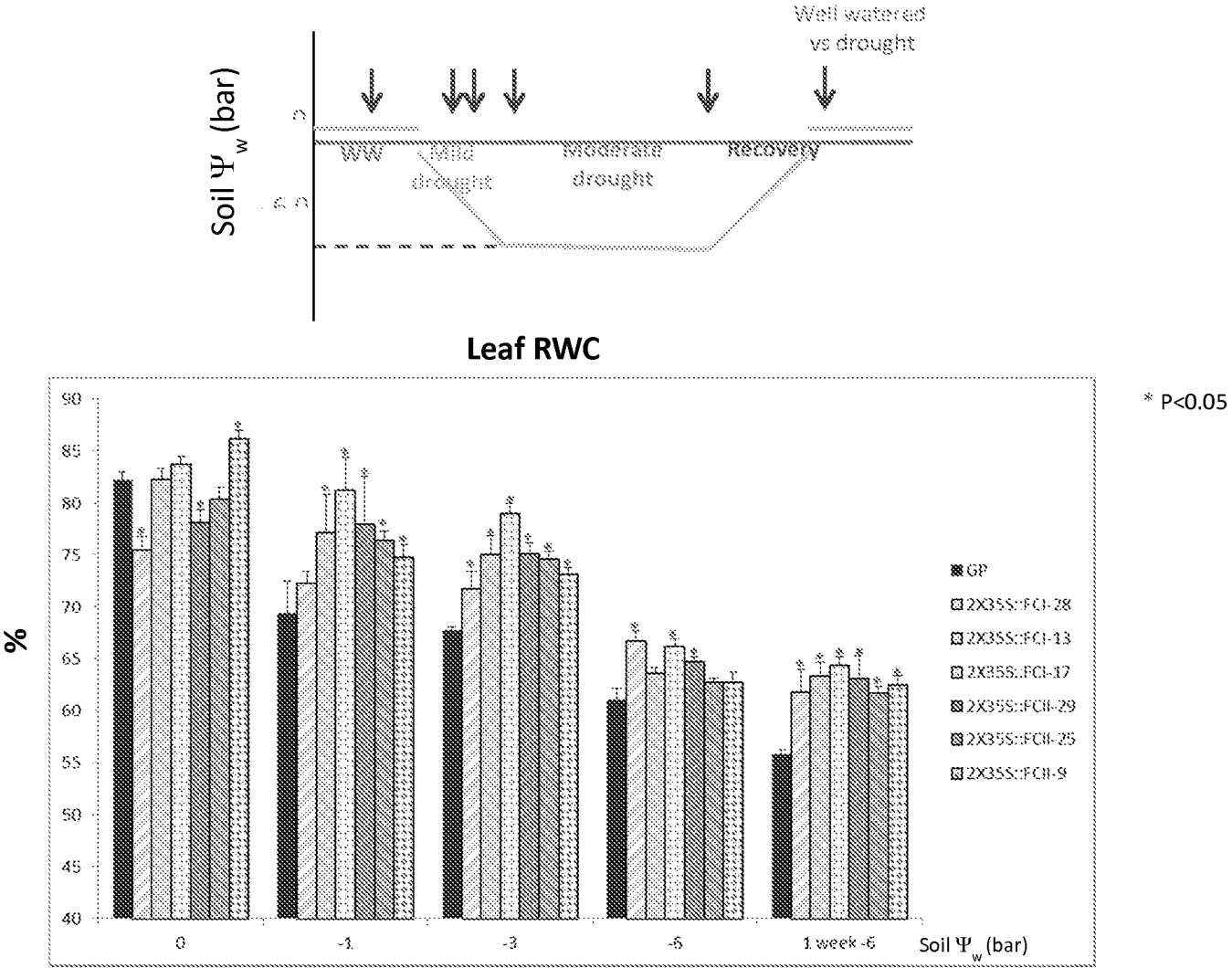
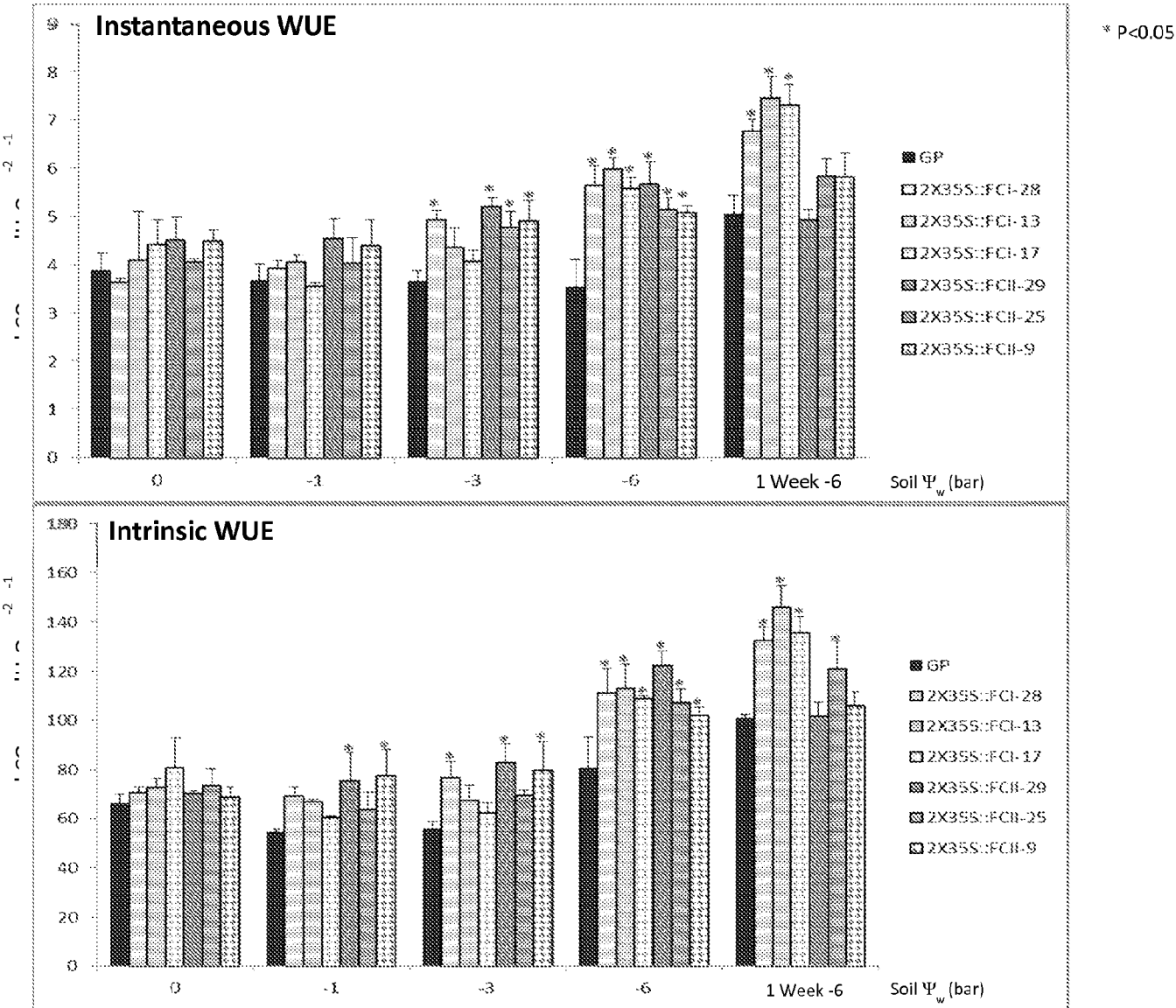
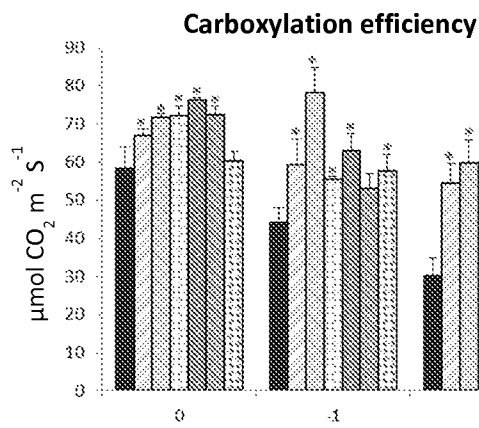
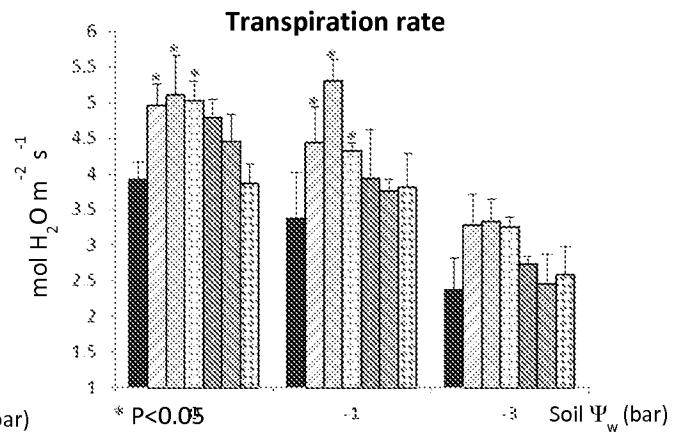
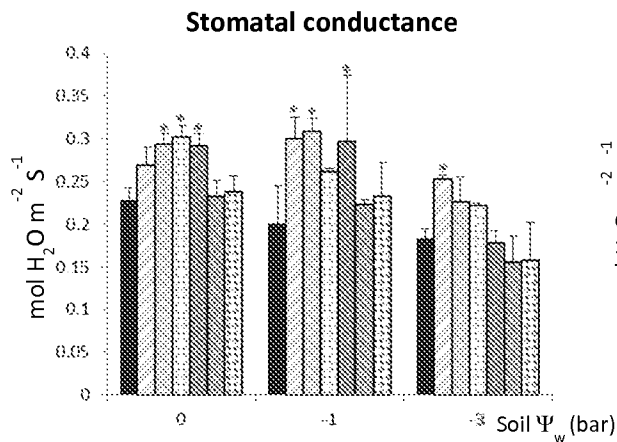
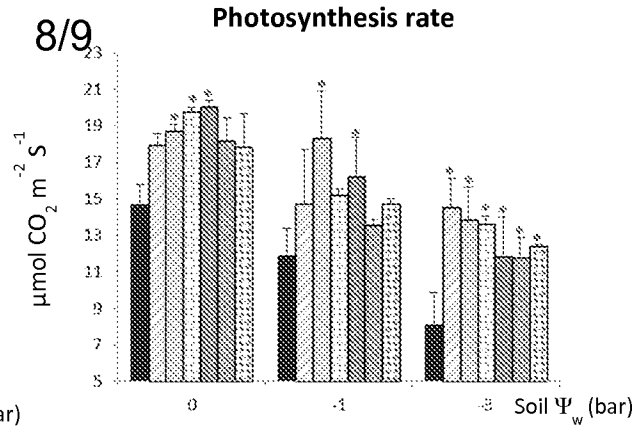
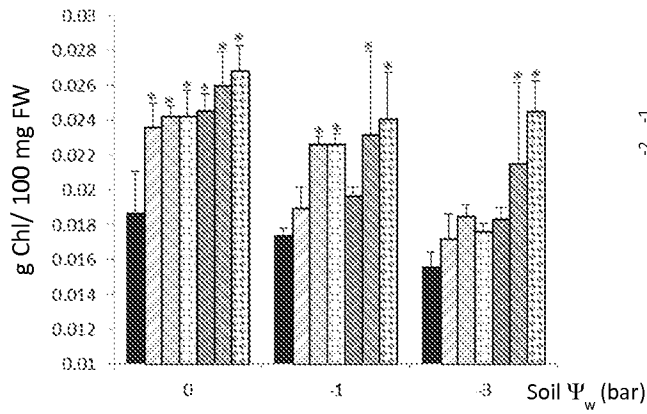


FIG. 6



**FIG. 7**



\* P<0.05

GP  
2X35S::FCI-28  
2X35S::FCI-13  
2X35S::FCI-17  
2X35S::FCII-29  
2X35S::FCII-25  
2X35S::FCII-9

Soil  $\Psi_w$  (bar)

FIG. 8

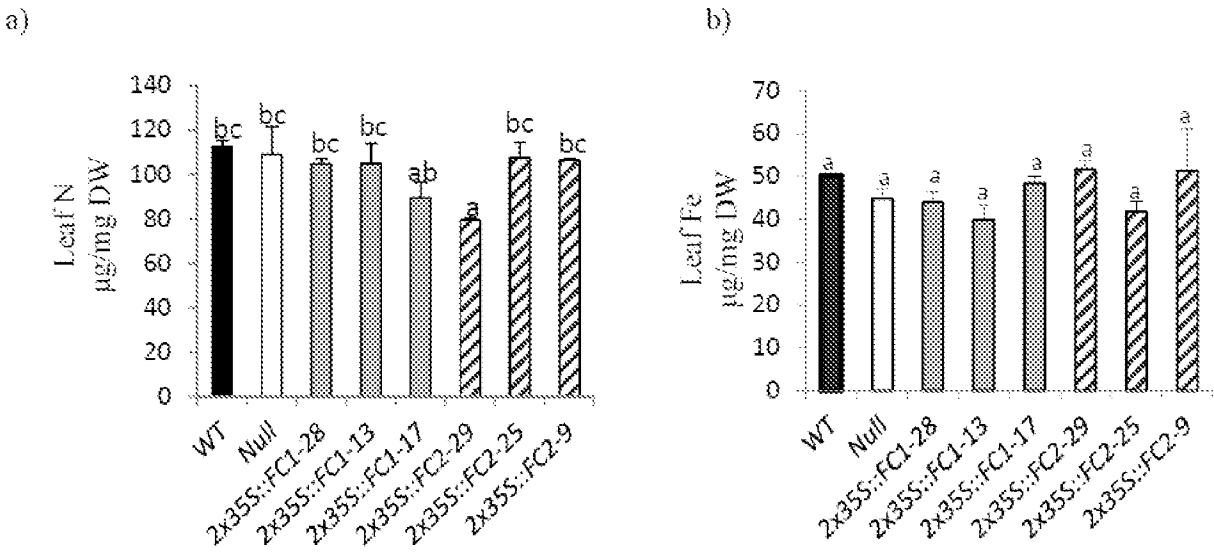


FIG. 9

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/US15/53629

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(8) - C12N 15/29, 15/113; C07K 14/415 (2016.01)

CPC - C12N 15/8261, 15/8216; C07K 14/415

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(8): C12N 15/29, 15/82, 15/87, 15/113, 5/05, 5/10; C07K 14/415; A01H 3/00, 5/00 (2016.01)

CPC: C07K 14/415; C12N 15/8261, 15/8216, 15/82, 15/87, 15/113; A01H 3/00, 5/00

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)

PatSeer (US, EP, WO, JP, DE, GB, CN, FR, KR, ES, AU, IN, CA, INPADOC Data); Google; Google Scholar; PubMed; ScienceDirect; improv\*, increas\*, agronomic, characteristic\*, phenotyp\*, plant\*, soybean\*, maize, rice, polynucleotid\*, 'amino acid,' modulat\*, transgenic, promot\*, recombinant, 'marker assisted selection,' 'allele', 'genomic region,' 'DNA,' polypeptid\*, 'grain yield,' 'grain weight'

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category*     | Citation of document, with indication, where appropriate, of the relevant passages                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | Relevant to claim No.                 |
|---------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------|
| X<br>---<br>Y | WO 2009/073605 A2 (E. I. DU PONT DE NEMOURS AND COMPANY, et al.) June 11, 2009; Table 4; abstract; page 1, line 16; page 2, lines 30, 32-33; page 4, line 7; page 11, line 19; page 14, lines 17-18; page 16, lines 13-14; page 17, lines 31-33; page 24, line 14; page 26, lines 22-25; page 27, line 6; page 28, line 32; page 29, line 1; page 31, lines 20-23; page 34, lines 2, 14-15, 2-26; page 35, line 11; page 37, line 27-33; page 39, lines 6-25; page 48, lines 19-29; page 49, lines 4-7, 21; page 52, line 26; page 63, line 17; page 86, line 25 | 1-6, 8-12, 16-20<br>-----<br>7, 13-15 |
| Y             | WO 2013/044742 A1 (INSTITUTE OF GENETICS AND DEVELOPMENTAL BIOLOGY) April 4, 2013; page 2, lines 15-16; page 2, line 22; page 2, lines 25-29; page 2, lines 30-37; page 3, lines 1-3; page 7, line 26; page 19, lines 28-30                                                                                                                                                                                                                                                                                                                                      | 7, 13-15                              |
| P, Y          | WO 2015/018359 A1 (INSTITUTE OF GENETICS AND DEVELOPMENTAL BIOLOGY, CHINESE ACADEMY OF SCIENCES) February 12, 2015; abstract                                                                                                                                                                                                                                                                                                                                                                                                                                     | 1-20                                  |

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

\* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&amp;" document member of the same patent family

Date of the actual completion of the international search

29 January 2016 (29.01.2016)

Date of mailing of the international search report

05 FEB 2016

Name and mailing address of the ISA/

Mail Stop PCT, Attn: ISA/US, Commissioner for Patents

P.O. Box 1450, Alexandria, Virginia 22313-1450

Facsimile No. 571-273-8300

Authorized officer

Shane Thomas

PCT Helpdesk: 571-272-4300

PCT OSP: 571-272-7774

# INTERNATIONAL SEARCH REPORT

International application No.

PCT/US15/53629

## 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:

-\*\*\* Please See Supplemental Page-\*\*\*-

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 additional fees, this Authority did not invite payment of additional fees.
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.:

-\*\*\*-Please See Supplemental Page-\*\*\*-

### 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  
Information on patent family members

International application No.

PCT/US15/53629

\*\*\*Continued from Box No. III: Observations Where Unity of Invention Is Lacking:

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees must be paid.

Groups I+: Claims 1-20 are directed toward a method of improving rice grain yield, the method comprising modulating the expression of (i) a polynucleotide encoding an amino acid sequence comprising SEQ ID NO: 1 or 2; a method of marker assisted selection of a plant; a method of identifying one or more alleles in a population of plants that are associated with increased grain yield; and a transgenic plant comprising a plant regulatory element that operably regulates the expression of a polynucleotide encoding a polypeptide comprising the amino acid sequence of SEQ ID NO: 1 or 2.

The methods and transgenic plant will be searched to the extent that the polynucleotide encodes an amino acid sequence encompasses SEQ ID NO: 1 (Hordeum vulgare amino acid sequence). Applicant is invited to elect additional polynucleotide sequence(s) and/or encoded polypeptide sequence(s), with specified SEQ ID NO: for each, to be searched. Additional polynucleotide and/or polypeptide sequence(s) will be searched upon the payment of additional fees. It is believed that Claims 1 (in-part), 2 (in-part), 3 (in-part), 4 (in-part), 5 (in-part), 6 (in-part), 7 (in-part), 8 (in-part), 9 (in-part), 10 (in-part), 11 (in-part), 12 (in-part), 13, 14 (in-part), 15 (in-part), 16 (in-part), 17 (in-part), 18 (in-part), 19 (in-part) and 20 (in-part) encompass this first named invention and thus these claims will be searched without fee to the extent that they encompass SEQ ID NO: 1 (Hordeum vulgare amino acid sequence). Failure to clearly identify how any paid additional invention fees are to be applied to the "+" group(s) will result in only the first claimed invention to be searched/examined. An Exemplary Election would be: SEQ ID NO: 2 (Hordeum vulgare amino acid sequence).

Groups I+ share the technical features including: a method of improving an agronomic characteristic of a plant, the method comprising modulating the expression of (i) a polynucleotide encoding an amino acid sequence comprising SEQ ID NO: 1 or 2 or an amino acid sequence that is at least 95% identical to one of SEQ ID NO: 1 or 2 (ii) a polynucleotide that hybridizes under stringent hybridization conditions to a fragment of polynucleotide comprising SEQ ID NO: 3 or 4, wherein the fragment comprises at least 100 contiguous nucleotides of SEQ ID NO: 3 or 4 (iii) a polynucleotide that encodes an amino acid sequence that is at least 90% identical to SEQ ID NO: 1 or 2, (iv) a polynucleotide encoding a polypeptide comprising one or more deletions or insertions or substitutions of amino acids compared to SEQ ID NO: 1 or 2. (a recombinant expression cassette, comprising the polynucleotide wherein the polynucleotide is operably linked to a regulatory element, wherein the expression cassette is functional in a plant cell; a host cell comprising the expression cassette; a transgenic plant comprising the recombinant expression cassette; a method of improving yield of a plant, the method comprising increasing the expression of a polynucleotide that encodes a polypeptide comprising the amino acid sequence selected from the group consisting of SEQ ID NOS: 1, 2 or an allelic variant thereof; a method of improving rice grain yield, the method comprising the expression of a polynucleotide that encodes a polypeptide comprising the amino acid sequence of SEQ ID NO: 1 or a variant thereof; a method of marker assisted selection of a plant for improved yield, the method comprising: a. performing marker-assisted selection of plants that have one or more variations in genomic regions encoding a protein comprising SEQ ID NO: 1 or 2 or a variant thereof or a regulatory sequence thereof; and b. identifying the plant that exhibits higher yield; a method of identifying one or more alleles in a population of plants that are associated with increased grain yield, the method comprising: a. evaluating in a population of plants for one or more allelic variations in (i) a genomic region, the genomic region encoding a polypeptide or (ii) the regulatory region controlling the expression of the polypeptide, wherein the polypeptide comprises the amino acid sequence of SEQ ID NO: 1 or 2 or a sequence that is 95% identical to SEQ ID NO: 1 or 2; b. obtaining phenotypic values of increased yield for the one or more plants in the population; c. associating the allelic variations in the genomic region with the phenotype; and d. identifying the one or more alleles that are associated with increased yield; and a transgenic plant comprising a plant regulatory element that operably regulates the expression of a polynucleotide encoding a polypeptide comprising the amino acid sequence of SEQ ID NO: 1 or 2 or a variant or an ortholog thereof, wherein the regulatory element is heterologous to the polynucleotide and wherein the transgenic plant exhibits substantially normal phenotype as compared to a control plant.

\*\*\*Continued on Next Supplemental Page\*\*\*

-----Continued from Previous Supplemental Page:

However, these shared technical features are previously disclosed by US 2011/0010798 A1 to Allen, et al. (hereinafter 'Allen') in view of US 2011/0258733 A1 (SEBASTIAN).

Allen discloses a method of improving an agronomic characteristic of a plant (paragraphs [0011], [0194], [0195]), the method comprising modulating the expression of (i) a polynucleotide encoding an amino acid sequence (the method comprising altering the level of the expression of a polypeptide encoded by a nucleic acid construct in a host cell; paragraphs [0011], [0229]) comprising an amino acid sequence that is at least 95% identical to one of SEQ ID NO: 1 (comprising SEQ ID NO: 48 of the Allen reference (comprising or an amino acid sequence that is at least 95% identical to one of SEQ ID NO: 1 of the instant PCT application); paragraph [0066]; Table 4, wherein SEQ ID NO: 48 of the Allen reference is 99.8% identical to SEQ ID NO: 1 of the instant PCT application); a recombinant expression cassette (a recombinant DNA construct for expression in a host organism (a recombinant expression cassette); paragraph [0163]), comprising the polynucleotide (paragraphs [0066], [0230]) wherein the polynucleotide is operably linked to a regulatory element (wherein the polynucleotide is operably-linked to a regulatory element; paragraphs [0162], [0163]), wherein the expression cassette is functional in a plant cell (wherein the expression construct (cassette) is functional in a plant cell; paragraph [0002]); a host cell comprising the expression cassette (a host cell comprising the expression construct (cassette); paragraphs [0117], [0163]); a transgenic plant comprising the recombinant expression cassette (a transgenic plant comprising the recombinant expression construct (cassette); paragraph [0010]); a method of improving yield of a plant (paragraphs [0011], [0194], [0195]), the method comprising increasing the expression of a polynucleotide that encodes a polypeptide (paragraphs [0011], [0195], [0229]) comprising the amino acid sequence selected from the group consisting of SEQ ID NO: 1 or an allelic variant thereof (comprising SEQ ID NO: 48 of the Allen reference (comprising the amino acid sequence selected from the group consisting of SEQ ID NO: 1 or an allelic variant thereof); paragraph [0066], Table 4, wherein SEQ ID NO: 48 of the Allen reference is a conservative variant of SEQ ID NO: 1 of the instant PCT application, with conservative substitutions for two amino acids relative to SEQ ID NO: 1 of the instant PCT application, S35N and R164K); a method of improving rice grain yield (a method of improving yield of a plant, including seed grain yield of rice; paragraphs [0011], [0089], [0182], [0194], [0195]), the method comprising the expression of a polynucleotide that encodes a polypeptide (paragraphs [0011], [0195], [0229]) comprising the amino acid sequence of SEQ ID NO: 1 or a variant thereof (comprising SEQ ID NO: 48 of the Allen reference (comprising the amino acid sequence selected from the group consisting of SEQ ID NO: 1 or a variant thereof); paragraph [0066]; Table 4, wherein SEQ ID NO: 48 of the Allen reference is a conservative variant of SEQ ID NO: 1 of the instant PCT application, with conservative substitutions for two amino acids relative to SEQ ID NO: 1 of the instant PCT application, S35N and R164K); and a transgenic plant (paragraph [0010]) comprising a plant regulatory element that operably regulates the expression of a polynucleotide encoding a polypeptide (comprising a construct including a plant regulatory element that regulates the expression of an encoded polypeptide; paragraphs [0010], [0163]) comprising the amino acid sequence of SEQ ID NO: 1 or a variant thereof (comprising SEQ ID NO: 48 of the Allen reference (comprising the amino acid sequence selected from the group consisting of SEQ ID NO: 1 of the instant PCT application or a variant thereof); paragraph [0066]; Table 4, wherein SEQ ID NO: 48 of the Allen reference is a conservative variant of SEQ ID NO: 1 of the instant PCT application, with conservative substitutions for two amino acids relative to SEQ ID NO: 1 of the instant PCT application, S35N and R164K), wherein the regulatory element is heterologous to the polynucleotide (wherein the promoter comprises sequences derived from different promoters found in nature or synthetic DNA segments (wherein the regulatory element is heterologous to the polynucleotide); paragraph [0172]); wherein the source organism for the polypeptide is barley (paragraph [0308]); and where color analysis following application of stress is employed for identifying transgenic plant lines (paragraph [0263]).

Allen does not disclose a method of marker assisted selection of a plant for improved yield, the method comprising: a. performing marker-assisted selection of plants that have one or more variations in genomic regions encoding a protein comprising SEQ ID NO: 1 or 2 or a variant thereof or a regulatory sequence thereof; and b. identifying the plant that exhibits higher yield; a method of identifying one or more alleles in a population of plants that are associated with increased grain yield, the method comprising: a. evaluating in a population of plants for one or more allelic variations in (i) a genomic region, the genomic region encoding a polypeptide or (ii) the regulatory region controlling the expression of the polypeptide, wherein the polypeptide comprises the amino acid sequence of SEQ ID NO: 1 or 2 or a sequence that is 95% identical to SEQ ID NO: 1 or 2; b. obtaining phenotypic values of increased yield for the one or more plants in the population; c. associating the allelic variations in the genomic region with the phenotype; and d. identifying the one or more alleles that are associated with increased yield; and wherein the transgenic plant exhibits substantially normal phenotype as compared to a control plant.

Sebastian discloses a method of marker assisted selection of a plant (paragraph [0001]) for improved yield (for increased (improved) yield; paragraph [0007]), the method comprising: a. performing marker-assisted selection of plants (paragraphs [0075], [0077], [0084]) that have one or more variations in genomic regions (that have one or more markers in chromosomal regions associated with identified alleles associated with superior agronomic performance (that have one or more variations in genomic regions); paragraphs [0075], [0077], [0084]); and b. identifying the plant that exhibits higher yield (paragraphs [0075], [0077]); a method of identifying one or more alleles in a population of plants that are associated with increased grain yield (a method of identifying one or more chromosomal segments encompassing genetic elements associated with alleles associated with a desired phenotype, including increased yield; paragraphs [0007], [0075], [0077], [0084]), the method comprising: a. evaluating in a population of plants for one or more allelic variations in (i) a genomic region (the method comprising: a. evaluating in a population of plants for one or more allelic variations in (i) a chromosomal (genomic) region; paragraphs [0075], [0077], [0084]); b. obtaining phenotypic values of increased yield for the one or more plants in the population (obtaining correlations of allelic frequency with agronomic performance, including increased yield, for plants in a population; paragraphs [0007], [0084]); c. associating the allelic variations in the genomic region with the phenotype (paragraphs [0075], [0077], [0084]); and d. identifying the one or more alleles that are associated with increased yield (paragraphs [0007], [0075], [0077], [0084]).

-----Continued on Next Supplemental Page-----

-\*\*\*-Continued from Previous Supplemental Page:

It would have been obvious to a person of ordinary skill in the art, at the time of the invention, to have readily understood the previous disclosure of Allen to have indicated wherein the transgenic plant exhibits substantially normal phenotype as compared to a control plant, since Allen discloses that the transgenic plants are not distinguishable from the control plants until after application of stress. Furthermore, it would have been obvious to a person of ordinary skill in the art, at the time of the invention, to have modified the disclosure of Allen, for performing marker assisted selection of plants, as disclosed by Sebastian, including source plants for sequences utilized for transformation of transgenic hosts, such as Barley, as disclosed by Allen, for identifying variants having increased yield, as disclosed by Allen, wherein the markers and associated alleles map to a nucleotide region encoding a protein comprising SEQ ID NO: 1 or a variant thereof or a regulatory sequence thereof, for identifying source variants of SEQ ID NO: 1 of the instant PCT application, which further enhance yield, for enabling the transgenic expression of the variants in hosts, as disclosed by Allen, in order to further enhance the seed grain yield of said host plants, such as rice. Moreover, it would have been obvious to a person of ordinary skill in the art, at the time of the invention, to have modified the previous disclosure of Allen, for providing a method of identifying one or more alleles, as disclosed by Sebastian, associated with a nucleotide region encoding a protein comprising SEQ ID NO: 1 or a variant thereof or a regulatory sequence thereof, as disclosed by Allen, within a population of source plants, such as barley, as disclosed by Allen, associated with increased grain yield, as disclosed by Sebastian, for identifying source variants of SEQ ID NO: 1 of the instant PCT application, which further enhance yield, for enabling the transgenic expression of the variants in hosts, as disclosed by Allen, in order to enhance the seed grain yield of said host plants, such as rice.

Since none of the special technical features of the Groups I+ inventions is found in more than one of the inventions, and since all of the shared technical features are previously disclosed by a combination of the Allen and Sebastian references, unity of invention is lacking.