

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
11 June 2009 (11.06.2009)

PCT

(10) International Publication Number
WO 2009/073605 A2

(51) International Patent Classification:

C12N 9/88 (2006.01) *A01H 5/00* (2006.01)
C12N 15/82 (2006.01) *A01H 5/10* (2006.01)

(21) International Application Number:

PCT/US2008/085134

(22) International Filing Date:

1 December 2008 (01.12.2008)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

60/991,859 3 December 2007 (03.12.2007) US

(71) Applicants (for all designated States except US): **E. I. DU PONT DE NEMOURS AND COMPANY** [US/US]; 1007 Market Street, Wilmington, Delaware 19898 (US). **PIO-NEER HI-BRED INTERNATIONAL, INC.** [US/US]; 800 Capital Square 400 Locust Street, Des Moines, Iowa 50309 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **ALLEN, Stephen,**

M. [US/US]; 2225 Rosewood Drive, Wilmington, Delaware 19810 (US). **LUCK, Stanley** [US/US]; 144 Ballymeade Drive, Wilmington, Delaware 19810-1449 (US). **MULLEN, Jeffrey** [US/US]; 6 Heathercroft Court, Chesterfield, Missouri 63017 (US). **SAKAI, Hajime** [DE/US]; 31 Bridle Brook Lane, Newark, Delaware 19711 (US). **TINGEY, Scott, V.** [US/US]; 2401 E. Heather Road, Wilmington, Delaware 19803 (US). **WILLIAMS, Robert, Wayne** [US/US]; 39 Pierson's Ridge, Hockessin, Delaware 19701 (US). **SIVANSANKAR, Shoba** [US/US]; 13916 Oak Brook Drive, Urbandale, Iowa 50323 (US).

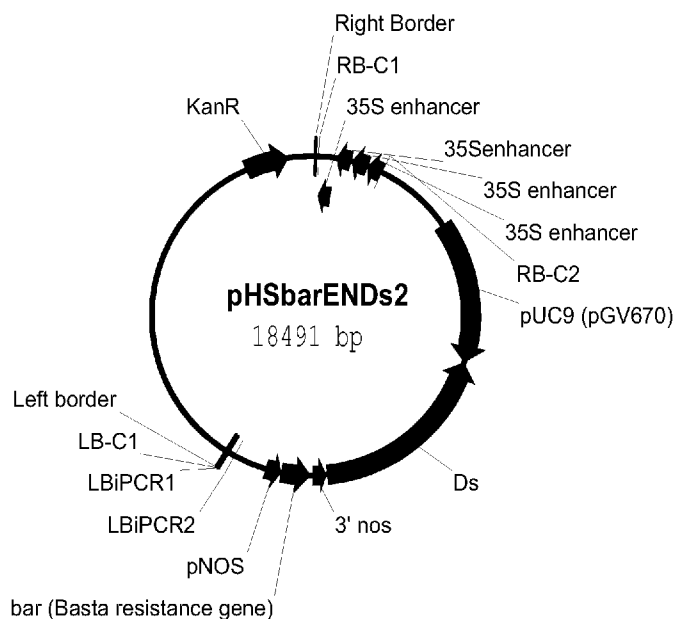
(74) Agent: **KURTZ, Bradley, S.**; E. I. du Pont de Nemours and Company, Legal Patent Records Center, 4417 Lancaster Pike, Wilmington, Delaware 19805 (US).

(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, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK,

[Continued on next page]

(54) Title: DROUGHT TOLERANT PLANTS AND RELATED CONSTRUCTS AND METHODS INVOLVING GENES ENCODING FERROCHELATASES

FIG. 1



(57) Abstract: Isolated polynucleotides and polypeptides and recombinant DNA constructs useful for conferring drought tolerance, compositions (such as plants or seeds) comprising these recombinant DNA constructs, and methods utilizing these recombinant DNA constructs. The recombinant DNA construct comprises a polynucleotide operably linked to a promoter that is functional in a plant, wherein said polynucleotide encodes a ferrochelatase.



LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI,

Published:

- *without international search report and to be republished upon receipt of that report*
- *with sequence listing part of description published separately in electronic form and available upon request from the International Bureau*

TITLE

DROUGHT TOLERANT PLANTS AND RELATED CONSTRUCTS AND METHODS INVOLVING GENES ENCODING FERROCHELATASES

5 This application claims the benefit of U.S. Provisional Application No. 60/991,859, filed December 3, 2007, the entire content of which is herein incorporated by reference.

FIELD OF THE INVENTION

10 The field of invention relates to plant breeding and genetics and, in particular, relates to recombinant DNA constructs useful in plants for conferring tolerance to drought.

BACKGROUND OF THE INVENTION

15 Abiotic stressors significantly limit crop production worldwide. Cumulatively, these factors are estimated to be responsible for an average 70% reduction in agricultural production (Bresson, 1999).

20 Drought stress, in particular, not only causes a reduction in the average yield for crops but also causes yield instability through high interannual yield variation. Globally, about 35-40% of arable land falls under arid or semiarid classification. Even in non-arid regions where soils are nutrient-rich, drought stress occurs regularly for brief periods or at moderate levels. Moreover, it has been predicted that in the coming years rainfall patterns will shift and become more variable due to increased global temperatures.

25 U.S. studies have shown that the ten most important kinds of cultivated plants (corn, soybeans, wheat, tomatoes, etc.) produced only about 50% of the genetically possible yields on average per year; two thirds of the losses were due to the frequent combination of heat stress and water shortage (G. Schütte, S. Stirn, and V. Beusmann, Transgene Pflanzen - Sicherheitsforschung, Risikoabschätzung und Nachzulassungs-Monitoring. Birkhäuser Verlag AG, Basel-Boston-Berlin, 2001).

30 Plants are sessile and have to adjust to the prevailing environmental conditions of their surroundings. This has led to their development of a great plasticity in gene regulation, morphogenesis, and metabolism. Adaptation and defense strategies involve the activation of genes encoding proteins important in the acclimation or defense towards the different stressors. Some

of the molecular responses to abiotic stress factors such as drought are specific, but it has also been shown that similar genes are activated by several stressors (Royal Society of London, *Transgenic Plants and World Agriculture*, 2000, National Academy Press, Washington, DC). It is believed that about 15 percent of a plant's genome is devoted to stress perception and adaptation (see e.g., Cushman and Bohnert, 2000).

Earlier work on molecular aspects of abiotic stress responses was accomplished by differential and/or subtractive analysis (e.g., see Bray, 1993, Shinozaki and Yamaguchi-Shinozaki, 1997, Zhu et al., 1997, Thomashow, 1999).

Other methods include selection of candidate genes (e.g., selection of genes from a particular known module and analyzing expression of such a gene or its active product under stresses, or by functional complementation in a stressor system that is well defined, see Xiong and Zhu, 2001). Additionally, forward and reverse genetic studies involving the identification and isolation of mutations in regulatory genes have also been used to provide evidence for observed changes in gene expression under stress or exposure (Xiong and Zhu, 2001).

Activation tagging can be utilized to identify genes with the ability to affect a trait. This approach has been used in the model plant species *Arabidopsis thaliana* (Weigel et al., *Plant Physiol.* 122:1003-1013 (2000)). Insertions of transcriptional enhancer elements can dominantly activate and/or elevate the expression of nearby endogenous genes. This method can be used to select genes involved in agronomically important phenotypes, including stress tolerance.

SUMMARY OF THE INVENTION

The present invention includes:

In one embodiment, a plant comprising in its genome a recombinant DNA construct comprising a polynucleotide operably linked to at least one regulatory element, wherein said polynucleotide encodes a polypeptide having an amino acid sequence of at least 50% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28 or 30, and wherein said plant exhibits increased drought tolerance when compared to a control plant not comprising said recombinant DNA construct.

In another embodiment, a method of increasing drought tolerance in a plant, comprising (a) introducing into a regenerable plant cell a recombinant DNA

construct comprising a polynucleotide operably linked to at least one regulatory sequence, wherein the polynucleotide encodes a polypeptide having an amino acid sequence of at least 50% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28 or 30; (b) regenerating a transgenic plant from the regenerable plant cell after step (a), wherein the transgenic plant comprises in its genome the recombinant DNA construct and exhibits increased drought tolerance when compared to a control plant not comprising the recombinant DNA construct; and optionally, (c) obtaining a progeny plant derived from the transgenic plant, wherein said progeny plant comprises in its genome the recombinant DNA construct and exhibits increased drought tolerance when compared to a control plant not comprising the recombinant DNA construct.

In another embodiment, a method of evaluating drought tolerance in a plant, comprising (a) introducing into a regenerable plant cell a recombinant DNA construct comprising a polynucleotide operably linked to at least one regulatory sequence, wherein the polynucleotide encodes a polypeptide having an amino acid sequence of at least 50% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28 or 30; (b) regenerating a transgenic plant from the regenerable plant cell after step (a), wherein the transgenic plant comprises in its genome the recombinant DNA construct; and (c) evaluating the transgenic plant for drought tolerance compared to a control plant not comprising the recombinant DNA construct; and optionally, (d) obtaining a progeny plant derived from the transgenic plant, wherein the progeny plant comprises in its genome the recombinant DNA construct; and optionally, (e) evaluating the progeny plant for drought tolerance compared to a control plant not comprising the recombinant DNA construct.

In another embodiment, a method of evaluating drought tolerance in a plant, comprising (a) introducing into a regenerable plant cell a recombinant DNA construct comprising a polynucleotide operably linked to at least one regulatory sequence, wherein the polynucleotide encodes a polypeptide having an amino acid sequence of at least 50% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28 or 30; (b) regenerating a transgenic plant from the regenerable plant cell after

step (a), wherein the transgenic plant comprises in its genome the recombinant DNA construct; (c) obtaining a progeny plant derived from the transgenic plant, wherein the progeny plant comprises in its genome the recombinant DNA construct; and (d) evaluating the progeny plant for drought tolerance compared to a control plant not comprising the recombinant DNA construct.

In another embodiment, the present invention includes any of the methods of the present invention wherein the plant is a maize plant or a soybean plant.

In another embodiment, the present invention includes an isolated polynucleotide comprising: (a) a nucleotide sequence encoding a ferrochelatase, wherein the polypeptide has an amino acid sequence of at least 90% or 95% sequence identity, based on the Clustal V method of alignment, when compared to one of SEQ ID NO:4, 6, 8, 10, 12, 14, 16, 18, 24, 26, 28 or 30, or (b) a full complement of the nucleotide sequence, wherein the full complement and the nucleotide sequence consist of the same number of nucleotides and are 100% complementary. The polypeptide may comprise the amino acid sequence of SEQ ID NO:4, 6, 8, 10, 12, 14, 16, 18, 24, 26, 28 or 30. The nucleotide sequence may comprise the nucleotide sequence of SEQ ID NO:3, 5, 7, 9, 11, 13, 15, 17, 23, 25, 27 or 29.

In another embodiment, the present invention concerns a recombinant DNA construct comprising any of the isolated polynucleotides of the present invention operably linked to at least one regulatory sequence, and a cell, a plant, and a seed comprising the recombinant DNA construct.

In another embodiment, the present invention includes a vector comprising any of the isolated polynucleotides of the present invention.

In another embodiment, the present invention concerns a cell, plant or seed comprising any of the recombinant DNA constructs of the present invention. The cell may be eukaryotic, e.g., a yeast, insect or plant cell, or prokaryotic, e.g., a bacterium.

BRIEF DESCRIPTION OF THE FIGURES AND SEQUENCE LISTING

The invention can be more fully understood from the following detailed description and the accompanying drawings and Sequence Listing which form a part of this application.

Figure 1 shows a schematic of the pHSbarENDs2 activation tagging construct (SEQ ID NO:31) used to make the *Arabidopsis* populations.

Figure 2 shows a map of the vector pDONRTM/Zeo (SEQ ID NO:32). The attP1 site is at nucleotides 570-801; the attP2 site is at nucleotides 2754-2985 (complementary strand).

Figure 3 shows a map of the vector pDONRTM221 (SEQ ID NO:33). The attP1 site is at nucleotides 570-801; the attP2 site is at nucleotides 2754-2985 (complementary strand).

Figure 4 shows a map of the vector pBC-yellow (SEQ ID NO:34), a destination vector for use in construction of expression vectors for *Arabidopsis*. The attR1 site is at nucleotides 11276-11399 (complementary strand); the attR2 site is at nucleotides 9695-9819 (complementary strand).

Figure 5 shows a map of PHP27840 (SEQ ID NO:35), a destination vector for use in construction of expression vectors for soybean. The attR1 site is at nucleotides 7310-7434; the attR2 site is at nucleotides 8890-9014.

Figure 6 shows a map of PHP23236 (SEQ ID NO:36), a destination vector for use in construction of expression vectors for Gaspé Flint derived maize lines. The attR1 site is at nucleotides 2006-2130; the attR2 site is at nucleotides 2899-3023.

Figure 7 shows a map of PHP10523 (SEQ ID NO:37), a plasmid DNA present in *Agrobacterium* strain LBA4404 (Komari et al., *Plant J.* 10:165-174 (1996); NCBI General Identifier No. 59797027).

Figure 8 shows a map of PHP23235 (SEQ ID NO:38), a vector used to construct the destination vector PHP23236.

Figure 9 shows a map of PHP28647 (SEQ ID NO:39), a destination vector for use with maize inbred-derived lines. The attR1 site is at nucleotides 2289-2413; the attR2 site is at nucleotides 3869-3993.

Figures 10A – 10C show the multiple alignment of the amino acid sequences of the ferrochelatase-I proteins of SEQ ID NOs:2, 4, 6, 8, 10, 12, 14, 16, 18, 48, 49 and 50. Residues that are identical to the residue of SEQ ID NO:2 at a given position are enclosed in a box. A consensus sequence is presented where a residue is shown if identical in all sequences, otherwise, a period is shown.

Figure 11 shows the percent sequence identity and the divergence values for each pair of amino acids sequences of ferrochelatases displayed in Figures 10A – 10C.

Figures 12A – 12C show the multiple alignment of the amino acid sequences of the ferrochelatase-II proteins of SEQ ID NOs:20, 22, 24, 26, 28, 30, 51, 52, 53 and 54. Residues that are identical to the residue of SEQ ID NO:20 at a given position are enclosed in a box. A consensus sequence is presented where a residue is shown if identical in all sequences, otherwise, a period is shown.

Figure 13 shows the percent sequence identity and the divergence values for each pair of amino acids sequences of ferrochelatases displayed in Figures 12A – 12C.

Figures 14A – 14B show an evaluation of individual Gaspe Flint derived maize lines transformed with PHP31419.

Figure 15 shows a summary evaluation of Gaspe Flint derived maize lines transformed with PHP31419.

Figures 16A – 16B show an evaluation of individual Gaspe Flint derived maize lines transformed with PHP33089.

Figure 17 shows a summary evaluation of Gaspe Flint derived maize lines transformed with PHP33089.

Figures 18A – 18B show an evaluation of individual Gaspe Flint derived maize lines transformed with PHP30829.

Figure 19 shows a summary evaluation of Gaspe Flint derived maize lines transformed with PHP30829.

Figures 20A – 20B show an evaluation of individual Gaspe Flint derived maize lines transformed with PHP30745.

Figure 21 shows a summary evaluation of Gaspe Flint derived maize lines transformed with PHP30745.

Figures 22A – 22B show an evaluation of individual Gaspe Flint derived maize lines transformed with PHP30761.

Figure 23 shows a summary evaluation of Gaspe Flint derived maize lines transformed with PHP30761.

SEQ ID NO:1 corresponds to NCBI GI No. 511080 which is the nucleotide sequence of a cDNA fragment encoding an *Arabidopsis* ferrochelatase-I protein (locus At5g26030).

5 SEQ ID NO:2 corresponds to NCBI GI No. 511081, which is the amino acid sequence of the *Arabidopsis* ferrochelatase-I protein encoded by SEQ ID NO:1.

TABLE 1

cDNAs Encoding Ferrochelatase-I (FeC-I)

Plant	Clone Designation	SEQ ID NO: (Nucleotide)	SEQ ID NO: (Amino Acid)
Sugar Beet	ebs1c.pk002.n16 (FIS)	3	4
Brassica	ebb1c.pk006.j11 (FIS)	5	6
Maize	cfp3n.pk004.f12 (FIS)	7	8
Maize	cfp5n.pk009.j16 (FIS)	9	10
Rice	rl0n.pk117.h21 (FIS)	11	12
Soybean	se3.pk0034.e10 (FIS)	13	14
Tulip	etb1n.pk002.n16 (FIS)	15	16
Wheat	wlp1c.pk002.p10 (FIS)	17	18

10 SEQ ID NO:19 corresponds to NCBI GI No. 30684569 which is the nucleotide sequence of a cDNA fragment encoding a first allele of an *Arabidopsis* ferrochelatase-II protein (locus At2g30390).

SEQ ID NO:20 corresponds to NCBI GI No. 15227742, which is the amino acid sequence of the *Arabidopsis* ferrochelatase-II protein encoded by SEQ ID NO:19.

15 SEQ ID NO:21 corresponds to NCBI GI No. 2623989, which is the nucleotide sequence of a cDNA fragment encoding a second allele of an *Arabidopsis* ferrochelatase-II protein (locus At2g30390).

20 SEQ ID NO:22 corresponds to NCBI GI No. 2623990, which is the amino acid sequence of the *Arabidopsis* ferrochelatase-II protein encoded by SEQ ID NO:21.

TABLE 2
cDNAs Encoding Ferrochelatase-II (FeC-II)

Plant	Clone Designation	SEQ ID NO: (Nucleotide)	SEQ ID NO: (Amino Acid)
Catmint	ecl1c.pk005.l15 (FIS)	23	24
Maize	cfp6n.pk072.n9 (FIS)	25	26
Wheat	Wlp1c.pk004.d13 (FIS)	27	28
Wheat	wpa1c.pk014.g4 (FIS)	29	30

SEQ ID NO:31 is the nucleotide sequence of the pHSbarENDs2 activation
5 tagging vector.

SEQ ID NO:32 is the nucleotide sequence of the GATEWAY® donor vector
pDONR™/Zeo.

SEQ ID NO:33 is the nucleotide sequence of the GATEWAY® donor vector
pDONR™221.

10 SEQ ID NO:34 is the nucleotide sequence of pBC-yellow, a destination
vector for use with *Arabidopsis*.

SEQ ID NO:35 is the nucleotide sequence of PHP27840, a destination vector
for use with soybean.

15 SEQ ID NO:36 is the nucleotide sequence of PHP23236, a destination vector
for use with Gaspé Flint derived maize lines.

SEQ ID NO:37 is the nucleotide sequence of PHP10523 (Komari et al., *Plant*
J. 10:165-174 (1996); NCBI General Identifier No. 59797027).

SEQ ID NO:38 is the nucleotide sequence of PHP23235, a destination vector
for use with Gaspé Flint derived lines.

20 SEQ ID NO:39 is the nucleotide sequence of PHP28647, a destination vector
for use with maize inbred-derived lines.

SEQ ID NO:40 is the nucleotide sequence of the attB1 site.

SEQ ID NO:41 is the nucleotide sequence of the attB2 site.

25 SEQ ID NO:42 is the nucleotide sequence of the At5g26030-5'attB forward
primer, containing the attB1 sequence, used to amplify the At5g26030 protein-
coding region.

SEQ ID NO:43 is the nucleotide sequence of the At5g26030-3'attB reverse primer, containing the attB2 sequence, used to amplify the At5g26030 protein-coding region.

5 SEQ ID NO:44 is the nucleotide sequence of the At2g30390-5'attB forward primer, containing the attB1 sequence, used to amplify the At2g30390 protein-coding region.

SEQ ID NO:45 is the nucleotide sequence of the At2g30390-3'attB reverse primer, containing the attB2 sequence, used to amplify the At2g30390 protein-coding region.

10 SEQ ID NO:46 is the nucleotide sequence of the VC062 primer, containing the T3 promoter and attB1 site, useful to amplify cDNA inserts cloned into a Bluescript® II SK(+) vector (Stratagene).

SEQ ID NO:47 is the nucleotide sequence of the VC063 primer, containing the T7 promoter and attB2 site, useful to amplify cDNA inserts cloned into a
15 Bluescript® II SK(+) vector (Stratagene).

SEQ ID NO:48 is the amino acid sequence of a *Hordeum vulgare* ferrochelatase-I protein (NCBI GI NO. 2460251).

SEQ ID NO:49 is the amino acid sequence of a rice ferrochelatase-I protein (NCBI GI NO. 113631036).

20 SEQ ID NO:50 is the amino acid sequence of a *Vitis vinifera* ferrochelatase-I protein (NCBI GI NO. 147818793).

SEQ ID NO:51 is the amino acid sequence of a rice ferrochelatase-II protein (NCBI GI NO. 115463419).

25 SEQ ID NO:52 is the amino acid sequence of a *Cucumis sativus* ferrochelatase-II protein (NCBI GI NO. 12082085).

SEQ ID NO:53 is the amino acid sequence of a *Nicotiana tabacum* ferrochelatase-II protein (NCBI GI NO. 15147828).

SEQ ID NO:54 is the amino acid sequence of a *Synechocystis* ferrochelatase protein (NCBI GI NO. 1708186).

30 SEQ ID NO:55 is the nucleotide acid sequence of plasmid PHP30949, an entry clone containing the maize ferrochelatase-I protein (SEQ ID NO:10).

SEQ ID NO:56 is the nucleotide sequence of the cDNA insert of cfp5n.pk064.n7 and encodes a maize ferrochelatase-II protein (SEQ ID NO:8).

SEQ ID NO:56 differs from the nucleotide sequence of cfp3n.pk004.f12 (SEQ ID NO:7) in that it contains a 12 base pair insertion in the 5'-UTR, 8 nucleotides before the ATG start codon.

SEQ ID NO:57 is the amino acid sequence presented in SEQ ID NO:240025 of US Patent Publication No. US2004031072.

SEQ ID NO:58 is the amino acid sequence presented in SEQ ID NO:20 of Japanese Patent Publication No. JP2001190168-A.

SEQ ID NO:59 is the amino acid sequence presented in SEQ ID NO:13029 of US Patent Publication No. US2006150283.

SEQ ID NO:60 is the amino acid sequence presented in SEQ ID NO:7745 of US Patent Publication No. US2006150283.

SEQ ID NO:61 is the amino acid sequence presented in SEQ ID NO:46156 of Japanese Patent Publication No. JP2005185101.

SEQ ID NO:62 is the amino acid sequence presented in SEQ ID NO:52154 of Japanese Patent Publication No. JP2005185101.

SEQ ID NO:63 is the amino acid sequence presented in SEQ ID NO:72746 of US Patent Publication No. US2004034888-A1.

The sequence descriptions and Sequence Listing attached hereto comply with the rules governing nucleotide and/or amino acid sequence disclosures in patent applications as set forth in 37 C.F.R. §1.821-1.825.

The Sequence Listing contains the one letter code for nucleotide sequence characters and the three letter codes for amino acids as defined in conformity with the IUPAC-IUBMB standards described in *Nucleic Acids Res.* 13:3021-3030 (1985) and in the *Biochemical J.* 219 (No. 2):345-373 (1984) which are herein incorporated by reference. The symbols and format used for nucleotide and amino acid sequence data comply with the rules set forth in 37 C.F.R. §1.822.

DETAILED DESCRIPTION OF EMBODIMENTS

The disclosure of each reference set forth herein is hereby incorporated by reference in its entirety.

As used herein and in the appended claims, the singular forms "a", "an", and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, reference to "a plant" includes a plurality of such plants, reference to "a

cell" includes one or more cells and equivalents thereof known to those skilled in the art, and so forth.

As used herein:

5 "Ferrochelatase" (protoheme ferrolyase; EC 4.99.1.1) is the terminal enzyme of the biosynthetic pathway of heme; it catalyses the chelation of ferrous ion into the protoporphyrin IX ring to form protoheme. In higher plants there is evidence for two isoforms of this enzyme, ferrochelatase-1 and ferrochelatase-2 (Suzuki et al. 2002 J Biol Chem 277:4731-4737). The terms "ferrochelatase-I", "ferrochelatase-1", "FeC-I" and "FeC-1" are used interchangeably herein. The terms "ferrochelatase-II",
10 "ferrochelatase-2", "FeC-II" and "FeC-2" are used interchangeably herein.

An "Expressed Sequence Tag" ("EST") is a DNA sequence derived from a cDNA library and therefore is a sequence which has been transcribed. An EST is typically obtained by a single sequencing pass of a cDNA insert. The sequence of an entire cDNA insert is termed the "Full-Insert Sequence" ("FIS"). A "Contig"
15 sequence is a sequence assembled from two or more sequences that can be selected from, but not limited to, the group consisting of an EST, FIS and PCR sequence. A sequence encoding an entire or functional protein is termed a "Complete Gene Sequence" ("CGS") and can be derived from an FIS or a contig.

"Agronomic characteristic" is a measurable parameter including but not
20 limited to, greenness, yield, growth rate, biomass, fresh weight at maturation, dry weight at maturation, fruit yield, seed yield, total plant nitrogen content, fruit nitrogen content, seed nitrogen content, nitrogen content in a vegetative tissue, total plant free amino acid content, fruit free amino acid content, seed free amino acid content, free amino acid content in a vegetative tissue, total plant protein content, fruit
25 protein content, seed protein content, protein content in a vegetative tissue, drought tolerance, nitrogen uptake, root lodging, harvest index, stalk lodging, plant height, ear height and ear length.

"Transgenic" refers to any cell, cell line, callus, tissue, plant part or plant, the genome of which has been altered by the presence of a heterologous nucleic acid,
30 such as a recombinant DNA construct, including those initial transgenic events as well as those created by sexual crosses or asexual propagation from the initial transgenic event. 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.

“Genome” as it applies to plant cells encompasses not only chromosomal DNA found within the nucleus, but organelle DNA found within subcellular components (e.g., mitochondrial, plastid) of the cell.

“Plant” includes reference to whole plants, plant organs, plant tissues, seeds and plant cells and progeny of same. Plant cells include, without limitation, cells from seeds, suspension cultures, embryos, meristematic regions, callus tissue, leaves, roots, shoots, gametophytes, sporophytes, pollen, and microspores.

“Progeny” comprises any subsequent generation of a plant.

“Transgenic plant” includes reference to a plant which comprises within its genome a heterologous polynucleotide. For example, 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 DNA construct.

“Heterologous” with respect to sequence means a sequence 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 intervention.

“Polynucleotide”, “nucleic acid sequence”, “nucleotide sequence”, or “nucleic acid fragment” are used interchangeably and is a polymer of RNA or DNA that is single- or double-stranded, optionally containing synthetic, non-natural or altered nucleotide bases. Nucleotides (usually found in their 5'-monophosphate form) are referred to by their single letter designation as follows: “A” for adenylate or deoxyadenylate (for RNA or DNA, respectively), “C” for cytidylate or deoxycytidylate, “G” for guanylate or deoxyguanylate, “U” for uridylate, “T” for deoxythymidylate, “R” for purines (A or G), “Y” for pyrimidines (C or T), “K” for G or T, “H” for A or C or T, “I” for inosine, and “N” for any nucleotide.

“Polypeptide”, “peptide”, “amino acid sequence” and “protein” are used interchangeably herein 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. The terms "polypeptide", "peptide", "amino acid sequence", and "protein" are also inclusive of modifications including, but not limited to, glycosylation, lipid attachment, sulfation, gamma-carboxylation of glutamic acid residues, hydroxylation and ADP-ribosylation.

5 "Messenger RNA (mRNA)" refers to the RNA that is without introns and that can be translated into protein by the cell.

"cDNA" refers to a DNA that is complementary to and synthesized from a mRNA template using the enzyme reverse transcriptase. The cDNA can be single-stranded or converted into the double-stranded form using the Klenow fragment of
10 DNA polymerase I.

"Mature" protein refers to a post-translationally processed polypeptide; i.e., one from which any pre- or pro-peptides present in the primary translation product have been removed.

"Precursor" protein refers to the primary product of translation of mRNA; i.e.,
15 with pre- and pro-peptides still present. Pre- and pro-peptides may be and are not limited to intracellular localization signals.

"Isolated" refers to materials, such as nucleic acid molecules and/or proteins, which are substantially free or otherwise removed from components that normally accompany or interact with the materials in a naturally occurring environment.
20 Isolated polynucleotides may be purified from a host cell in which they naturally occur. Conventional nucleic acid purification methods known to skilled artisans may be used to obtain isolated polynucleotides. The term also embraces recombinant polynucleotides and chemically synthesized polynucleotides.

"Recombinant" refers to an artificial combination of two otherwise separated
25 segments of sequence, e.g., by chemical synthesis or by the manipulation of isolated segments of nucleic acids by genetic engineering techniques.

"Recombinant" also includes reference to a cell or vector, that has been modified by the introduction of a heterologous nucleic acid or a cell derived from a cell so modified, but does not encompass the alteration of the cell or vector by naturally
30 occurring events (e.g., spontaneous mutation, natural transformation/transduction/transposition) such as those occurring without deliberate human intervention.

“Recombinant DNA construct” refers to a combination of nucleic acid fragments that are not normally found together in nature. Accordingly, a recombinant DNA construct may comprise regulatory sequences and coding sequences that are derived from different sources, or regulatory sequences and coding sequences derived from the same source, but arranged in a manner different than that normally found in nature.

The terms “entry clone” and “entry vector” are used interchangeably herein.

“Regulatory sequences” refer to nucleotide sequences located upstream (5' non-coding sequences), within, or downstream (3' non-coding sequences) of a coding sequence, and which influence the transcription, RNA processing or stability, or translation of the associated coding sequence. Regulatory sequences may include, but are not limited to, promoters, translation leader sequences, introns, and polyadenylation recognition sequences. The terms “regulatory sequence” and “regulatory element” are used interchangeably herein.

“Promoter” refers to a nucleic acid fragment capable of controlling transcription of another nucleic acid fragment.

“Promoter functional in a plant” is a promoter capable of controlling transcription in plant cells whether or not its origin is from a plant cell.

“Tissue-specific promoter” and “tissue-preferred promoter” are used interchangeably, and refer to a promoter that is expressed predominantly but not necessarily exclusively in one tissue or organ, but that may also be expressed in one specific cell.

“Developmentally regulated promoter” refers to a promoter whose activity is determined by developmental events.

“Operably linked” refers to the association of nucleic acid fragments in a single fragment so that the function of one is regulated by the other. For example, a promoter is operably linked with a nucleic acid fragment when it is capable of regulating the transcription of that nucleic acid fragment.

“Expression” refers to the production of a functional product. For example, expression of a nucleic acid fragment may refer to transcription of the nucleic acid fragment (e.g., transcription resulting in mRNA or functional RNA) and/or translation of mRNA into a precursor or mature protein.

“Phenotype” means the detectable characteristics of a cell or organism.

"Introduced" in the context of inserting a nucleic acid fragment (e.g., a recombinant DNA construct) into a cell, means "transfection" or "transformation" or "transduction" and includes reference to the incorporation of a nucleic acid fragment into a eukaryotic or prokaryotic cell where the nucleic acid fragment 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).

A "transformed cell" is any cell into which a nucleic acid fragment (e.g., a recombinant DNA construct) has been introduced.

"Transformation" as used herein refers to both stable transformation and transient transformation.

"Stable transformation" refers to the introduction of a nucleic acid fragment into a genome of a host organism resulting in genetically stable inheritance. Once stably transformed, the nucleic acid fragment is stably integrated in the genome of the host organism and any subsequent generation.

"Transient transformation" refers to the introduction of a nucleic acid fragment into the nucleus, or DNA-containing organelle, of a host organism resulting in gene expression without genetically stable inheritance.

"Allele" is one of several alternative forms of a gene occupying a given locus on a chromosome. When the alleles present at a given locus on a pair of homologous chromosomes in a diploid plant are the same that plant is homozygous at that locus. If the alleles present at a given locus on a pair of homologous chromosomes in a diploid plant differ that plant is heterozygous at that locus. If a transgene is present on one of a pair of homologous chromosomes in a diploid plant that plant is hemizygous at that locus.

Sequence alignments and percent identity calculations may be determined using a variety of comparison methods designed to detect homologous sequences including, but not limited to, the Megalign® program of the LASERGENE® bioinformatics computing suite (DNASTAR® Inc., Madison, WI). Unless stated otherwise, multiple alignment of the sequences provided herein were performed using the Clustal V method of alignment (Higgins and Sharp (1989) *CABIOS* 5:151-153) with the default parameters (GAP PENALTY=10, GAP LENGTH PENALTY=10). Default parameters for pairwise alignments and calculation of

percent identity of protein sequences using the Clustal V method are KTUPLE=1, GAP PENALTY=3, WINDOW=5 and DIAGONALS SAVED=5. For nucleic acids these parameters are KTUPLE=2, GAP PENALTY=5, WINDOW=4 and DIAGONALS SAVED=4. After alignment of the sequences, using the Clustal V program, it is possible to obtain "percent identity" and "divergence" values by viewing the "sequence distances" table on the same program; unless stated otherwise, percent identities and divergences provided and claimed herein were calculated in this manner.

Standard recombinant DNA and molecular cloning techniques used herein are well known in the art and are described more fully in Sambrook, J., Fritsch, E.F. and Maniatis, T. *Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, 1989 (hereinafter "Sambrook").

Turning now to the embodiments:

Embodiments include isolated polynucleotides and polypeptides, recombinant DNA constructs useful for conferring drought tolerance, compositions (such as plants or seeds) comprising these recombinant DNA constructs, and methods utilizing these recombinant DNA constructs.

Isolated Polynucleotides and Polypeptides:

The present invention includes the following isolated polynucleotides and polypeptides:

An isolated polynucleotide comprising: (i) a nucleic acid sequence encoding a polypeptide having an amino acid sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO:4, 6, 8, 10, 12, 14, 16, 18, 24, 26, 28 or 30; or (ii) a full complement of the nucleic acid sequence of (i), wherein the full complement and the nucleic acid sequence of (i) consist of the same number of nucleotides and are 100% complementary. Any of the foregoing isolated polynucleotides may be utilized in any recombinant DNA constructs (including suppression DNA constructs) of the present invention. The polypeptide is preferably a ferrochelataase.

An isolated polypeptide having an amino acid sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 56%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO:4, 6, 8, 10, 12, 14, 16, 18, 24, 26, 28 or 30. The polypeptide is preferably a ferrochelatase.

An isolated polynucleotide comprising (i) a nucleic acid sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 56%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO:3, 5, 7, 9, 11, 13, 15, 17, 23, 25, 27 or 29; or (ii) a full complement of the nucleic acid sequence of (i).

Any of the foregoing isolated polynucleotides may be utilized in any recombinant DNA constructs (including suppression DNA constructs) of the present invention. The isolated polynucleotide preferably encodes a ferrochelatase.

Recombinant DNA Constructs and Suppression DNA Constructs:

In one aspect, the present invention includes recombinant DNA constructs (including suppression DNA constructs).

In one embodiment, a recombinant DNA construct comprises a polynucleotide operably linked to at least one regulatory sequence (e.g., a promoter functional in a plant), wherein the polynucleotide comprises (i) a nucleic acid sequence encoding an amino acid sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 56%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28 or 30; or (ii) a full complement of the nucleic acid sequence of (i).

In another embodiment, a recombinant DNA construct comprises a polynucleotide operably linked to at least one regulatory sequence (e.g., a promoter functional in a plant), wherein said polynucleotide comprises (i) a nucleic acid

sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 56%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27 or 29; or (ii) a full complement of the nucleic acid sequence of (i).

Figures 10A-10C show the multiple alignment of the amino acid sequences of the ferrochelataes of SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 48, 49 and 50. The multiple alignment of the sequences was performed using the Megalign® program of the LASERGENE® bioinformatics computing suite (DNASTAR® Inc., Madison, WI); in particular, using the Clustal V method of alignment (Higgins and Sharp (1989) *CABIOS*. 5:151-153) with the multiple alignment default parameters of GAP PENALTY=10 and GAP LENGTH PENALTY=10, and the pairwise alignment default parameters of KTUPLE=1, GAP PENALTY=3, WINDOW=5 and DIAGONALS SAVED=5.

Figure 11 shows the percent sequence identity and the divergence values for each pair of amino acids sequences displayed in Figures 10A-10C.

Figures 12A-12C show the multiple alignment of the amino acid sequences of the ferrochelataes of SEQ ID NO:20, 22, 24, 26, 28, 30, 51, 52, 53 and 54. The multiple alignment of the sequences was performed using the Megalign® program of the LASERGENE® bioinformatics computing suite (DNASTAR® Inc., Madison, WI); in particular, using the Clustal V method of alignment (Higgins and Sharp (1989) *CABIOS*. 5:151-153) with the multiple alignment default parameters of GAP PENALTY=10 and GAP LENGTH PENALTY=10, and the pairwise alignment default parameters of KTUPLE=1, GAP PENALTY=3, WINDOW=5 and DIAGONALS SAVED=5.

Figure 13 shows the percent sequence identity and the divergence values for each pair of amino acids sequences displayed in Figures 12A-12C.

In another embodiment, a recombinant DNA construct comprises a polynucleotide operably linked to at least one regulatory sequence (e.g., a promoter functional in a plant), wherein said polynucleotide encodes a ferrochelatae. The

ferrochelatase may be from *Arabidopsis thaliana*, *Zea mays*, *Glycine max*, *Glycine tabacina*, *Glycine soja* and *Glycine tomentella*.

In another aspect, the present invention includes suppression DNA constructs.

5 A suppression DNA construct may comprise at least one regulatory sequence (e.g., a promoter functional in a plant) operably linked to (a) all or part of: (i) a nucleic acid sequence encoding a polypeptide having an amino acid sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 56%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 10 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28 or 30, or (ii) a full complement of the nucleic acid sequence of (a)(i); or (b) a region derived from all or part of a sense 15 strand or antisense strand of a target gene of interest, said region having a nucleic acid sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 56%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% 20 sequence identity, based on the Clustal V method of alignment, when compared to said all or part of a sense strand or antisense strand from which said region is derived, and wherein said target gene of interest encodes a ferrochelatase; or (c) all or part of: (i) a nucleic acid sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 56%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 25 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27 or 29, or (ii) a full complement of the nucleic acid sequence of (c)(i). The suppression 30 DNA construct may comprise a cosuppression construct, antisense construct, viral-suppression construct, hairpin suppression construct, stem-loop suppression construct, double-stranded RNA-producing construct, RNAi construct, or small RNA construct (e.g., an siRNA construct or an miRNA construct).

It is understood, as those skilled in the art will appreciate, that the invention 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 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 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.

“Suppression DNA construct” is a recombinant DNA construct which when transformed or stably integrated into the genome of the plant, results in “silencing” of a target gene in the plant. The target gene may be endogenous or transgenic to the plant. “Silencing,” as used herein with respect to the target gene, refers generally to the suppression of levels of mRNA or protein/enzyme expressed by the target gene, and/or the level of the enzyme activity or protein functionality. The terms “suppression”, “suppressing” and “silencing”, used interchangeably herein, include lowering, reducing, declining, decreasing, inhibiting, eliminating or preventing. “Silencing” or “gene silencing” does not specify mechanism and is inclusive, and not limited to, anti-sense, cosuppression, viral-suppression, hairpin suppression, stem-loop suppression, RNAi-based approaches, and small RNA-based approaches.

A suppression DNA construct may comprise a region derived from a target gene of interest and may comprise all or part of the nucleic acid sequence of the sense strand (or antisense strand) of the target gene of interest. Depending upon the approach to be utilized, the region may be 100% identical or less than 100% identical (e.g., at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 56%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%,

88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical) to all or part of the sense strand (or antisense strand) of the gene of interest.

Suppression DNA constructs are well-known in the art, are readily constructed once the target gene of interest is selected, and include, without
5 limitation, cosuppression constructs, antisense constructs, viral-suppression constructs, hairpin suppression constructs, stem-loop suppression constructs, double-stranded RNA-producing constructs, and more generally, RNAi (RNA interference) constructs and small RNA constructs such as siRNA (short interfering RNA) constructs and miRNA (microRNA) constructs.

10 “Antisense inhibition” refers to the production of antisense RNA transcripts capable of suppressing the expression of the target gene or gene product.

“Antisense RNA” refers to an RNA transcript that is complementary to all or part of a target primary transcript or mRNA and that blocks the expression of a target isolated nucleic acid fragment (U.S. Patent No. 5,107,065). The complementarity of an
15 antisense RNA may be with any part of the specific gene transcript, i.e., at the 5' non-coding sequence, 3' non-coding sequence, introns, or the coding sequence.

“Cosuppression” refers to the production of sense RNA transcripts capable of suppressing the expression of the target gene or gene product. “Sense” RNA refers to RNA transcript that includes the mRNA and can be translated into protein within a
20 cell or *in vitro*. Cosuppression constructs in plants have been previously designed by focusing on overexpression of a nucleic acid sequence having homology to a native mRNA, in the sense orientation, which results in the reduction of all RNA having homology to the overexpressed sequence (see Vaucheret et al., *Plant J.* 16:651-659 (1998); and Gura, *Nature* 404:804-808 (2000)).

25 Another variation describes the use of plant viral sequences to direct the suppression of proximal mRNA encoding sequences (PCT Publication No. WO 98/36083 published on August 20, 1998).

Previously described is the use of “hairpin” structures that incorporate all, or part, of an mRNA encoding sequence in a complementary orientation that results in
30 a potential “stem-loop” structure for the expressed RNA (PCT Publication No. WO 99/53050 published on October 21, 1999). In this case the stem is formed by polynucleotides corresponding to the gene of interest inserted in either sense or anti-sense orientation with respect to the promoter and the loop is formed by some

polynucleotides of the gene of interest, which do not have a complement in the construct. This increases the frequency of cosuppression or silencing in the recovered transgenic plants. For review of hairpin suppression see Wesley, S.V. et al. (2003) *Methods in Molecular Biology, Plant Functional Genomics: Methods and*
5 *Protocols* 236:273-286.

A construct where the stem is formed by at least 30 nucleotides from a gene to be suppressed and the loop is formed by a random nucleotide sequence has also effectively been used for suppression (PCT Publication No. WO 99/61632 published on December 2, 1999).

10 The use of poly-T and poly-A sequences to generate the stem in the stem-loop structure has also been described (PCT Publication No. WO 02/00894 published January 3, 2002).

Yet another variation includes using synthetic repeats to promote formation of a stem in the stem-loop structure. Transgenic organisms prepared with such
15 recombinant DNA fragments have been shown to have reduced levels of the protein encoded by the nucleotide fragment forming the loop as described in PCT Publication No. WO 02/00904, published January 3, 2002.

RNA interference refers to the process of sequence-specific post-transcriptional gene silencing in animals mediated by short interfering RNAs
20 (siRNAs) (Fire et al., *Nature* 391:806 (1998)). The corresponding process in plants is commonly referred to as post-transcriptional gene silencing (PTGS) or RNA silencing and is also referred to as quelling in fungi. The process of post-transcriptional gene silencing is thought to be an evolutionarily-conserved cellular defense mechanism used to prevent the expression of foreign genes and is
25 commonly shared by diverse flora and phyla (Fire et al., *Trends Genet.* 15:358 (1999)). Such protection from foreign gene expression may have evolved in response to the production of double-stranded RNAs (dsRNAs) derived from viral infection or from the random integration of transposon elements into a host genome via a cellular response that specifically destroys homologous single-stranded RNA
30 of viral genomic RNA. The presence of dsRNA in cells triggers the RNAi response through a mechanism that has yet to be fully characterized.

The presence of long dsRNAs in cells stimulates the activity of a ribonuclease III enzyme referred to as dicer. Dicer is involved in the processing of

the dsRNA into short pieces of dsRNA known as short interfering RNAs (siRNAs) (Berstein et al., *Nature* 409:363 (2001)). Short interfering RNAs derived from dicer activity are typically about 21 to about 23 nucleotides in length and comprise about 19 base pair duplexes (Elbashir et al., *Genes Dev.* 15:188 (2001)). Dicer has also
5 been implicated in the excision of 21- and 22-nucleotide small temporal RNAs (stRNAs) from precursor RNA of conserved structure that are implicated in translational control (Hutvagner et al., *Science* 293:834 (2001)). The RNAi response also features an endonuclease complex, commonly referred to as an RNA-induced silencing complex (RISC), which mediates cleavage of single-
10 stranded RNA having sequence complementarity to the antisense strand of the siRNA duplex. Cleavage of the target RNA takes place in the middle of the region complementary to the antisense strand of the siRNA duplex. In addition, RNA interference can also involve small RNA (e.g., miRNA) mediated gene silencing, presumably through cellular mechanisms that regulate chromatin structure and
15 thereby prevent transcription of target gene sequences (see, e.g., Allshire, *Science* 297:1818-1819 (2002); Volpe et al., *Science* 297:1833-1837 (2002); Jenuwein, *Science* 297:2215-2218 (2002); and Hall et al., *Science* 297:2232-2237 (2002)). As such, miRNA molecules of the invention can be used to mediate gene silencing via interaction with RNA transcripts or alternately by interaction with particular gene
20 sequences, wherein such interaction results in gene silencing either at the transcriptional or post-transcriptional level.

RNAi has been studied in a variety of systems. Fire et al. (*Nature* 391:806 (1998)) were the first to observe RNAi in *Caenorhabditis elegans*. Wianny and Goetz (*Nature Cell Biol.* 2:70 (1999)) describe RNAi mediated by dsRNA in mouse
25 embryos. Hammond et al. (*Nature* 404:293 (2000)) describe RNAi in *Drosophila* cells transfected with dsRNA. Elbashir et al., (*Nature* 411:494 (2001)) describe RNAi induced by introduction of duplexes of synthetic 21-nucleotide RNAs in cultured mammalian cells including human embryonic kidney and HeLa cells.

Small RNAs play an important role in controlling gene expression. Regulation
30 of many developmental processes, including flowering, is controlled by small RNAs. It is now possible to engineer changes in gene expression of plant genes by using transgenic constructs which produce small RNAs in the plant.

Small RNAs appear to function by base-pairing to complementary RNA or DNA target sequences. When bound to RNA, small RNAs trigger either RNA cleavage or translational inhibition of the target sequence. When bound to DNA target sequences, it is thought that small RNAs can mediate DNA methylation of the target sequence. The consequence of these events, regardless of the specific mechanism, is that gene expression is inhibited.

It is thought that sequence complementarity between small RNAs and their RNA targets helps to determine which mechanism, RNA cleavage or translational inhibition, is employed. It is believed that siRNAs, which are perfectly complementary with their targets, work by RNA cleavage. Some miRNAs have perfect or near-perfect complementarity with their targets, and RNA cleavage has been demonstrated for at least a few of these miRNAs. Other miRNAs have several mismatches with their targets, and apparently inhibit their targets at the translational level. Again, without being held to a particular theory on the mechanism of action, a general rule is emerging that perfect or near-perfect complementarity causes RNA cleavage, whereas translational inhibition is favored when the miRNA/target duplex contains many mismatches. The apparent exception to this is microRNA 172 (miR172) in plants. One of the targets of miR172 is APETALA2 (AP2), and although miR172 shares near-perfect complementarity with AP2 it appears to cause translational inhibition of AP2 rather than RNA cleavage.

MicroRNAs (miRNAs) are noncoding RNAs of about 19 to about 24 nucleotides (nt) in length that have been identified in both animals and plants (Lagos-Quintana et al., *Science* 294:853-858 (2001), Lagos-Quintana et al., *Curr. Biol.* 12:735-739 (2002); Lau et al., *Science* 294:858-862 (2001); Lee and Ambros, *Science* 294:862-864 (2001); Llave et al., *Plant Cell* 14:1605-1619 (2002); Mourelatos et al., *Genes. Dev.* 16:720-728 (2002); Park et al., *Curr. Biol.* 12:1484-1495 (2002); Reinhart et al., *Genes. Dev.* 16:1616-1626 (2002)). They are processed from longer precursor transcripts that range in size from approximately 70 to 200 nt, and these precursor transcripts have the ability to form stable hairpin structures. In animals, the enzyme involved in processing miRNA precursors is called dicer, an RNase III-like protein (Grishok et al., *Cell* 106:23-34 (2001); Hutvagner et al., *Science* 293:834-838 (2001); Ketting et al., *Genes. Dev.* 15:2654-2659 (2001)). Plants also have a dicer-like enzyme, DCL1 (previously named

CARPEL FACTORY/SHORT INTEGUMENTS1/ SUSPENSOR1), and recent evidence indicates that it, like dicer, is involved in processing the hairpin precursors to generate mature miRNAs (Park et al., *Curr. Biol.* 12:1484-1495 (2002); Reinhart et al., *Genes Dev.* 16:1616-1626 (2002)). Furthermore, it is becoming clear from recent work that at least some miRNA hairpin precursors originate as longer polyadenylated transcripts, and several different miRNAs and associated hairpins can be present in a single transcript (Lagos-Quintana et al., *Science* 294:853-858 (2001); Lee et al., *EMBO J.* 21:4663-4670 (2002)). Recent work has also examined the selection of the miRNA strand from the dsRNA product arising from processing of the hairpin by DICER (Schwartz et al., *Cell* 115:199-208 (2003)). It appears that the stability (i.e. G:C versus A:U content, and/or mismatches) of the two ends of the processed dsRNA affects the strand selection, with the low stability end being easier to unwind by a helicase activity. The 5' end strand at the low stability end is incorporated into the RISC complex, while the other strand is degraded.

MicroRNAs (miRNAs) appear to regulate target genes by binding to complementary sequences located in the transcripts produced by these genes. In the case of lin-4 and let-7, the target sites are located in the 3' UTRs of the target mRNAs (Lee et al., *Cell* 75:843-854 (1993); Wightman et al., *Cell* 75:855-862 (1993); Reinhart et al., *Nature* 403:901-906 (2000); Slack et al., *Mol. Cell* 5:659-669 (2000)), and there are several mismatches between the lin-4 and let-7 miRNAs and their target sites. Binding of the lin-4 or let-7 miRNA appears to cause downregulation of steady-state levels of the protein encoded by the target mRNA without affecting the transcript itself (Olsen and Ambros, *Dev. Biol.* 216:671-680 (1999)). On the other hand, recent evidence suggests that miRNAs can in some cases cause specific RNA cleavage of the target transcript within the target site, and this cleavage step appears to require 100% complementarity between the miRNA and the target transcript (Hutvagner and Zamore, *Science* 297:2056-2060 (2002); Llave et al., *Plant Cell* 14:1605-1619 (2002)). It seems likely that miRNAs can enter at least two pathways of target gene regulation: (1) protein downregulation when target complementarity is <100%; and (2) RNA cleavage when target complementarity is 100%. MicroRNAs entering the RNA cleavage pathway are analogous to the 21-25 nt short interfering RNAs (siRNAs) generated during RNA interference (RNAi) in animals and posttranscriptional gene silencing (PTGS) in

plants, and likely are incorporated into an RNA-induced silencing complex (RISC) that is similar or identical to that seen for RNAi.

Identifying the targets of miRNAs with bioinformatics has not been successful in animals, and this is probably due to the fact that animal miRNAs have a low degree of complementarity with their targets. On the other hand, bioinformatic approaches have been successfully used to predict targets for plant miRNAs (Llave et al., *Plant Cell* 14:1605-1619 (2002); Park et al., *Curr. Biol.* 12:1484-1495 (2002); Rhoades et al., *Cell* 110:513-520 (2002)), and thus it appears that plant miRNAs have higher overall complementarity with their putative targets than do animal miRNAs. Most of these predicted target transcripts of plant miRNAs encode members of transcription factor families implicated in plant developmental patterning or cell differentiation.

Regulatory Sequences:

A recombinant DNA construct (including a suppression DNA construct) of the present invention may comprise at least one regulatory sequence.

A regulatory sequence may be a promoter.

A number of promoters can be used in recombinant DNA constructs of the present invention. The promoters can be selected based on the desired outcome, and may include constitutive, tissue-specific, inducible, or other promoters for expression in the host organism.

High level, constitutive expression of the candidate gene under control of the 35S or UBI promoter may have pleiotropic effects, although candidate gene efficacy may be estimated when driven by a constitutive promoter. Use of tissue-specific and/or stress-specific promoters may eliminate undesirable effects but retain the ability to enhance drought tolerance. This effect has been observed in *Arabidopsis* (Kasuga et al. (1999) *Nature Biotechnol.* 17:287-91).

Suitable constitutive promoters for use in a plant host cell include, for example, the core promoter of the Rsyn7 promoter and other constitutive promoters disclosed in WO 99/43838 and U.S. Patent No. 6,072,050; the core CaMV 35S promoter (Odell et al., *Nature* 313:810-812 (1985)); rice actin (McElroy et al., *Plant Cell* 2:163-171 (1990)); ubiquitin (Christensen et al., *Plant Mol. Biol.* 12:619-632 (1989) and Christensen et al., *Plant Mol. Biol.* 18:675-689 (1992)); pEMU (Last et al., *Theor. Appl. Genet.* 81:581-588 (1991)); MAS (Velten et al., *EMBO J.* 3:2723-

2730 (1984)); ALS promoter (U.S. Patent No. 5,659,026), and the like. Other constitutive promoters include, for example, those discussed in U.S. Patent Nos. 5,608,149; 5,608,144; 5,604,121; 5,569,597; 5,466,785; 5,399,680; 5,268,463; 5,608,142; and 6,177,611.

5 In choosing a promoter to use in the methods of the invention, it may be desirable to use a tissue-specific or developmentally regulated promoter.

A tissue-specific or developmentally regulated promoter is a DNA sequence which regulates the expression of a DNA sequence selectively in the cells/tissues of a plant critical to tassel development, seed set, or both, and limits the expression of
10 such a DNA sequence to the period of tassel development or seed maturation in the plant. Any identifiable promoter may be used in the methods of the present invention which causes the desired temporal and spatial expression.

Promoters which are seed or embryo-specific and may be useful in the invention include soybean Kunitz trypsin inhibitor (Kti3, Jofuku and Goldberg, Plant
15 Cell 1:1079-1093 (1989)), patatin (potato tubers) (Rocha-Sosa, M., et al. (1989) EMBO J. 8:23-29), convicilin, vicilin, and legumin (pea cotyledons) (Rerie, W.G., et al. (1991) Mol. Gen. Genet. 259:149-157; Newbigin, E.J., et al. (1990) Planta 180:461-470; Higgins, T.J.V., et al. (1988) Plant. Mol. Biol. 11:683-695), zein (maize endosperm) (Schemthaner, J.P., et al. (1988) EMBO J. 7:1249-1255), phaseolin
20 (bean cotyledon) (Segupta-Gopalan, C., et al. (1985) Proc. Natl. Acad. Sci. U.S.A. 82:3320-3324), phytohemagglutinin (bean cotyledon) (Voelker, T. et al. (1987) EMBO J. 6:3571-3577), B-conglycinin and glycinin (soybean cotyledon) (Chen, Z-L, et al. (1988) EMBO J. 7:297- 302), glutelin (rice endosperm), hordein (barley endosperm) (Marris, C., et al. (1988) Plant Mol. Biol. 10:359-366), glutenin and
25 gliadin (wheat endosperm) (Colot, V., et al. (1987) EMBO J. 6:3559-3564), and sporamin (sweet potato tuberous root) (Hattori, T., et al. (1990) Plant Mol. Biol. 14:595-604). Promoters of seed-specific genes operably linked to heterologous coding regions in chimeric gene constructions maintain their temporal and spatial expression pattern in transgenic plants. Such examples include *Arabidopsis thaliana*
30 2S seed storage protein gene promoter to express enkephalin peptides in *Arabidopsis* and *Brassica napus* seeds (Vanderkerckhove et al., Bio/Technology 7:L929-932 (1989)), bean lectin and bean beta-phaseolin promoters to express luciferase (Riggs et al., Plant Sci. 63:47-57 (1989)), and wheat glutenin promoters to

express chloramphenicol acetyl transferase (Colot et al., EMBO J 6:3559- 3564 (1987)).

Inducible promoters selectively express an operably linked DNA sequence in response to the presence of an endogenous or exogenous stimulus, for example by chemical compounds (chemical inducers) or in response to environmental, hormonal, chemical, and/or developmental signals. Inducible or regulated promoters include, for example, promoters regulated by light, heat, stress, flooding or drought, phytohormones, wounding, or chemicals such as ethanol, jasmonate, salicylic acid, or safeners.

Promoters for use in the current invention include the following: 1) the stress-inducible RD29A promoter (Kasuga et al. (1999) Nature Biotechnol. 17:287-91); 2) the barley promoter, B22E; expression of B22E is specific to the pedicel in developing maize kernels ("Primary Structure of a Novel Barley Gene Differentially Expressed in Immature Aleurone Layers". Klemsdal, S.S. et al., Mol. Gen. Genet. 228(1/2):9-16 (1991)); and 3) maize promoter, Zag2 ("Identification and molecular characterization of ZAG1, the maize homolog of the *Arabidopsis* floral homeotic gene AGAMOUS", Schmidt, R.J. et al., Plant Cell 5(7):729-737 (1993); "Structural characterization, chromosomal localization and phylogenetic evaluation of two pairs of AGAMOUS-like *MADS-box* genes from maize", Theissen et al. Gene 156(2):155-166 (1995); NCBI GenBank Accession No. X80206)). Zag2 transcripts can be detected 5 days prior to pollination to 7 to 8 days after pollination ("DAP"), and directs expression in the carpel of developing female inflorescences and Cim1 which is specific to the nucleus of developing maize kernels. Cim1 transcript is detected 4 to 5 days before pollination to 6 to 8 DAP. Other useful promoters include any promoter which can be derived from a gene whose expression is maternally associated with developing female florets.

Additional promoters for regulating the expression of the nucleotide sequences of the present invention in plants are stalk-specific promoters. Such stalk-specific promoters include the alfalfa S2A promoter (GenBank Accession No. EF030816; Abrahams et al., Plant Mol. Biol. 27:513-528 (1995)) and S2B promoter (GenBank Accession No. EF030817) and the like, herein incorporated by reference.

Promoters may be derived in their entirety from a native gene, or be composed of different elements derived from different promoters found in nature, or

even comprise synthetic DNA segments. It is understood by those skilled in the art that different promoters may direct the expression of a gene in different tissues or cell types, or at different stages of development, or in response to different environmental conditions. It is further recognized that since in most cases the exact boundaries of regulatory sequences have not been completely defined, DNA fragments of some variation may have identical promoter activity. Promoters that cause a gene to be expressed in most cell types at most times are commonly referred to as “constitutive promoters”. New promoters of various types useful in plant cells are constantly being discovered; numerous examples may be found in the compilation by Okamura, J. K., and Goldberg, R. B., *Biochemistry of Plants* 15:1-82 (1989).

Promoters for use in the current invention may include: RIP2, mLIP15, ZmCOR1, Rab17, CaMV 35S, RD29A, B22E, Zag2, SAM synthetase, ubiquitin, CaMV 19S, nos, Adh, sucrose synthase, R-allele, the vascular tissue preferred promoters S2A (Genbank accession number EF030816) and S2B (Genbank accession number EF030817), and the constitutive promoter GOS2 from *Zea mays*. Other promoters include root preferred promoters, such as the maize NAS2 promoter, the maize Cyclo promoter (US 2006/0156439, published July 13, 2006), the maize ROOTMET2 promoter (WO05063998, published July 14, 2005), the CR1BIO promoter (WO06055487, published May 26, 2006), the CRWAQ81 (WO05035770, published April 21, 2005) and the maize ZRP2.47 promoter (NCBI accession number: U38790; GI No. 1063664),

Recombinant DNA constructs of the present invention may also include other regulatory sequences, including but not limited to, translation leader sequences, introns, and polyadenylation recognition sequences. In another embodiment of the present invention, a recombinant DNA construct of the present invention further comprises an enhancer or silencer.

An intron sequence can be added to the 5' untranslated region, the protein-coding region or the 3' untranslated region to increase the amount of the mature message that accumulates in the cytosol. Inclusion of a spliceable intron in the transcription unit in both plant and animal expression constructs has been shown to increase gene expression at both the mRNA and protein levels up to 1000-fold. Buchman and Berg, *Mol. Cell Biol.* 8:4395-4405 (1988); Callis et al., *Genes Dev.*

1:1183-1200 (1987). Such intron enhancement of gene expression is typically greatest when placed near the 5' end of the transcription unit. Use of maize introns Adh1-S intron 1, 2, and 6, the Bronze-1 intron are known in the art. See generally, *The Maize Handbook*, Chapter 116, Freeling and Walbot, Eds., Springer, New York (1994).

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

A translation leader sequence is a DNA sequence located between the promoter sequence of a gene and the coding sequence. The translation leader sequence is present in the fully processed mRNA upstream of the translation start sequence. The translation leader sequence may affect processing of the primary transcript to mRNA, mRNA stability or translation efficiency. Examples of translation leader sequences have been described (Turner, R. and Foster, G. D. (1995) *Molecular Biotechnology* 3:225).

Any plant can be selected for the identification of regulatory sequences and ferrochelatase genes to be used in recombinant DNA constructs of the present invention. Examples of suitable plant targets for the isolation of genes and regulatory sequences would include but are not limited to alfalfa, apple, apricot, *Arabidopsis*, artichoke, arugula, asparagus, avocado, banana, barley, beans, beet, blackberry, blueberry, broccoli, brussels sprouts, cabbage, canola, cantaloupe, carrot, cassava, castorbean, cauliflower, celery, cherry, chicory, cilantro, citrus, clementines, clover, coconut, coffee, corn, cotton, cranberry, cucumber, Douglas fir, eggplant, endive, escarole, eucalyptus, fennel, figs, garlic, gourd, grape, grapefruit, honey dew, jicama, kiwifruit, lettuce, leeks, lemon, lime, Loblolly pine, linseed, mango, melon, mushroom, nectarine, nut, oat, oil palm, oil seed rape, okra, olive, onion, orange, an ornamental plant, palm, papaya, parsley, parsnip, pea, peach, peanut, pear, pepper, persimmon, pine, pineapple, plantain, plum, pomegranate, poplar, potato, pumpkin, quince, radiata pine, radicchio, radish, rapeseed, raspberry, rice, rye, sorghum, Southern pine, soybean, spinach, squash, strawberry,

sugarbeet, sugarcane, sunflower, sweet potato, sweetgum, tangerine, tea, tobacco, tomato, triticale, turf, turnip, a vine, watermelon, wheat, yams, and zucchini.

Compositions:

A composition of the present invention is a plant comprising in its genome
5 any of the recombinant DNA constructs (including any of the suppression DNA constructs) of the present invention (such as any of the constructs discussed above). Compositions also include any progeny of the plant, and any seed obtained from the plant or its progeny, wherein the progeny or seed comprises within its genome the recombinant DNA construct (or suppression DNA construct). Progeny
10 includes subsequent generations obtained by self-pollination or out-crossing of a plant. Progeny also includes hybrids and inbreds.

In hybrid seed propagated crops, mature transgenic plants can be self-pollinated to produce a homozygous inbred plant. The inbred plant produces seed containing the newly introduced recombinant DNA construct (or suppression DNA
15 construct). These seeds can be grown to produce plants that would exhibit an altered agronomic characteristic (e.g., an increased agronomic characteristic optionally under water limiting conditions), or used in a breeding program to produce hybrid seed, which can be grown to produce plants that would exhibit such an altered agronomic characteristic. The seeds may be maize seeds.

20 The plant may be a monocotyledonous or dicotyledonous plant, for example, a maize or soybean plant, such as a maize hybrid plant or a maize inbred plant. The plant may also be sunflower, sorghum, canola, wheat, alfalfa, cotton, rice, barley or millet.

The recombinant DNA construct may be stably integrated into the genome of
25 the plant.

Particularly embodiments include but are not limited to the following:

1. A plant (for example, a maize or soybean plant) comprising in its genome a recombinant DNA construct comprising a polynucleotide operably linked to at least one regulatory sequence, wherein said polynucleotide encodes a
30 polypeptide having an amino acid sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 56%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%,

97%, 98%, 99%, or 100% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28 or 30, and wherein said plant exhibits increased drought tolerance when compared to a control plant not comprising said recombinant DNA construct. The plant may further exhibit an alteration of at least one agronomic characteristic when compared to the control plant.

2. A plant (for example, a maize or soybean plant) comprising in its genome a recombinant DNA construct comprising a polynucleotide operably linked to at least one regulatory sequence, wherein said polynucleotide encodes a ferrochelatase, and wherein said plant exhibits increased drought tolerance when compared to a control plant not comprising said recombinant DNA construct. The plant may further exhibit an alteration of at least one agronomic characteristic when compared to the control plant.

3. A plant (for example, a maize or soybean plant) comprising in its genome a recombinant DNA construct comprising a polynucleotide operably linked to at least one regulatory sequence, wherein said polynucleotide encodes a ferrochelatase, and wherein said plant exhibits an alteration of at least one agronomic characteristic when compared to a control plant not comprising said recombinant DNA construct.

4. A plant (for example, a maize or soybean plant) comprising in its genome a recombinant DNA construct comprising a polynucleotide operably linked to at least one regulatory element, wherein said polynucleotide encodes a polypeptide having an amino acid sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 56%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28 or 30, and wherein said plant exhibits an alteration of at least one agronomic characteristic when compared to a control plant not comprising said recombinant DNA construct.

5. A plant (for example, a maize or soybean plant) comprising in its genome a suppression DNA construct comprising at least one regulatory element

operably linked to a region derived from all or part of a sense strand or antisense strand of a target gene of interest, said region having a nucleic acid sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 56%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity, based on the Clustal V method of alignment, when compared to said all or part of a sense strand or antisense strand from which said region is derived, and wherein said target gene of interest encodes a ferrochelatase, and wherein said plant exhibits an alteration of at least one agronomic characteristic when compared to a control plant not comprising said suppression DNA construct.

6. A plant (for example, a maize or soybean plant) comprising in its genome a suppression DNA construct comprising at least one regulatory element operably linked to all or part of (a) a nucleic acid sequence encoding a polypeptide having an amino acid sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 56%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28 or 30, or (b) a full complement of the nucleic acid sequence of (a), and wherein said plant exhibits an alteration of at least one agronomic characteristic when compared to a control plant not comprising said suppression DNA construct.

7. Any progeny of the above plants in embodiments 1-6, any seeds of the above plants in embodiments 1-6, any seeds of progeny of the above plants in embodiments 1-6, and cells from any of the above plants in embodiments 1-6 and progeny thereof.

In any of the foregoing embodiments 1-7 or any other embodiments of the present invention, the ferrochelatase may be from *Arabidopsis thaliana*, *Zea mays*, *Glycine max*, *Glycine tabacina*, *Glycine soja* or *Glycine tomentella*.

In any of the foregoing embodiments 1-7 or any other embodiments of the present invention, the recombinant DNA construct (or suppression DNA construct) may comprise at least a promoter functional in a plant as a regulatory sequence.

In any of the foregoing embodiments 1-7 or any other embodiments of the present invention, the alteration of at least one agronomic characteristic is either an increase or decrease.

5 In any of the foregoing embodiments 1-7 or any other embodiments of the present invention, the at least one agronomic characteristic may be selected from the group consisting of greenness, yield, growth rate, biomass, fresh weight at maturation, dry weight at maturation, fruit yield, seed yield, total plant nitrogen content, fruit nitrogen content, seed nitrogen content, nitrogen content in a vegetative tissue, total plant free amino acid content, fruit free amino acid content, 10 seed free amino acid content, free amino acid content in a vegetative tissue, total plant protein content, fruit protein content, seed protein content, protein content in a vegetative tissue, drought tolerance, nitrogen uptake, root lodging, harvest index, stalk lodging, plant height, ear height and ear length. For example, the alteration of at least one agronomic characteristic may be an increase in yield, greenness or 15 biomass.

In any of the foregoing embodiments 1-7 or any other embodiments of the present invention, the plant may exhibit the alteration of at least one agronomic characteristic when compared, under water limiting conditions, to a control plant not comprising said recombinant DNA construct (or said suppression DNA construct).

20 "Drought" refers to a decrease in water availability to a plant that, especially when prolonged, can cause damage to the plant or prevent its successful growth (e.g., limiting plant growth or seed yield).

"Drought tolerance" is a trait of a plant to survive under drought conditions over prolonged periods of time without exhibiting substantial physiological or 25 physical deterioration.

"Increased drought tolerance" of a plant is measured relative to a reference or control plant, and is a trait of the plant to survive under drought conditions over prolonged periods of time, without exhibiting the same degree of physiological or physical deterioration relative to the reference or control plant grown under similar 30 drought conditions. Typically, when a transgenic plant comprising a recombinant DNA construct or suppression DNA construct in its genome exhibits increased drought tolerance relative to a reference or control plant, the reference or control

plant does not comprise in its genome the recombinant DNA construct or suppression DNA construct.

One of ordinary skill in the art is familiar with protocols for simulating drought conditions and for evaluating drought tolerance of plants that have been subjected to simulated or naturally-occurring drought conditions. For example, one can simulate drought conditions by giving plants less water than normally required or no water over a period of time, and one can evaluate drought tolerance by looking for differences in physiological and/or physical condition, including (but not limited to) vigor, growth, size, or root length, or in particular, leaf color or leaf area size. Other techniques for evaluating drought tolerance include measuring chlorophyll fluorescence, photosynthetic rates and gas exchange rates.

A drought stress experiment may involve a chronic stress (i.e., slow dry down) and/or may involve two acute stresses (i.e., abrupt removal of water) separated by a day or two of recovery. Chronic stress may last 8 – 10 days. Acute stress may last 3 – 5 days. The following variables may be measured during drought stress and well watered treatments of transgenic plants and relevant control plants:

The variable “% area chg_start chronic - acute2” is a measure of the percent change in total area determined by remote visible spectrum imaging between the first day of chronic stress and the day of the second acute stress

The variable “% area chg_start chronic - end chronic” is a measure of the percent change in total area determined by remote visible spectrum imaging between the first day of chronic stress and the last day of chronic stress

The variable “% area chg_start chronic – harvest” is a measure of the percent change in total area determined by remote visible spectrum imaging between the first day of chronic stress and the day of harvest

The variable “% area chg_start chronic - recovery24hr” is a measure of the percent change in total area determined by remote visible spectrum imaging between the first day of chronic stress and 24 hrs into the recovery (24hrs after acute stress 2)

The variable “psii_acute1” is a measure of Photosystem II (PSII) efficiency at the end of the first acute stress period. It provides an estimate of the efficiency at

which light is absorbed by PSII antennae and is directly related to carbon dioxide assimilation within the leaf.

The variable "psii_acute2" is a measure of Photosystem II (PSII) efficiency at the end of the second acute stress period. It provides an estimate of the efficiency at which light is absorbed by PSII antennae and is directly related to carbon dioxide assimilation within the leaf.

The variable "fv/fm_acute1" is a measure of the optimum quantum yield (Fv/Fm) at the end of the first acute stress - (variable fluorescence difference between the maximum and minimum fluorescence / maximum fluorescence)

The variable "fv/fm_acute2" is a measure of the optimum quantum yield (Fv/Fm) at the end of the second acute stress - (variable fluorescence difference between the maximum and minimum fluorescence / maximum fluorescence)

The variable "leaf rolling_harvest" is a measure of the ratio of top image to side image on the day of harvest.

The variable "leaf rolling_recovery24hr" is a measure of the ratio of top image to side image 24 hours into the recovery.

The variable "Specific Growth Rate (SGR)" represents the change in total plant surface area (as measured by Lemna Tec Instrument) over a single day ($Y(t) = Y_0 \cdot e^{r \cdot t}$). $Y(t) = Y_0 \cdot e^{r \cdot t}$ is equivalent to % change in $Y/\Delta t$ where the individual terms are as follows: $Y(t)$ = Total surface area at t ; Y_0 = Initial total surface area (estimated); r = Specific Growth Rate day^{-1} , and t = Days After Planting ("DAP")

The variable "shoot dry weight" is a measure of the shoot weight 96 hours after being placed into a 104 °C oven

The variable "shoot fresh weight" is a measure of the shoot weight immediately after being cut from the plant

The Examples below describe some representative protocols and techniques for simulating drought conditions and/or evaluating drought tolerance.

One can also evaluate drought tolerance by the ability of a plant to maintain sufficient yield (at least 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% yield) in field testing under simulated or naturally-occurring drought conditions (e.g., by measuring for substantially equivalent yield under drought

conditions compared to non-drought conditions, or by measuring for less yield loss under drought conditions compared to a control or reference plant).

One of ordinary skill in the art would readily recognize a suitable control or reference plant to be utilized when assessing or measuring an agronomic characteristic or phenotype of a transgenic plant in any embodiment of the present invention in which a control plant is utilized (e.g., compositions or methods as described herein). For example, by way of non-limiting illustrations:

1. Progeny of a transformed plant which is hemizygous with respect to a recombinant DNA construct (or suppression DNA construct), such that the progeny are segregating into plants either comprising or not comprising the recombinant DNA construct (or suppression DNA construct): the progeny comprising the recombinant DNA construct (or suppression DNA construct) would be typically measured relative to the progeny not comprising the recombinant DNA construct (or suppression DNA construct) (i.e., the progeny not comprising the recombinant DNA construct (or the suppression DNA construct) is the control or reference plant).

2. Introgression of a recombinant DNA construct (or suppression DNA construct) into an inbred line, such as in maize, or into a variety, such as in soybean: the introgressed line would typically be measured relative to the parent inbred or variety line (i.e., the parent inbred or variety line is the control or reference plant).

3. Two hybrid lines, where the first hybrid line is produced from two parent inbred lines, and the second hybrid line is produced from the same two parent inbred lines except that one of the parent inbred lines contains a recombinant DNA construct (or suppression DNA construct): the second hybrid line would typically be measured relative to the first hybrid line (i.e., the first hybrid line is the control or reference plant).

4. A plant comprising a recombinant DNA construct (or suppression DNA construct): the plant may be assessed or measured relative to a control plant not comprising the recombinant DNA construct (or suppression DNA construct) but otherwise having a comparable genetic background to the plant (e.g., sharing at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity of nuclear genetic material compared to the plant comprising the recombinant DNA construct (or suppression DNA construct)). There are many

laboratory-based techniques available for the analysis, comparison and characterization of plant genetic backgrounds; among these are Isozyme Electrophoresis, Restriction Fragment Length Polymorphisms (RFLPs), Randomly Amplified Polymorphic DNAs (RAPDs), Arbitrarily Primed Polymerase Chain Reaction (AP-PCR), DNA Amplification Fingerprinting (DAF), Sequence Characterized Amplified Regions (SCARs), Amplified Fragment Length Polymorphisms (AFLP®s), and Simple Sequence Repeats (SSRs) which are also referred to as Microsatellites.

Furthermore, one of ordinary skill in the art would readily recognize that a suitable control or reference plant to be utilized when assessing or measuring an agronomic characteristic or phenotype of a transgenic plant would not include a plant that had been previously selected, via mutagenesis or transformation, for the desired agronomic characteristic or phenotype.

Methods:

Methods include but are not limited to methods for increasing drought tolerance in a plant, methods for evaluating drought tolerance in a plant, methods for altering an agronomic characteristic in a plant, methods for determining an alteration of an agronomic characteristic in a plant, and methods for producing seed. The plant may be a monocotyledonous or dicotyledonous plant, for example, a maize or soybean plant. The plant may also be sunflower, sorghum, canola, wheat, alfalfa, cotton, rice, barley or millet. The seed is may be a maize or soybean seed, for example, a maize hybrid seed or maize inbred seed.

Methods include but are not limited to the following:

A method for transforming a cell comprising transforming a cell with any of the isolated polynucleotides of the present invention. The cell transformed by this method is also included. In particular embodiments, the cell is eukaryotic cell, e.g., a yeast, insect or plant cell, or prokaryotic, e.g., a bacterium.

A method for producing a transgenic plant comprising transforming a plant cell with any of the isolated polynucleotides or recombinant DNA constructs of the present invention and regenerating a transgenic plant from the transformed plant cell. The invention is also directed to the transgenic plant produced by this method, and transgenic seed obtained from this transgenic plant.

A method for isolating a polypeptide of the invention from a cell or culture medium of the cell, wherein the cell comprises a recombinant DNA construct comprising a polynucleotide of the invention operably linked to at least one regulatory sequence, and wherein the transformed host cell is grown under conditions that are suitable for expression of the recombinant DNA construct.

A method of altering the level of expression of a polypeptide of the invention in a host cell comprising: (a) transforming a host cell with a recombinant DNA construct of the present invention; and (b) growing the transformed host cell under conditions that are suitable for expression of the recombinant DNA construct wherein expression of the recombinant DNA construct results in production of altered levels of the polypeptide of the invention in the transformed host cell.

A method of increasing drought tolerance in a plant, comprising: (a) introducing into a regenerable plant cell a recombinant DNA construct comprising a polynucleotide operably linked to at least one regulatory sequence (for example, a promoter functional in a plant), wherein the polynucleotide encodes a polypeptide having an amino acid sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 56%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28 or 30; and (b) regenerating a transgenic plant from the regenerable plant cell after step (a), wherein the transgenic plant comprises in its genome the recombinant DNA construct and exhibits increased drought tolerance when compared to a control plant not comprising the recombinant DNA construct. The method may further comprise (c) obtaining a progeny plant derived from the transgenic plant, wherein said progeny plant comprises in its genome the recombinant DNA construct and exhibits increased drought tolerance when compared to a control plant not comprising the recombinant DNA construct.

A method of increasing drought tolerance in a plant, comprising: (a) introducing into a regenerable plant cell a suppression DNA construct comprising at least one regulatory sequence (for example, a promoter functional in a plant) operably linked to all or part of (i) a nucleic acid sequence encoding a polypeptide

having an amino acid sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 56%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28 or 30, or (ii) a full complement of the nucleic acid sequence of (a)(i); and (b) regenerating a transgenic plant from the regenerable plant cell after step (a), wherein the transgenic plant comprises in its genome the suppression DNA construct and exhibits increased drought tolerance when compared to a control plant not comprising the suppression DNA construct. The method may further comprise (c) obtaining a progeny plant derived from the transgenic plant, wherein said progeny plant comprises in its genome the suppression DNA construct and exhibits increased drought tolerance when compared to a control plant not comprising the suppression DNA construct..

A method of increasing drought tolerance in a plant, comprising: (a) introducing into a regenerable plant cell a suppression DNA construct comprising at least one regulatory sequence (for example, a promoter functional in a plant) operably linked to a region derived from all or part of a sense strand or antisense strand of a target gene of interest, said region having a nucleic acid sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 56%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity, based on the Clustal V method of alignment, when compared to said all or part of a sense strand or antisense strand from which said region is derived, and wherein said target gene of interest encodes a ferrochelatase; and (b) regenerating a transgenic plant from the regenerable plant cell after step (a), wherein the transgenic plant comprises in its genome the suppression DNA construct and exhibits increased drought tolerance when compared to a control plant not comprising the suppression DNA construct. The method may further comprise (c) obtaining a progeny plant derived from the transgenic plant, wherein said progeny plant comprises in its genome the suppression DNA construct and exhibits

increased drought tolerance when compared to a control plant not comprising the suppression DNA construct..

A method of evaluating drought tolerance in a plant, comprising (a) introducing into a regenerable plant cell a recombinant DNA construct comprising a polynucleotide operably linked to at least one regulatory sequence (for example, a promoter functional in a plant), wherein the polynucleotide encodes a polypeptide having an amino acid sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28 or 30; (b) regenerating a transgenic plant from the regenerable plant cell after step (a), wherein the transgenic plant comprises in its genome the recombinant DNA construct; and (c) evaluating the transgenic plant for drought tolerance compared to a control plant not comprising the recombinant DNA construct. The method may further comprise (d) obtaining a progeny plant derived from the transgenic plant, wherein the progeny plant comprises in its genome the recombinant DNA construct; and (e) evaluating the progeny plant for drought tolerance compared to a control plant not comprising the recombinant DNA construct.

A method of evaluating drought tolerance in a plant, comprising (a) introducing into a regenerable plant cell a suppression DNA construct comprising at least one regulatory sequence (for example, a promoter functional in a plant) operably linked to all or part of (i) a nucleic acid sequence encoding a polypeptide having an amino acid sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28 or 30, or (ii) a full complement of the nucleic acid sequence of (a)(i); (b) regenerating a transgenic plant from the regenerable plant cell after step (a), wherein the transgenic plant comprises in its genome the suppression DNA construct; and (c)

evaluating the transgenic plant for drought tolerance compared to a control plant not comprising the suppression DNA construct. The method may further comprise (d) obtaining a progeny plant derived from the transgenic plant, wherein the progeny plant comprises in its genome the suppression DNA construct; and (e) evaluating the progeny plant for drought tolerance compared to a control plant not comprising the suppression DNA construct.

A method of evaluating drought tolerance in a plant, comprising (a) introducing into a regenerable plant cell a suppression DNA construct comprising at least one regulatory sequence (for example, a promoter functional in a plant) operably linked to a region derived from all or part of a sense strand or antisense strand of a target gene of interest, said region having a nucleic acid sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 56%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity, based on the Clustal V method of alignment, when compared to said all or part of a sense strand or antisense strand from which said region is derived, and wherein said target gene of interest encodes a ferrochelatase; (b) regenerating a transgenic plant from the regenerable plant cell after step (a), wherein the transgenic plant comprises in its genome the suppression DNA construct; and (c) evaluating the transgenic plant for drought tolerance compared to a control plant not comprising the suppression DNA construct. The method may further comprise (d) obtaining a progeny plant derived from the transgenic plant, wherein the progeny plant comprises in its genome the suppression DNA construct; and (e) evaluating the progeny plant for drought tolerance compared to a control plant not comprising the suppression DNA construct.

A method of evaluating drought tolerance in a plant, comprising (a) introducing into a regenerable plant cell a recombinant DNA construct comprising a polynucleotide operably linked to at least one regulatory sequence (for example, a promoter functional in a plant), wherein said polynucleotide encodes a polypeptide having an amino acid sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 56%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%,

85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28 or 30; (b) regenerating a transgenic plant from the regenerable plant cell after step (a),
5 wherein the transgenic plant comprises in its genome the recombinant DNA construct; (c) obtaining a progeny plant derived from said transgenic plant, wherein the progeny plant comprises in its genome the recombinant DNA construct; and (d) evaluating the progeny plant for drought tolerance compared to a control plant not comprising the recombinant DNA construct.

10 A method of evaluating drought tolerance in a plant, comprising (a) introducing into a regenerable plant cell a suppression DNA construct comprising at least one regulatory sequence (for example, a promoter functional in a plant) operably linked to all or part of (i) a nucleic acid sequence encoding a polypeptide having an amino acid sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%,
15 57%, 58%, 59%, 60%, 56%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28 or 30, or (ii)
20 a full complement of the nucleic acid sequence of (a)(i); (b) regenerating a transgenic plant from the regenerable plant cell after step (a), wherein the transgenic plant comprises in its genome the suppression DNA construct; (c) obtaining a progeny plant derived from said transgenic plant, wherein the progeny plant comprises in its genome the suppression DNA construct; and (d) evaluating
25 the progeny plant for drought tolerance compared to a control plant not comprising the suppression DNA construct.

A method of evaluating drought tolerance in a plant, comprising (a) introducing into a regenerable plant cell a suppression DNA construct comprising at least one regulatory sequence (for example, a promoter functional in a plant)
30 operably linked to a region derived from all or part of a sense strand or antisense strand of a target gene of interest, said region having a nucleic acid sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 56%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%,

77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity, based on the Clustal V method of alignment, when compared to said all or part of a sense strand or antisense strand from which said region is derived, and wherein
5 said target gene of interest encodes a ferrochelatase; (b) regenerating a transgenic plant from the regenerable plant cell after step (a), wherein the transgenic plant comprises in its genome the suppression DNA construct; (c) obtaining a progeny plant derived from the transgenic plant, wherein the progeny plant comprises in its genome the suppression DNA construct; and (d) evaluating the progeny plant for
10 drought tolerance compared to a control plant not comprising the suppression DNA construct.

A method of determining an alteration of an agronomic characteristic in a plant, comprising (a) introducing into a regenerable plant cell a recombinant DNA construct comprising a polynucleotide operably linked to at least one regulatory
15 sequence (for example, a promoter functional in a plant), wherein said polynucleotide encodes a polypeptide having an amino acid sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%,
20 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28 or 30; (b) regenerating a transgenic plant from the regenerable plant cell after step (a), wherein the transgenic plant comprises in its genome said recombinant DNA construct; and (c) determining whether the
25 transgenic plant exhibits an alteration in at least one agronomic characteristic when compared, optionally under water limiting conditions, to a control plant not comprising the recombinant DNA construct. The method may further comprise (d) obtaining a progeny plant derived from the transgenic plant, wherein the progeny plant comprises in its genome the recombinant DNA construct; and (e) determining
30 whether the progeny plant exhibits an alteration in at least one agronomic characteristic when compared, optionally under water limiting conditions, to a control plant not comprising the recombinant DNA construct.

A method of determining an alteration of an agronomic characteristic in a plant, comprising (a) introducing into a regenerable plant cell a suppression DNA construct comprising at least one regulatory sequence (for example, a promoter functional in a plant) operably linked to all or part of (i) a nucleic acid sequence encoding a polypeptide having an amino acid sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28 or 30, or (ii) a full complement of the nucleic acid sequence of (i); (b) regenerating a transgenic plant from the regenerable plant cell after step (a), wherein the transgenic plant comprises in its genome the suppression DNA construct; and (c) determining whether the transgenic plant exhibits an alteration in at least one agronomic characteristic when compared, optionally under water limiting conditions, to a control plant not comprising the suppression DNA construct. The method may further comprise (d) obtaining a progeny plant derived from the transgenic plant, wherein the progeny plant comprises in its genome the suppression DNA construct; and (e) determining whether the progeny plant exhibits an alteration in at least one agronomic characteristic when compared, optionally under water limiting conditions, to a control plant not comprising the suppression DNA construct.

A method of determining an alteration of an agronomic characteristic in a plant, comprising (a) introducing into a regenerable plant cell a suppression DNA construct comprising at least one regulatory sequence (for example, a promoter functional in a plant) operably linked to a region derived from all or part of a sense strand or antisense strand of a target gene of interest, said region having a nucleic acid sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity, based on the Clustal V method of alignment, when compared to said all or part of a sense strand or antisense strand from which said region is

derived, and wherein said target gene of interest encodes a ferrochelatase; (b) regenerating a transgenic plant from the regenerable plant cell after step (a), wherein the transgenic plant comprises in its genome the suppression DNA construct; and (c) determining whether the transgenic plant exhibits an alteration in at least one agronomic characteristic when compared, optionally under water limiting conditions, to a control plant not comprising the suppression DNA construct. The method may further comprise (d) obtaining a progeny plant derived from the transgenic plant, wherein the progeny plant comprises in its genome the suppression DNA construct; and (e) determining whether the progeny plant exhibits an alteration in at least one agronomic characteristic when compared, optionally under water limiting conditions, to a control plant not comprising the suppression DNA construct.

A method of determining an alteration of an agronomic characteristic in a plant, comprising (a) introducing into a regenerable plant cell a recombinant DNA construct comprising a polynucleotide operably linked to at least one regulatory sequence (for example, a promoter functional in a plant), wherein said polynucleotide encodes a polypeptide having an amino acid sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28 or 30; (b) regenerating a transgenic plant from the regenerable plant cell after step (a), wherein the transgenic plant comprises in its genome said recombinant DNA construct; (c) obtaining a progeny plant derived from said transgenic plant, wherein the progeny plant comprises in its genome the recombinant DNA construct; and (d) determining whether the progeny plant exhibits an alteration in at least one agronomic characteristic when compared, optionally under water limiting conditions, to a control plant not comprising the recombinant DNA construct.

A method of determining an alteration of an agronomic characteristic in a plant, comprising (a) introducing into a regenerable plant cell a suppression DNA construct comprising at least one regulatory sequence (for example, a promoter

functional in a plant) operably linked to all or part of (i) a nucleic acid sequence encoding a polypeptide having an amino acid sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 56%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28 or 30, or (ii) a full complement of the nucleic acid sequence of (i); (b) regenerating a transgenic plant from the regenerable plant cell after step (a), wherein the transgenic plant comprises in its genome the suppression DNA construct; (c) obtaining a progeny plant derived from said transgenic plant, wherein the progeny plant comprises in its genome the suppression DNA construct; and (d) determining whether the progeny plant exhibits an alteration in at least one agronomic characteristic when compared, optionally under water limiting conditions, to a control plant not comprising the suppression DNA construct.

A method of determining an alteration of an agronomic characteristic in a plant, comprising (a) introducing into a regenerable plant cell a suppression DNA construct comprising at least one regulatory sequence (for example, a promoter functional in a plant) operably linked to a region derived from all or part of a sense strand or antisense strand of a target gene of interest, said region having a nucleic acid sequence of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 56%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity, based on the Clustal V method of alignment, when compared to said all or part of a sense strand or antisense strand from which said region is derived, and wherein said target gene of interest encodes a ferrochelatase; (b) regenerating a transgenic plant from the regenerable plant cell after step (a), wherein the transgenic plant comprises in its genome the suppression DNA construct; (c) obtaining a progeny plant derived from said transgenic plant, wherein the progeny plant comprises in its genome the suppression DNA construct; and (d) determining whether the progeny plant exhibits an alteration in at least one

agronomic characteristic when compared, optionally under water limiting conditions, to a control plant not comprising the suppression DNA construct.

A method of producing seed (for example, seed that can be sold as a drought tolerant product offering) comprising any of the preceding methods, and further comprising obtaining seeds from said progeny plant, wherein said seeds comprise in their genome said recombinant DNA construct (or suppression DNA construct).

In any of the preceding methods or any other embodiments of methods of the present invention, in said introducing step said regenerable plant cell may comprise a callus cell, an embryogenic callus cell, a gametic cell, a meristematic cell, or a cell of an immature embryo. The regenerable plant cells may derive from an inbred maize plant.

In any of the preceding methods or any other embodiments of methods of the present invention, said regenerating step may comprise the following: (i) culturing said transformed plant cells in a media comprising an embryogenic promoting hormone until callus organization is observed; (ii) transferring said transformed plant cells of step (i) to a first media which includes a tissue organization promoting hormone; and (iii) subculturing said transformed plant cells after step (ii) onto a second media, to allow for shoot elongation, root development or both.

In any of the preceding methods or any other embodiments of methods of the present invention, the at least one agronomic characteristic may be selected from the group consisting of greenness, yield, growth rate, biomass, fresh weight at maturation, dry weight at maturation, fruit yield, seed yield, total plant nitrogen content, fruit nitrogen content, seed nitrogen content, nitrogen content in a vegetative tissue, total plant free amino acid content, fruit free amino acid content, seed free amino acid content, amino acid content in a vegetative tissue, total plant protein content, fruit protein content, seed protein content, protein content in a vegetative tissue, drought tolerance, nitrogen uptake, root lodging, harvest index, stalk lodging, plant height, ear height and ear length. The alteration of at least one agronomic characteristic may be an increase in yield, greenness or biomass.

In any of the preceding methods or any other embodiments of methods of the present invention, the plant may exhibit the alteration of at least one agronomic characteristic when compared, under water limiting conditions, to a control plant not comprising said recombinant DNA construct (or said suppression DNA construct).

In any of the preceding methods or any other embodiments of methods of the present invention, alternatives exist for introducing into a regenerable plant cell a recombinant DNA construct comprising a polynucleotide operably linked to at least one regulatory sequence. For example, one may introduce into a regenerable plant cell a regulatory sequence (such as one or more enhancers, optionally as part of a transposable element), and then screen for an event in which the regulatory sequence is operably linked to an endogenous gene encoding a polypeptide of the instant invention.

The introduction of recombinant DNA constructs of the present invention into plants may be carried out by any suitable technique, including but not limited to direct DNA uptake, chemical treatment, electroporation, microinjection, cell fusion, infection, vector-mediated DNA transfer, bombardment, or *Agrobacterium*-mediated transformation.

Techniques are set forth below in the Examples below for transformation of maize plant cells and soybean plant cells.

Other methods for transforming dicots, primarily by use of *Agrobacterium tumefaciens*, and obtaining transgenic plants include those published for cotton (U.S. Patent No. 5,004,863, U.S. Patent No. 5,159,135, U.S. Patent No. 5,518, 908); soybean (U.S. Patent No. 5,569,834, U.S. Patent No. 5,416,011, McCabe et al., *BiolTechnology* 6:923 (1988), Christou et al., *Plant Physiol.* 87:671 674 (1988)); Brassica (U.S. Patent No. 5,463,174); peanut (Cheng et al., *Plant Cell Rep.* 15:653 657 (1996), McKently et al., *Plant Cell Rep.* 14:699 703 (1995)); papaya; and pea (Grant et al., *Plant Cell Rep.* 15:254 258, (1995)).

Transformation of monocotyledons using electroporation, particle bombardment, and *Agrobacterium* have also been reported, for example, transformation and plant regeneration as achieved in asparagus (Bytebier et al., *Proc. Natl. Acad. Sci. (USA)* 84:5354, (1987)); barley (Wan and Lemaux, *Plant Physiol* 104:37 (1994)); maize (Rhodes et al., *Science* 240:204 (1988), Gordon-Kamm et al., *Plant Cell* 2:603 618 (1990), Fromm et al., *BiolTechnology* 8:833 (1990), Koziel et al., *BiolTechnology* 11: 194, (1993), Armstrong et al., *Crop Science* 35:550 557 (1995)); oat (Somers et al., *BiolTechnology* 10: 15 89 (1992)); orchard grass (Horn et al., *Plant Cell Rep.* 7:469 (1988)); rice (Toriyama et al., *TheorAppl. Genet.* 205:34, (1986); Part et al., *Plant Mol. Biol.* 32:1135 1148, (1996); Abedinia et

al., Aust. J. Plant Physiol. 24:133 141 (1997); Zhang and Wu, Theor. Appl. Genet. 76:835 (1988); Zhang et al. Plant Cell Rep. 7:379, (1988); Battraw and Hall, Plant Sci. 86:191 202 (1992); Christou et al., Bio/Technology 9:957 (1991)); rye (De la Pena et al., Nature 325:274 (1987)); sugarcane (Bower and Birch, Plant J. 2:409 (1992)); tall fescue (Wang et al., Bio/Technology 10:691 (1992)), and wheat (Vasil et al., Bio/Technology 10:667 (1992); U.S. Patent No. 5,631,152).

There are a variety of methods for the regeneration of plants from plant tissue. The particular method of regeneration will depend on the starting plant tissue and the particular plant species to be regenerated.

The regeneration, development, and cultivation of plants from single plant protoplast transformants or from various transformed explants is well known in the art (Weissbach and Weissbach, In: Methods for Plant Molecular Biology, (Eds.), Academic Press, Inc. San Diego, CA, (1988)). This regeneration and growth process typically includes the steps of selection of transformed cells, culturing those individualized cells through the usual stages of embryonic development through the rooted plantlet stage. Transgenic embryos and seeds are similarly regenerated. The resulting transgenic rooted shoots are thereafter planted in an appropriate plant growth medium such as soil.

The development or regeneration of plants containing the foreign, exogenous isolated nucleic acid fragment that encodes a protein of interest is well known in the art. The regenerated plants may be self-pollinated to provide homozygous transgenic plants. Otherwise, pollen obtained from the regenerated plants is crossed to seed-grown plants of agronomically important lines. Conversely, pollen from plants of these important lines is used to pollinate regenerated plants. A transgenic plant of the present invention containing a desired polypeptide is cultivated using methods well known to one skilled in the art.

EXAMPLES

The present invention is further illustrated in the following Examples, in which parts and percentages are by weight and degrees are Celsius, unless otherwise stated. It should be understood that these Examples, while indicating embodiments of the invention, are given by way of illustration only. From the above discussion and these Examples, one skilled in the art can ascertain the essential characteristics of this invention, and without departing from the spirit and scope thereof, can make

various changes and modifications of the invention to adapt it to various usages and conditions. Thus, various modifications of the invention in addition to those shown and described herein will be apparent to those skilled in the art from the foregoing description. Such modifications are also intended to fall within the scope of the appended claims.

EXAMPLE 1

Creation of an *Arabidopsis* Population with Activation-Tagged Genes

An 18.5-kb T-DNA based binary construct was created, pHsbarENDs2 (FIG. 1; SEQ ID NO:31), that contains four multimerized enhancer elements derived from the Cauliflower Mosaic Virus 35S promoter (corresponding to sequences -341 to -64, as defined by Odell et al., *Nature* 313:810-812 (1985)). The construct also contains vector sequences (pUC9) and a polylinker to allow plasmid rescue, transposon sequences (Ds) to remobilize the T-DNA, and the bar gene to allow for glufosinate selection of transgenic plants. In principle, only the 10.8-kb segment from the right border (RB) to left border (LB) inclusive will be transferred into the host plant genome. Since the enhancer elements are located near the RB, they can induce cis-activation of genomic loci following T-DNA integration.

Arabidopsis activation-tagged populations were created by whole plant *Agrobacterium* transformation. The pHsbarENDs2 construct was transformed into *Agrobacterium tumefaciens* strain C58, grown in LB at 25 °C to OD600 ~1.0. Cells were then pelleted by centrifugation and resuspended in an equal volume of 5% sucrose/0.05% Silwet L-77 (OSI Specialties, Inc). At early bolting, soil grown *Arabidopsis thaliana* ecotype Col-0 were top watered with the *Agrobacterium* suspension. A week later, the same plants were top watered again with the same *Agrobacterium* strain in sucrose/Silwet. The plants were then allowed to set seed as normal. The resulting T1 seed were sown on soil, and transgenic seedlings were selected by spraying with glufosinate (Finale®; AgrEvo; Bayer Environmental Science). A total of 100,000 glufosinate resistant T1 seedlings were selected. T2 seed from each line was kept separate.

EXAMPLE 2

Screens to Identify Lines with Enhanced Drought Tolerance

Quantitative Drought Screen: From each of 96,000 separate T1 activation-tagged lines, nine glufosinate resistant T2 plants are sown, each in a single pot on

Scotts® Metro-Mix® 200 soil. Flats are configured with 8 square pots each. Each of the square pots is filled to the top with soil. Each pot (or cell) is sown to produce 9 glufosinate resistant seedlings in a 3x3 array.

5 The soil is watered to saturation and then plants are grown under standard conditions (i.e., 16 hour light, 8 hour dark cycle; 22°C; ~60% relative humidity). No additional water is given.

Digital images of the plants are taken at the onset of visible drought stress symptoms. Images are taken once a day (at the same time of day), until the plants appear dessicated. Typically, four consecutive days of data is captured.

10 Color analysis is employed for identifying potential drought tolerant lines. Color analysis can be used to measure the increase in the percentage of leaf area that falls into a yellow color bin. Using hue, saturation and intensity data ("HSI"), the yellow color bin consists of hues 35 to 45.

Maintenance of leaf area is also used as another criterion for identifying
15 potential drought tolerant lines, since *Arabidopsis* leaves wilt during drought stress. Maintenance of leaf area can be measured as reduction of rosette leaf area over time.

Leaf area is measured in terms of the number of green pixels obtained using the LemnaTec imaging system. Activation-tagged and control (e.g., wild-type)
20 plants are grown side by side in flats that contain 72 plants (9 plants/pot). When wilting begins, images are measured for a number of days to monitor the wilting process. From these data wilting profiles are determined based on the green pixel counts obtained over four consecutive days for activation-tagged and accompanying control plants. The profile is selected from a series of measurements over the four
25 day period that gives the largest degree of wilting. The ability to withstand drought is measured by the tendency of activation-tagged plants to resist wilting compared to control plants.

LemnaTec HTSBonitUV software is used to analyze CCD images. Estimates of the leaf area of the *Arabidopsis* plants are obtained in terms of the number of
30 green pixels. The data for each image is averaged to obtain estimates of mean and standard deviation for the green pixel counts for activation-tagged and wild-type plants. Parameters for a noise function are obtained by straight line regression of the squared deviation versus the mean pixel count using data for all images in a

batch. Error estimates for the mean pixel count data are calculated using the fit parameters for the noise function. The mean pixel counts for activation-tagged and wild-type plants are summed to obtain an assessment of the overall leaf area for each image. The four-day interval with maximal wilting is obtained by selecting the interval that corresponds to the maximum difference in plant growth. The individual wilting responses of the activation-tagged and wild-type plants are obtained by normalization of the data using the value of the green pixel count of the first day in the interval. The drought tolerance of the activation-tagged plant compared to the wild-type plant is scored by summing the weighted difference between the wilting response of activation-tagged plants and wild-type plants over day two to day four; the weights are estimated by propagating the error in the data. A positive drought tolerance score corresponds to an activation-tagged plant with slower wilting compared to the wild-type plant. Significance of the difference in wilting response between activation-tagged and wild-type plants is obtained from the weighted sum of the squared deviations.

Lines with a significant delay in yellow color accumulation and/or with significant maintenance of rosette leaf area, when compared to the average of the whole flat, are designated as Phase 1 hits. Phase 1 hits are re-screened in duplicate under the same assay conditions. When either or both of the Phase 2 replicates show a significant difference (score of greater than 0.9) from the whole flat mean, the line is then considered a validated drought tolerant line.

EXAMPLE 3

Identification of Activation-Tagged Genes

Genes flanking the T-DNA insert in drought tolerant lines are identified using one, or both, of the following two standard procedures: (1) thermal asymmetric interlaced (TAIL) PCR (Liu et al., (1995), *Plant J.* 8:457-63); and (2) SAIFF PCR (Siebert et al., (1995) *Nucleic Acids Res.* 23:1087-1088). In lines with complex multimerized T-DNA inserts, TAIL PCR and SAIFF PCR may both prove insufficient to identify candidate genes. In these cases, other procedures, including inverse PCR, plasmid rescue and/or genomic library construction, can be employed.

A successful result is one where a single TAIL or SAIFF PCR fragment contains a T-DNA border sequence and *Arabidopsis* genomic sequence.

Once a tag of genomic sequence flanking a T-DNA insert is obtained, candidate genes are identified by alignment to publicly available *Arabidopsis* genome sequence.

Specifically, the annotated gene nearest the 35S enhancer elements/T-DNA RB are candidates for genes that are activated.

To verify that an identified gene is truly near a T-DNA and to rule out the possibility that the TAIL/SAIFF fragment is a chimeric cloning artifact, a diagnostic PCR on genomic DNA is done with one oligo in the T-DNA and one oligo specific for the candidate gene. Genomic DNA samples that give a PCR product are interpreted as representing a T-DNA insertion. This analysis also verifies a situation in which more than one insertion event occurs in the same line, e.g., if multiple differing genomic fragments are identified in TAIL and/or SAIFF PCR analyses.

EXAMPLE 4A

Identification of Activation-Tagged

Ferrochelatase-I (FeC-I) Gene

An activation-tagged line (No. 105998) showing drought tolerance was further analyzed. DNA from the line was extracted, and genes flanking the T-DNA insert in the mutant line were identified using SAIFF PCR (Siebert et al., *Nucleic Acids Res.* 23:1087-1088 (1995)). A PCR amplified fragment was identified that contained T-DNA border sequence and *Arabidopsis* genomic sequence. Genomic sequences flanking the T-DNA insert was obtained, and the candidate gene was identified by alignment to the completed *Arabidopsis* genome. For a given T-DNA integration event, the annotated gene nearest the 35S enhancer elements/T-DNA RB was the candidate for gene that is activated in the line. In the case of line 105998, the gene nearest the 35S enhancers at the integration site was At5g26030 (SEQ ID NO:1; NCBI GI No. 511080), encoding a ferrochelatase-I protein (SEQ ID NO:2; NCBI GI No. 511081).

EXAMPLE 4B

Assay for Expression Level of Candidate Drought Tolerance Genes

A functional activation-tagged allele should result in either up-regulation of the candidate gene in tissues where it is normally expressed, ectopic expression in tissues that do not normally express that gene, or both.

Expression levels of the candidate genes in the cognate mutant line vs. wild-type are compared. A standard RT-PCR procedure, such as the QuantiTect® Reverse Transcription Kit from Qiagen®, is used. RT-PCR of the actin gene is used as a control to show that the amplification and loading of samples from the mutant line and wild-type are similar.

Assay conditions are optimized for each gene. Expression levels are checked in mature rosette leaves. If the activation-tagged allele results in ectopic expression in other tissues (e.g., roots), it is not detected by this assay. As such, a positive result is useful but a negative result does not eliminate a gene from further analysis.

EXAMPLE 5A

Validation of *Arabidopsis* Candidate Gene At5g26030 (Ferrochelatase-I) via Transformation into *Arabidopsis*

Candidate genes can be transformed into *Arabidopsis* and overexpressed under the 35S promoter. If the same or similar phenotype is observed in the transgenic line as in the parent activation-tagged line, then the candidate gene is considered to be a validated “lead gene” in *Arabidopsis*.

The candidate *Arabidopsis* ferrochelatase-I gene (At5g26030; SEQ ID NO:1; NCBI GI No. 511080) was tested for its ability to confer drought tolerance in the following manner.

A 16.8-kb T-DNA based binary vector, called pBC-yellow (SEQ ID NO:34; FIG. 4), was constructed with a 1.3-kb 35S promoter immediately upstream of the INVITROGEN™ GATEWAY® C1 conversion insert. The vector also contains the RD29a promoter driving expression of the gene for ZS-Yellow (INVITROGEN™), which confers yellow fluorescence to transformed seed.

The At5g26030 cDNA protein-coding region was amplified by RT-PCR with the following primers:

(1) At5g26030-5'attB forward primer (SEQ ID NO:42):

TTAAACAAGTTTGTACAAAAAGCAGGCTCAACAATGCAGGCAACGG
CTTTATCA

(2) At5g26030-3'attB reverse primer (SEQ ID NO:43):

TTAAACCACTTTGTACAAGAAAGCTGGGTCTATAGGTTCCGGAACGC
ATG

The forward primer contains the attB1 sequence (ACAAGTTTGTACAAAAAAGCAGGCT; SEQ ID NO:40) and a consensus Kozak sequence (CAACA) adjacent to the first 21 nucleotides of the protein-coding region, beginning with the ATG start codon, of said cDNA.

5 The reverse primer contains the attB2 sequence (ACCACTTTGTACAAGAAAGCTGGGT; SEQ ID NO:41) adjacent to the reverse complement of the last 21 nucleotides of the protein-coding region, beginning with the reverse complement of the stop codon, of said cDNA.

Using the INVITROGEN™ GATEWAY® CLONASE™ technology, a BP
10 Recombination Reaction was performed with pDONR™/Zeo (SEQ ID NO:32; FIG. 2). This process removed the bacteria lethal ccdB gene, as well as the chloramphenicol resistance gene (CAM) from pDONR™/Zeo and directionally cloned the PCR product with flanking attB1 and attB2 sites creating an entry clone, PHP31052. This entry clone was used for a subsequent LR Recombination
15 Reaction with a destination vector, as follows.

A 16.8-kb T-DNA based binary vector (destination vector), called pBC-yellow (SEQ ID NO:34; FIG. 4), was constructed with a 1.3-kb 35S promoter immediately upstream of the INVITROGEN™ GATEWAY® C1 conversion insert, which contains the bacterial lethal ccdB gene as well as the chloramphenicol resistance gene
20 (CAM) flanked by attR1 and attR2 sequences. The vector also contains the RD29a promoter driving expression of the gene for ZS-Yellow (INVITROGEN™), which confers yellow fluorescence to transformed seed. Using the INVITROGEN™ GATEWAY® technology, an LR Recombination Reaction was performed on the PHP31052 entry clone, containing the directionally cloned PCR product, and pBC-
25 yellow. This allowed for rapid and directional cloning of the candidate gene behind the 35S promoter in pBC-yellow to create the 35S promoter::At5g26030 expression construct, pBC-Yellow-At5g26030.

Applicants then introduced the 35S promoter::At5g26030 expression construct into wild-type *Arabidopsis* ecotype Col-0, using the same *Agrobacterium*-
30 mediated transformation procedure described in Example 1. Transgenic T1 seeds were selected by yellow fluorescence, and T1 seeds were plated next to wild-type seeds and grown under water limiting conditions. Growth conditions and imaging analysis were as described in Example 2. It was found that the original drought

tolerance phenotype from activation tagging could be recapitulated in wild-type *Arabidopsis* plants that were transformed with a construct where At5g26030 was directly expressed by the 35S promoter. The drought tolerance score, as determined by the method of Example 2, was 5.3.

5

EXAMPLE 5B

Validation of *Arabidopsis* Candidate Gene At2g30390 (Ferrochelatase-II) via Transformation into *Arabidopsis*

A cDNA encoding ferrochelatase-I from *Arabidopsis* was previously cloned by functional complementation of a yeast mutant (Smith et al. 1994 J Biol Chem
10 269:13405-13413). Subsequently, a second ferrochelatase isoform, ferrochelatase-II, was found in *Arabidopsis* (Chow et al. 1998 Plant J 15:531-541). Two forms of ferrochelatase also have been identified in cucumber, *Cucumis sativus* (Miyamoto et al. 1994 Plant Physiol 105:769-770; and Suzuki et al. 2002 J Biol Chem 277:4731-4737). The C-terminal region of ferrochelatase-II, but not ferrochelatase-I, contains
15 a conserved motif found in light-harvesting chlorophyll proteins (Suzuki et al. 2002 J Biol Chem 277:4731-4737). The C-terminal region of the ferrochelatase from the cyanobacteria *Synechocystis* (NCBI GI NO. 1708186) has sequence homology to this conserved motif.

The *Arabidopsis* ferrochelatase-II gene (At2g30390) was selected as a
20 second candidate gene, and was tested for its ability to confer drought tolerance in the following manner.

A 16.8-kb T-DNA based binary vector, called pBC-yellow (SEQ ID NO:34; FIG. 4), was constructed with a 1.3-kb 35S promoter immediately upstream of the INVITROGEN™ GATEWAY® C1 conversion insert. The vector also contains the
25 RD29a promoter driving expression of the gene for ZS-Yellow (INVITROGEN™), which confers yellow fluorescence to transformed seed.

The At2g30390 cDNA protein-coding region was amplified by RT-PCR with the following primers:

(3) At2g30390-5'attB forward primer (SEQ ID NO:44):

30 TTAACAAGTTTGTACAAAAAAGCAGGCTCAACAATGAATTGCCAG
CCATGACT

(4) At2g30390-3'attB reverse primer (SEQ ID NO:45):

TTAAACCACTTTGTACAAGAAAGCTGGGTTTATAATGAAGGCAAGAT
GCC

The forward primer contains the attB1 sequence

5 (ACAAGTTTGTACAAAAAAGCAGGCT; SEQ ID NO:40) and a consensus Kozak sequence (CAACA) adjacent to the first 21 nucleotides of the protein-coding region, beginning with the ATG start codon, of said cDNA.

The reverse primer contains the attB2 sequence

10 (ACCACTTTGTACAAGAAAGCTGGGT; SEQ ID NO:41) adjacent to the reverse complement of the last 21 nucleotides of the protein-coding region, beginning with the reverse complement of the stop codon, of said cDNA.

Using the INVITROGEN™ GATEWAY® CLONASE™ technology, a BP Recombination Reaction was performed with pDONR™/Zeo (SEQ ID NO:32; FIG. 2). This process removed the bacteria lethal ccdB gene, as well as the
15 chloramphenicol resistance gene (CAM) from pDONR™/Zeo and directionally cloned the PCR product with flanking attB1 and attB2 sites creating an entry clone, PHP-Entry-At2g30390. This entry clone was used for a subsequent LR Recombination Reaction with a destination vector, as follows.

A 16.8-kb T-DNA based binary vector (destination vector), called pBC-yellow
20 (SEQ ID NO:34; FIG. 4), was constructed with a 1.3-kb 35S promoter immediately upstream of the INVITROGEN™ GATEWAY® C1 conversion insert, which contains the bacterial lethal ccdB gene as well as the chloramphenicol resistance gene (CAM) flanked by attR1 and attR2 sequences. The vector also contains the RD29a promoter driving expression of the gene for ZS-Yellow (INVITROGEN™), which
25 confers yellow fluorescence to transformed seed. Using the INVITROGEN™ GATEWAY® technology, an LR Recombination Reaction was performed on the PHP-Entry-At2g30390 entry clone, containing the directionally cloned PCR product, and pBC-yellow. This allowed for rapid and directional cloning of the candidate gene behind the 35S promoter in pBC-yellow to create the 35S
30 promoter::At2g30390 expression construct, pBC-Yellow-At2g30390.

Applicants then introduced the 35S promoter::At2g30390 expression construct into wild-type *Arabidopsis* ecotype Col-0, using the same *Agrobacterium*-mediated transformation procedure described in Example 1. Transgenic T1 seeds

were selected by yellow fluorescence, and T1 seeds were plated next to wild-type seeds and grown under water limiting conditions. Growth conditions and imaging analysis were as described in Example 2. It was found that the original drought tolerance phenotype from activation tagging could be recapitulated in wild-type *Arabidopsis* plants that were transformed with a construct where At2g30390 was directly expressed by the 35S promoter. The drought tolerance score, as determined by the method of Example 2, was 3.7.

EXAMPLE 6A

Preparation of cDNA Libraries and

Isolation and Sequencing of cDNA Clones

cDNA libraries may be prepared by any one of many methods available. For example, the cDNAs may be introduced into plasmid vectors by first preparing the cDNA libraries in UNI-ZAP™ XR vectors according to the manufacturer's protocol (Stratagene Cloning Systems, La Jolla, CA). The UNI-ZAP™ XR libraries are converted into plasmid libraries according to the protocol provided by Stratagene. Upon conversion, cDNA inserts will be contained in the plasmid vector pBluescript®. In addition, the cDNAs may be introduced directly into pre-cut Bluescript® II SK(+) vectors (Stratagene) using T4 DNA ligase (New England Biolabs), followed by transfection into DH10B cells according to the manufacturer's protocol (GIBCO BRL Products). Once the cDNA inserts are in plasmid vectors, plasmid DNAs are prepared from randomly picked bacterial colonies containing recombinant pBluescript® plasmids, or the insert cDNA sequences are amplified via polymerase chain reaction using primers specific for vector sequences flanking the inserted cDNA sequences. Amplified insert DNAs or plasmid DNAs are sequenced in dye-primer sequencing reactions to generate partial cDNA sequences (expressed sequence tags or "ESTs"; see Adams et al., (1991) *Science* 252:1651-1656). The resulting ESTs are analyzed using a Perkin Elmer Model 377 fluorescent sequencer.

Full-insert sequence (FIS) data is generated utilizing a modified transposition protocol. Clones identified for FIS are recovered from archived glycerol stocks as single colonies, and plasmid DNAs are isolated via alkaline lysis. Isolated DNA templates are reacted with vector primed M13 forward and reverse oligonucleotides in a PCR-based sequencing reaction and loaded onto automated sequencers.

Confirmation of clone identification is performed by sequence alignment to the original EST sequence from which the FIS request is made.

Confirmed templates are transposed via the Primer Island transposition kit (PE Applied Biosystems, Foster City, CA) which is based upon the *Saccharomyces cerevisiae* Ty1 transposable element (Devine and Boeke (1994) *Nucleic Acids Res.* 22:3765-3772). The *in vitro* transposition system places unique binding sites randomly throughout a population of large DNA molecules. The transposed DNA is then used to transform DH10B electro-competent cells (Gibco BRL/Life Technologies, Rockville, MD) via electroporation. The transposable element contains an additional selectable marker (named DHFR; Fling and Richards (1983) *Nucleic Acids Res.* 11:5147-5158), allowing for dual selection on agar plates of only those subclones containing the integrated transposon. Multiple subclones are randomly selected from each transposition reaction, plasmid DNAs are prepared via alkaline lysis, and templates are sequenced (ABI Prism® dye-terminator ReadyReaction mix) outward from the transposition event site, utilizing unique primers specific to the binding sites within the transposon.

Sequence data is collected (ABI Prism® Collections) and assembled using Phred and Phrap (Ewing et al. (1998) *Genome Res.* 8:175-185; Ewing and Green (1998) *Genome Res.* 8:186-194). Phred is a public domain software program which re-reads the ABI sequence data, re-calls the bases, assigns quality values, and writes the base calls and quality values into editable output files. The Phrap sequence assembly program uses these quality values to increase the accuracy of the assembled sequence contigs. Assemblies are viewed by the Consed sequence editor (Gordon et al. (1998) *Genome Res.* 8:195-202).

In some of the clones the cDNA fragment may correspond to a portion of the 3'-terminus of the gene and does not cover the entire open reading frame. In order to obtain the upstream information one of two different protocols is used. The first of these methods results in the production of a fragment of DNA containing a portion of the desired gene sequence while the second method results in the production of a fragment containing the entire open reading frame. Both of these methods use two rounds of PCR amplification to obtain fragments from one or more libraries. The libraries some times are chosen based on previous knowledge that the specific gene should be found in a certain tissue and some times are randomly-chosen.

Reactions to obtain the same gene may be performed on several libraries in parallel or on a pool of libraries. Library pools are normally prepared using from 3 to 5 different libraries and normalized to a uniform dilution. In the first round of amplification both methods use a vector-specific (forward) primer corresponding to a portion of the vector located at the 5'-terminus of the clone coupled with a gene-specific (reverse) primer. The first method uses a sequence that is complementary to a portion of the already known gene sequence while the second method uses a gene-specific primer complementary to a portion of the 3'-untranslated region (also referred to as UTR). In the second round of amplification a nested set of primers is used for both methods. The resulting DNA fragment is ligated into a pBluescript® vector using a commercial kit and following the manufacturer's protocol. This kit is selected from many available from several vendors including INVITROGEN™ (Carlsbad, CA), Promega Biotech (Madison, WI), and Gibco-BRL (Gaithersburg, MD). The plasmid DNA is isolated by alkaline lysis method and submitted for sequencing and assembly using Phred/Phrap, as above.

EXAMPLE 7

Identification of cDNA Clones

cDNA clones encoding ferrochelatases can be identified by conducting BLAST (Basic Local Alignment Search Tool; Altschul et al. (1993) *J. Mol. Biol.* 215:403-410; see also the explanation of the BLAST algorithm on the world wide web site for the National Center for Biotechnology Information at the National Library of Medicine of the National Institutes of Health) searches for similarity to amino acid sequences contained in the BLAST "nr" database (comprising all non-redundant GenBank CDS translations, sequences derived from the 3-dimensional structure Brookhaven Protein Data Bank, the last major release of the SWISS-PROT protein sequence database, EMBL, and DDBJ databases). The DNA sequences from clones can be translated in all reading frames and compared for similarity to all publicly available protein sequences contained in the "nr" database using the BLASTX algorithm (Gish and States (1993) *Nat. Genet.* 3:266-272) provided by the NCBI. The polypeptides encoded by the cDNA sequences can be analyzed for similarity to all publicly available amino acid sequences contained in the "nr" database using the BLASTP algorithm provided by the National Center for Biotechnology Information (NCBI). For convenience, the

P-value (probability) or the E-value (expectation) of observing a match of a cDNA-encoded sequence to a sequence contained in the searched databases merely by chance as calculated by BLAST are reported herein as "pLog" values, which represent the negative of the logarithm of the reported P-value or E-value.

- 5 Accordingly, the greater the pLog value, the greater the likelihood that the cDNA-encoded sequence and the BLAST "hit" represent homologous proteins.

ESTs sequences can be compared to the Genbank database as described above. ESTs that contain sequences more 5- or 3-prime can be found by using the BLASTn algorithm (Altschul et al (1997) *Nucleic Acids Res.* 25:3389-3402.) against
 10 the Du Pont proprietary database comparing nucleotide sequences that share common or overlapping regions of sequence homology. Where common or overlapping sequences exist between two or more nucleic acid fragments, the sequences can be assembled into a single contiguous nucleotide sequence, thus extending the original fragment in either the 5 or 3 prime direction. Once the most
 15 5-prime EST is identified, its complete sequence can be determined by Full Insert Sequencing as described above. Homologous genes belonging to different species can be found by comparing the amino acid sequence of a known gene (from either a proprietary source or a public database) against an EST database using the tBLASTn algorithm. The tBLASTn algorithm searches an amino acid query against
 20 a nucleotide database that is translated in all 6 reading frames. This search allows for differences in nucleotide codon usage between different species, and for codon degeneracy.

EXAMPLE 8

Characterization of cDNA Clones Encoding Ferrochetalases

25 cDNA libraries representing mRNAs from various tissues of Sugar Beet, Canola, Maize, Rice, Soybean, Wheat and Catmint were prepared and cDNA clones encoding ferrochelatasases were identified. The characteristics of the libraries are described below.

TABLE 3

30 cDNA Libraries from Sugar Beet, Canola, Maize, Rice, Soybean, Wheat and Catmint

Library	Description	Clone
ebs1c	Sugar Beet, shoot and phloem specific genes	ebs1c.pk002.n16:fis

ebb1c	Immature buds of Canola, Rf gene knock out mutant line, 02SM2	ebb1c.pk006.j11: fis
cfp3n	Maize Ear, pooled V10-V14-v16-VT, Full-length enriched normalized	cfp3n.pk004.f12: fis
cfp5n	Maize Kernel, pooled stages, Full-length enriched, normalized	cfp5n.pk009.j16: fis
rl0n	Rice 15 Day Old Leaf*	rl0n.pk117.h21: fis
se3	Soybean Embryo, 17 Days After Flowering	se3.pk0034.e10: fis
etb1n	Normalized tulip bulb library	etb1n.pk002.n16: fis
wlp1c	Wheat (Triticum aestivum, Hi Line) lemma and palea	wlp1c.pk002.p10: fis
ecl1c	Catmint (Nepeta racemosa) cDNA library from young leaf tissue	ecl1c.pk005.l15: fis
cfp6n	Maize Leaf and Seed pooled, Full-length enriched normalized	cfp6n.pk072.n9: fis
wlp1c	Wheat (Triticum aestivum, Hi Line) lemma and palea	wlp1c.pk004.d13: fis
wpa1c	Wheat (Triticum aestivum) pre-meiotic anthers	wpa1c.pk014.g4: fis

*These libraries were normalized essentially as described in U.S. Pat. No. 5,482,845

Two of the cDNA clones listed in Table 3 had anomalies. The clone ebb1c.pk006.j11: fis contains an unspliced intron at nucleotides 386-481 of SEQ ID NO:5. The polypeptide encoded by the two exons surrounding the intron is given in SEQ ID NO:6. Additionally, the clone ecl1c.pk005.l15: fis has a single base deletion at nucleotide 442, which results in a frame-shift in the corresponding translation. Based on comparison to a highly similar clone cDNA clone from grape, veb1c.pk008.p6 (411 base pairs), a "T" nucleotide was added to the ecl1c.pk005.l15: fis sequence at position 442. The modified nucleotide sequence is presented in SEQ ID NO:23, which encodes the amino acid sequence of SEQ ID NO:24.

The BLAST search using the sequences from clones listed in Table 3 revealed similarity of the polypeptides encoded by the cDNAs to the ferrochelatases from various organisms. As shown in Table 4 and Figures 10A-10C, certain cDNAs encoded polypeptides similar to ferrochelatase-I proteins from *Arabidopsis* (GI No. 511081; SEQ ID NO:2; FeC-I), barley (GI No. 2460251; SEQ ID NO:48; FeC-I), rice (GI No. 113631036; SEQ ID NO:49; FeC-I), and grape (GI No. 147818793; SEQ ID NO:50; FeC-I). As shown in Table 6 and Figures 12A-12C, certain cDNAs encoded

polypeptides similar to ferrochelatase-II proteins from *Arabidopsis* (GI No. 15227742; SEQ ID NO:20; FeC-II), rice (GI No. 115463419; SEQ ID NO:51; FeC-II), cucumber (GI No. 12082085; SEQ ID NO:52; FeC-II), tobacco (GI No. 15147828; SEQ ID NO:53; FeC-II) and a ferrochelatase-II-like protein from

5 *Synechocystis* (GI No. 1708186; SEQ ID NO:54; FeC-II-like).

Shown in Tables 4 and 6 (non-patent literature) and Tables 5 and 7 (patent literature) are the BLASTP results for the amino acid sequences derived from the nucleotide sequences of the entire cDNA inserts ("Full-Insert Sequence" or "FIS") of the clones listed in Table 3. Each cDNA insert encodes an entire or functional

10 protein ("Complete Gene Sequence" or "CGS"). Also shown in Tables 4-7 are the percent sequence identity values for each pair of amino acid sequences:

TABLE 4
BLASTP Results for Ferrochelatase-I Polypeptides

Sequence (SEQ ID NO)	Status	NCBI GI No. (SEQ ID NO)	BLASTP pLog of E-value	Percent Sequence Identity
ebs1c.pk002.n16 (FIS) (SEQ ID NO:4)	CGS	147818793 (SEQ ID NO:50)	>180	72.6
ebb1c.pk006.j11 (FIS) (SEQ ID NO:6)	CGS	511081 (SEQ ID NO:2)	>180	83.3
cfp3n.pk004.f12 (FIS) (SEQ ID NO:8)	CGS	113631036 (SEQ ID NO:49)	>180	84.0
cfp5n.pk009.j16 (FIS) (SEQ ID NO:10)	CGS	113631036 (SEQ ID NO:49)	>180	81.7
rl0n.pk117.h21 (FIS) (SEQ ID NO:12)	CGS	113631036 (SEQ ID NO:49)	>180	100
se3.pk0034.e10 (FIS) (SEQ ID NO:14)	CGS	147818793 (SEQ ID NO:50)	>180	74.2
etb1n.pk002.n16 (FIS) (SEQ ID NO:16)	CGS	113631036 (SEQ ID NO:49)	180	69.5
wlp1c.pk002.p10 (FIS) (SEQ ID NO:18)	CGS	2460251 (SEQ ID NO:48)	>180	98.3

TABLE 5**BLASTP Results for Ferrochelatase-I Polypeptides**

Sequence (SEQ ID NO)	Status	Reference (SEQ ID NO)	BLASTP pLog of E-value	Percent Sequence Identity
ebs1c.pk002.n16 (FIS) (SEQ ID NO:4)	CGS	SEQ ID NO:240025 of US2004031072 (SEQ ID NO:57)	>180	68.8
ebb1c.pk006.j11 (FIS) (SEQ ID NO:6)	CGS	SEQ ID NO:20 of JP2001190168-A (SEQ ID NO:58)	>180	83.3
cfp3n.pk004.f12 (FIS) (SEQ ID NO:8)	CGS	SEQ ID NO:13029 of US2006150283 (SEQ ID NO:59)	>180	99.4
cfp5n.pk009.j16 (FIS) (SEQ ID NO:10)	CGS	SEQ ID NO:7745 of US2006150283 (SEQ ID NO:60)	>180	100
rl0n.pk117.h21 (FIS) (SEQ ID NO:12)	CGS	SEQ ID NO:46156 of JP2005185101 (SEQ ID NO:61)	>180	100
se3.pk0034.e10 (FIS) (SEQ ID NO:14)	CGS	SEQ ID NO:240025 of US2004031072 (SEQ ID NO:57)	>180	100
etb1n.pk002.n16 (FIS) (SEQ ID NO:16)	CGS	SEQ ID NO:46156 of JP2005185101 (SEQ ID NO:61)	>180	69.5
wlp1c.pk002.p10 (FIS) (SEQ ID NO:18)	CGS	SEQ ID NO:46156 of JP2005185101 (SEQ ID NO:61)	>180	84.6

TABLE 6**BLASTP Results for Ferrochelatase-II Polypeptides**

Sequence (SEQ ID NO)	Status	NCBI GI No. (SEQ ID NO)	BLASTP pLog of E- value	Percent Sequence Identity
ec1c.pk005.l15 (FIS) (SEQ ID NO:24)	CGS	12082085 (SEQ ID NO:52)	>180	78.6
cfp6n.pk072.n9 (FIS) (SEQ ID NO:26)	CGS	115463419 (SEQ ID NO:51)	>180	90.5
wlp1c.pk004.d13 (FIS) (SEQ ID NO:28)	CGS	115463419 (SEQ ID NO:51)	>180	82.9
wpa1c.pk014.g4 (FIS) (SEQ ID NO:30)	CGS	115463419 (SEQ ID NO:51)	>180	83.3

TABLE 7**BLASTP Results for Ferrochelatase-II Polypeptides**

Sequence (SEQ ID NO)	Status	Reference (SEQ ID NO)	BLASTP pLog of E-value	Percent Sequence Identity
ec1c.pk005.l15 (FIS) (SEQ ID NO:24)	CGS	SEQ ID NO:52154 of JP2005185101 (SEQ ID NO:62)	>180	77.4
cfp6n.pk072.n9 (FIS) (SEQ ID NO:26)	CGS	SEQ ID NO:72746 of US2004034888 (SEQ ID NO:63)	>180	99.6
wlp1c.pk004.d13 (FIS) (SEQ ID NO:28)	CGS	SEQ ID NO:52154 of JP2005185101 (SEQ ID NO:62)	>180	82.9
wpa1c.pk014.g4 (FIS) (SEQ ID NO:30)	CGS	SEQ ID NO:52154 of JP2005185101 (SEQ ID NO:62)	>180	83.3

Figures 10A-10C present an alignment of the amino acid sequences of ferrochelatase-I proteins set forth in SEQ ID NOs:2, 4, 6, 8, 10, 12, 14, 16, 18, 48, 49 and 50. Figure 11 presents the percent sequence identities and divergence values for each sequence pair presented in Figures 10A and 10B.

Figures 12A-12C present an alignment of the amino acid sequences of ferrochelatase-II proteins set forth in SEQ ID NOs:20, 22, 24, 26, 28, 30, 51, 52, 53 and 54. Figure 13 presents the percent sequence identities and divergence values for each sequence pair presented in Figures 12A and 12B.

Sequence alignments and percent identity calculations were performed using the Megalign® program of the LASERGENE® bioinformatics computing suite (DNASTAR® Inc., Madison, WI). Multiple alignment of the sequences was performed using the Clustal V method of alignment (Higgins and Sharp (1989) *CABIOS*. 5:151-153) with the default parameters (GAP PENALTY=10, GAP LENGTH PENALTY=10). Default parameters for pairwise alignments using the Clustal method were KTUPLE=1, GAP PENALTY=3, WINDOW=5 and DIAGONALS SAVED=5.

Import experiments into isolated chloroplasts and mitochondria showed that the *Arabidopsis* ferrochelatase-II gene encodes a precursor which is imported solely into the chloroplast, in contrast to *Arabidopsis* ferrochelatase-I which is targeted to both organelles (Chow et al. 1997 *J Biol Chem* 272:27565-27571; Chow et al. 1998 *Plant J* 15:531-541). The ferrochelatase from *Synechocystis* (GI No. 1708186; SEQ ID NO:54) is more similar to the ferrochelatase-II isoform.

Sequence alignments and BLAST scores and probabilities indicate that the nucleic acid fragments comprising the instant cDNA clones encode ferrochelatases.

EXAMPLE 9

Preparation of a Plant Expression Vector

Containing a Homolog to the *Arabidopsis* Lead Gene

Sequences homologous to the *Arabidopsis* ferrochelatase-I polypeptide can be identified using sequence comparison algorithms such as BLAST (Basic Local Alignment Search Tool; Altschul et al., *J. Mol. Biol.* 215:403-410 (1993); see also the explanation of the BLAST algorithm on the world wide web site for the National Center for Biotechnology Information at the National Library of Medicine of the National Institutes of Health). Sequences encoding homologous ferrochelatases can be PCR-amplified by either of the following methods.

Method 1 (RNA-based): If the 5' and 3' sequence information for the protein-coding region of a gene encoding a ferrochelatase homolog is available, gene-specific primers can be designed as outlined in Example 5. RT-PCR can be used

with plant RNA to obtain a nucleic acid fragment containing the protein-coding region flanked by attB1 (SEQ ID NO:40) and attB2 (SEQ ID NO:41) sequences. The primer may contain a consensus Kozak sequence (CAACA) upstream of the start codon.

5 Method 2 (DNA-based): Alternatively, if a cDNA clone is available for a gene encoding a ferrochelatase homolog, the entire cDNA insert (containing 5' and 3' non-coding regions) can be PCR amplified. Forward and reverse primers can be designed that contain either the attB1 sequence and vector-specific sequence that precedes the cDNA insert or the attB2 sequence and vector-specific sequence that
10 follows the cDNA insert, respectively. For a cDNA insert cloned into the vector pBulescript SK+, the forward primer VC062 (SEQ ID NO:46) and the reverse primer VC063 (SEQ ID NO:47) can be used.

Methods 1 and 2 can be modified according to procedures known by one skilled in the art. For example, the primers of Method 1 may contain restriction sites
15 instead of attB1 and attB2 sites, for subsequent cloning of the PCR product into a vector containing attB1 and attB2 sites. Additionally, Method 2 can involve amplification from a cDNA clone, a lambda clone, a BAC clone or genomic DNA.

A PCR product obtained by either method above can be combined with the GATEWAY® donor vector, such as pDONR™/Zeo (INVITROGEN™; FIG. 2; SEQ
20 ID NO:32) or pDONR™221 (INVITROGEN™; FIG. 3; SEQ ID NO:33), using a BP Recombination Reaction. This process removes the bacteria lethal ccdB gene, as well as the chloramphenicol resistance gene (CAM) from pDONR™221 and directionally clones the PCR product with flanking attB1 and attB2 sites to create an entry clone. Using the INVITROGEN™ GATEWAY® CLONASE™ technology, the
25 sequence encoding the homologous ferrochelatase from the entry clone can then be transferred to a suitable destination vector, such as pBC-Yellow (FIG. 4; SEQ ID NO:34), PHP27840 (FIG. 5; SEQ ID NO:35) or PHP23236 (FIG. 6; SEQ ID NO:36), to obtain a plant expression vector for use with *Arabidopsis*, soybean and corn, respectively.

30 The attP1 and attP2 sites of donor vectors pDONR™/Zeo or pDONR™221 are shown in Figures 2 and 3, respectively. The attR1 and attR2 sites of destination vectors pBC-Yellow, PHP27840 and PHP23236 are shown in Figures 4, 5 and 6, respectively.

Alternatively a MultiSite GATEWAY® LR recombination reaction between multiple entry clones and a suitable destination vector can be performed to create an expression vector.

EXAMPLE 10

Preparation of Soybean Expression Vectors and Transformation of Soybean with Validated *Arabidopsis* Lead Genes

Soybean plants can be transformed to overexpress a validated *Arabidopsis* lead gene or the corresponding homologs from various species in order to examine the resulting phenotype.

The same GATEWAY® entry clone described in Example 5 can be used to directionally clone each gene into the PHP27840 vector (SEQ ID NO:35; FIG. 5) such that expression of the gene is under control of the SCP1 promoter.

Soybean embryos may then be transformed with the expression vector comprising sequences encoding the instant polypeptides.

To induce somatic embryos, cotyledons, 3-5 mm in length dissected from surface sterilized, immature seeds of the soybean cultivar A2872, can be cultured in the light or dark at 26 °C on an appropriate agar medium for 6-10 weeks. Somatic embryos, which produce secondary embryos, are then excised and placed into a suitable liquid medium. After repeated selection for clusters of somatic embryos which multiply as early, globular staged embryos, the suspensions are maintained as described below.

Soybean embryogenic suspension cultures can be maintained in 35 mL liquid media on a rotary shaker, 150 rpm, at 26 °C with florescent lights on a 16:8 hour day/night schedule. Cultures are subcultured every two weeks by inoculating approximately 35 mg of tissue into 35 mL of liquid medium.

Soybean embryogenic suspension cultures may then be transformed by the method of particle gun bombardment (Klein et al. (1987) *Nature* (London) 327:70-73, U.S. Patent No. 4,945,050). A DUPONT™ BIOLISTIC™ PDS1000/HE instrument (helium retrofit) can be used for these transformations.

A selectable marker gene which can be used to facilitate soybean transformation is a chimeric gene composed of the 35S promoter from cauliflower mosaic virus (Odell et al. (1985) *Nature* 313:810-812), the hygromycin phosphotransferase gene from plasmid pJR225 (from *E. coli*; Gritz et al. (1983)

Gene 25:179-188) and the 3' region of the nopaline synthase gene from the T-DNA of the Ti plasmid of *Agrobacterium tumefaciens*. Another selectable marker gene which can be used to facilitate soybean transformation is an herbicide-resistant acetolactate synthase (ALS) gene from soybean or *Arabidopsis*. ALS is the first common enzyme in the biosynthesis of the branched-chain amino acids valine, leucine and isoleucine. Mutations in ALS have been identified that convey resistance to some or all of three classes of inhibitors of ALS (US Patent No. 5,013,659; the entire contents of which are herein incorporated by reference). Expression of the herbicide-resistant ALS gene can be under the control of a SAM synthetase promoter (U.S. Patent Application No. US-2003-0226166-A1; the entire contents of which are herein incorporated by reference).

To 50 μ L of a 60 mg/mL 1 μ m gold particle suspension is added (in order): 5 μ L DNA (1 μ g/ μ L), 20 μ L spermidine (0.1 M), and 50 μ L CaCl_2 (2.5 M). The particle preparation is then agitated for three minutes, spun in a microfuge for 10 seconds and the supernatant removed. The DNA-coated particles are then washed once in 400 μ L 70% ethanol and resuspended in 40 μ L of anhydrous ethanol. The DNA/particle suspension can be sonicated three times for one second each. Five μ L of the DNA-coated gold particles are then loaded on each macro carrier disk.

Approximately 300-400 mg of a two-week-old suspension culture is placed in an empty 60x15 mm petri dish and the residual liquid removed from the tissue with a pipette. For each transformation experiment, approximately 5-10 plates of tissue are normally bombarded. Membrane rupture pressure is set at 1100 psi and the chamber is evacuated to a vacuum of 28 inches mercury. The tissue is placed approximately 3.5 inches away from the retaining screen and bombarded three times. Following bombardment, the tissue can be divided in half and placed back into liquid and cultured as described above.

Five to seven days post bombardment, the liquid media may be exchanged with fresh media, and eleven to twelve days post bombardment with fresh media containing 50 mg/mL hygromycin. This selective media can be refreshed weekly. Seven to eight weeks post bombardment, green, transformed tissue may be observed growing from untransformed, necrotic embryogenic clusters. Isolated green tissue is removed and inoculated into individual flasks to generate new,

clonally propagated, transformed embryogenic suspension cultures. Each new line may be treated as an independent transformation event. These suspensions can then be subcultured and maintained as clusters of immature embryos or regenerated into whole plants by maturation and germination of individual somatic embryos.

T1 plants can be subjected to a soil-based drought stress. Using image analysis, plant area, volume, growth rate and color analysis can be taken at multiple times before and during drought stress. Overexpression constructs that result in a significant delay in wilting or leaf area reduction, yellow color accumulation and/or increased growth rate during drought stress will be considered evidence that the *Arabidopsis* gene functions in soybean to enhance drought tolerance.

Soybean plants transformed with validated genes can then be assayed under more vigorous field-based studies to study yield enhancement and/or stability under well-watered and water-limiting conditions.

EXAMPLE 11

Transformation of Maize with Validated *Arabidopsis* Lead Genes Using Particle Bombardment

Maize plants can be transformed to overexpress a validated *Arabidopsis* lead gene or the corresponding homologs from various species in order to examine the resulting phenotype.

The same GATEWAY® entry clone described in Example 5 can be used to directionally clone each gene into a maize transformation vector. Expression of the gene in the maize transformation vector can be under control of a constitutive promoter such as the maize ubiquitin promoter (Christensen et al., (1989) *Plant Mol. Biol.* 12:619-632 and Christensen et al., (1992) *Plant Mol. Biol.* 18:675-689)

The recombinant DNA construct described above can then be introduced into corn cells by the following procedure. Immature corn embryos can be dissected from developing caryopses derived from crosses of the inbred corn lines H99 and LH132. The embryos are isolated 10 to 11 days after pollination when they are 1.0 to 1.5 mm long. The embryos are then placed with the axis-side facing down and in contact with agarose-solidified N6 medium (Chu et al. (1975) *Sci. Sin. Peking* 18:659-668). The embryos are kept in the dark at 27°C. Friable embryogenic callus consisting of undifferentiated masses of cells with somatic proembryoids and

embryoids borne on suspensor structures proliferates from the scutellum of these immature embryos. The embryogenic callus isolated from the primary explant can be cultured on N6 medium and sub-cultured on this medium every 2 to 3 weeks.

The plasmid, p35S/Ac (obtained from Dr. Peter Eckes, Hoechst Ag, Frankfurt, Germany) may be used in transformation experiments in order to provide for a selectable marker. This plasmid contains the *Pat* gene (see European Patent Publication 0 242 236) which encodes phosphinothricin acetyl transferase (PAT). The enzyme PAT confers resistance to herbicidal glutamine synthetase inhibitors such as phosphinothricin. The *pat* gene in p35S/Ac is under the control of the 35S promoter from cauliflower mosaic virus (Odell et al. (1985) *Nature* 313:810-812) and the 3' region of the nopaline synthase gene from the T-DNA of the Ti plasmid of *Agrobacterium tumefaciens*.

The particle bombardment method (Klein et al. (1987) *Nature* 327:70-73) may be used to transfer genes to the callus culture cells. According to this method, gold particles (1 μm in diameter) are coated with DNA using the following technique. Ten μg of plasmid DNAs are added to 50 μL of a suspension of gold particles (60 mg per mL). Calcium chloride (50 μL of a 2.5 M solution) and spermidine free base (20 μL of a 1.0 M solution) are added to the particles. The suspension is vortexed during the addition of these solutions. After 10 minutes, the tubes are briefly centrifuged (5 sec at 15,000 rpm) and the supernatant removed. The particles are resuspended in 200 μL of absolute ethanol, centrifuged again and the supernatant removed. The ethanol rinse is performed again and the particles resuspended in a final volume of 30 μL of ethanol. An aliquot (5 μL) of the DNA-coated gold particles can be placed in the center of a KAPTON™ flying disc (Bio-Rad Labs). The particles are then accelerated into the corn tissue with a DUPONT™ BIOLISTIC™ PDS-1000/He (Bio-Rad Instruments, Hercules CA), using a helium pressure of 1000 psi, a gap distance of 0.5 cm and a flying distance of 1.0 cm.

For bombardment, the embryogenic tissue is placed on filter paper over agarose-solidified N6 medium. The tissue is arranged as a thin lawn and covers a circular area of about 5 cm in diameter. The petri dish containing the tissue can be placed in the chamber of the PDS-1000/He approximately 8 cm from the stopping

screen. The air in the chamber is then evacuated to a vacuum of 28 inches of Hg. The macrocarrier is accelerated with a helium shock wave using a rupture membrane that bursts when the He pressure in the shock tube reaches 1000 psi.

Seven days after bombardment the tissue can be transferred to N6 medium that contains bialaphos (5 mg per liter) and lacks casein or proline. The tissue continues to grow slowly on this medium. After an additional 2 weeks the tissue can be transferred to fresh N6 medium containing bialaphos. After 6 weeks, areas of about 1 cm in diameter of actively growing callus can be identified on some of the plates containing the bialaphos-supplemented medium. These calli may continue to grow when sub-cultured on the selective medium.

Plants can be regenerated from the transgenic callus by first transferring clusters of tissue to N6 medium supplemented with 0.2 mg per liter of 2,4-D. After two weeks the tissue can be transferred to regeneration medium (Fromm et al. (1990) *Bio/Technology* 8:833-839). Transgenic T0 plants can be regenerated and their phenotype determined following high throughput ("HTP") procedures. T1 seed can be collected.

T1 plants can be subjected to a soil-based drought stress. Using image analysis, plant area, volume, growth rate and color analysis can be taken at multiple times before and during drought stress. Overexpression constructs that result in a significant delay in wilting or leaf area reduction, yellow color accumulation and/or increased growth rate during drought stress will be considered evidence that the *Arabidopsis* gene functions in maize to enhance drought tolerance.

EXAMPLE 12

Electroporation of *Agrobacterium tumefaciens* LBA4404

Electroporation competent cells (40 μ L), such as *Agrobacterium tumefaciens* LBA4404 containing PHP10523 (FIG. 7; SEQ ID NO:37), are thawed on ice (20-30 min). PHP10523 contains VIR genes for T-DNA transfer, an *Agrobacterium* low copy number plasmid origin of replication, a tetracycline resistance gene, and a Cos site for in vivo DNA bimolecular recombination. Meanwhile the electroporation cuvette is chilled on ice. The electroporator settings are adjusted to 2.1 kV. A DNA aliquot (0.5 μ L parental DNA at a concentration of 0.2 μ g -1.0 μ g in low salt buffer or twice distilled H₂O) is mixed with the thawed *Agrobacterium tumefaciens* LBA4404 cells while still on ice. The mixture is transferred to the bottom of electroporation

cuvette and kept at rest on ice for 1-2 min. The cells are electroporated (Eppendorf electroporator 2510) by pushing the "pulse" button twice (ideally achieving a 4.0 millisecond pulse). Subsequently, 0.5 mL of room temperature 2xYT medium (or SOC medium) are added to the cuvette and transferred to a 15 mL snap-cap tube (e.g., FALCON™ tube). The cells are incubated at 28-30 °C, 200-250 rpm for 3 h.

Aliquots of 250 µL are spread onto plates containing YM medium and 50 µg/mL spectinomycin and incubated three days at 28-30 °C. To increase the number of transformants one of two optional steps can be performed:

Option 1: Overlay plates with 30 µL of 15 mg/mL rifampicin. LBA4404 has a chromosomal resistance gene for rifampicin. This additional selection eliminates some contaminating colonies observed when using poorer preparations of LBA4404 competent cells.

Option 2: Perform two replicates of the electroporation to compensate for poorer electrocompetent cells.

Identification of transformants:

Four independent colonies are picked and streaked on plates containing AB minimal medium and 50 µg/mL spectinomycin for isolation of single colonies. The plates are incubated at 28 °C for two to three days. A single colony for each putative co-integrate is picked and inoculated with 4 mL of 10 g/L bactopectone, 10 g/L yeast extract, 5 g/L sodium chloride and 50 mg/L spectinomycin. The mixture is incubated for 24 h at 28 °C with shaking. Plasmid DNA from 4 mL of culture is isolated using Qiagen® Miniprep and an optional Buffer PB wash. The DNA is eluted in 30 µL. Aliquots of 2 µL are used to electroporate 20 µL of DH10b + 20 µL of twice distilled H₂O as per above. Optionally a 15 µL aliquot can be used to transform 75-100 µL of INVITROGEN™ Library Efficiency DH5α. The cells are spread on plates containing LB medium and 50 µg/mL spectinomycin and incubated at 37 °C overnight.

Three to four independent colonies are picked for each putative co-integrate and inoculated 4 mL of 2xYT medium (10 g/L bactopectone, 10 g/L yeast extract, 5 g/L sodium chloride) with 50 µg/mL spectinomycin. The cells are incubated at 37 °C overnight with shaking. Next, isolate the plasmid DNA from 4 mL of culture using QIAprep® Miniprep with optional Buffer PB wash (elute in 50 µL). Use 8 µL for digestion with Sall (using parental DNA and PHP10523 as controls). Three more

digestions using restriction enzymes BamHI, EcoRI, and HindIII are performed for 4 plasmids that represent 2 putative co-integrates with correct Sall digestion pattern (using parental DNA and PHP10523 as controls). Electronic gels are recommended for comparison.

EXAMPLE 13

Transformation of Maize Using *Agrobacterium*

Maize plants can be transformed to overexpress a validated *Arabidopsis* lead gene or the corresponding homologs from various species in order to examine the resulting phenotype.

Agrobacterium-mediated transformation of maize is performed essentially as described by Zhao et al. in *Meth. Mol. Biol.* 318:315-323 (2006) (see also Zhao et al., *Mol. Breed.* 8:323-333 (2001) and U.S. Patent No. 5,981,840 issued November 9, 1999, incorporated herein by reference). The transformation process involves bacterium inoculation, co-cultivation, resting, selection and plant regeneration.

1. Immature Embryo Preparation:

Immature maize embryos are dissected from caryopses and placed in a 2 mL microtube containing 2 mL PHI-A medium.

2. *Agrobacterium* Infection and Co-Cultivation of Immature Embryos:

2.1 Infection Step:

PHI-A medium of (1) is removed with 1 mL micropipettor, and 1 mL of *Agrobacterium* suspension is added. The tube is gently inverted to mix. The mixture is incubated for 5 min at room temperature.

2.2 Co-culture Step:

The *Agrobacterium* suspension is removed from the infection step with a 1 mL micropipettor. Using a sterile spatula the embryos are scraped from the tube and transferred to a plate of PHI-B medium in a 100x15 mm Petri dish. The embryos are oriented with the embryonic axis down on the surface of the medium. Plates with the embryos are cultured at 20 °C, in darkness, for three days. L-Cysteine can be used in the co-cultivation phase. With the standard binary vector, the co-cultivation medium supplied with 100-400 mg/L L-cysteine is critical for recovering stable transgenic events.

3. Selection of Putative Transgenic Events:

To each plate of PHI-D medium in a 100x15 mm Petri dish, 10 embryos are transferred, maintaining orientation and the dishes are sealed with parafilm. The plates are incubated in darkness at 28 °C. Actively growing putative events, as pale yellow embryonic tissue, are expected to be visible in six to to eight weeks.

- 5 Embryos that produce no events may be brown and necrotic, and little friable tissue growth is evident. Putative transgenic embryonic tissue is subcultured to fresh PHI-D plates at two-three week intervals, depending on growth rate. The events are recorded.

4. Regeneration of T0 plants:

- 10 Embryonic tissue propagated on PHI-D medium is subcultured to PHI-E medium (somatic embryo maturation medium), in 100x25 mm Petri dishes and incubated at 28 °C, in darkness, until somatic embryos mature, for about ten to eighteen days. Individual, matured somatic embryos with well-defined scutellum and coleoptile are transferred to PHI-F embryo germination medium and incubated
15 at 28 °C in the light (about 80 µE from cool white or equivalent fluorescent lamps). In seven to ten days, regenerated plants, about 10 cm tall, are potted in horticultural mix and hardened-off using standard horticultural methods.

Media for Plant Transformation:

- 20 1. PHI-A: 4g/L CHU basal salts, 1.0 mL/L 1000X Eriksson's vitamin mix, 0.5 mg/L thiamin HCl, 1.5 mg/L 2,4-D, 0.69 g/L L-proline, 68.5 g/L sucrose, 36 g/L glucose, pH 5.2. Add 100 µM acetosyringone (filter-sterilized).
- 25 2. PHI-B: PHI-A without glucose, increase 2,4-D to 2 mg/L, reduce sucrose to 30 g/L and supplemente with 0.85 mg/L silver nitrate (filter-sterilized), 3.0 g/L Gelrite[®], 100 µM acetosyringone (filter-sterilized), pH 5.8.
- 30 3. PHI-C: PHI-B without Gelrite[®] and acetosyringonee, reduce 2,4-D to 1.5 mg/L and supplemente with 8.0 g/L agar, 0.5 g/L 2-[N-morpholino]ethane-sulfonic acid (MES) buffer, 100 mg/L carbenicillin (filter-sterilized).
4. PHI-D: PHI-C supplemented with 3 mg/L bialaphos (filter-sterilized).
5. PHI-E: 4.3 g/L of Murashige and Skoog (MS) salts, (Gibco, BRL 11117-074), 0.5 mg/L nicotinic acid, 0.1 mg/L thiamine HCl, 0.5 mg/L

pyridoxine HCl, 2.0 mg/L glycine, 0.1 g/L myo-inositol, 0.5 mg/L zeatin (Sigma, Cat. No. Z-0164), 1 mg/L indole acetic acid (IAA), 26.4 µg/L abscisic acid (ABA), 60 g/L sucrose, 3 mg/L bialaphos (filter-sterilized), 100 mg/L carbenicillin (filter-sterilized), 8 g/L agar, pH 5.6.

6. PHI-F: PHI-E without zeatin, IAA, ABA; reduce sucrose to 40 g/L; replacing agar with 1.5 g/L Gelrite®; pH 5.6.

Plants can be regenerated from the transgenic callus by first transferring clusters of tissue to N6 medium supplemented with 0.2 mg per liter of 2,4-D. After two weeks the tissue can be transferred to regeneration medium (Fromm et al., *Bio/Technology* 8:833-839 (1990)).

Transgenic T0 plants can be regenerated and their phenotype determined. T1 seed can be collected.

Furthermore, a recombinant DNA construct containing a validated *Arabidopsis* gene can be introduced into an elite maize inbred line either by direct transformation or introgression from a separately transformed line.

Transgenic plants, either inbred or hybrid, can undergo more vigorous field-based experiments to study yield enhancement and/or stability under water limiting and water non-limiting conditions.

Subsequent yield analysis can be done to determine whether plants that contain the validated *Arabidopsis* lead gene have an improvement in yield performance (under water limiting or non-limiting conditions), when compared to the control (or reference) plants that do not contain the validated *Arabidopsis* lead gene. Specifically, water limiting conditions can be imposed during the flowering and/or grain fill period for plants that contain the validated *Arabidopsis* lead gene and the control plants. Plants containing the validated *Arabidopsis* lead gene would have less yield loss relative to the control plants, for example, at least 50% less yield loss, under water limiting conditions, or would have increased yield relative to the control plants under water non-limiting conditions.

EXAMPLE 14A

Preparation of *Arabidopsis* Lead Gene (At5g26030)

Expression Vector for Transformation of Maize

Using INVITROGEN's™ GATEWAY® technology, an LR Recombination Reaction was performed with an entry clone (PHP31052) and a destination vector (PHP28647) to create the precursor plasmid PHP31079. The vector PHP31079 contains the following expression cassettes:

- 5 1. Ubiquitin promoter::moPAT::PinII terminator; cassette expressing the PAT herbicide resistance gene used for selection during the transformation process.
2. LTP2 promoter::DS-RED2::PinII terminator; cassette expressing the DS-RED color marker gene used for seed sorting.
3. Ubiquitin promoter::At5g26030::PinII terminator; cassette overexpressing
- 10 the gene of interest, Arabidopsis ferrochelatase-I.

EXAMPLE 14B

Transformation of Maize with the *Arabidopsis* Lead Gene (At5g26030) Using *Agrobacterium*

The ferrochelatase-I expression cassette present in vector PHP31079 can be
15 introduced into a maize inbred line, or a transformable maize line derived from an elite maize inbred line, using *Agrobacterium*-mediated transformation as described in Examples 12 and 13.

Vector PHP31079 can be electroporated into the LBA4404 *Agrobacterium* strain containing vector PHP10523 (FIG. 7; SEQ ID NO:37) to create the co-
20 integrate vector PHP31217. The co-integrate vector is formed by recombination of the 2 plasmids, PHP31079 and PHP10523, through the COS recombination sites contained on each vector. The co-integrate vector PHP31217 contains the same 3 expression cassettes as above (Example 14A) in addition to other genes (TET, TET, TRFA, ORI terminator, CTL, ORI V, VIR C1, VIR C2, VIR G, VIR B) needed for the
25 *Agrobacterium* strain and the *Agrobacterium*-mediated transformation.

EXAMPLE 15

Preparation of the Destination Vector PHP23236 for Transformation Into Gaspe Flint Derived Maize Lines

Destination vector PHP23236 (FIG. 6 , SEQ ID NO:36) was obtained by
30 transformation of *Agrobacterium* strain LBA4404 containing plasmid PHP10523 (FIG. 7, SEQ ID NO:37) with plasmid PHP23235 (FIG. 8, SEQ ID NO:38) and isolation of the resulting co-integration product. Destination vector PHP23236, can be used in a recombination reaction with an entry clone as described in Example 16

to create a maize expression vector for transformation of Gaspe Flint-derived maize lines.

EXAMPLE 16

Preparation of Plasmids for Transformation

into Gaspe Flint Derived Maize Lines

Using the INVITROGEN™ GATEWAY® LR Recombination technology, the same entry clone described in Example 5A, PHP31052, was directionally cloned into the destination vector PHP23236 (SEQ ID NO:36; FIG. 6) to create an expression vector, PHP31419. This expression vector contains the cDNA of interest, encoding AtFeC-I, under control of the UBI promoter and is a T-DNA binary vector for *Agrobacterium*-mediated transformation into corn as described, but not limited to, the examples described herein.

Using the INVITROGEN™ GATEWAY® LR Recombination technology, the same entry clone described in Example 5B, PHP-Entry-At2g30390, was directionally cloned into the destination vector PHP29634 to create an expression vector, PHP33089. Destination vector PHP29634 is similar to destination vector PHP23236, however, destination vector PHP29634 has site-specific recombination sites FRT1 and FRT87 and also encodes the GAT4602 selectable marker protein for selection of transformants using glyphosate. This expression vector contains the cDNA of interest, encoding AtFeC-II, under control of the UBI promoter and is a T-DNA binary vector for *Agrobacterium*-mediated transformation into corn as described, but not limited to, the examples described herein.

EXAMPLE 17

Transformation of Gaspe Flint Derived Maize Lines

with a Validated *Arabidopsis* Lead Gene

Maize plants can be transformed to overexpress the *Arabidopsis* lead gene or the corresponding homologs from other species in order to examine the resulting phenotype.

Recipient Plants:

Recipient plant cells can be from a uniform maize line having a short life cycle ("fast cycling"), a reduced size, and high transformation potential. Typical of these plant cells for maize are plant cells from any of the publicly available Gaspe Flint (GBF) line varieties. One possible candidate plant line variety is the F1 hybrid

of GBF x QTM (Quick Turnaround Maize, a publicly available form of Gaspe Flint selected for growth under greenhouse conditions) disclosed in Tomes et al. U.S. Patent Application Publication No. 2003/0221212. Transgenic plants obtained from this line are of such a reduced size that they can be grown in four inch pots (1/4 the space needed for a normal sized maize plant) and mature in less than 2.5 months. (Traditionally 3.5 months is required to obtain transgenic T0 seed once the transgenic plants are acclimated to the greenhouse.) Another suitable line is a double haploid line of GS3 (a highly transformable line) X Gaspe Flint. Yet another suitable line is a transformable elite inbred line carrying a transgene which causes early flowering, reduced stature, or both.

Transformation Protocol:

Any suitable method may be used to introduce the transgenes into the maize cells, including but not limited to inoculation type procedures using *Agrobacterium* based vectors. Transformation may be performed on immature embryos of the recipient (target) plant.

Precision Growth and Plant Tracking:

The event population of transgenic (T0) plants resulting from the transformed maize embryos is grown in a controlled greenhouse environment using a modified randomized block design to reduce or eliminate environmental error. A randomized block design is a plant layout in which the experimental plants are divided into groups (e.g., thirty plants per group), referred to as blocks, and each plant is randomly assigned a location with the block.

For a group of thirty plants, twenty-four transformed, experimental plants and six control plants (plants with a set phenotype) (collectively, a "replicate group") are placed in pots which are arranged in an array (a.k.a. a replicate group or block) on a table located inside a greenhouse. Each plant, control or experimental, is randomly assigned to a location with the block which is mapped to a unique, physical greenhouse location as well as to the replicate group. Multiple replicate groups of thirty plants each may be grown in the same greenhouse in a single experiment. The layout (arrangement) of the replicate groups should be determined to minimize space requirements as well as environmental effects within the greenhouse. Such a layout may be referred to as a compressed greenhouse layout.

An alternative to the addition of a specific control group is to identify those transgenic plants that do not express the gene of interest. A variety of techniques such as RT-PCR can be applied to quantitatively assess the expression level of the introduced gene. T0 plants that do not express the transgene can be compared to those which do.

Each plant in the event population is identified and tracked throughout the evaluation process, and the data gathered from that plant is automatically associated with that plant so that the gathered data can be associated with the transgene carried by the plant. For example, each plant container can have a machine readable label (such as a Universal Product Code (UPC) bar code) which includes information about the plant identity, which in turn is correlated to a greenhouse location so that data obtained from the plant can be automatically associated with that plant.

Alternatively any efficient, machine readable, plant identification system can be used, such as two-dimensional matrix codes or even radio frequency identification tags (RFID) in which the data is received and interpreted by a radio frequency receiver/processor. See U.S. Published Patent Application No. 2004/0122592, incorporated herein by reference.

Phenotypic Analysis Using Three-Dimensional Imaging:

Each greenhouse plant in the T0 event population, including any control plants, is analyzed for agronomic characteristics of interest, and the agronomic data for each plant is recorded or stored in a manner so that it is associated with the identifying data (see above) for that plant. Confirmation of a phenotype (gene effect) can be accomplished in the T1 generation with a similar experimental design to that described above.

The T0 plants are analyzed at the phenotypic level using quantitative, non-destructive imaging technology throughout the plant's entire greenhouse life cycle to assess the traits of interest. A digital imaging analyzer may be used for automatic multi-dimensional analyzing of total plants. The imaging may be done inside the greenhouse. Two camera systems, located at the top and side, and an apparatus to rotate the plant, are used to view and image plants from all sides. Images are acquired from the top, front and side of each plant. All three images together

provide sufficient information to evaluate the biomass, size and morphology of each plant.

Due to the change in size of the plants from the time the first leaf appears from the soil to the time the plants are at the end of their development, the early stages of plant development are best documented with a higher magnification from the top. This may be accomplished by using a motorized zoom lens system that is fully controlled by the imaging software.

In a single imaging analysis operation, the following events occur: (1) the plant is conveyed inside the analyzer area, rotated 360 degrees so its machine readable label can be read, and left at rest until its leaves stop moving; (2) the side image is taken and entered into a database; (3) the plant is rotated 90 degrees and again left at rest until its leaves stop moving, and (4) the plant is transported out of the analyzer.

Plants are allowed at least six hours of darkness per twenty four hour period in order to have a normal day/night cycle.

Imaging Instrumentation:

Any suitable imaging instrumentation may be used, including but not limited to light spectrum digital imaging instrumentation commercially available from LemnaTec GmbH of Wurselen, Germany. The images are taken and analyzed with a LemnaTec Scanalyzer HTS LT-0001-2 having a 1/2" IT Progressive Scan IEE CCD imaging device. The imaging cameras may be equipped with a motor zoom, motor aperture and motor focus. All camera settings may be made using LemnaTec software. For example, the instrumental variance of the imaging analyzer is less than about 5% for major components and less than about 10% for minor components.

Software:

The imaging analysis system comprises a LemnaTec HTS Bonit software program for color and architecture analysis and a server database for storing data from about 500,000 analyses, including the analysis dates. The original images and the analyzed images are stored together to allow the user to do as much reanalyzing as desired. The database can be connected to the imaging hardware for automatic data collection and storage. A variety of commercially available software systems (e.g. Matlab, others) can be used for quantitative interpretation of

the imaging data, and any of these software systems can be applied to the image data set.

Conveyor System:

A conveyor system with a plant rotating device may be used to transport the plants to the imaging area and rotate them during imaging. For example, up to four plants, each with a maximum height of 1.5 m, are loaded onto cars that travel over the circulating conveyor system and through the imaging measurement area. In this case the total footprint of the unit (imaging analyzer and conveyor loop) is about 5 m x 5 m.

The conveyor system can be enlarged to accommodate more plants at a time. The plants are transported along the conveyor loop to the imaging area and are analyzed for up to 50 seconds per plant. Three views of the plant are taken. The conveyor system, as well as the imaging equipment, should be capable of being used in greenhouse environmental conditions.

Illumination:

Any suitable mode of illumination may be used for the image acquisition. For example, a top light above a black background can be used. Alternatively, a combination of top- and backlight using a white background can be used. The illuminated area should be housed to ensure constant illumination conditions. The housing should be longer than the measurement area so that constant light conditions prevail without requiring the opening and closing of doors. Alternatively, the illumination can be varied to cause excitation of either transgene (e.g., green fluorescent protein (GFP), red fluorescent protein (RFP)) or endogenous (e.g. Chlorophyll) fluorophores.

Biomass Estimation Based on Three-Dimensional Imaging:

For best estimation of biomass the plant images should be taken from at least three axes, for example, the top and two side (sides 1 and 2) views. These images are then analyzed to separate the plant from the background, pot and pollen control bag (if applicable). The volume of the plant can be estimated by the calculation:

$$Volume(voxels) = \sqrt{TopArea(pixels)} \times \sqrt{Side1Area(pixels)} \times \sqrt{Side2Area(pixels)}$$

In the equation above the units of volume and area are “arbitrary units”. Arbitrary units are entirely sufficient to detect gene effects on plant size and growth in this system because what is desired is to detect differences (both positive-larger and negative-smaller) from the experimental mean, or control mean. The arbitrary units of size (e.g. area) may be trivially converted to physical measurements by the addition of a physical reference to the imaging process. For instance, a physical reference of known area can be included in both top and side imaging processes. Based on the area of these physical references a conversion factor can be determined to allow conversion from pixels to a unit of area such as square centimeters (cm²). The physical reference may or may not be an independent sample. For instance, the pot, with a known diameter and height, could serve as an adequate physical reference.

Color Classification:

The imaging technology may also be used to determine plant color and to assign plant colors to various color classes. The assignment of image colors to color classes is an inherent feature of the LemnaTec software. With other image analysis software systems color classification may be determined by a variety of computational approaches.

For the determination of plant size and growth parameters, a useful classification scheme is to define a simple color scheme including two or three shades of green and, in addition, a color class for chlorosis, necrosis and bleaching, should these conditions occur. A background color class which includes non plant colors in the image (for example pot and soil colors) is also used and these pixels are specifically excluded from the determination of size. The plants are analyzed under controlled constant illumination so that any change within one plant over time, or between plants or different batches of plants (e.g. seasonal differences) can be quantified.

In addition to its usefulness in determining plant size growth, color classification can be used to assess other yield component traits. For these other yield component traits additional color classification schemes may be used. For instance, the trait known as “staygreen”, which has been associated with improvements in yield, may be assessed by a color classification that separates

shades of green from shades of yellow and brown (which are indicative of senescing tissues). By applying this color classification to images taken toward the end of the T0 or T1 plants' life cycle, plants that have increased amounts of green colors relative to yellow and brown colors (expressed, for instance, as Green/Yellow Ratio) may be identified. Plants with a significant difference in this Green/Yellow ratio can be identified as carrying transgenes which impact this important agronomic trait.

The skilled plant biologist will recognize that other plant colors arise which can indicate plant health or stress response (for instance anthocyanins), and that other color classification schemes can provide further measures of gene action in traits related to these responses.

Plant Architecture Analysis:

Transgenes which modify plant architecture parameters may also be identified using the present invention, including such parameters as maximum height and width, internodal distances, angle between leaves and stem, number of leaves starting at nodes and leaf length. The LemnaTec system software may be used to determine plant architecture as follows. The plant is reduced to its main geometric architecture in a first imaging step and then, based on this image, parameterized identification of the different architecture parameters can be performed. Transgenes that modify any of these architecture parameters either singly or in combination can be identified by applying the statistical approaches previously described.

Pollen Shed Date:

Pollen shed date is an important parameter to be analyzed in a transformed plant, and may be determined by the first appearance on the plant of an active male flower. To find the male flower object, the upper end of the stem is classified by color to detect yellow or violet anthers. This color classification analysis is then used to define an active flower, which in turn can be used to calculate pollen shed date.

Alternatively, pollen shed date and other easily visually detected plant attributes (e.g. pollination date, first silk date) can be recorded by the personnel responsible for performing plant care. To maximize data integrity and process efficiency this data is tracked by utilizing the same barcodes utilized by the LemnaTec light spectrum digital analyzing device. A computer with a barcode

reader, a palm device, or a notebook PC may be used for ease of data capture recording time of observation, plant identifier, and the operator who captured the data.

Orientation of the Plants:

5 Mature maize plants grown at densities approximating commercial planting often have a planar architecture. That is, the plant has a clearly discernable broad side, and a narrow side. The image of the plant from the broadside is determined. To each plant a well defined basic orientation is assigned to obtain the maximum difference between the broadside and edgewise images. The top image is used to
10 determine the main axis of the plant, and an additional rotating device is used to turn the plant to the appropriate orientation prior to starting the main image acquisition.

EXAMPLE 18

Screening of Gaspe Flint Derived

Maize Lines for Drought Tolerance

15 Transgenic Gaspe Flint derived maize lines containing the candidate gene can be screened for tolerance to drought stress in the following manner.

Transgenic maize plants are subjected to well-watered conditions (control) and to drought-stressed conditions. Transgenic maize plants are screened at the
20 T1 stage or later.

Stress is imposed starting at 10 to 14 days after sowing (DAS) or 7 days after transplanting, and is continued through to silking. Pots are watered by an automated system fitted to timers to provide watering at 25 or 50% of field capacity during the entire period of drought-stress treatment. The intensity and duration of
25 this stress will allow identification of the impact on vegetative growth as well as on the anthesis-silking interval.

Potting mixture: A mixture of 1/3 TURFACE® (Profile Products LLC, IL, USA), 1/3 sand and 1/3 SB300 (Sun Gro Horticulture, WA, USA) can be used. The SB300 can be replaced with *Fafard Fine-Germ* (Conrad Fafard, Inc., MA, USA) and
30 the proportion of sand in the mixture can be reduced. Thus, a final potting mixture can be 3/8 (37.5%) TURFACE®, 3/8 (37.5%) *Fafard* and 1/4 (25%) sand.

Field Capacity Determination: The weight of the soil mixture (w1) to be used in one S200 pot (minus the pot weight) is measured. If all components of the soil

mix are not dry, the soil is dried at 100°C to constant weight before determining w1. The soil in the pot is watered to full saturation and all the gravitational water is allowed to drain out. The weight of the soil (w2) after all gravitational water has seeped out (minus the pot weight) is determined. Field capacity is the weight of the water remaining in the soil obtained as w2-w1. It can be written as a percentage of the oven-dry soil weight.

Stress Treatment: During the early part of plant growth (10 DAS to 21 DAS), the well-watered control has a daily watering of 75% field capacity and the drought-stress treatment has a daily watering of 25% field capacity, both as a single daily dose at or around 10 AM. As the plants grow bigger, by 21 DAS, it will become necessary to increase the daily watering of the well-watered control to full field capacity and the drought stress treatment to 50% field capacity.

Nutrient Solution: A modified Hoagland's solution at 1/16 dilution with tap water is used for irrigation.

TABLE 8

Preparation of 20 L of Modified Hoagland's Solution Using the Following Recipe:

Component	Amount/20 L
10X Micronutrient Solution	16 mL
KH ₂ PO ₄ (MW: 136.02)	22 g
MgSO ₄ (MW: 120.36)	77 g
KNO ₃ (MW: 101.2)	129.5 g
Ca(NO ₃) ₂ ·4H ₂ O (MW: 236.15)	151 g
NH ₄ NO ₃ (MW: 80.04)	25.6 g
Sprint 330 (Iron chelate)	32 g

TABLE 9

Preparation of 1L of 10X Micronutrient Solution Using the Following Recipe:

Component	mg/L	Concentration
H ₃ BO ₃	1854	30 mM
MnCl ₂ ·4H ₂ O	1980	10 mM
ZnSO ₄ ·7H ₂ O	2874	10 mM
CuSO ₄ ·5H ₂ O	250	1 mM
H ₂ MoO ₄ ·H ₂ O	242	1 mM

Fertilizer grade KNO_3 is used.

It is useful to add half a teaspoon of OSMOCOTE® (NPK 15:9:12) to the pot at the time of transplanting or after emergence (The Scotts Miracle-Gro Company, OH, USA).

5 *Border plants:* Place a row of border plants on bench-edges adjacent to the glass walls of the greenhouse or adjacent to other potential causes of microenvironment variability such as a cooler fan.

10 *Automation:* Watering can be done using PVC pipes with drilled holes to supply water to systematically positioned pots using a siphoning device. Irrigation scheduling can be done using timers.

Statistical analysis: Mean values for plant size, color and chlorophyll fluorescence recorded on transgenic events under different stress treatments will be exported to Spotfire (Spotfire, Inc., MA, USA). Treatment means will be evaluated for differences using Analysis of Variance.

15 *Replications:* Eight to ten individual plants are used per treatment per event.

Observations Made: Lemnatec measurements are made three times a week throughout growth to capture plant-growth rate. Leaf color determinations are made three times a week throughout the stress period using Lemnatec. Chlorophyll fluorescence is recorded as PhiPSII (which is indicative of the operating quantum efficiency of photosystem II photochemistry) and F_v'/F_m' (which is the maximum efficiency of photosystem II) two to four times during the experimental period, starting at 11 AM on the measurement days, using the Hansatech FMS2 instrument (LemnaTec GmbH, Wurselen, Germany). Measurements are started during the stress period at the beginning of visible drought stress symptoms, namely, leaf greying and the start of leaf rolling until the end of the experiment and measurements are recorded on the youngest most fully expanded leaf. The dates of tasseling and silking on individual plants are recorded, and the ASI is computed.

25 The above methods may be used to select transgenic plants with increased drought tolerance when compared to a control plant not comprising said recombinant DNA construct.

EXAMPLE 19

Yield Analysis of Maize Lines with the *Arabidopsis* Lead Gene

A recombinant DNA construct containing a validated *Arabidopsis* gene can
5 be introduced into an elite maize inbred line either by direct transformation or
introgression from a separately transformed line.

Transgenic plants, either inbred or hybrid, can undergo more vigorous field-
based experiments to study yield enhancement and/or stability under well-watered
and water-limiting conditions.

10 Subsequent yield analysis can be done to determine whether plants that
contain the validated *Arabidopsis* lead gene have an improvement in yield
performance under water-limiting conditions, when compared to the control plants
that do not contain the validated *Arabidopsis* lead gene. Specifically, drought
conditions can be imposed during the flowering and/or grain fill period for plants that
15 contain the validated *Arabidopsis* lead gene and the control plants. Reduction in
yield can be measured for both. Plants containing the validated *Arabidopsis* lead
gene have less yield loss relative to the control plants, for example, at least 50%
less yield loss.

The above method may be used to select transgenic plants with increased
20 yield, under water-limiting conditions and/or well-watered conditions, when
compared to a control plant not comprising said recombinant DNA construct.

EXAMPLE 20A

Preparation of Maize Ferrochelatase-I Lead Gene Expression Vector for Transformation of Maize

25 Clone cfp5n.pk009.j16 encodes a maize ferrochelatase-I protein designated
“ZmFeC-Ia” (SEQ ID NO:10). The protein-coding region of clone cfp5n.pk009.j16
was introduced into the INVITROGEN™ vector pENTR/D-TOPO® to create entry
clone PHP30949 (SEQ ID NO:55).

Using INVITROGEN's™ GATEWAY® technology, an LR Recombination
30 Reaction was performed with an entry clone (PHP30949) and a destination vector
(PHP28647) to create the precursor plasmid PHP30963. The vector PHP30963
contains the following expression cassettes:

1. Ubiquitin promoter::moPAT::PinII terminator; cassette expressing the PAT herbicide resistance gene used for selection during the transformation process.

2. LTP2 promoter::DS-RED2::PinII terminator; cassette expressing the DS-RED color marker gene used for seed sorting.

5 3. Ubiquitin promoter::ZmFeC-Ia::PinII terminator; cassette overexpressing the gene of interest, maize ferrochelatase-I.

EXAMPLE 20B

Transformation of Maize with Maize Ferrochelatase-I

Lead Gene Using *Agrobacterium*

10 The ZmFeC-Ia expression cassette present in vector PHP30963 can be introduced into a maize inbred line, or a transformable maize line derived from an elite maize inbred line, using *Agrobacterium*-mediated transformation as described in Examples 12 and 13.

Vector PHP30963 can be electroporated into the LBA4404 *Agrobacterium*
15 strain containing vector PHP10523 (FIG. 7; SEQ ID NO:37) to create the co-integrate vector PHP30976. The co-integrate vector is formed by recombination of the 2 plasmids, PHP30963 and PHP10523, through the COS recombination sites contained on each vector. The co-integrate vector PHP30976 contains the same 3 expression cassettes as above (Example 14A) in addition to other genes (TET, TET,
20 TRFA, ORI terminator, CTL, ORI V, VIR C1, VIR C2, VIR G, VIR B) needed for the *Agrobacterium* strain and the *Agrobacterium*-mediated transformation.

EXAMPLE 21

Preparation of Maize Expression Plasmids for Transformation into Gaspe Flint Derived Maize Lines

25 Clone cfp5n.pk009.j16 encodes a complete maize ferrochelatase-I homolog designated "ZmFeC-Ia" (SEQ ID NO:10). Clone cfp3n.pk004.f12 encodes a different complete maize ferrochelatase-I homolog designated "ZmFeC-Ib" (SEQ ID NO:8). Clone cfp5n.pk064.n7 also encodes the ZmFeC-Ib protein (SEQ ID NO:8), however the nucleotide sequence of cfp5n.pk064.n7 (SEQ ID NO:56) differs from
30 the nucleotide sequence of cfp3n.pk004.f12 (SEQ ID NO:7) in that it contains a 12 base pair insertion in the 5'-UTR, 8 nucleotides before the ATG start codon.

Using the INVITROGEN™ GATEWAY® Recombination technology described in Example 9, the three clones encoding maize ferrochelatase-I homologs

were directionally cloned into the destination vector PHP23236 (SEQ ID NO:36; FIG. 6) to create the expression vectors listed in Table 10. Each expression vector contains the cDNA of interest under control of the UBI promoter and is a T-DNA binary vector for *Agrobacterium*-mediated transformation into corn as described, but not limited to, the examples described herein.

TABLE 10**Maize Ferrochelatase-I Expression Vectors**

Protein	Clone Origin	SEQ ID NO: (Amino Acid)	Expression Vector
ZmFeC-Ia	cfp5n.pk009.j16 (FIS)	10	PHP30829
ZmFeC-Ib	cfp3n.pk004.f12 (FIS)	8	PHP30745
ZmFeC-Ib	cfp5n.pk064.n7 (FIS)	8	PHP30761

EXAMPLE 22**Transformation and Evaluation of Gaspe Flint Derived****Maize Lines for Drought Tolerance**

Gaspe Flint derived maize lines were transformed via *Agrobacterium* using the following plasmid DNAs: PHP31419 (AtFeC-I; At5g26030); PHP33089 (AtFeC-II; At2g30390); PHP30829 (ZmFeC-Ia; cfp5n.pk009.j16); PHP30745 (ZmFeC-Ib; cfp3n.pk004.f12) and PHP30761 (ZmFeC-Ib; cfp5n.pk064.n7). Four transformation events for each plasmid construct were evaluated for drought tolerance in the following manner. For plant growth, the soil mixture consisted of ⅓ TURFACE®, ⅓ SB300 and ⅓ sand. All pots were filled with the same amount of soil ± 10 grams. Pots were brought up to 100% field capacity ("FC") by hand watering. All plants were maintained at 60% FC using a 20-10-20 (N-P-K) 125 ppm N nutrient solution. Throughout the experiment pH was monitored at least three times weekly for each table. Starting at 13 days after planting (DAP), the experiment was divided into two treatment groups, well watered and reduce watered. All plants comprising the reduced watered treatment were maintained at 40% FC while plants in the well watered treatment were maintained at 80% FC. Reduced watered plants were grown for 10 days under chronic drought stress conditions (40% FC). All plants were imaged daily throughout chronic stress period. Plants were sampled for

metabolic profiling analyses at the end of chronic drought period, 22 DAP . At the conclusion of the chronic stress period all plants were imaged and measured for chlorophyll fluorescence. Reduced watered plants were subjected to a severe drought stress period followed by a recovery period, 23 – 31 DAP and 32 – 34 DAP respectively. During the severe drought stress, water and nutrients were withheld until the plants reached 8% FC. At the conclusion of severe stress and recovery periods all plants were again imaged and measured for chlorophyll fluorescence. The probability of a greater Student's t Test was calculated for each transgenic mean compared to the appropriate null mean (either segregant null or construct null). A minimum ($P < t$) of 0.1 was used as a cut off for a statistically significant result.

Tables 11-20 show the variables for each transgenic event that were significantly altered, as compared to the segregant nulls. A "positive effect" was defined as statistically significant improvement in that variable for the transgenic event relative to the null control. A "negative effect" was defined as a statistically significant improvement in that variable for the null control relative to the transgenic event. Tables 11, 13, 15, 17, and 19 present the number of variable with a significant change for individual events transformed with each of the five plasmid DNA constructs. Tables 12, 14, 16, 18 and 20 present the number of events for each construct that showed a significant change for each individual variable.

TABLE 11
Number of Variables with a Significant Change* for Individual Events
Transformed with PHP31419 Encoding AtFeC-I (At5g26030)

Event	Reduced Water		Well Watered	
	Positive Effect	Negative Effect	Positive Effect	Negative Effect
EA2392.498.1.2	2	1	4	0
EA2392.498.1.3	3	1	1	3
EA2392.498.1.5	0	0	3	0
EA2392.498.1.6	12	0	3	0

*P-value less than or equal to 0.1

TABLE 12
Number of Events Transformed with PHP31419 Encoding AtFeC-I (At5g26030) with
a Significant Change* for Individual Variables

Variable	Reduced Water		Well Watered	
	Positive Effect	Negative Effect	Positive Effect	Negative Effect
% area chg_start chronic – acute2	1	0	1	0
% area chg_start chronic – end chronic	1	0	2	0
% area chg_start chronic – harvest	1	0	1	0
%area chg_start chronic – recovery24hr	1	0	1	1
fv/fm_acute1	2	0	2	0
fv/fm_acute2	1	0	1	0
leaf rolling_harvest	2	0	0	0
leaf rolling_recovery24hr	2	0	1	0
psii_acute1	2	0	2	0
psii_acute2	1	0	0	0
sgr – r2 > 0.9	2	0	0	0
shoot dry weight	0	1	0	1
shoot fresh weight	1	1	0	1

5 *P-value less than or equal to 0.1

TABLE 13
Number of Variables with a Significant Change* for Individual Events
Transformed with PHP33089 Encoding AtFeC-II (At2g30390)

Event	Reduced Water		Well Watered	
	Positive Effect	Negative Effect	Positive Effect	Negative Effect
EA2534.088.2.2	0	3	1	0
EA2534.088.2.4	3	1	2	2
EA2534.088.2.5	6	0	2	2
EA2534.088.2.8	1	3	0	2

10 *P-value less than or equal to 0.1

TABLE 14
 Number of Events Transformed with PHP33089 Encoding AtFeC-II (At2g30390)
 with a Significant Change* for Individual Variables

Variable	Reduced Water		Well Watered	
	Positive Effect	Negative Effect	Positive Effect	Negative Effect
% area chg_start chronic – end chronic	2	0	1	0
% area chg_start chronic – end severe	1	2	0	0
% area chg_start chronic – recovery48hr	0	1	0	0
fv/fm_end severe	1	2	1	0
fv/fm_recovery48hr	1	0	0	1
fv/fm_start severe	1	0	0	0
leaf rolling_recovery24hr	0	0	1	1
psii_end severe	1	2	1	0
psii_recovery48hr	1	0	0	2
psii_start severe	1	0	1	2
sgr – r2 > 0.9	0	0	0	0
shoot dry weight	1	0	0	0
shoot fresh weight	0	0	0	0

*P-value less than or equal to 0.1

5

TABLE 15
 Number of Variables with a Significant Change* for Individual Events
 Transformed with PHP30829 Encoding MzFeC-Ia (cfp5n.pk009.j16)

Event	Reduced Water		Well Watered	
	Positive Effect	Negative Effect	Positive Effect	Negative Effect
EA2391.472.1.10	0	5	3	3
EA2391.472.1.2	0	2	6	1
EA2391.472.1.3	0	4	2	0
EA2391.472.1.4	2	2	3	0

*P-value less than or equal to 0.1

10

TABLE 16

Number of Events Transformed with PHP30829 Encoding MzFeC-Ia (cfp5n.pk009.j16) with a Significant Change* for Individual Variables

Variable	Reduced Water		Well Watered	
	Positive Effect	Negative Effect	Positive Effect	Negative Effect
% area chg_start chronic - end chronic	0	2	0	0
% area chg_start chronic - harvest	0	1	3	0
% area chg_start chronic - recovery24hr	0	0	0	0
% area chg_start chronic - recovery48hr	0	1	0	0
fv/fm_acute1	0	0	2	1
fv/fm_acute2	1	1	2	0
leaf rolling_recovery24hr	0	1	1	0
leaf rolling_recovery48hr	0	2	1	0
psii_acute1	0	0	2	1
psii_acute2	1	1	1	1
sgr - r2 > 0.9	0	2	1	0
shoot dry weight	0	1	0	1
shoot fresh weight	0	1	1	0

*P-value less than or equal to 0.1

5

TABLE 17

Number of Variables with a Significant Change* for Individual Events Transformed with PHP30745 Encoding MzFeC-Ib (cfp3n.pk004.f12)

Event	Reduced Water		Well Watered	
	Positive Effect	Negative Effect	Positive Effect	Negative Effect
EA2392.447.1.1	4	1	4	0
EA2392.447.1.3	0	3	2	7
EA2392.447.1.5	0	0	2	1
EA2392.447.1.9	2	4	4	1

*P-value less than or equal to 0.1

10

TABLE 18

Number of Events Transformed with PHP30745 Encoding MzFeC-Ib (cfp3n.pk004.f12) with a Significant Change* for Individual Variables

Variable	Reduced Water		Well Watered	
	Positive Effect	Negative Effect	Positive Effect	Negative Effect
% area chg_start chronic - end chronic	1	0	3	1
% area chg_start chronic - end severe	0	0	2	1
% area chg_start chronic - recovery48hr	0	0	2	1
fv/fm_end severe	2	0	0	0
fv/fm_recovery48hr	1	1	0	1
fv/fm_start severe	0	1	0	1
leaf rolling_recovery48hr	0	0	0	0
psii_end severe	0	0	0	0
psii_recovery48hr	0	1	1	2
psii_start severe	0	2	0	1
shoot dry weight	2	0	2	1
shoot fresh weight	2	1	2	0

*P-value less than or equal to 0.1

5

TABLE 19

Number of Variables with a Significant Change* for Individual Events Transformed with PHP30761 Encoding MzFeC-Ib (cfp5n.pk064.n7)

Event	Reduced Water		Well Watered	
	Positive Effect	Negative Effect	Positive Effect	Negative Effect
EA2392.441.1.1	0	0	0	1
EA2392.441.1.2	0	0	2	0
EA2392.441.1.4	5	0	7	3
EA2392.441.1.7	5	3	0	0

*P-value less than or equal to 0.1

10

TABLE 20
 Number of Events Transformed with PHP30761 Encoding MzFeC-Ib
 (cfp5n.pk064.n7) with a Significant Change* for Individual Variables

Variable	Reduced Water		Well Watered	
	Positive Effect	Negative Effect	Positive Effect	Negative Effect
% area chg_start chronic - end chronic	0	1	0	1
% area chg_start chronic - end severe	0	1	0	0
% area chg_start chronic - recovery48hr	0	0	0	0
fv/fm_end severe	2	0	1	0
fv/fm_recovery48hr	2	0	1	0
fv/fm_start severe	0	0	2	0
leaf rolling_recovery48hr	1	0	1	0
psii_end severe	2	0	1	0
psii_recovery48hr	2	0	1	0
psii_start severe	1	0	1	0
sgr - r2 > 0.9	0	0	0	1
shoot dry weight	0	1	1	2
shoot fresh weight	0	0	0	0

*P-value less than or equal to 0.1

5

For each of the five constructs evaluated, the statistical value associated with each improved variable is presented in Figures 14A – 23. A significant positive result had a P-value of less than or equal to 0.1. The results for individual transformed maize lines are presented in Figures 14A-14B, 16A-16B, 18A-18B, 20A-20B and 22A-22B. The summary evaluations for each of the five constructs are presented in Figures 15, 17, 19, 21, and 23.

10

EXAMPLE 23

Screening of Inbred Derived

Maize Lines for Drought Tolerance

A transformable maize line derived from an elite maize inbred line was transformed with PHP30829 which encodes the maize ferrochelatase-I protein, MzFeC-1a (SEQ ID NO:10). Seed of transgenic events from the PHP30829

15

transformation were separated into transgenic and null seed using a seed color marker. The Fv'/Fm' and Phi PSII data were collected from a drought seedling assay following a procedure similar to the one described in Example 18 with the following modifications: an elite maize hybrid seedling was used for the assay instead of a Gaspe Flint derived maize line; and the chlorophyll fluorescence data were collected only during recovery from moderate drought stress (or chronic drought stress). The experiment was designed and analyzed as a Randomized Complete Block Design with the events including the construct null as treatments. A number of individual events were found to have higher Fv'/Fm' and/or Phi PSII values during drought stress relative to the null segregant control. The data is presented in the Table below.

TABLE 21

Fv'/Fm' and Phi PSII Values For Individual Events Transformed
With PHP30829 Encoding MzFeC-1a (cfp5n.pk009.j16)

Event	Fv'/Fm'	Phi PSII
E7733.78.1.1	0.421	0.330 ^a
E7733.78.1.2	0.462	0.407 ^a
E7733.78.2.20	0.520 ^a	0.420 ^a
E7733.78.4.1	0.428	0.339 ^a
E7733.78.2.10	0.475 ^a	0.398 ^a
E7733.78.2.2	0.360 ^b	0.284
E7733.78.2.6	0.373 ^b	0.289
E7733.78.2.7	0.469 ^a	0.392 ^a
30976-CN ^c	0.392	0.314

^a Positive and statistically significant at the 0.1 level of confidence

^b Negative and statistically significant at the 0.1 level of confidence

^c Null segregant used as a control

EXAMPLE 24

Transformation and Evaluation of Soybean
with Soybean Homologs of Validated Lead Genes

Based on homology searches, one or several candidate soybean homologs of validated *Arabidopsis* lead genes can be identified and also be assessed for their ability to enhance drought tolerance in soybean. Vector construction, plant

transformation and phenotypic analysis will be similar to that in previously described Examples.

EXAMPLE 25

Transformation of *Arabidopsis* with

Maize and Soybean Homologs of Validated Lead Genes

Soybean and maize homologs to validated *Arabidopsis* lead genes can be transformed into *Arabidopsis* under control of the 35S promoter and assessed for their ability to enhance drought tolerance in *Arabidopsis*. Vector construction, plant transformation and phenotypic analysis will be similar to that in previously described Examples.

CLAIMS

What is claimed is:

1. A plant comprising in its genome a recombinant DNA construct comprising a polynucleotide operably linked to at least one regulatory element,
5 wherein said polynucleotide encodes a polypeptide having an amino acid sequence of at least 50% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28 or 30, and wherein said plant exhibits increased drought tolerance when compared to a control plant not comprising said recombinant DNA construct.
- 10 2. The plant of claim 1, wherein the plant is a maize plant or a soybean plant.
3. A method of increasing drought tolerance in a plant, comprising:
 - (a) introducing into a regenerable plant cell a recombinant DNA construct comprising a polynucleotide operably linked to at least one regulatory
15 sequence, wherein the polynucleotide encodes a polypeptide having an amino acid sequence of at least 50% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28 or 30; and
 - (b) regenerating a transgenic plant from the regenerable plant cell after
20 step (a), wherein the transgenic plant comprises in its genome the recombinant DNA construct and exhibits increased drought tolerance when compared to a control plant not comprising the recombinant DNA construct.
4. The method of claim 3, further comprising:
 - (c) obtaining a progeny plant derived from the transgenic plant, wherein
25 said progeny plant comprises in its genome the recombinant DNA construct and exhibits increased drought tolerance when compared to a control plant not comprising the recombinant DNA construct.
5. A method of evaluating drought tolerance in a plant, comprising:
 - (a) introducing into a regenerable plant cell a recombinant DNA
30 construct comprising a polynucleotide operably linked to at least one regulatory sequence, wherein the polynucleotide encodes a polypeptide having an amino acid sequence of at least 50% sequence identity, based on

the Clustal V method of alignment, when compared to SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28 or 30;

(b) regenerating a transgenic plant from the regenerable plant cell after step (a), wherein the transgenic plant comprises in its genome the recombinant

5 DNA construct; and

(c) evaluating the transgenic plant for drought tolerance compared to a control plant not comprising the recombinant DNA construct.

6. The method of claim 5, further comprising:

(d) obtaining a progeny plant derived from the transgenic plant, wherein
10 the progeny plant comprises in its genome the recombinant DNA construct; and

(e) evaluating the progeny plant for drought tolerance compared to a control plant not comprising the recombinant DNA construct.

7. A method of evaluating drought tolerance in a plant, comprising:

(a) introducing into a regenerable plant cell a recombinant DNA
15 construct comprising a polynucleotide operably linked to at least one regulatory sequence, wherein the polynucleotide encodes a polypeptide having an amino acid sequence of at least 50% sequence identity, based on the Clustal V method of alignment, when compared to SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28 or 30;

(b) regenerating a transgenic plant from the regenerable plant cell after
20 step (a), wherein the transgenic plant comprises in its genome the recombinant DNA construct;

(c) obtaining a progeny plant derived from the transgenic plant, wherein the progeny plant comprises in its genome the recombinant DNA construct; and

(d) evaluating the progeny plant for drought tolerance compared to a
25 control plant not comprising the recombinant DNA construct.

8. The method of claim 3, wherein the plant is a maize plant or a soybean plant.

9. The method of claim 4, wherein the plant is a maize plant or a soybean
30 plant.

10. The method of claim 5, wherein the plant is a maize plant or a soybean plant.

11. The method of claim 6, wherein the plant is a maize plant or a soybean plant.

12. The method of claim 7, wherein the plant is a maize plant or a soybean plant.

5 13. An isolated polynucleotide comprising:

(a) a nucleotide sequence encoding a polypeptide with ferrochelatase activity, wherein the polypeptide has an amino acid sequence of at least 90% sequence identity, based on the Clustal V method of alignment with pairwise alignment default parameters of KTUPLE=1, GAP PENALTY=3, WINDOW=5 and DIAGONALS SAVED=5, when compared to SEQ ID NO:4, 6, 8, 10, 12, 14, 16, 18, 24, 26, 28 or 30, or

(b) the full complement of the nucleotide sequence of (a).

14. The polynucleotide of Claim 13, wherein the polypeptide has an amino acid sequence of at least 95% sequence identity, based on the Clustal V method of alignment with the pairwise alignment default parameters, when compared to SEQ ID NO:4, 6, 8, 10, 12, 14, 16, 18, 24, 26, 28 or 30.

15. The polynucleotide of Claim 28, wherein the amino acid sequence of the polypeptide comprises SEQ ID NO:4, 6, 8, 10, 12, 14, 16, 18, 24, 26, 28 or 30.

16. The polynucleotide of Claim 13 wherein the nucleotide sequence comprises SEQ ID NO:3, 5, 7, 9, 11, 13, 15, 17, 23, 25, 27 or 29.

17. A vector comprising the polynucleotide of Claim 13.

18. A recombinant DNA construct comprising the isolated polynucleotide of Claim 13 operably linked to at least one regulatory sequence.

19. A cell comprising the recombinant DNA construct of Claim 18, wherein the cell is selected from the group consisting of a bacterial cell, a yeast cell, and insect cell and a plant cell.

20. A plant or seed comprising the recombinant DNA construct of Claim 18.

FIG. 1

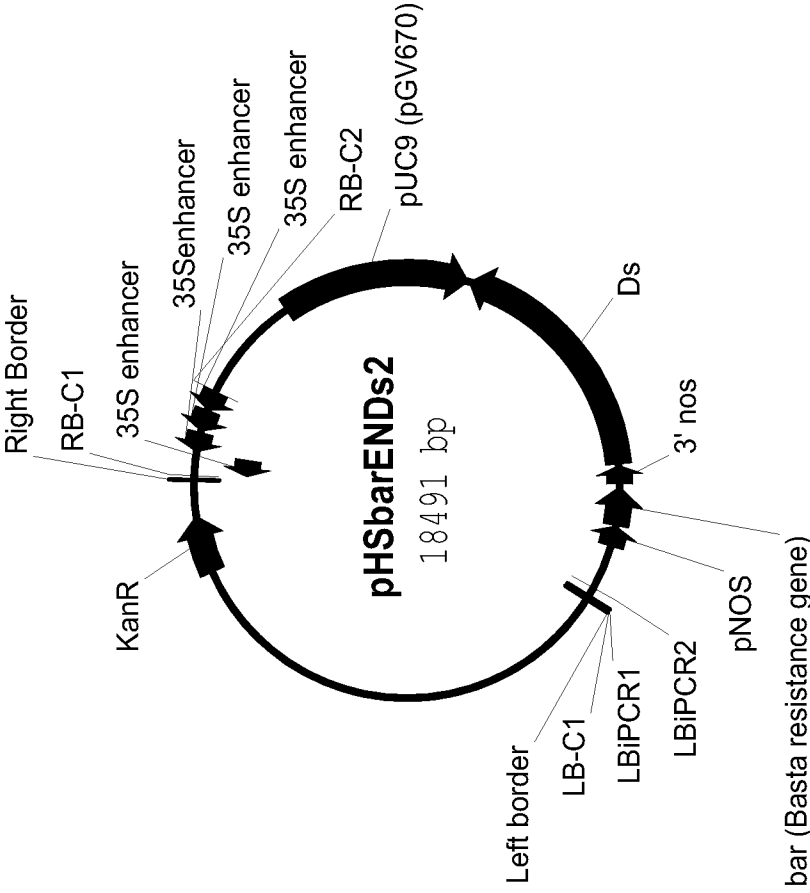


FIG. 2

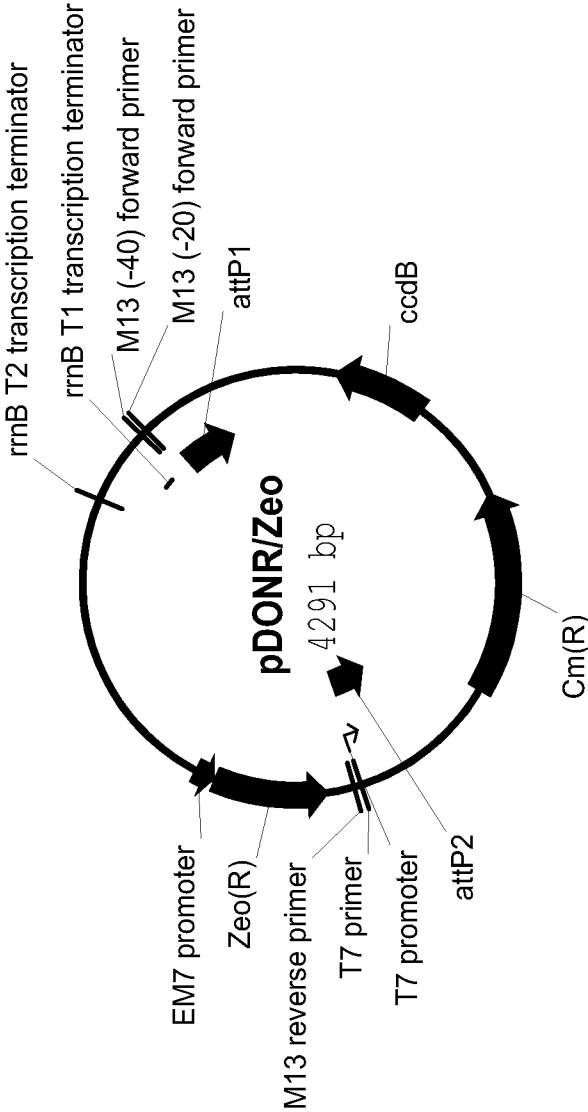


FIG. 3

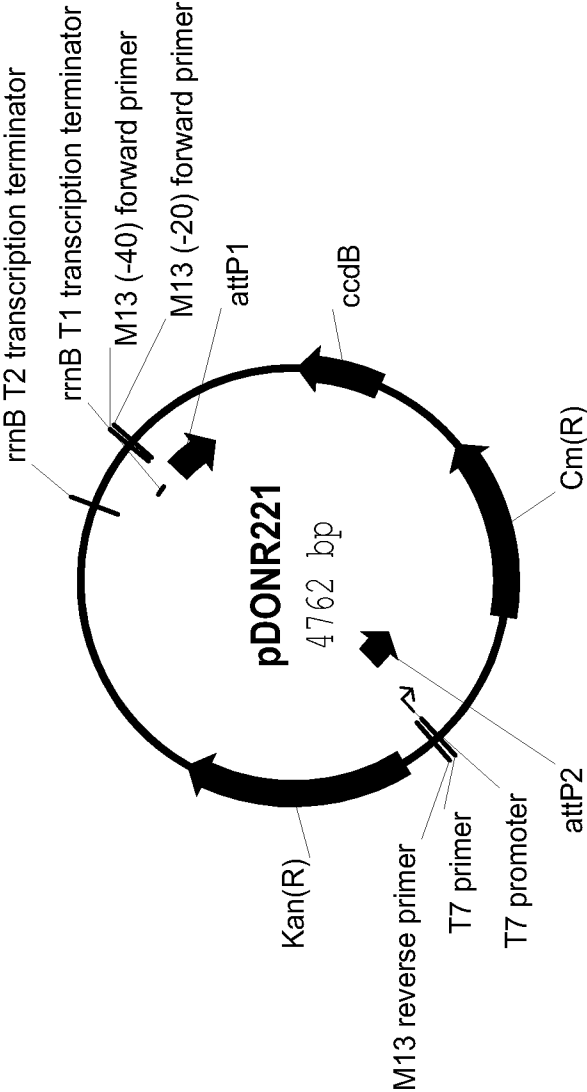
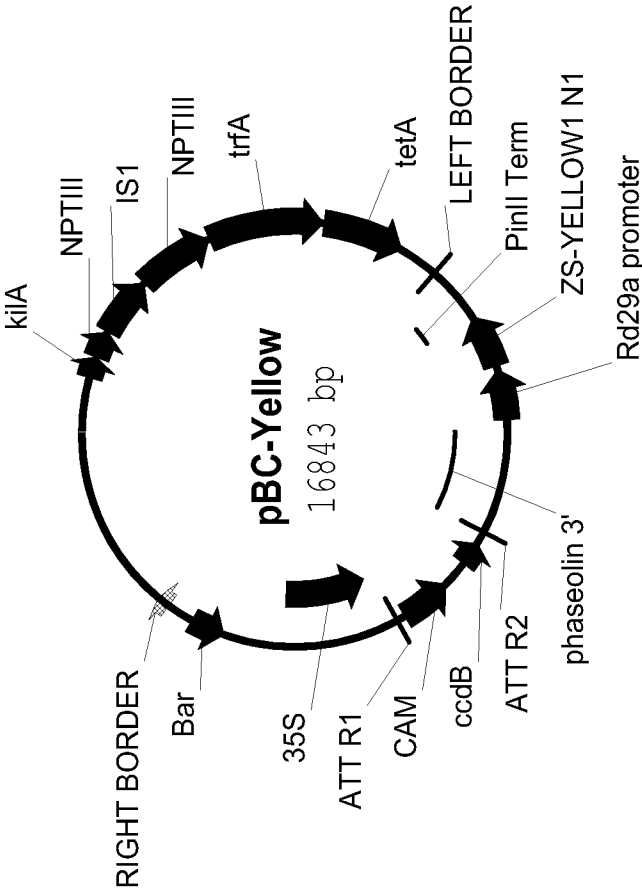
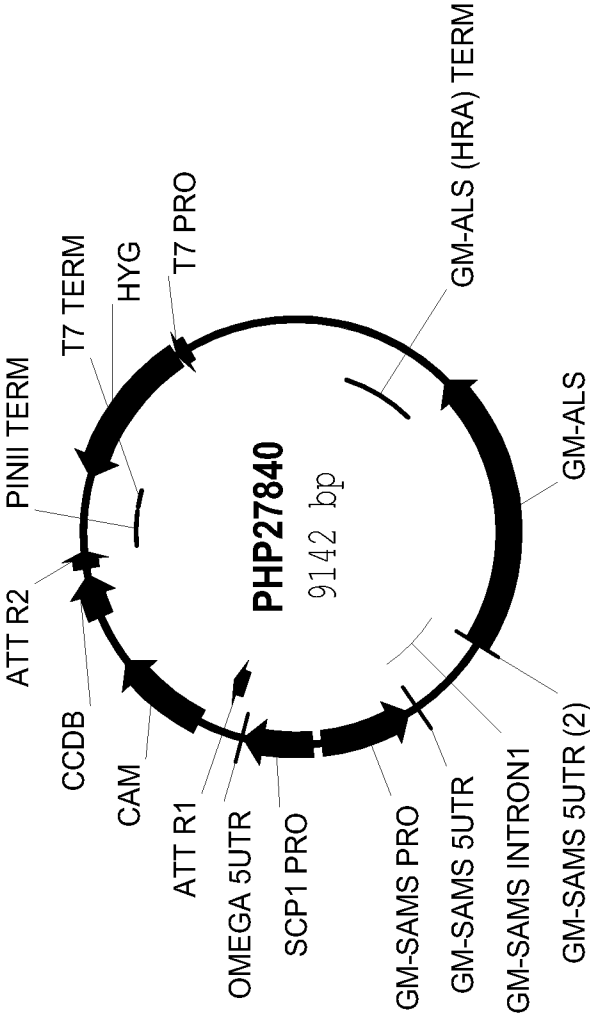


FIG. 4



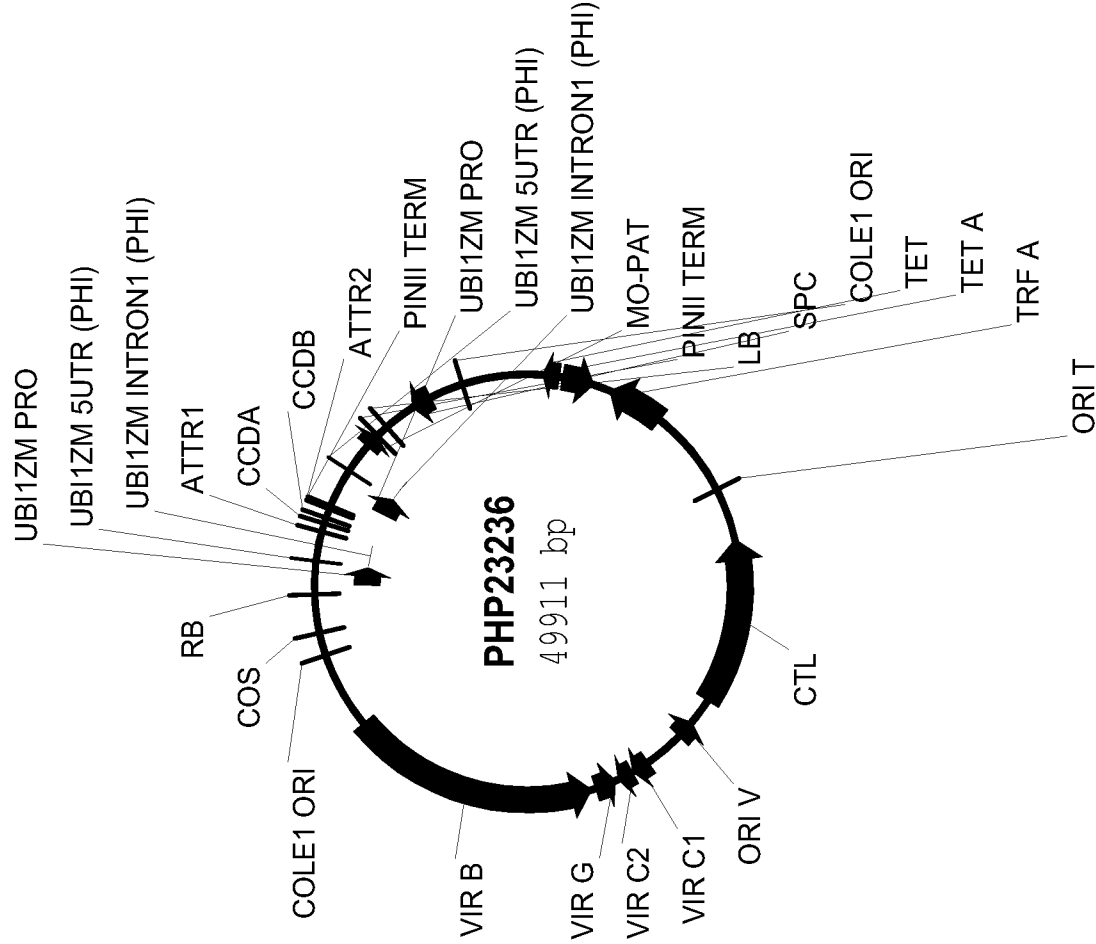
5/27

FIG. 5



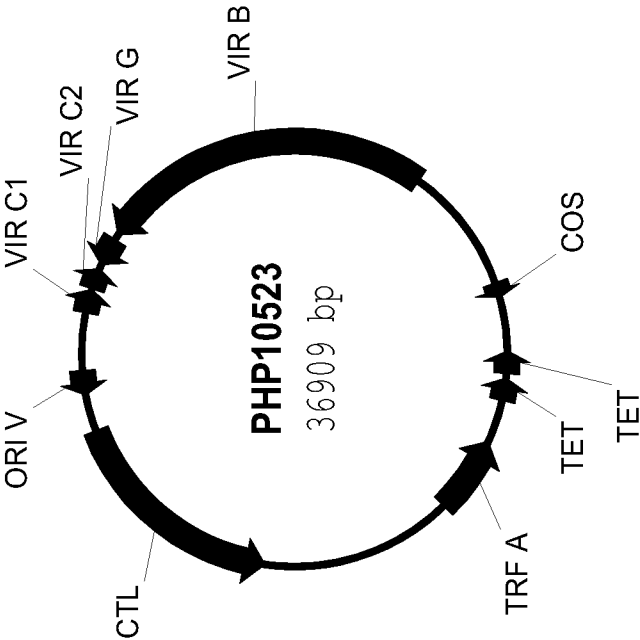
6/27

FIG. 6



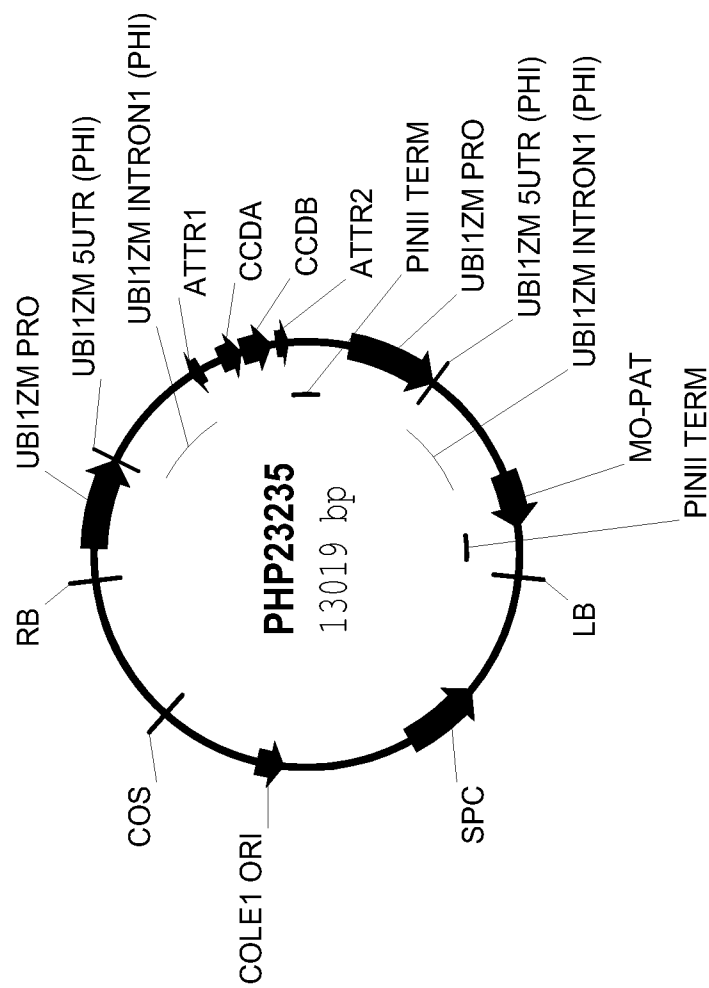
7/27

FIG. 7



8/27

FIG. 8



9/27

FIG. 9

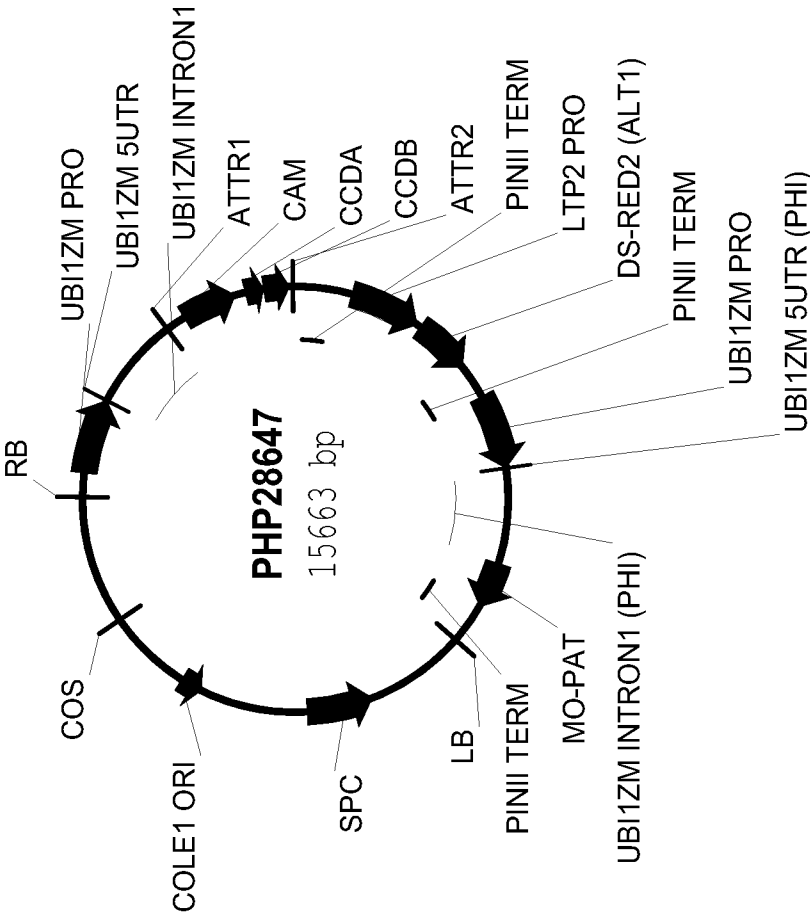


FIG. 10A

Multiple Alignment of Ferrochelatase-I Proteins

[illegible]

FIG. 10C
Multiple Alignment of Ferrochelatase-I Proteins

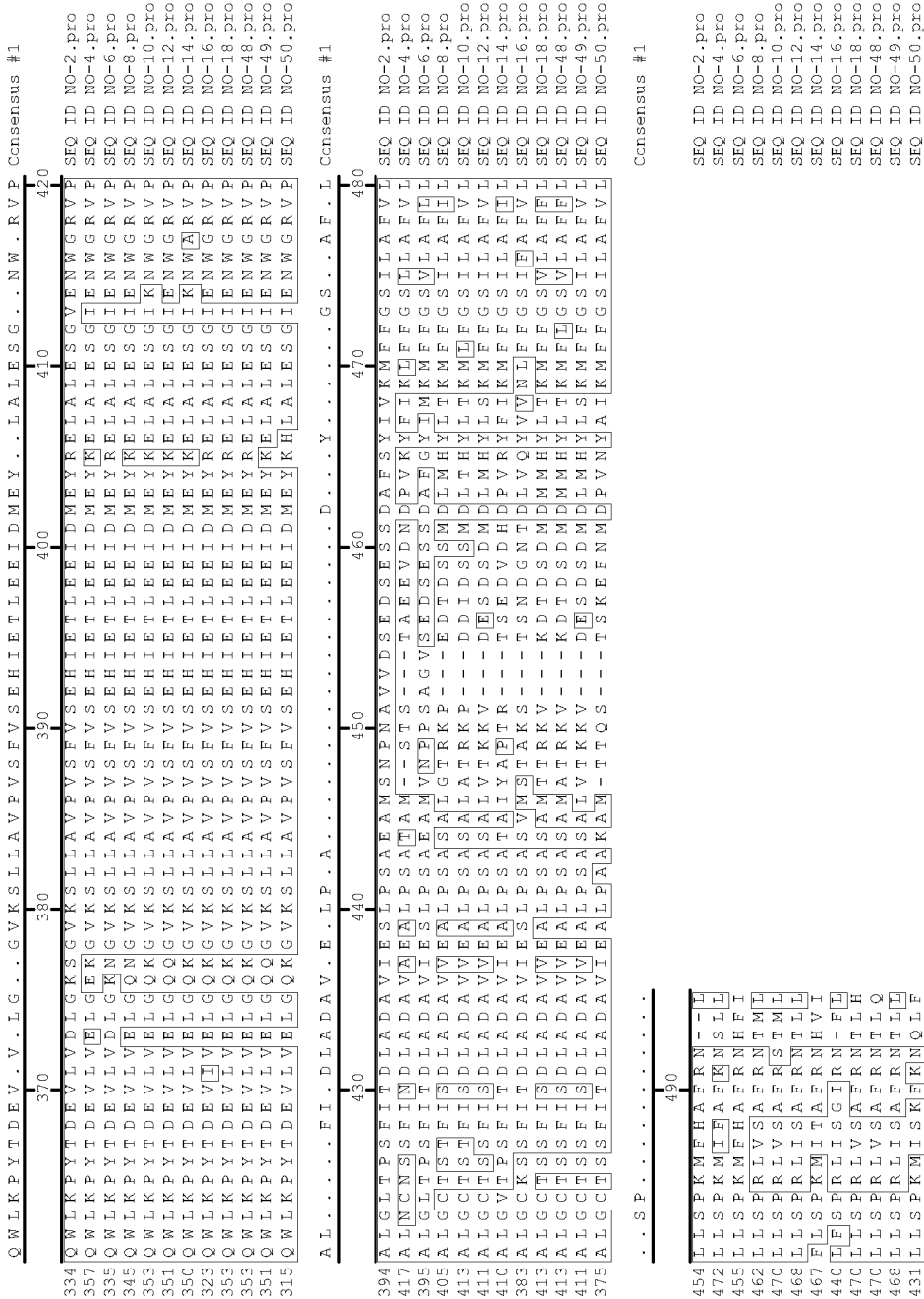


FIG. 11
Percent Sequence Identities Between Ferrochelatase-I Proteins

		Percent Identity												Divergence	
		1	2	3	4	5	6	7	8	9	10	11	12		
1	████	████	64.6	83.3	62.0	62.2	63.9	66.5	63.1	63.9	63.9	63.9	67.6	1	SEQ ID NO-2.pro
2	44.6	████	████	63.5	64.7	63.6	65.8	68.8	65.8	64.0	64.3	65.8	72.6	2	SEQ ID NO-4.pro
3	14.6	47.3	████	████	61.2	59.9	61.8	64.2	61.6	62.7	62.7	61.8	65.8	3	SEQ ID NO-6.pro
4	43.7	43.3	47.2	████	87.6	84.0	84.0	64.1	67.3	79.6	79.4	84.0	67.2	4	SEQ ID NO-8.pro
5	46.9	44.9	50.6	12.8	████	81.7	81.7	64.0	67.8	77.1	77.5	81.7	67.6	5	SEQ ID NO-10.pro
6	42.5	41.5	45.8	16.8	19.4	████	64.4	69.5	69.5	84.6	84.9	100.0	70.3	6	SEQ ID NO-12.pro
7	42.5	36.3	44.5	46.5	47.5	45.4	████	65.1	65.1	63.4	63.8	64.4	74.2	7	SEQ ID NO-14.pro
8	42.7	43.7	45.6	37.4	39.6	35.0	43.5	████	68.2	68.2	68.4	69.5	67.2	8	SEQ ID NO-16.pro
9	43.8	44.2	45.0	20.6	24.5	14.2	48.1	36.7	████	98.3	98.3	84.6	69.0	9	SEQ ID NO-18.pro
10	43.8	44.6	45.0	20.9	23.9	14.0	47.3	36.4	1.7	████	84.9	84.9	69.0	10	SEQ ID NO-48.pro
11	42.5	41.5	45.8	16.8	19.4	0.0	45.4	35.0	35.0	14.2	14.0	████	70.3	11	SEQ ID NO-49.pro
12	36.2	31.2	38.5	38.0	39.5	35.9	27.6	36.0	36.0	37.2	36.8	35.9	████	12	SEQ ID NO-50.pro
	1	2	3	4	5	6	7	8	9	10	11	12			

FIG. 12C

Multiple Alignment of Ferrochelatase-II Proteins

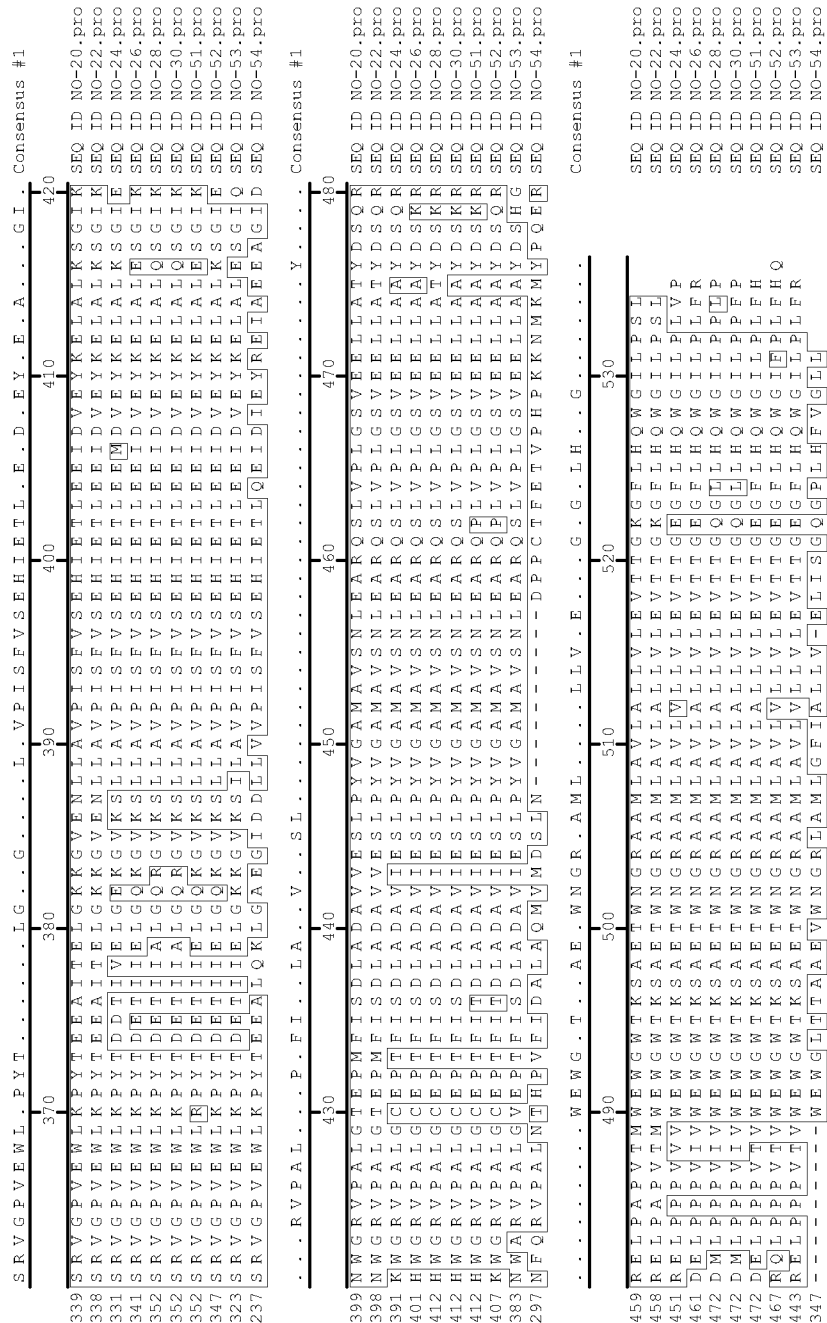


FIG. 13
Percent Sequence Identities Between Ferrochelatase-II Proteins

		Percent Identity										Divergence	
		1	2	3	4	5	6	7	8	9	10		
1	1	100	98.8	74.5	73.6	72.9	72.5	73.4	72.7	75.7	57.4		SEQ ID NO-20.pro
	2	0.6	100	73.1	72.4	71.6	71.2	72.2	71.6	73.8	56.3		SEQ ID NO-22.pro
	3	27.5	28.5	100	76.2	75.0	75.0	77.4	78.6	78.3	50.9		SEQ ID NO-24.pro
	4	29.8	30.8	23.5	100	82.7	83.1	90.5	74.6	77.3	56.1		SEQ ID NO-26.pro
	5	32.3	33.3	26.9	19.2	100	98.7	82.9	72.2	75.7	56.8		SEQ ID NO-28.pro
	6	33.5	34.5	26.9	18.7	1.3	100	83.3	72.4	75.9	56.6		SEQ ID NO-30.pro
	7	31.7	32.7	24.9	9.5	19.5	19.0	100	75.7	77.9	56.1		SEQ ID NO-51.pro
	8	31.7	32.1	23.0	29.7	34.2	33.3	28.3	100	77.1	55.3		SEQ ID NO-52.pro
	9	25.4	26.3	19.3	24.7	27.0	26.7	24.4	25.3	100	53.2		SEQ ID NO-53.pro
	10	54.6	54.8	59.9	57.7	56.7	57.2	57.7	58.8	57.2	100		SEQ ID NO-54.pro
		1	2	3	4	5	6	7	8	9	10		

18/27

FIG. 14A
Evaluation of Individual Gaspe Flint Derived Maize Lines Transformed with PHP31419*

Treatment	Event name	% area chg_start chronic - acute2	% area chg_start chronic - end chronic	% area chg_start chronic - harvest	% area chg_start chronic - recovery 24hr	fv/fm_acute1	fv/fm_acute2
Reduce Water	EA2392.498.1.2	ns	ns	ns	ns	ns	ns
Reduce Water	EA2392.498.1.3	ns	ns	ns	ns	0.089	ns
Reduce Water	EA2392.498.1.5	ns	ns	ns	ns	ns	ns
Reduce Water	EA2392.498.1.6	0.034	0.072	0.0004	0.079	0.014	0.003
Well Water	EA2392.498.1.2	ns	0.037	ns	0.085	ns	0.030
Well Water	EA2392.498.1.3	ns	ns	0.093	(0.032)	ns	ns
Well Water	EA2392.498.1.5	ns	0.083	ns	ns	0.016	ns
Well Water	EA2392.498.1.6	0.051	ns	ns	ns	0.001	ns

* Significant positive effect has P-value less than or equal to 0.1; significant negative effect is in parenthesis; "ns" when difference not significant

FIG. 14B
Evaluation of Individual Gaspe Flint Derived Maize Lines Transformed with PHP31419*

Treatment	Event name	leaf rolling_ harvest	leaf rolling_ recovery 24hr	psii_acute1	psii_acute2	sgr - r2 > 0.9	shoot dry weight	shoot fresh weight
Reduce Water	EA2392.498.1.2	0.065	ns	ns	ns	0.046	(0.028)	ns
Reduce Water	EA2392.498.1.3	ns	0.019	0.024	ns	ns	ns	(0.037)
Reduce Water	EA2392.498.1.5	ns	ns	ns	ns	ns	ns	ns
Reduce Water	EA2392.498.1.6	0.068	0.011	0.072	0.003	0.015	ns	0.009
Well Water	EA2392.498.1.2	ns	0.083	ns	ns	ns	ns	ns
Well Water	EA2392.498.1.3	ns	ns	ns	ns	ns	(0.047)	(0.031)
Well Water	EA2392.498.1.5	ns	ns	0.022	ns	ns	ns	ns
Well Water	EA2392.498.1.6	ns	ns	0.007	ns	ns	ns	ns

* Significant positive effect has P-value less than or equal to 0.1; significant negative effect is in parenthesis; "ns" when difference not significant

20/27

FIG. 16A
Evaluation of Individual Gaspe Flint Derived Maize Lines Transformed with PHP33089*

Treatment	Event name	% area chg_ start chronic - end chronic	% area chg_ start chronic - end severe	% area chg_ start chronic - recovery 48hr	fv/fm_end severe	fv/fm_recovery 48hr	fv/fm_start severe
Reduce Water	EA2534.088.2.2	ns	(0.076)	ns	(0.074)	ns	ns
Reduce Water	EA2534.088.2.4	0.095	Ns	ns	(0.025)	0.041	ns
Reduce Water	EA2534.088.2.5	0.068	0.063	ns	0.065	ns	0.036
Reduce Water	EA2534.088.2.8	ns	(0.012)	(0.032)	ns	ns	ns
Well Water	EA2534.088.2.2	ns	Ns	ns	ns	ns	ns
Well Water	EA2534.088.2.4	0.030	Ns	ns	ns	ns	ns
Well Water	EA2534.088.2.5	ns	Ns	ns	0.01	ns	ns
Well Water	EA2534.088.2.8	ns	Ns	ns	ns	(0.012)	ns

* Significant positive effect has P-value less than or equal to 0.1; significant negative effect is in parenthesis; "ns" when difference not significant

FIG. 16B
Evaluation of Individual Gaspe Flint Derived Maize Lines Transformed with PHP33089*

Treatment	Event name	leaf rolling_ recovery 48hr	Psii_end severe	psii_ recovery 48hr	psii_start severe	sgr - r2 > 0.9	shoot dry weight	shoot fresh weight
Reduce Water	EA2534.088.2.2	ns	(0.016)	ns	ns	ns	ns	ns
Reduce Water	EA2534.088.2.4	ns	ns	0.043	ns	ns	ns	ns
Reduce Water	EA2534.088.2.5	ns	0.008	ns	0.022	ns	ns	ns
Reduce Water	EA2534.088.2.8	ns	(0.011)	ns	ns	ns	0.005	ns
Well Water	EA2534.088.2.2	ns	ns	ns	0.058	ns	ns	ns
Well Water	EA2534.088.2.4	0.072	ns	(0.017)	(0.060)	ns	ns	ns
Well Water	EA2534.088.2.5	(0.051)	0.006	ns	(0.029)	ns	ns	ns
Well Water	EA2534.088.2.8	ns	ns	(0.048)	ns	ns	ns	ns

* Significant positive effect has P-value less than or equal to 0.1; significant negative effect is in parenthesis; "ns" when difference not significant

22/27

FIG. 18A
Evaluation of Individual Gaspe Flint Derived Maize Lines Transformed with PHP30829*

Treatment	Event name	% area chg_start chronic - end chronic	% area chg_start chronic - harvest	% area chg_start chronic - recovery 24hr	% area chg_start chronic - recovery 48hr	fv/fm acute1	fv/fm acute2
Reduce Water	EA2391.472.1.10	(0.013)	(0.002)	ns	(0.002)	ns	ns
Reduce Water	EA2391.472.1.2	ns	ns	ns	ns	ns	(0.028)
Reduce Water	EA2391.472.1.3	(0.036)	ns	ns	ns	ns	ns
Reduce Water	EA2391.472.1.4	ns	ns	ns	ns	ns	0.0022
Well Water	EA2391.472.1.10	ns	0.001	ns	ns	(0.007)	ns
Well Water	EA2391.472.1.2	ns	0.058	ns	ns	0.008	0.001
Well Water	EA2391.472.1.3	ns	0.059	ns	ns	ns	ns
Well Water	EA2391.472.1.4	ns	ns	ns	ns	0.006	0.095

* Significant positive effect has P-value less than or equal to 0.1; significant negative effect is in parenthesis; "ns" when difference not significant

FIG. 18B
Evaluation of Individual Gaspe Flint Derived Maize Lines Transformed with PHP30829*

Treatment	Event name	leaf rolling_ recovery 24hr	leaf rolling_ recovery 48hr	psii_acute1	psii_acute2	sgr - r2 > 0.9	shoot dry weight	shoot fresh weight
Reduce Water	EA2391.472.1.10	ns	(0.012)	ns	ns	(0.025)	ns	ns
Reduce Water	EA2391.472.1.2	ns	ns	ns	(0.021)	ns	ns	ns
Reduce Water	EA2391.472.1.3	ns	(0.096)	ns	ns	ns	(0.079)	(0.056)
Reduce Water	EA2391.472.1.4	(0.081)	ns	ns	0.012	(0.058)	ns	ns
Well Water	EA2391.472.1.10	ns	ns	(0.018)	(0.071)	0.034	ns	0.012
Well Water	EA2391.472.1.2	ns	0.069	0.026	0.016	ns	(0.046)	ns
Well Water	EA2391.472.1.3	0.091	ns	ns	ns	ns	ns	ns
Well Water	EA2391.472.1.4	ns	ns	0.04	ns	ns	ns	ns

* Significant positive effect has P-value less than or equal to 0.1; significant negative effect is in parenthesis; "ns" when difference not significant

FIG. 19
Summary Evaluation of Gaspe Flint Derived Maize Lines Transformed with PHP30829

Treatment	% area		% area		% area		chronic - - end	chronic - harvest	chronic - recovery 24hr	chronic - recovery 48hr	fv/fm_		leaf rolling_		leaf rolling_		psii_		sgr - r2		shoot	
	chronic	start	chronic	start	chronic	start					acute1	acute2	recovery 24hr	recovery 48hr	recovery 24hr	recovery 48hr	acute1	acute2	> 0.9	dry weight	fresh weight	weight
Reduce																						
Water	(0.056)		ns		ns		(0.087)		ns		ns		ns	(0.087)		ns		ns	(0.056)		ns	ns
Well																						
Water	0.032		0.0002		ns		ns		0.027		ns		ns	ns		ns		ns	0.081		ns	ns

* Significant positive effect has P-value less than or equal to 0.1; significant negative effect is in parenthesis; "ns" when difference not significant

24/27

FIG. 20A
Evaluation of Individual Gaspe Flint Derived Maize Lines Transformed with PHP30745*

Treatment	Event name	% area start chronic - end chronic	% area chg_start chronic - end severe	% area chg_start chronic - recovery 48hr	fv/fm_end severe	fv/fm_recovery 48hr	fv/fm_start severe
Reduce Water	EA2392.447.1.1	ns	ns	ns	(0.071)	0.039	ns
Reduce Water	EA2392.447.1.3	ns	ns	ns	ns	(0.087)	ns
Reduce Water	EA2392.447.1.5	ns	ns	ns	ns	ns	ns
Reduce Water	EA2392.447.1.9	(0.059)	ns	ns	(0.091)	ns	0.003
Well Water	EA2392.447.1.1	0.003	0.007	0.007	ns	ns	ns
Well Water	EA2392.447.1.3	(0.054)	(0.037)	(0.025)	ns	(0.004)	(0.001)
Well Water	EA2392.447.1.5	0.098	ns	ns	ns	ns	ns
Well Water	EA2392.447.1.9	0.001	0.002	0.005	ns	ns	ns

* Significant positive effect has P-value less than or equal to 0.1; significant negative effect is in parenthesis; "ns" when difference not significant

FIG. 20B
Evaluation of Individual Gaspe Flint Derived Maize Lines Transformed with PHP30745*

Treatment	Event name	leaf rolling_ recovery 48hr	psii_end severe	psii_recovery 48hr	psii_start severe	shoot dry weight	shoot fresh weight
Reduce Water	EA2392.447.1.1	ns	ns	0.01	0.054	ns	0.006
Reduce Water	EA2392.447.1.3	ns	ns	ns	ns	(0.028)	(0.081)
Reduce Water	EA2392.447.1.5	ns	ns	ns	ns	ns	ns
Reduce Water	EA2392.447.1.9	ns	ns	ns	0.019	(0.096)	(0.012)
Well Water	EA2392.447.1.1	ns	ns	ns	ns	0.051	ns
Well Water	EA2392.447.1.3	ns	ns	(0.051)	(0.0003)	0.0012	0.003
Well Water	EA2392.447.1.5	ns	ns	0.002	ns	(0.026)	ns
Well Water	EA2392.447.1.9	ns	ns	(0.014)	ns	ns	0.066

* Significant positive effect has P-value less than or equal to 0.1; significant negative effect is in parenthesis; "ns" when difference not significant

FIG. 21
Summary Evaluation of Gaspe Flint Derived Maize Lines Transformed with PHP30745*

Treatment	% area		% area		% area		chg_start		chg_end		chronic - recovery		chronic - severe		fv/fm_ recovery 48hr		fv/fm_ end severe		fv/fm_ start severe		leaf rolling_ recovery 48hr		psii_ end severe		psii_ recovery 48hr		psii_ start severe		shoot dry weight		shoot fresh weight	
	chg_start	chg_end	chg_start	chg_end	chg_start	chg_end	chg_start	chg_end	chg_start	chg_end	chg_start	chg_end	chg_start	chg_end	chg_start	chg_end	chg_start	chg_end	chg_start	chg_end	chg_start	chg_end	chg_start	chg_end	chg_start	chg_end	chg_start	chg_end	chg_start	chg_end	chg_start	chg_end
Reduce																																
Water	ns		ns		ns		ns		ns		ns		ns		ns		ns		0.081		ns		ns		0.035		0.091		ns		ns	
Well																																
Water	ns		ns		ns		ns		ns		ns		ns		ns		0.034		ns		ns		ns		ns		ns		ns		ns	

* Significant positive effect has P-value less than or equal to 0.1; significant negative effect is in parenthesis; "ns" when difference not significant

FIG. 22A
Evaluation of Individual Gaspe Flint Derived Maize Lines Transformed with PHP30761*

Treatment	Event name	% area chg_ start chronic - end chronic	% area chg_ start chronic - end severe	% area chg_ start chronic - recovery 48hr	fv/fm_end severe	fv/fm_recovery 48hr	fv/fm_start severe
Reduce Water	EA2392.441.1.1	ns	ns	ns	ns	ns	ns
Reduce Water	EA2392.441.1.2	ns	ns	ns	ns	ns	ns
Reduce Water	EA2392.441.1.4	ns	ns	ns	0.02	0.038	ns
Reduce Water	EA2392.441.1.7	(0.034)	(0.054)	ns	0.047	0.042	ns
Well Water	EA2392.441.1.1	ns	ns	ns	ns	ns	ns
Well Water	EA2392.441.1.2	ns	ns	ns	ns	ns	0.03
Well Water	EA2392.441.1.4	(0.004)	ns	ns	0.000	0.008	0.001
Well Water	EA2392.441.1.7	ns	ns	ns	ns	ns	ns

* Significant positive effect has P-value less than or equal to 0.1; significant negative effect is in parenthesis; "ns" when difference not significant

FIG. 22B
Evaluation of Individual Gaspe Flint Derived Maize Lines Transformed with PHP30761*

Treatment	Event name	leaf rolling_ recovery 48hr	psii_end severe	psii_ recovery 48hr	psii_start severe	sgr - r2 > 0.9	shoot dry weight	shoot fresh weight
Reduce Water	EA2392.441.1.1	ns	ns	ns	ns	ns	ns	ns
Reduce Water	EA2392.441.1.2	ns	ns	ns	ns	ns	ns	ns
Reduce Water	EA2392.441.1.4	0.058	0.071	0.072	ns	ns	ns	ns
Reduce Water	EA2392.441.1.7	ns	0.098	0.037	0.068	ns	(0.001)	ns
Well Water	EA2392.441.1.1	ns	ns	ns	ns	ns	(0.048)	ns
Well Water	EA2392.441.1.2	ns	ns	ns	ns	ns	0.1	ns
Well Water	EA2392.441.1.4	0.001	0.002	0.000	0.006	(0.0002)	(0.050)	ns
Well Water	EA2392.441.1.7	ns	ns	ns	ns	ns	ns	ns

* Significant positive effect has P-value less than or equal to 0.1; significant negative effect is in parenthesis; "ns" when difference not significant

FIG. 23

Summary Evaluation of Gaspe Flint Derived Maize Lines Transformed with PHP30761*

Treatment	% area chg_start			% area chg_end			% area chg_recovery			% area chg_severe			% area chg_weight		
	chronic	chronic - end	chronic severe	chronic	chronic - end	chronic severe	chronic	chronic - end	chronic severe	chronic	chronic - end	chronic severe	chronic	chronic - end	chronic severe
Reduce Water Well	(0.050)	ns	ns	ns	0.065	0.000	ns	ns	ns	ns	0.024	ns	ns	ns	ns
Water Well	(0.001)	(0.063)	ns	ns	ns	ns	ns	ns	0.032	ns	0.071	(0.002)	(0.005)	ns	ns

* Significant positive effect has P-value less than or equal to 0.1; significant negative effect is in parenthesis; "ns" when difference not significant