

(51) International Patent Classification:
C12N 15/82 (2006.01) A01H 1/00 (2006.01)(21) International Application Number:
PCT/US2010/037391(22) International Filing Date:
4 June 2010 (04.06.2010)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/184,588 5 June 2009 (05.06.2009) US(71) Applicant (for all designated States except US):
MENDEL BIOTECHNOLOGY, INC. [US/US]; 3935 Point Eden Way, Hayward, California 94545 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **ARMSTRONG, Joshua I.** [US/US]; 322 24th Avenue, San Mateo, California 94403 (US). **TIWARI, Shiv B.** [IN/US]; 1501 Decoto Road, Apt. 196, Union City, California 94587 (US). **RATCLIFFE, Oliver J.** [GB/US]; 814 East 21st Street, Oakland, California 94606 (US).(74) Agents: **LIBBY, Jeffrey** et al.; Mendel Biotechnology, Inc., 3935 Point Eden Way, Hayward, California 94545 (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, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

- with international search report (Art. 21(3))
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))
- with sequence listing part of description (Rule 5.2(a))

(54) Title: REGULATION OF GENE EXPRESSION WITH THE CONSTANS RESPONSE ELEMENT

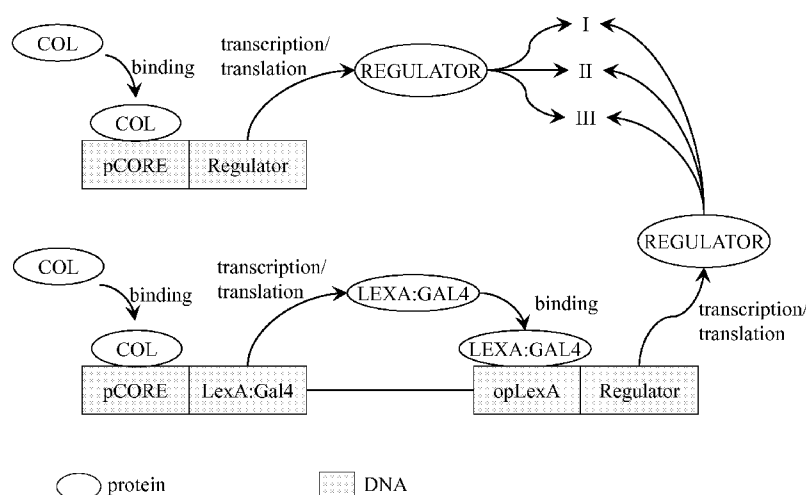


FIG. 4

(57) **Abstract:** CORE promoter sequences were identified that respond to activation by CONSTANS or CONSTANS-like (COL) proteins and which can be used to regulate gene expression in response to those proteins. These promoters may be used to produce transgenic plants that have an altered trait relative to control plants, particularly in response to floral cues during or after flowering stages.. Any of these CORE promoters may be incorporated into a nucleic acid construct to regulate the expression of a polynucleotide that encodes a polypeptide that, when ectopically expressed, confers a desired or improved trait in plants.

REGULATION OF GENE EXPRESSION WITH THE
CONSTANS RESPONSE ELEMENT

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims benefit of U.S. provisional application no. 61/184,588, filed June
5 5, 2009, the entire content of which is herein incorporated by reference.

FIELD OF THE INVENTION

The present invention relates to plant genomics, and more specifically to promoters that mediate gene expression in a plant.

BACKGROUND OF THE INVENTION

10 The transition of plants from a vegetative state to a reproductive stage is complex, involving alterations on multiple levels. In many plant species, this transition is controlled by day length, which is perceived in leaves and induces a systemic signal, called florigen, which moves through the phloem to the shoot apex (Turck et al., 2008). Flowering is integrated with several physiological pathways including nutrient sensing (sugar signaling), development and
15 the perception of environmental stress. The timing of the floral transition has a significant impact on both biomass production and grain yields, with early flowering sometimes being associated with reduced grain production and plant biomass. In contrast, late flowering may prevent the development of hybrid lines, impeding breeding of new out-crossed plant varieties of agricultural interest. Additionally, many crop species show an increased sensitivity to
20 environmental stresses such as water deficit during the reproductive versus the vegetative phase of the life cycle (Edmeades et al., 2000). The ability to manipulate directly the timing of the floral transition or to deliver transgene products that confer stress tolerance during particular developmental stages could therefore have direct implications on the yield of bioenergy and food crops.

25 Currently, the manipulation of the floral transition can be achieved via the genetic regulation of critical components of the flowering pathway. Current transgenic technologies for enhanced stress tolerance typically rely on constitutive promoters or tissue-specific promoters driving ectopic expression of a gene (or genes). For instance, there are numerous examples in the art whereby constitutive, high-level expression or knock-out or knock-down of *CONSTANS*
30 (*CO*), *FLOWERING LOCUS C* (*FLC*), *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1* (*SOC1*) and other genes result in dramatic changes in the timing of the floral transition (Ratcliffe and Riechmann, 2002). Of the aforementioned genes, the transcription factor

CONSTANS has a critical role in the triggering of flowering in response to photoperiod (Putterill et al., 2004). CO is a member of a family of related transcription factors, which are conserved across plant species, and which are known as CONSTANS LIKE PROTEINS (COL); in Arabidopsis the COL family comprises around 17 members (Robson et al., 2001). However, these shifts in flowering time may also be accompanied by undesirable alterations in morphology and physiology due to side effects on developmental processes such as organ growth. Similar results have been obtained with the constitutive, high-level expression of genes associated with stress tolerance, where a beneficial resistance to stress is observed but which is often coupled with undesirable effects on development, morphology and yield (Zhang et al., 2004). Tissue-specific promoters or inducible promoters are used to limit the expression of the target gene(s) to specific cell-types or to regulate the expression to specific stimuli. However, these promoters are often not fine-tunable such that they deliver enriched expression of the genes that they control during specific phases of development, e.g., post-flowering.

The present application describes the identification of key genetic elements involved in the signaling cascade that culminates in the transition from a vegetative to a reproductive growth stage (the floral transition). The promoter elements described in the present invention enable the regulation of target gene expression such that it is enriched during the floral transition, a key advantage for minimizing off-types resulting from high-level constitutive expression of a transgene and ensuring delivery of the desired trait during the decisive developmental stage. In particular, most row crops, especially maize and soy, are most susceptible to grain yield losses as a result of environmental stress following the floral transition. The system provided herein can offer the ability to deliver enriched expression of transgenes that confer stress tolerance at these later stages and thereby avoid any negative effects that would arise from the high-levels constitutive expression of such protective proteins throughout the lifecycle of the plants. The invention described herein also offers methods for granular control over the floral transition via the synchronized activation/repression of the floral transition using a feed-forward pathway that amplifies the floral transition signal. Numerous examples of how this element can be employed to regulate flowering and control gene expression in response to floral induction cues are provided.

The present invention thus relates to methods and compositions for producing transgenic plants, where the plants are transformed with nucleic acid construct containing a promoter sequence that regulate target gene expression during a specific and fundamental developmental shift (the floral transition). The other aspects and embodiments are described below and can be derived from the teachings of this disclosure as a whole.

SUMMARY OF THE INVENTION

The present invention is directed to isolated polynucleotide sequences that comprise promoter sequences and promoter control elements that may be used to transform a plant.

Promoter or promoter control element sequences of the present invention are capable of
 5 modulating preferential transcription.

The invention provides an isolated polynucleotide comprising a promoter sequence that includes any of the promoter sequences provided by SEQ ID NOs: 101, 102, 1 - 4, 7-10, 13- 21 . The functional part of the promoter that is capable of regulating transcription when operably linked to a transcribable polynucleotide may have about 8, 9, 10, 13, 17, 20, 30, 35, 40, 50, 75,
 10 100, 125, 150, 191, 200, 251, 311, 491, 496, 525, 550, 575, 600, 650, 700, 750, 800, 900, 950, 1000, 1050, 1100, 1150, 1200, 1500, 1633, 1639 or 1800 contiguous nucleotides of the nucleic acid sequences of SEQ ID NOs: 1, 14, 15, or 21, as well as all lengths of contiguous nucleotides within such sizes, or have multimeric sequences selected from of SEQ ID NO: 101, 102, 7-10, 13, 16-20 or from any combinations thereof.

The invention is also directed to an isolated polynucleotide that comprises a promoter sequence or promoter control element, where the promoter sequence or promoter control element has at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or
 20 100% sequence identity to SEQ ID NOs: 101, 102, 1 - 4, 7-10, 13- 21 . The promoter sequences comprising one or more copies of the polynucleotide SEQ ID NO: 16 or its complementary sequences are referred to as CORE promoters (CORE REGULATORY ELEMENT PROMOTERS) in the present application. These promoter sequences, upon binding by active COL polypeptides accumulated during floral transition stage, can regulate the transcription of a
 25 polynucleotide that confer an improved trait to a plant during and following the floral transition stage. For example, the CORE promoters can be used to regulate the expression of a polynucleotide sequence that encodes a polypeptide that can confer an improved and/or desirable trait during and following the floral transition. Thus, expression of the polypeptide may be enriched in a specific developmental stage during or after flowering.

The invention also pertains to nucleic acid constructs (for example, expression vectors or expression cassettes) that comprise a CORE promoter sequence. The CORE promoter may comprise any of SEQ ID NOs: 101, 102, 1 - 4, 7-10, 13- 21, or a functional part thereof, provided the functional part also includes a CORE promoter function. The promoter comprises a transcription initiation domain having an RNA polymerase binding site. In one embodiment, the

present invention provides a nucleic acid construct that comprises a promoter as described above, located 5' relative to and operably linked to a coding sequence encoding a polypeptide that confers to a plant desired gene and/or protein regulation upon COL binding, such as in response to an inductive photoperiod. In another embodiment, the present invention also provides a two-component system to drive higher levels of regulator or reporter gene expression following activation of the CORE motif. For example, the CORE promoter could drive expression of a translational fusion of the LexA DNA binding domain and the GAL4 transcriptional activation domain. A desired regulator sequence or a reporter gene would be driven from the opLexA promoter such that activation of the CORE promoter would result in LEXA:GAL4-mediated expression of the regulator or the reporter gene.

This alternative activation system may deliver more consistent and high levels of the "regulator" protein. Nucleic acid constructs that comprise a promoter of any of SEQ ID NOs: 101, 102, 1 - 4, 7-10, 13- 21, may be introduced into plants, and as a result of the promoter's regulatory control of a gene of interest, the plants may have an improved or desirable trait relative to a control plant. In some cases, the transformed plants may show wild-type or near-wild type morphology and development. This may be of significant utility in that many polypeptides that confer improved traits upon their expression can also cause undesirable morphological and/or developmental traits when the polypeptides are constitutively expressed at high levels. Non-constitutive regulation of expression, such as by regulation of flowering transitional signal induced by a lengthening photoperiod, may be used to confer the improved traits while mitigating the undesirable morphological and/or developmental effects.

The present invention is also directed to a host plant cell comprising a CORE promoter, comprising any of SEQ ID NOs: 101, 102, 1 - 4, 7-10, 13- 21 or a functional part thereof, wherein the functional part includes a promoter function.

The invention also is also directed to a transgenic plant comprising a CORE promoter, comprising any of SEQ ID NOs: 101, 102, 1 - 4, 7-10, 13- 21 or a functional part thereof, wherein the functional part includes a promoter function, and transgenic seed produced by the transgenic plant.

The present invention also provides methods for producing a transgenic plant having regulated transgene expression that leads to altered traits in response to active COL proteins, relative to a control plant provided. In one embodiment, the method steps include the generation of a nucleic acid construct (e.g., an expression vector or cassette) that comprises a promoter sequence of any of SEQ ID NOs: 101, 102, 1 - 4, 7-10, 13- 21 or a functional part thereof, wherein the functional part includes a promoter function that responds to COL proteins. In one

embodiment, the promoter sequence is operably linked to a nucleotide sequence that encodes a polypeptide that improves a trait, or produces a desirable trait, in a plant, and the promoter sequence drives expression of the nucleotide sequence that encodes the polypeptide during or after the floral stage. A target plant is then transformed with the nucleic acid construct to produce a transgenic plant. When the polypeptide is overexpressed in the transformed plant in response to the presence of active COL proteins, such as in response to inductive photoperiods, the transformed plant will exhibit the desired trait relative to the control plant. A transgenic plant that is produced by this method may be crossed with itself, a plant from the same line as the transgenic plant, a non-transgenic plant, a wild-type plant, or another transgenic plant from a different transgenic line of plants, to produce a transgenic seed that comprises the expression vector. In a second embodiment, a first nucleic acid construct harboring a selection marker and comprised of a CORE promoter and operably linked to the LexA:Gal4 expression cassette (SEQ ID NO: 22) is transformed into a target plant 1. A second nucleic acid construct harboring a different selection marker and comprised of the opLexA promoter sequence (SEQ ID NO: 24) operably linked to a desired regulator coding sequence is transformed into plant 2. A transgenic plant comprising both expression constructs may be selected after crossing transgenic plants 1 and 2. The CORE promoter sequence thus drives the expression of a transcriptional activation complex LEXA:GAL4, which induces expression of any desired “regulator in response to inductive day length. This two-component activation system may deliver more consistent and high levels of the “regulator” protein to thus selected transgenic plants.

This application provides key genetic elements involved in the signaling cascade that culminates during the floral transition. Numerous examples of how these elements can be employed to regulate flowering and control gene expression in response to floral induction cues are provided in the following description.

Brief Description of the Sequence Listing and Drawings

The Sequence Listing provides exemplary polynucleotide and polypeptide sequences. The traits associated with the use of the sequences are included in the Examples.

Incorporation of the Sequence Listing. The Sequence Listing provides exemplary polynucleotide and polypeptide sequences. The copy of the Sequence Listing, being submitted electronically with this patent application, provided under 37 CFR §1.821-1.825, is a read-only memory computer-readable file in ASCII text format. The Sequence Listing is named “MBI-0092PCT_ST25.txt”, the electronic file of the Sequence Listing was created on May 17, 2010, and is 220,112 bytes in size (measured in MS-WINDOWS). The Sequence Listing is herein incorporated by reference in its entirety.

Figure 1A illustrates the promoter analysis strategy leading to the identification of the CONSTANS REGULATORY ELEMENTS (CORE) motif. Progressively shorter fragments of the FT promoter region (FT1, FT2, FT3, FT4, FT5 and FT6) were used to identify candidate CO binding sequences CORE1 and CORE2; comparison with the SOC1 promoter enabled the
 5 identification of a conserved CORE motif, indicated in boxes. Polynucleotide start and end positions relative to the start codons are indicated in parentheses after the sequence name. SEQ ID NOs of sequences are also found within the parentheses. Figure 1B shows that a conserved sequence, SEQ ID NO: 87, which is complementary to the CORE motif, is present in the PAP2 and AOX1 promoters. The conserved sequences are found within the boxes.

10 Figures 2A, 2B and 2C show a comparison of synthetic promoter elements that comprise multiple copies of the CORE1 or CORE2 sequences (or mutated versions thereof) and a hybrid promoter sequence comprised of multimeric copies of both the CORE1 and CORE2 motifs (CORE3). Asterisks indicate the positions where the nucleotides differ among the aligned sequences. The broken lines in Figs. 2B and 2C represent the regions where the nucleotide
 15 sequence is absent. These multimeric sequences were used to characterize the CORE motif. SEQ ID NOs of sequences are found within the parentheses after the sequence names.

Figure 3 shows direct binding of the CONSTANS (CO) protein with the CORE motif in an electromobility shift assay (EMSA). The EMSA experiment used the 4XCORE2B sequence and an epitope tagged variant of the CO protein. CO protein was bound specifically to the
 20 4XCORE2B sequence (middle lane). The CO protein failed to bind to the 4XCORE2BM1 sequence with a mutated CORE motif (TGTG to TATA) (the last lane). SEQ ID NOs of sequences are found within the parentheses next to the sequence names. This result confirms that the CORE motif is a direct target of CO and plays a critical role in the regulation of the FT promoter and flowering.

25 Figure 4 schematically represents delivery of a desired trait during and following the floral transition in response to presence of a COL protein in one or more plant cells, harboring a DNA construct containing the CORE promoter. Two possible schematics are presented for the activation of a “regulator” gene that confers a desired trait. pCORE refers to promoter sequence containing one or more copies of the CORE motif. Regulator refers to a gene encoding a protein
 30 (for example, a transcription factor) that regulates a downstream genetic pathway. COL refers to a member of the family of transcription factors that are homologous and phylogenetically related to CONSTANS (CO), SEQ ID NO: 61. LEXA:GAL4 is a translational fusion of the LEXA DNA binding and GAL4 transcriptional activation protein domains. OpLEXA is a promoter sequence which is bound by the LEXA protein. I, II and III represent the examples of the desired

traits that may be conferred by the regulators; they are: biotic/abiotic stress tolerance, promotion of flowering and inhibition of flowering.

DETAILED DESCRIPTION

The present invention relates to polynucleotides and polypeptides for modifying plant phenotypes, particularly promoter sequences recognized by transcription factors of the CONSTANS (CO) and CONSTANS-like (COL) family of proteins, and which may regulate an improved trait in transgenic plants with respect to a control plant at an important developmental stage, the floral transition stage. COL family proteins have been identified in a wide range of plants (Robson et al., 2001; Millar et al., 2008; Chia et al., 2008; Nemoto et al., 2008). Examples of control plants include, for example, genetically unaltered or non-transgenic plants such as wild-type plants of the same species, or non-transformed plants, or transgenic plant lines that comprise an empty expression vector. The present invention also provides methods of modifying, producing and using the same. Throughout this disclosure, various information sources are referred to and/or are specifically incorporated. The information sources include scientific journal articles, patent documents, textbooks, and World Wide Web page addresses. While the reference to these information sources clearly indicates that they can be used by one of skill in the art, each and every one of the information sources cited herein are specifically incorporated in their entirety, whether or not a specific mention of "incorporation by reference" is noted. The contents and teachings of each and every one of the information sources can be relied on and used to make and use embodiments of the invention.

As used herein and in the appended claims of the invention, the singular forms "a", "an", and "the" include the plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "a stress" is a reference to one or more stresses and equivalents thereof known to those skilled in the art, and so forth.

DEFINITIONS

"Nucleic acid molecule" refers to an oligonucleotide, polynucleotide or any fragment thereof. It may be DNA or RNA of genomic or synthetic origin, double-stranded or single-stranded, and combined with carbohydrate, lipids, protein, or other materials to perform a particular activity such as transformation or form a useful composition such as a peptide nucleic acid (PNA).

"Polynucleotide" is a nucleic acid molecule comprising a plurality of polymerized nucleotides, e.g., at least about 15 consecutive polymerized nucleotides. A polynucleotide may

be a nucleic acid, oligonucleotide, nucleotide, or any fragment thereof. In many instances, a polynucleotide comprises a nucleotide sequence encoding a polypeptide (or protein) or a domain or fragment thereof. Additionally, the polynucleotide may comprise a promoter, an intron, an enhancer region, a polyadenylation site, a translation initiation site, 5' or 3' untranslated regions, a reporter gene, a selectable marker, or the like. The polynucleotide can be single-stranded or double-stranded DNA or RNA. The polynucleotide optionally comprises modified bases or a modified backbone. The polynucleotide can be, e.g., genomic DNA or RNA, a transcript (such as an mRNA), a cDNA, a PCR product, a cloned DNA, a synthetic DNA or RNA, or the like. The polynucleotide can be combined with carbohydrate, lipids, protein, or other materials to perform a particular activity such as transformation or form a useful composition such as a peptide nucleic acid (PNA). The polynucleotide can comprise a sequence in either sense or antisense orientations. "Oligonucleotide" is substantially equivalent to the terms amplicon, primer, oligomer, element, target, and probe and is preferably single-stranded.

A "recombinant polynucleotide" is a polynucleotide that is not in its native state, e.g., the polynucleotide comprises a nucleotide sequence not found in nature, or the polynucleotide is in a context other than that in which it is naturally found, e.g., separated from nucleotide sequences with which it typically is in proximity in nature, or adjacent (or contiguous with) nucleotide sequences with which it typically is not in proximity. For example, the sequence at issue can be cloned into a vector, or otherwise recombined with one or more additional nucleic acid.

An "isolated polynucleotide" is a polynucleotide, whether naturally occurring or recombinant, that is present outside the cell in which it is typically found in nature, whether purified or not. Optionally, an isolated polynucleotide is subject to one or more enrichment or purification procedures, e.g., cell lysis, extraction, centrifugation, precipitation, or the like.

"Gene" or "gene sequence" refers to the partial or complete coding sequence of a gene, its complement, and its 5' or 3' untranslated regions. A gene is also a functional unit of inheritance, and in physical terms is a particular segment or sequence of nucleotides along a molecule of DNA (or RNA, in the case of RNA viruses) involved in producing a polypeptide chain. The latter may be subjected to subsequent processing such as chemical modification or folding to obtain a functional protein or polypeptide. A gene may be isolated, partially isolated, or found with an organism's genome. By way of example, a transcription factor gene encodes a transcription factor polypeptide, which may be functional or require processing to function as an initiator of transcription.

Operationally, genes may be defined by the cis-trans test, a genetic test that determines whether two mutations occur in the same gene and that may be used to determine the limits of

the genetically active unit (Rieger et al. (1976)). A gene generally includes regions preceding (“leaders”; upstream) and following (“trailers”; downstream) the coding region. A gene may also include intervening, non-coding sequences, referred to as “introns”, located between individual coding segments, referred to as “exons”. Most genes have an associated promoter region, a regulatory sequence 5' of the transcription initiation codon (there are some genes that do not have an identifiable promoter). The function of a gene may also be regulated by enhancers, operators, and other regulatory elements.

A “promoter” or “promoter region” or “promoter sequence” refers to a segment of DNA containing an RNA polymerase binding site, generally found upstream or 5' relative to a coding sequence under the regulatory control of the promoter. The promoter will generally comprise response elements that are recognized by transcription factors. Transcription factors bind to the promoter sequences, recruiting RNA polymerase, which synthesizes RNA from the coding region. Dissimilarities in promoter sequences account for different efficiencies of transcription initiation and hence different relative expression levels of different genes.

A “COL polypeptide or COL protein” refers to a member of the family of transcriptional regulators with homology and phylogenetic relatedness to the CONSTANS protein, as exemplified in Robson et al., 2001.

A “COL” refers to a polynucleotide that encodes a COL polypeptide.

“CORE promoter” herein refers to a promoter sequence comprising at least one CORE response element (SEQ ID NO: 16 or the complement thereof) that can be recognized by at least one COL polypeptide.

“Promoter function” includes regulating expression of the coding sequences under a promoter’s control by providing a recognition site for RNA polymerase and/or other factors, such as transcription factors, all of which are necessary for the start of transcription at a transcription initiation site. A “promoter function” may also include the extent to which a gene coding sequence is transcribed to the extent determined by a promoter sequence.

“CORE promoter function” herein refers to the regulation of transcription of a transcribable DNA sequence by any of the COL polypeptides through binding to its cognate site in a promoter.

The term “inductive day length” or “inductive photoperiod” refers to a light period of a duration that induces flowering in a plant. For example, so-called “long day plants” are typically induced to flower by day lengths of 16 hours duration whereas “short day plants” are typically induced to flower by day lengths of 8-10 hours in duration.

A promoter or promoter region may include variations of promoters found in the present Sequence Listing, which may be derived by ligation to other regulatory sequences, random mutagenesis, controlled mutagenesis, and/or by the addition or duplication of enhancer sequences. Promoters disclosed in the present Sequence Listing and biologically functional
5 equivalents or variations thereof may drive the transcription of operably-linked coding sequences when comprised within an expression vector and introduced into a host plant. Promoters such as those found in the Sequence Listing (i.e., SEQ ID NOs: 101, 102, 1-4, 7, 8, 9, 10, 13, 14, 15, 18, 20 and 21) may be used to generate similarly functional promoters containing essential promoter elements. Functional promoters may also include a functional part or
10 fragment of any of SEQ ID NO: 101, 102, 1 - 4, 7-10, 13-15, 18, 20 or 21 provided the functional part also includes a CORE promoter function.

A "polypeptide" is an amino acid sequence comprising a plurality of consecutive polymerized amino acid residues e.g., at least about 15 consecutive polymerized amino acid residues. In some of the instances referred to in this application, a polypeptide comprises a
15 polymerized amino acid residue sequence that is a transcription factor or a domain or portion or fragment thereof. Additionally, the transcription factor may comprise: (i) a localization domain; (ii) an activation domain; (iii) a repression domain; (iv) an oligomerization domain; (v) a DNA-binding domain; or the like. The polypeptide optionally comprises modified amino acid residues, naturally occurring amino acid residues not encoded by a codon, non-naturally
20 occurring amino acid residues.

"Protein" refers to an amino acid sequence, oligopeptide, peptide, polypeptide or portions thereof whether naturally occurring or synthetic.

A "recombinant polypeptide" is a polypeptide produced by translation of a recombinant polynucleotide. A "synthetic polypeptide" is a polypeptide created by consecutive
25 polymerization of isolated amino acid residues using methods well known in the art. An "isolated polypeptide," whether a naturally occurring or a recombinant polypeptide, is more enriched in (or out of) a cell than the polypeptide in its natural state in a wild-type cell, e.g., more than about 5% enriched, more than about 10% enriched, or more than about 20%, or more than about 50%, or more, enriched, i.e., alternatively denoted: 105%, 110%, 120%, 150% or
30 more, enriched relative to wild type standardized at 100%. Such enrichment is not the result of a natural response of a wild-type plant. Alternatively, or additionally, the isolated polypeptide is separated from other cellular components with which it is typically associated, e.g., by any of the various protein purification methods herein.

"Homology" refers to sequence similarity between a reference sequence and at least a fragment of a newly sequenced clone insert or its encoded amino acid sequence.

"Identity" or "similarity" refers to sequence similarity between two polynucleotide sequences or between two polypeptide sequences, with identity being a more strict comparison.

5 The phrases "percent identity" and "% identity" refer to the percentage of sequence similarity found in a comparison of two or more polynucleotide sequences or two or more polypeptide sequences.

"Sequence similarity" refers to the percent similarity in base pair sequence (as determined by any suitable method) between two or more polynucleotide sequences. Two or
10 more sequences can be anywhere from 0-100% similar, or any integer value there between. Identity or similarity can be determined by comparing a position in each sequence that may be aligned for purposes of comparison. When a position in the compared sequence is occupied by the same nucleotide base or amino acid, then the molecules are identical at that position. A degree of similarity or identity between polynucleotide sequences is a function of the number of
15 identical, matching or corresponding nucleotides at positions shared by the polynucleotide sequences. A degree of identity of polypeptide sequences is a function of the number of identical amino acids at corresponding positions shared by the polypeptide sequences. A degree of homology or similarity of polypeptide sequences is a function of the number of amino acids at corresponding positions shared by the polypeptide sequences.

20 "Complementary" refers to the natural hydrogen bonding by base pairing between purines and pyrimidines. For example, the sequence A-C-G-T (5' -> 3') forms hydrogen bonds with its complements A-C-G-T (5' -> 3') or A-C-G-U (5' -> 3'). Two single-stranded molecules may be considered partially complementary, if only some of the nucleotides bond, or "completely complementary" if all of the nucleotides bond. The degree of complementarity
25 between nucleic acid strands affects the efficiency and strength of hybridization and amplification reactions. "Fully complementary" refers to the case where bonding occurs between every base pair and its complement in a pair of sequences, and the two sequences have the same number of nucleotides.

The terms "paralog" and "ortholog" are defined as evolutionarily related genes that have
30 similar sequences and functions. Orthologs are structurally related genes in different species that are derived by a speciation event. Paralogs are structurally related genes within a single species that are derived by a duplication event.

The term "equivalog" describes members of a set of homologous proteins that are conserved with respect to function since their last common ancestor (Haft et al., 2003, Nucleic

Acids Res. 31: 371–373). Related proteins are grouped into equivalent families, and otherwise into protein families with other hierarchically defined homology types.

In general, the term "variant" refers to molecules with some differences, generated synthetically or naturally, in their base or amino acid sequences as compared to a reference (native) polynucleotide or polypeptide, respectively. These differences include substitutions, insertions, deletions or any desired combinations of such changes in a native polynucleotide of amino acid sequence.

Promoters that are similar to those listed in the Sequence Listing: SEQ ID NOs: 101, 102, 1 - 4, 7-10, 13-15, 18, 20 or 21 may be made that have some alterations in the nucleotide sequence and yet retain the function of the listed sequences. One preferred method of alteration of a polynucleotide sequence is to use PCR to modify selected nucleotides or regions of sequences. These methods are well known to those of skill in the art. Sequences can be modified, for example by insertion, deletion, or replacement of template sequences in a PCR-based DNA modification approach. A "promoter variant" or "variant promoter" is a promoter containing changes in which one or more nucleotides of an original promoter is deleted, added, and/or substituted, preferably while substantially maintaining promoter function. For example, one or more base pairs may be deleted from the 5' or 3' end of a promoter to produce a "truncated" promoter. One or more base pairs can also be inserted, deleted, or substituted internally to a promoter. In the case of a promoter fragment, variant promoters can include changes affecting the transcription of a minimal promoter to which it is operably linked. Variant promoters can be produced, for example, by standard DNA mutagenesis techniques or by chemically synthesizing the variant promoter or a portion thereof.

Also within the claimed scope is a variant of a gene promoter listed in the Sequence Listing, that is, one having a sequence that differs from one of the polynucleotide sequences in the Sequence Listing, or a complementary sequence.

With regard to polynucleotide variants of coding sequences that encode polypeptides, differences between presently disclosed polynucleotides and polynucleotide variants are limited so that the nucleotide sequences of the former and the latter are closely similar overall and, in many regions, identical. Due to the degeneracy of the genetic code, differences between the former and latter nucleotide sequences may be silent (i.e., the amino acids encoded by the polynucleotide are the same, and the variant polynucleotide sequence encodes the same amino acid sequence as the presently disclosed polynucleotide). Variant nucleotide sequences of coding sequences may encode different amino acid sequences, in which case such nucleotide differences will result in amino acid substitutions, additions, deletions, insertions, truncations or

fusions with respect to the similar disclosed polynucleotide sequences. These variations may result in polynucleotide variants encoding polypeptides that share at least one functional characteristic. The degeneracy of the genetic code also dictates that many different variant polynucleotides can encode identical and/or substantially similar polypeptides in addition to those sequences illustrated in the Sequence Listing.

The term "plant" includes whole plants, shoot vegetative organs/structures (for example, leaves, stems and tubers), roots, flowers and floral organs/structures (for example, bracts, sepals, petals, stamens, carpels, anthers and ovules), seed (including embryo, endosperm, and seed coat) and fruit (the mature ovary), plant tissue (for example, vascular tissue, ground tissue, and the like) and cells (for example, guard cells, egg cells, and the like), and progeny of same. The class of plants that can be used in the instant method is generally as broad as the class of higher and lower plants amenable to transformation techniques, including angiosperms (monocotyledonous and dicotyledonous plants), gymnosperms, ferns, horsetails, psilophytes, lycophytes, bryophytes, and multicellular algae.

A "control plant" as used in the present invention refers to a plant cell, seed, plant component, plant tissue, plant organ or whole plant used to compare against transgenic or genetically modified plant for the purpose of identifying an enhanced phenotype in the transgenic or genetically modified plant. A control plant may in some cases be a transgenic plant line that comprises an empty vector or marker gene, but does not contain the recombinant polynucleotide of the present invention that is expressed in the transgenic or genetically modified plant being evaluated. In general, a control plant is a plant of the same line or variety as the transgenic or genetically modified plant being tested. A suitable control plant would include a genetically unaltered or non-transgenic plant of the parental line used to generate a transgenic plant herein.

A "transgenic plant" refers to a plant that contains genetic material not found in a wild-type plant of the same species, variety or cultivar. The genetic material may include a transgene, an insertional mutagenesis event (such as by transposon or T-DNA insertional mutagenesis), an activation tagging sequence, a mutated sequence, a homologous recombination event or a sequence modified by chimeraplasty. Typically, the foreign genetic material has been introduced into the plant by human manipulation, but any method can be used as one of skill in the art recognizes.

A transgenic plant may contain a nucleic acid construct (e.g., an expression vector or cassette). The nucleic acid construct typically comprises a polypeptide-encoding sequence operably linked (i.e., under regulatory control of) to an inducible regulatory sequence, such as a

promoter, that allows for the controlled expression of polypeptide. The nucleic acid construct can be introduced into a plant by transformation or by breeding after transformation of a parent plant. A plant refers to a whole plant as well as to a plant part, such as seed, fruit, leaf, or root, plant tissue, plant cells or any other plant material, e.g., a plant explant, as well as to progeny
5 thereof, and to *in vitro* systems that mimic biochemical or cellular components or processes in a cell.

"Wild type" or "wild-type", as used herein, refers to a plant cell, seed, plant component, plant tissue, plant organ or whole plant that has not been genetically modified or treated in an experimental sense. Wild-type cells, seed, components, tissue, organs or whole plants may be
10 used as controls to compare levels of expression and the extent and nature of trait modification with cells, tissue or plants of the same species in which expression of a polypeptide, such as a transcription factor polypeptide, is altered, e.g., in that it has been overexpressed or ectopically expressed.

A "trait" refers to a physiological, morphological, biochemical, or physical characteristic
15 of a plant or particular plant material or cell. In some instances, this characteristic is visible to the human eye, such as seed or plant size, or can be measured by biochemical techniques, such as detecting the protein, starch, or oil content of seed or leaves, or by observation of a metabolic or physiological process, e.g., by measuring tolerance to a form of stress, such as water deficit or water deprivation, or particular salt or sugar concentrations, or by the observation of the
20 expression level of a gene or genes, e.g., by employing Northern analysis, RT-PCR, microarray gene expression assays, or reporter gene expression systems, or by agricultural observations such as extent of wilting, turgor, hyperosmotic stress tolerance or in a preferred embodiment, yield. Any technique can be used to measure the amount of, comparative level of, or difference in any selected chemical compound or macromolecule in the transgenic plants, however.

"Trait modification" refers to a detectable difference in a characteristic in a plant
25 ectopically expressing a polynucleotide or polypeptide of the present invention relative to a plant not doing so, such as a wild-type plant. In some cases, the trait modification can be evaluated quantitatively. For example, the trait modification can entail at least about a 2% increase or decrease, or an even greater difference, in an observed trait as compared with a control or wild-
30 type plant. It is known that there can be a natural variation in the modified trait. Therefore, the trait modification observed entails a change of the normal distribution and magnitude of the trait in the plants as compared to control or wild-type plants.

When two or more plants are "morphologically similar" they have comparable forms or appearances, including analogous features such as dimension, height, width, mass, root mass,

shape, glossiness, color, stem diameter, leaf size, leaf dimension, leaf density, internode distance, branching, root branching, number and form of inflorescences, and other macroscopic characteristics at a particular stage of growth. If the plants are morphologically similar at all stages of growth, they are also "developmentally similar". It may be difficult to distinguish two
5 plants that are genotypically distinct but morphologically similar based on morphological characteristics alone.

"Ectopic expression" or "altered expression" in reference to a polynucleotide indicates that the pattern of expression in, e.g., a transgenic plant or plant tissue, is different from the expression pattern in a wild-type plant or a reference plant of the same species. The pattern of
10 expression may also be compared with a reference expression pattern in a wild-type plant of the same species. For example, the polynucleotide or polypeptide is expressed in a cell or tissue type other than a cell or tissue type in which the sequence is expressed in the wild-type plant, or by expression at a time other than at the time the sequence is expressed in the wild-type plant, or by a response to different inducible agents, such as hormones or environmental signals, or at
15 different expression levels (either higher or lower) compared with those found in a wild-type plant. The term also refers to altered expression patterns that are produced by lowering the levels of expression to below the detection level or completely abolishing expression. The resulting expression pattern can be transient or stable, constitutive or inducible. In reference to a
20 polypeptide, the term "ectopic expression or altered expression" further may relate to altered activity levels resulting from the interactions of the polypeptides with exogenous or endogenous modulators or from interactions with factors or as a result of the chemical modification of the polypeptides.

The term "overexpression" as used herein refers to a greater expression level of a gene in a plant, plant cell or plant tissue, compared to expression in a wild-type plant, cell or tissue, at
25 any developmental or temporal stage for the gene. Overexpression can occur when, for example, the genes encoding one or more proteins are under the control of a strong promoter (e.g., the cauliflower mosaic virus 35S transcription initiation region). Overexpression may also under the control of an inducible promoter such as a CORE promoter. Thus, overexpression may occur throughout a plant or in the presence of particular environmental signals, depending on the
30 promoter used.

Overexpression may take place in plant cells normally lacking expression of polypeptides functionally equivalent or identical to a polypeptide that can confer an improved trait, for example, increased stress tolerance or improved yield. Overexpression may also occur in plant cells where endogenous expression of the present proteins that confer an improved trait,

for example, improved stress tolerance, or functionally equivalent molecules, normally occurs, but such normal expression is at a lower level. Overexpression thus results in a greater than normal production, or "overproduction" of the protein that confers the improved trait in the plant, cell or tissue.

5 The term "transcription regulating region" refers to a DNA regulatory sequence that regulates expression of one or more genes in a plant when a polypeptide having one or more specific binding domains binds to the DNA regulatory sequence. Polypeptides, for example, transcription factors, may possess a conserved domain. Transcription factors may also comprise an amino acid subsequence that forms a transcription activation domain that regulates
10 expression of one or more stress resistance genes in a plant when the transcription factor binds to the regulating region.

 The term "regulator" herein refers to a polynucleotide or polypeptide sequence that regulates expression of one or more genes.

 The phrases "coding sequence," "structural sequence," and "transcribable polynucleotide
15 sequence" refer to a physical structure comprising an orderly arrangement of nucleic acids. The nucleic acids are arranged in a series of nucleic acid triplets that each form a codon. Each codon encodes for a specific amino acid. Thus the coding sequence, structural sequence, and transcribable polynucleotide sequence encode a series of amino acids forming a protein, polypeptide, or peptide sequence. The coding sequence, structural sequence, and transcribable
20 polynucleotide sequence may be contained, without limitation, within a larger nucleic acid molecule, vector, etc. In addition, the orderly arrangement of nucleic acids in these sequences may be depicted, without limitation, in the form of a sequence listing, figure, table, electronic medium, etc.

 The term "isolated", indicates that the molecule referenced is not in its native
25 environment, that is, not normally found in the genome of a particular host cell, or a DNA not normally found in the host genome in an identical context, or any two sequences adjacent to each other that are not normally or naturally adjacent to each other.

 The term "operably linked" refers to a first polynucleotide molecule, such as a promoter, connected with a second transcribable polynucleotide molecule, such as a gene of interest, where
30 the polynucleotide molecules are so arranged that the first polynucleotide molecule affects the function of the second polynucleotide molecule. The two polynucleotide molecules may be part of a single contiguous polynucleotide molecule and may be adjacent. For example, a promoter is operably linked to a gene of interest if the promoter modulates transcription of the gene of interest in a cell.

DESCRIPTION OF THE SPECIFIC EMBODIMENTS

The control of flowering by photoperiod is an important adaptive characteristic in plants. Studies of the model dicot *Arabidopsis* have shown that the *CO* (*CONSTANS*) gene has an important role in the photoperiod pathway, which is one of four regulatory pathways controlling the timing of flowering (Martinez-Zapater et al., 2004; Putterill et al., 1995; Mouradov et al., 2002; Simpson et al., 2002).

The *Arabidopsis* transcription factor *CONSTANS* (*CO*) promotes flowering in *Arabidopsis* in response to a lengthening photoperiod via a mechanism involving tight regulation at both the DNA and protein level (see Turck et al, 2008). Ectopic, constitutive expression of *CO* prematurely triggers flowering independently of day length. *CO* protein has two-tandem B-box zinc finger domains at the amino terminus and contains a conserved region at the carboxy-terminus known as the CCT motif. *Arabidopsis* has 16 additional proteins with either one or two B-boxes at the amino terminus and one CCT motif at the carboxy- terminus. These proteins are named *CONSTANS*-like or *COL* proteins (*COLs*): *COL1* to *COL16* (Robson et al., 2001). A subset of the *COL* proteins has been demonstrated to participate in the regulation of the floral transition while other family members may have no apparent role in the process. There are numerous other *Arabidopsis* proteins in the B-box zinc finger family. These proteins contain one or two B-box at their amino-terminus but lack the conserved CCT motif at their carboxy- terminus. Both classes of B-box protein families (with or without CCT motif) regulate flowering, circadian rhythm and light mediated growth and developmental processes in plants. The function of the conserved B-box and CCT domains has not been fully elucidated; however, the existence of proteins containing only one of these domains suggests that these motifs do not require each other for their function and act independently. In animal systems, the B-box is present alone or as a part of a tripartite motif comprised of a zinc-binding RING finger, one or two B-boxes followed by a coil-coil domain (RBCC). The RBCC motif is implicated in protein-protein interactions and is present in various transcription factors, ubiquitin ligases, receptor proteins and other structural protein classes. It is unclear if the B-box proteins in plants, with a distinct composition of protein structural motifs, function similarly as the animal proteins. *COL* proteins, including *CO*, have been shown to function as transcriptional regulators that control the expression of various genes including the flowering modulator *FLOWERING LOCUS T* (*FT*).

In *Arabidopsis*, the *CO* protein is a positive regulator of the *FT* gene; in rice, *CO* acts as a floral repressor. Rice is a facultative short day plant, under long days (for example, 16 h of light) the *CO* protein is stabilized and negatively regulates the expression of *FT* resulting in a repression of flowering (Kojima et al., 2002; Tamaki et al., 2007). Another key transcription

factor that negatively regulates flowering time is the repressor FLOWERING LOCUS C (FLC). However, unlike CO, FLC regulates flowering as part of the "autonomous" pathway of flowering control, and in Arabidopsis the protein has a native function which is largely independent of light period duration (Martinez-Zapater et al., 2004). FLC binds to the CArG box within the FT promoter and negatively regulates the expression of the FT gene. In addition to CO and FLC, numerous other transcription factors directly or indirectly affect the expression of FT, but CO is likely to represent the major component in regulating the floral transition in response to photoperiod.

The present application provides the promoter sequences that can be bound by COL proteins, including CO. These CORE promoters can regulate expression of useful proteins and may be of significant value for a number of reasons, including, but not limited to, the following:

1. CORE promoters are capable of causing, in response to changes in amount of active COLs with a plant or plant cell, sufficient expression of an exogenous gene so that the exogenous protein encoded by the exogenous gene will be produced at a level sufficient to confer an improved trait in a transformed plant, or suppression or inactivity of one or more endogenous proteins in a plant resulting in an improved trait in a plant. Examples of altered traits include, but are not limited to, increased yield, increased disease resistance, altered timing of flowering, sterility, reduced sensitivity to light, greater early season growth, greater height, greater stem diameter, increased biomass, increased photosynthetic rate, increased resistance to lodging, increased internode length, increased secondary rooting, greater cold tolerance, greater tolerance to water deprivation, greater tolerance to salt, greater tolerance to heat, altered sugar sensing, reduced stomatal conductance, altered C/N sensing, increased low nitrogen tolerance, increased low phosphorus tolerance, increased tolerance to hyperosmotic stress, greater late season growth and vigor, increased number of mainstem nodes, or greater canopy coverage.

2. CORE promoters can be used to regulate timing of the floral transition. The transition of plants from a vegetative state to a reproductive stage is a complex process involving alterations on multiple levels. Flowering is integrated with several physiological pathways including nutrient sensing (sugar signaling), development and the perception of environmental stress. The timing of the floral transition has a significant impact on both biomass production and grain yields, with early flowering sometimes being associated with reduced grain production and plant biomass. Conversely, delayed flowering can result in the accumulation of vegetative biomass and an increase in photosynthetic capacity, which can lead to enhanced yield. However, the absence of flowering, or photoperiod dependent flowering, or late flowering can pose barriers in breeding programs. For example, if two parental lines have widely differing

flowering times, this can impede the development of hybrid lines, which restricts the ability to breed new out-crossed plant varieties of agricultural interest. CORE promoters can be used to drive the expression of an activator or inhibitor of flowering as the regulator protein resulting in a positive or negative feedback loop to rapidly up-regulate or repress the florigen signal relative to the endogenous gene, and consequently regulate the timing of floral transition, which will have direct implications on the yield of bioenergy and food crops.

3. Some regulator genes have negative impact to plants during the seedling and vegetative growth stages of a plant life cycle, but have reduced detrimental effects if expressed at high levels at and following the transition to flowering. CORE promoters can be used to drive enriched expression of such regulators during or post-flowering stage, which can impart desired traits while minimizing the negative effects on plants.

4. Ectopic expression of useful polypeptides in transgenic plants without significant adverse morphological effects would make these polypeptides available as effective commercial tools for improved traits such as, for example, improved abiotic stress tolerance, improved disease resistance, improved yield, and the like. One such means is the use of CORE promoters that can confer improved traits while mitigating undesirable effects that arise during constitutive overexpression of proteins of interest. The CORE promoter driving expression of a regulator gene would enable the creation of plants with enhanced traits during and following the floral transition, for example, enhanced abiotic stress tolerance or enhanced disease resistance for crops that are sensitive to abiotic stress or disease during reproductive development.

5. CORE promoters can be used to drive expression of a polypeptide that results in sterility of the transgenic plant through either direct or indirect mechanisms. The production of the sterility regulator protein only following the floral transition would be beneficial in that it would reduce any off-types that would be associated with the production of such a protein during vegetative growth. The induced sterility would also control genetic drift from the transgenic plants to wild-type plants in the field.

6. CORE promoters driving the expression of selectable / visible markers are valuable in studying COL involved regulatory pathways, for example, photoperiodic flowering pathways. The expression of such a marker will be altered in plants that are defective in flowering signaling in response to changes in day length. Plants transformed with CORE-promoter::marker constructs can be used to screen for genetic mutations which may lead to changes in the expression pattern or in amplitude of a quantifiable marker signal, for example, β -glucuronidase (GUS).

7. CORE promoters can be used in a two-component activation system as illustrated in Figure 4 to deliver more consistent and high levels of the “regulator” protein that can confer desired traits. For example, the CORE promoter sequence could be used to drive the expression of a transcriptional activation complex such as LEXA:GAL4, where LEXA represents a DNA
 5 binding domain and GAL4 represents a transcriptional activation domain. The LEXA:GAL4 protein fusion binds specifically to the opLexA promoter to drive expression of a desired “regulator”. This alternative activation system may deliver more consistent and high levels of the “regulator” protein.

Promoters are provided as SEQ ID NO: 1 - 4, 7-10, 13-15, 18, 20 or 21, and expression
 10 vectors that may be constructed using these promoters may be introduced into plants for the purpose of regulating expression of polypeptides of interest to confer improved traits. In some embodiments of the invention, a polypeptides of interest is encoded by a transcribable polynucleotide that is heterogonous to a promoter sequence of the invention. The invention also encompasses a CORE promoter that comprises a functional part of any of SEQ ID NOs: 101,
 15 102, 1 - 4, 7-10, 13- 21, provided that the functional part of the promoter also includes a CORE promoter function. 8, 9, 10, 13, 17, 20, 30, 35, 40, 50, 75, 100, 125, 150, 191, 200, 251, 311, 491, 496, 525, 550, 575, 600, 650, 700, 750, 800, 900, 950, 1000, 1050, 1100, 1150, 1200, 1500, 1633, 1639 or 1800 contiguous nucleotides of the nucleic acid sequences of SEQ ID NOs: 1, 14, 15, or 21, as well as all lengths of contiguous nucleotides within such sizes, or have multimeric
 20 copies of sequences selected from of SEQ ID NO: 101, 102, 7-10, 13, 16-20 or from any combinations thereof, provided that the functional part of the promoter includes a CORE promoter function. These promoters contain the consensus sequence TGTGN(1-3)ATG, (SEQ ID NO:16) (Fig. 1A) or a complement thereof (Fig. 1B).

To select a CORE promoter, a system that displays changes in amplitude of a
 25 quantifiable marker signal was used. For example, co-transfection into plant protoplasts of (1) a reporter construct comprised of the FT promoter sequences (FT1-FT6, SEQ ID NOs 1-6) fused with the GUS reporter gene and (2) a construct encoding the CO protein enabled the direct measurement of CO transcriptional activity. Critical regions required for the CORE promoter function have been identified through this system, for example, in FT promoter, a key activation
 30 motif must exist between -146 and -190 (relative to the start codon) with additional elements likely to be present in the region between -190 and -251. The two related motifs, CORE1 and CORE2 were identified in these regions (Figure 1A). A similar element GATTGTGCATG (SEQ ID NO: 96) is present in the promoter of another downstream target of CO, the *SOC1* gene (Figure 1A). A sequence motif CATN(1-3)CACA (SEQ ID NO: 87), which is complementary to

TGTGN(1-3)ATG (SEQ ID NO: 16) is present in both AOX1 and PAP2 (Figure1B). Promoters that are similar to those listed in the Sequence Listing: SEQ ID NOs: 101, 102, 1 - 4, 7-10, 13-15, 18, 20 or 21 may be made that have some alterations in the nucleotide sequence and yet retain the function of the listed sequences. At the nucleotide level, the promoter sequences will typically share at least about at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% nucleotide sequence identity with any of SEQ ID NOs: 101, 102, 1 - 4, 7-10, 13- 21.

To determine the percent identity of two amino acid sequences or of two nucleic acids, the sequences are aligned for optimal comparison purposes. The percent identity between the two sequences is a function of the number of identical positions shared by the sequences (i.e., $\text{percent identity} = \frac{\text{number of identical positions}}{\text{total number of positions}} \times 100$ (e.g., overlapping positions).times.100). In one embodiment, the total number of positions is the total number of nucleotides or amino acid residues contained in the entire length of one of the optimally aligned sequences. The percent identity between two sequences can be determined using techniques similar to those described below, with or without allowing gaps.

Percent identity can be determined electronically, e.g., by using the MEGALIGN program (DNASTAR, Inc. Madison, Wis.). The MEGALIGN program can create alignments between two or more sequences according to different methods, for example, the clustal method (see, for example, Higgins and Sharp (1988). The clustal algorithm groups sequences into clusters by examining the distances between all pairs. The clusters are aligned pairwise and then in groups. Other alignment algorithms or programs may be used, including Accelrys Gene, FASTA, BLAST, or ENTREZ, FASTA and BLAST, and which may be used to calculate percent similarity. These are available as a part of the GCG sequence analysis package (University of Wisconsin, Madison, WI), and can be used with or without default settings. ENTREZ is available through the National Center for Biotechnology Information. In one embodiment, the percent identity of two sequences can be determined by the GCG program with a gap weight of 1 (see USPN 6,262,333).

The percent identity between two polypeptide sequences can also be determined using Accelrys Gene v2.5 (2006) with default parameters: Pairwise Matrix: GONNET; Align Speed: Slow; Open Gap Penalty: 10.000; Extended Gap Penalty: .100; Multiple Matrix: GONNET; Multiple Open Gap Penalty: 10.000; Multiple Extended Gap Penalty: .05; Delay Divergent: 30; Gap Separation Distance: 8; End Gap Separation: false; Residue Specific Penalties: false;

Hydrophilic Penalties: false; Hydrophilic Residues: GPSNDQEKR. The default parameters for determining percent identity between two polynucleotide sequences using Accelrys Gene are: Align Speed: Slow; Open Gap Penalty: 10.000; Extended Gap Penalty: 5.000; Multiple Open Gap Penalty: 10.000; Multiple Extended Gap Penalty: 5.000; Delay Divergent: 40; Transition:
5 Weighted.

Software for performing BLAST analyses is publicly available, e.g., through the National Center for Biotechnology Information (see internet website at www.ncbi.nlm.nih.gov/). This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued
10 threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul (1990); Altschul (1993)). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are then extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, for
15 nucleotide sequences, the parameters M (reward score for a pair of matching residues; always > 0) and N (penalty score for mismatching residues; always < 0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation
20 of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) of 10, a cutoff of 100, M=5, N=-4, and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength (W) of 3, an
25 expectation (E) of 10, and the BLOSUM62 scoring matrix (*see* Henikoff & Henikoff (1989)). Unless otherwise indicated for comparisons of predicted polynucleotides, "sequence identity" refers to the % sequence identity generated from a tblastx using the NCBI version of the algorithm at the default settings using gapped alignments with the filter "off" (see, for example, internet website at www.ncbi.nlm.nih.gov/).

30 Novel chimeric promoters can be designed or engineered based on the promoters disclosed in the present invention by a number of methods. Many promoters contain cis-elements that activate, enhance or define the strength and/or specificity of the promoter. For example promoters may contain "TATA" boxes defining the site of transcription initiation and other cis-elements located upstream of the transcription initiation site that modulate transcription

levels. For example, a chimeric promoter may be produced by fusing a first promoter fragment containing the activator cis-element from one promoter to a second promoter fragment containing the activator cis-element from another promoter; the resultant chimeric promoter may cause an increase in expression of an operably linked transcribable polynucleotide molecule.

5 Promoters can be constructed such that promoter fragments or elements are operably linked, for example, by placing such a fragment upstream of a minimal promoter. The cis-elements and fragments of the present invention can be used for the construction of such chimeric promoters. Methods for construction of chimeric and variant promoters of the present invention include, but are not limited to, combining control elements of different promoters or duplicating portions or
10 regions of a promoter (see for example, U.S. Pat. Nos. 4,990,607; 5,110,732; and 5,097,025, all of which are herein incorporated by reference). Those of skill in the art are familiar with the standard resource materials that describe specific conditions and procedures for the construction, manipulation, and isolation of macromolecules (e.g., polynucleotide molecules, plasmids, etc.), as well as the generation of recombinant organisms and the screening and isolation of
15 polynucleotide molecules.

Sequence Variations

It will readily be appreciated by those of skill in the art, that the invention includes any of a variety of polynucleotide sequences provided in the Sequence Listing or capable of encoding polypeptides that function similarly to those provided in the Sequence Listing. Due to
20 the degeneracy of the genetic code, many different polynucleotides can encode identical and/or substantially similar polypeptides in addition to those sequences illustrated in the Sequence Listing. Nucleic acids having a sequence that differs from the sequences shown in the Sequence Listing, or complementary sequences, that encode functionally equivalent peptides (that is, peptides having some degree of equivalent or similar biological activity) but differ in sequence
25 from the sequence shown in the sequence listing due to degeneracy in the genetic code, are also within the scope of the invention.

Altered polynucleotide sequences encoding polypeptides include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polynucleotide encoding a polypeptide with at least one functional characteristic of the instant polypeptides.
30 Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding the instant polypeptides, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding the instant polypeptides.

Sequence alterations that do not change the amino acid sequence encoded by the polynucleotide are termed “silent” variations. With the exception of the codons ATG and TGG, encoding methionine and tryptophan, respectively, any of the possible codons for the same amino acid can be substituted by a variety of techniques, for example, site-directed mutagenesis,
 5 available in the art. Accordingly, any and all such variations of a sequence selected from the above table are a feature of the invention.

In addition to silent variations, other conservative variations that alter one, or a few amino acids in the encoded polypeptide, can be made without altering the function of the polypeptide. For example, substitutions, deletions and insertions introduced into the sequences
 10 provided in the Sequence Listing are also envisioned. Such sequence modifications can be engineered into a sequence by site-directed mutagenesis (for example, Olson et al., Smith et al., Zhao et al., and other articles in Wu (ed.) *Meth. Enzymol.* (1993) vol. 217, Academic Press) or the other methods known in the art or noted herein. Amino acid substitutions are typically of single residues; insertions usually will be on the order of about from 1 to 10 amino acid residues;
 15 and deletions will range about from 1 to 30 residues. In preferred embodiments, deletions or insertions are made in adjacent pairs, for example, a deletion of two residues or insertion of two residues. Substitutions, deletions, insertions or any combination thereof can be combined to arrive at a sequence. The mutations that are made in the polynucleotide encoding the transcription factor should not place the sequence out of reading frame and should not create
 20 complementary regions that could produce secondary mRNA structure. Preferably, the polypeptide encoded by the DNA performs the desired function.

Conservative substitutions are those in which at least one residue in the amino acid sequence has been removed and a different residue inserted in its place. Such substitutions generally are made in accordance with the Table 1 when it is desired to maintain the activity of
 25 the protein. Table 1 shows amino acids which can be substituted for an amino acid in a protein and which are typically regarded as conservative substitutions.

Table 1. Possible conservative amino acid substitutions

Amino Acid Residue	Conservative substitutions
Ala	Ser
Arg	Lys
Asn	Gln; His
Asp	Glu
Gln	Asn
Cys	Ser
Glu	Asp
Gly	Pro

His	Asn; Gln
Ile	Leu, Val
Leu	Ile; Val
Lys	Arg; Gln
Met	Leu; Ile
Phe	Met; Leu; Tyr
Ser	Thr; Gly
Thr	Ser; Val
Trp	Tyr
Tyr	Trp; Phe
Val	Ile; Leu

The polypeptides provided in the Sequence Listing have a novel activity, such as, for example, regulatory activity. Although all conservative amino acid substitutions (for example, one basic amino acid substituted for another basic amino acid) in a polypeptide will not necessarily result in the polypeptide retaining its activity, it is expected that many of these conservative mutations would result in the polypeptide retaining its activity. Most mutations, conservative or non-conservative, made to a protein but outside of a conserved domain required for function and protein activity will not affect the activity of the protein to any great extent.

The following examples are included to demonstrate preferred embodiments of the invention. It should be appreciated by those of skill in the art that the techniques disclosed in the examples that follow represent techniques discovered by the inventors to function well in the practice of the invention. However, those of skill in the art should, in light of the present disclosure, appreciate that many changes can be made in the specific embodiments that are disclosed and still obtain a like or similar result without departing from the spirit and scope of the invention, therefore all matter set forth or shown in the accompanying drawings is to be interpreted as illustrative and not in a limiting sense.

EXAMPLES

Example I. Identification of CORE promoters of Native Genes

CO and other COL proteins have been shown to function as transcriptional regulators that control the expression of various genes including the flowering modulator FLOWERING LOCUS T (FT). In *Arabidopsis*, the CO protein is a positive regulator of the FT gene; however, the binding sites of the CO protein within the FT promoter had not been identified. A series of transcriptional activation experiments were performed in *Arabidopsis* protoplasts using progressively shorter sequences derived from the FT promoter (Figure 1A and Table 2). Co-transfection of (1) a reporter construct comprised of the FT promoter sequences (FT1-FT6, SEQ ID NOs: 1-6) fused with the GUS reporter gene and (2) an expression construct encoding the

CO protein enabled the direct measurement of transcriptional activity conferred by CO in the plant cells. FT1-FT4 contains one promoter motif between -146 and -190 (relative to the start codon), and a second CORE motif between -190 and -251 conferred the normal CORE function as indicated by the reporter activity, while FT5 and FT6 showed largely reduced or a loss of activity.

AOX1a protein plays a role in photorespiration, and PAP2 protein is responsive to sugar levels in the plant cell. Both the AOX1 promoter, SEQ ID NO: 15, and the PAP2 promoter, SEQ ID NO: 14 contain the CORE motif in the complementary strand. Reporter constructs comprised of transcriptional fusions of the promoters and the GUS gene were generated and used in transient transfection assays with plant protoplasts. Co-transfection of the reporter construct and construct driving constitutive CO expression resulted in the activation of the GUS reporter gene from both the AOX1a and PAP2 promoters, demonstrating that these promoters contain CORE activity. In particular, this result indicates that the CORE motif is active within alternative promoter sequence environments in addition to FT promoter (Table 2). The values in Table 2 represent relative promoter activity, i.e., fold induction of GUS activity for the indicated combination of protein and promoter sequence relative to the reporter activity caused by a control protein on the same promoter.

Table 2. Relative promoter activity

Protein	CO	COL9	COL15
FT1	8.8	NT	NT
FT2	6.7	NT	NT
FT3	7.1	NT	NT
FT4	5.7	NT	NT
FT5	1.7	NT	NT
FT6	1.2	NT	NT
3XCORE1A	2.8	NT	NT
3XCORE1B	4.9	1.2	5.0
2XCORE2A	5.0	NT	NT
4XCORE2B	6.3	10.3	7.9
4XCORE2BM1	0.6	NT	NT
4XCORE2BM2	1.0	NT	NT
CORE3	16.0	7.0	3.9
PAP2	2.0	8.5	1.5
AOX1a	1.6	2.5	1.7

NT: not tested

Example II. Analysis of sequence motif in the CORE promoters

Sequence alignment of the promoters with CORE function revealed the presence of one or more copies of the conserved CORE motif (SEQ ID NO: 16); FT1-FT4, SEQ ID NOs: 1-4, contains CORE1 motif (SEQ ID NO: 17) between -146 and -190 (relative to the start codon), and CORE2 motif (SEQ ID NO: 19) between -190 and -251 (Figure 1A). Both AOX1 promoter, SEQ ID NO: 15, and PAP2 promoter, SEQ ID NO: 14 contain the complementary sequence of the CORE motif, SEQ ID NO: 88 (Figure 1B). To analyze the function of the CORE motif, synthetic promoter elements comprised of multiple copies of the CORE1 or CORE2 sites (or mutated versions thereof) and a hybrid promoter sequence comprised of multimeric copies of both the CORE1 and CORE2 motifs (CORE3) upstream of the minimal Cauliflower Mosaic Virus (-46S TATA) promoter (SEQ ID NO: 97) (Figures 2A, 2B and 2C) were also tested for CORE activity. These promoters were used to generate transcriptional fusions to the GUS reporter gene. The synthetic reporter constructs demonstrated high levels of CO-dependent induction in plant protoplasts based studies as shown in Table 2. The CORE3 reporter demonstrated higher levels of reporter induction than constructs containing multiple copies of either single motif.

Specific binding assays were also performed to address whether CO is capable of directly binding the CORE motif, for example, an EMSA experiment was performed using the 4XCORE2B sequence and an epitope tagged variant of the CO protein. It was found that CO protein was bound specifically to the 4XCORE2B sequence (Figure 3). The CO protein failed to bind to the 4XCORE2BM1 sequence with a mutated CORE motif (TGTG to TATA). This result confirmed that the CORE motif is a direct target of CO and therefore likely plays a critical role in the regulation of the FT promoter and flowering.

Both the TGTG and ATG components of the CORE motif are necessary for the CORE function of the promoter. Mutant variants of the synthetic promoters where these individual components are disrupted (CORE2BM1 and CORE2BM2) lost their CORE activity (Table 2).

Example III. Other CONSTANS homologs that activate the CORE promoter

COL genes have been widely identified across the plant kingdom. In two cases (*Brassica napus BnCOa1*, Robert et al., 1998; and *Pharbitis nil PnCO*, Liu et al., 2001), *COL* genes have been shown to complement a *constans* mutant in *Arabidopsis*, demonstrating functional equivalence. The *Hd1* (*Heading date 1*) gene of rice (*Oryza sativa*) is also homologous to *CO* (Yano et al., 2000). Conservation between short-day (SD) plants (rice and *P. nil*) and long-day (LD) plants (*Arabidopsis* and *B. napus*) suggests that *CO* is involved in a conserved pathway regulating flowering in response to inductive day length and *CO*-like genes are likely to be involved in flowering time control in other cereals (Griffiths et al., 2003). These

orthologous CONSTANS proteins are likely to function through binding their cognitive CORE motif, as does the *Arabidopsis* CONSTANS.

In addition, two *Arabidopsis* CONSTANS homologs, COL9 and COL15, demonstrated capability of activating the promoters AOX1a, PAP2 and synthetic promoters to varying levels (Table 2) using the reporter gene analysis as described above, indicating possible roles in tissue-specific or developmentally-specific or environmentally-specific (e.g. stress responses) downstream signaling pathways.

Example IV. Transcribable polynucleotides of interest that can be regulated by CORE promoter

The nature of the polynucleotide to be transcribed is not limited. Specifically, the polynucleotide may include sequences that will have activity as RNA as well as sequences that result in a polypeptide product. These sequences may include, but are not limited to, antisense sequences, ribozyme sequences, microRNAs and their precursors, spliceosomes, amino acid coding sequences, and fragments thereof. Polypeptide products encoded by these amino acid coding sequences may include, but are not limited to, endogenous proteins or fragments thereof, or heterologous proteins including marker genes or fragments thereof, including SEQ ID Nos: 25-69 in the sequence listing. Examples of polypeptides and the traits they can confer to a plant are listed in Table 3. These promoter and control elements are useful to modulate flowering development when operably linked to the polynucleotide, of which transcription can regulate floral transition, for example, polynucleotides SEQ ID NOs: 48, 52, 61-67, or 90 provided in the sequence listing. Promoters and control elements of the present invention operably linked to polynucleotides are also useful for creation of plants that have improved traits, such as enhanced stress tolerance (SEQ ID NOs: 25-58, or 91-95), improved disease resistance (SEQ ID NOs: 41, 43, 68, or 69) or greater yield (SEQ ID NOs: 36, 37, 59 or 60). These promoters, when linked to a polynucleotide encoding a protein that results in sterility of the transgenic plant through either direct or indirect mechanisms, will be beneficial to reduce any off-types that would be associated with protein production during vegetative growth. The induced sterility would also control genetic drift from the transgenic plants to wild-type plants in the field.

Table 3. Polypeptides that confer altered traits relative to controls

Gene identifier	SEQ ID NO:	Trait(s)
G481	25	Enhanced tolerance to abiotic stress
G1073	26	Enhanced tolerance to abiotic stress
G1274	27	Enhanced tolerance to abiotic stress
G2133	28	Enhanced tolerance to abiotic stress
G2932	29	Enhanced tolerance to abiotic stress

G47	30	Enhanced tolerance to abiotic stress
G1753	31	Enhanced tolerance to abiotic stress
G1760	32	Enhanced tolerance to abiotic stress
G682	33	Enhanced tolerance to abiotic stress
G867	34	Enhanced tolerance to abiotic stress
G9	35	Enhanced tolerance to abiotic stress
G1543	36	Enhanced tolerance to abiotic stress, greater yield
G1988	37	Enhanced tolerance to abiotic stress, greater yield
G2982	38	Enhanced tolerance to abiotic stress
G2999	39	Enhanced tolerance to abiotic stress
G545	40	Enhanced tolerance to abiotic stress
G1792	41	abiotic stress tolerance, improved disease resistance
G1945	42	Enhanced tolerance to abiotic stress
G28	43	Enhanced tolerance to abiotic stress, improved disease resistance
G462	44	Enhanced tolerance to abiotic stress
G664	45	Enhanced tolerance to abiotic stress
G1335	46	Enhanced tolerance to abiotic stress
G153	47	Enhanced tolerance to abiotic stress
G154	48	Enhanced tolerance to abiotic stress
G652	49	Enhanced tolerance to abiotic stress
G740	50	Enhanced tolerance to abiotic stress
G2555	51	Enhanced tolerance to abiotic stress
G3086	52	abiotic stress tolerance, altered floral induction
G3866	53	Enhanced tolerance to abiotic stress
G3876	54	Enhanced tolerance to abiotic stress
G40	55	Enhanced tolerance to abiotic stress
G41	56	Enhanced tolerance to abiotic stress
G913	57	Enhanced tolerance to abiotic stress
G42	58	Enhanced tolerance to abiotic stress
SOS2	91	Enhanced tolerance to abiotic stress
NCED2	92	Enhanced tolerance to abiotic stress
LOS5	93	Enhanced tolerance to abiotic stress
NHX1	94	Enhanced tolerance to abiotic stress
NPK1	95	Enhanced tolerance to abiotic stress
G1266	59	greater yield
G155	60	greater yield
G2443 (CO)	61	altered floral induction
G3967 (GI)	62	altered floral induction
G4305 (FT)	63	altered floral induction
G4642 (FCA)	90	altered floral induction
G1759 (FLC)	64	altered floral induction
G2007	65	altered floral induction
G748	66	altered floral induction
TFL1	67	altered floral induction
G1795	68	improved disease resistance

G30	69	improved disease resistance
-----	----	-----------------------------

Alternatively, expression constructs can be used to inhibit expression of these peptides and polypeptides by incorporating the promoters in constructs for antisense use, RNAi use, co-suppression use or for the production of dominant negative mutations.

5 Example V. Plant Transformation and Analysis

The above-identified promoters may be used to regulate expression of genes of interest that confer an improved trait to a plant during and following a plant's floral transition. Plants transformed with the claimed promoters may be prepared using the following methods, although these examples are not intended to limit the invention.

10 Promoter cloning. For genes showing regulated activity by COLs, approximately 1.2 kb of upstream sequence are cloned by polymerase chain reaction (unless this region contains another gene, in which case the upstream sequence up to the next gene is cloned). Alternatively, a DNA sequence comprising the CORE promoter may be chemically synthesized. Each promoter is cloned into a nucleic acid construct (e.g., an expression vector or cassette) in front of
15 either a polynucleotide encoding a marker of gene expression such as green fluorescent protein (GFP) (SEQ ID NO: 100), GUS (SEQ ID NO: 98) and Luciferase (SEQ ID NO: 99), or in front of a polynucleotide encoding a polypeptide of interest, for example, a polypeptide found in the Sequence Listing, such as SEQ ID NOs: 25-69. In some of these cases, the polypeptide may produce deleterious morphological effects in the plants when they are constitutively
20 overexpressed, but which effects can be mitigated to some extent, or entirely, when expression of the polypeptide is regulated by a CORE promoter.

Transformation. Technology for introduction of DNA into cells is well known to those of skill in the art. Methods and materials for transforming plant cells by introducing a plant polynucleotide construct into a plant genome in the practice of this invention can include any of
25 the well-known and demonstrated methods including: (1) chemical methods (Graham et al., 1973; Zatloukal, et al., 1992); (2) physical methods such as microinjection (Capecci et al., 1980), electroporation (Wong et al., 1982; Fromm et al., 1985; U.S. Pat No. 5,384,253, herein incorporated by reference) particle acceleration (Johnston et al., 1994; Fynan et al., 1993) and microprojectile bombardment (as illustrated in U.S. Pat. No. 5,015,580; U.S. Pat. No. 5,550,318;
30 U.S. Pat. No. 5,538,880; U.S. Pat. No. 6,160,208; U.S. Pat. No. 6,399,861; and U.S. Pat. No. 6,403,865, all of which are herein incorporated by reference); (3) viral vectors (Clapp, 1993; Lu, et al., 1993; Eglitis et al., 1988); (4) receptor-mediated mechanisms (Curiel et al., 1992; Wagner, et al., 1992), and (5) bacterial mediated mechanisms such as *Agrobacterium*-mediated

transformation (as illustrated in U.S. Pat. No. 5,824,877; U.S. Pat. No. 5,591,616; U.S. Pat. No. 5,981,840; and U.S. Pat. No. 6,384,301, all of which are herein incorporated by reference); (6) Nucleic acids can be directly introduced into pollen by directly injecting a plant's reproductive organs (Zhou, et al., 1983; Hess, 1987; Luo, et al., 1988; Pena, et al., 1987). (7) Protoplast transformation, as illustrated in U.S. Pat. No. 5,508,184 (herein incorporated by reference). (8) The nucleic acids may also be injected into immature embryos (Neuhaus, et al., 1987).

Plant preparation. In the case of the model plant species *Arabidopsis*, seeds are sown on mesh covered pots. The seedlings are thinned so that 6-10 evenly spaced plants remain on each pot 10 days after planting. The primary bolts are cut off a week before transformation to break apical dominance and encourage auxiliary shoots to form. Transformation is typically performed at 4-5 weeks after sowing.

Bacterial culture preparation. *Agrobacterium* stocks are inoculated from single colony plates or from glycerol stocks and grown with the appropriate antibiotics and grown until saturation. On the morning of transformation, the saturated cultures are centrifuged and bacterial pellets are re-suspended in Infiltration Medium (0.5X MS, 1X B5 Vitamins, 5% sucrose, 1 mg/ml benzylaminopurine riboside, 200 µl/L Silwet L77) until an A600 reading of 0.8 is reached.

Transformation and seed harvest. The *Agrobacterium* solution is poured into dipping containers. All flower buds and rosette leaves of the plants are immersed in this solution for 30 seconds. The plants are laid on their side and wrapped to keep the humidity high. The plants are kept this way overnight at 4 °C and then the pots are turned upright, unwrapped, and moved to the growth racks.

The plants are maintained on the growth rack under 24-hour light until seeds are ready to be harvested. Seeds are harvested when 80% of the siliques of the transformed plants are ripe (approximately 5 weeks after the initial transformation). This seed is deemed T0 seed, since it is obtained from the T0 generation, and is later plated on selection plates (either kanamycin or sulfonamide). Resistant plants that are identified on such selection plates comprise the T1 generation.

For polynucleotides encoding polypeptides used in these experiments (e.g., SEQ ID NOs: 25-69), RT-PCR may be performed to confirm the ability of cloned promoter fragments to drive expression of the polypeptide transgene in plants transformed with the vectors.

T1 plants transformed with promoter::TF combinations comprised within a nucleic acid construct are subjected to morphological analysis. Promoters that produce a substantial amelioration of the negative effects of TF overexpression are subjected to further analysis by

propagation into the T2 generation, where the plants are analyzed for an altered trait relative to a control plant under inductive day length conditions. Examples of altered traits include, but are not limited to, the traits listed in Table 3, and/or sterility, reduced sensitivity to light, greater early season growth, greater height, greater stem diameter, increased biomass, increased
5 photosynthetic rate, increased resistance to lodging, increased internode length, increased secondary rooting, greater cold tolerance, greater tolerance to water deprivation, greater tolerance to salt, greater tolerance to heat, altered sugar sensing, reduced stomatal conductance, altered C/N sensing, increased low nitrogen tolerance, increased low phosphorus tolerance, increased tolerance to hyperosmotic stress, greater late season growth and vigor, increased
10 number of mainstem nodes, or greater canopy coverage.

Example VI. Transformation of dicots to produce improved traits

Crop species including tomato and soybean plants that overexpress genes encoding polypeptides of interest may produce plants with improved or desirable traits when placed under the regulatory control of CORE promoters found in the sequence listing, or related sequences
15 with similar regulatory function. Such genes, when overexpressed, will result in improved quality and larger yields than non-transformed plants in non-stressed or stressed conditions; the latter may occur in the field to even a low, imperceptible degree at any time in the growing season.

Thus, promoter sequences listed in the Sequence Listing recombined into, for example,
20 a nucleic acid construct, or another suitable expression vector, may be transformed into a plant for the purpose of regulating responses to floral cues and modifying plant traits for the purpose of improving yield and/or quality. The cloning vector may be introduced into a variety of plants by means well known in the art such as, for example, direct DNA transfer or *Agrobacterium tumefaciens*-mediated transformation. It is now routine to produce transgenic plants using most
25 dicot plants (see Weissbach and Weissbach, (1989); Gelvin et al. (1990); Herrera-Estrella et al. (1983); Bevan (1984); and Klee (1985). Methods for analysis of traits are routine in the art and disclosed in these Examples.

Numerous protocols for the transformation of tomato and soy plants have been previously described, and are well known in the art. Gruber et al. (1993), and Glick and
30 Thompson (1993) describe several expression vectors and culture methods that may be used for cell or tissue transformation and subsequent regeneration. For soybean transformation, methods are described by Miki et al. (1993); and U.S. Pat. No. 5,563,055, (Townsend and Thomas), issued Oct. 8, 1996.

There are a substantial number of alternatives to *Agrobacterium*-mediated transformation protocols, other methods for the purpose of transferring exogenous genes into soybeans or tomatoes. One such method is microprojectile-mediated transformation, in which DNA on the surface of microprojectile particles is driven into plant tissues with a biolistic device (see, for example, Sanford et al. (1987); Christou et al. (1992); Sanford (1993); Klein et al. (1987); U.S. Pat. No. 5,015,580 (Christou et al), issued May 14, 1991; and U.S. Pat. No. 5,322,783 (Tomes et al.), issued Jun. 21, 1994).

Alternatively, sonication methods (see, for example, Zhang et al. (1991); direct uptake of DNA into protoplasts using CaCl₂ precipitation, polyvinyl alcohol or poly-L-ornithine (Hain et al. (1985); Draper et al. (1982); liposome or spheroplast fusion (see, for example, Deshayes et al. (1985); Christou et al. (1987); and electroporation of protoplasts and whole cells and tissues (see, for example, Donn et al.(1990); D'Halluin et al. (1992); and Spencer et al. (1994), have been used to introduce foreign DNA and expression vectors into plants.

After a plant or plant cell is transformed (and the latter regenerated into a plant), the transformed plant may be crossed with itself or a plant from the same line, a non-transformed or wild-type plant, or another transformed plant from a different transgenic line of plants. Crossing provides the advantages of producing new and often stable transgenic varieties. Genes and the traits they confer that have been introduced into a tomato or soybean line may be moved into distinct line of plants using traditional backcrossing techniques well known in the art.

Transformation of tomato plants may be conducted using the protocols of Koornneef et al (1986), and in U.S. Patent 6,613,962, the latter method described in brief here. Eight day old cotyledon explants are precultured for 24 hours in Petri dishes containing a feeder layer of *Petunia hybrida* suspension cells plated on MS medium with 2% (w/v) sucrose and 0.8% agar supplemented with 10 μ M α -naphthalene acetic acid and 4.4 μ M 6-benzylaminopurine. The explants are then infected with a diluted overnight culture of *Agrobacterium tumefaciens* containing an expression vector comprising a polynucleotide for 5-10 minutes, blotted dry on sterile filter paper and cocultured for 48 hours on the original feeder layer plates. Culture conditions are as described above. Overnight cultures of *Agrobacterium tumefaciens* are diluted in liquid MS medium with 2% (w/v) sucrose, pH 5.7) to an OD₆₀₀ of 0.8.

Following cocultivation, the cotyledon explants are transferred to Petri dishes with selective medium comprising MS medium with 4.56 μ M zeatin, 67.3 μ M vancomycin, 418.9 μ M cefotaxime and 171.6 μ M kanamycin sulfate, and cultured under the culture conditions described above. The explants are subcultured every three weeks onto fresh medium. Emerging shoots are dissected from the underlying callus and transferred to glass jars with selective

medium without zeatin to form roots. The formation of roots in a kanamycin sulfate-containing medium is a positive indication of a successful transformation.

Transformation of soybean plants may be conducted using the methods found in, for example, U.S. Patent 5,563,055 (Townsend et al., issued October 8, 1996), described in brief
5 here. In this method soybean seed is surface sterilized by exposure to chlorine gas evolved in a glass bell jar. Seeds are germinated by plating on 1/10 strength agar solidified medium that consists of B5 basal medium with minimal organics, Sigma Chemical Co., cat. no. G5893, 0.32 gm/L; sucrose, 0.2% w/v and 2-[N-morpholino]ethanesulfonic acid (MES), 3.0 mM. without
10 plant growth regulators and culturing at 28° C. with a 16 hour day length. After three or four days, seed may be prepared for cocultivation. The seedcoat is removed and the elongating radicle removed 3-4 mm below the cotyledons.

Overnight cultures of *Agrobacterium tumefaciens* harboring the expression vector comprising a polynucleotide are grown to log phase, pooled, and concentrated by centrifugation. Inoculations are conducted in batches such that each plate of seed is treated with a newly
15 resuspended pellet of *Agrobacterium*. The pellets are resuspended in 20 ml inoculation medium that consists of B5 salts (G5893), 3.2 gm/L; sucrose, 2.0% w/v. 6-benzylaminopurine (BAP), 44 .mu.M; indolebutyric acid (IBA), 0.5 .mu.M; acetosyringone (AS), 100 .mu.M and was buffered to pH 5.5 with MES, 10 mM.. The inoculum is poured into a Petri dish containing prepared seed and the cotyledonary nodes are macerated with a surgical blade. After 30 minutes the explants
20 are transferred to plates of the same medium that has been solidified. Explants are embedded with the adaxial side up and level with the surface of the medium and cultured at 22° C. for three days under white fluorescent light. These plants may then be regenerated according to methods well established in the art, such as by moving the explants after three days to a liquid counter-selection medium (see U.S. Patent 5,563,055).

25 The explants may then be picked, embedded and cultured in solidified selection medium. After one month on selective medium transformed tissue becomes visible as green sectors of regenerating tissue against a background of bleached, less healthy tissue. Explants with green sectors are transferred to an elongation medium that consists of B5 salts (G5893), 3.2 gm/L; sucrose, 2.0% w/v; IBA, 3.3 .mu.M; gibberellic acid, 1.7 .mu.M; vancomycin, 100 .mu.g/ml; cefotaxime, 30 .mu.g/ml; and timentin, 30 .mu.g/ml, buffered to pH 5.7 with MES, 3.0 mM].
30 Elongation medium was solidified with gelrite, 0.2% w/v.. Culture is continued on this medium with transfers to fresh plates every two weeks. When shoots are 0.5 cm in length, they may be excised at the base and placed in a rooting medium that consists of B5 salts (G5893), 3.2 gm/L;

sucrose, 15 gm/L; nicotinic acid, 20 .mu.M; pyroglutamic acid (PGA), 900 mg/L and IBA, 10 .mu.M. It was buffered to pH 5.7 with MES, 3.0 mM and solidified with Gelrite, 0.2% w/v..

Protocols for the transformation of canola plants have also been previously described. See, for example, Pua et al. (1987); Charest et al. (1988); Radke et al. (1988); De Block et al. (1989); or Stewart et al. (1996) who teach *Agrobacterium*-mediated transformation of canola, or
5 (1989); or Stewart et al. (1996) who teach *Agrobacterium*-mediated transformation of canola, or Cardoza et al. (2003), who teach a method of *Agrobacterium*-mediated transformation of canola using hypocotyls as explant tissue.

Example VII: Transformation of monocots to produce improved traits

Cereal plants and other grasses such as, but not limited to, corn, wheat, rice, sorghum,
10 barley, *Miscanthus*, and switchgrass may be transformed with the present promoter sequences such as those presented in the present Sequence Listing, cloned into a vector such as pGA643 and containing a kanamycin-resistance marker, and inducibly express a polypeptide, for example, a transcription factor, that confers an improved or desirable trait. The expression vectors may be one found in the Sequence Listing, or any other suitable expression vector that
15 incorporates a CORE promoter sequence, may be similarly used. For example, pMEN020 may be modified to replace the NptII coding region with the BAR gene of *Streptomyces hygroscopicus* that confers resistance to phosphinothricin. The KpnI and BglII sites of the Bar gene are removed by site-directed mutagenesis with silent codon changes.

The cloning vector may be introduced into a variety of cereal plants by means well
20 known in the art including direct DNA transfer or *Agrobacterium tumefaciens*-mediated transformation. The latter approach may be accomplished by a variety of means, including, for example, that of U.S. Patent No. 5,591,616, in which monocotyledon callus is transformed by contacting dedifferentiating tissue with the *Agrobacterium* containing the cloning vector.

The sample tissues are immersed in a suspension of 3×10^9 cells of *Agrobacterium*
25 containing the cloning vector for 3-10 minutes. The callus material is cultured on solid medium at 25° C in the dark for several days. The calli grown on this medium are transferred to Regeneration medium. Transfers are continued every 2-3 weeks (2 or 3 times) until shoots develop. Shoots are then transferred to Shoot-Elongation medium every 2-3 weeks. Healthy looking shoots are transferred to rooting medium and after roots have developed, the plants are
30 placed into moist potting soil.

The transformed plants are then analyzed for the presence of the NPTII gene/ kanamycin resistance by ELISA, using the ELISA NPTII kit from 5Prime-3Prime Inc. (Boulder, CO).

It is also routine to use other methods to produce transgenic plants of most cereal crops (Vasil (1994), such as corn, wheat, rice, sorghum (Cassas et al. (1993), and barley (Wan and

Lemeaux (1994). DNA transfer methods such as the microprojectile method can be used for corn (Fromm et al. (1990); Gordon-Kamm et al. (1990); Ishida (1990); wheat (Vasil et al. (1992); Vasil et al. (1993); Weeks et al. (1993); and rice (Christou (1991); Hiei et al. (1994); Aldemita and Hodges (1996); and Hiei et al. (1997). For most cereal plants, embryogenic cells
 5 derived from immature scutellum tissues are the preferred cellular targets for transformation (Hiei et al. (1997) *supra*; Vasil (1994) *supra*). For transforming corn embryogenic cells derived from immature scutellar tissue using microprojectile bombardment, the A188XB73 genotype is the preferred genotype (Fromm et al. (1990) *supra*; Gordon-Kamm et al. (1990) *supra*). After microprojectile bombardment the tissues are selected on phosphinothricin to identify the
 10 transgenic embryogenic cells (Gordon-Kamm et al. (1990) *supra*). Transgenic plants are regenerated by standard corn regeneration techniques (Fromm et al. (1990) *supra*; Gordon-Kamm et al. (1990) *supra*). *Agrobacterium*-mediated transformation of switchgrass has also been reported by Somleva et al. (2002).

Example VIII. Morphological and Physiological Analysis

15 In these Examples, unless otherwise indicated, morphological and physiological traits are disclosed in comparison to control plants, including, for example, wild-type plants, plants that have not been transformed, or plants transformed with an “empty” expression vector (lacking a polynucleotide that has been introduced into an experimental plant) under the identical photoperiodic conditions where CO or other COLs can be activated, for example, the pots
 20 containing transgenic lines and control lines of *Arabidopsis* plants are maintained in a growth room under 8-hour light conditions (18 - 23°C, and 90 - 100 $\mu\text{E m}^{-2} \text{s}^{-1}$) for a period of 14 days. The light conditions are then shifted to 16h light (18 - 23°C, and 90 - 100 $\mu\text{E m}^{-2} \text{s}^{-1}$)/8h dark for an additional 4-7 days. Thus, a transformed plant that is described as large and/or drought tolerant is large and more tolerant to drought with respect to a control plant, the latter including
 25 wild-type plants, parental lines and lines transformed with a vector that does not contain a sequence of interest. When a plant is said to have a better performance than controls, it generally is larger, have greater yield, and/or show less stress symptoms than control plants. The better performing lines may, for example, have produced less anthocyanin, or are larger, greener, or more vigorous in response to a particular stress, as noted below. Better performance generally
 30 implies greater size or yield, or tolerance to a particular biotic or abiotic stress, less sensitivity to ABA, or better recovery from a stress (as in the case of a soil-based drought treatment) than controls.

Morphological analysis

Morphological analysis is performed to determine whether changes in transcription factor levels affect plant growth and development. This is primarily carried out on the T1 generation, when typically 10-20 independent lines are examined, in the case of Arabidopsis studies, for example. However, in cases where a phenotype requires confirmation or detailed
5 characterization, plants from subsequent generations are also analyzed.

Primary transformants are selected on Murashige and Skoog (MS) medium (Murashige T and Skoog F, 1962) with 0.3% sucrose and 50 mg/l kanamycin. T2 and later generation plants are selected in the same manner, except that kanamycin is used at 35 mg/l. In cases where lines carry a sulfonamide marker (as in all lines generated by super-transformation), seeds are
10 selected on MS medium with 0.3% sucrose and 1.5 mg/l sulfonamide. KO lines are usually germinated on plates without a selection. Seeds are cold-treated (stratified) on plates for 3 days in the dark (in order to increase germination efficiency) prior to transfer to growth cabinets. Initially, plates are incubated at 22°C under a light intensity of approximately 100 microEinsteins for 7 days. At this stage, transformants are green, possess the first two true
15 leaves, and are easily distinguished from bleached kanamycin or sulfonamide-susceptible seedlings. Resistant seedlings are then transferred onto soil (Sunshine potting mix). Following transfer to soil, trays of seedlings are covered with plastic lids for 2-3 days to maintain humidity while they become established. Plants are grown on soil under fluorescent light at an intensity of 70-95 microEinsteins and a temperature of 18-23°C. Light conditions consist of a 24-hour
20 photoperiod unless otherwise stated. In instances where alterations in flowering time are apparent, flowering may be re-examined under both 12-hour and 24-hour light to assess whether the phenotype is photoperiod dependent. Under typical 24-hour light growth conditions, the typical generation time (seed to seed) for Arabidopsis is approximately 14 weeks.

Because many aspects of *Arabidopsis* development are dependent on localized
25 environmental conditions, in all cases plants are evaluated in comparison to controls in the same flat. Controls for transgenic lines are transgenic plants harboring the empty transformation vector, selected on kanamycin or sulfonamide. Careful examination is made at the following stages: seedling (1 week), rosette (2-3 weeks), flowering (4-7 weeks), and late seed set (8-12 weeks). Seed is also inspected. Plants having no or few seeds are sterile. Seedling morphology is
30 assessed on selection plates. At all other stages, plants are macroscopically evaluated while growing on soil. All significant differences (including alterations in growth rate, size, leaf and flower morphology, coloration and flowering time) are recorded, but routine measurements are not be taken if no differences are apparent. In certain cases, stem sections are stained to reveal lignin distribution. In these instances, hand-sectioned stems are mounted in phloroglucinol

saturated 2M HCl (which stains lignin pink) and viewed immediately under a dissection microscope.

For a given gene-promoter combination, GAL4 fusion lines, RNAi lines etc., ten lines are typically examined in subsequent plate based physiology assays.

5 Plate Assays. Different plate-based physiological assays (shown below), representing a variety of abiotic and water-deprivation-stress related conditions, are used as a pre-screen to identify top performing lines (i.e. lines from transformation with a particular construct), that are generally then tested in subsequent soil based assays. Typically, ten lines are subjected to plate assays, from which the best three lines are selected for subsequent soil based assays.

10 In addition, a nutrient limitation assay can be used to find genes that allow more plant growth upon deprivation of nitrogen. Nitrogen is a major nutrient affecting plant growth and development that ultimately impacts yield and stress tolerance. These assays monitor primarily root but also rosette growth on nitrogen deficient media. In all higher plants, inorganic nitrogen is first assimilated into glutamate, glutamine, aspartate and asparagine, the four amino acids
15 used to transport assimilated nitrogen from sources (e.g. leaves) to sinks (e.g. developing seeds). This process may be regulated by light, as well as by C/N metabolic status of the plant. A C/N sensing assay is thus used to look for alterations in the mechanisms plants use to sense internal levels of carbon and nitrogen metabolites which could activate signal transduction cascades that regulate the transcription of N-assimilatory genes. To determine whether these mechanisms are
20 altered, we exploit the observation that wild-type plants grown on media containing high levels of sucrose (3%) without a nitrogen source accumulate high levels of anthocyanins. This sucrose-induced anthocyanin accumulation can be relieved by the addition of either inorganic or organic nitrogen. Glutamine is used as a nitrogen source since it also serves as a compound used to transport N in plants.

25 Growth assays. The following growth assays can be conducted with *Arabidopsis* transformed with expression construct comprising CORE promoter linked to a polynucleotide that is able to confer a desired trait: severe desiccation (a type of water deprivation assay), growth in cold conditions at 8° C, root development (visual assessment of lateral and primary roots, root hairs and overall growth), and phosphate limitation. For the nitrogen limitation assay, plants are
30 grown in 80% Murashige and Skoog (MS) medium in which the nitrogen source is reduced to 20 mg/L of NH₄NO₃. Note that 80% MS normally has 1.32 g/L NH₄NO₃ and 1.52 g/L KNO₃. For phosphate limitation assays, seven day old seedlings are germinated on phosphate-free MS medium in which KH₂PO₄ is replaced by K₂SO₄.

Unless otherwise stated, all experiments are to be performed with the *Arabidopsis thaliana* ecotype Columbia (col-0), soybean or maize plants. Assays are usually conducted on non-selected segregating T2 populations (in order to avoid the extra stress of selection). Control plants for assays on lines containing direct promoter-fusion constructs are Col-0 plants
5 transformed an empty transformation vector (pMEN65). Controls for 2-component lines (generated by supertransformation) are the background promoter-driver lines (i.e. promoter::LexA-GAL4TA lines), into which the supertransformations are initially performed.

Procedures

Growth assays may be conducted with *Arabidopsis* or other plant species (e.g., soy, maize, etc.). For example, growth assays may assess tolerance to severe desiccation (a type of
10 water deprivation assay), growth in cold conditions at 8° C, root development (visual assessment of lateral and primary roots, root hairs and overall growth), and phosphate limitation. Assays are usually conducted on non-selected segregating T2 populations in order to avoid the extra stress of selection. Control plants for assays on lines may include wild-type plants or plants
15 transformed with an empty transformation vector.

For chilling growth assays, seeds are germinated and grown for seven days on MS + Vitamins + 1% sucrose at 22° C and then transferred to chilling conditions at 8° C and evaluated after another 10 days and 17 days.

For severe desiccation (plate-based water deprivation) assays, seedlings are grown for 14
20 days on MS+ Vitamins + 1% Sucrose at 22° C. Plates are opened in the sterile hood for 3 h. for hardening and then seedlings are removed from the media and let dry for two hours in the hood. After this time the plants are transferred back to plates and incubated at 22° C for recovery. The plants are then evaluated after five days.

For the polyethylene glycol (PEG) hyperosmotic stress tolerance screen, plant seeds are
25 gas sterilized with chlorine gas for 2 h. The seeds are plated on each plate containing 3% PEG, 1/2 X MS salts, 1% phytigel, and 10 µg/ml glufosinate-ammonium (BASTA). Two replicate plates per seed line are planted. The plates are placed at 4° C for 3 days to stratify seeds. The plates are held vertically for 11 additional days at temperatures of 22° C (day) and 20° C (night). The photoperiod is 16 h. with an average light intensity of about 120 µmol/m²/s. The racks
30 holding the plates are rotated daily within the shelves of the growth chamber carts. At 11 days, root length measurements are made. At 14 days, seedling status is determined, root length is measured, growth stage is recorded, the visual color is assessed, pooled seedling fresh weight is measured, and a whole plate photograph is taken.

Germination assays may also be carried out with NaCl (150 mM, to measure tolerance to salt), sucrose (9.4%, to measure altered sugar sensing), cold (8° C) or heat (32° C). All germination assays are performed in aseptic conditions. Growing the plants under controlled temperature and humidity on sterile medium produces uniform plant material that has not been exposed to additional stresses (such as water stress) which could cause variability in the results obtained.

Prior to plating, seed for all experiments were surface sterilized in the following manner: (1) 5 minute incubation with mixing in 70% ethanol, (2) 20 minute incubation with mixing in 30% bleach, 0.01% triton-X 100, (3) 5X rinses with sterile water, (4) Seeds are re-suspended in 0.1% sterile agarose and stratified at 4° C for 3-4 days. All germination assays follow modifications of the same basic protocol. Sterile seeds may be sown on conditional media that has a basal composition of 80% MS + Vitamins. Plates may be incubated at 22° C under 24-hour light (120-130 $\mu\text{E m}^{-2} \text{s}^{-1}$) in a growth chamber. Evaluation of germination and seedling vigor may be performed five days after planting.

Chlorophyll content, an indicator of photosynthetic capacity, may be measured with a SPAD meter.

Wilt screen assay. Transgenic and wild-type soybean plants are grown in 5" pots in growth chambers. After the seedlings reach the V1 stage (the V1 stage occurs when the plants have one trifoliolate, and the unifoliolate and first trifoliolate leaves are unrolled), water is withheld and the drought treatment thus started. A drought injury phenotype score is recorded, in increasing severity of effect, as 1 to 4, with 1 designated no obvious effect and 4 indicating a dead plant. Drought scoring is initiated as soon as one plant in one growth chamber had a drought score of 1.5. Scoring continues every day until at least 90% of the wild type plants achieve scores of 3.5 or more. At the end of the experiment the scores for both transgenic and wild type soybean seedlings are statistically analyzed using Risk Score and Survival analysis methods (Glantz (2001); Hosmer and Lemeshow (1999)).

Water use efficiency (WUE). WUE is estimated by exploiting the observation that elements can exist in both stable and unstable (radioactive) forms. Most elements of biological interest (including C, H, O, N, and S) have two or more stable isotopes, with the lightest of these being present in much greater abundance than the others. For example, ^{12}C is more abundant than ^{13}C in nature ($^{12}\text{C} = 98.89\%$, $^{13}\text{C} = 1.11\%$, $^{14}\text{C} = <10^{-10}\%$). Because ^{13}C is slightly larger than ^{12}C , fractionation of CO_2 during photosynthesis occurs at two steps:

1. $^{12}\text{CO}_2$ diffuses through air and into the leaf more easily;

2. $^{12}\text{CO}_2$ is preferred by the enzyme in the first step of photosynthesis, ribulose biphosphate carboxylase/oxygenase.

WUE has been shown to be negatively correlated with carbon isotope discrimination during photosynthesis in several C3 crop species. Carbon isotope discrimination has also been
5 linked to drought tolerance and yield stability in drought-prone environments and has been successfully used to identify genotypes with better drought tolerance. $^{13}\text{C}/^{12}\text{C}$ content is measured after combustion of plant material and conversion to CO_2 , and analysis by mass spectroscopy. With comparison to a known standard, ^{13}C content is altered in such a way as to suggest that the expression of the polynucleotide under the CORE promoter improves water use
10 efficiency.

Another potential indicator of WUE is stomatal conductance, that is, the extent to which stomata are open.

Data interpretation

At the time of evaluation, plants are typically given one of the following scores:

15 (++) Substantially enhanced performance compared to controls. The phenotype is very consistent and growth is significantly above the normal levels of variability observed for that assay.

 (+) Enhanced performance compared to controls. The response is consistent but is only moderately above the normal levels of variability observed for that assay.

20 (wt) No detectable difference from wild-type controls.

 (-) Impaired performance compared to controls. The response is consistent but is only moderately above the normal levels of variability observed for that assay.

 (- -) Substantially impaired performance compared to controls. The phenotype is consistent and growth is significantly above the normal levels of variability observed for
25 that assay.

 (n/d) Experiment failed, data not obtained, or assay not performed.

Soil Drought (Clay Pot)

The soil drought assay (performed in clay pots) is based on that described by Haake et al. (2002).

30 Procedures. In the current procedure, seedlings are first germinated on selection plates containing either kanamycin or sulfonamide. Seeds are sterilized by a 2 minute ethanol treatment followed by 20 minutes in 30% bleach / 0.01% Tween and five washes in distilled water. Seeds are sown to MS agar in 0.1% agarose and stratified for three days at 4° C, before transfer to growth cabinets with a temperature of 22 °C. After seven days of growth on selection

plates, seedlings are transplanted to 3.5 inch diameter clay pots containing 80 grams of a 50:50 mix of vermiculite:perlite topped with 80 grams of ProMix. Typically, each pot contains 14 seedlings, and plants of the transgenic line being tested are in separate pots to the wild-type controls. Pots containing the transgenic line versus control pots are interspersed in the growth room, maintained under 24-hour light conditions (18 - 23°C, and 90 - 100 $\mu\text{E m}^{-2} \text{s}^{-1}$) and watered for a period of 14 days. Water is then withheld and pots are placed on absorbent paper for a period of 8-10 days to apply a drought treatment. After this period, a visual qualitative “drought score” from 0-6 is assigned to record the extent of visible drought stress symptoms. A score of “6” corresponds to no visible symptoms whereas a score of “0” corresponds to extreme wilting and the leaves having a “crispy” texture. At the end of the drought period, pots are re-watered and scored after 5-6 days; the number of surviving plants in each pot is counted, and the proportion of the total plants in the pot that survive is calculated.

Analysis of results. In a given experiment, 6 or more pots of a transgenic line with 6 or more pots of the appropriate control are typically compared. The mean drought score and mean proportion of plants surviving (survival rate) are calculated for both the transgenic line and the wild-type pots. In each case a p -value* is calculated, which indicates the significance of the difference between the two mean values.

Calculation of p-values. For the assays where control and experimental plants are in separate pots, survival is analyzed with a logistic regression to account for the fact that the random variable is a proportion between 0 and 1. The reported p -value is the significance of the experimental proportion contrasted to the control, based upon regressing the logit-transformed data.

Drought score, being an ordered factor with no real numeric meaning, is analyzed with a non-parametric test between the experimental and control groups. The p -value is calculated with a Mann-Whitney rank-sum test.

Disease Resistance

Resistance to pathogens, such as *Sclerotinia sclerotiorum* and *Botrytis cinerea*, can be assessed in plate-based assays. Unless otherwise stated, all experiments are performed with the *Arabidopsis thaliana* ecotype Columbia (Col-0). Control plants for assays on lines containing direct promoter-fusion constructs are wild-type plants or Col-0 plants transformed an empty transformation vector (pMEN65).

Prior to plating, seed for all experiments are surface sterilized in the following manner: (1) 5 minute incubation with mixing in 70 % ethanol; (2) 20 minute incubation with mixing in

30% bleach, 0.01% Triton X-100; (3) five rinses with sterile water. Seeds are resuspended in 0.1% sterile agarose and stratified at 4 °C for 2-4 days.

Sterile seeds are sown on starter plates (15 mm deep) containing 50% MS solution, 1% sucrose, 0.05% MES, and 1% Bacto-Agar. 40 to 50 seeds are sown on each plate. Plates are incubated at 22 °C under 24-hour light (95-110 $\mu\text{E m}^{-2} \text{s}^{-1}$) in a germination growth chamber. On day 10, seedlings are transferred to assay plates (25 mm deep plates with medium minus sucrose). On day 14, seedlings are inoculated (specific method below). After inoculation, plates are put in a growth chamber under a 12-hour light/12-hour dark schedule. Light intensity is lowered to 70-80 $\mu\text{E m}^{-2} \text{s}^{-1}$ for the disease assay.

Sclerotinia inoculum preparation. A *Sclerotinia* liquid culture is started three days prior to plant inoculation by cutting a small agar plug (1/4 sq. inch) from a 14- to 21-day old *Sclerotinia* plate (on Potato Dextrose Agar; PDA) and placing it into 100 ml of half-strength Potato Dextrose Broth. The culture is allowed to grow in the Potato Dextrose Broth at room temperature under 24-hour light for three days. On the day of seedling inoculation, the hyphal ball is retrieved from the medium, weighed, and ground in a blender with water (50 ml/gm tissue). After grinding, the mycelial suspension is filtered through two layers of cheesecloth and the resulting suspension is diluted 1:5 in water. Plants are inoculated by spraying to run-off with the mycelial suspension using a Preval aerosol sprayer.

Botrytis inoculum preparation. *Botrytis* inoculum is prepared on the day of inoculation. Spores from a 14- to 21-day old plate (on PDA) are resuspended in a solution of 0.05% glucose, 0.03M KH_2PO_4 to a final concentration of 10^4 spores/ml. Seedlings are inoculated with a Preval aerosol sprayer, as with *Sclerotinia* inoculation.

Resistance to *Erysiphe cichoracearum* is assessed in a soil-based assay. *Erysiphe cichoracearum* is propagated on a *pad4* mutant line in the Col-0 background, which is highly susceptible to *Erysiphe* (Reuber et al. (1998) *Plant J.* 16: 473-485), or on squash plants, since this particular strain also parasitizes squash. Inocula are maintained by using a small paintbrush to dust conidia from a 2-3 week old culture onto 4-week old plants. For the assay, seedlings are grown on plates for one week under 24-hour light in a germination chamber, then transplanted to soil and grown in a walk-in growth chamber under a 12-hour light/12-hour dark light regimen, 70% humidity. Each line is transplanted to two 13 cm square pots, nine plants per pot. In addition, three control plants are transplanted to each pot, for direct comparison with the test line. Approximately 3.5 weeks after transplanting, plants are inoculated using settling towers, as described by Reuber et al., (1998). Generally, three to four heavily infested leaves are used per pot for the disease assay. Level of fungal growth is evaluated eight to ten days after inoculation.

Example IX. Regulating expression of polynucleotides encoding RNA species which act at a non-protein level

In addition to use of the CORE promoters to regulate the expression of a polynucleotide encoding a polypeptide, these promoters can also be used to regulate the expression of a polynucleotide encoding a non-coding RNA species (that is, one which acts at a non-protein level), such as a microRNA, a microRNA precursor, or a sequence designed to act through RNA interference (RNAi). For example, a substantial number of microRNA (miRNA) species have been implicated in stress responses and these molecules have been shown to be involved in the control of many aspects of plant growth and development (Bartel and Bartel (2003); Aukerman and Sakai (2003).; Bartel (2004); Juarez et al. (2004); Bowman (2004); Sunkar and Zhu (2004)).

It should be noted that, for particular families of highly related plant polypeptides such as transcription factors, overexpression of one or more of the family members produces a comparable phenotype to that obtained from reducing expression (for example, by mutation or knockdown approaches such as antisense or RNA interference) of one or more of the family members. For instance, overexpression of the CBF family proteins has been widely demonstrated to confer tolerance to drought and low temperature stress (e.g., Jaglo et al. (2001). Nonetheless, Novillo et al. (2004) show that homozygous *cbf2* mutant *Arabidopsis* plants carrying a disruption in the CBF2 gene also exhibit enhanced freezing tolerance. Such results can be accounted for by cross-regulation between the genes encoding different transcription factor family members. In the study by Novillo et al, (2004) *supra*, CBF2 was shown to be a negative transcriptional regulator of the *CBF1* and *CBF3* genes. Comparable mechanisms likely account for the fact that stress tolerance has been observed from both overexpression and from knockdown approaches with certain NF-Y family genes.

EXAMPLE X. Field plot designs, harvesting and yield measurements of soybean and maize.

A field plot of soybeans with any of various configurations and/or planting densities may be used to measure crop yield. For example, 30-inch-row trial plots consisting of multiple rows, for example, four to six rows, may be used for determining yield measurements. The rows may be approximately 20 feet long or less, or 20 meters in length or longer. The plots may be seeded at a measured rate of seeds per acre, for example, at a rate of about 100,000, 200,000, or 250,000 seeds/acre, or about 100,000-250,000 seeds per acre (the latter range is about 250,000 to 620,000 seeds/hectare).

Harvesting may be performed with a small plot combine or by hand harvesting. Harvest yield data are generally collected from inside rows of each plot of soy plants to measure yield,

for example, the innermost inside two rows. Soybean yield may be reported in bushels (60 pounds) per acre. Grain moisture and test weight are determined; an electronic moisture monitor may be used to determine the moisture content, and yield is then adjusted for a moisture content of 13 percent (130 g/kg) moisture. Yield is typically expressed in bushels per acre or tonnes per hectare. Seed may be subsequently processed to yield component parts such as oil or carbohydrate, and this may also be expressed as the yield of that component per unit area.

For determining yield of maize, varieties are commonly planted at a rate of 15,000 to 40,000 seeds per acre (about 37,000 to 100,000 seeds per hectare), often in 30 inch rows. A common sampling area for each maize variety tested is with rows of 30 in. per row by 50 or 100 or more feet. At physiological maturity, maize grain yield may also be measured from each of number of defined area grids, for example, in each of 100 grids of, for example, 4.5 m² or larger. Yield measurements may be determined using a combine equipped with an electronic weigh bucket, or a combine harvester fitted with a grain-flow sensor. Generally, center rows of each test area (for example, center rows of a test plot or center rows of a grid) are used for yield measurements. Yield is typically expressed in bushels per acre or tonnes per hectare. Seed may be subsequently processed to yield component parts such as oil or carbohydrate, and this may also be expressed as the yield of that component per unit area.

Example XI: Sequences that confer significant improvements to non-*Arabidopsis* species

The function of CORE promoter sequences has been analyzed and may be further characterized and the sequences may be incorporated into crop plants. The ectopic overexpression of nucleotide sequences encoding a polypeptide, or any other sequence that may confer an improved or desirable trait may be regulated using CORE regulatory elements found in the Sequence Listing. In addition to these sequences, it is expected that newly discovered polynucleotide sequences from, for example, other species having similar sequences, may be closely related to polynucleotide sequences found in the Sequence Listing can also confer improved traits in a similar manner to the sequences found in the Sequence Listing, when transformed into a any of a considerable variety of plants of different species, and including dicots and monocots. The polynucleotide and polypeptide sequences derived from monocots (e.g., the rice sequences) may be used to transform both monocot and dicot plants, and those derived from dicots (e.g., the *Arabidopsis* and soy genes) may be used to transform either group, although a preferred embodiment may include a sequence may transformed into a plant from the same major clades of angiosperm as that from which the sequence is derived.

The results presented in these Examples indicate that polypeptides that confer an improved or desirable trait may do so when they are expressed under the regulatory control of a

CORE promoter sequence, without having a significant adverse impact on plant morphology and/or development. The lines that display useful traits may be selected for further study or commercial development.

Monocotyledonous plants, including rice, corn, wheat, rye, sorghum, barley, switchgrass, *Miscanthus*, and others, or eudicots, including soy, cotton, canola, tomato, alfalfa, poplar, and others, may be transformed with a plasmid containing a polynucleotide of interest. The polynucleotide sequence may include dicot or monocot-derived sequences such as those presented herein. These polynucleotide sequences may be cloned into an expression vector containing a kanamycin-resistance marker, and then expressed in an inducible manner under the regulatory control of a CORE promoter sequence.

The cloning vector may be introduced into monocots or eudicots by, for example, means described in the previous Examples, including direct DNA transfer or *Agrobacterium tumefaciens*-mediated transformation. The latter approach may be accomplished by a variety of means, including, for example, that of U.S. Patent No. 5,591,616, in which monocotyledon callus is transformed by contacting de-differentiating tissue with the *Agrobacterium* containing the cloning vector.

The sample tissues are immersed in a suspension of 3×10^9 cells of *Agrobacterium* containing the cloning vector for 3-10 minutes. The callus material is cultured on solid medium at 25° C in the dark for several days. The calli grown on this medium are transferred to Regeneration medium. Transfers are continued every 2-3 weeks (2 or 3 times) until shoots develop. Shoots are then transferred to Shoot-Elongation medium every 2-3 weeks. Healthy looking shoots are transferred to rooting medium and after roots have developed, the plants are placed into moist potting soil.

The transformed plants are then analyzed for the presence of the NPTII gene/ kanamycin resistance by ELISA, using the ELISA NPTII kit from 5Prime-3Prime Inc. (Boulder, CO).

Northern blot analysis, RT-PCR or microarray analysis of the regenerated, transformed plants may be used to show expression of a polynucleotide that is capable of conferring an improved trait, or increased yield or quality, in the transformed plants.

To verify the ability to confer an improved or desirable trait, mature plants or seedling progeny of these plants expressing an equilog gene may be challenged using methods described herein. By comparing control plants and the transgenic plants, the latter are identified as having the improved or desirable trait. As an example of a first step to determine, for example, a water deficit- or water deprivation-related tolerance, seeds of transgenic plants may be subjected to germination assays to measure sucrose sensing. For example, sterile eudicot

seeds including, but not limited to soybean and alfalfa, are sown on 80% MS medium plus vitamins with 9.4% sucrose; control media lack sucrose. All assay plates are then incubated at 22° C under 24-hour light, 120-130 μ Ein/m²/s, in a growth chamber. Evaluation of germination and seedling vigor is then conducted three days after planting. Plants overexpressing proteins that confer improved tolerance to water deficit or water deprivation, where the proteins are under the regulatory control of CORE promoters, may be found to be more tolerant to high sucrose by having better germination, longer radicles, and more cotyledon expansion, than control plants in the presence of the sugar concentration. It is expected that closely related and structurally similar promoter sequences, may also confer the improved trait, such as altered sugar sensing or improved hyperosmotic stress tolerance of this example.

It is expected that the same methods may be applied to identify other useful and valuable promoter sequences, and the sequences may be derived from a diverse range of species.

All publications and patent applications mentioned in this specification are herein incorporated by reference to the same extent as if each individual publication or patent application was specifically and individually indicated to be incorporated by reference.

References

- Aldemita and Hodges (1996) *Planta* 199: 612-617
 Aukerman and Sakai (2003) *Plant Cell* 15: 2730-2741
 Bartel (2004) *Cell* 116: 281-297
 Bartel and Bartel (2003) *Plant Physiol.* 132: 709-717
 Bevan (1984) *Nucleic Acids Res.* 12: 8711-8721
 Boote, eds (Madison, WI.: Crop Sci. Soc. America and Amer. Soc. Agron).
 Bowman (2004) *Bioessays* 26: 938-942
 Capecchi, *Cell*, 22(2): 479-488, 1980
 Cardoza et al. (2003) *Plant Cell Rep.* 21: 599-604
 Cassas et al. (1993) *Proc. Natl. Acad. Sci. USA* 90: 11212-11216
 Charest et al. (1988) *Theor. Appl. Genet.* 75: 438-445
 Chia et al. (2008). *J. Exp. Bot.* 59, 2735-2748.
 Christou et al. (1987) *Proc. Natl. Acad. Sci. USA*
 Christou (1991) *Bio/Technol.* 9:957-962
 Christou et al. (1992) *Plant. J.* 2: 275-281
 Clapp (1993) *Clin. Perinatol.* 20(1): 155-168
 Curiel et al. (1992) *Hum. Gen. Ther.* 3(2):147-154
 Daly et al. (2001) *Plant Physiol.* 127: 1328-1333

- De Block et al. (1987) *Plant Physiol.* 91: 694-701
- Deshayes et al. (1985) *EMBO J.* 4: 2731-2737
- Donn et al. (1990) in Abstracts of VIIth International Congress on Plant Cell and Tissue Culture IAPTC, A2-38: 53
- 5 Draper et al. (1982) *Plant Cell Physiol.* 23: 451-458
- D'Halluin et al. (1992) *Plant Cell* 4: 1495-1505
- Edmeades et al. (2000). In Physiology and Modeling Kernel Set in Maize., M.E. Westgate and K.J. Eglitis and Anderson, *Biotechniques*, 6(7): 608-614, 1988
- Fromm et al. (1985) *Proc. Natl. Acad. Sci. USA*, 82(17): 5824-5828
- 10 Fromm et al. (1990) *Bio/Technol.* 8: 833-839
- Fynan et al. (1993) *Proc. Natl. Acad. Sci. USA*, 90(24): 11478-11482
- Gelvin et al. (1990) Plant Molecular Biology Manual, Kluwer Academic Publishers
- Gordon-Kamm et al. (1990) *Plant Cell* 2: 603-618
- Griffiths, S., Dunford, R. P., Coupland, G. & Laurie, D. A. (2003), *Plant Physiol*, 131: 1855-67.
- 15 Gruber et al. (1993) in Methods in Plant Molecular Biology and Biotechnology, p. 89-119
- Glick and Thompson (1993) Methods in Plant Molecular Biology and Biotechnology. CRC Press., Boca Raton, FL
- Glantz (2001) Relative risk and risk score, in Primer of Biostatistics. 5th ed., McGraw Hill/Appleton and Lange, pub.
- 20 Graham and Van der Eb (1973) *Virology*, 54(2): 536-539
- Haake et al. (2002) *Plant Physiol.* 130: 639-648
- Hess, *Intern Rev. Cytol.* (1987) 107: 367
- Herrera-Estrella et al. (1983) *Nature* 303: 209
- Hiei et al. (1997) *Plant Mol. Biol.* 35:205-218
- 25 Hain et al. (1985) *Mol. Gen. Genet.* 199: 161-168
- Hiei et al. (1994) *Plant J.* 6:271-282
- Hiei et al. (1997) *Plant Mol. Biol.* 35:205-218
- Helliwell, C. A., Wood, C. C., Robertson, M., James Peacock, W. & Dennis, E. S. (2006) *Plant J.* 46 183-92.
- 30 Hosmer and Lemeshow (1999) Applied Survival Analysis: Regression Modeling of Time to Event Data. John Wiley & Sons, Inc., Publisher.
- Johnston and Tang (1994) *Methods Cell Biol.* 43(A): 353-365
- Ishida (1990) *Nature Biotechnol.* 14: 745-750
- Jaglo et al. (2001) *Plant Physiol.* 127: 910-917

- Juarez et al. (2004) *Nature* 428: 84-88
- Klee (1985) *Bio/Technology* 3: 637-642
- Klein et al. (1987) *Nature* 327: 70-73
- Kojima, S., Takahashi, Y., Kobayashi, Y., Monna, L., Sasaki, T., Araki, T. & Yano, M. (2002)
- 5 *Plant Cell Physiol*, 43: 1096-105
- Koornneef et al. (1986) in Tomato Biotechnology: Alan R. Liss, Inc., 169-178
- Kumagai et al. (2008) *Biosci.Biotechnol.Biochem* 72: 1539-1549
- Liu et al. (2001) *Plant Physiol* 125: 1821-1830
- Lu, et al. (1993) *J. Exp. Med.* 178(6): 2089-2096
- 10 Luo, et al. (1988) *Plant Mol Biol. Reporter*, 6: 165
- Martinez-Zapater et al. (1994) In: Meyerowitz EM, Somerville CR (eds) Arabidopsis. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, pp 403-433
- Miki et al. (1993) in Methods in Plant Molecular Biology and Biotechnology, p. 67-88, Glick and Thompson, eds., CRC Press, Inc., Boca Raton;
- 15 Michaels, S. D., Himmelblau, E., Kim, S. Y., Schomburg, F. M. & Amasino, R. M. (2005) *Plant Physiol.* 137: 149-56
- Miller et al. (2008). *Planta* 227: 1377-1388.
- Mouradov A, Cremer F, Coupland G (2002) *Plant Cell Suppl* S111-S130
- Murashige T and Skoog F (1962) *Physiol Plant* 15(3): 473-497
- 20 Novillo et al. (2004) *Proc. Natl. Acad. Sci. USA* 101: 3985-3990
- Nemoto et al. (2003). *Plant Journal* 36, 82-93
- Neuhaus, et al. (1987) *Theor. Appl. Genet.* 75: 30
- Pena, et al. (1987) *Nature* 325: 274
- Pua et al. (1987) *Biotechnol.* 5: 815-817
- 25 Putterill et al. (1995) *Cell* 80:847-857.
- Putterill et al., (2004) *Bioessays* 26, 363-73.
- Radke et al. (1988) *Theor. Appl. Genet.* 75: 685-694
- Ratcliffe and Riechmann (2002) *Curr. Issue Mol. Biol.* 77-91
- Robson et al. (2001) *Plant Journal* 28: 919-631
- 30 Reuber et al. *Plant J.* (1998) 16: 473-485
- Rieger et al. (1976) *Glossary of Genetics and Cytogenetics: Classical and Molecular*, 4th ed.
- Robert et al., (1998) Conserved structure and function of the *Arabidopsis* flowering time gene *CONSTANS* in *Brassica napus*. *Plant Mol Biol* 37: 763-772 Springer Verlag, Berlin
- Sanford et al. (1987) *Part. Sci. Technol.* 5:27-37

- Sanford (1993) *Methods Enzymol.* 217: 483-509
- Sheldon, C. C., Rouse, D. T., Finnegan, E. J., Peacock, W. J. & Dennis, E. S. (2000) *Proc Natl Acad Sci U S A*, 97: 3753-8
- Simpson GG, Dean C (2002) *Arabidopsis*, the rosetta stone of flowering time. *Science* 296: 285–
5 289
- Spencer et al. (1994) *Plant Mol. Biol.* 24: 51-61
- Somleva et al. (2002) *Crop Sci.* 42: 2080-2087.
- Stewart et al. (1996) *Plant Physiol.* 112: 115-120
- Sunkar and Zhu (2004) *Plant Cell* 16: 2001-2019
- 10 Suarez-Lopez et al. (2001) *Nature* 410:1116–1120
- Tamaki, S., Matsuo, S., Wong, H. L., Yokoi, S. & Shimamoto, K., *Science*, 316 (2007) 1033-6.
- Turck, F., Fornara, F. & Coupland, G., *Annu Rev Plant Biol*, 59 (2008) 573-94.
- Torok, M. & Etkin, L. D. (2001) *Differentiation* 67: 63-71.
- Vasil et al. (1992) *Bio/Technol.* 10:667-674
- 15 Vasil et al. (1993) *Bio/Technol.* 11:1553-1558
- Vasil (1994) *Plant Mol. Biol.* 25: 925-937
- Wagner et al. (1992) *Proc. Natl. Acad. Sci. USA*, 89(13): 6099-6103
- Weissbach and Weissbach (1989) *Methods for Plant Molecular Biology*, Academic Press
- Wan and Lemeaux (1994) *Plant Physiol.* 104: 37-48
- 20 Wenkel et. al. (2006) *Plant Cell* 18, 2971-84
- Weeks et al. (1993) *Plant Physiol.* 102:1077-1084
- Wong and Neumann (1982) *Biochim. Biophys. Res. Commun.* 107(2): 584-587
- Yano M, Katayose Y, Ashikara M, Yamanouchi U, Monna L, Fuse T, Baba T, Yamamoto K,
Umehara Y, Nagamura Y et al. (2000) *Plant Cell* 12: 2473–2483
- 25 Zatloukal, et al. (1992) *Ann. N.Y. Acad. Sci.* 660: 136-153
- Zhang et al. (1991) *Bio/Technology* 9: 996-997
- Zhang et al. (2004) *Plant Physiol* 135:615–621.
- Zhou, et al. (1983) *Methods in Enzymology* 101: 433

30 All publications and patent applications mentioned in this specification are herein incorporated by reference to the same extent as if each individual publication or patent application was specifically and individually indicated to be incorporated by reference.

The present invention is not limited by the specific embodiments described herein. The invention now being fully described, it will be apparent to one of ordinary skill in the art that

many changes and modifications can be made thereto without departing from the spirit or scope of the appended claims. Modifications that become apparent from the foregoing description and accompanying figures fall within the scope of the claims.

What is claimed is:

1. An isolated polynucleotide comprising a promoter sequence having one or more subsequences, any of which has a percentage identity with any of SEQ ID NOs: 101, 102, 7-10, 13, 16-20,

5 wherein the percentage identity is selected from the group consisting of at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% and 100%;

10 wherein the promoter sequence can be bound by a COL polypeptide and as a result of said binding, regulates transcription.

2. An recombinant polynucleotide operably linked to a transcribable polynucleotide, wherein the recombinant polynucleotide comprises a continuous region of at least 8, 9, 10, 13, 17, 20, 30, 35, 40, 50, 75, 100, 125, 150, 191, 200, 251, 311, 491, 496, 525, 550, 575, 600, 650, 700, 750, 800, 900, 950, 1000, 1050, 1100, 1150, 1200, 1500, 1633, 1639 or 1800 base pairs of
15 SEQ ID NO: 1, 14, 15, or 21.

3. The isolated polynucleotide of claim 1, wherein said promoter sequence comprises the nucleotide sequence set forth in SEQ ID NO: 16 or a complement thereof.

4. The isolated polynucleotide of claim 1, wherein the promoter sequence regulates expression of a transcribable polynucleotide of interest in a plant cell.

20 5. An expression vector comprising the polynucleotide of claim 3, wherein the expression vector regulates expression of a transcribable polynucleotide of interest in a plant cell.

6. The expression vector of claim 5, wherein the transcribable polynucleotide encodes a polypeptide.

25 7. The expression vector of claim 5, wherein the transcribable polynucleotide is a reporter gene.

8. The expression vector of claim 6, wherein the polypeptide controls the phenotype of a plant trait selected from the group consisting of enhanced resistance to abiotic stress, enhanced resistance to biotic stress, enhanced disease resistance, enhanced floral induction, enhanced
30 floral repression and reduced fertility.

9. The expression vector of claim 6, wherein the polypeptide is selected from the group consisting of SEQ ID NOs: 25-69.

10. An expression vector comprising said promoter sequence of claim 3 operably linked to SEQ ID NO: 22 that encodes SEQ ID NO: 23, wherein SEQ ID NO: 23 regulates the transcription of a polynucleotide sequence of interest operably linked to SEQ ID NO: 24.

11. An expression vector comprising a promoter sequence operably linked to a transcribable polynucleotide, wherein the promoter sequence comprises one or multiple copies of SEQ ID NO: 101, 102, 7-10, 13, 16-20 or any combinations thereof; wherein said promoter sequence regulates expression of a transcribable polynucleotide of interest in a plant cell.

12. A transgenic plant transformed with an isolated polynucleotide comprising a promoter sequence having a percentage identity with any of SEQ ID NOs: 101, 102, 1 - 4, 7-10, 13- 21 or a functional part thereof;

wherein the percentage identity is selected from the group consisting of at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% and 100%;

wherein the functional part is a DNA sequence that can be bound by a COL polypeptide and as a result of said binding, regulates transcription.

13. A transgenic plant cell made from the transgenic plant in claim 12.

14. A transgenic seed from the transgenic plant in claim 12, wherein the seed comprises a promoter sequence having a percentage identity with any of SEQ ID NOs: 101, 102, 1 - 4, 7-10, 13- 21 or a functional part thereof,

wherein the percentage identity is selected from the group consisting of at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% and 100%;

wherein the functional part is a DNA sequence that can be bound by a COL polypeptide and as a result of said binding, regulates transcription.

15. A transgenic plant transformed with the expression vector in claim 11, wherein the expression of the transcribable polynucleotide of interest confers a desired trait to the transgenic plant.

16. A progeny plant from the transgenic plant in claim 12, wherein the progeny comprises a promoter sequence having a percentage identity with any SEQ ID NOs: 101, 102, 1 - 4, 7-10, 13- 21 or a functional part thereof,

5 wherein the percentage identity is selected from the group consisting of at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% and 100%;

wherein the functional part is a DNA sequence that can be bound by a COL polypeptide and as a result of said binding, regulates transcription.

10 17. A method of producing a transgenic plant with an altered trait relative to a control plant, the method comprising introducing into a plant an expression vector comprising a polynucleotide that comprises a promoter sequence having one or more subsequences, any of which has a percentage identity with any of SEQ ID NOs: 101, 102, 1 - 4, 7-10, 13- 21 or a functional part thereof;

15 wherein the percentage identity is selected from the group consisting of at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% and 100%;

20 wherein the functional part is a DNA sequence that can be bound by a COL polypeptide and as a result of said binding, regulates transcription;

wherein the polynucleotide regulates expression of a transcribable polynucleotide molecule that confers an altered trait.

18. The method of claim 17, wherein the altered trait is the enhanced abiotic stress tolerance or biotic stress resistance that occurs following transition to reproductive development.

25 19. A method of claim 17, wherein the altered trait is selected from the group of enhanced floral induction, enhanced floral repression, reduced fertility, and enhanced disease resistance.

30 20. A method of conferring a desired trait to transgenic plants, the method comprising introducing into a plant an expression vector comprising a promoter sequence operably linked to a transcribable polynucleotide,

wherein the promoter sequence comprises one or multiple copies of SEQ ID NO: 101, 102, 7-10, 13, 16-20 or any combination thereof;

wherein the promoter sequence regulates expression of a transcribable polynucleotide of interest in a plant cell.

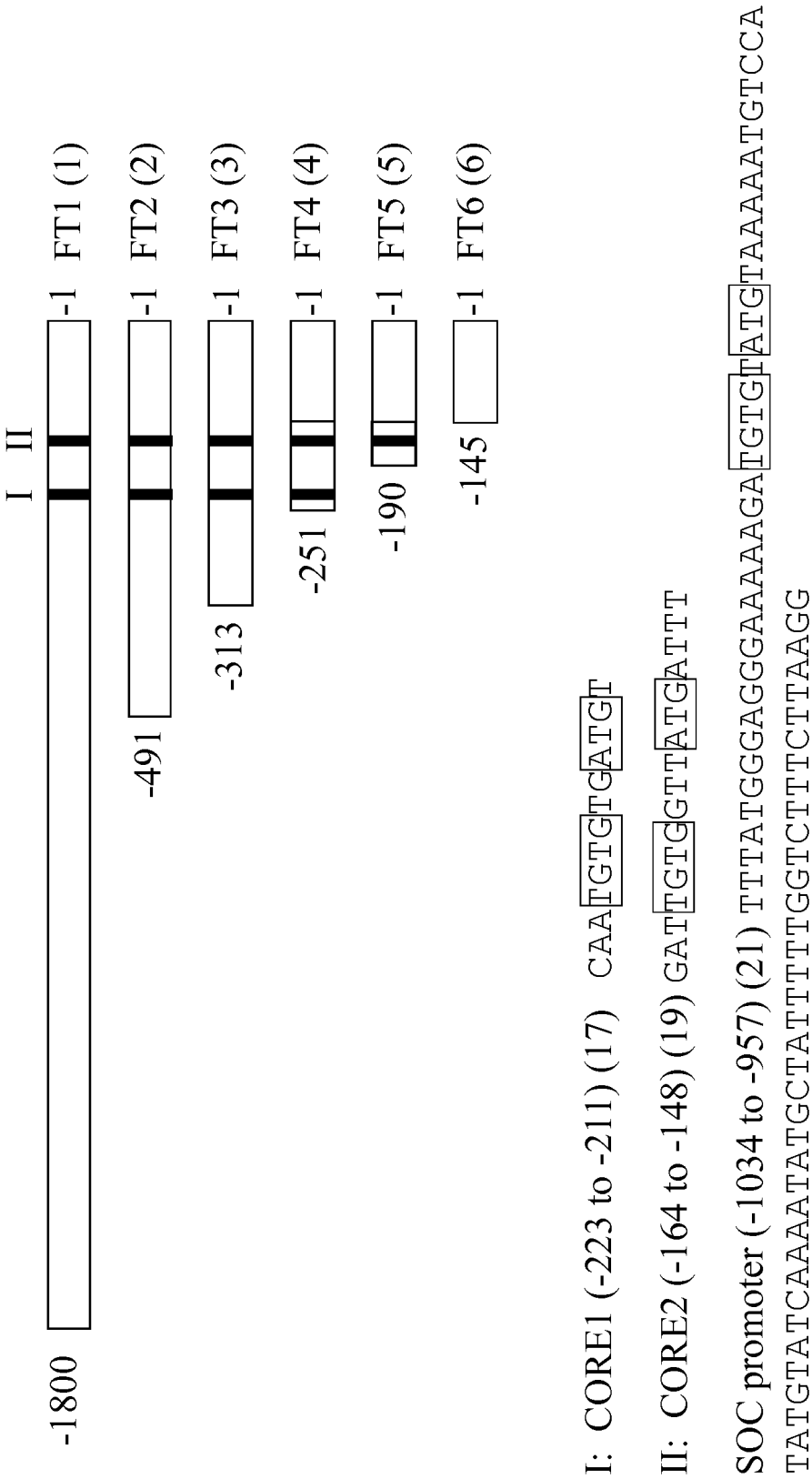


FIG. 1A

PAP2 promoter (-141 to -10, relative to start codon) (88)

CCTATTGCTCA[CATGCTCACAT]TTATCAATAAGTCTCCAGGATTTG
GGAAGCCACAATAACCCCTATTCCTCGGCCTTTTTAAAGTTTTT
AGAAATAATCCGATAAAATACTTTTATATAATTTTT

AOX1 promoter (-297 to -110, relative to start codon) (89)

CCTCAGTCAAAGAAAGTCGTAAACATAGTCTCTCT[CATGACCACA]AGGG
TAACACAGTCATCCCTAAATATAAACCCACACAAGAAACTGTTACTTTA
TACACGTGTCATAGTCTCATTTACATCTACGTGAAGAGTTTCGATCATCAA
CCGTTCGTTTCTTACTATATAAACCTTGCTCGAGACC

Complementary sequence of the CORE motif : CATN₍₁₋₃₎ CACA (87)

FIG. 1B

3XCORE1A	(7)	GCAATGTGTGATGTA	CGTAGCAATGTGTGATGTA	CGTAGCAATGTGTGATGTA	CGTAGG
3XCORE1B	(8)	GCAATGTGTGATGTA	AA TAGCAATGTGTGATGTA	AA TAGCAATGTGTGATGTA	AA TAGG

FIG. 2A

CORE1	(17)	CAATGTGTGATGT	-----
4xCORE1	(18)	CAATGTGTGATGT	CAATGTGTGATGTCAATGTGTGATGT-----
CORE3	(13)	GCAATGTGTGATGT	CAATGTGTGATGTGCAATGTGTGATGTGATGTGGATT
CORE2	(19)	-----	-----GATT
4xCORE2	(20)	-----	-----GATT
CORE1	(17)	-----	-----
4xCORE1	(18)	-----	-----
CORE3	(13)	GTGGTTATGATTT	GATTGTGGTTATGATTTGGATTGTGGTTATGATTTGATTGTGGTTA
CORE2	(19)	GTGGTTATGATTT	-----
4xCORE2	(20)	GTGGTTATGATTT	GATTGTGGTTATGATTTGGATTGTGGTTATGATTTGATTGTGGTTA
CORE1	(17)	-----	-----
4xCORE1	(18)	-----	-----
CORE3	(13)	TGATTTG	-----
CORE2	(19)	-----	-----
4xCORE2	(20)	TGATTT	-----

FIG. 2C

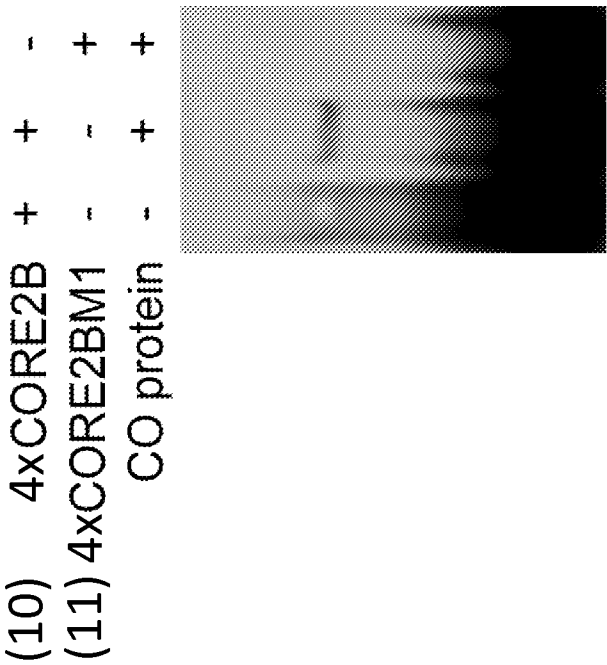


FIG. 3

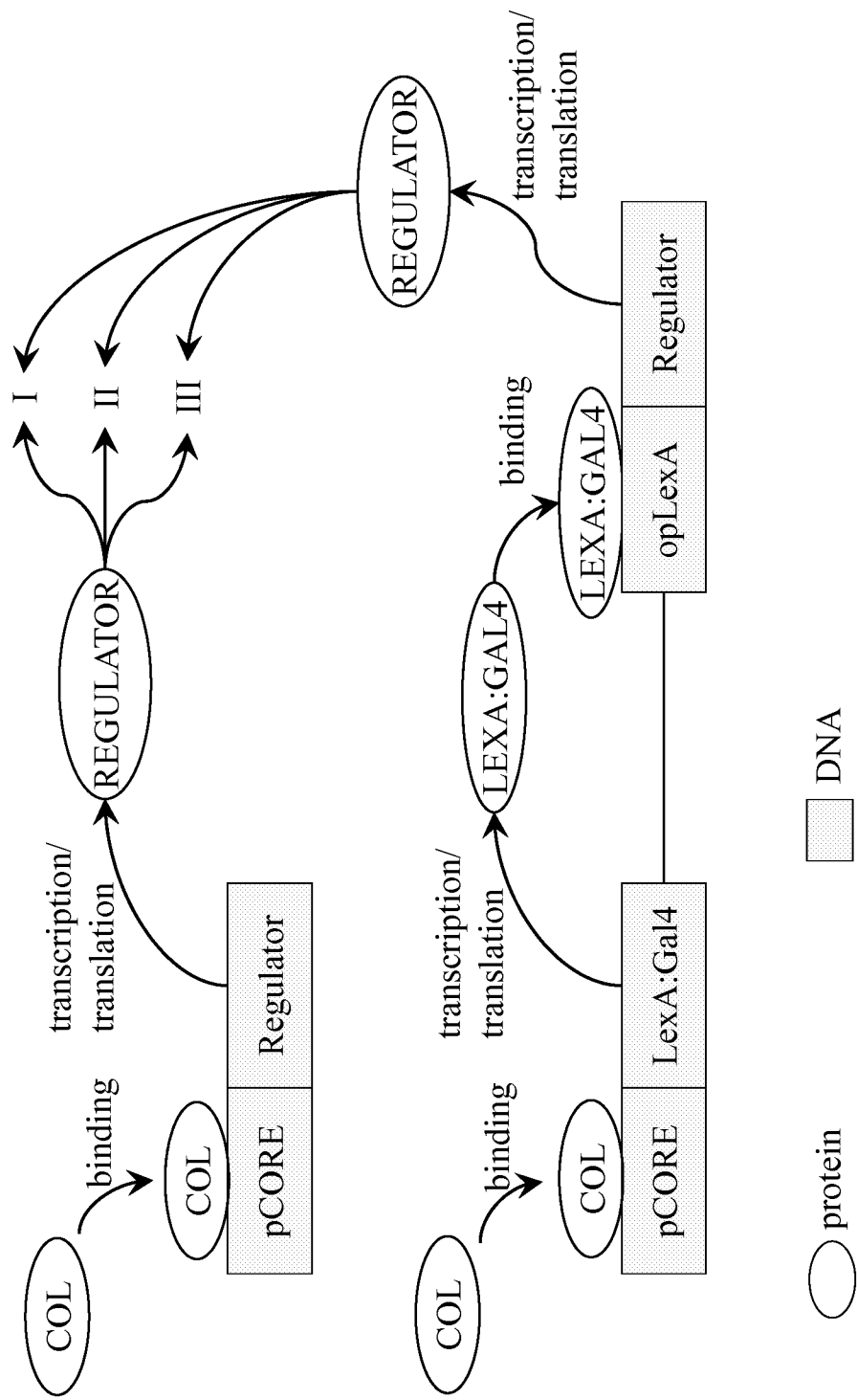


FIG. 4

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 10/37391

A. CLASSIFICATION OF SUBJECT MATTER

IPC(8) - C12N 15/82, A01H 1/00 (2010.01)

USPC - 435/419, 435/320.1, 800/287

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC(8):C12N 15/82, A01H 1/00 (2010.01)

USPC:435/419, 435/320.1, 800/287

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

USPC: 435/468, 800/279, 800/289, 800/290, 800/278

(keyword limited; terms below)

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

PubWest (USPT,PGPB,EPAB,JPAB), DialogPRO--Chemical Engineering and Biotechnology Abstracts, INSPEC, NTIS (National Technical Information Service), PASCAL, Current Contents Search, MEDLINE search terms: promoter, constans like protein, constans bind, plant, vector, polynucleotide

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	BEN-NAIM et al., The CCAAT binding factor can mediate interactions between CONSTANS-like proteins and DNA. Plant J., May 2006, Vol 46, No 3, pg 462-76, pg 463, col 2, para 4-pg 465, para 1; pg 469-470; abstract.	1, 4, 11-20
A	GenBank: AC157779.22, Medicago truncatula chromosome 6 clone mth2-156d20, complete sequence, first submitted 01-MAR-2005 [online]. [Retrieved on 2010-09-21]. Retrieved from the Internet: <URL: http://www.ncbi.nlm.nih.gov/nuccore/AC157779 >	1, 4, 11-20

☐ Further documents are listed in the continuation of Box C.

* Special categories of cited documents:

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

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

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

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

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

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search

22 September 2010 (22.11.2010)

Date of mailing of the international search report

04 OCT 2010

Name and mailing address of the ISA/US

Mail Stop PCT, Attn: ISA/US, Commissioner for Patents
P.O. Box 1450, Alexandria, Virginia 22313-1450

Facsimile No. 571-273-3201

Authorized officer:

Lee W. Young

PCT Helpdesk: 571-272-4300
PCT OSP: 571-272-7774

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 10/37391

Box No. 1 Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing filed or furnished:

a. (means)

☐

on paper

☒

in electronic form

b. (time)

☐

in the international application as filed

☒

together with the international application in electronic form

☐

subsequently to this Authority for the purposes of search

2. ☐ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that in the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 10/37391

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

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

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

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

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

Group I: Claims 1, 4, and 11-20, drawn to methods, plants, expression vectors and isolated polynucleotides comprising a promoter sequence having percentage identity to SEQ ID NO: 101.

Group II-XI: Claims 1, 4, and 11-20, drawn to methods, plants, expression vectors and isolated polynucleotides comprising a promoter sequence having percentage identity with a sequence selected from SEQ ID NOs: 102, 7-10, 13, and 17-20. If Applicant elects to have this group searched, Applicant must specify the specific promoter sequence(s) to be searched. Each unrelated sequence constitutes an inventive concept.

-----please see continuation on extra sheet---

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1, 4, and 11-20, limited to SEQ ID NO: 101

Remark on Protest

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

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 10/37391

Continuation of:

Box No. III Observations where unity of invention is lacking

Group XII+: Claims 1 and 3-20, drawn to methods, plants, expression vectors and isolated polynucleotides comprising a promoter sequence having percentage identity to SEQ ID NO:16 and a polypeptide selected from the group consisting of SEQ ID NOs: 25-69, and an expression vector comprising the promoter operably linked to SEQ ID NO:22 that encodes SEQ ID NO:23, wherein SEQ ID NO:23 regulates the transcription of a polynucleotide operably linked to SEQ ID NO:24. If Applicant elects to have this group searched, Applicant must specify the specific polypeptide sequence(s) (selected from SEQ ID NOs: 25-69) to be searched. Each unrelated sequence constitutes an inventive concept.

Group XIII: Claims 2, 12-14, 16-19, drawn to methods, plants, and recombinant polynucleotides operably linked to a transcribable polynucleotide comprising a sequence selected from one of SEQ ID NOs: 1, 14, 15, or 21. If Applicant elects to have this group searched, Applicant must specify the specific promoter sequence(s) to be searched. Each unrelated sequence constitutes an inventive concept.

Group XIV: Claims 12-14 and 16-19, drawn to methods and plants comprising a promoter sequence which has percentage identity with any of SEQ ID NOs: 2-4. If Applicant elects to have this group searched, Applicant must specify the specific promoter sequence(s) to be searched. Each unrelated sequence constitutes an inventive concept.

The inventions listed as Groups I-XIV+ do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

A technical feature shared by each of the inventions listed as Groups I-XIV+ is a promoter sequence that can be bound by a COL polypeptide and as a result of said binding, regulate transcription. This shared special technical feature does not provide a contribution over the prior art, as evidenced by the article entitled "The CCAAT binding factor can mediate interactions between CONSTANS-like proteins and DNA" by Ben-Naim et al. (Plant J., May 2006, Vol 46, No 3, pp 462-76). Ben-Naim et al. teaches a promoter sequence that can be bound by a COL polypeptide and as a result of said binding, regulate transcription (abstract and p 469-470). In the absence of a contribution over the prior art, the shared technical feature is not a shared special technical feature. Without a shared special technical feature, the inventions lack unity with one another.

Further, another technical feature shared by each of the inventions listed as Groups I-XIV+ is the specific promoter sequences recited therein. Significant structural similarities cannot readily be ascertained among the promoter sequences. Without significant structural similarities, the promoter sequences do not have a shared special technical feature. In the absence of a shared special technical feature, the inventions lack unity with one another.

In this case, the first named sequences that will be searched without additional fees is SEQ ID NO: 101. In order for more sequences to be examined, the appropriate additional examination fees must be paid and the desired sequences to be searched clearly identified.