



(43) International Publication Date
6 March 2014 (06.03.2014)

(10) International Publication Number
WO 2014/036048 A1

- (51) **International Patent Classification:**
C07K 14/415 (2006.01) *C12N 15/82* (2006.01)
- (21) **International Application Number:**
PCT/US2013/056931
- (22) **International Filing Date:**
28 August 2013 (28.08.2013)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**
61/694,811 30 August 2012 (30.08.2012) US
- (71) **Applicant:** E. I. DU PONT DE NEMOURS AND COMPANY [US/US]; 1007 Market Street, Wilmington, Delaware 19898 (US).
- (72) **Inventors:** MCGONIGLE, Brian; 1707 N. Union Street, Wilmington, Delaware 19806-2501 (US). MOLOSHOK, Thomas, D.; 19 Whippain Drive, Ambler, Pennsylvania 19002 (US).
- (74) **Agent:** STOOP, Johan, M.; E. I. du Pont de Nemours and Company, Legal Patent Records Center, 4417 Lancaster Pike, Wilmington, Delaware 19805 (US).
- (81) **Designated States** (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY,

BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, 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, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

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

Published:

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

(54) **Title:** LONG INTERGENIC NON-CODING RNAS IN MAIZE

(57) **Abstract:** The invention provides methods and compositions useful in altering gene expression. The invention provides polynucleotide constructs comprising useful for altering gene expression, as well as cells, plants and seeds comprising the polynucleotides. The invention also provides methods for using lincRNAs to alter gene expression.



WO 2014/036048 A1

TITLE

Long Intergenic Non-Coding RNAs in maize

This application claims the benefit of U.S. Provisional Application
5 No. 61/694811, filed August 30, 2012, the entire content of which is herein
incorporated by reference.

FIELD OF THE INVENTION

The field of the present invention relates generally to plant molecular biology.
More specifically it relates to compositions, constructs and methods of altering gene
10 expression.

BACKGROUND

Recent data obtained using whole-genome tiling arrays and RNA sequencing
have shown that the transcription landscape in eukaryotes is much more complex
than was previously anticipated, with a high proportion of novel transcripts
15 generated from intergenic regions and promoters of annotated genes (Zhu et al.
2012, Genes 3:176-190; Jacquier et al. 2009, Nat. Rev. Genet. 10:833–844).

A large proportion of the eukaryotic genome produces RNA molecules that
have no protein-coding capacity, namely non-coding RNAs (ncRNAs) (Zhu and
Wang, 2012, Genes 3:176-190). ncRNAs are arbitrarily grouped into short (<200 nt)
20 and long ncRNAs (lncRNAs; >200 nt). Long non-coding RNAs that are transcribed
from non-coding DNA sequences between protein-coding genes are referred to as
long intergenic non-coding RNAs (lincRNA). The importance of short ncRNAs,
including siRNAs, miRNAs and piRNAs, in transcriptional and posttranscriptional
regulation of gene expression has been well recognized. In contrast, the regulatory
25 roles of lncRNAs are beginning to be recognized (Zhu et al. 2012, Genes 3:176-190).

Recent studies of lncRNAs in mammalian systems have shown that they are
key regulators of a variety of biological processes. In comparison to animals, plants
have fewer lncRNAs identified and functionally characterized. However, the
regulatory functions of plant lncRNAs appear largely similar to animal lncRNAs (Zhu
30 and Wang, 2012, Genes 3:176-190). LincRNAs have been shown to act as
transcriptional regulators and competing endogenous RNAs (ceRNAs), to serve as
molecular cargos for protein re-localization and as modular scaffolds to recruit the
assembly of multiple protein complexes for chromatin modifications. Some of these

functions have been found to be conserved in plants. (Zhu et al. 2012, Genes 3:176-190).

Although some lncRNAs have been discovered from plants and characterized, there is an ongoing interest in the isolation of more novel lncRNAs given the important role they play in diverse array of biological processes.

SUMMARY OF THE INVENTION

The present invention relates, *inter alia*, to an isolated polynucleotide comprising (a) a nucleotide sequence selected from the group consisting of SEQ ID NOs:1-203,370; or, (b) a full-length complement of (a); or, (c) a nucleotide sequence comprising a sequence having at least 95% sequence identity, based on the Clustal V method of alignment with pairwise alignment default parameters (KTUPLE=2, GAP PENALTY=5, WINDOW=4 and DIAGONALS SAVED=4), when compared to the nucleotide sequence of (a); wherein said nucleotide sequence is a long intergenic non-coding RNA (lincRNA).

In a second embodiment, the invention relates to an isolated polynucleotide of embodiment 1, comprising a nucleotide sequence having at least 98% sequence identity, based on the Clustal V method of alignment with pairwise alignment default parameters (KTUPLE=2, GAP PENALTY=5, WINDOW=4 and DIAGONALS SAVED=4), when compared to the nucleotide sequence of (a).

In a third embodiment, the invention relates to a recombinant DNA construct comprising any of the isolated polynucleotides discussed herein operably linked to at least one regulatory sequence.

In a fourth embodiment, the invention relates to a plant comprising in its genome the DNA expression constructs discussed herein. Such plants can be selected from the group consisting of corn, rice, sorghum, sunflower, millet, soybean, canola, wheat, barley, oat, beans, and nuts.

In a fifth embodiment, the invention relates to transgenic seeds obtained from a plant comprising in its genome the recombinant DNA constructs discussed herein. Also within the scope of the invention are transformed plant tissues or a plant cell comprising in its genome recombinant DNA constructs discussed herein.

In a sixth embodiment, the invention relates to a method of expressing a lincRNA in a plant comprising:

- a) introducing the recombinant DNA construct of the instant invention into the plant, wherein the at least one heterologous nucleotide sequence comprises a lincRNA;
- b) growing the plant of step a); and,
- c) selecting a plant displaying expression of the lincRNA.

BRIEF DESCRIPTION OF THE DRAWINGS AND SEQUENCE LISTINGS

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

A Sequence Listing and Table are provided herewith on Compact Disk. The contents of the Compact Disk containing the Sequence Listing and Table are hereby incorporated by reference in compliance with 37 CFR 1.52(e). The Compact Disks are submitted in duplicate and are identical to one another. The disks are labeled "Copy 1 – Sequence Listing and Table ", "Copy 2 - Sequence Listing and Table", and CRF. The disks contain the following files: BB2041US PSP Sequence Listing having the following size: 174,979,072 bytes and BB2041 US PSP Table 1 having the following size: 10,672,128 bytes which were created August 29, 2012.

The sequence descriptions summarize the Sequence Listing attached hereto. The Sequence Listing contains one letter codes for nucleotide sequence characters and the single and three letter codes for amino acids as defined in the IUPAC-IUB standards described in *Nucleic Acids Research* 13:3021-3030 (1985) and in the *Biochemical Journal* 219(2):345-373 (1984).

SEQ ID NOs: 1-203,370 represent individual long intergenic non-coding RNAs (lincRNAs) from maize (*Zea mays*). SEQ ID NOs: 1-203,370 are shown as DNA sequences and represent a conversion of the RNA nucleotide "U" to the DNA nucleotide "T". The SEQ ID NO:s are described in Table 1 listing the SEQ ID NO in column 1, and the description of the lincRNA in column 2.

Table 1 lists the sequence number of the lincRNA in column one followed by a description in column two. For example, SEQ ID NO:1 represents the DNA sequence of the A1710005-ZmChr8-90483850-90484319 lincRNA. A1710005-ZmChr8-90483850-90484319 refers to genotype A1710005 on the *Zea mays* chromosome 8 at base pair location 90483850-90484319.

DETAILED DESCRIPTION

The disclosure of all patents, patent applications, and publications cited herein are incorporated by reference in their entirety.

As used herein and in the appended claims, the singular forms “a”, “an”, and “the” include plural reference unless the context clearly dictates otherwise. Thus, for example, reference to “a plant” includes a plurality of such plants; reference to “a cell” includes one or more cells and equivalents thereof known to those skilled in the art, and so forth.

In the context of this disclosure, a number of terms shall be utilized.

As used herein “contiguous” sequences or domains refer to sequences that are sequentially linked without added nucleotides intervening between the domains.

As used herein, “domain” or “functional domain” refers to nucleic acid sequence(s) that are capable of eliciting a biological response in plants. The present invention concerns lincRNAs composed of at least 200 nucleotide sequences acting either individually, or in concert with other nucleotide sequences, therefore a domain could refer to either individual lincRNAs or groups of lincRNAs. As used herein, “subdomains” or “functional subdomains” refer to subsequences of domains that are capable of eliciting a biological response in plants

As used herein, “encodes” or “encoding” refers to a DNA sequence which can be processed to generate an RNA and/or polypeptide.

As used herein, “expression” or “expressing” refers to production of a functional product, such as, the generation of an RNA transcript from an introduced construct, an endogenous DNA sequence, or a stably incorporated heterologous DNA sequence. The term may also refer to a polypeptide produced from an mRNA generated from any of the above DNA precursors. Thus, expression of a nucleic acid fragment may refer to transcription of the nucleic acid fragment (e.g., transcription resulting in mRNA or other functional RNA) and/or translation of RNA into a precursor or mature protein (polypeptide).

A “functional fragment” refers to a portion or subsequence of the nucleotide sequence of the present invention in which the ability to function as a lincRNA is maintained. Fragments can be obtained via methods such as site-directed mutagenesis and synthetic construction.

The term "genome" refers to the following: (1) the entire complement of genetic material (genes and non-coding sequences) present in each cell of an organism, or virus or organelle; (2) a complete set of chromosomes inherited as a (haploid) unit from one parent. "Genome" as it applies to plant cells encompasses not only chromosomal DNA found within the nucleus, but organelle DNA found within subcellular components (e.g., mitochondrial, plastid) of the cell.

As used herein, "heterologous" with respect to a sequence means a sequence that originates from a foreign species, or, if from the same species, is substantially modified from its native form in composition and/or genomic locus by deliberate human intervention. For example, with respect to a nucleic acid, it can be a nucleic acid that originates from a foreign species, or is synthetically designed, or, if from the same species, is substantially modified from its native form in composition and/or genomic locus by deliberate human intervention. A heterologous protein may originate from a foreign species or, if from the same species, is substantially modified from its original form by deliberate human intervention.

A "heterologous nucleotide sequence" refers to a sequence that is not naturally occurring with the nucleotide sequence of the invention. While this nucleotide sequence is heterologous to the nucleotide sequence of the invention, it may be homologous, or native, or heterologous, or foreign, to the plant host. The terms "heterologous nucleotide sequence", "heterologous sequence", "heterologous nucleic acid fragment", and "heterologous nucleic acid sequence" are used interchangeably herein.

The term "host cell" refers to a cell which contains or into which is introduced a nucleic acid construct and supports the replication and/or expression of the construct. Host cells may be prokaryotic cells such as *E. coli*, or eukaryotic cells such as fungi, yeast, insect, amphibian, nematode, or mammalian cells. Alternatively, the host cells are monocotyledonous or dicotyledonous plant cells. An example of a monocotyledonous host cell is a maize host cell.

The term "introduced" means providing a nucleic acid (e.g., expression construct) or protein into a cell. Introduced includes reference to the incorporation of a nucleic acid into a eukaryotic or prokaryotic cell where the nucleic acid may be incorporated into the genome of the cell, and includes reference to the transient

provision of a nucleic acid or protein to the cell. Introduced includes reference to stable or transient transformation methods, as well as sexually crossing. Thus, "introduced" in the context of inserting a nucleic acid fragment (e.g., a recombinant DNA construct/expression construct) into a cell, means "transfection" or

5 "transformation" or "transduction" and includes reference to the incorporation of a nucleic acid fragment into a eukaryotic or prokaryotic cell where the nucleic acid fragment may be incorporated into the genome of the cell (e.g., chromosome, plasmid, plastid or mitochondrial DNA), converted into an autonomous replicon, or transiently expressed (e.g., transfected mRNA).

10 The term "isolated" refers to material, such as a nucleic acid or a protein, which is: (1) substantially or essentially free from components which normally accompany or interact with the material as found in its naturally occurring environment or (2) if the material is in its natural environment, the material has been altered by deliberate human intervention to a composition and/or placed at a locus
15 in the cell other than the locus native to the material.

As used herein, microRNA or "miRNA" refers to an oligoribonucleic acid, which regulates expression of a polynucleotide comprising the target sequence. A "mature miRNA" refers to the miRNA generated from the processing of a miRNA precursor. A "miRNA template" is an oligonucleotide region, or regions, in a nucleic
20 acid construct which encodes the miRNA. The "backside" region of a miRNA is a portion of a polynucleotide construct which is substantially complementary to the miRNA template and is predicted to base pair with the miRNA template. The miRNA template and backside may form a double-stranded polynucleotide, including a hairpin structure.

25 By "nucleic acid library" is meant a collection of isolated DNA or RNA molecules which comprise and substantially represent the entire transcribed fraction of a genome of a specified organism or of a tissue from that organism. Construction of exemplary nucleic acid libraries, such as genomic and cDNA libraries, is taught in standard molecular biology references such as Berger and Kimmel, *Guide to*
30 *Molecular Cloning Techniques, Methods in Enzymology*, Vol. 152, Academic Press, Inc., San Diego, CA (Berger); Sambrook et al., *Molecular Cloning - A Laboratory Manual*, 2nd ed., Vol. 1-3 (1989); and *Current Protocols in Molecular Biology*, F.M.

Ausubel et al., Eds., Current Protocols, a joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc. (1994).

As used herein "operably linked" includes reference to a functional linkage of at least two sequences. Operably linked includes linkage between a promoter and a second sequence, wherein the promoter sequence initiates and mediates transcription of the DNA sequence corresponding to the second sequence.

"PCR" or "Polymerase Chain Reaction" is a technique for the synthesis of large quantities of specific DNA segments, consisting of a series of repetitive cycles (Perkin Elmer Cetus Instruments, Norwalk, CT). Typically, the double stranded DNA is heat denatured, the two primers complementary to the 3' boundaries of the target segment are annealed at low temperature and then extended at an intermediate temperature. One set of these three consecutive steps comprises a cycle.

The terms "plasmid", "vector" and "cassette" refer a small nucleic acid molecule (plasmid, virus, bacteriophage, artificial or cut DNA molecule) that can be used to deliver a polynucleotide of the invention into a host cell. Vectors are capable of being replicated and contain cloning sites for introduction of a foreign polynucleotide. Thus, expression vectors permit transcription of a nucleic acid inserted therein.

One aspect of the invention concerns a plant, cell, and seed comprising a recombinant DNA construct expressing a lincRNA and/or a lincRNA of the current disclosure. Typically, the cell will be a cell from a plant, but other prokaryotic or eukaryotic cells are also contemplated, including but not limited to viral, bacterial, fungus, yeast, insect, nematode, or animal cells. Plant cells include cells from monocots and dicots. The invention also provides plants and seeds comprising the recombinant DNA construct and/or a lincRNA of the current disclosure.

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

The terms "monocot" and "monocotyledonous plant" are used interchangeably herein. A monocot of the current invention includes the Gramineae.

The terms “dicot” and “dicotyledonous plant” are used interchangeably herein. A dicot of the current invention includes the following families: Asteraceae, Brassicaceae, Leguminosae, and Solanaceae.

The term “plant parts” includes differentiated and undifferentiated tissues including, but not limited to the following: roots, stems, shoots, leaves, pollen, seeds, tumor tissue and various forms of cells and culture (e.g., single cells, protoplasts, embryos and callus tissue). The plant tissue may be in plant or in a plant organ, plant organelle, tissue culture or cell culture.

The term “plant organ” refers to plant tissue or group of tissues that constitute a morphologically and functionally distinct part of a plant.

The terms “polynucleotide”, “polynucleotide sequence”, “nucleic acid sequence”, “nucleic acid fragment”, and “isolated nucleic acid fragment” are used interchangeably herein. These terms encompass nucleotide sequences and the like. A polynucleotide may be a polymer of RNA or DNA that is single- or double-stranded, that optionally contains synthetic, non-natural or altered nucleotide bases. A polynucleotide in the form of a polymer of DNA may be comprised of one or more segments of cDNA, genomic DNA, synthetic DNA, or mixtures thereof. Nucleotides (usually found in their 5'-monophosphate form) are referred to by a single letter designation as follows: “A” for adenylate or deoxyadenylate (for RNA or DNA, respectively), “C” for cytidylate or deoxycytidylate, “G” for guanylate or deoxyguanylate, “U” for uridylate, “T” for deoxythymidylate, “R” for purines (A or G), “Y” for pyrimidines (C or T), “K” for G or T, “H” for A or C or T, “I” for inosine, and “N” for any nucleotide.

An “isolated polynucleotide” refers to a polymer of ribonucleotides (RNA) or deoxyribonucleotides (DNA) that is single- or double-stranded, optionally containing synthetic, non-natural or altered nucleotide bases. An isolated polynucleotide in the form of DNA may be comprised of one or more segments of cDNA, genomic DNA or synthetic DNA.

“Progeny” comprises any subsequent generation of a plant.

As used herein, “polypeptide” means proteins, protein fragments, modified proteins, amino acid sequences and synthetic amino acid sequences. The polypeptide can be glycosylated or not.

As used herein, "promoter" refers to a nucleic acid fragment, e.g., a region of DNA, that is involved in recognition and binding of an RNA polymerase and other proteins to initiate transcription. In other words, this nucleic acid fragment is capable of controlling transcription of another nucleic acid fragment.

5 A number of promoters can be used, these promoters can be selected based on the desired outcome. It is recognized that different applications will be enhanced by the use of different promoters in plant expression cassettes to modulate the timing, location and/or level of expression of the lincRNA. Such plant expression cassettes may also contain, if desired, a promoter regulatory region (e.g., one
10 conferring inducible, constitutive, environmentally- or developmentally-regulated, or cell- or tissue-specific/selective expression), a transcription initiation start site, a ribosome binding site, an RNA processing signal, a transcription termination site, and/or a polyadenylation signal.

Constitutive, tissue-preferred or inducible promoters can be employed.

15 Examples of constitutive promoters include the cauliflower mosaic virus (CaMV) 35S transcription initiation region, the 1'- or 2'- promoter derived from T-DNA of *Agrobacterium tumefaciens*, the ubiquitin 1 promoter, the Smas promoter, the cinnamyl alcohol dehydrogenase promoter (U.S. Patent No. 5,683,439), the Nos promoter, the pEmu promoter, the rubisco promoter, the GRP1-8 promoter and
20 other transcription initiation regions from various plant genes known to those of skill. If low level expression is desired, weak promoter(s) may be used. Weak constitutive promoters include, for example, the core promoter of the Rsyn7 promoter (WO 99/43838 and U.S. Patent No. 6,072,050), the core 35S CaMV promoter, and the like. Other constitutive promoters include, for example, U.S.
25 Patent Nos. 5,608,149; 5,608,144; 5,604,121; 5,569,597; 5,466,785; 5,399,680; 5,268,463; and 5,608,142. See also, U.S. Patent No. 6,177,611, herein incorporated by reference.

Examples of inducible promoters are the Adh1 promoter which is inducible by hypoxia or cold stress, the Hsp70 promoter which is inducible by heat stress, the
30 PPKK promoter and the pepcarboxylase promoter which are both inducible by light. Also useful are promoters which are chemically inducible, such as the In2-2 promoter which is safener induced (U.S. patent 5,364,780), the ERE promoter which

is estrogen induced, and the Axig1 promoter which is auxin induced and tapetum specific but also active in callus (PCT US01/22169).

Examples of promoters under developmental control include promoters that initiate transcription preferentially in certain tissues, such as leaves, roots, fruit,

5 seeds, or flowers. An exemplary promoter is the anther specific promoter 5126 (U.S. Patent Nos. 5,689,049 and 5,689,051). Examples of seed-preferred promoters include, but are not limited to, 27 kD gamma zein promoter and waxy promoter, Boronat, A. et al. (1986) *Plant Sci.* 47:95-102; Reina, M. et al. *Nucl. Acids Res.* 18(21):6426; and Kloesgen, R.B. et al. (1986) *Mol. Gen. Genet.* 203:237-244.

10 Promoters that express in the embryo, pericarp, and endosperm are disclosed in US patent 6,225,529 and PCT publication WO 00/12733. The disclosures each of these are incorporated herein by reference in their entirety.

In some embodiments it will be beneficial to express a gene of interest and/or a lincRNA using an inducible promoter, particularly from a pathogen-inducible

15 promoter. Such promoters include those from pathogenesis-related proteins (PR proteins), which are induced following infection by a pathogen; e.g., PR proteins, SAR proteins, beta-1,3-glucanase, chitinase, etc. See, for example, Redolfi et al. (1983) *Neth. J. Plant Pathol.* 89:245-254; Uknes et al. (1992) *Plant Cell* 4:645-656; and Van Loon (1985) *Plant Mol. Virol.* 4:111-116. See also WO 99/43819, herein

20 incorporated by reference.

Of interest are promoters that are expressed locally at or near the site of pathogen infection. See, for example, Marineau et al. (1987) *Plant Mol. Biol.* 9:335-342; Matton et al. (1989) *Molecular Plant-Microbe Interactions* 2:325-331; Somsisch et al. (1986) *Proc. Natl. Acad. Sci. USA* 83:2427-2430; Somsisch et al. (1988) *Mol.*

25 *Gen. Genet.* 2:93-98; and Yang (1996) *Proc. Natl. Acad. Sci. USA* 93:14972-14977. See also, Chen et al. (1996) *Plant J.* 10:955-966; Zhang et al. (1994) *Proc. Natl. Acad. Sci. USA* 91:2507-2511; Warner et al. (1993) *Plant J.* 3:191-201; Siebertz et al. (1989) *Plant Cell* 1:961-968; U.S. Patent No. 5,750,386 (nematode-inducible);

and the references cited therein. Of particular interest is the inducible promoter for the maize PRms gene, whose expression is induced by the pathogen *Fusarium moniliforme* (see, for example, Cordero et al. (1992) *Physiol. Mol. Plant Path.* 41:189-200).

Additionally, as pathogens find entry into plants through wounds or insect damage, a wound-inducible promoter may be used in the constructions of the polynucleotides. Such wound-inducible promoters include potato proteinase inhibitor (pin II) gene (Ryan (1990) *Ann. Rev. Phytopath.* 28:425-449; Duan et al. (1996) *Nature Biotech.* 14:494-498); *wun1* and *wun2*, U.S. Patent No. 5,428,148; *win1* and *win2* (Stanford et al. (1989) *Mol. Gen. Genet.* 215:200-208); systemin (McGurl et al. (1992) *Science* 225:1570-1573); WIP1 (Rohmeier et al. (1993) *Plant Mol. Biol.* 22:783-792; Eckelkamp et al. (1993) *FEBS Lett.* 323:73-76); MPI gene (Corderok et al. (1994) *Plant J.* 6(2):141-150); and the like, herein incorporated by reference.

Chemical-regulated promoters can be used to modulate the expression of a gene and /or a lincRNA in a plant through the application of an exogenous chemical regulator. Depending upon the objective, the promoter may be a chemical-inducible promoter, where application of the chemical induces gene expression, or a chemical-repressible promoter, where application of the chemical represses gene expression. Chemical-inducible promoters are known in the art and include, but are not limited to, the maize *ln2-2* promoter, which is activated by benzenesulfonamide herbicide safeners, the maize GST promoter, which is activated by hydrophobic electrophilic compounds that are used as pre-emergent herbicides, and the tobacco PR-1a promoter, which is activated by salicylic acid. Other chemical-regulated promoters of interest include steroid-responsive promoters (see, for example, the glucocorticoid-inducible promoter in Schena et al. (1991) *Proc. Natl. Acad. Sci. USA* 88:10421-10425 and McNellis et al. (1998) *Plant J.* 14(2):247-257) and tetracycline-inducible and tetracycline-repressible promoters (see, for example, Gatz et al. (1991) *Mol. Gen. Genet.* 227:229-237, and U.S. Patent Nos. 5,814,618 and 5,789,156), herein incorporated by reference.

Tissue-preferred promoters can be utilized to target enhanced expression of a sequence of interest within a particular plant tissue. Tissue-preferred promoters include Yamamoto et al. (1997) *Plant J.* 12(2):255-265; Kawamata et al. (1997) *Plant Cell Physiol.* 38(7):792-803; Hansen et al. (1997) *Mol. Gen. Genet.* 254(3):337-343; Russell et al. (1997) *Transgenic Res.* 6(2):157-168; Rinehart et al. (1996) *Plant Physiol.* 112(3):1331-1341; Van Camp et al. (1996) *Plant Physiol.* 112(2):525-535; Canevascini et al. (1996) *Plant Physiol.* 112(2):513-524; Yamamoto et al. (1994)

Plant Cell Physiol. 35(5):773-778; Lam (1994) *Results Probl. Cell Differ.* 20:181-196; Orozco et al. (1993) *Plant Mol Biol.* 23(6):1129-1138; Matsuoka et al. (1993) *Proc Natl. Acad. Sci. USA* 90(20):9586-9590; and Guevara-Garcia et al. (1993) *Plant J.* 4(3):495-505. Such promoters can be modified, if necessary, for weak expression.

Leaf-preferred promoters are known in the art. See, for example, Yamamoto et al. (1997) *Plant J.* 12(2):255-265; Kwon et al. (1994) *Plant Physiol.* 105:357-67; Yamamoto et al. (1994) *Plant Cell Physiol.* 35(5):773-778; Gotor et al. (1993) *Plant J.* 3:509-18; Orozco et al. (1993) *Plant Mol. Biol.* 23(6):1129-1138; and Matsuoka et al. (1993) *Proc. Natl. Acad. Sci. USA* 90(20):9586-9590. In addition, the promoters of cab and rubisco can also be used. See, for example, Simpson et al. (1958) *EMBO J* 4:2723-2729 and Timko et al. (1988) *Nature* 318:57-58.

Root-preferred promoters are known and can be selected from the many available from the literature or isolated de novo from various compatible species. See, for example, Hire et al. (1992) *Plant Mol. Biol.* 20(2):207-218 (soybean root-specific glutamine synthetase gene); Keller and Baumgartner (1991) *Plant Cell* 3(10):1051-1061 (root-specific control element in the GRP 1.8 gene of French bean); Sanger et al. (1990) *Plant Mol. Biol.* 14(3):433-443 (root-specific promoter of the mannopine synthase (MAS) gene of *Agrobacterium tumefaciens*); and Miao et al. (1991) *Plant Cell* 3(1):11-22 (full-length cDNA clone encoding cytosolic glutamine synthetase (GS), which is expressed in roots and root nodules of soybean). See also Bogusz et al. (1990) *Plant Cell* 2(7):633-641, where two root-specific promoters isolated from hemoglobin genes from the nitrogen-fixing nonlegume *Parasponia andersonii* and the related non-nitrogen-fixing nonlegume *Trema tomentosa* are described. The promoters of these genes were linked to a β -glucuronidase reporter gene and introduced into both the nonlegume *Nicotiana tabacum* and the legume *Lotus corniculatus*, and in both instances root-specific promoter activity was preserved. Leach and Aoyagi (1991) describe their analysis of the promoters of the highly expressed rolC and rolD root-inducing genes of *Agrobacterium rhizogenes* (see *Plant Science* (Limerick) 79(1):69-76). They concluded that enhancer and tissue-preferred DNA determinants are dissociated in those promoters. Teeri et al. (1989) used gene fusion to lacZ to show that the *Agrobacterium* T-DNA gene encoding octopine synthase is especially active in the epidermis of the root tip and

that the TR2' gene is root specific in the intact plant and stimulated by wounding in leaf tissue, an especially desirable combination of characteristics for use with an insecticidal or larvicidal gene (see *EMBO J.* 8(2):343-350). The TR1' gene, fused to nptII (neomycin phosphotransferase II) showed similar characteristics. Additional
5 root-preferred promoters include the VfENOD-GRP3 gene promoter (Kuster et al. (1995) *Plant Mol. Biol.* 29(4):759-772); and rolB promoter (Capana et al. (1994) *Plant Mol. Biol.* 25(4):681-691. See also U.S. Patent Nos. 5,837,876; 5,750,386; 5,633,363; 5,459,252; 5,401,836; 5,110,732; and 5,023,179. The phaseolin gene (Murai et al. (1983) *Science* 23:476-482 and Sengopta-Gopalen et al. (1988) *PNAS*
10 82:3320-3324.

The term "recombinant DNA construct" or "recombinant expression construct" is used interchangeably and refers to a discrete polynucleotide into which a nucleic acid sequence or fragment can be moved. Preferably, it is a plasmid vector or a fragment thereof comprising the promoters of the present invention. The
15 choice of plasmid vector is dependent upon the method that will be used to transform host plants. The skilled artisan is well aware of the genetic elements that must be present on the plasmid vector in order to successfully transform, select and propagate host cells containing the chimeric gene. The skilled artisan will also recognize that different independent transformation events will result in different
20 levels and patterns of expression (Jones et al., *EMBO J.* 4:2411-2418 (1985); De Almeida et al., *Mol. Gen. Genetics* 218:78-86 (1989)), and thus that multiple events must be screened in order to obtain lines displaying the desired expression level and pattern. Such screening may be accomplished by PCR and Southern analysis of DNA, RT-PCR and Northern analysis of mRNA expression, Western
25 analysis of protein expression, or phenotypic analysis. This recombinant DNA construct may comprise any combination of deoxyribonucleotides, ribonucleotides, and/or modified nucleotides. The construct may be transcribed to form an RNA, wherein the RNA may be capable of forming a double-stranded RNA and/or hairpin structure. This construct may be expressed in the cell, or isolated or synthetically
30 produced. The construct may further comprise a promoter, or other sequences which facilitate manipulation or expression of the construct.

RNA interference (RNAi) refers to the process of sequence-specific post-transcriptional gene silencing in animals mediated by short interfering RNAs

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

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

transcripts or alternately by interaction with particular gene sequences, wherein such interaction results in gene silencing either at the transcriptional or post-transcriptional level.

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

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

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

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

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

15 In animals, there is direct evidence indicating a role for specific miRNAs in development. The lin-4 and let-7 miRNAs in *C. elegans* have been found to control temporal development, based on the phenotypes generated when the genes producing the lin-4 and let-7 miRNAs are mutated (Lee et al., *Cell* 75:843-854 1993; Reinhart et al., *Nature* 403:901-906 2000). In addition, both miRNAs display a
20 temporal expression pattern consistent with their roles in developmental timing. Other animal miRNAs display developmentally regulated patterns of expression, both temporal and tissue-specific (Lagos-Quintana et al., *Science* 294:853-853 2001, Lagos-Quintana et al., *Curr. Biol.* 12:735-739 2002; Lau et al., *Science* 294:858-862 2001; Lee and Ambros, *Science* 294:862-864 2001), leading to the
25 hypothesis that miRNAs may, in many cases, be involved in the regulation of important developmental processes. Likewise, in plants, the differential expression patterns of many miRNAs suggests a role in development (Llave et al., *Plant Cell* 14:1605-1619 2002; Park et al., *Curr. Biol.* 12:1484-1495 2002; Reinhart et al., *Genes. Dev.* 16:1616-1626 2002). However, a developmental role for miRNAs has
30 not been directly proven in plants, because to date there has been no report of a developmental phenotype associated with a specific plant miRNA.

MicroRNAs appear to regulate target genes by binding to complementary sequences located in the transcripts produced by these genes. In the case of lin-4

and let-7, the target sites are located in the 3' UTRs of the target mRNAs (Lee et al., *Cell* 75:843-854 1993; Wightman et al., *Cell* 75:855-862 1993; Reinhart et al., *Nature* 403:901-906 2000; Slack et al., *Mol. Cell* 5:659-669 2000), and there are several mismatches between the lin-4 and let-7 miRNAs and their target sites.

5 Binding of the lin-4 or let-7 miRNA appears to cause down-regulation of steady-state levels of the protein encoded by the target mRNA without affecting the transcript itself (Olsen and Ambros, *Dev. Biol.* 216:671-680 1999). On the other hand, recent evidence suggests that miRNAs can, in some cases, cause specific RNA cleavage of the target transcript within the target site, and this cleavage step appears to
10 require 100% complementarity between the miRNA and the target transcript (Hutvagner and Zamore, *Science* 297:2056-2060 2002; Llave et al., *Plant Cell* 14:1605-1619 2002). It seems likely that miRNAs can enter at least two pathways of target gene regulation: Protein down-regulation when target complementarity is <100%, and RNA cleavage when target complementarity is 100%. MicroRNAs
15 entering the RNA cleavage pathway are analogous to the 21-25 nt short interfering RNAs (siRNAs) generated during RNA interference (RNAi) in animals and posttranscriptional gene silencing (PTGS) in plants (Hamilton and Baulcombe 1999; Hammond et al., 2000; Zamore et al., 2000; Elbashir et al., 2001), and likely are incorporated into an RNA-induced silencing complex (RISC) that is similar or
20 identical to that seen for RNAi.

The term "long non-coding RNAs" refers to non-protein coding transcripts longer than 200 nucleotides. The 200 nucleotide limit distinguishes long ncRNAs from small regulatory RNAs such as microRNAs (miRNAs), short interfering RNAs (siRNAs), Piwi-interacting RNAs (piRNAs) and small nucleolar RNAs (snoRNAs).

25 "long non-coding RNAs", "long ncRNAs" and "lncRNA" are used interchangeably herein.

As used herein "long intergenic non-coding RNAs" refer to long non-coding RNAs that are transcribed from non-coding DNA sequences between protein-coding genes. "long intergenic non-coding RNAs" and "lincRNA" are used interchangeably
30 herein.

LnRNAs and LincRNAs have been shown to be involved in a multitude of biological processes such as, but not limited to, modulation of gene expression, modulation of the function of transcription factors, regulating basal transcription

machinery, slicing of mRNA, translation repression, Epigenetic modifications, including histone and DNA methylation, histone acetylation and sumoylation, and controlling various aspects of post-transcriptional mRNA processing. Alteration of lincRNA expression will thus allow modulation of these activities and the associated phenotypes.

Long ncRNA can also act cooperatively as ligands to modulate the activities of RNA binding proteins. Furthermore, it has been shown that lncRNAs can form extended intramolecular hairpins that may be processed into siRNAs (Czech et al. 2008, Nature 453 (7196): 798–802).

Some lncRNAs have been functionally annotated in lncRNAdb (a database of literature described lncRNAs) (Amaral, P. P. et al Nucleic Acids Research 39 (Database issue): D146–D151. doi:10.1093/nar/gkq1138. PMC 3013714. PMID 21112873.). Long ncRNAs can play a role in gene-specific transcription where they can target transcriptional activators or repressors, different components of the transcription reaction including RNA polymerase (RNAP) II and even the DNA duplex to regulate gene transcription and expression (Goodrich et al. 2006, Nature Reviews Molecular Cell Biology 7 (8): 612–6.). In combination these ncRNAs may comprise a regulatory network that, including transcription factors, finely control gene expression in complex eukaryotes. Alternation of lincRNA expression will allow modulation of these gene expression patterns.

Several molecular functions have been attributed to lincRNAs. LincRNA may function as a natural miRNA target mimic, guiding recruitment of chromatin modifiers, and as molecular cargo for protein re-localization (Zhu et al. Genes 2012 3:176-190). Some lincRNAs attach to messenger RNA to block protein production. Many different lincRNAs are needed to prevent an embryonic stem cell from differentiating.

Standard recombinant DNA and molecular cloning techniques used herein are well known in the art and are described more fully in Sambrook, J. et al., In Molecular Cloning: A Laboratory Manual; 2nd ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, New York, 1989 (hereinafter “Sambrook et al., 1989”) or Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. and Struhl, K., Eds.; In Current Protocols in Molecular Biology; John Wiley and Sons: New York, 1990 (hereinafter “Ausubel et al., 1990”).

The term “suppression” or “silencing” or “inhibition” are used interchangeably to denote the down-regulation of the expression of a product of a target sequence relative to its normal expression level in a wild type organism. Suppression includes expression that is decreased by about 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%,
5 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or 100% relative to the wild type expression level.

As used herein, the terms “sequence of interest” or “target sequence” are used interchangeably. Target sequence is used to mean the nucleic acid sequence that is selected for alteration (e.g., suppression) of expression, and is not limited to
10 polynucleotides encoding polypeptides. The target sequence includes, but is not limited to, RNA, DNA, or a polynucleotide comprising the target sequence.

General categories of sequences of interest include, for example, those genes involved in regulation or information, such as zinc fingers, transcription factors, homeotic genes, or cell cycle and cell death modulators, those involved in
15 communication, such as kinases, and those involved in housekeeping, such as heat shock proteins.

Target sequences further include coding regions and non-coding regions such as promoters, enhancers, terminators, introns and the like, which may be modified in order to alter the expression of a gene of interest. For example, an
20 intron sequence can be added to the 5' region to increase the amount of mature message that accumulates (see for example Buchman and Berg, *Mol. Cell Biol.* 8:4395-4405 (1988); and Callis et al., *Genes Dev.* 1:1183-1200 (1987)).

The target sequence may be an endogenous sequence, or may be an introduced heterologous sequence, or transgene. For example, the methods may
25 be used to alter the regulation or expression of a transgene, or to remove a transgene or other introduced sequence such as an introduced site-specific recombination site. The target sequence may also be a sequence from a pathogen, for example, the target sequence may be from a plant pathogen such as a virus, a mold or fungus, an insect, or a nematode. A lincRNA could be expressed in a plant
30 which, upon infection or infestation, would target the pathogen and confer some degree of resistance to the plant.

In plants, other categories of target sequences include genes affecting agronomic traits, insect resistance, disease resistance, herbicide resistance,

sterility, grain characteristics, and commercial products. Genes of interest also included those involved in oil, starch, carbohydrate, or nutrient metabolism as well as those affecting, for example, kernel size, sucrose loading, and the like. The quality of grain is reflected in traits such as levels and types of oils, saturated and unsaturated, quality and quantity of essential amino acids, and levels of cellulose. For example, genes of the phytic acid biosynthetic pathway could be suppressed to generate a high available phosphorous phenotype. See, for example, phytic acid biosynthetic enzymes including inositol polyphosphate kinase-2 polynucleotides, disclosed in WO 02/059324, inositol 1,3,4-trisphosphate 5/6-kinase polynucleotides, disclosed in WO 03/027243, and myo-inositol 1-phosphate synthase and other phytate biosynthetic polynucleotides, disclosed in WO 99/05298, all of which are herein incorporated by reference. Genes in the lignification pathway could be suppressed to enhance digestibility or energy availability. Genes affecting cell cycle or cell death could be suppressed to affect growth or stress response. Genes affecting DNA repair and/or recombination could be suppressed to increase genetic variability. Genes affecting flowering time could be suppressed, as well as genes affecting fertility. Any target sequence could be suppressed in order to evaluate or confirm its role in a particular trait or phenotype, or to dissect a molecular, regulatory, biochemical, or proteomic pathway or network.

Other commercially desirable traits are genes and proteins conferring cold, heat, salt, and drought resistance.

Disease and /or insect resistance genes may encode resistance to pests that have great yield drag such as for example, anthracnose, soybean mosaic virus, soybean cyst nematode, root-knot nematode, brown leaf spot, Downy mildew, purple seed stain, seed decay and seedling diseases caused commonly by the fungi - *Pythium* sp., *Phytophthora* sp., *Rhizoctonia* sp., *Diaporthe* sp.. Bacterial blight caused by the bacterium *Pseudomonas syringae* pv. *Glycinea*. Genes conferring insect resistance include, for example, *Bacillus thuringiensis* toxic protein genes (U.S. Patent Nos. 5,366,892; 5,747,450; 5,737,514; 5,723,756; 5,593,881; and Geiser et al (1986) Gene 48:109); lectins (Van Damme et al. (1994) Plant Mol. Biol. 24:825); and the like.

Herbicide resistance traits may include genes coding for resistance to herbicides that act to inhibit the action of acetolactate synthase (ALS), in particular

the sulfonylurea-type herbicides (e.g., the acetolactate synthase ALS gene containing mutations leading to such resistance, in particular the S4 and/or HRA mutations). The ALS-gene mutants encode resistance to the herbicide chlorsulfuron. Glyphosate acetyl transferase (GAT) is an N-acetyltransferase from
5 Bacillus licheniformis that was optimized by gene shuffling for acetylation of the broad spectrum herbicide, glyphosate, forming the basis of a novel mechanism of glyphosate tolerance in transgenic plants (Castle et al. (2004) Science 304, 1151-1154).

Antibiotic resistance genes include, for example, neomycin
10 phosphotransferase (npt) and hygromycin phosphotransferase (hpt). Two neomycin phosphotransferase genes are used in selection of transformed organisms: the neomycin phosphotransferase I (nptI) gene and the neomycin phosphotransferase II (nptII) gene. The second one is more widely used. It was initially isolated from the transposon Tn5 that was present in the bacterium strain Escherichia coli K12. The
15 gene codes for the aminoglycoside 3'-phosphotransferase (denoted aph(3')-II or NPTII) enzyme, which inactivates by phosphorylation a range of aminoglycoside antibiotics such as kanamycin, neomycin, geneticin and paromomycin. NPTII is widely used as a selectable marker for plant transformation. It is also used in gene expression and regulation studies in different organisms in part because N-terminal
20 fusions can be constructed that retain enzyme activity. NPTII protein activity can be detected by enzymatic assay. In other detection methods, the modified substrates, the phosphorylated antibiotics, are detected by thin-layer chromatography, dot-blot analysis or polyacrylamide gel electrophoresis. Plants such as maize, cotton, tobacco, Arabidopsis, flax, soybean and many others have been successfully
25 transformed with the nptII gene.

The hygromycin phosphotransferase (denoted hpt, hph or aphIV) gene was originally derived from Escherichia coli. The gene codes for hygromycin phosphotransferase (HPT), which detoxifies the aminocyclitol antibiotic hygromycin B. A large number of plants have been transformed with the hpt gene and
30 hygromycin B has proved very effective in the selection of a wide range of plants, including monocotyledonous. Most plants exhibit higher sensitivity to hygromycin B than to kanamycin, for instance cereals. Likewise, the hpt gene is used widely in selection of transformed mammalian cells. The sequence of the hpt gene has been

modified for its use in plant transformation. Deletions and substitutions of amino acid residues close to the carboxy (C)-terminus of the enzyme have increased the level of resistance in certain plants, such as tobacco. At the same time, the hydrophilic C-terminus of the enzyme has been maintained and may be essential for the strong activity of HPT. HPT activity can be checked using an enzymatic assay. A non-destructive callus induction test can be used to verify hygromycin resistance.

Genes involved in plant growth and development have been identified in plants. One such gene, which is involved in cytokinin biosynthesis, is isopentenyl transferase (IPT). Cytokinin plays a critical role in plant growth and development by stimulating cell division and cell differentiation (Sun et al. (2003), *Plant Physiol.* 131: 167-176).

Calcium-dependent protein kinases (CDPK), a family of serine-threonine kinase found primarily in the plant kingdom, are likely to function as sensor molecules in calcium-mediated signaling pathways. Calcium ions are important second messengers during plant growth and development (Harper et al. *Science* 252, 951-954 (1993); Roberts et al. *Curr. Opin. Cell Biol.* 5, 242-246 (1993); Roberts et al. *Annu. Rev. Plant Mol. Biol.* 43, 375-414 (1992)).

Nematode responsive protein (NRP) is produced by soybean upon the infection of soybean cyst nematode. NRP has homology to a taste-modifying glycoprotein miraculin and the NF34 protein involved in tumor formation and hyper response induction. NRP is believed to function as a defense-inducer in response to nematode infection (Tenhaken et al. *BMC Bioinformatics* 6:169 (2005)).

The quality of seeds and grains is reflected in traits such as levels and types of fatty acids or oils, saturated and unsaturated, quality and quantity of essential amino acids, and levels of carbohydrates. Therefore, commercial traits can also be encoded on a gene or genes that could increase for example methionine and cysteine, two sulfur containing amino acids that are present in low amounts in soybeans. Cystathionine gamma synthase (CGS) and serine acetyl transferase (SAT) are proteins involved in the synthesis of methionine and cysteine, respectively.

Other commercial traits can encode genes to increase for example monounsaturated fatty acids, such as oleic acid, in oil seeds. Soybean oil for example contains high levels of polyunsaturated fatty acids and is more prone to

oxidation than oils with higher levels of monounsaturated and saturated fatty acids. High oleic soybean seeds can be prepared by recombinant manipulation of the activity of oleoyl 12 desaturase (Fad2). High oleic soybean oil can be used in applications that require a high degree of oxidative stability, such as cooking for a long period of time at an elevated temperature.

Raffinose saccharides accumulate in significant quantities in the edible portion of many economically significant crop species, such as soybean (*Glycine max* L. Merrill), sugar beet (*Beta vulgaris*), cotton (*Gossypium hirsutum* L.), canola (*Brassica* sp.) and all of the major edible leguminous crops including beans (*Phaseolus* sp.), chick pea (*Cicer arietinum*), cowpea (*Vigna unguiculata*), mung bean (*Vigna radiata*), peas (*Pisum sativum*), lentil (*Lens culinaris*) and lupine (*Lupinus* sp.). Although abundant in many species, raffinose saccharides are an obstacle to the efficient utilization of some economically important crop species.

Down regulation of the expression of the enzymes involved in raffinose saccharide synthesis, such as galactinol synthase for example, would be a desirable trait.

In certain embodiments, the present invention contemplates the transformation of a recipient cell with more than one advantageous transgene. Two or more transgenes can be supplied in a single transformation event using either distinct transgene-encoding vectors, or a single vector incorporating two or more gene coding sequences. Any two or more transgenes of any description, such as those conferring herbicide, insect, disease (viral, bacterial, fungal, and nematode), or drought resistance, oil quantity and quality, or those increasing yield or nutritional quality may be employed as desired.

The term "selectively hybridizes" includes reference to hybridization, under stringent hybridization conditions, of a nucleic acid sequence to a specified nucleic acid target sequence to a detectably greater degree (e.g., at least 2-fold over background) than its hybridization to non-target nucleic acid sequences and to the substantial exclusion of non-target nucleic acids. Selectively hybridizing sequences typically have about at least 80% sequence identity, or 90% sequence identity, up to and including 100% sequence identity (i.e., fully complementary) with each other.

The skilled artisan recognizes that substantially similar nucleic acid sequences encompassed by this invention are also defined by their ability to

hybridize, under moderately stringent conditions (for example, 0.5 X SSC, 0.1% SDS, 60 °C) with the sequences exemplified herein, or to any portion of the nucleotide sequences reported herein and which are functionally equivalent to the lincRNA of the invention. Estimates of such homology are provided by either DNA
5 DNA or DNA RNA hybridization under conditions of stringency as is well understood by those skilled in the art (Hames and Higgins, Eds.; In Nucleic Acid Hybridisation; IRL Press: Oxford, U.K., 1985). Stringency conditions can be adjusted to screen for moderately similar fragments, such as homologous sequences from distantly related organisms, to highly similar fragments, such as genes that duplicate functional
10 enzymes from closely related organisms. Post-hybridization washes partially determine stringency conditions.

The term "stringent conditions" or "stringent hybridization conditions" includes reference to conditions under which a probe will selectively hybridize to its target sequence. Stringent conditions are sequence-dependent and will be different in
15 different circumstances. By controlling the stringency of the hybridization and/or washing conditions, target sequences can be identified which are 100% complementary to the probe (homologous probing). Alternatively, stringency conditions can be adjusted to allow some mismatching in sequences so that lower degrees of similarity are detected (heterologous probing). Generally, a probe is less
20 than about 1000 nucleotides in length, optionally less than 500 nucleotides in length.

Typically, stringent conditions will be those in which the salt concentration is less than about 1.5 M Na ion, typically about 0.01 to 1.0 M Na ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes (*e.g.*, 10 to 50 nucleotides) and at least about 60°C for long probes (*e.g.*,
25 greater than 50 nucleotides). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide. Exemplary low stringency conditions include hybridization with a buffer solution of 30 to 35% formamide, 1 M NaCl, 1% SDS (sodium dodecyl sulphate) at 37°C, and a wash in 1X to 2X SSC (20X SSC = 3.0 M NaCl/0.3 M trisodium citrate) at 50 to 55°C. Exemplary moderate
30 stringency conditions include hybridization in 40 to 45% formamide, 1 M NaCl, 1% SDS at 37°C, and a wash in 0.5X to 1X SSC at 55 to 60°C. Exemplary high stringency conditions include hybridization in 50% formamide, 1 M NaCl, 1% SDS at 37°C, and a wash in 0.1X SSC at 60 to 65°C.

Specificity is typically the function of post-hybridization washes, the critical factors being the ionic strength and temperature of the final wash solution. For DNA-DNA hybrids, the T_m can be approximated from the equation of Meinkoth and Wahl, *Anal. Biochem.*, 138:267-284 (1984): $T_m = 81.5\text{ }^{\circ}\text{C} + 16.6 (\log M) + 0.41$
5 $(\%GC) - 0.61 (\% \text{ form}) - 500/L$; where M is the molarity of monovalent cations, $\%GC$ is the percentage of guanosine and cytosine nucleotides in the DNA, $\% \text{ form}$ is the percentage of formamide in the hybridization solution, and L is the length of the hybrid in base pairs. The T_m is the temperature (under defined ionic strength and pH) at which 50% of a complementary target sequence hybridizes to a perfectly
10 matched probe. T_m is reduced by about 1°C for each 1% of mismatching; thus, T_m , hybridization and/or wash conditions can be adjusted to hybridize to sequences of the desired identity. For example, if sequences with $\geq 90\%$ identity are sought, the T_m can be decreased 10°C . Generally, stringent conditions are selected to be about 5°C lower than the thermal melting point (T_m) for the specific sequence and its
15 complement at a defined ionic strength and pH. However, severely stringent conditions can utilize a hybridization and/or wash at 1, 2, 3, or 4°C lower than the thermal melting point (T_m); moderately stringent conditions can utilize a hybridization and/or wash at 6, 7, 8, 9, or 10°C lower than the thermal melting point (T_m); low stringency conditions can utilize a hybridization and/or wash at 11, 12, 13, 14, 15, or
20 20°C lower than the thermal melting point (T_m). Using the equation, hybridization and wash compositions, and desired T_m , those of ordinary skill will understand that variations in the stringency of hybridization and/or wash solutions are inherently described. If the desired degree of mismatching results in a T_m of less than 45°C (aqueous solution) or 32°C (formamide solution) it is preferred to increase the SSC
25 concentration so that a higher temperature can be used. An extensive guide to the hybridization of nucleic acids is found in Tijssen, *Laboratory Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Acid Probes*, Part I, Chapter 2 "Overview of principles of hybridization and the strategy of nucleic acid probe assays", Elsevier, New York (1993); and *Current Protocols in Molecular*
30 *Biology*, Chapter 2, Ausubel et al., Eds., Greene Publishing and Wiley-Interscience, New York (1995). Hybridization and/or wash conditions can be applied for at least 10, 30, 60, 90, 120, or 240 minutes.

Preferred substantially similar nucleic acid sequences encompassed by this invention are those sequences that are 80% identical to the nucleic acid fragments reported herein or which are 80% identical to any portion of the nucleotide sequences reported herein. More preferred are nucleic acid fragments which are 90% identical to the nucleic acid sequences reported herein, or which are 90% identical to any portion of the nucleotide sequences reported herein. Most preferred are nucleic acid fragments which are 95% identical to the nucleic acid sequences reported herein, or which are 95% identical to any portion of the nucleotide sequences reported herein. It is well understood by one skilled in the art that many levels of sequence identity are useful in identifying related polynucleotide sequences. Useful examples of percent identities are those listed above, or also preferred is any integer percentage from 80 % to 100%, such as 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% and 100%.

A “substantially homologous sequence” refers to variants of the disclosed sequences such as those that result from site-directed mutagenesis, as well as synthetically derived sequences. A substantially homologous sequence of the present invention also refers to those fragments of a particular promoter nucleotide sequence disclosed herein that operate to promote the constitutive expression of an operably linked heterologous nucleic acid fragment. These promoter fragments will comprise at least about 20 contiguous nucleotides, preferably at least about 50 contiguous nucleotides, more preferably at least about 75 contiguous nucleotides, even more preferably at least about 100 contiguous nucleotides of the particular promoter nucleotide sequence disclosed herein. The nucleotides of such fragments will usually comprise the TATA recognition sequence of the particular promoter sequence. Such fragments may be obtained by use of restriction enzymes to cleave the naturally occurring promoter nucleotide sequences disclosed herein; by synthesizing a nucleotide sequence from the naturally occurring promoter DNA sequence; or may be obtained through the use of PCR technology. See particularly, Mullis et al., *Methods Enzymol.* 155:335-350 (1987), and Higuchi, R. In *PCR Technology: Principles and Applications for DNA Amplifications*; Erlich, H.A., Ed.; Stockton Press Inc.: New York, 1989. Again, variants of these promoter fragments,

such as those resulting from site-directed mutagenesis, are encompassed by the compositions of the present invention.

“Codon degeneracy” refers to divergence in the genetic code permitting variation of the nucleotide sequence without affecting the amino acid sequence of an encoded polypeptide. Accordingly, the instant invention relates to any nucleic acid fragment comprising a nucleotide sequence that encodes all or a substantial portion of the amino acid sequences set forth herein. The skilled artisan is well aware of the “codon-bias” exhibited by a specific host cell in usage of nucleotide codons to specify a given amino acid. Therefore, when synthesizing a nucleic acid fragment for improved expression in a host cell, it is desirable to design the nucleic acid fragment such that its frequency of codon usage approaches the frequency of preferred codon usage of the host cell.

Sequence alignments and percent identity calculations may be determined using a variety of comparison methods designed to detect homologous sequences including, but not limited to, the Megalign® program of the LASERGENE® bioinformatics computing suite (DNASTAR® Inc., Madison, WI). Unless stated otherwise, multiple alignment of the sequences provided herein were performed using the Clustal V method of alignment (Higgins and Sharp (1989) CABIOS. 5:151-153) with the default parameters (GAP PENALTY=10, GAP LENGTH PENALTY=10). Default parameters for pairwise alignments and calculation of percent identity of protein sequences using the Clustal V method are KTUPLE=1, GAP PENALTY=3, WINDOW=5 and DIAGONALS SAVED=5. For nucleic acids these parameters are KTUPLE=2, GAP PENALTY=5, WINDOW=4 and DIAGONALS SAVED=4. After alignment of the sequences, using the Clustal V program, it is possible to obtain “percent identity” and “divergence” values by viewing the “sequence distances” table on the same program; unless stated otherwise, percent identities and divergences provided and claimed herein were calculated in this manner.

Alternatively, the Clustal W method of alignment may be used. The Clustal W method of alignment (described by Higgins and Sharp, CABIOS. 5:151-153 (1989); Higgins, D. G. et al., Comput. Appl. Biosci. 8:189-191 (1992)) can be found in the MegAlign™ v6.1 program of the LASERGENE® bioinformatics computing suite (DNASTAR® Inc., Madison, Wis.). Default parameters for multiple alignment

correspond to GAP PENALTY=10, GAP LENGTH PENALTY=0.2, Delay Divergent Sequences=30%, DNA Transition Weight=0.5, Protein Weight Matrix=Gonnet Series, DNA Weight Matrix=IUB. For pairwise alignments the default parameters are Alignment=Slow-Accurate, Gap Penalty=10.0, Gap Length=0.10, Protein Weight
5 Matrix=Gonnet 250 and DNA Weight Matrix=IUB. After alignment of the sequences using the Clustal W program, it is possible to obtain "percent identity" and "divergence" values by viewing the "sequence distances" table in the same program. In one embodiment the % sequence identity is determined over the entire length of the molecule (nucleotide or amino acid).

10 A "substantial portion" of an amino acid or nucleotide sequence comprises enough of the amino acid sequence of a polypeptide or the nucleotide sequence of a gene to afford putative identification of that polypeptide or gene, either by manual evaluation of the sequence by one skilled in the art, or by computer-automated sequence comparison and identification using algorithms such as BLAST (Altschul,
15 S. F. et al., J. Mol. Biol. 215:403 410 (1993)) and Gapped Blast (Altschul, S. F. et al., Nucleic Acids Res. 25:3389 3402 (1997)). BLASTN refers to a BLAST program that compares a nucleotide query sequence against a nucleotide sequence database.

"Transgenic" refers to any cell, cell line, callus, tissue, plant part or plant, the
20 genome of which has been altered by the presence of a heterologous nucleic acid, such as a recombinant DNA construct, including those initial transgenic events as well as those created by sexual crosses or asexual propagation from the initial transgenic event. The heterologous polynucleotide may be integrated into the genome alone or as part of a recombinant DNA construct. The term "transgenic" as
25 used herein does not encompass the alteration of the genome (chromosomal or extra-chromosomal) by conventional plant breeding methods or by naturally occurring events such as random cross-fertilization, non-recombinant viral infection, non-recombinant bacterial transformation, non-recombinant transposition, or spontaneous mutation.

30 "Transgenic plant" includes reference to a plant which comprises within its genome a heterologous polynucleotide. For example, the heterologous polynucleotide is stably integrated within the genome such that the polynucleotide is

passed on to successive generations. The heterologous polynucleotide may be integrated into the genome alone or as part of a recombinant DNA construct.

“Transient expression” refers to the temporary expression of a gene of interest. These include but are not limited to reporter genes such as β -glucuronidase (GUS), fluorescent protein genes ZS-GREEN1, ZS-YELLOW1 N1, AM-CYAN1, DS-RED in certain cell types of the host organism in which the transgenic gene is introduced temporally by a transformation method.

Units, prefixes, and symbols may be denoted in their SI accepted form.

Unless otherwise indicated, nucleic acids are written left to right in 5' to 3'

orientation; amino acid sequences are written left to right in amino to carboxyl orientation, respectively. Numeric ranges recited within the specification are

inclusive of the numbers defining the range and include each integer within the defined range. Amino acids may be referred to herein by either commonly known three letter symbols or by the one-letter symbols recommended by the IUPAC-IUB

Biochemical Nomenclature Commission. Nucleotides, likewise, may be referred to

by their commonly accepted single-letter codes. Unless otherwise provided for, software, electrical, and electronics terms as used herein are as defined in The New IEEE Standard Dictionary of Electrical and Electronics Terms (5th edition, 1993).

The terms defined below are more fully defined by reference to the specification as

a whole.

The present invention concerns methods and compositions useful in altering gene expression in plants by up or down-regulating lincRNAs to obtain useful phenotypes. Up regulation of a lincRNA, or expressing a lincRNA in a tissue where it is not normally expressed, will enhance the normal effect of that lincRNA. Down-

regulationlinc RNAs, through the methods described above, will decrease the normal effect of that lincRNA. Since lincRNAs are involved in multiple plant

biochemical or signal transduction pathways, and since individual lincRNAs may impact expression of multiple genes, up or down regulation of lincRNAs will provide a powerful method to alter gene regulation to obtain useful phenotypes such as

those described herein.

In one embodiment, the present invention relates, *inter alia*, to an isolated polynucleotide comprising an lincRNA selected from the group consisting of SEQ ID NOs:1-203,370.

In another embodiment, the invention relates to a full-length complement of the lincRNA of selected from the group consisting of SEQ ID NOs:1-203,370 or a nucleotide sequence capable of hybridizing to these a fore mention lincRNA

Computational identification of lincRNAs was accomplished from RNA
 5 sequence analysis of 15 inbred lines of *Zea mays*. Assembled transcripts similar to known coding RNAs and ncRNAs were removed computationally, retaining only those meeting size criteria (greater than, or equal to, 200 nt in length and incapable of encoding a polypeptide of greater than, or equal to, 100 amino acids in length.

Transformation protocols as well as protocols for introducing nucleotide
 10 sequences into plants may vary depending on the type of plant or plant cell, *i.e.*, monocot or dicot, targeted for transformation. Suitable methods of introducing the DNA construct include microinjection (Crossway et al. (1986) *Biotechniques* 4:320-334; and U.S. Patent No. 6,300,543), sexual crossing, electroporation (Riggs et al. (1986) *Proc. Natl. Acad. Sci. USA* 83:5602-5606), *Agrobacterium*-mediated
 15 transformation (Townsend et al., U.S. Pat No. 5,563,055; and U.S. Patent No. 5,981,840), direct gene transfer (Paszkowski et al. (1984) *EMBO J.* 3:2717-2722), and ballistic particle acceleration (see, for example, Sanford et al., U.S. Patent No. 4,945,050; Tomes et al., U.S. Patent No. 5,879,918; Tomes et al., U.S. Patent No. 5,886,244; Bidney et al., U.S. Patent No. 5,932,782; Tomes et al. (1995) "Direct
 20 DNA Transfer into Intact Plant Cells via Microprojectile Bombardment," in *Plant Cell, Tissue, and Organ Culture: Fundamental Methods*, ed. Gamborg and Phillips (Springer-Verlag, Berlin); and McCabe et al. (1988) *Biotechnology* 6:923-926). Also see Weissinger et al. (1988) *Ann. Rev. Genet.* 22:421-477; Sanford et al. (1987) *Particulate Science and Technology* 5:27-37 (onion); Christou et al. (1988) *Plant
 25 Physiol.* 87:671-674 (soybean); Finer and McMullen (1991) *In Vitro Cell Dev. Biol.* 27P:175-182 (soybean); Singh et al. (1998) *Theor. Appl. Genet.* 96:319-324 (soybean); Datta et al. (1990) *Biotechnology* 8:736-740 (rice); Klein et al. (1988) *Proc. Natl. Acad. Sci. USA* 85:4305-4309 (maize); Klein et al. (1988) *Biotechnology* 6:559-563 (maize); Tomes, U.S. Patent No. 5,240,855; Busing et al., U.S. Patent
 30 Nos. 5,322,783 and 5,324,646; Klein et al. (1988) *Plant Physiol.* 91:440-444 (maize); Fromm et al. (1990) *Biotechnology* 8:833-839 (maize); Hooykaas-Van Slogteren et al. (1984) *Nature* (London) 311:763-764; Bowen *et al.*, U.S. Patent No. 5,736,369 (cereals); Bytebier et al. (1987) *Proc. Natl. Acad. Sci. USA* 84:5345-5349

(Liliaceae); De Wet et al. (1985) in *The Experimental Manipulation of Ovule Tissues*, ed. Chapman et al. (Longman, New York), pp. 197-209 (pollen); Kaeppler et al. (1990) *Plant Cell Reports* 9:415-418 and Kaeppler et al. (1992) *Theor. Appl. Genet.* 84:560-566 (whisker-mediated transformation); D'Halluin et al. (1992) *Plant Cell* 4:1495-1505 (electroporation); Li et al. (1993) *Plant Cell Reports* 12:250-255 and Christou and Ford (1995) *Annals of Botany* 75:407-413 (rice); Osjoda et al. (1996) *Nature Biotechnology* 14:745-750 (maize via *Agrobacterium tumefaciens*); and U.S. Patent No. 5,736,369 (meristem transformation), all of which are herein incorporated by reference.

The nucleotide constructs may be introduced into plants by contacting plants with a virus or viral nucleic acids. Generally, such methods involve incorporating a nucleotide construct of the invention within a viral DNA or RNA molecule. Further, it is recognized that useful promoters encompass promoters utilized for transcription by viral RNA polymerases. Methods for introducing nucleotide constructs into plants and expressing a protein encoded therein, involving viral DNA or RNA molecules, are known in the art. See, for example, U.S. Patent Nos. 5,889,191, 5,889,190, 5,866,785, 5,589,367 and 5,316,931; herein incorporated by reference.

In some embodiments, transient expression may be desired. In those cases, standard transient transformation techniques may be used. Such methods include, but are not limited to viral transformation methods, and microinjection of DNA or RNA, as well other methods well known in the art.

The cells from the plants that have stably incorporated the nucleotide sequence may be grown into plants in accordance with conventional ways. See, for example, McCormick et al. (1986) *Plant Cell Reports* 5:81-84. These plants may then be grown, and either pollinated with the same transformed strain or different strains, and the resulting hybrid having constitutive expression of the desired phenotypic characteristic imparted by the nucleotide sequence of interest and/or the genetic markers contained within the target site or transfer cassette. Two or more generations may be grown to ensure that expression of the desired phenotypic characteristic is stably maintained and inherited and then seeds harvested to ensure expression of the desired phenotypic characteristic has been achieved.

Initial identification and selection of cells and/or plants comprising the recombinant DNA constructs may be facilitated by the use of marker genes. Gene

targeting can be performed without selection if there is a sensitive method for identifying recombinants, for example if the targeted gene modification can be easily detected by PCR analysis, or if it results in a certain phenotype. However, in most cases, identification of gene targeting events will be facilitated by the use of markers. Useful markers include positive and negative selectable markers as well as markers that facilitate screening, such as visual markers. Selectable markers include genes carrying resistance to an antibiotic such as spectinomycin (e.g. the *aadA* gene, Svab et al. 1990 *Plant Mol. Biol.* 14:197), streptomycin (e.g., *aadA*, or SPT, Svab et al. 1990 *Plant Mol. Biol.* 14:197; Jones et al. 1987 *Mol. Gen. Genet.* 210:86), kanamycin (e.g., *nptII*, Fraley et al. 1983 *PNAS* 80:4803), hygromycin (e.g., HPT, Vanden Elzen et al. 1985 *Plant Mol. Biol.* 5:299), gentamycin (Hayford et al. 1988 *Plant Physiol.* 86:1216), phleomycin, zeocin, or bleomycin (Hille et al. 1986 *Plant Mol. Biol.* 7:171), or resistance to a herbicide such as phosphinothricin (bar gene), or sulfonylurea (acetolactate synthase (ALS)) (Charest et al. (1990) *Plant Cell Rep.* 8:643), genes that fulfill a growth requirement on an incomplete media such as *HIS3*, *LEU2*, *URA3*, *LYS2*, and *TRP1* genes in yeast, and other such genes known in the art. Negative selectable markers include cytosine deaminase (*codA*) (Stougaard 1993 *Plant J.* 3:755-761), *tms2* (DePicker et al. 1988 *Plant Cell Rep.* 7:63-66), nitrate reductase (Nussame et al. 1991 *Plant J.* 1:267-274), *SU1* (O'Keefe et al. 1994 *Plant Physiol.* 105:473-482), *aux-2* from the Ti plasmid of *Agrobacterium*, and thymidine kinase. Screenable markers include fluorescent proteins such as green fluorescent protein (GFP) (Chalfie et al., 1994 *Science* 263:802; US 6,146,826; US 5,491,084; and WO 97/41228), reporter enzymes such as - glucuronidase (GUS) (Jefferson R.A. 1987 *Plant Mol. Biol. Rep.* 5:387; US 5,599,670; and US 5,432,081), -galactosidase (*lacZ*), alkaline phosphatase (AP), glutathione S-transferase (GST) and luciferase (US 5,674,713; and Ow et al. 1986 *Science* 234(4778):856-859), visual markers like anthocyanins such as CRC (Ludwig et al. (1990) *Science* 247(4841):449-450) R gene family (e.g. Lc, P, S), A, C, R-nj, body and/or eye color genes in *Drosophila*, coat color genes in mammalian systems, and others known in the art.

One or more markers may be used in order to select and screen for gene targeting events. One common strategy for gene disruption involves using a target modifying polynucleotide in which the target is disrupted by a promoterless

selectable marker. Since the selectable marker lacks a promoter, random integration events are unlikely to lead to transcription of the gene. Gene targeting events will put the selectable marker under control of the promoter for the target gene. Gene targeting events are identified by selection for expression of the selectable marker. Another common strategy utilizes a positive-negative selection scheme. This scheme utilizes two selectable markers, one that confers resistance (R+) coupled with one that confers a sensitivity (S+), each with a promoter. When this polynucleotide is randomly inserted, the resulting phenotype is R+/S+. When a gene targeting event is generated, the two markers are uncoupled and the resulting phenotype is R+/S-. Examples of using positive-negative selection are found in Thykjær et al. (1997) *Plant Mol. Biol.* 35:523-530; and WO 01/66717, which are herein incorporated by reference.

Non-limiting examples of methods and compositions disclosed herein are as follows:

1. An isolated polynucleotide comprising:

- (a) a nucleotide sequence selected from the group consisting of SEQ ID NOs:1-203,370; or,
- (b) a full-length complement of (a); or,
- (c) a nucleotide sequence comprising a sequence having at least 95% sequence identity, based on the Clustal V method of alignment with pairwise alignment default parameters (KTUPLE=2, GAP PENALTY=5, WINDOW=4 and DIAGONALS SAVED=4), when compared to the nucleotide sequence of (a);

wherein said nucleotide sequence is a long intergenic non-coding RNA (lincRNA).

2. The isolated polynucleotide of embodiment 1, comprising a nucleotide sequence having at least 98% sequence identity, based on the Clustal V method of alignment with pairwise alignment default parameters (KTUPLE=2, GAP PENALTY=5, WINDOW=4 and DIAGONALS SAVED=4), when compared to the nucleotide sequence of (a);
3. A recombinant DNA construct comprising the isolated polynucleotide of Embodiment 1 operably linked to at least one regulatory sequence.
4. A vector comprising the recombinant DNA construct of embodiment 3.
5. A cell comprising the recombinant DNA construct of embodiment 3.

6. The cell of embodiment 5 wherein the cell is a plant cell.
7. A transgenic plant comprising the recombinant DNA construct of embodiment 3.
8. The transgenic plant of embodiment 7 wherein said plant is a monocot plant.
- 5 9. The transgenic plant of embodiment 7 wherein the plant is selected from the group consisting of maize, rice, sorghum, sunflower, millet, soybean, canola, wheat, barley, oat, beans, and nuts.
- 10 10. A transgenic seed obtained from the plant of Embodiment 7.
11. Progeny plants obtained from the seeds of Embodiment 10.
- 10 12. Transformed plant tissue or a plant cell comprising in its genome the recombinant DNA construct of Embodiment 3.
13. The transformed plant tissue or plant cell of Embodiment 12 wherein the plant is selected from the group consisting of corn, rice, sorghum, millet, soybean, canola, wheat, barley, oat, beans, and nuts.
- 15 14. A method of expressing a lincRNA in a plant comprising:
 - a) introducing the recombinant DNA construct of embodiment 3 into the plant, wherein the at least one heterologous nucleotide sequence comprises a lincRNA;
 - b) growing the plant of step a); and,
 - 20 c) selecting a plant displaying expression of the lincRNA.
15. A method of transgenically altering a marketable plant trait, comprising:
 - a) introducing a recombinant DNA construct of embodiment 3 into the plant;
 - b) growing a fertile, mature plant resulting from step a); and,
 - 25 c) selecting a plant expressing the at least one heterologous nucleotide sequence in at least one plant tissue based on the altered marketable trait.
16. The method of embodiment 15 wherein the marketable trait is selected from the group consisting of: disease resistance, herbicide resistance, insect resistance carbohydrate metabolism, fatty acid metabolism, amino acid metabolism, plant development, plant growth regulation, yield improvement, drought resistance, cold resistance, heat resistance, and salt resistance.
- 30 17. A method for altering expression of a gene of interest in a plant comprising:

- (a) transforming a plant cell with the recombinant DNA construct of Embodiment 3;
- (b) growing fertile mature plants from transformed plant cell of step (a); and
- (c) selecting plants containing the transformed plant cell wherein the expression of the gene of interest is increased or decreased.

EXAMPLES

The following are non-limiting examples intended to illustrate the invention. Although the present invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be obvious that certain changes and modifications may be practiced within the scope of the appended claims.

EXAMPLE 1

Transformation of Plants

Described in this example are methods one may use for introduction of a polynucleotide or polypeptide into a plant cell.

A. Maize particle-mediated DNA delivery

A DNA construct can be introduced into maize cells capable of growth on suitable maize culture medium. Such competent cells can be from maize suspension culture, callus culture on solid medium, freshly isolated immature embryos or meristem cells. Immature embryos of the Hi-II genotype can be used as the target cells. Ears are harvested at approximately 10 days post-pollination, and 1.2-1.5mm immature embryos are isolated from the kernels, and placed scutellum-side down on maize culture medium.

The immature embryos are bombarded from 18-72 hours after being harvested from the ear. Between 6 and 18 hours prior to bombardment, the immature embryos are placed on medium with additional osmoticum (MS basal medium, Musashige and Skoog, 1962, *Physiol. Plant* 15:473-497, with 0.25 M sorbitol). The embryos on the high-osmotic medium are used as the bombardment target, and are left on this medium for an additional 18 hours after bombardment.

For particle bombardment, plasmid DNA (described above) is precipitated onto 1.8 mm tungsten particles using standard CaCl₂- spermidine chemistry (see, for example, Klein et al., 1987, *Nature* 327:70-73). Each plate is bombarded once

at 600 PSI, using a DuPont Helium Gun (Lowe et al., 1995, *Bio/Technol* 13:677-682). For typical media formulations used for maize immature embryo isolation, callus initiation, callus proliferation and regeneration of plants, see Armstrong, C., 1994, In "The Maize Handbook", M. Freeling and V. Walbot, eds. Springer Verlag, NY, pp 663-671.

Within 1-7 days after particle bombardment, the embryos are moved onto N6-based culture medium containing 3 mg/l of the selective agent bialaphos. Embryos, and later callus, are transferred to fresh selection plates every 2 weeks. The calli developing from the immature embryos are screened for the desired phenotype.

After 6-8 weeks, transformed calli are recovered.

B. Soybean transformation

Soybean embryogenic suspension cultures are maintained in 35 ml liquid media SB196 or SB172 in 250 ml Erlenmeyer flasks on a rotary shaker, 150 rpm, 26°C with cool white fluorescent lights on 16:8 hr day/night photoperiod at light intensity of 30-35 uE/m²s. Cultures are subcultured every two weeks by inoculating approximately 35 mg of tissue into 35 ml of fresh liquid media. Alternatively, cultures are initiated and maintained in 6-well Costar plates.

SB 172 media is prepared as follows: (per liter), 1 bottle Murashige and Skoog Medium (Duchefa # M 0240), 1 ml B5 vitamins 1000X stock, 1 ml 2,4-D stock (Gibco 11215-019), 60 g sucrose, 2 g MES, 0.667 g L-Asparagine anhydrous (GibcoBRL 11013-026), pH 5.7. SB 196 media is prepared as follows: (per liter) 10ml MS FeEDTA, 10ml MS Sulfate, 10ml FN-Lite Halides, 10ml FN-Lite P,B,Mo, 1ml B5 vitamins 1000X stock, 1 ml 2,4-D, (Gibco 11215-019), 2.83g KNO₃, 0.463g (NH₄)₂SO₄, 2g MES, 1g Asparagine Anhydrous, Powder (Gibco 11013-026), 10g Sucrose, pH 5.8. 2,4-D stock concentration 10 mg/ml is prepared as follows: 2,4-D is solubilized in 0.1 N NaOH, filter-sterilized, and stored at -20°C. B5 vitamins 1000X stock is prepared as follows: (per 100 ml) - store aliquots at -20°C, 10 g myo-inositol, 100 mg nicotinic acid, 100 mg pyridoxine HCl, 1 g thiamin.

Soybean embryogenic suspension cultures are transformed with various plasmids by the method of particle gun bombardment (Klein et al., 1987 *Nature* 327:70). To prepare tissue for bombardment, approximately two flasks of suspension culture tissue that has had approximately 1 to 2 weeks to recover since its most recent subculture is placed in a sterile 60 x 20 mm petri dish containing 1 sterile

filter paper in the bottom to help absorb moisture. Tissue (*i.e.* suspension clusters approximately 3-5 mm in size) is spread evenly across each petri plate. Residual liquid is removed from the tissue with a pipette, or allowed to evaporate to remove excess moisture prior to bombardment. Per experiment, 4 - 6 plates of tissue are
5 bombarded. Each plate is made from two flasks.

To prepare gold particles for bombardment, 30 mg gold is washed in ethanol, centrifuged and resuspended in 0.5 ml of sterile water. For each plasmid combination (treatments) to be used for bombardment, a separate micro-centrifuge tube is prepared, starting with 50 µl of the gold particles prepared above. Into each
10 tube, the following are also added; 5µl of plasmid DNA (at 1µg/µl), 50µl CaCl₂, and 20µl 0.1 M spermidine. This mixture is agitated on a vortex shaker for 3 minutes, and then centrifuged using a microcentrifuge set at 14,000 RPM for 10 seconds. The supernatant is decanted and the gold particles with attached, precipitated DNA are washed twice with 400 µl aliquots of ethanol (with a brief centrifugation as above
15 between each washing). The final volume of 100% ethanol per each tube is adjusted to 40µl, and this particle/DNA suspension is kept on ice until being used for bombardment.

Immediately before applying the particle/DNA suspension, the tube is briefly dipped into a sonicator bath to disperse the particles, and then 5 L of DNA prep is
20 pipetted onto each flying disk and allowed to dry. The flying disk is then placed into the DuPont Biolistics PDS1000/HE. Using the DuPont Biolistic PDS1000/HE instrument for particle-mediated DNA delivery into soybean suspension clusters, the following settings are used. The membrane rupture pressure is 1100 psi. The chamber is evacuated to a vacuum of 27-28 inches of mercury. The tissue is placed
25 approximately 3.5 inches from the retaining/stopping screen (3rd shelf from the bottom). Each plate is bombarded twice, and the tissue clusters are rearranged using a sterile spatula between shots.

Following bombardment, the tissue is re-suspended in liquid culture medium, each plate being divided between 2 flasks with fresh SB196 or SB172 media and
30 cultured as described above. Four to seven days post-bombardment, the medium is replaced with fresh medium containing a selection agent. The selection media is refreshed weekly for 4 weeks and once again at 6 weeks. Weekly replacement after 4 weeks may be necessary if cell density and media turbidity is high.

Four to eight weeks post-bombardment, green, transformed tissue may be observed growing from untransformed, necrotic embryogenic clusters. Isolated, green tissue is removed and inoculated into 6-well microtiter plates with liquid medium to generate clonally-propagated, transformed embryogenic suspension cultures.

Each embryogenic cluster is placed into one well of a Costar 6-well plate with 5mls fresh SB196 media with selection agent. Cultures are maintained for 2-6 weeks with fresh media changes every 2 weeks. When enough tissue is available, a portion of surviving transformed clones are subcultured to a second 6-well plate as a back-up to protect against contamination.

To promote in vitro maturation, transformed embryogenic clusters are removed from liquid SB196 and placed on solid agar media, SB 166, for 2 weeks. Tissue clumps of 2 - 4 mm size are plated at a tissue density of 10 to 15 clusters per plate. Plates are incubated in diffuse, low light ($< 10 \mu\text{E}$) at $26 \pm 1^\circ\text{C}$. After two weeks, clusters are subcultured to SB 103 media for 3 - 4 weeks.

SB 166 is prepared as follows: (per liter), 1 pkg. MS salts (Gibco/ BRL - Cat# 11117-017), 1 ml B5 vitamins 1000X stock, 60 g maltose, 750 mg MgCl_2 hexahydrate, 5 g activated charcoal, pH 5.7, 2 g gelrite. SB 103 media is prepared as follows: (per liter), 1 pkg. MS salts (Gibco/BRL - Cat# 11117-017), 1 ml B5 vitamins 1000X stock, 60 g maltose, 750 mg MgCl_2 hexahydrate, pH 5.7, 2 g gelrite. After 5-6 week maturation, individual embryos are desiccated by placing embryos into a 100 X 15 petri dish with a 1cm² portion of the SB103 media to create a chamber with enough humidity to promote partial desiccation, but not death.

Approximately 25 embryos are desiccated per plate. Plates are sealed with several layers of parafilm and again are placed in a lower light condition. The duration of the desiccation step is best determined empirically, and depends on size and quantity of embryos placed per plate. For example, small embryos or few embryos/plate require a shorter drying period, while large embryos or many embryos/plate require a longer drying period. It is best to check on the embryos after about 3 days, but proper desiccation will most likely take 5 to 7 days. Embryos will decrease in size during this process.

Desiccated embryos are planted in SB 71-1 or MSO medium where they are left to germinate under the same culture conditions described for the suspension

cultures. When the plantlets have two fully-expanded trifoliate leaves, germinated and rooted embryos are transferred to sterile soil and watered with MS fertilizer. Plants are grown to maturity for seed collection and analysis. Healthy, fertile transgenic plants are grown in the greenhouse.

- 5 SB 71-1 is prepared as follows: 1 bottle Gamborg's B5 salts w/ sucrose (Gibco/BRL - Cat# 21153-036), 10 g sucrose, 750 mg MgCl₂ hexahydrate, pH 5.7, 2 g gelrite. MSO media is prepared as follows: 1 pkg Murashige and Skoog salts (Gibco 11117-066), 1 ml B5 vitamins 1000X stock, 30 g sucrose, pH 5.8, 2g Gelrite.

10

EXAMPLE 2

Identification of long intergenic non-coding RNAs (lincRNAs) in maize

Computational identification of lincRNAs was accomplished from RNA sequence analysis of 15 inbred lines of *Zea mays*. Assembled transcripts similar to known coding RNAs and ncRNAs were removed by identification by BLAST

15 analyses (ref: Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schaffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Res. 25:3389-3402.) with an "Expectation value" cut-off of 1e-9 and low-complexity filtering off.

20

Of those sequences with little-to-no similarity to known coding RNAs and ncRNAs, BLAST was used to identify candidate transcripts mapping to intergenic regions. Intergenic regions are defined as those not specified as mapped gene components in ZmB73_refgen_v2 (MaizeSequence.org). Computational methods were then used to identify those meeting size criteria (greater than, or

25 equal to, 200 nt in length) and coding criteria (encoding a polypeptide of less than, or equal to, 100 amino acids in length).

CLAIMS

WHAT IS CLAIMED

1. An isolated polynucleotide comprising:

- 5 a. a nucleotide sequence selected from the group consisting of SEQ ID
 NOs:1-203,370; or,
 b. a full-length complement of (a); or,
 c. a nucleotide sequence comprising a sequence having at least 95%
 sequence identity, based on the Clustal V method of alignment with
10 pairwise alignment default parameters (KTUPLE=2, GAP
 PENALTY=5, WINDOW=4 and DIAGONALS SAVED=4), when
 compared to the nucleotide sequence of (a);

 wherein said nucleotide sequence is a long intergenic non-coding RNA
 (lincRNA).

- 15 2. The isolated polynucleotide of claim 1, comprising a nucleotide sequence
 having at least 98% sequence identity, based on the Clustal V method of
 alignment with pairwise alignment default parameters (KTUPLE=2, GAP
 PENALTY=5, WINDOW=4 and DIAGONALS SAVED=4), when compared to
20 the nucleotide sequence of (a);

3. A recombinant DNA construct comprising the isolated polynucleotide of
 Claim 1 operably linked to at least one regulatory sequence.

- 25 4. A vector comprising the recombinant DNA construct of claim 3.

5. A cell comprising the recombinant DNA construct of claim 3.

6. The cell of claim 5 wherein the cell is a plant cell.

- 30 7. A transgenic plant comprising the recombinant DNA construct of claim 3.

8. The transgenic plant of claim 7 wherein said plant is a monocot plant.

9. The transgenic plant of claim 7 wherein the plant is selected from the group consisting of maize, rice, sorghum, sunflower, millet, soybean, canola, wheat, barley, oat, beans, and nuts.
- 5 10. A transgenic seed obtained from the plant of Claim 7.
11. Progeny plants obtained from the seeds of Claim 10.
- 10 12. Transformed plant tissue or a plant cell comprising in its genome the recombinant DNA construct of Claim 3.
13. The transformed plant tissue or plant cell of Claim 12 wherein the plant is selected from the group consisting of corn, rice, sorghum, millet, soybean, canola, wheat, barley, oat, beans, and nuts.
- 15 14. A method of expressing a lincRNA in a plant comprising:
- a) introducing the recombinant DNA construct of claim 3 into the plant, wherein the at least one heterologous nucleotide sequence comprises a lincRNA;
- 20 b) growing the plant of step a); and,
- c) selecting a plant displaying expression of the lincRNA.
15. A method of transgenically altering a marketable plant trait, comprising:
- a) introducing a recombinant DNA construct of claim 3 into the plant;
- 25 b) growing a fertile, mature plant resulting from step a); and,
- c) selecting a plant expressing the at least one heterologous nucleotide sequence in at least one plant tissue based on the altered marketable trait.
- 30 16. The method of claim 15 wherein the marketable trait is selected from the group consisting of: disease resistance, herbicide resistance, insect resistance carbohydrate metabolism, fatty acid metabolism, amino acid metabolism, plant development, plant growth regulation, yield

improvement, drought resistance, cold resistance, heat resistance, and salt resistance.

- 5 17. A method for altering expression of a gene of interest in a plant comprising:
- a. transforming a plant cell with the recombinant DNA construct of Claim 3;
 - b. growing fertile mature plants from transformed plant cell of step (a); and,
 - c. selecting plants containing the transformed plant cell wherein the expression of the gene of interest is increased or decreased.

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2013/056931

A. CLASSIFICATION OF SUBJECT MATTER
INV. C07K14/415 C12N15/82
ADD.

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)
C07K C12N

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

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

EPO-Internal, Sequence Search, BIOSIS, EMBASE, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE EMBL [Online] 7 August 2007 (2007-08-07), "Zea mays chromosome 8 clone CH201-261E9, *** SEQUENCING IN PROGRESS ***, 4 unordered pieces.", XP002716527, retrieved from EBI accession no. EM NEW:AC207791 Database accession no. AC207791 see nucleotides 126792 to 127261; sequence ----- -/-	1-6



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents :

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

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

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

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

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

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search

18 November 2013

Date of mailing of the international search report

18/02/2014

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Oderwald, Harald

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2013/056931

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	BOERNER SUSAN ET AL: "Computational Identification and Functional Predictions of Long Noncoding RNA in Zea mays", PLOS ONE, vol. 7, no. 8, 16 August 2012 (2012-08-16), XP002716528, the whole document	1-17
A	----- ZHU Q -H ET AL: "Molecular functions of long non-coding RNAs in plants", GENES 2012 MDPI AG CHE, vol. 3, no. 1, March 2012 (2012-03), pages 176-190, XP002716529, ISSN: 2073-4425 cited in the application the whole document	1
A	----- AU PHIL CHI KHANG ET AL: "Long non-coding RNA-mediated mechanisms independent of the RNAi pathway in animals and plants", RNA BIOLOGY, vol. 8, no. 3, May 2011 (2011-05), pages 404-414, XP002716530, the whole document	1
A	----- JACQUIER ALAIN: "The complex eukaryotic transcriptome: unexpected pervasive transcription and novel small RNAs", NATURE REVIEWS GENETICS, vol. 10, no. 12, December 2009 (2009-12), pages 833-844, XP002716531, ISSN: 1471-0056 cited in the application the whole document -----	1

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2013/056931

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

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

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

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

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

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of additional 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.:

1-17(partially)

Remark on Protest

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

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. claims: 1-17(partially)

Polynucleotide comprising a nucleotide sequence as shown in
SEQ ID NO: 1. Vector, cell, transgenic plant, seed, progeny,
plant tissue, and methods relating to said polynucleotide.

2-204. claims: 1-17(partially)

same as invention but relating to a polynucleotide shown in
SEQ ID NO: 2 (invention 2),... or SEQ ID NO: 370 (invention
204), respectively.
