(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
18 January 2001 (18.01.2001)

PCT

(10) International Publication Number
WO 01/04285 A2

(51) International Patent Classification[7]: C12N 15/00

(21) International Application Number: PCT/US00/40371

(22) International Filing Date: 13 July 2000 (13.07.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/143,437 13 July 1999 (13.07.1999) US

(71) Applicant *(for all designated States except US)*: **PIONEER HI-BRED INTERNATIONAL, INC.** [US/US]; 800 Capital Square, 400 Locust Street, Des Moines, IA 50309 (US).

(72) Inventor; and

(75) Inventor/Applicant *(for US only)*: **BALDWIN, Donald, Adelphi** [US/US]; 1113 Douglas Avenue, Des Moines, IA 50313 (US).

(74) Agents: **WILLIAMSON, Kelly, J.** et al.; Alston & Bird LLP, P.O. Drawer 34009, Charlotte, NC 28234-4009 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AT (utility model), AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, CZ (utility model), DE, DE (utility model), DK, DK (utility model), DM, DZ, EE, EE (utility model), ES, FI, FI (utility model), GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SK (utility model), SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— *Without international search report and to be republished upon receipt of that report.*

— *With (an) indication(s) in relation to deposited biological material furnished under Rule 13bis separately from the description.*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/04285 A2

(54) Title: MAIZE MSI POLYNUCLEOTIDES AND METHODS OF USE

(57) Abstract: Methods and compositions for modulating plant cell cycle, DNA metabolism, and pathogen defense are provided. Nucleotide sequences encoding maize retinoblastoma binding proteins are provided. The sequence can be used in expression cassettes for modulating DNA replication, DNA repair, recombination, and pathogen defense. Transformed plants, plant cells, tissues, and seed are also provided.

MAIZE MSI POLYNUCLEOTIDES AND METHODS OF USE

FIELD OF THE INVENTION

The invention relates to the genetic manipulation of plants, particularly the modulation of gene activity in plants and increased disease resistance.

BACKGROUND OF THE INVENTION

Disease in plants is caused by biotic and abiotic causes. Biotic causes include fungi, viruses, bacteria, and nematodes. Among the causative agents of infectious disease of crop plants, phytopathogenic fungi play the dominant role. Phytopathogenic fungi cause significant annual crop yield losses as well as devastating epidemics. Plant disease outbreaks have resulted in catastrophic crop failures that have triggered famines and caused major social change. All of the approximately 300,000 species of flowering plants are attacked by pathogenic fungi; however, a single plant species can be host to only a few fungal species, and similarly, most fungi usually have a limited host range. Generally, the best strategy for plant disease control is to use resistant cultivars selected or developed by plant breeders for this purpose. However, the potential for serious crop disease epidemics persists today, as evidenced by outbreaks of the Victoria blight of oats and southern corn leaf blight. Molecular methods of crop protection have the potential to implement novel mechanisms for disease resistance and can also be implemented more quickly than traditional breeding methods. Accordingly, molecular methods are needed to supplement traditional breeding methods to protect plants from pathogen attack.

A host of cellular processes enable plants to defend themselves against disease caused by pathogenic agents. These defense mechanisms are activated by initial pathogen infection in a process known as elicitation. In elicitation, the host plant recognizes a pathogen-derived compound known as an elicitor; the plant then activates disease gene expression to limit further spread of the invading microorganism. It is generally believed that to overcome these plant defense mechanisms, plant pathogens must find a way to suppress elicitation.

One necrotrophic pathogen, the filamentous fungus *Cochliobolus carbonum* race 1, synthesizes a cyclic tetrapeptide known as HC-toxin which is absolutely required for pathogenicity. Resistant maize genotypes produce an HC-toxin reductase encoded by the nuclear *Hm* locus which abolishes toxin activity by reducing the ketone group. HC-toxin acts in a cytostatic manner; it is not toxic to plant cells and does not determine pathogenicity by simply killing host cells prior to colonization. HC-toxin is a specific inhibitor of histone deacetylase (HD) activity.

Histones are proteins which bind DNA to form a complex termed the nucleosome. Nucleosomes structurally organize and compact chromosomal DNA into chromatin. In chromatin, the degree of interaction between histones and DNA varies with transcriptional activity. Chromatin regions containing active promoters often have histones which are hyperacetylated. Hyperacetylated histones are thought to adopt a chromatin structure that allows other proteins to bind promoter DNA and activate transcription. In contrast, hypoacetylated histones are associated with inactive promoters, and deacetylation of histones in normally active chromatin can repress transcription in that region. Histone deacetylase (HD) is responsible for removing acetyl modifications from histones. HD may be localized to promoters targeted for repression by other proteins that associate with HD and specifically bind regulatory elements in promoter DNA. HD and other proteins bind retinoblastoma protein (pRb) to help form the multiprotein retinoblastoma complex. When the retinoblastoma complex is localized to its DNA binding site in a regulated promoter, transcription is repressed by a number of mechanisms including chromatin histone deacetylation. The retinoblastoma complex is a key regulator of cell cycle progression through the G1 phase and also in cellular differentiation; loss of retinoblastoma function plays a role in many cancers.

Thus, histone acetylation plays a key role in fundamental cellular processes such as transcription and cell cycle progression as well as processes such as plant resistance to pathogen invasion. Mechanisms are therefore needed to modulate acetylation in order to control gene activities, cancer, and plant disease resistance.

SUMMARY OF THE INVENTION

The present invention provides nucleotide and amino acid sequences that find use in modulating development, developmental pathways, and the plant pathogen defense system. Particularly, the nucleotide and amino acid sequence for two maize retinoblastoma-associated-like proteins (Rb-Ap-like or MSI-like) proteins are provided.

The methods and compositions of the present invention can be used to modulate development in a host cell. For instance, the methods and compositions of the invention can be used to modulate the plant pathogen defense system. More specifically, methods and compositions may be used for enhancing resistance to plant pathogens including fungal pathogens, plant viruses, and the like. The compositions and methods of the invention can also be used to alter metabolic states of host cells. In particular, methods are provided to modulate cell division, differentiation, as well as cellular processes controlling or modulating, for example, gene expression, DNA metabolism, DNA replication, DNA repair, recombination, and chromatin structure and function in host cells. Additionally, the methods can be used to promote cell death particularly in an inducible or tissue-preferred manner.

The MSI-like proteins of the invention additionally find use in manipulating these processes in any host cell, particularly plant cells. Thus, transformed plants, plant cells, and seeds, as well as methods for making such plants, plant cells and seeds are provided. It is recognized that a variety of promoters will be useful in the invention, the choice of which will depend in part upon the desired level of expression of the disclosed nucleotide sequences. As such, the levels of expression can be controlled to modulate the disease resistance pathway resulting in levels of immunity in the plant, which impart resistance in the plant to the pathogen or to induce cell death.

DETAILED DESCRIPTION OF THE INVENTION

The compositions of the invention comprise two maize MSI-like nucleotide and amino acid sequences. Particularly, the nucleotide and amino acid sequence for two maize MSI-like proteins, ZmMSIa and ZmMSIb, are provided and set forth

in SEQ ID NO:1 and SEQ ID NO:3, respectively. The present invention further provides for isolated nucleic acid molecules comprising nucleotide sequences encoding the amino acid sequences shown in SEQ ID NO:2 and SEQ ID NO:4, or the nucleotide sequences encoding the DNA sequences deposited in a bacterial host as Patent Deposit No. PTA-317. Further provided are polypeptides having an amino acid sequence encoded by a nucleic acid molecule described herein, for example those set forth in SEQ ID NO:1 and SEQ ID NO:3, those deposited as Patent Deposit No. PTA-317, and fragments and variants thereof.

Plasmids containing the nucleotide sequences of the invention were deposited in a bacterial host with the Patent Depository of the American Type Culture Collection (ATCC), Manassas, Virginia, on July 9, 1999, and assigned Patent Deposit No. PTA-317. This deposit will be maintained under the terms of the Budapest Treaty on the international recognition of the deposit of microorganisms for the purposes of patent procedure. This deposit was made merely as a convenience for those of skill in the art and is not an admission that a deposit is required under 35 U.S.C. §112.

The MSI-like sequences of the invention are members of a family of molecules having sequence identity to the MSI1 (multicopy suppressor of *ira1*) protein of yeast. Yeast MSI-1 influences the Ras-cAMP signaling pathway and therefore plays a role in controlling cellular growth. Other members of the MSI-like family include, for example, Retinoblastoma-associated proteins (Rb-Ap) from a variety of organisms, including both mammals and plants. The term "family" when referring to the proteins and nucleic acid molecules of the invention is intended to mean two or more proteins or nucleic acid molecules having sufficient amino acid or nucleotide sequence identity as defined herein.

The MSI-like sequences of the invention share sequence identity with other members of the MSI-like family. For instance, the ZmMSIa sequence of the invention has 76% identity from nucleotide 90 to 1321 to the WD-40 repeat protein (MSI1) from *Arabidopsis thaliana* (GenBank Accession no. AF016846) and 76% identity from nucleotide 90 to 1327 to the WD-40 repeat protein (LeMSI1) from *Lycopersicon esculentum* (Accession No. AF016845). ZmMISa also shares sequence identity to the retinoblastoma-binding protein (mRbAp46) from *Mus musculus* (Accession No. U35124) (i.e. 66% identity from nucleotides 722-1318,

60% identity from nucleotide 304-710, and 68% identity from nucleotides 86 to 301); the retinoblastoma binding protein from *Rattus norvegicus* (Accession no. AF090306) (i.e. 65% identity from nucleotides 722-1318, 59% identity from nucleotides 304-710, and 67% from nucleotides 86-301); and the retinoblastoma-binding protein RbAp46 from *Homo sapiens* (Accession No. U35143) (i.e. 65% identity from nucleotides 722-1318, 59% identity from nucleotides 304-710, and 66% identity from 86-301).

The MSI-like sequence, ZmMSIb, of the present invention shares sequence identity, for example, to the WD-40 repeat protein MSI4 from *Arabidopsis thaliana* (Accession no. AF028711) (i.e. 78% identity from nucleotides 194-775, 74% identity from nucleotides 797 to 1273, and 75 % identity from nucleotides 1281 –1540). ZmMSIb also shares sequence identity to the WD-40 repeat protein MSI1 from *Arabidopsis thaliana* (Accession no. AF016846) (i.e. 55% identity from nucleotide 485-631).

MSI family members play a role in the conserved retinoblastoma pathway. Retinoblastoma protein (Rb) is a central part of an evolutionarily conserved regulatory system. Ach *et al.* ((1997) *Plant Cell 9*:1595-1606) showed that the maize retinoblastoma protein, RRB1, shares significant biochemical similarity with the human retinoblastoma protein pRB, including binding of cell-cycle regulators such as D-cyclins and viral proteins, for example, simian virus SV40 (T-antigen) and adenovirus protein E1A. Mutations in conserved regions of the maize protein have similar effects on function with respect to the human protein. Members of the MSI-like family are components of the highly conserved retinoblastoma (Rb) complex. Other components of the complex include Rb, RbAp (Rb-associated protein, known as MSI in yeast) and histone deacetylase (HD). This complex is a key regulator of cell cycle progression. When the complex is localized to its DNA binding site in a regulated promoter, transcription is repressed by a number of mechanisms including chromatin histone deacetylation.

Components of the retinoblastoma complex, including the Rb-Ap genes to which the present invention bears sequence identity, have been shown to interact with fungal and viral genes of known plant pathogens. In particular, interaction has been demonstrated between histone deacetylase (HD) and the toxin of the maize fungal pathogen *Cochliobolus carbonum*, which inhibits HD activity.

Accordingly, the compositions and methods of the present invention are useful in protecting plants and plant cells against fungal pathogens, viruses, nematodes, insects and the like; also, the compositions can be used in formulations so as to fully implement their antimicrobial activities.

Further, proteins of the present invention are highly conserved with respect to a retinoblastoma-binding protein in Arabidopsis and tomato. These Arabidopsis and tomato genes have been shown to bind both the human and maize retinoblastoma homologs. Thus, the proteins of the present invention are expected to bind retinoblastoma homologs and function to modulate cell cycle, DNA metabolism, and division. Hence, the present invention provides utility in such exemplary applications as modification, control, or modulation of various aspects of the cell cycle, cell division, gene expression, DNA metabolism, DNA replication, DNA repair, and recombination as well as cell, organ, or whole plant differentiation and/or disease resistance.

Native retinoblastoma proteins in maize have been shown to localize strictly to the nucleus in maize. Sequence comparison of retinoblastoma proteins from Arabidopsis and maize show conserved regions including WD-40 repeats, A and B binding pockets, N- and C-terminal domains, conserved potential phosphorylation sites, and cysteine residues (see Ach *et al.* (1997) *Plant Cell 9*:1595-1606; Ach *et al.* (1997) *Mol. Cell. Biol. 17*:5077-5086). WD-40 repeats have been described in Marszalek *et al.* (1999) *J. Cell. Biol. 145*:469-479, Gonzalez-Aseguinolaza *et al.* (1999) *Eur. J. Biochem. 259*:909-916, Mathias *et al.* (1999) *Mol. Cell. Biol. 19*:1759-1767, Kamura *et al.* (1998) *Genes Dev. 12*:3872-3881, Torii *et al.* (1998) *EMBO J. 17*:5577-5587, and Iwasaki *et al.* (1995) *Plant Cell. Physiol. 36*:505-510. As the proteins of the present invention may interact with retinoblastoma proteins from various organisms, the present invention may provide a mechanism for modulating the functions performed by any or all of these conserved motifs, which may perform their roles in the native protein or as fragments spliced together to form a component of a designed protein. One skilled in the art will recognize that such proteins may be used alone or in conjunction with other proteins or methods to modulate cellular function. Further, the genes to which the present invention bears homology have been shown to bind to tomato

golden mosaic virus AL1 protein, which induces gene transcription in differentiated cells.

The present invention provides compositions and methods for modulating the total levels of proteins of the present invention and/or altering their ratios in host cells, particularly plant cells. The methods also comprise modulating the activity of the MSI-like proteins of the invention. The functional or biological activity of the MSI-like proteins refers to an activity exerted by the MSI-like polypeptide or nucleic acid sequence on a host cell. Such activities include, for example, modulation of cell cycle, differentiation, gene expression, DNA metabolism, DNA replication, DNA repair, recombination, and chromatin structure.

The invention encompasses isolated or substantially purified nucleic acid or protein compositions. An "isolated" or "purified" nucleic acid molecule or protein, or biologically active portion thereof, is substantially free of other cellular material, or culture medium when produced by recombinant techniques, or substantially free of chemical precursors or other chemicals when chemically synthesized. Preferably, an "isolated" nucleic acid is free of sequences that naturally flank the nucleic acid (i.e., sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated nucleic acid molecule can contain less than about 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0.5 kb, or 0.1 kb of nucleotide sequences that naturally flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived (and also referred to as "flanking DNA"). A protein that is substantially free of cellular material includes preparations of protein having less than about 30%, 20%, 10%, or 5% (by dry weight) of contaminating protein. When the protein of the invention or biologically active portion thereof is recombinantly produced, preferably, culture medium represents less than about 30%, 20%, 10%, or 5% (by dry weight) of chemical precursors or non-protein of interest chemicals.

Fragments and variants of the disclosed nucleotide sequences and proteins encoded thereby are also encompassed by the present invention. By "fragment" is intended a portion of the nucleotide sequence or a portion of the amino acid sequence and hence protein encoded thereby. Fragments of a nucleotide sequence

may encode protein fragments that retain the biological activity of the native protein and hence modulate cell cycle, cell division, gene expression, DNA metabolism, DNA replication, DNA repair, and recombination as well as cell, organ, or whole plant differentiation and/or disease resistance. Alternatively, fragments of a nucleotide sequence that are useful as hybridization probes generally do not encode fragment proteins retaining biological activity. Thus, fragments of a nucleotide sequence may range from at least about 30, 50, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1,000, 1,100, 1,200, 1,300, 1,400, or 1,500 nucleotides, or up to the full-length nucleotide sequence encoding the proteins of the invention.

A fragment of an MSI-like nucleotide sequence that encodes a biologically active portion of an MSI-like protein of the invention will encode at least 30, 50, 100, 150, 200, 250, 300, 350, or 400 contiguous amino acids, or up to the total number of amino acids present in a full-length MSI-like protein of the invention (for example, 431 total amino acids for SEQ ID NO:2 or 453 total amino acids for SEQ ID NO:4). Fragments of an MSI-like nucleotide sequence that are useful as hybridization probes for PCR primers generally need not encode a biologically active portion of an MSI-like protein.

Thus, a fragment of an MSI-like nucleotide sequence may encode a biologically active portion of an MSI-like protein, or it may be a fragment that can be used as a hybridization probe or PCR primer using methods disclosed below. A biologically active portion of an MSI-like protein can be prepared by isolating a portion of one of the MSI-like nucleotide sequences of the invention, expressing the encoded portion of the MSI-like protein (e.g., by recombinant expression *in vitro*), and assessing the activity of the encoded portion of the MSI-like protein. Nucleic acid molecules that are fragments of an MSI-like nucleotide sequence comprise at least 30, 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 800, 900, 1,000, 1,100, 1,200, 1,300, 1,400 or 1,500 nucleotides, or up to the number of nucleotides present in a full-length MSI-like nucleotide sequence disclosed herein (for example, 1584 nucleotides for SEQ ID NO:1 and 1777 nucleotides for SEQ ID NO:3).

By "variants" is intended substantially similar sequences. For nucleotide sequences, conservative variants include those sequences that, because of the

degeneracy of the genetic code, encode the amino acid sequence of one of the MSI-like polypeptides of the invention. Naturally occurring allelic variants such as these can be identified with the use of well-known molecular biology techniques, as, for example, with polymerase chain reaction (PCR) and hybridization techniques as outlined below. Variant nucleotide sequences also include synthetically derived nucleotide sequences, such as those generated, for example, by using site-directed mutagenesis but which still encode a MSI-like protein of the invention. Generally, variants of a particular nucleotide sequence of the invention will have at least about 40%, 50%, 60%, 65%, 70%, generally at least about 75%, 80%, 85%, preferably at least about 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, and more preferably at least about 98%, 99% or more sequence identity to that particular nucleotide sequence as determined by sequence alignment programs described elsewhere herein using default parameters.

By "variant" protein is intended a protein derived from the native protein by deletion (so-called truncation) or addition of one or more amino acids to the N-terminal and/or C-terminal end of the native protein; deletion or addition of one or more amino acids at one or more sites in the native protein; or substitution of one or more amino acids at one or more sites in the native protein. Variant proteins encompassed by the present invention are biologically active, that is they continue to possess the desired biological activity of the native protein, that is, cell cycle control and DNA metabolism activity as described herein. Such variants may result from, for example, genetic polymorphism or from human manipulation. Biologically active variants of a native MSI-like protein of the invention will have at least about 40%, 50%, 60%, 65%, 70%, generally at least about 75%, 80%, 85%, preferably at least about 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, and more preferably at least about 98%, 99% or more sequence identity to the amino acid sequence for the native protein as determined by sequence alignment programs described elsewhere herein using default parameters. A biologically active variant of a protein of the invention may differ from that protein by as few as 1-15 amino acid residues, as few as 1-10, such as 6-10, as few as 5, as few as 4, 3, 2, or even 1 amino acid residue.

The proteins of the invention may be altered in various ways including amino acid substitutions, deletions, truncations, and insertions. Methods for such

manipulations are generally known in the art. For example, amino acid sequence variants of the MSI-like proteins can be prepared by mutations in the DNA. Methods for mutagenesis and nucleotide sequence alterations are well known in the art. See, for example, Kunkel (1985) *Proc. Natl. Acad. Sci. USA 82*:488-492; Kunkel *et al.* (1987) *Methods in Enzymol. 154*:367-382; US Patent No. 4,873,192; Walker and Gaastra, eds. (1983) *Techniques in Molecular Biology* (MacMillan Publishing Company, New York) and the references cited therein. Guidance as to appropriate amino acid substitutions that do not affect biological activity of the protein of interest may be found in the model of Dayhoff *et al.* (1978) *Atlas of Protein Sequence and Structure* (Natl. Biomed. Res. Found., Washington, D.C.), herein incorporated by reference. Conservative substitutions, such as exchanging one amino acid with another having similar properties, may be preferred.

Thus, the genes and nucleotide sequences of the invention include both the naturally occurring sequences as well as mutant forms. Likewise, the proteins of the invention encompass both naturally occurring proteins as well as variations and modified forms thereof. Such variants will continue to possess the desired activity (i.e., modulating cell cycle regulation, modulating gene activity, modulating disease resistance). Assays for these various cellular events are known in the art and are discussed in more detail below. Obviously, the mutations that will be made in the DNA encoding the variant must not place the sequence out of reading frame and preferably will not create complementary regions that could produce secondary mRNA structure. See, EP Patent Application Publication No. 75,444.

The deletions, insertions, and substitutions of the protein sequences encompassed herein are not expected to produce radical changes in the characteristics of the protein. However, when it is difficult to predict the exact effect of the substitution, deletion, or insertion in advance of doing so, one skilled in the art will appreciate that the effect will be evaluated by routine screening assays. That is, the activity can be evaluated by assays including suppression of yeast Ras signaling pathway mutations. For example, the IRA1 gene product negatively regulates the Ras-cAMP pathway in yeast. Over expression of yeast MSI-1, and other MSI-like family members, suppress the *ira1* mutant phenotype. In addition, the $ras2^{val19}$ yeast mutant, which has a reduced intrinsic GTPase activity, is also suppressed by over expression of several of the MSI-like family

members. Therefore the activity of the sequences of the invention may be assayed by transforming *ira1* or *ras2*^{val19} yeast strains and assaying for suppression of the mutant phenotype. See, for example, Ach *et al.* (1997) *The Plant Cell 9*:1595-1606, herein incorporated by reference. Further, in light of the role of MSI proteins in chromatin remodeling, assays to detect alteration of chromatin or histones may be used to evaluate activity. Alternatively, assays measuring the modulations in cell cycle regulation and the plant/pathogen defense system may be used to assay for the activity of the MSI-like sequences of the invention.

Variant nucleotide sequences and proteins also encompass sequences and proteins derived from a mutagenic and recombinogenic procedure such as DNA shuffling. With such a procedure, one or more different MSI coding sequences can be manipulated to create a new MSI sequence or protein possessing the desired properties. In this manner, libraries of recombinant polynucleotides are generated from a population of related sequence polynucleotides comprising sequence regions that have substantial sequence identity and can be homologously recombined *in vitro* or *in vivo*. For example, using this approach, sequence motifs encoding a domain of interest may be shuffled between the MSI gene of the invention and other known retinoblastoma-associated protein genes to obtain a new gene coding for a protein with an improved property of interest, such as an increased K_m in the case of an enzyme. Strategies for such DNA shuffling are known in the art. See, for example, Stemmer (1994) *Proc. Natl. Acad. Sci. USA 91*:10747-10751; Stemmer (1994) *Nature 370*:389-391; Crameri *et al.* (1997) *Nature Biotech. 15*:436-438; Moore *et al.* (1997) *J. Mol. Biol. 272*:336-347; Zhang *et al.* (1997) *Proc. Natl. Acad. Sci. USA 94*:4504-4509; Crameri *et al.* (1998) *Nature 391*:288-291; and U.S. Patent Nos. 5,605,793 and 5,837,458.

In this manner, the present invention encompasses the proteins as well as components, fragments, and variants thereof. That is, it is recognized that component polypeptides or fragments of the proteins may be produced which retain biological activity that modifies, modulates, or controls cell processes or development in a plant or plant cell as disclosed herein. These fragments include truncated sequences, as well as N-terminal, C-terminal, internal and internally deleted amino acid sequences of the proteins.

The nucleotide sequences of the invention can be used to isolate corresponding sequences from other organisms, particularly other plants, more particularly other monocots. In this manner, methods such as PCR, hybridization, and the like can be used to identify such sequences based on their sequence homology to the sequences set forth herein. Sequences isolated based on their sequence identity to the entire MSI-like sequences set forth herein or to fragments thereof are encompassed by the present invention. Such sequences include sequences that are of the disclosed sequences. By "orthologs" is intended genes derived from a common ancestral gene and which are found in different species as a result of speciation. Genes found in different species are considered orthologs when their nucleotide sequences and/or their encoded protein sequences have a high percentage of sequence identity and/or similarity. Functions of orthologs are often highly conserved among species.

In a PCR approach, oligonucleotide primers can be designed for use in PCR reactions to amplify corresponding DNA sequences from cDNA or genomic DNA extracted from any plant or organism of interest. Methods for designing PCR primers and PCR cloning are generally known in the art and are disclosed in Sambrook *et al.* (1989) *Molecular Cloning: A Laboratory Manual* (2d ed., Cold Spring Harbor Laboratory Press, Plainview, New York). See also Innis *et al.*, eds. (1990) *PCR Protocols: A Guide to Methods and Applications* (Academic Press, New York); Innis and Gelfand, eds. (1995) *PCR Strategies* (Academic Press, New York); and Innis and Gelfand, eds. (1999) *PCR Methods Manual* (Academic Press, New York). Known methods of PCR include, but are not limited to, methods using paired primers, nested primers, single specific primers, degenerate primers, gene-specific primers, vector-specific primers, partially-mismatched primers, and the like.

In hybridization techniques, all or part of a known nucleotide sequence is used as a probe that selectively hybridizes to other corresponding nucleotide sequences present in a population of cloned genomic DNA fragments or cDNA fragments (*i.e.*, genomic or cDNA libraries) from a chosen organism. The hybridization probes may be genomic DNA fragments, cDNA fragments, RNA fragments, or other oligonucleotides, and may be labeled with a detectable group such as ^{32}P, or any other detectable marker. Thus, for example, probes for

hybridization can be made by labeling synthetic oligonucleotides based on the sequence of the invention. Methods for preparation of probes for hybridization and for construction of cDNA and genomic libraries are generally known in the art and are disclosed in Sambrook *et al.* (1989) *Molecular Cloning: A Laboratory Manual* (2d ed., Cold Spring Harbor Laboratory Press, Plainview, New York).

For example, the entire sequence disclosed herein, or one or more portions thereof, may be used as a probe capable of specifically hybridizing to corresponding sequences and messenger RNAs. To achieve specific hybridization under a variety of conditions, such probes include sequences that are unique among MSI sequences and are preferably at least about 10 nucleotides in length, and most preferably at least about 20 nucleotides in length. Such probes may be used to amplify corresponding MSI-like sequences from a chosen organism by PCR. This technique may be used to isolate additional coding sequences from a desired organism or as a diagnostic assay to determine the presence of coding sequences in an organism. Hybridization techniques include hybridization screening of plated DNA libraries (either plaques or colonies; see, for example, Sambrook *et al.* (1989) *Molecular Cloning: A Laboratory Manual* (2d ed., Cold Spring Harbor Laboratory Press, Plainview, New York).

Hybridization of such sequences may be carried out under stringent conditions. By "stringent conditions" or "stringent hybridization conditions" is intended conditions under which a probe will hybridize to its target sequence to a detectably greater degree than to other sequences (e.g., at least 2-fold over background). Stringent conditions are sequence-dependent and will be different in different circumstances. By controlling the stringency of the hybridization and/or washing conditions, target sequences that are 100% complementary to the probe can be identified (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 than about 1000 nucleotides in length, preferably 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.,

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 stringency conditions include hybridization in 40 to 45% formamide, 1.0 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. Duration of hybridization is generally less than about 24 hours, usually about 4 to about 12 hours.

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 (1984) *Anal. Biochem. 138*:267-284: $T_m = 81.5°C + 16.6$ (log M) + 0.41 (%GC) - 0.61 (% form) - 500/L; where M is the molarity of monovalent cations, %GC is the percentage of guanosine and cytosine nucleotides in the DNA, % 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 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 ≥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 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°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 concentration so that a higher temperature can be used. An extensive guide to the hybridization of nucleic acids is found in Tijssen (1993) *Laboratory Techniques in Biochemistry and Molecular Biology—Hybridization with Nucleic Acid Probes*, Part I, Chapter 2 (Elsevier, New York); and Ausubel *et al.*, eds. (1995) *Current Protocols in Molecular Biology*, Chapter 2 (Greene Publishing and Wiley-Interscience, New York). See Sambrook *et al.* (1989) *Molecular Cloning: A Laboratory Manual* (2d ed., Cold Spring Harbor Laboratory Press, Plainview, New York).

Thus, isolated sequences that encode an MSI-like protein and which hybridize under stringent conditions to the MSI-like nucleotide sequences disclosed herein, or to fragments thereof, are encompassed by the present invention. Such sequences will be at least about 40% to 50% homologous, about 60%, 65%, or 70% homologous, and even at least about 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more homologous with the disclosed sequences. That is, the sequence identity of sequences may range, sharing at least about 40% to 50%, about 60%, 65%, or 70%, and even at least about 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more sequence identity.

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

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

(b) As used herein, "comparison window" makes reference to a contiguous and specified segment of a polynucleotide sequence, wherein the polynucleotide sequence in the comparison window may comprise additions or deletions (i.e., gaps) compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. Generally, the

comparison window is at least 20 contiguous nucleotides in length, and optionally can be 30, 40, 50, 100, or longer. Those of skill in the art understand that to avoid a high similarity to a reference sequence due to inclusion of gaps in the polynucleotide sequence a gap penalty is typically introduced and is subtracted from the number of matches.

Methods of alignment of sequences for comparison are well known in the art. Thus, the determination of percent sequence identity between any two sequences can be accomplished using a mathematical algorithm. Non-limiting examples of such mathematical algorithms are the algorithm of Myers and Miller (1988) *CABIOS 4*:11-17; the local homology algorithm of Smith *et al.* (1981) *Adv. Appl. Math. 2*:482; the homology alignment algorithm of Needleman and Wunsch (1970) *J. Mol. Biol. 48*:443-453; the search-for-similarity-method of Pearson and Lipman (1988) *Proc. Natl. Acad. Sci. 85*:2444-2448; the algorithm of Karlin and Altschul (1990) *Proc. Natl. Acad. Sci. USA 87*:2264, modified as in Karlin and Altschul (1993) *Proc. Natl. Acad. Sci. USA 90*:5873-5877.

Computer implementations of these mathematical algorithms can be utilized for comparison of sequences to determine sequence identity. Such implementations include, but are not limited to: CLUSTAL in the PC/Gene program (available from Intelligenetics, Mountain View, California); the ALIGN program (Version 2.0) and GAP, BESTFIT, BLAST, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Version 8 (available from Genetics Computer Group (GCG), 575 Science Drive, Madison, Wisconsin, USA). Alignments using these programs can be performed using the default parameters. The CLUSTAL program is well described by Higgins *et al.* (1988) *Gene 73*:237-244 (1988); Higgins *et al.* (1989) *CABIOS 5*:151-153; Corpet *et al.* (1988) *Nucleic Acids Res. 16*:10881-90; Huang *et al.* (1992) *CABIOS 8*:155-65; and Pearson *et al.* (1994) *Meth. Mol. Biol. 24*:307-331. The ALIGN program is based on the algorithm of Myers and Miller (1988) *supra*. A PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4 can be used with the ALIGN program when comparing amino acid sequences. The BLAST programs of Altschul *et al* (1990) *J. Mol. Biol. 215*:403 are based on the algorithm of Karlin and Altschul (1990) *supra*. BLAST nucleotide searches can be performed with the BLASTN program, score = 100, wordlength = 12, to obtain nucleotide sequences

homologous to a nucleotide sequence encoding a protein of the invention. BLAST protein searches can be performed with the BLASTX program, score = 50, wordlength = 3, to obtain amino acid sequences homologous to a protein or polypeptide of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST (in BLAST 2.0) can be utilized as described in Altschul *et al.* (1997) *Nucleic Acids Res. 25*:3389. Alternatively, PSI-BLAST (in BLAST 2.0) can be used to perform an iterated search that detects distant relationships between molecules. See Altschul *et al.* (1997) *supra.* When utilizing BLAST, Gapped BLAST, PSI-BLAST, the default parameters of the respective programs (e.g., BLASTN for nucleotide sequences, BLASTX for proteins) can be used. See http://www.ncbi.nlm.nih.gov. Alignment may also be performed manually by inspection.

(c) As used herein, "sequence identity" or "identity" in the context of two nucleic acid or polypeptide sequences makes reference to the residues in the two sequences that are the same when aligned for maximum correspondence over a specified comparison window. When percentage of sequence identity is used in reference to proteins it is recognized that residue positions which are not identical often differ by conservative amino acid substitutions, where amino acid residues are substituted for other amino acid residues with similar chemical properties (e.g., charge or hydrophobicity) and therefore do not change the functional properties of the molecule. When sequences differ in conservative substitutions, the percent sequence identity may be adjusted upwards to correct for the conservative nature of the substitution. Sequences that differ by such conservative substitutions are said to have "sequence similarity" or "similarity". Means for making this adjustment are well known to those of skill in the art. Typically this involves scoring a conservative substitution as a partial rather than a full mismatch, thereby increasing the percentage sequence identity. Thus, for example, where an identical amino acid is given a score of 1 and a non-conservative substitution is given a score of zero, a conservative substitution is given a score between zero and 1. The scoring of conservative substitutions is calculated, e.g., as implemented in the program PC/GENE (Intelligenetics, Mountain View, California).

(d) As used herein, "percentage of sequence identity" means the value determined by comparing two optimally aligned sequences over a comparison

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

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

Another indication that nucleotide sequences are substantially identical is if two molecules hybridize to each other under stringent conditions. Generally, stringent conditions are selected to be about 5°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. However, stringent conditions encompass temperatures in the range of about 1°C to about 20°C, depending upon the desired degree of stringency as otherwise qualified herein. Nucleic acids that do not hybridize to each other under stringent conditions are still substantially identical if the polypeptides they encode are substantially identical. This may occur, e.g., when a copy of a nucleic acid is created using the maximum codon degeneracy permitted by the genetic code. One indication that two nucleic acid sequences are substantially identical is when the polypeptide encoded by the first nucleic acid is immunologically cross reactive with the polypeptide encoded by the second nucleic acid.

(e)(ii) The term "substantial identity" in the context of a peptide indicates that a peptide comprises a sequence with at least 70% sequence identity to a reference sequence, preferably 80%, more preferably 85%, most preferably at least 90% or 95% sequence identity to the reference sequence over a specified comparison window. Preferably, optimal alignment is conducted using the homology alignment algorithm of Needleman *et al.* (1970) *J. Mol. Biol. 48*:443. An indication that two peptide sequences are substantially identical is that one peptide is immunologically reactive with antibodies raised against the second peptide. Thus, a peptide is substantially identical to a second peptide, for example, where the two peptides differ only by a conservative substitution. Peptides that are "substantially similar" share sequences as noted above except that residue positions that are not identical may differ by conservative amino acid changes. Peptide alignment and homology can be determined with the CLUSTAL W program using the default values for relevant parameters: gap opening penalty 10.00, gap extension penalty 0.05, delay divergent sequences 40%, residue-specific and hydrophilic penalties ON, hydrophilic = GPSNDQEKR, gap separation distance 8.

The nucleotide sequences of the invention are provided in expression cassettes for expression in a host cell of interest, particularly a plant cell. The cassette will include 5' and 3' regulatory sequences operably linked to a sequence of the invention. By "operably linked" is intended a functional 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. Generally, operably linked means that the nucleic acid sequences being linked are contiguous and, where necessary to join two protein coding regions, contiguous and in the same reading frame. The cassette may additionally contain at least one additional gene to be cotransformed into the organism. Alternatively, the additional gene(s) can be provided on multiple expression cassettes. Such an expression cassette is provided with a plurality of restriction sites for insertion of the MSI-like sequence to be under the transcriptional regulation of the regulatory regions. The expression cassette may additionally contain selectable marker genes.

The expression cassette will include in the 5'-3' direction of transcription, a transcriptional and translational initiation region, a nucleotide sequence of the

invention, and a transcriptional and translational termination region functional in plants. The transcriptional initiation region, the promoter, may be native or analogous or foreign or heterologous to the host cell. Additionally, the promoter may be the natural sequence or alternatively a synthetic sequence. By "foreign" is intended that the transcriptional initiation region is not found in the native plant into which the transcriptional initiation region is introduced. As used herein, a chimeric gene comprises a coding sequence operably linked to a transcription initiation region that is heterologous to the coding sequence. By "heterologous" is intended the transcription initiation region is not native to the coding sequence.

While it may be preferable to express the sequences using heterologous promoters, the native promoter sequences may be used. When such a construct is used to transform a plant, the expression levels of MSI-like sequences in the plant or plant cell would change. Thus, the phenotype of the plant or plant cell is altered.

The termination region may be native with the transcriptional initiation region, may be native with the operably linked DNA sequence of interest, or may be derived from another source. Convenient termination regions are available from the Ti-plasmid of *A. tumefaciens*, such as the octopine synthase and nopaline synthase termination regions. See also Guerineau *et al.* (1991) *Mol. Gen. Genet. 262*:141-144; Proudfoot (1991) *Cell 64*:671-674; Sanfacon *et al.* (1991) *Genes Dev. 5*:141-149; Mogen *et al.* (1990) *Plant Cell 2*:1261-1272; Munroe *et al.* (1990) *Gene 91*:151-158; Ballas *et al.* (1989) *Nucleic Acids Res. 17*:7891-7903; and Joshi *et al.* (1987) *Nucleic Acid Res. 15*:9627-9639.

Where appropriate, the gene(s) may be optimized for increased expression in the transformed plant. That is, the genes can be synthesized using plant-preferred codons for improved expression. Methods are available in the art for synthesizing plant-preferred genes. See, for example, U.S. Patent Nos. 5,380,831, and 5,436,391, and Murray *et al.* (1989) *Nucleic Acids Res. 17*:477-498, herein incorporated by reference.

Additional sequence modifications are known to enhance gene expression in a cellular host. These include elimination of sequences encoding spurious polyadenylation signals, exon-intron splice site signals, transposon-like repeats, and other such well-characterized sequences that may be deleterious to gene

expression. The G-C content of the sequence may be adjusted to levels average for a given cellular host, as calculated by reference to known genes expressed in the host cell. When possible, the sequence is modified to avoid predicted hairpin secondary mRNA structures.

The expression cassettes may additionally contain 5' leader sequences in the expression cassette construct. Such leader sequences can act to enhance translation. Translation leaders are known in the art and include: picornavirus leaders, for example, EMCV leader (Encephalomyocarditis 5' noncoding region) (Elroy-Stein *et al.* (1989) *PNAS USA 86*:6126-6130); potyvirus leaders, for example, TEV leader (Tobacco Etch Virus) (Allison *et al.* (1986); MDMV leader (Maize Dwarf Mosaic Virus); *Virology 154*:9-20), and human immunoglobulin heavy-chain binding protein (BiP), (Macejak *et al.* (1991) *Nature 353*:90-94); untranslated leader from the coat protein mRNA of alfalfa mosaic virus (AMV RNA 4) (Jobling *et al.* (1987) *Nature 325*:622-625); tobacco mosaic virus leader (TMV) (Gallie *et al.* (1989) in *Molecular Biology of RNA*, ed. Cech (Liss, New York), pp. 237-256); and maize chlorotic mottle virus leader (MCMV) (Lommel *et al.* (1991) *Virology 81*:382-385). See also, Della-Cioppa *et al.* (1987) *Plant Physiol. 84*:965-968. Other methods known to enhance translation can also be utilized, for example, introns, and the like.

In preparing the expression cassette, the various DNA fragments may be manipulated, so as to provide for the DNA sequences in the proper orientation and, as appropriate, in the proper reading frame. Toward this end, adapters or linkers may be employed to join the DNA fragments or other manipulations may be involved to provide for convenient restriction sites, removal of superfluous DNA, removal of restriction sites, or the like. For this purpose, *in vitro* mutagenesis, primer repair, restriction, annealing, resubstitutions, e.g., transitions and transversions, may be involved.

A number of promoters can be used in the practice of the invention. The promoters can be selected based on the desired outcome. The nucleic acids can be combined with constitutive, tissue-preferred, or other promoters for expression in plants.

Such constitutive promoters include, for example, the core promoter of the Rsyn7 promoter and other constitutive promoters disclosed in WO 99/43838; the

core CaMV 35S promoter (Odell *et al.* (1985) *Nature 313*:810-812); rice actin (McElroy *et al.* (1990) *Plant Cell 2*:163-171); ubiquitin (Christensen *et al.* (1989) *Plant Mol. Biol. 12*:619-632 and Christensen *et al.* (1992) *Plant Mol. Biol. 18*:675-689); pEMU (Last *et al.* (1991) *Theor. Appl. Genet. 81*:581-588); MAS (Velten *et al.* (1984) *EMBO J. 3*:2723-2730); ALS promoter (U.S. Patent No. 5,659,026), and the like. Other constitutive promoters include, for example, U.S. Patent Nos. 5,608,149; 5,608,144; 5,604,121; 5,569,597; 5,466,785; 5,399,680; 5,268,463; and 5,608,142.

Generally, it will be beneficial to express the gene from an inducible promoter, particularly from a pathogen-inducible 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 the copending applications entitled "Inducible Maize Promoters", U.S. Application Serial No. 09/257,583, filed February 25, 1999, which is herein 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. 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 invention. Such wound-inducible promoters include potato proteinase inhibitor (pin II) gene (Ryan (1990) *Ann. Rev. Phytopath. 28*:425-449; Duan *et al.* (1996)

Nature Biotechnology 14:494-498); wun1 and wun2, US 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 Letters 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 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 In2-2 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 MSI expression 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-specific 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.

Root-specific 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 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.

"Seed-preferred" promoters include both "seed-specific" promoters (those promoters active during seed development such as promoters of seed storage proteins) as well as "seed-germinating" promoters (those promoters active during seed germination). See Thompson *et al.* (1989) *BioEssays 10*:108, herein incorporated by reference. Such seed-preferred promoters include, but are not limited to, Cim1 (cytokinin-induced message); cZ19B1 (maize 19 kDa zein); and celA (cellulose synthase). Gama-zein is a preferred endosperm-specific promoter. Glob-1 is a preferred embryo-specific promoter. For dicots, seed-specific promoters include, but are not limited to, bean β-phaseolin, napin, β-conglycinin, soybean lectin, cruciferin, and the like. For monocots, seed-specific promoters include, but are not limited to, maize 15 kDa zein, 22 kDa zein, 27 kDa zein, g-zein, waxy, shrunken 1, shrunken 2, globulin 1, etc.

Generally, the expression cassette will comprise a selectable marker gene for the selection of transformed cells. Selectable marker genes are utilized for the selection of transformed cells or tissues. Marker genes include genes encoding antibiotic resistance, such as those encoding neomycin phosphotransferase II (NEO) and hygromycin phosphotransferase (HPT), as well as genes conferring resistance to herbicidal compounds, such as glufosinate ammonium, bromoxynil, imidazolinones, and 2,4-dichlorophenoxyacetate (2,4-D). See generally, Yarranton (1992) *Curr. Opin. Biotech. 3*:506-511; Christopherson *et al.* (1992) *Proc. Natl. Acad. Sci. USA 89*:6314-6318; Yao *et al.* (1992) *Cell 71*:63-72; Reznikoff (1992) *Mol. Microbiol. 6*:2419-2422; Barkley *et al.* (1980) in *The Operon*, pp. 177-220; Hu *et al.* (1987) *Cell 48*:555-566; Brown *et al.* (1987) *Cell 49*:603-612; Figge *et al.* (1988) *Cell 52*:713-722; Deuschle *et al.* (1989) *Proc. Natl. Acad. Aci. USA 86*:5400-5404; Fuerst *et al.* (1989) *Proc. Natl. Acad. Sci. USA 86*:2549-2553;

Deuschle *et al.* (1990) *Science 248*:480-483; Gossen (1993) Ph.D. Thesis, University of Heidelberg; Reines *et al.* (1993) *Proc. Natl. Acad. Sci. USA 90*:1917-1921; Labow *et al.* (1990) *Mol. Cell. Biol. 10*:3343-3356; Zambretti *et al.* (1992) *Proc. Natl. Acad. Sci. USA 89*:3952-3956; Baim *et al.* (1991) *Proc. Natl. Acad. Sci. USA 88*:5072-5076; Wyborski *et al.* (1991) *Nucleic Acids Res. 19*:4647-4653; Hillenand-Wissman (1989) *Topics Mol. Struc. Biol. 10*:143-162; Degenkolb *et al.* (1991) *Antimicrob. Agents Chemother. 35*:1591-1595; Kleinschnidt *et al.* (1988) *Biochemistry 27*:1094-1104; Bonin (1993) Ph.D. Thesis, University of Heidelberg; Gossen *et al.* (1992) *Proc. Natl. Acad. Sci. USA 89*:5547-5551; Oliva *et al.* (1992) *Antimicrob. Agents Chemother. 36*:913-919; Hlavka *et al.* (1985) *Handbook of Experimental Pharmacology*, Vol. 78 (Springer-Verlag, Berlin); Gill *et al.* (1988) *Nature 334*:721-724. Such disclosures are herein incorporated by reference. The above list of selectable marker genes is not meant to be limiting. Any selectable marker gene can be used in the present invention.

It is expected that those of skill in the art are knowledgeable in the numerous expression systems available for expression of a nucleic acid encoding a protein of the present invention. No attempt to describe in detail the various methods known for the expression of proteins in prokaryotes or eukaryotes will be made.

The nucleic acids of the present invention can also be employed for use in sense or antisense suppression of one or more genes of the invention in a host cell, tissue, or plant. In this embodiment nucleotide sequences, antisense constructions, complementary to at least a portion of the messenger RNA (mRNA) for the MSI-like sequences can be constructed. Antisense nucleotides are constructed to hybridize with the corresponding mRNA. Modifications of the antisense sequences may be made as long as the sequences hybridize to and interfere with expression of the corresponding mRNA. In this manner, antisense constructions having 70%, preferably 80%, more preferably 85% sequence identity to the corresponding antisensed sequences may be used. Furthermore, portions of the antisense nucleotides may be used to disrupt the expression of the target gene. Generally, sequences of at least 50 nucleotides, 100 nucleotides, 200 nucleotides, or greater may be used.

The nucleotide sequences of the present invention may also be used in the sense orientation to suppress the expression of endogenous genes in plants. Methods for suppressing gene expression in plants using nucleotide sequences in the sense orientation are known in the art. The methods generally involve transforming plants with a DNA construct comprising a promoter that drives expression in a plant operably linked to at least a portion of a nucleotide sequence that corresponds to the transcript of the endogenous gene. Typically, such a nucleotide sequence has substantial sequence identity to the sequence of the transcript of the endogenous gene, preferably greater than about 65% sequence identity, more preferably greater than about 85% sequence identity, most preferably greater than about 95% sequence identity. See, U.S. Patent Nos. 5,283,184 and 5,034,323; herein incorporated by reference.

The sequences of the invention can be used to transform any host cell, particularly plant cells. It is recognized that one skilled in the art is knowledgeable in the numerous methods available for both transient and stable expression of sequences in both prokaryotic and eukaryotic host cells. Transformation protocols as well as protocols for introducing nucleotide 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 nucleotide sequences into plant cells and subsequent insertion into the plant genome include microinjection (Crossway *et al.* (1986) *Biotechniques 4*:320-334), electroporation (Riggs *et al.* (1986) *Proc. Natl. Acad. Sci. USA 83*:5602-5606, *Agrobacterium*-mediated transformation (Townsend *et al.*, U.S. Pat No. 5,563,055; Zhao *et al.*, 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.* (1995) "Direct 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 Physiol. 87*:671-674 (soybean); McCabe *et al.* (1988) *Bio/Technology 6*:923-926 (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; Buising *et al.*, U.S. Patent Nos. 5,322,783 and 5,324,646; Tomes *et al.* (1995) "Direct DNA Transfer into Intact Plant Cells via Microprojectile Bombardment," in *Plant Cell, Tissue, and Organ Culture: Fundamental Methods*, ed. Gamborg (Springer-Verlag, Berlin) (maize); 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; 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*); all of which are herein incorporated by reference.

The cells that have been transformed 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 identified. Two or more generations may be grown to ensure that constitutive expression of the desired phenotypic characteristic is stably maintained and inherited and then seeds harvested to ensure constitutive expression of the desired phenotypic characteristic has been achieved.

The present invention may be used for transformation of any plant species, including, but not limited to, corn (*Zea mays*), *Brassica* sp. (e.g., *B. napus*, *B. rapa*, *B. juncea*), particularly those *Brassica* species useful as sources of seed oil, alfalfa (*Medicago sativa*), rice (*Oryza sativa*), rye (*Secale cereale*), sorghum (*Sorghum bicolor*, *Sorghum vulgare*), millet (e.g., pearl millet (*Pennisetum glaucum*), proso millet (*Panicum miliaceum*), foxtail millet (*Setaria italica*), finger millet (*Eleusine*

coracana)), sunflower (*Helianthus annuus*), safflower (*Carthamus tinctorius*), wheat (*Triticum aestivum*), soybean (*Glycine max*), tobacco (*Nicotiana tabacum*), potato (*Solanum tuberosum*), peanuts (*Arachis hypogaea*), cotton (*Gossypium barbadense*, *Gossypium hirsutum*), sweet potato (*Ipomoea batatus*), cassava (*Manihot esculenta*), coffee (*Coffea* spp.), coconut (*Cocos nucifera*), pineapple (*Ananas comosus*), citrus trees (*Citrus* spp.), cocoa (*Theobroma cacao*), tea (*Camellia sinensis*), banana (*Musa* spp.), avocado (*Persea americana*), fig (*Ficus casica*), guava (*Psidium guajava*), mango (*Mangifera indica*), olive (*Olea europaea*), papaya (*Carica papaya*), cashew (*Anacardium occidentale*), macadamia (*Macadamia integrifolia*), almond (*Prunus amygdalus*), sugar beets (*Beta vulgaris*), sugarcane (*Saccharum* spp.), oats, barley, vegetables, ornamentals, and conifers.

Vegetables include tomatoes (*Lycopersicon esculentum*), lettuce (e.g., *Lactuca sativa*), green beans (*Phaseolus vulgaris*), lima beans (*Phaseolus limensis*), peas (*Lathyrus* spp.), and members of the genus Cucumis such as cucumber (*C. sativus*), cantaloupe (*C. cantalupensis*), and musk melon (*C. melo*). Ornamentals include azalea (*Rhododendron* spp.), hydrangea (*Macrophylla hydrangea*), hibiscus (*Hibiscus rosasanensis*), roses (*Rosa* spp.), tulips (*Tulipa* spp.), daffodils (*Narcissus* spp.), petunias (*Petunia hybrida*), carnation (*Dianthus caryophyllus*), poinsettia (*Euphorbia pulcherrima*), and chrysanthemum. Conifers that may be employed in practicing the present invention include, for example, pines such as loblolly pine (*Pinus taeda*), slash pine (*Pinus elliotii*), ponderosa pine (*Pinus ponderosa*), lodgepole pine (*Pinus contorta*), and Monterey pine (*Pinus radiata*); Douglas-fir (*Pseudotsuga menziesii*); Western hemlock (*Tsuga canadensis*); Sitka spruce (*Picea glauca*); redwood (*Sequoia sempervirens*); true firs such as silver fir (*Abies amabilis*) and balsam fir (*Abies balsamea*); and cedars such as Western red cedar (*Thuja plicata*) and Alaska yellow-cedar (*Chamaecyparis nootkatensis*). Preferably, plants of the present invention are crop plants (for example, corn, alfalfa, sunflower, *Brassica*, soybean, cotton, safflower, peanut, sorghum, wheat, millet, tobacco, etc.), more preferably corn and soybean plants, yet more preferably corn plants.

The isolated nucleic acids of the present invention can also be used for recombinant expression of polypeptides, or for use as immunogens in the

preparation and/or screening of antibodies. Attachment of chemical agents, which bind, intercalate, cleave and/or crosslink to the isolated nucleic acids of the present invention can also be used to modulate transcription or translation. Further, using a primer specific to an insertion sequence (e.g., transposon) and a primer which specifically hybridizes to an isolated nucleic acid of the present invention, one can use nucleic acid amplification to identity insertion sequence inactivated genes of the invention from a cDNA library prepared from insertion sequence mutagenized plants. Progeny seed from the plants comprising the desired inactivated gene can be grown to a plant to study the phenotypic changes characteristic of that inactivation. See, *Tools to Determine the Function of Genes*, 1995 Proceedings of the Fiftieth Annual Corn and Sorghum Industry Research Conference, American Seed Trade Association, Washington, D.C., 1995. Additionally, non-translated 5' or 3' regions of the polynucleotides of the present invention can be used to modulate turnover of heterologous mRNAs and/or protein synthesis. Further, the codon preference characteristic of the polynucleotides of the present invention can be employed in heterologous sequences, or altered in homologous or heterologous sequences, to modulate translational level and/or rates.

In one embodiment, the sequences of the invention find use in modulating gene activity. As described above, the sequences of the invention share sequence identity to human Retinoblastoma-associated proteins, RbAp48 and RbAp46. These proteins contain WD-40 domains and were first reported as retinoblastoma binding proteins. Recently, human RbAp48 and the yeast homolog MSI1 have been found associated with proteins involved in histone metabolism (chromatin assembly factor-I; histone acetyltransferase; histone deacetylase.) Thus, these proteins may regulate gene silencing. For instance, deletion of MSI1 (renamed CAC3) reduces silencing of genes adjacent to telomeric DNA. Hence, the sequences of the invention can be used to modulate gene activity in host cells, particularly plant cells. By "modulating gene activity" is intended the increase or decrease in the activity state of a gene or gene regions. Molecular events that result in the modulation of gene activity include, for example, histone acetylation, heterochromatin formation, and chromatin assembly. Furthermore, "modulating gene activity" also encompasses a general means of preparing a gene for transcription. Thus, the nucleotide sequences of the invention can act to modulate

gene activity in various manners. Increased histone acetylation may enhance the ability of transcription factors to bind to DNA when contained in a nucleosome. In this manner, transformed plants having altered gene activity and enhanced disease resistance can be obtained.

Assays to detect modulations in gene activity are known in the art. For instance, the MSI-like nucleotide sequences of the invention, as well as corresponding antisense sequences, may be used in expression cassettes to transform host cells, particularly plant cells. Generally, the transformation and expression of such sequences in the host cell may lead to an increase or alternatively, a decrease in gene activity. It is further recognized that the constructs of the invention may globally modulate gene activity or alternatively, may target particular regions of the chromosome. Assays such as RNAase protection, Northern Blot, and mRNA profiling are known in the art and may be used to identify either global alterations in gene activity or may be used to identify specific chromosomal loci influenced by the sequences of the invention. Other assays are available for determining activity. See, generally, Lusser *et al.* (1997) *Science 277*:88-91; Rundlett *et al.* (1996) *PNAS 93*:14503-14508; DeRubertis *et al.* (1996) *Nature 384*:589-591; Pazin *et al.* (1997) *Cell 89*:325-328; herein incorporated by reference. See also, Walton *et al.* (1993) *Ann. Rev. Phytopathol. 31*:275-303; Brosch *et al.* (1995) *Plant Cell 7*:1941-1950; Walton *et al.* (1985) *Experientia 41*:348-350; Yoshida *et al.* (1995) *Bioessays 17*:423; Taunton *et al.* (1996) *Science 272*:408-411; Pazin *et al.* (1997) *Cell 89*:325-328; Verreault *et al.* (1996) *Cell 87*:95-104; Kaufman *et al.* (1997) *Genes Dev. 11*:345-357; Parthun *et al.* (1996) *Cell 87*:85-94; Ciuffetti *et al.* (1995) *Physiol. Mol. Pl. Pathol. 46*:61-70; Rasmussen *et al.* (1988) *Physiol. Mol. Pl. Pathol. 32*:283-292; Ciuffetti *et al.* (1983) *Biochem. 22*:3507-3510; Wolf *et al.* (1990) *Plant Sci. 70*:127-137; Ach *et al.* (1997) *Plant Cell 9*:1595-1606. Additionally, function of the MSI sequences can be elucidated by the characterization of mutants isolated by TUSC screening (Benson *et al.* (1995) *Plant Cell 7*:75-84; Mena *et al.* (1996) *Science 274*:1537-1540; U.S. Patent Application No. 08/835,638, which is a continuation of U.S. Patent Application No. 08/262,056).

In another embodiment, the sequences of the invention can be used to modulate the cell cycle. Manipulation of the cell cycle can have effects on cellular

processes including DNA metabolism, DNA replication, DNA repair, and recombination. Accordingly, the compositions and methods of the present invention find use in modulating the cell cycle, DNA replication, DNA repair, and DNA recombination in a host cell or organism.

While the invention is not bound by any particular mechanism of action, the gene products, probably proteins or polypeptides, function to modulate DNA metabolism. It is recognized that the present invention is not dependent upon a particular mechanism of modulation. Rather, the genes and methods of the invention work to alter the cell cycle, DNA metabolism, DNA replication, DNA repair, and recombination independently of how such modulation is achieved. The mechanisms described herein may be used alone or in combination with other proteins or agents to modulate the processes or end results affected.

Assays for monitoring modulations in the cell cycle are known in the art. For example, plant cell suspensions can be synchronized by chemical agents, which arrest the cell cycle by acting on, for example, CDK enzymes or regulators of the cell cycle apparatus. See, for example, Planchaisa *et al.* (2000) *FEBS lett 476*:78-83 herein incorporated by reference. Subsequently, cells synchronized to arrest in mitosis are released from the cell cycle arrest and provided BrdU and an MSI-like amino acid sequence of the invention. The level of BrdU incorporation into the cells is reflective of progression into S phase thereby allowing the effect of the MSI-like sequences on the G-S phase transition to be assayed. *In vitro* growth inhibition or acceleration assays may also be performed on plant cell cultures. In these assays cells are incubated with ^{3}H-thymidine in the presence and absence of the MSI-like sequence of the invention. Cells are harvested and assayed for ^{3}H-thymidine incorporation as described in *Proc. Natl. Acad. Sci. USA* (1986) *83*:4749-4753. *In vivo* assays to measure cell proliferation rates in plants by confocal microscopy are provided in Lauf *et al.* (1998) *Plant Cell 10*:1375-1389, herein incorporated by reference.

Because the sequences of the invention modulate cell cycle and DNA metabolism they may find use in transformation and culture protocols. The sequences may work to increase transformation efficiency, particularly in plants recalcitrant to transformation and promote the establishment of plant cultures, particularly in plants which are difficult to culture.

It is recognized by one skilled in the art that regulating proteins may bind to a variety of other proteins to achieve their regulatory effect, and that the regulatory effect of a protein may depend on more than one other protein or complex or mode of action. For example, one promoter influenced by the Rb complex controls expression of E2F, a gene required for cell cycling through G1 into S phase. The E2F gene is expressed when cyclin-dependent kinase phosphorylates Rb, disrupting the complex and dissociating RbAp's and histone deacetylase (HD). HD has no inherent ability to bind to DNA and must be directed to a targeted promoter by co-regulatory factors. HD is known to interact directly with both Rb and RbAP. E2F and Rb have potential use in therapy of human diseases resulting from inappropriate cell proliferation, such as cancer. See U.S. Patent No. 5,885,833, "Nucleic acid constructs for the cell-cycle-regulated expression of genes and therapeutic methods utilizing such constructs"; U.S. Patent No. 5,851,991, "Therapeutic use of the retinoblastoma susceptibility gene product." See also U.S. Patent No. 5, 756,476, "Inhibition of cell proliferation using antisense oligonucleotides." Thus, the sequences of the invention may have potential use in such therapies by modulating the effects of E2F and Rb.

Other conserved components of this regulatory complex have also been found in plants, including homologs of plant D-type cyclins. It is understood by those skilled in the art that components of conserved regulatory complexes may differ slightly in function across organisms, and that such variations may be useful in achieving desired effects.

In another embodiment, altered expression of the genes or proteins of the present invention may be used to alter flower development. Affecting RbAp48/MSI1 expression by co-suppression severely disrupts normal inflorescence development, causing flowers to progressively lose their normal morphology. However, different effects may be achieved with different expression levels of the present invention. For example, a maize line containing a Mutator (Mu) transposable-element-tagged allele has been identified but has no obvious developmental phenotype. It is possible, and highly likely, that the lack of obvious phenotypes in the msi1-1::Mu1/msi1-2::Mu1 homozygotes is a consequence of the functional redundancy within the MSI gene family in maize; therefore, a

phenotype may result if mutant alleles of the other MSI family members are isolated and crossed into the msi1-1 or msi1-2 mutant background.

It is recognized that the present invention is not premised upon any particular mechanism of action. It is sufficient for purposes of the invention that the genes and proteins are involved in gene expression and inflorescence development and can be used to alter such processes and characteristics. The methods and compositions described herein may be used alone or in combination with other methods and/or compositions.

Dominant negatives of the sequences may also be used. Deletion or point mutations could be expressed to produce altered or truncated proteins that retain the ability to interact with retinoblastoma but disrupt the normal function of the complex. For example, expression of the Rb binding domain may prevent recruitment of histone deacetylase by out-competing endogenous MSI with a peptide that does not intereact with HD. In this manner, the Rb complex is specifically targeted without affecting HD function in other complexes, such as, for example, when HC-toxin is used.

Extrapolating from mutation studies of retinoblastoma in mice, phenotypes of maize MSI mutants might include embryo lethality, stunted growth as the result of reduced MSI or other gene dosage, or abnormal cellular differentiation and zones of cell proliferation. Assays to measure cell proliferation rates within zones of the plant using confocal sectioning are described in Laufs *et al.* (1998) *Plant Cell 10*:1375-1389.

In another embodiment, the present invention improves the overall frequency and/or homologous recombination frequency of genetic transformation events. The invention does not depend for its effect on a particular mechanism of increasing the overall frequency or homologous recombination frequency. There is suggestion in the prior art that DNA metabolism and/or chromatin state is interrelated with both overall gene expression and frequency of recombination and/or genetic transformation. See, for example: Akhmedov *et al.* (1998) *J. Biol. Chem. 273*: 24088-24094 (showing that structural-maintenance-of-chromosomes proteins interact with DNA in chromosome condensation, DNA recombination, and gene dosage compensation); Singh *et al.* (1998) *Mol. Cell. Biol. 18*: 5511-5522 (proposing that effects on gene silencing are indirect consequences of changes in

chromatin structure); Piruat and Aguilera (1998) *EMBO J. 17*: 4859-4872 (showing that novel yeast gene THO2 is involved in RNA polymerase II transcription and providing new evidence for transcriptional elongation-associated recombination); and Rieth *et al.* (1999) *Mol. Reprod. Dev. 53*: 1-7 (use of bovine satellite sequences to increase transgene integration by homologous recombination in bovine embryos).

It is understood in the art that plant DNA viruses and fungal pathogens remodel the control of the host replication and gene expression machinery to accomplish their own replication and effective infection. The present invention may be useful in preventing such corruption of the cell. Therefore, the compositions and methods of the invention find use in controlling pathogenic agents. The anti-pathogenic compositions comprise polynucleotides and proteins, particularly, the maize MSI-like nucleotide and amino acid of the present invention and fragments and variants thereof. Accordingly, the compositions and methods are also useful in protecting plants against fungal pathogens, viruses, nematodes, insects and the like.

By "disease resistance" is intended that the plants avoid the disease symptoms which are the outcome of plant-pathogen interactions. That is, pathogens are prevented from causing plant diseases and the associated disease symptoms, or alternatively, the disease symptoms caused by the pathogen is minimized or lessened. The methods of the invention can be utilized to protect plants from disease, particularly those diseases that are caused by plant pathogens. By "anti-pathogenic compositions" is intended that the compositions of the invention are capable of suppressing, controlling, and/or killing the invading pathogenic organism.

The HC toxin of the maize fungal pathogen *Cochliobolus carbonum* has been found to be a potent inhibitor of histone deacetylase (HD) activity. The HC-toxin of the maize pathogen *C. carbonum* and related cyclic tetrapeptides inhibit HDs and cause hyperacetylation of histones in susceptible, but not in resistant, maize strains. Perhaps the inhibition of histone deacetylation interferes with the induction of plant defense genes mediated by RNA polymerase II transcription. Also, inhibition of deacetylation by HC-toxin may lead to a rather general inhibition of host rRNA transcription, owing to inhibition of nucleolar HD2. Thus,

the sequences of the present invention provides methods for modulation of HD proteins and may create new possibilities for engineering fungal pathogen resistance. It is not known whether the ability of HC-toxin to overcome the maize defense response is due to prevention of defense gene expression or more general toxicity such as disruption of cell cycle control.

The MSI-like sequences of the present invention additionally comprises a new tool for controlling replication of agriculturally important DNA viruses. If the model for Retinoblastoma function in mammalian cells is applicable to plants, we can expect that the Retinoblastoma pathway found in maize and other plants has a central role in plant cell division and differentiation as well, and that it is the target for inactivation during replication of plant DNA viruses. See for example, Gutierrez *et al.* (2000) *EMBO Journal 19*:792-799, herein incorporated by reference.

The methods involve stably transforming a plant with a DNA construct comprising an anti-pathogenic nucleotide sequence of the invention operably linked to promoter that drives expression in a plant. Such methods find use in agriculture particularly in limiting the impact of plant pathogens on crop plants. While the choice of promoter will depend on the desired timing and location of expression of the anti-pathogenic nucleotide sequences, preferred promoters include constitutive and pathogen-inducible promoters.

Inhibition of endogenous gene expression is deleterious to the cell, organism, or plant. Thus, the constructs of the invention can be used to selectively kill target cells or tissues. This can be accomplished through the use of inducible or tissue-preferred promoters. In this manner, the sequences of the invention may find use in enhancing pathogen resistance. For example, in one embodiment an antisense construct for the MSI coding sequence is operably linked to a pathogen inducible promoter. Upon contact with the pathogen, the MSI antisense construct is expressed resulting in disruption of gene expression leading to cell death and effectively preventing the invasion of the pathogen.

Additionally, the compositions can be used in formulation use for their antimicrobial activities. The proteins of the invention can be formulated with an acceptable carrier into a pesticidal composition(s) that is for example, a suspension, a solution, an emulsion, a dusting powder, a dispersible granule, a

wettable powder, and an emulsifiable concentrate, an aerosol, an impregnated granule, an adjuvant, a coatable paste, and also encapsulations in, for example, polymer substances.

The MSI-like sequences of the invention can also be used to control resistance to pathogens by enhancing the defense mechanisms in a plant. While the exact function of the MSI-like sequences is not known, they are involved in influencing the expression of defense-related proteins. It is recognized that the present invention is not premised upon any particular mechanism of action of the MSI-like sequences. It is sufficient for purposes of the invention that the genes and proteins are involved in the plant defense system and can be used to increase resistance levels in the plant to pathogens. Assays to determine a modulation in the plant pathogen defense system in response to various pathogens are known in the art. For example, resistance to *C. carbonum* is characterized by the induction of pathogen defense genes and the appearance of only non-expanding lesions. Alternatively, sensitivity to *C. carbonum* is easily assayed by measuring lesion expansion.

The plant defense mechanisms described herein may be used alone or in combination with other methods, proteins or agents to protect against plant diseases and pathogens. Other plant defense proteins include those described in copending applications entitled "*Methods for Enhancing Disease Resistance in Plants*", U.S. Application Serial No. 09/256,898, filed February 24, 1999, and copending application entitled "*Genes for Activation of Plant Pathogen Defense Systems*", U.S. Application Serial No. 09/256,158, filed February 24, 1999, all of which are herein incorporated by reference. Although any one of a variety of second nucleotide sequences may be utilized, some embodiments of the invention encompass those second nucleotide sequences that, when expressed in a plant, help to increase the resistance of a plant to pathogens. It is recognized that such second nucleotide sequences may be used in either the sense or antisense orientation depending on the desired outcome.

The invention is drawn to compositions and methods for inducing resistance in a plant to plant pests. Accordingly, the compositions and methods are also useful in protecting plants against fungal pathogens, viruses, nematodes, insects and the like. Pathogens of the invention include, but are not limited to,

viruses or viroids, bacteria, insects, nematodes, fungi, and the like. Viruses include any plant virus, for example, tobacco or cucumber mosaic virus, ringspot virus, necrosis virus, maize dwarf mosaic virus, etc. Specific fungal and viral pathogens for the major crops include: Soybeans: *Phytophthora megasperma* fsp. *glycinea*, *Macrophomina phaseolina*, *Rhizoctonia solani*, *Sclerotinia sclerotiorum*, *Fusarium oxysporum*, *Diaporthe phaseolorum* var. *sojae* (*Phomopsis sojae*), *Diaporthe phaseolorum* var. *caulivora*, *Sclerotium rolfsii*, *Cercospora kikuchii*, *Cercospora sojina*, *Peronospora manshurica*, *Colletotrichum dematium* (*Colletotichum truncatum*), *Corynespora cassiicola*, *Septoria glycines*, *Phyllosticta sojicola*, *Alternaria alternata*, *Pseudomonas syringae* p.v. *glycinea*, *Xanthomonas campestris* p.v. *phaseoli*, *Microsphaera diffusa*, *Fusarium semitectum*, *Phialophora gregata*, Soybean mosaic virus, *Glomerella glycines*, Tobacco Ring spot virus, Tobacco Streak virus, *Phakopsora pachyrhizi*, *Pythium aphanidermatum*, *Pythium ultimum*, *Pythium debaryanum*, Tomato spotted wilt virus, *Heterodera glycines Fusarium solani*; Canola: *Albugo candida*, *Alternaria brassicae*, *Leptosphaeria maculans*, *Rhizoctonia solani*, *Sclerotinia sclerotiorum*, *Mycosphaerella brassiccola*, *Pythium ultimum*, *Peronospora parasitica*, *Fusarium roseum*, *Alternaria alternata*; Alfalfa: *Clavibater michiganese subsp. insidiosum*, *Pythium ultimum*, *Pythium irregulare*, *Pythium splendens*, *Pythium debaryanum*, *Pythium aphanidermatum*, *Phytophthora megasperma*, *Peronospora trifoliorum*, *Phoma medicaginis* var. *medicaginis*, *Cercospora medicaginis*, *Pseudopeziza medicaginis*, *Leptotrochila medicaginis*, *Fusarium*, *Xanthomonas campestris* p.v. *alfalfae*, *Aphanomyces euteiches*, *Stemphylium herbarum*, *Stemphylium alfalfae*; Wheat: *Pseudomonas syringae* p.v. *atrofaciens*, *Urocystis agropyri*, *Xanthomonas campestris* p.v. *translucens*, *Pseudomonas syringae* p.v. *syringae*, *Alternaria alternata*, *Cladosporium herbarum*, *Fusarium graminearum*, *Fusarium avenaceum*, *Fusarium culmorum*, *Ustilago tritici*, *Ascochyta tritici*, *Cephalosporium gramineum*, *Collotetrichum graminicola*, *Erysiphe graminis* f.sp. *tritici*, *Puccinia graminis* f.sp. *tritici*, *Puccinia recondita* f.sp. *tritici*, *Puccinia striiformis*, *Pyrenophora tritici-repentis*, *Septoria nodorum*, *Septoria tritici*, *Septoria avenae*, *Pseudocercosporella herpotrichoides*, *Rhizoctonia solani*, *Rhizoctonia cerealis*, *Gaeumannomyces graminis* var. *tritici*, *Pythium aphanidermatum*, *Pythium arrhenomanes*, *Pythium ultimum*, *Bipolaris*

sorokiniana, Barley Yellow Dwarf Virus, Brome Mosaic Virus, Soil Borne Wheat Mosaic Virus, Wheat Streak Mosaic Virus, Wheat Spindle Streak Virus, American Wheat Striate Virus, *Claviceps purpurea*, *Tilletia tritici*, *Tilletia laevis*, *Ustilago tritici*, *Tilletia indica*, *Rhizoctonia solani*, *Pythium arrhenomannes*, *Pythium gramicola*, *Pythium aphanidermatum*, High Plains Virus, European wheat striate virus; Sunflower: *Plasmophora halstedii*, *Sclerotinia sclerotiorum*, Aster Yellows, *Septoria helianthi*, *Phomopsis helianthi*, *Alternaria helianthi*, *Alternaria zinniae*, *Botrytis cinerea*, *Phoma macdonaldii*, *Macrophomina phaseolina*, *Erysiphe cichoracearum*, *Rhizopus oryzae*, *Rhizopus arrhizus*, *Rhizopus stolonifer*, *Puccinia helianthi*, *Verticillium dahliae*, *Erwinia carotovorum* pv. *carotovora*, *Cephalosporium acremonium*, *Phytophthora cryptogea*, *Albugo tragopogonis*; Corn: *Fusarium moniliforme* var. *subglutinans*, *Erwinia stewartii*, *Fusarium moniliforme*, *Gibberella zeae* (*Fusarium graminearum*), *Stenocarpella maydi* (*Diplodia maydis*), *Pythium irregulare*, *Pythium debaryanum*, *Pythium graminicola*, *Pythium splendens*, *Pythium ultimum*, *Pythium aphanidermatum*, *Aspergillus flavus*, *Bipolaris maydis* O, T (*Cochliobolus heterostrophus*), *Helminthosporium carbonum* I, II & III (*Cochliobolus carbonum*), *Exserohilum turcicum* I, II & III, *Helminthosporium pedicellatum*, *Physoderma maydis*, *Phyllosticta maydis*, *Kabatiella-maydis*, *Cercospora sorghi*, *Ustilago maydis*, *Puccinia sorghi*, *Puccinia polysora*, *Macrophomina phaseolina*, *Penicillium oxalicum*, *Nigrospora oryzae*, *Cladosporium herbarum*, *Curvularia lunata*, *Curvularia inaequalis*, *Curvularia pallescens*, *Clavibacter michiganense* subsp. *nebraskense*, *Trichoderma viride*, Maize Dwarf Mosaic Virus A & B, Wheat Streak Mosaic Virus, Maize Chlorotic Dwarf Virus, *Claviceps sorghi*, *Pseudonomas avenae*, *Erwinia chrysanthemi* pv. *zea*, *Erwinia carotovora*, *Corn stunt spiroplasma*, *Diplodia macrospora*, *Sclerophthora macrospora*, *Peronosclerospora sorghi*, *Peronosclerospora philippinensis*, *Peronosclerospora maydis*, *Peronosclerospora sacchari*, *Sphacelotheca reiliana*, *Physopella zeae*, *Cephalosporium maydis*, *Cephalosporium acremonium*, Maize Chlorotic Mottle Virus, High Plains Virus, Maize Mosaic Virus, Maize Rayado Fino Virus, Maize Streak Virus, Maize Stripe Virus, Maize Rough Dwarf Virus; Sorghum: *Exserohilum turcicum*, *Colletotrichum graminicola* (*Glomerella graminicola*), *Cercospora sorghi*, *Gloeocercospora sorghi*, *Ascochyta sorghina*, *Pseudomonas*

syringae p.v. *syringae, Xanthomonas campestris* p.v. *holcicola, Pseudomonas andropogonis, Puccinia purpurea, Macrophomina phaseolina, Perconia circinata, Fusarium moniliforme, Alternaria alternata, Bipolaris sorghicola, Helminthosporium sorghicola, Curvularia lunata, Phoma insidiosa, Pseudomonas avenae* (*Pseudomonas alboprecipitans*), *Ramulispora sorghi, Ramulispora sorghicola, Phyllachara sacchari, Sporisorium reilianum* (*Sphacelotheca reiliana*), *Sphacelotheca cruenta, Sporisorium sorghi*, Sugarcane mosaic H, Maize Dwarf Mosaic Virus A & B, *Claviceps sorghi, Rhizoctonia solani, Acremonium strictum, Sclerophthona macrospora, Peronosclerospora sorghi, Peronosclerospora philippinensis, Sclerospora graminicola, Fusarium graminearum, Fusarium oxysporum, Pythium arrhenomanes, Pythium graminicola*, etc.

Nematodes include parasitic nematodes such as root-knot, cyst, lesion, and renniform nematodes, etc.

Insect pests include insects selected from the orders Coleoptera, Diptera, Hymenoptera, Lepidoptera, Mallophaga, Homoptera, Hemiptera, Orthoptera, Thysanoptera, Dermaptera, Isoptera, Anoplura, Siphonaptera, Trichoptera, etc., particularly Coleoptera and Lepidoptera. Insect pests of the invention for the major crops include: Maize: *Ostrinia nubilalis*, European corn borer; *Agrotis ipsilon*, black cutworm; *Helicoverpa zea*, corn earworm; *Spodoptera frugiperda*, fall armyworm; *Diatraea grandiosella*, southwestern corn borer; *Elasmopalpus lignosellus*, lesser cornstalk borer; *Diatraea saccharalis*, surgarcane borer; *Diabrotica virgifera*, western corn rootworm; *Diabrotica longicornis barberi*, northern corn rootworm; *Diabrotica undecimpunctata howardi*, southern corn rootworm; *Melanotus spp.*, wireworms; *Cyclocephala borealis*, northern masked chafer (white grub); *Cyclocephala immaculata*, southern masked chafer (white grub); *Popillia japonica*, Japanese beetle; *Chaetocnema pulicaria*, corn flea beetle; *Sphenophorus maidis*, maize billbug; *Rhopalosiphum maidis*, corn leaf aphid; *Anuraphis maidiradicis*, corn root aphid; *Blissus leucopterus leucopterus*, chinch bug; *Melanoplus femurrubrum*, redlegged grasshopper; *Melanoplus sanguinipes*, migratory grasshopper; *Hylemya platura*, seedcorn maggot; *Agromyza parvicornis*, corn blot leafminer; *Anaphothrips obscrurus*, grass thrips; *Solenopsis milesta*, thief ant; *Tetranychus urticae*, twospotted spider mite; Sorghum: *Chilo partellus*, sorghum borer; *Spodoptera frugiperda*, fall armyworm; *Helicoverpa zea*, corn

earworm; *Elasmopalpus lignosellus*, lesser cornstalk borer; *Feltia subterranea*, granulate cutworm; *Phyllophaga crinita*, white grub; *Eleodes*, *Conoderus*, and *Aeolus* spp., wireworms; *Oulema melanopus*, cereal leaf beetle; *Chaetocnema pulicaria*, corn flea beetle; *Sphenophorus maidis*, maize billbug; *Rhopalosiphum maidis*; corn leaf aphid; *Sipha flava*, yellow sugarcane aphid; *Blissus leucopterus leucopterus*, chinch bug; *Contarinia sorghicola*, sorghum midge; *Tetranychus cinnabarinus*, carmine spider mite; *Tetranychus urticae*, twospotted spider mite; Wheat: *Pseudaletia unipunctata*, army worm; *Spodoptera frugiperda*, fall armyworm; *Elasmopalpus lignosellus*, lesser cornstalk borer; *Agrotis orthogonia*, western cutworm; *Elasmopalpus lignosellus*, lesser cornstalk borer; *Oulema melanopus*, cereal leaf beetle; *Hypera punctata*, clover leaf weevil; *Diabrotica undecimpunctata howardi*, southern corn rootworm; Russian wheat aphid; *Schizaphis graminum*, greenbug; *Macrosiphum avenae*, English grain aphid; *Melanoplus femurrubrum*, redlegged grasshopper; *Melanoplus differentialis*, differential grasshopper; *Melanoplus sanguinipes*, migratory grasshopper; *Mayetiola destructor*, Hessian fly; *Sitodiplosis mosellana*, wheat midge; *Meromyza americana*, wheat stem maggot; *Hylemya coarctata*, wheat bulb fly; *Frankliniella fusca*, tobacco thrips; *Cephus cinctus*, wheat stem sawfly; *Aceria tulipae*, wheat curl mite; Sunflower: *Suleima helianthana*, sunflower bud moth; *Homoeosoma electellum*, sunflower moth; *zygogramma exclamationis*, sunflower beetle; *Bothyrus gibbosus*, carrot beetle; *Neolasioptera murtfeldtiana*, sunflower seed midge; Cotton: *Heliothis virescens*, cotton budworm; *Helicoverpa zea*, cotton bollworm; *Spodoptera exigua*, beet armyworm; *Pectinophora gossypiella*, pink bollworm; *Anthonomus grandis grandis*, boll weevil; *Aphis gossypii*, cotton aphid; *Pseudatomoscelis seriatus*, cotton fleahopper; *Trialeurodes abutilonea*, bandedwinged whitefly; *Lygus lineolaris*, tarnished plant bug; *Melanoplus femurrubrum*, redlegged grasshopper; *Melanoplus differentialis*, differential grasshopper; *Thrips tabaci*, onion thrips; *Franklinkiella fusca*, tobacco thrips; *Tetranychus cinnabarinus*, carmine spider mite; *Tetranychus urticae*, twospotted spider mite; Rice: *Diatraea saccharalis*, sugarcane borer; *Spodoptera frugiperda*, fall armyworm; *Helicoverpa zea*, corn earworm; *Colaspis brunnea*, grape colaspis; *Lissorhoptrus oryzophilus*, rice water weevil; *Sitophilus oryzae*, rice weevil; *Nephotettix nigropictus*, rice leafhopper; *Blissus leucopterus leucopterus*, chinch

bug; *Acrosternum hilare*, green stink bug; Soybean: *Pseudoplusia includens*, soybean looper; *Anticarsia gemmatalis*, velvetbean caterpillar; *Plathypena scabra*, green cloverworm; *Ostrinia nubilalis*, European corn borer; *Agrotis ipsilon*, black cutworm; *Spodoptera exigua*, beet armyworm; *Heliothis virescens*, cotton budworm; *Helicoverpa zea*, cotton bollworm; *Epilachna varivestis*, Mexican bean beetle; *Myzus persicae*, green peach aphid; *Empoasca fabae*, potato leafhopper; *Acrosternum hilare*, green stink bug; *Melanoplus femurrubrum*, redlegged grasshopper; *Melanoplus differentialis*, differential grasshopper; *Hylemya platura*, seedcorn maggot; *Sericothrips variabilis*, soybean thrips; *Thrips tabaci*, onion thrips; *Tetranychus turkestani*, strawberry spider mite; *Tetranychus urticae*, twospotted spider mite; Barley: *Ostrinia nubilalis*, European corn borer; *Agrotis ipsilon*, black cutworm; *Schizaphis graminum*, greenbug; *Blissus leucopterus leucopterus*, chinch bug; *Acrosternum hilare*, green stink bug; *Euschistus servus*, brown stink bug; *Delia platura*, seedcorn maggot; *Mayetiola destructor*, Hessian fly; *Petrobia latens*, brown wheat mite; Oil Seed Rape: *Brevicoryne brassicae*, cabbage aphid; *Phyllotreta cruciferae*, Flea beetle; *Mamestra configurata*, Bertha armyworm; *Plutella xylostella*, Diamond-back moth; *Delia* ssp., Root maggots.

The present invention also provides a method of genotyping a plant using a polynucleotide of the present invention. Preferably, the plant is a monocot, such as maize or sorghum. Genotyping provides a means of distinguishing homologs of a chromosome pair and can be used to differentiate segregants in a plant population. Molecular marker methods can be used for phylogenetic studies, characterizing genetic relationships among crop varieties, identifying crosses or somatic hybrids, localizing chromosomal segments affecting monogenic traits, map based cloning, and the study of quantitative inheritance. See, e.g., *Plant Molecular Biology: A Laboratory Manual*, Chapter 7, Clark, Ed., Springer-Verlag, Berlin (1997). For molecular marker methods, see generally, *The DNA Revolution* by Andrew H. Paterson 1996 (Chapter 2) in: *Genome Mapping in Plants* (ed. Andrew H. Paterson) by Academic Press/R. G. Landis Company, Austin, Texas, pp.7-21.

The particular method of genotyping in the present invention may employ any number of molecular marker analytic techniques such as, but not limited to, restriction fragment length polymorphisms (RFLPs). RFLPs are the product of allelic differences between DNA restriction fragments caused by nucleotide

sequence variability. As is well known to those of skill in the art, RFLPs are typically detected by extraction of genomic DNA and digestion with a restriction enzyme. Generally, the resulting fragments are separated according to size and hybridized with a probe; single copy probes are preferred. Restriction fragments from homologous chromosomes are revealed. Differences in fragment size among alleles represent an RFLP. Thus, the present invention further provides a means to follow segregation of an MSI-like gene or nucleic acid of the present invention as well as chromosomal sequences genetically linked to these genes or nucleic acids using such techniques as RFLP analysis. Linked chromosomal sequences are within 50 centiMorgans (cM), often within 40 or 30 cM, preferably within 20 or 10 cM, more preferably within 5, 3, 2, or 1 cM of an MSI gene.

In another embodiment of the invention, the nucleotide sequences for MSI-like can be utilized to produce the enzyme with greater purity. Such enzyme preparations can be utilized for assays of enzymatic activity as well as to produce anti-MSI-like antibodies which may have pharmaceutical or therapeutic use in treatment of cell proliferation diseases such as cancer. Mechanisms for antibody production are known in the art. See, for example, Harlow and Lane (1988) *Antibodies, A Laboratory Manual* (Cold Spring Harbor Publications, New York) and the references cited therein. Such antibodies are useful to immunoprecipitate MSI-like from cell extracts and isolate members of regulatory co-factor complexes associated with MSI and Rb *in vivo*.

Antibodies can be raised to a protein of the present invention, including individual, allelic, strain, or species variants, and fragments thereof, both in their naturally occurring (full-length) forms and in recombinant forms. Additionally, antibodies are raised to these proteins in either their native configurations or in non-native configurations. Anti-idiotypic antibodies can also be generated. Many methods of making antibodies are known to persons of skill; however, one of skill will recognize that many variations upon such methods are known.

Monoclonal antibodies are prepared from cells secreting the desired antibody. Monoclonals antibodies are screened for binding to a protein from which the immunogen was derived. Specific monoclonal and polyclonal antibodies will usually have an antibody binding site with an affinity constant for its cognate monovalent antigen at least between 10^6-10^7, usually at least 10^8,

preferably at least 10^9, more preferably at least 10^{10}, and most preferably at least 10^{11} liters/mole.

The following examples are offered by way of illustration and not by way of limitation.

EXPERIMENTAL

Example 1: Transformation and Regeneration of Transgenic Plants

Immature maize embryos from greenhouse donor plants are bombarded with a plasmid containing ZmMSIa or ZmMSIb sequence operably linked to a ubiquitin promoter plus a plasmid containing the selectable marker gene PAT (Wohlleben *et al.* (1988) *Gene 70*:25-37) that confers resistance to the herbicide Bialaphos. Transformation is performed as follows. Media recipes follow below.

Preparation of Target Tissue

The ears are surface sterilized in 30% Chlorox bleach plus 0.5% Micro detergent for 20 minutes, and rinsed two times with sterile water. The immature embryos are excised and placed embryo axis side down (scutellum side up), 25 embryos per plate, on 560Y medium for 4 hours and then aligned within the 2.5-cm target zone in preparation for bombardment.

Preparation of DNA

A plasmid vector comprising the ZmMSIb or ZmMSIa nucleotide sequence operably linked to a ubiquitin promoter is made. This plasmid DNA plus plasmid DNA containing a PAT selectable marker is precipitated onto 1.1 μm (average diameter) tungsten pellets using a $CaCl_2$ precipitation procedure as follows:

100 μl prepared tungsten particles in water
10 μl (1 μg) DNA in TrisEDTA buffer (1 μg total)
100 μl 2.5 M $CaCl_2$
10 μl 0.1 M spermidine

Each reagent is added sequentially to the tungsten particle suspension, while maintained on the multitube vortexer. The final mixture is sonicated briefly

and allowed to incubate under constant vortexing for 10 minutes. After the precipitation period, the tubes are centrifuged briefly, liquid removed, washed with 500 ml 100% ethanol, and centrifuged for 30 seconds. Again the liquid is removed, and 105 µl 100% ethanol is added to the final tungsten particle pellet. For particle gun bombardment, the tungsten/DNA particles are briefly sonicated and 10 µl spotted onto the center of each macrocarrier and allowed to dry about 2 minutes before bombardment.

Particle Gun Treatment

The sample plates are bombarded at level #4 in particle gun #HE34-1 or #HE34-2. All samples receive a single shot at 650 PSI, with a total of ten aliquots taken from each tube of prepared particles/DNA.

Subsequent Treatment

Following bombardment, the embryos are kept on 560Y medium for 2 days, then transferred to 560R selection medium containing 3 mg/liter Bialaphos, and subcultured every 2 weeks. After approximately 10 weeks of selection, selection-resistant callus clones are transferred to 288J medium to initiate plant regeneration. Following somatic embryo maturation (2-4 weeks), well-developed somatic embryos are transferred to medium for germination and transferred to the lighted culture room. Approximately 7-10 days later, developing plantlets are transferred to 272V hormone-free medium in tubes for 7-10 days until plantlets are well established. Plants are then transferred to inserts in flats (equivalent to 2.5" pot) containing potting soil and grown for 1 week in a growth chamber, subsequently grown an additional 1-2 weeks in the greenhouse, then transferred to classic 600 pots (1.6 gallon) and grown to maturity.

Plants are monitored and scored for the relevant trait affected by the present invention. For example, the plants are inoculated with *C. carbonum*. Modulation of the plant pathogen response can be assayed by measuring lesion expansion. When the pathogen defense genes are induced, non-expanding lesions are produced, while sensitivity to *C. carbonum* is characterized by lesion expansion. Alternatively, upon infection of the plant with a pathogen, standard gene expression assays (i.e., RNase protection, Northern blot, and RNA profiling) can

be employed to determine if expression of plant pathogen defense genes are induced or repressed.

Example 2: *Agrobacterium*-Mediated Transformation of Plants

For *Agrobacterium*-mediated transformation of maize with an MSI-like sequence sequence, preferably the method of Zhao is employed (U.S. Patent No. 5,981,840, and PCT patent publication WO98/32326; the contents of which are hereby incorporated by reference). Briefly, immature embryos are isolated from maize and the embryos contacted with a suspension of *Agrobacterium*, where the bacteria are capable of transferring the MSI-like sequences to at least one cell of at least one of the immature embryos (step 1: the infection step). In this step the immature embryos are preferably immersed in an *Agrobacterium* suspension for the initiation of inoculation. The embryos are co-cultured for a time with the *Agrobacterium* (step 2: the co-cultivation step). Preferably the immature embryos are cultured on solid medium following the infection step. Following this co-cultivation period an optional "resting" step is contemplated. In this resting step, the embryos are incubated in the presence of at least one antibiotic known to inhibit the growth of *Agrobacterium* without the addition of a selective agent for plant transformants (step 3: resting step). Preferably the immature embryos are cultured on solid medium with antibiotic, but without a selecting agent, for elimination of *Agrobacterium* and for a resting phase for the infected cells. Next, inoculated embryos are cultured on medium containing a selective agent and growing transformed callus is recovered (step 4: the selection step). Preferably, the immature embryos are cultured on solid medium with a selective agent resulting in the selective growth of transformed cells. The callus is then regenerated into plants (step 5: the regeneration step), and preferably calli grown on selective medium are cultured on solid medium to regenerate the plants.

Example 3: Soybean Embryo Transformation Prophetic Example

Soybean embryos are bombarded with a plasmid containing the MSI-like nucleotide sequence operably linked to a ubiquitin promoter as follows. To induce somatic embryos, cotyledons, 3-5 mm in length dissected from surface-sterilized, immature seeds of the soybean cultivar A2872, are cultured in the light or dark at

26°C on an appropriate agar medium for six to ten weeks. Somatic embryos producing secondary embryos are then excised and placed into a suitable liquid medium. After repeated selection for clusters of somatic embryos that multiplied as early, globular-staged embryos, the suspensions are maintained as described below.

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

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

A selectable marker gene that can be used to facilitate soybean transformation is a transgene composed of the 35S promoter from Cauliflower Mosaic Virus (Odell *et al.* (1985) *Nature 313*:810-812), the hygromycin phosphotransferase gene from plasmid pJR225 (from *E. coli*; Gritz *et al.* (1983) *Gene 25*:179-188), and the 3' region of the nopaline synthase gene from the T-DNA of the Ti plasmid of *Agrobacterium tumefaciens*. The expression cassette comprising the MSI gene operably linked to the ubiquitin promoter can be isolated as a restriction fragment. This fragment can then be inserted into a unique restriction site of the vector carrying the marker gene.

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

Approximately 300-400 mg of a two-week-old suspension culture is placed in an empty 60x15 mm petri dish and the residual liquid removed from the tissue with a pipette. For each transformation experiment, approximately 5-10 plates of

tissue are normally bombarded. Membrane rupture pressure is set at 1100 psi, and the chamber is evacuated to a vacuum of 28 inches mercury. The tissue is placed approximately 3.5 inches away from the retaining screen and bombarded three times. Following bombardment, the tissue can be divided in half and placed back into liquid and cultured as described above.

Five to seven days post bombardment, the liquid media may be exchanged with fresh media, and eleven to twelve days post-bombardment with fresh media containing 50 mg/ml hygromycin. This selective media can be refreshed weekly. Seven to eight weeks post-bombardment, green, transformed tissue may be observed growing from untransformed, necrotic embryogenic clusters. Isolated green tissue is removed and inoculated into individual flasks to generate new, clonally propagated, transformed embryogenic suspension cultures. Each new line may be treated as an independent transformation event. These suspensions can then be subcultured and maintained as clusters of immature embryos or regenerated into whole plants by maturation and germination of individual somatic embryos.

Example 4: Sunflower Meristem Tissue Transformation Prophetic Example

Sunflower meristem tissues are transformed with an expression cassette containing the MSI-like sequence operably linked to ubiquitin promoter as follows (see also European Patent Number EP 0 486233, herein incorporated by reference, and Malone-Schoneberg *et al.* (1994) *Plant Science 103*:199-207). Mature sunflower seed (*Helianthus annuus* L.) are dehulled using a single wheat-head thresher. Seeds are surface sterilized for 30 minutes in a 20% Chlorox bleach solution with the addition of two drops of Tween 20 per 50 ml of solution. The seeds are rinsed twice with sterile distilled water.

Split embryonic axis explants are prepared by a modification of procedures described by Schrammeijer *et al.* (Schrammeijer *et al.*(1990) *Plant Cell Rep. 9*:55-60). Seeds are imbibed in distilled water for 60 minutes following the surface sterilization procedure. The cotyledons of each seed are then broken off, producing a clean fracture at the plane of the embryonic axis. Following excision of the root tip, the explants are bisected longitudinally between the primordial leaves. The two halves are placed, cut surface up, on GBA medium consisting of

Murashige and Skoog mineral elements (Murashige *et al.* (1962) *Physiol. Plant. 15*:473-497), Shepard's vitamin additions (Shepard (1980) in *Emergent Techniques for the Genetic Improvement of Crops* (University of Minnesota Press, St. Paul, Minnesota), 40 mg/l adenine sulfate, 30 g/l sucrose, 0.5 mg/l 6-benzyl-aminopurine (BAP), 0.25 mg/l indole-3-acetic acid (IAA), 0.1 mg/l gibberellic acid (GA_3), pH 5.6, and 8 g/l Phytagar.

The explants are subjected to microprojectile bombardment prior to *Agrobacterium* treatment (Bidney *et al.* (1992) *Plant Mol. Biol. 18*:301-313). Thirty to forty explants are placed in a circle at the center of a 60 X 20 mm plate for this treatment. Approximately 4.7 mg of 1.8 mm tungsten microprojectiles are resuspended in 25 ml of sterile TE buffer (10 mM Tris HCl, 1 mM EDTA, pH 8.0) and 1.5 ml aliquots are used per bombardment. Each plate is bombarded twice through a 150 mm nytex screen placed 2 cm above the samples in a PDS 1000® particle acceleration device.

Disarmed *Agrobacterium tumefaciens* strain EHA105 is used in all transformation experiments. A binary plasmid vector comprising the expression cassette that contains the MSI-like operably linked to the ubiquitin promoter is introduced into *Agrobacterium* strain EHA105 via freeze-thawing as described by Holsters *et al.* (1978) *Mol. Gen. Genet. 163*:181-187. This plasmid further comprises a kanamycin selectable marker gene (i.e, *nptII*). Bacteria for plant transformation experiments are grown overnight (28°C and 100 RPM continuous agitation) in liquid YEP medium (10 gm/l yeast extract, 10 gm/l Bactopeptone, and 5 gm/l NaCl, pH 7.0) with the appropriate antibiotics required for bacterial strain and binary plasmid maintenance. The suspension is used when it reaches an OD_{600} of about 0.4 to 0.8. The *Agrobacterium* cells are pelleted and resuspended at a final OD_{600} of 0.5 in an inoculation medium comprised of 12.5 mM MES pH 5.7, 1 gm/l NH_4Cl, and 0.3 gm/l $MgSO_4$.

Freshly bombarded explants are placed in an *Agrobacterium* suspension, mixed, and left undisturbed for 30 minutes. The explants are then transferred to GBA medium and co-cultivated, cut surface down, at 26°C and 18-hour days. After three days of co-cultivation, the explants are transferred to 374B (GBA medium lacking growth regulators and a reduced sucrose level of 1%) supplemented with 250 mg/l cefotaxime and 50 mg/l kanamycin sulfate. The

explants are cultured for two to five weeks on selection and then transferred to fresh 374B medium lacking kanamycin for one to two weeks of continued development. Explants with differentiating, antibiotic-resistant areas of growth that have not produced shoots suitable for excision are transferred to GBA medium containing 250 mg/l cefotaxime for a second 3-day phytohormone treatment. Leaf samples from green, kanamycin-resistant shoots are assayed for the presence of NPTII by ELISA and for the presence of transgene expression by assaying for MSI protein activity as described in Lusser *et al.* (1997) *Science 277*:88-91; Rundlett *et al.* (1996) *PNAS 93*:14503-14508; DeRubertis *et al.* (1996) *Nature 384*:589-591; and Pazin *et al.* (1997) *Cell 89*:325-328.

NPTII-positive shoots are grafted to Pioneer® hybrid 6440 *in vitro*-grown sunflower seedling rootstock. Surface sterilized seeds are germinated in 48-0 medium (half-strength Murashige and Skoog salts, 0.5% sucrose, 0.3% gelrite, pH 5.6) and grown under conditions described for explant culture. The upper portion of the seedling is removed, a 1 cm vertical slice is made in the hypocotyl, and the transformed shoot inserted into the cut. The entire area is wrapped with parafilm to secure the shoot. Grafted plants can be transferred to soil following one week of *in vitro* culture. Grafts in soil are maintained under high humidity conditions followed by a slow acclimatization to the greenhouse environment. Transformed sectors of T_0 plants (parental generation) maturing in the greenhouse are identified by NPTII ELISA and/or by analysis of MSI-like protein activity in leaf extracts while transgenic seeds harvested from NPTII-positive T_0 plants are identified by analysis of MSI protein activity in small portions of dry seed cotyledon.

An alternative sunflower transformation protocol allows the recovery of transgenic progeny without the use of chemical selection pressure. Seeds are dehulled and surface-sterilized for 20 minutes in a 20% Chlorox bleach solution with the addition of two to three drops of Tween 20 per 100 ml of solution, then rinsed three times with distilled water. Sterilized seeds are imbibed in the dark at 26°C for 20 hours on filter paper moistened with water. The cotyledons and root radical are removed, and the meristem explants are cultured on 374E (GBA medium consisting of MS salts, Shepard vitamins, 40 mg/l adenine sulfate, 3% sucrose, 0.5 mg/l 6-BAP, 0.25 mg/l IAA, 0.1 mg/l GA, and 0.8% Phytagar at pH 5.6) for 24 hours under the dark. The primary leaves are removed to expose the

apical meristem, around 40 explants are placed with the apical dome facing upward in a 2 cm circle in the center of 374M (GBA medium with 1.2% Phytagar), and then cultured on the medium for 24 hours in the dark.

Approximately 18.8 mg of 1.8 μm tungsten particles are resuspended in 150 μl absolute ethanol. After sonication, 8 μl of it is dropped on the center of the surface of macrocarrier. Each plate is bombarded twice with 650 psi rupture discs in the first shelf at 26 mm of Hg helium gun vacuum.

The plasmid of interest is introduced into *Agrobacterium tumefaciens* strain EHA105 via freeze thawing as described previously. The pellet of overnight-grown bacteria at 28°C in a liquid YEP medium (10 g/l yeast extract, 10 g/l Bactopeptone, and 5 g/l NaCl, pH 7.0) in the presence of 50 μg/l kanamycin is resuspended in an inoculation medium (12.5 mM 2-mM 2-(N-morpholino) ethanesulfonic acid, MES, 1 g/l NH_4Cl and 0.3 g/l $MgSO_4$ at pH 5.7) to reach a final concentration of 4.0 at OD 600. Particle-bombarded explants are transferred to GBA medium (374E), and a droplet of bacteria suspension is placed directly onto the top of the meristem. The explants are co-cultivated on the medium for 4 days, after which the explants are transferred to 374C medium (GBA with 1% sucrose and no BAP, IAA, GA3 and supplemented with 250 μg/ml cefotaxime). The plantlets are cultured on the medium for about two weeks under 16-hour day and 26°C incubation conditions.

Explants (around 2 cm long) from two weeks of culture in 374C medium are screened for MSI protein activity using assays known in the art such as those described in Lusser *et al.* (1997) *Science 277*:88-91; Rundlett *et al.* (1996) *PNAS 93*:14503-14508; DeRubertis *et al.* (1996) *Nature 384*:589-591; Pazin *et al.* (1997) *Cell 89*:325-328. After positive (i.e., for MSI protein expression) explants are identified, those shoots that fail to exhibit MSI protein activity are discarded, and every positive explant is subdivided into nodal explants. One nodal explant contains at least one potential node. The nodal segments are cultured on GBA medium for three to four days to promote the formation of auxiliary buds from each node. Then they are transferred to 374C medium and allowed to develop for an additional four weeks. Developing buds are separated and cultured for an additional four weeks on 374C medium. Pooled leaf samples from each newly recovered shoot are screened again by the appropriate protein activity assay. At

this time, the positive shoots recovered from a single node will generally have been enriched in the transgenic sector detected in the initial assay prior to nodal culture.

Recovered shoots positive for MSI protein expression are grafted to Pioneer hybrid 6440 *in vitro*-grown sunflower seedling rootstock. The rootstocks are prepared in the following manner. Seeds are dehulled and surface-sterilized for 20 minutes in a 20% Chlorox bleach solution with the addition of two to three drops of Tween 20 per 100 ml of solution, and are rinsed three times with distilled water. The sterilized seeds are germinated on the filter moistened with water for three days, then they are transferred into 48 medium (half-strength MS salt, 0.5% sucrose, 0.3% gelrite pH 5.0) and grown at 26°C under the dark for three days, then incubated at 16-hour-day culture conditions. The upper portion of selected seedling is removed, a vertical slice is made in each hypocotyl, and a transformed shoot is inserted into a V-cut. The cut area is wrapped with parafilm. After one week of culture on the medium, grafted plants are transferred to soil. In the first two weeks, they are maintained under high humidity conditions to acclimatize to a greenhouse environment.

Bombardment and Culture Media

Bombardment medium (560Y) comprises 4.0 g/l N6 basal salts (SIGMA C-1416), 1.0 ml/l Eriksson's Vitamin Mix (1000X SIGMA-1511), 0.5 mg/l thiamine HCl, 120.0 g/l sucrose, 1.0 mg/l 2,4-D, and 2.88 g/l L-proline (brought to volume with D-I H_20 following adjustment to pH 5.8 with KOH); 2.0 g/l Gelrite (added after bringing to volume with D-I H_20); and 8.5 mg/l silver nitrate (added after sterilizing the medium and cooling to room temperature). Selection medium (560R) comprises 4.0 g/l N6 basal salts (SIGMA C-1416), 1.0 ml/l Eriksson's Vitamin Mix (1000X SIGMA-1511), 0.5 mg/l thiamine HCl, 30.0 g/l sucrose, and 2.0 mg/l 2,4-D (brought to volume with D-I H_20 following adjustment to pH 5.8 with KOH); 3.0 g/l Gelrite (added after bringing to volume with D-I H_20); and 0.85 mg/l silver nitrate and 3.0 mg/l bialaphos(both added after sterilizing the medium and cooling to room temperature).

Plant regeneration medium (288J) comprises 4.3 g/l MS salts (GIBCO 11117-074), 5.0 ml/l MS vitamins stock solution (0.100 g nicotinic acid, 0.02 g/l thiamine HCL, 0.10 g/l pyridoxine HCL, and 0.40 g/l glycine brought to volume

with polished D-I H_20) (Murashige and Skoog (1962) *Physiol. Plant. 15*:473), 100 mg/l myo-inositol, 0.5 mg/l zeatin, 60 g/l sucrose, and 1.0 ml/l of 0.1 mM abscisic acid (brought to volume with polished D-I H_20 after adjusting to pH 5.6); 3.0 g/l Gelrite (added after bringing to volume with D-I H_20); and 1.0 mg/l indoleacetic acid and 3.0 mg/l bialaphos (added after sterilizing the medium and cooling to 60°C). Hormone-free medium (272V) comprises 4.3 g/l MS salts (GIBCO 11117-074), 5.0 ml/l MS vitamins stock solution (0.100 g/l nicotinic acid, 0.02 g/l thiamine HCL, 0.10 g/l pyridoxine HCL, and 0.40 g/l glycine brought to volume with polished D-I H_20), 0.1 g/l myo-inositol, and 40.0 g/l sucrose (brought to volume with polished D-I H_20 after adjusting pH to 5.6); and 6 g/l bacto-agar (added after bringing to volume with polished D-I H_20), sterilized and cooled to 60°C.

All publications and patent applications mentioned in the specification are indicative of the level of those skilled in the art to which this invention pertains. All publications and patent applications are herein incorporated by reference to the same extent as if each individual publication or patent application was specifically and individually indicated to be incorporated by reference.

Although the foregoing 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.

THAT WHICH IS CLAIMED:

1. An isolated polypeptide selected from the group consisting of:

a) a polypeptide comprising an amino acid sequence set forth in SEQ ID NO:2 or 4;

b) a polypeptide encoded by a nucleotide sequence comprising the sequence set forth in SEQ ID NO:1 or 3;

c) a polypeptide having 40% identity to the sequence set forth in SEQ ID NO:2 or 4;

d) a polypeptide encoded by a nucleotide sequence that hybridizes under stringent conditions to the sequence set forth in SEQ ID NO:1 or 3; and,

e) a polypeptide encoded by the nucleotide sequence deposited in a bacterial host as Patent Deposit No. PTA-317.

2. An isolated nucleic acid molecule selected from the group consisting of:

a) a nucleotide sequence comprising the sequence set forth in SEQ ID NO:1 or SEQ ID NO:3;

b) a nucleotide sequence comprising the DNA sequences deposited in a bacterial host as Patent Deposit No. PTA-317;

c) a nucleotide sequence encoding a polypeptide comprising the amino acid sequence set forth in SEQ ID NO:2 or SEQ ID NO:4;

d) a nucleotide sequence having 40% identity to the sequence set forth in SEQ ID NO:1 or 3;

e) a nucleotide sequence comprising at least 16 contiguous nucleotides of a sequence of a), b), c), or d);

f) a nucleotide sequence comprising an antisense sequence corresponding to a sequence of a), b), c), d), or e); and,

g) a nucleotide sequence that hybridizes under stringent conditions to the nucleotide sequences of a), b), c), d), e), or f).

3. A vector for delivery of a nucleotide sequence into a host cell, the vector comprising at least one nucleotide sequence of claim 2.

4. A host cell containing the vector of claim 3.

5. A DNA construct comprising a nucleotide sequence of claim 2, wherein said nucleotide sequence is operably linked to a promoter.

6. The DNA construct of claim 5, wherein said promoter drives expression in a plant cell.

7. A transformed plant cell having stably incorporated into its genome at least one nucleotide sequence, said nucleotide sequence operably linked to a heterologous promoter that drives expression in a plant cell, wherein said nucleotide sequence is selected from the group consisting of:

a) a nucleotide sequence comprising the sequence set forth in SEQ ID NO:1 or SEQ ID NO:3;

b) a nucleotide sequence comprising the DNA sequences deposited in a bacterial host as Patent Deposit No. PTA-317;

c) a nucleotide sequence encoding a polypeptide comprising the amino acid sequence set forth in SEQ ID NO:2 or SEQ ID NO:4;

d) a nucleotide sequence having 40% identity to the sequence set forth in SEQ ID NO:1 or 3;

e) nucleotide sequence comprising at least 16 contiguous nucleotides of a sequence of a), b), c), or d);

f) nucleotide sequence comprising an antisense sequence corresponding to a sequence of a), b), c), d), or e); and,

g) a nucleotide sequence that hybridizes under stringent conditions to the nucleotide sequences of a), b), c), d), e), or f).

8. The plant cell of claim 7, wherein said promoter is selected from the group consisting of a constitutive promoter, a tissue-preferred promoter, an inducible promoter and a pathogen-specific promoter.

9. The plant cell of claim 7, wherein said promoter is a pathogen-inducible promoter and said nucleotide sequence is an antisense construct.

10. A plant having stably incorporated into its genome at least one nucleotide sequence, said nucleotide sequence operably linked to a heterologous promoter that drives expression in a plant cell, wherein said nucleotide sequence is selected from the group consisting of:

a) a nucleotide sequence comprising the sequence set forth in SEQ ID NO:1 or SEQ ID NO:3;

b) a nucleotide sequence comprising the DNA sequences deposited in a bacterial host as Patent Deposit No. PTA-317;

c) a nucleotide sequence encoding a polypeptide comprising the amino acid sequence set forth in SEQ ID NO:2 or SEQ ID NO:4;

d) a nucleotide sequence having 40% identity to the sequence set forth in SEQ ID NO:1 or 3;

e) a nucleotide sequence comprising at least 16 contiguous nucleotides of a sequence of a), b), c), or d);

f) a nucleotide sequence comprising an antisense sequence corresponding to a sequence of a), b), c), d), or e); and,

g) a nucleotide sequence that hybridizes under stringent conditions to the nucleotide sequences of a), b), c), d), e), or f).

11. The plant of claim 10, wherein said promoter is selected from the group consisting of a constitutive promoter, a tissue-preferred promoter, an inducible promoter and a pathogen-specific promoter.

12. The plant of claim 10, wherein said promoter is a pathogen-inducible promoter and said nucleotide sequence is an antisense construct.

13. The plant of claim 10, wherein said plant is a monocot.

14. The plant of claim 13, wherein said monocot is maize, wheat, rice, barley, sorghum, or rye.

15. The plant of claim 10, wherein said plant is a dicot.

16. Transformed seed of the plant of claim 10.

17. Transformed seed of the plant of claim 13.

18. Transformed seed of the plant of claim 15.

19. A method for producing a polypeptide comprising culturing the host cell of claim 4 under conditions in which a nucleic acid molecule encoding the polypeptide is expressed, wherein said nucleic acid is selected from the group consisting of:

a) a nucleotide sequence comprising the sequence set forth in SEQ ID NO:1 or SEQ ID NO:3;

b) a nucleotide sequence comprising the DNA sequences deposited in a bacterial host as Patent Deposit No. PTA-317;

c) a nucleotide sequence encoding a polypeptide comprising the amino acid sequence set forth in SEQ ID NO:2 or SEQ ID NO:4;

d) a nucleotide sequence having 40% identity to the sequence set forth in SEQ ID NO:1 or 3;

e) a nucleotide sequence comprising at least 16 contiguous nucleotides of a sequence of a), b), c), or d);

f) a nucleotide sequence comprising an antisense sequence corresponding to a sequence of a), b), c), d), or e); and,

g) a nucleotide sequence that hybridizes under stringent conditions to the nucleotide sequences of a), b), c), d), e), or f).

20. The method of claim 19, wherein said host cell is a plant cell.

21. A method for modulating the activity of a polypeptide, said method comprising stably integrating into the genome of a plant cell at least one nucleotide sequence operably linked to a heterologous promoter that drives expression in said plant cell, wherein said nucleotide sequence is selected from the group consisting of:

a) a nucleotide sequence comprising the sequence set forth in SEQ ID NO:1 or SEQ ID NO:3;

b) a nucleotide sequence comprising the DNA sequences deposited in a bacterial host as Patent Deposit No. PTA-317;

c) a nucleotide sequence encoding a polypeptide comprising the amino acid sequence set forth in SEQ ID NO:2 or SEQ ID NO:4;

d) a nucleotide sequence having 40% identity to the sequence set forth in SEQ ID NO:1 or 3;

e) a nucleotide sequence comprising at least 16 contiguous nucleotides of a sequence of a), b), c), or d);

f) a nucleotide sequence comprising an antisense sequence corresponding to a sequence of a), b), c), d), or e); and,

g) a nucleotide sequence that hybridizes under stringent conditions to the nucleotide sequences of a), b), c), d), e), or f).

22. The method of claim 21 wherein said activity modulates the cell cycle.

23. A method for increasing pathogen resistance in a plant, said method comprises stably integrating into the genome of a plant cell at least one nucleotide sequence operably linked to a heterologous promoter that drives expression in said plant cell, wherein said nucleotide sequence is selected from the group consisting of:

a) a nucleotide sequence comprising the sequence set forth in SEQ ID NO:1 or SEQ ID NO:3;

b) a nucleotide sequence comprising the DNA sequences deposited in a bacterial host as Patent Deposit No. PTA-317;

c) a nucleotide sequence encoding a polypeptide comprising the amino acid sequence set forth in SEQ ID NO:2 or SEQ ID NO:4;

d) a nucleotide sequence having 40% identity to the sequence set forth in SEQ ID NO:1 or 3;

e) a nucleotide sequence comprising at least 16 contiguous nucleotides of a sequence of a), b), c), or d);

f) a nucleotide sequence comprising an antisense sequence corresponding to a sequence of a), b), c), d), or e); and,

g) a nucleotide sequence that hybridizes under stringent conditions to the nucleotide sequences of a), b), c), d), e), or f).

SEQUENCE LISTING

<110> Baldwin, Donald

<120> Maize MSI Polynucleotides and Methods of Use

<130> 5718-101-1-PC

<150> US 60/143,437
<151> 1999-07-13

<160> 4

<170> FastSEQ for Windows Version 3.0

<210> 1
<211> 1584
<212> DNA
<213> Zea mays

<220>
<221> CDS
<222> (63)...(1355)

<400> 1

```
ctcaaacccc ctccgctctc aaaccctaac cctagcccta gcccgctccc ggcggaccga      60
cg atg ccg aag gca ggc ggc ggg gcg gcg gac gaa gag gag ttc cgc         107
   Met Pro Lys Ala Gly Gly Gly Ala Ala Asp Glu Glu Glu Phe Arg
    1               5                  10                  15

gcc gag gtg gag gag cgc ctc atc aac gag gag tac aag atc tgg aag       155
Ala Glu Val Glu Glu Arg Leu Ile Asn Glu Glu Tyr Lys Ile Trp Lys
                20                  25                  30

aag aac aca ccc ttc ctc tac gac ctc gtc atc acc cac gcg ctc gaa       203
Lys Asn Thr Pro Phe Leu Tyr Asp Leu Val Ile Thr His Ala Leu Glu
            35                  40                  45

tgg ccc tcc ctt acc gtg cag tgg ctc ccc gac cgc acc gag ccg ccg       251
Trp Pro Ser Leu Thr Val Gln Trp Leu Pro Asp Arg Thr Glu Pro Pro
        50                  55                  60

ggg aag gac cac tcc gtc cag aag atg atc ctt ggc acg cac acc tct       299
Gly Lys Asp His Ser Val Gln Lys Met Ile Leu Gly Thr His Thr Ser
    65                  70                  75

gac aac gag ccc aac tac ctc atg ctc gcg cag gtc cag ctg ccc ctc       347
Asp Asn Glu Pro Asn Tyr Leu Met Leu Ala Gln Val Gln Leu Pro Leu
80                  85                  90                  95

gac gac gcc gag gcc gac gcc cgc cac tac gac gat gac cac gcc gac       395
Asp Asp Ala Glu Ala Asp Ala Arg His Tyr Asp Asp Asp His Ala Asp
                100                 105                 110

atc ggt ggg ttt ggc gcc gcc tcc ggc aag gtg caa att gtt cag cag       443
Ile Gly Gly Phe Gly Ala Ala Ser Gly Lys Val Gln Ile Val Gln Gln
            115                 120                 125

ata aat cat gat gga gag gtc aat cga gct cgc tat atg ccc caa aat       491
```

```
Ile Asn His Asp Gly Glu Val Asn Arg Ala Arg Tyr Met Pro Gln Asn
        130                 135                 140

tca ttt ata att gct act aag aca gtt agc gca aaa gta tat gtt ttt      539
Ser Phe Ile Ile Ala Thr Lys Thr Val Ser Ala Lys Val Tyr Val Phe
    145                 150                 155

gac tat agc aag cac cca tct aag cct cca cta gat ggt gca tgc aac      587
Asp Tyr Ser Lys His Pro Ser Lys Pro Pro Leu Asp Gly Ala Cys Asn
160                 165                 170                 175

cct gat tta cgg ctg aag gga cat aat tct gaa gga tat ggc tta tca      635
Pro Asp Leu Arg Leu Lys Gly His Asn Ser Glu Gly Tyr Gly Leu Ser
                180                 185                 190

tgg agt att ttt aaa gag gga cat ttg tta agt ggg tct gac gat gct      683
Trp Ser Ile Phe Lys Glu Gly His Leu Leu Ser Gly Ser Asp Asp Ala
            195                 200                 205

caa att tgc ttg tgg gac att aaa gca aat agt aga aac aaa agt ctt      731
Gln Ile Cys Leu Trp Asp Ile Lys Ala Asn Ser Arg Asn Lys Ser Leu
        210                 215                 220

gac gcc ttg cag att ttt aag cat cat gat ggt gtc gtt gaa gat gtt      779
Asp Ala Leu Gln Ile Phe Lys His His Asp Gly Val Val Glu Asp Val
    225                 230                 235

gct tgg cac ttg agg cat gag tac tta ttt ggg tca gtt ggt gac gat      827
Ala Trp His Leu Arg His Glu Tyr Leu Phe Gly Ser Val Gly Asp Asp
240                 245                 250                 255

tat cat ctt ttg att tgg gac ctg cgg tct ccc gcc cct act aaa cct      875
Tyr His Leu Leu Ile Trp Asp Leu Arg Ser Pro Ala Pro Thr Lys Pro
                260                 265                 270

gtt cag tca gtg gtg gcg cac cag ggt gag gtg aac tgc ctg gct ttt      923
Val Gln Ser Val Val Ala His Gln Gly Glu Val Asn Cys Leu Ala Phe
            275                 280                 285

aac ccg ttc aat gaa tgg gtt gtt gca act ggt tct act gac aag act      971
Asn Pro Phe Asn Glu Trp Val Val Ala Thr Gly Ser Thr Asp Lys Thr
        290                 295                 300

gtc aaa tta ttt gat ctt agg aag att gat act tct ctg cac acc ttt     1019
Val Lys Leu Phe Asp Leu Arg Lys Ile Asp Thr Ser Leu His Thr Phe
    305                 310                 315

gac tgt cac aaa gag gaa gtt ttt caa gtt gga tgg agt cca aag aat     1067
Asp Cys His Lys Glu Glu Val Phe Gln Val Gly Trp Ser Pro Lys Asn
320                 325                 330                 335

gaa act gta ctt gca tcc tgt tgt ctg ggc aga agg ctc atg gtc tgg     1115
Glu Thr Val Leu Ala Ser Cys Cys Leu Gly Arg Arg Leu Met Val Trp
                340                 345                 350

aat cta agc agg atc gac caa gaa caa aca cca gag gat gct gaa gat     1163
Asn Leu Ser Arg Ile Asp Gln Glu Gln Thr Pro Glu Asp Ala Glu Asp
            355                 360                 365

ggt cct cca gag ctc atg ttc atc cat gga ggc cat acc agc aag atc     1211
Gly Pro Pro Glu Leu Met Phe Ile His Gly Gly His Thr Ser Lys Ile
```

```
       370                 375                 380

tct gat ttc tcc tgg aac ccc tgc gag gac tgg gtg gtt gct agc gtc     1259
Ser Asp Phe Ser Trp Asn Pro Cys Glu Asp Trp Val Val Ala Ser Val
    385                 390                 395

gct gag gac aac atc ctc cag ata tgg cag atg gct gag aac atc tac     1307
Ala Glu Asp Asn Ile Leu Gln Ile Trp Gln Met Ala Glu Asn Ile Tyr
400                 405                 410                 415

cat gac gaa gat gat ctg cca att agc gat gag ccg gca aaa act tct     1355
His Asp Glu Asp Asp Leu Pro Ile Ser Asp Glu Pro Ala Lys Thr Ser
                420                 425                 430

tgaagcaaca catcacaaaa ggaatcttaa ctcctatcga ctgttggtgc ccctagatcg   1415
gaaccctgct cgccggagaa tcttaaccga gtgagagttg gctgtaggtt gcttccgtgg   1475
tcgtaacttt atgcttgacg ggttggtcgt ttcttcgacc gtttgtatgt gtattcgtcg   1535
aacctttagc attatacata tgctgattga ctaaaaaaaa aaaaaaaaa              1584
```

<210> 2
<211> 431
<212> PRT
<213> Zea mays

<400> 2

```
Met Pro Lys Ala Gly Gly Gly Ala Ala Asp Glu Glu Glu Phe Arg Ala
 1               5                  10                  15
Glu Val Glu Glu Arg Leu Ile Asn Glu Glu Tyr Lys Ile Trp Lys Lys
            20                  25                  30
Asn Thr Pro Phe Leu Tyr Asp Leu Val Ile Thr His Ala Leu Glu Trp
        35                  40                  45
Pro Ser Leu Thr Val Gln Trp Leu Pro Asp Arg Thr Glu Pro Pro Gly
    50                  55                  60
Lys Asp His Ser Val Gln Lys Met Ile Leu Gly Thr His Thr Ser Asp
65                  70                  75                  80
Asn Glu Pro Asn Tyr Leu Met Leu Ala Gln Val Gln Leu Pro Leu Asp
                85                  90                  95
Asp Ala Glu Ala Asp Ala Arg His Tyr Asp Asp Asp His Ala Asp Ile
            100                 105                 110
Gly Gly Phe Gly Ala Ala Ser Gly Lys Val Gln Ile Val Gln Gln Ile
        115                 120                 125
Asn His Asp Gly Glu Val Asn Arg Ala Arg Tyr Met Pro Gln Asn Ser
    130                 135                 140
Phe Ile Ile Ala Thr Lys Thr Val Ser Ala Lys Val Tyr Val Phe Asp
145                 150                 155                 160
Tyr Ser Lys His Pro Ser Lys Pro Pro Leu Asp Gly Ala Cys Asn Pro
                165                 170                 175
Asp Leu Arg Leu Lys Gly His Asn Ser Glu Gly Tyr Gly Leu Ser Trp
            180                 185                 190
Ser Ile Phe Lys Glu Gly His Leu Leu Ser Gly Ser Asp Asp Ala Gln
        195                 200                 205
Ile Cys Leu Trp Asp Ile Lys Ala Asn Ser Arg Asn Lys Ser Leu Asp
    210                 215                 220
Ala Leu Gln Ile Phe Lys His His Asp Gly Val Val Glu Asp Val Ala
225                 230                 235                 240
Trp His Leu Arg His Glu Tyr Leu Phe Gly Ser Val Gly Asp Asp Tyr
                245                 250                 255
His Leu Leu Ile Trp Asp Leu Arg Ser Pro Ala Pro Thr Lys Pro Val
            260                 265                 270
Gln Ser Val Val Ala His Gln Gly Glu Val Asn Cys Leu Ala Phe Asn
        275                 280                 285
```

```
Pro Phe Asn Glu Trp Val Val Ala Thr Gly Ser Thr Asp Lys Thr Val
    290                 295                 300
Lys Leu Phe Asp Leu Arg Lys Ile Asp Thr Ser Leu His Thr Phe Asp
305                 310                 315                 320
Cys His Lys Glu Glu Val Phe Gln Val Gly Trp Ser Pro Lys Asn Glu
                325                 330                 335
Thr Val Leu Ala Ser Cys Cys Leu Gly Arg Arg Leu Met Val Trp Asn
            340                 345                 350
Leu Ser Arg Ile Asp Gln Glu Gln Thr Pro Glu Asp Ala Glu Asp Gly
        355                 360                 365
Pro Pro Glu Leu Met Phe Ile His Gly Gly His Thr Ser Lys Ile Ser
    370                 375                 380
Asp Phe Ser Trp Asn Pro Cys Glu Asp Trp Val Val Ala Ser Val Ala
385                 390                 395                 400
Glu Asp Asn Ile Leu Gln Ile Trp Gln Met Ala Glu Asn Ile Tyr His
                405                 410                 415
Asp Glu Asp Asp Leu Pro Ile Ser Asp Glu Pro Ala Lys Thr Ser
            420                 425                 430
```

<210> 3
<211> 1777
<212> DNA
<213> Zea mays

<220>
<221> CDS
<222> (164)...(1522)

<400> 3

```
agcacaggga ctgcctccgg agaccggagc aggccgcaag cgtccaggtt cgagacgcca      60
gcacgctctc gcagtacagc agcattatca gggagcgtgc cgtgggagtg ggaggttgcc     120
tcaggtgcgg agcagtcgcc ggcggcgacc atcggtgtct gag atg aag gag aga       175
                                                Met Lys Glu Arg
                                                 1

ggc ggc tcc agg gcg gca gtg gac gag cgc tac gcg cag tgg aag tcg       223
Gly Gly Ser Arg Ala Ala Val Asp Glu Arg Tyr Ala Gln Trp Lys Ser
 5                  10                  15                  20

ctc atc ccc gtt ctc tac gac tgg ttt gcc aat cac aac ctc gtt tgg       271
Leu Ile Pro Val Leu Tyr Asp Trp Phe Ala Asn His Asn Leu Val Trp
                25                  30                  35

cca tcc ctc tcc tgc cgg tgg ggg cca cag ttt gag aaa gct acc tac       319
Pro Ser Leu Ser Cys Arg Trp Gly Pro Gln Phe Glu Lys Ala Thr Tyr
            40                  45                  50

aag aat cgt caa cgc ctt tac cta tct gaa cag acc gat ggg agt gtg       367
Lys Asn Arg Gln Arg Leu Tyr Leu Ser Glu Gln Thr Asp Gly Ser Val
        55                  60                  65

cct aat act ctg gtt atc gca aac tgt gaa gtt gtg aaa cca agg gtt       415
Pro Asn Thr Leu Val Ile Ala Asn Cys Glu Val Val Lys Pro Arg Val
    70                  75                  80

gca gct gct gaa cat atc tcg cag ttt aac gag gaa gca cga tca cct       463
Ala Ala Ala Glu His Ile Ser Gln Phe Asn Glu Glu Ala Arg Ser Pro
 85                  90                  95                  100

ttt gta aag aag tac aag act ata gtt cat cct ggt gag gtt aac aga       511
Phe Val Lys Lys Tyr Lys Thr Ile Val His Pro Gly Glu Val Asn Arg
```

```
                105                 110                 115

atc agg gag ctt cca cag aac agt aag atc ata gcc acg cac act gac      559
Ile Arg Glu Leu Pro Gln Asn Ser Lys Ile Ile Ala Thr His Thr Asp
            120                 125                 130

agt cca gat gta ctt gtt tgg gat gtt gaa gca caa cca aat aga cat      607
Ser Pro Asp Val Leu Val Trp Asp Val Glu Ala Gln Pro Asn Arg His
        135                 140                 145

gct gtc tta gga gct tct gaa tct cgt cct gat ctg ata tta acg gga      655
Ala Val Leu Gly Ala Ser Glu Ser Arg Pro Asp Leu Ile Leu Thr Gly
    150                 155                 160

cac aag gaa gat gca gaa ttt gcg ctt gcc atg tgt cca gca gaa cca      703
His Lys Glu Asp Ala Glu Phe Ala Leu Ala Met Cys Pro Ala Glu Pro
165                 170                 175                 180

tat gtc ctg tca gga gga aag gac aaa ttt gtt gtc ttg tgg agc atc      751
Tyr Val Leu Ser Gly Gly Lys Asp Lys Phe Val Val Leu Trp Ser Ile
                185                 190                 195

caa gac cac ata tct gcc ctt ggg gat tcc tcg tct tct cct gga gca      799
Gln Asp His Ile Ser Ala Leu Gly Asp Ser Ser Ser Ser Pro Gly Ala
            200                 205                 210

tct ggc agc aag cag tct ggc aaa act gca aat gaa aag gag agt cct      847
Ser Gly Ser Lys Gln Ser Gly Lys Thr Ala Asn Glu Lys Glu Ser Pro
        215                 220                 225

aaa gtt gat cct aga ggt ata ttc cat ggc cat gac agt act gtt gaa      895
Lys Val Asp Pro Arg Gly Ile Phe His Gly His Asp Ser Thr Val Glu
    230                 235                 240

gat gtt cag ttc tgc cct tcc agt gcg cag gag ttt tgt agt gtg ggt      943
Asp Val Gln Phe Cys Pro Ser Ser Ala Gln Glu Phe Cys Ser Val Gly
245                 250                 255                 260

gat gat gct tgt ctt att ctg tgg gat gct cga act ggc act gac cca      991
Asp Asp Ala Cys Leu Ile Leu Trp Asp Ala Arg Thr Gly Thr Asp Pro
                265                 270                 275

gct gtt aag gtt gag aaa gct cac agt gga gat gtt cat tgt gtt gat     1039
Ala Val Lys Val Glu Lys Ala His Ser Gly Asp Val His Cys Val Asp
            280                 285                 290

tgg aat ccc ctt gat gtt aac tat atc tta act ggt tct gcc gat aac     1087
Trp Asn Pro Leu Asp Val Asn Tyr Ile Leu Thr Gly Ser Ala Asp Asn
        295                 300                 305

tct gtc cga atg tgg gat cgt cgc aat ctg ggt tcg gga gga gct ggt     1135
Ser Val Arg Met Trp Asp Arg Arg Asn Leu Gly Ser Gly Gly Ala Gly
    310                 315                 320

tct cca att cac aaa ttt gag ggc cat aaa gct gct gtt ctt tgt gtt     1183
Ser Pro Ile His Lys Phe Glu Gly His Lys Ala Ala Val Leu Cys Val
325                 330                 335                 340

cag tgg tca cct gac aga gca tct gtt ttt gga agt tct gca gaa gat     1231
Gln Trp Ser Pro Asp Arg Ala Ser Val Phe Gly Ser Ser Ala Glu Asp
                345                 350                 355
```

```
ggt ttc tta aac gtt tgg gac cat gag aag gtt ggg aag aag aaa aat     1279
Gly Phe Leu Asn Val Trp Asp His Glu Lys Val Gly Lys Lys Lys Asn
            360                 365                 370

tct aat gtc cca gct ggg ctt ttc ttt cag cat gct ggt cat agg gat     1327
Ser Asn Val Pro Ala Gly Leu Phe Phe Gln His Ala Gly His Arg Asp
        375                 380                 385

aag atc gta gac ttc cac tgg aat tcg tca gat cct tgg aca att gtt     1375
Lys Ile Val Asp Phe His Trp Asn Ser Ser Asp Pro Trp Thr Ile Val
    390                 395                 400

agt gtc tca gat gat ggt gag agc act ggt gga ggt gga aca ctg cag     1423
Ser Val Ser Asp Asp Gly Glu Ser Thr Gly Gly Gly Gly Thr Leu Gln
405                 410                 415                 420

ata tgg cga atg agt gat ttg atc tac cgc cca gag gat gaa gtt ctt     1471
Ile Trp Arg Met Ser Asp Leu Ile Tyr Arg Pro Glu Asp Glu Val Leu
                425                 430                 435

aca gag ctg gag aat ttc aag gct cac ctg gcc agt tgc gct ccg agg     1519
Thr Glu Leu Glu Asn Phe Lys Ala His Leu Ala Ser Cys Ala Pro Arg
            440                 445                 450

aat tgagtcatgt gaaaactctg agtaggttat gtagtttcgt ggccccaatt          1572
Asn


gacattggat ctgtcttctc atgctcctag agctgtaccc tcgaggtgct tgtgcctttg  1632
tgagcttgtg ttgaaccatg ttgtatgact aaatagtccg caaacgagtc agctggatgg  1692
ctaatctgta ggccaactat tgggatattg ttggatatta tcaacaaccg gttcagatta  1752
aaaaaaaaaa aaaaaaaaaa aaaaa                                        1777
```

<210> 4
<211> 453
<212> PRT
<213> Zea mays

<400> 4

```
Met Lys Glu Arg Gly Gly Ser Arg Ala Ala Val Asp Glu Arg Tyr Ala
 1               5                  10                  15
Gln Trp Lys Ser Leu Ile Pro Val Leu Tyr Asp Trp Phe Ala Asn His
            20                  25                  30
Asn Leu Val Trp Pro Ser Leu Ser Cys Arg Trp Gly Pro Gln Phe Glu
        35                  40                  45
Lys Ala Thr Tyr Lys Asn Arg Gln Arg Leu Tyr Leu Ser Glu Gln Thr
    50                  55                  60
Asp Gly Ser Val Pro Asn Thr Leu Val Ile Ala Asn Cys Glu Val Val
65                  70                  75                  80
Lys Pro Arg Val Ala Ala Ala Glu His Ile Ser Gln Phe Asn Glu Glu
                85                  90                  95
Ala Arg Ser Pro Phe Val Lys Lys Tyr Lys Thr Ile Val His Pro Gly
            100                 105                 110
Glu Val Asn Arg Ile Arg Glu Leu Pro Gln Asn Ser Lys Ile Ile Ala
        115                 120                 125
Thr His Thr Asp Ser Pro Asp Val Leu Val Trp Asp Val Glu Ala Gln
    130                 135                 140
Pro Asn Arg His Ala Val Leu Gly Ala Ser Glu Ser Arg Pro Asp Leu
145                 150                 155                 160
Ile Leu Thr Gly His Lys Glu Asp Ala Glu Phe Ala Leu Ala Met Cys
```

```
                165                 170                 175
Pro Ala Glu Pro Tyr Val Leu Ser Gly Gly Lys Asp Lys Phe Val Val
            180                 185                 190
Leu Trp Ser Ile Gln Asp His Ile Ser Ala Leu Gly Asp Ser Ser Ser
        195                 200                 205
Ser Pro Gly Ala Ser Gly Ser Lys Gln Ser Gly Lys Thr Ala Asn Glu
    210                 215                 220
Lys Glu Ser Pro Lys Val Asp Pro Arg Gly Ile Phe His Gly His Asp
225                 230                 235                 240
Ser Thr Val Glu Asp Val Gln Phe Cys Pro Ser Ser Ala Gln Glu Phe
                245                 250                 255
Cys Ser Val Gly Asp Asp Ala Cys Leu Ile Leu Trp Asp Ala Arg Thr
            260                 265                 270
Gly Thr Asp Pro Ala Val Lys Val Glu Lys Ala His Ser Gly Asp Val
        275                 280                 285
His Cys Val Asp Trp Asn Pro Leu Asp Val Asn Tyr Ile Leu Thr Gly
    290                 295                 300
Ser Ala Asp Asn Ser Val Arg Met Trp Asp Arg Arg Asn Leu Gly Ser
305                 310                 315                 320
Gly Gly Ala Gly Ser Pro Ile His Lys Phe Glu Gly His Lys Ala Ala
                325                 330                 335
Val Leu Cys Val Gln Trp Ser Pro Asp Arg Ala Ser Val Phe Gly Ser
            340                 345                 350
Ser Ala Glu Asp Gly Phe Leu Asn Val Trp Asp His Glu Lys Val Gly
        355                 360                 365
Lys Lys Lys Asn Ser Asn Val Pro Ala Gly Leu Phe Phe Gln His Ala
    370                 375                 380
Gly His Arg Asp Lys Ile Val Asp Phe His Trp Asn Ser Ser Asp Pro
385                 390                 395                 400
Trp Thr Ile Val Ser Val Ser Asp Asp Gly Glu Ser Thr Gly Gly Gly
                405                 410                 415
Gly Thr Leu Gln Ile Trp Arg Met Ser Asp Leu Ile Tyr Arg Pro Glu
            420                 425                 430
Asp Glu Val Leu Thr Glu Leu Glu Asn Phe Lys Ala His Leu Ala Ser
        435                 440                 445
Cys Ala Pro Arg Asn
    450
```