

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
20 July 2006 (20.07.2006)

PCT

(10) International Publication Number
WO 2006/076189 A2

(51) International Patent Classification:
C12N 15/82 (2006.01) A01H 5/00 (2006.01)
C07K 14/415 (2006.01)

(74) Agent: MOXON, Claire L.; Pioneer Hi-Bred International, Inc., 7100 N.W. 62nd Avenue, Johnston, Iowa 50131 (US).

(21) International Application Number:
PCT/US2006/000153

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(22) International Filing Date: 6 January 2006 (06.01.2006)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/643,720 13 January 2005 (13.01.2005) US

(71) Applicants (for all designated States except US): PIONEER HI-BRED INTERNATIONAL, INC. [US/US]; 7250 N.w. 62nd Avenue, Johnston, Iowa 50131-0552 (US). E. I. DU PONT DE NEMOURS & COMPANY [US/US]; 1007 Market Street, Wilmington, Delaware 19898 (US).

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

(72) Inventors; and

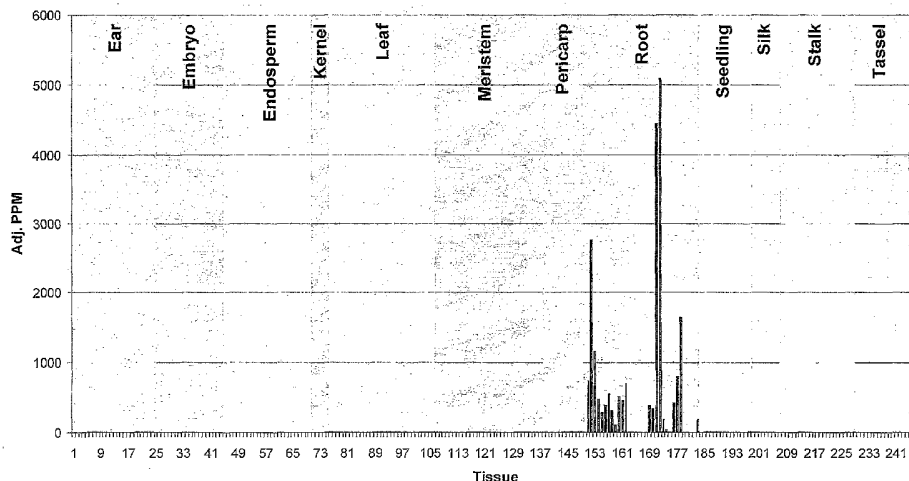
(75) Inventors/Applicants (for US only): DIEHN, Scott [US/US]; 4720 Tamara Lane, West Des Moines, Iowa 50265 (US). LU, Albert L. [CA/US]; 14 Charles Pointe, Newark, Delaware 19702 (US). SIMS, Lynne E. [US/US]; 816 W. Broadway Street, Polk City, Iowa 50226 (US). WARD, Kim R. [US/US]; 18 Airdrie Drive, Bear, Delaware 19701 (US).

Published:
— without international search report and to be republished upon receipt of that report

[Continued on next page]

(54) Title: MAIZE CYCLO1 GENE AND PROMOTER

Cyclo1 Expression Profile



(57) Abstract: The present invention provides compositions and methods for regulating expression of heterologous nucleotide sequences in a plant. Compositions include a novel nucleotide sequence for a root-preferred promoter for the gene encoding Cyclo1. A method for expressing a heterologous nucleotide sequence in a plant using the promoter sequences disclosed herein is provided. The method comprises stably incorporating into the genome of a plant cell a nucleotide sequence operably linked to the root-preferred promoter of the present invention and regenerating a stably transformed plant that expresses the nucleotide sequence. The present invention also relates to isolated nucleic acids encoding plant cyclotides. The invention relates to the construction of a chimeric gene encoding all or a portion of the plant cyclotides, in sense or antisense orientation, wherein expression of the chimeric gene results in the production of altered levels of plant cyclotides in a transformed host cell.

WO 2006/076189 A2



— *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.

MAIZE CYCLO1 GENE AND PROMOTER

FIELD OF THE INVENTION

The present invention relates to the field of plant molecular biology, more particularly to regulation of gene expression in plants.

BACKGROUND OF THE INVENTION

Recent advances in plant genetic engineering have enabled the engineering of plants having improved characteristics or traits, such as disease resistance, insect resistance, herbicide resistance, enhanced stability or shelf-life of the ultimate consumer product obtained from the plants and improvement of the nutritional quality of the edible portions of the plant. Thus, one or more desired genes from a source different than the plant, but engineered to impart different or improved characteristics or qualities, can be incorporated into the plant's genome. New gene(s) can then be expressed in the plant cell to exhibit the desired phenotype such as a new trait or characteristic.

The proper regulatory signals must be present and be in the proper location with respect to the gene in order to obtain expression of the newly inserted gene in the plant cell. These regulatory signals may include a promoter region, a 5' non-translated leader sequence and a 3' transcription termination/polyadenylation sequence.

A promoter is a DNA sequence that directs cellular machinery of a plant to produce RNA from the contiguous coding sequence downstream (3') of the promoter. The promoter region influences the rate, developmental stage, and cell type in which the RNA transcript of the gene is made. The RNA transcript is processed to produce messenger RNA (mRNA) which serves as a template for translation of the RNA sequence into the amino acid sequence of the encoded polypeptide. The 5' non-translated leader sequence is a region of the mRNA upstream of the protein coding region that may play a role in initiation and translation of the mRNA. The 3' transcription termination/polyadenylation signal is a non-translated region downstream of the protein coding region that functions in

the plant cells to cause termination of the RNA transcript and the addition of polyadenylate nucleotides to the 3' end of the RNA.

Expression of heterologous DNA sequences in a plant host is dependent upon the presence of an operably linked promoter that is functional within the plant host. The type of promoter sequence chosen is based on when and where within the organism expression of the heterologous DNA is desired. Where expression in specific tissues or organs is desired, tissue-preferred promoters may be used. Where gene expression in response to a stimulus is desired, inducible promoters are the regulatory element of choice. In contrast, where continuous expression is desired throughout the cells of a plant, constitutive promoters are utilized.

An inducible promoter is a promoter that is capable of directly or indirectly activating transcription of one or more DNA sequences or genes in response to an inducer. In the absence of an inducer, the DNA sequences or genes will not be transcribed or will be transcribed at a level lower than in an induced state. The inducer can be a chemical agent, such as a metabolite, growth regulator, herbicide or phenolic compound, or a physiological stress directly imposed upon the plant such as cold, heat, salt, drought, or toxins. In the case of fighting plant pests, it is also desirable to have a promoter which is induced by plant pathogens, including plant insect pests, nematodes or disease agents such as a bacterium, virus or fungus. Contact with the pathogen will induce activation of transcription, such that a pathogen-fighting protein will be produced at a time when it will be effective in defending the plant. A pathogen-induced promoter may also be used to detect contact with a pathogen, for example by expression of a detectable marker, so that the need for application of pesticides can be assessed. A plant cell containing an inducible promoter may be exposed to an inducer by externally applying the inducer to the cell or plant such as by spraying, watering, heating, or by exposure to the operative pathogen.

A constitutive promoter is a promoter that directs expression of a gene throughout the various parts of a plant and continuously throughout plant development. Examples of some constitutive promoters that are widely used for inducing the expression of heterologous genes in transgenic plants include the nopaline synthase (NOS) gene promoter, from *Agrobacterium tumefaciens* (U.S. Patent No. 5,034,322), the cauliflower mosaic virus (CaMv) 35S and 19S

promoters (U.S. Patent No. 5,352,605), those derived from any of the several actin genes, which are known to be expressed in most cells types (U.S. Patent No. 6,002,068), and the ubiquitin promoter (Christensen *et al.* (1989) *Plant Mol. Biol.* 12:619-632 and Christensen *et al.* (1992) *Plant Mol. Biol.* 18:675-689), which is a
5 gene product known to accumulate in many cell types.

Additional regulatory sequences upstream and/or downstream from the core promoter sequence may be included in expression constructs of transformation vectors to bring about varying levels of expression of heterologous nucleotide sequences in a transgenic plant. Genetically altering plants through the
10 use of genetic engineering techniques to produce plants with useful traits thus requires the availability of a variety of promoters.

In order to maximize the commercial application of transgenic plant technology, it is important to direct the expression of the introduced DNA in a site-specific manner. For example, it is desirable to produce toxic defensive
15 compounds in tissues subject to pathogen attack, but not in tissues that are to be harvested and eaten by consumers. By site-directing the synthesis or storage of desirable proteins or compounds, plants can be manipulated as factories, or production systems, for a tremendous variety of compounds with commercial utility. Cell-specific promoters provide the ability to direct the synthesis of
20 compounds, spatially and temporally, to highly specialized tissues or organs, such as roots, leaves, vascular tissues, embryos, seeds, or flowers.

Alternatively, it might be desirable to inhibit expression of a native DNA sequence within a plant's tissues to achieve a desired phenotype. Such inhibition might be accomplished with transformation of the plant to comprise a tissue-
25 preferred promoter operably linked to an antisense nucleotide sequence, such that expression of the antisense sequence produces an RNA transcript that interferes with translation of the mRNA of the native DNA sequence.

To date, the regulation of gene expression in plant roots has not been adequately studied despite the root's importance to plant development. To some
30 degree this is attributable to a lack of readily available, root-specific biochemical functions whose genes may be cloned, studied, and manipulated. Several genes that are preferentially expressed in plant root tissues have been identified. See, for example, Takahashi *et al.* (1991) *Plant J.* 1:327-332; Takahashi *et al.* (1990)

Proc. Natl. Acad. Sci. USA 87:8013-8016; Hertig *et al.* (1991) *Plant Mol Biol.* 16:171-174; Xu *et al.* (1995) *Plant Mol. Biol.* 27:237-248; Capone *et al.* (1994) *Plant Mol. Biol.* 25:681-691; Masuda *et al.* (1999) *Plant Cell Physiol.* 40(11):1177-81; Luschnig *et al.* (1998) *Genes Dev.* 12(14):2175-87; Goddemeier *et al.* (1998) *Plant Mol. Biol.* 36(5):799-802; and Yamamoto *et al.* (1991) *Plant Cell* 3(4):371-82. Though root-specific promoters have been characterized in several types of plants, no root specific promoters from maize have been described in the literature.

Constitutive expression of some heterologous proteins, such as insecticides, leads to undesirable phenotypic and agronomic effects. Limiting expression of insecticidal proteins, for example, to the target tissues of insect feeding (root, in this case), allows the plant to devote more energy to normal growth rather than toward expression of the protein throughout the plant. Using root-preferred promoters, one can also limit expression of the protein in non-desirable portions of the plant. However, many of the root-preferred promoters that have been isolated do not direct the expression of sufficient amounts of transgene for efficacy in plants. Thus, the isolation and characterization of tissue-preferred, particularly root-preferred, promoters that can direct transcription of a sufficiently high level of a desired heterologous nucleotide sequence is needed.

Since the patterns of expression of one or more chimeric genes introduced into a plant are controlled using promoters, there is an ongoing interest in the isolation and identification of novel promoters which are capable of controlling expression of chimeric gene(s).

SUMMARY OF THE INVENTION

Compositions and methods for regulating gene expression in a plant are provided. Compositions comprise novel nucleotide sequences for a promoter that initiates transcription in a root-preferred manner. More particularly, a transcriptional initiation region isolated from maize is provided. Further embodiments of the invention comprise the nucleotide sequence set forth in SEQ ID NO:1, a fragment of the nucleotide sequence set forth in SEQ ID NO:1, and the plant promoter sequences deposited in bacterial hosts as Patent Deposit No. NRRL B-30794. The compositions of the embodiments further comprise

nucleotide sequences having at least 95% sequence identity to the sequences set forth in SEQ ID NO:1, and which drive root-preferred expression of an operably linked nucleotide sequence. Also included are functional fragments of the sequence set forth as SEQ ID NO:1 which drive root-preferred expression of an operably linked nucleotide sequence.

Compositions of the embodiments also include DNA constructs comprising a promoter of the embodiments operably linked to a heterologous nucleotide sequence of interest wherein the promoter is capable of driving expression of the nucleotide sequence of interest in a plant cell and also wherein the promoter comprises the nucleotide sequences of the embodiments. The embodiments further provide expression vectors, and plants or plant cells having stably incorporated into their genomes a DNA construct mentioned above. Additionally, compositions include transgenic seed of such plants.

Methods of the embodiments comprise a means for selectively expressing a nucleotide sequence in a plant root, comprising transforming a plant cell with a DNA construct, and regenerating a transformed plant from the transformed plant cell, wherein the DNA construct comprises a promoter and a heterologous nucleotide sequence operably linked to the promoter, further wherein the promoter initiates root-preferred transcription of the nucleotide sequence in a plant cell. In this manner, the promoter sequences are useful for controlling the expression of operably linked coding sequences in a root-preferred manner.

Downstream from and under the transcriptional initiation regulation of the promoter will be a sequence of interest that will provide for modification of the phenotype of the plant. Such modification includes modulating the production of an endogenous product, as to amount, relative distribution, or the like, or production of an exogenous expression product to provide for a novel function or product in the plant. For example, a heterologous nucleotide sequence that encodes a gene product that confers herbicide, salt, cold, drought, pathogen or insect resistance is encompassed.

In a further aspect, methods of the embodiments relate to a method for modulating expression of a gene in the root of a stably transformed plant comprising the steps of (a) transforming a plant cell with an DNA construct comprising the promoter of the embodiments operably linked to at least one

nucleotide sequence; (b) growing the plant cell under plant growing conditions and (c) regenerating a stably transformed plant from the plant cell wherein expression of the nucleotide sequence alters the phenotype of the plant.

Embodiments include a cyclotide sequence which finds use in enhancing
5 the plant pathogen defense system. Other embodiments are directed to a cyclizable molecule and its linear precursor; cyclic peptides, polypeptides or proteins; and additionally includes the linear forms of non-cyclic structural homologues of the cyclic peptides, polypeptides and proteins. Also included are derivative forms of the cyclized molecule and their linear precursors encoded by
10 the subject nucleic acid molecules. The cyclic and linear peptides, polypeptides or proteins may be naturally occurring or may be modified by the insertion or substitution of heterologous amino acid sequences.

Embodiments also include a method involving stably transforming a plant with a cyclotide nucleotide sequence capable of modulating the plant pathogen
15 defense system operably linked with a promoter capable of driving expression of a gene in a plant cell. Transformed plants, plant cells, and seeds, as well as methods for making such plants, plant cells, and seeds, are additionally provided.

Embodiments of the invention concern an isolated polynucleotide comprising a nucleotide sequence set forth in SEQ ID NO: 11 ; a nucleotide
20 sequence that encodes a polypeptide having the amino acid sequence set forth in SEQ ID NOs: 12 and 13, a nucleotide sequence characterized by at least 85% sequence identity to the nucleotide sequences set forth in SEQ ID NO: 11; a nucleotide sequence characterized by at least 90% sequence identity to the nucleotide sequences set forth in SEQ ID NO: 11; a nucleotide sequence
25 characterized by at least 95% sequence identity to the nucleotide sequences set forth in SEQ ID NO: 11; and a nucleotide sequence that comprises the complement of any one of the above. A further embodiment is the complement of the nucleotide sequences disclosed herein.

An embodiment of the invention also provides an isolated polypeptide
30 selected from the group consisting of: a polypeptide comprising an amino acid sequence set forth in SEQ ID NOs: 12 and 13; a polypeptide characterized by at least 90% identity to SEQ ID NOs: 12 and 13; a polypeptide characterized by at least 95% identity to SEQ ID NOs: 12 and 13; a polypeptide characterized by at

least 97% identity to SEQ ID NOs: 12 and 13; a polypeptide characterized by at least 98% identity to SEQ ID NOs: 12 and 13; and a polypeptide characterized by at least 99% identity to SEQ ID NOs: 12 and 13.

5

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows the Lynx MPSS profile of the Cyclo1 gene which demonstrates root preferential expression.

Figure 2 shows the sequence of the maize Cyclo1 promoter. The positions of the TATA box, the transcriptional start site (TSS) mapped by 5' RACE, and other motifs of interest in the promoter sequence are indicated.

10

Figure 3 is an alignment of the cyclotide polypeptide encoded by the Cyclo1 gene compared to other known cyclotides. Identical residues are shown in bold face type. Residues which are similar and considered to be conservative substitutions are shown in italic type.

15

DETAILED DESCRIPTION OF THE INVENTION

The compositions of the embodiments comprise novel nucleotide sequences for plant promoters, particularly a root-preferred promoter for a maize Cyclo1 gene, more particularly, the Cyclo1 gene promoter. In particular, the embodiments provide for isolated nucleic acid molecules comprising the nucleotide sequence set forth in SEQ ID NO:1, and the plant promoter sequence deposited in bacterial hosts as Patent Deposit No: NRRL B-30794, and fragments, variants, and complements thereof.

20

Plasmids containing the plant promoter nucleotide sequences of the embodiments were deposited on December 1, 2004, with the Patent Depository of the Agricultural Research Service Culture Collection of the National Center for Agricultural Utilization Research, at 1815 N. University Street, Peoria, IL, 61604, and assigned Patent Deposit No. NRRL B-30794. 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

25

30

admission that a deposit is required under 35 U.S.C. §112. The deposit will irrevocably and without restriction or condition be available to the public upon issuance of a patent. However, it should be understood that the availability of a deposit does not constitute a license to practice the subject invention in derogation of patent rights granted by government action.

The promoter sequences of the embodiments are useful for expressing operably linked nucleotide sequences in a tissue-preferred, particularly a root-preferred manner. Therefore, the Cyclo1 promoter sequences find use in the root-preferred expression of an operably linked nucleotide sequence of interest. The sequences of the embodiments also find use in the construction of expression vectors for subsequent transformation into plants of interest, as probes for the isolation of other Cyclo1 gene promoters, as molecular markers, and the like.

The Cyclo1 promoter of the embodiments was isolated from maize genomic DNA. The specific method used to obtain the Cyclo1 promoter of the embodiments is described in Example 3 in the experimental section of this application.

The embodiments encompass isolated or substantially purified nucleic acid compositions. An "isolated" or "purified" nucleic acid molecule, 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. Generally, an "isolated" nucleic acid is free of sequences (for example, protein encoding 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.

The Cyclo1 gene (SEQ ID NO: 11) may be a defense-related gene. The predicted amino acid sequence of the Cyclo1 protein product (SEQ ID NOs: 12 and 13) shows homology to a class of defense-related proteins with a diverse set of activities including antimicrobial and insecticidal functions. This class is broadly known as cyclotides, which have been identified in several plant species (see US

Provisional Patent Applications 60/575,571 and 60/616,190). The Cyclo1 protein contains a similar conserved cysteine framework with additional conserved amino acid residues recognized to be a signature of this group of cyclic peptides. The maize Cyclo1 gene is preferentially expressed in maize root tissue as
5 demonstrated by gene tissue profile comparisons derived from Lynx Massively Parallel Signature Sequencing (MPSS), as further discussed in Example 1.

Small cysteine-rich proteins that have been implicated in host defense and isolated from plant sources include defensins, thionins, and small antimicrobial proteins (AMP's). Cyclotides, also cysteine-rich molecules, have been recognized
10 and characterized as being involved in host defense (Craig *et al.* (1999), *J. Mol. Biol.* 294: 1327-1336; Craig *et al.* (2000), *Toxicon* 39: 43-60). Cyclotide polypeptides are encoded by gene sequences, are produced as linear precursors, are cysteine-rich, and are capable of being cyclized via a peptide bond. Cyclotides display a diverse range of biological activities such as antibacterial
15 activity, antifungal activity, anti-HIV activity, and uterotonic activity (Craig (2001), *Toxicon* 39: 1809-1813). Cyclotides have additionally been shown to possess insecticidal activity (Jennings *et al.* (2001) *Proc. Natl. Acad. Sci. U.S.A.* 98:10614-10619). Cyclized cyclotides differ from classical proteins in that they have no free N- or C- terminus due to their amide-circularized backbone.

20 Cyclotide polypeptides are derived from longer precursor proteins and thus both cleavage and cyclization steps are involved in the production of the cyclic backbone. The cyclic backbone of the cyclotide molecule typically ranges in size from 29 to 37 amino acid residues and has three disulfide bonds that form a cystine knot motif where two disulfide bonds and their connecting backbone
25 strands form a ring that is threaded by the third disulfide bond. The mechanism(s) inherent to backbone cyclization is currently not known. One possibility is enzymatic or chemical involvement in both the backbone cleavage of the mature domain and the subsequent cyclization. The combined features of the cyclic cystine knot produces a unique protein fold that is topologically complex and has
30 exceptional chemical and biological stability.

The majority of the plant cyclotides have been isolated from Rubiaceae and Violaceae plants (Gustafson *et al.* (1994), *J. Nat. Prod.* 116: 9337-9338; Gustafson *et al.* (2000), *J. Nat. Prod.* 63: 176-178; Witherup *et al.* (1994), *J. Nat.*

Prod. 57: 1619-1625; Saether *et al.* (1995), *Biochemistry* 34, 4147-4158; Bokesch *et al.* (2001), *J. Nat. Prod.* 64: 249-250; Schöpke *et al.* (1993), *Sci. Pharm.* 61: 145-153; Claeson *et al.* (1998), *J. Nat. Prod.* 61: 77-81; Göransson *et al.* (1999), *J. Nat. Prod.* 62: 283-286; Hallock *et al.* (2000), *J. Org. Chem.* 65: 124-128; Broussalis *et al.* (2001), *Phytochemistry* 58: 47-51). Recently, two members of a new sub-class of the cyclotide family have been discovered in Curcubitaceae (Hernandez *et al.* (2000), *Biochemistry* 39: 5722-5730.; Felizmenio-Quimio *et al.* (2001), *J. Biol. Chem.* 276: 22875-22882; Heitz *et al.* (2001), *Biochemistry* 40: 7973-7983; Trabi and Craik, (2002), *Trends in Biochem. Sci.* 27: 132-138).

Cyclotides may be used in transgenic plants in order to produce plants with increased resistance to pathogens such as fungi, viruses, bacteria, nematodes, and insects. Thus, embodiments of the present invention may be used for the enhancement of a plant's defensive response via a molecularly based mechanism which can be quickly incorporated into commercial crops.

The compositions of the embodiments include isolated nucleic acid molecules comprising the promoter nucleotide sequence set forth in SEQ ID NO:1. The term "promoter" is intended to mean a regulatory region of DNA usually comprising a TATA box capable of directing RNA polymerase II to initiate RNA synthesis at the appropriate transcription initiation site for a particular coding sequence. A promoter may additionally comprise other recognition sequences generally positioned upstream or 5' to the TATA box, referred to as upstream promoter elements, which influence the transcription initiation rate. It is recognized that having identified the nucleotide sequences for the promoter regions disclosed herein, it is within the state of the art to isolate and identify further regulatory elements in the 5' untranslated region upstream from the particular promoter regions identified herein. Thus, for example, the promoter regions disclosed herein may further comprise upstream regulatory elements such as those responsible for tissue and temporal expression of the coding sequence, enhancers, and the like. See particularly Australian Patent No. AU-A-77751/94 and U.S. Patent Nos. 5,466,785 and 5,635,618. In the same manner, the promoter elements that enable expression in the desired tissue such as the root, can be identified, isolated, and used with other core promoters to confer root-

preferred expression. In this aspect of the embodiments, a "core promoter" is intended to mean a promoter without promoter elements.

In the context of this disclosure, the term "regulatory element" also refers to a sequence of DNA, usually, but not always, upstream (5') to the coding sequence of a structural gene, which includes sequences which control the expression of the coding region by providing the recognition for RNA polymerase and/or other factors required for transcription to start at a particular site. An example of a regulatory element that provides for the recognition for RNA polymerase or other transcriptional factors to ensure initiation at a particular site is a promoter element. A promoter element comprises a core promoter element, responsible for the initiation of transcription, as well as other regulatory elements (as discussed elsewhere in this application) that modify gene expression. It is to be understood that nucleotide sequences, located within introns, or 3' of the coding region sequence may also contribute to the regulation of expression of a coding region of interest. Examples of suitable introns include, but are not limited to, the maize IVS6 intron, or the maize actin intron. A regulatory element may also include those elements located downstream (3') to the site of transcription initiation, or within transcribed regions, or both. In the context of the embodiments, a post-transcriptional regulatory element may include elements that are active following transcription initiation, for example translational and transcriptional enhancers, translational and transcriptional repressors, and mRNA stability determinants.

The regulatory elements, or fragments thereof, of the embodiments may be operatively associated with heterologous regulatory elements or promoters in order to modulate the activity of the heterologous regulatory element. Such modulation includes: (a) enhancing or repressing transcriptional activity of the heterologous regulatory element; (b) modulating post-transcriptional events, or both (a) and (b). For example, one or more regulatory elements, or fragments thereof, of the embodiments may be operatively associated with constitutive, inducible, or tissue specific promoters or fragments thereof, to modulate the activity of such promoters within desired tissues within plant cells.

The maize Cyclo1 root-preferred promoter sequence of the embodiments, when assembled within a DNA construct such that the promoter is operably linked to a nucleotide sequence of interest, enables expression of the nucleotide

sequence in the cells of a plant stably transformed with this DNA construct. The term "operably linked" is intended to mean that the transcription or translation of the heterologous nucleotide sequence is under the influence of the promoter sequence. "Operably linked" is also intended to mean the joining of two nucleotide sequences such that the coding sequence of each DNA fragment remain in the proper reading frame. In this manner, the nucleotide sequences for the promoters of the embodiments are provided in DNA constructs along with the nucleotide sequence of interest, typically a heterologous nucleotide sequence, for expression in the plant of interest. The term "heterologous nucleotide sequence" is intended to mean a sequence that is not naturally operably linked with the promoter sequence. While this nucleotide sequence is heterologous to the promoter sequence, it may be homologous (native) or heterologous (foreign), to the plant host.

The regulatory sequences of the embodiments, when operably linked to a heterologous nucleotide sequence of interest and stably incorporated into the plant genome drive "root-preferred" expression of the heterologous nucleotide sequence. The term, "root-preferred" is intended to mean that expression of the heterologous nucleotide sequence is most abundant in the root. The term "root" is intended to mean any part of the root structure, including but not limited to, the root cap, apical meristem, protoderm, ground meristem, procambium, endodermis, cortex, vascular cortex, epidermis, and the like. While some level of expression of the heterologous nucleotide sequence may occur in other plant tissue types, expression occurs most abundantly in the root; which may include, but is not limited to primary, lateral, and adventitious roots.

It is recognized that the promoters of the embodiments thereof may be used with their native coding sequences to increase or decrease expression, thereby resulting in a change in phenotype of the transformed plant.

Modifications of the isolated promoter sequences of the embodiments can provide for a range of expression of the heterologous nucleotide sequence. Thus, they may be modified to be weak promoters or strong promoters. Generally, a "weak promoter" is intended to mean a promoter that drives expression of a coding sequence at a low level. A "low level" of expression is intended to mean expression at levels of about 1/10,000 transcripts to about 1/100,000 transcripts to

about 1/500,000 transcripts. Conversely, a strong promoter drives expression of a coding sequence at a high level, or at about 1/10 transcripts to about 1/100 transcripts to about 1/1,000 transcripts.

Fragments and variants of the disclosed promoter sequences are also encompassed by the embodiments. A "fragment" is intended to mean a portion of the promoter sequence. Fragments of a promoter sequence may retain biological activity and hence encompass fragments capable of driving root-preferred expression of an operably linked nucleotide sequence. Thus, for example, less than the entire promoter sequence disclosed herein may be utilized to drive expression of an operably linked nucleotide sequence of interest, such as a nucleotide sequence encoding a heterologous protein. It is within skill in the art to determine whether such fragments decrease expression levels or alter the nature of expression, i.e., constitutive or inducible expression. Alternatively, fragments of a promoter nucleotide sequence that are useful as hybridization probes, such as described below, generally do not retain this regulatory activity. Thus, fragments of a nucleotide sequence may range from at least about 20 nucleotides, about 50 nucleotides, about 100 nucleotides, and up to the full-length nucleotide sequence of the embodiments.

Thus, a fragment of a Cyclo1 promoter nucleotide sequence may encode a biologically active portion of the Cyclo1 promoter 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 a Cyclo1 promoter can be prepared by isolating a portion of the Cyclo1 promoter nucleotide sequence of the embodiments and assessing the activity of that portion of the Cyclo1 promoter. Nucleic acid molecules that are fragments of a promoter nucleotide sequence comprise at least 15, 20, 25, 30, 35, 40, 45, 50, 75, 100, 325, 350, 375, 400, 425, 450, 500, 550, 600, 650, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, or up to the number of nucleotides present in the full-length promoter nucleotide sequence disclosed herein, e.g. 1140 nucleotides for SEQ ID NO:1.

The nucleotides of such fragments will usually comprise the TATA recognition sequence of the particular promoter sequence. Such fragments may be obtained by use of restriction enzymes to cleave the naturally occurring promoter nucleotide sequence disclosed herein; by synthesizing a nucleotide

sequence from the naturally occurring sequence of the promoter DNA sequence; or may be obtained through the use of PCR technology. See particularly, Mullis *et al.* (1987) *Methods Enzymol.* 155:335-350, and Erlich, ed. (1989) *PCR Technology* (Stockton Press, New York). Variants of these promoter fragments, such as those resulting from site-directed mutagenesis and a procedure such as DNA "shuffling," are also encompassed by the compositions of the embodiments.

An "analogue" of the regulatory elements of the embodiments includes any substitution, deletion, or addition to the sequence of a regulatory element provided that said analogue maintains at least one regulatory property associated with the activity of the regulatory element of the embodiments. Such properties include directing organ specificity, tissue specificity, temporal activity, developmental activity, or any combination thereof.

The term "variants" is intended to mean sequences having substantial similarity with a promoter sequence disclosed herein. For nucleotide sequences, naturally occurring 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. Generally, variants of a particular nucleotide sequence of the embodiments will have at least 40%, 50%, 60%, 65%, 70%, generally at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, to 95%, 96%, 97%, 98%, 99% or more sequence identity to that particular nucleotide sequence as determined by sequence alignment programs described elsewhere herein using default parameters. Biologically active variants are also encompassed by the embodiments. Biologically active variants include, for example, the native promoter sequence of the embodiments having one or more nucleotide substitutions, deletions, or insertions. Promoter activity may be measured by using techniques such as northern blot analysis, reporter activity measurements taken from transcriptional fusions, and the like. See, for example, Sambrook *et al.* (1989) *Molecular Cloning: A Laboratory Manual* (2d ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York), hereinafter "Sambrook," herein incorporated by reference. Alternatively, levels of a reporter gene such as green fluorescent protein (GFP) or the like produced under the

control of a promoter fragment or variant can be measured. See, for example, U.S. Patent No. 6,072,050, herein incorporated by reference.

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; U.S. Patent No. 4,873,192; Walker and Gaastra, eds. (1983) *Techniques in Molecular Biology* (MacMillan Publishing Company, New York) and the references cited therein.

Variant promoter nucleotide sequences also encompass sequences derived from a mutagenic and recombinogenic procedure such as DNA shuffling. With such a procedure, one or more different promoter sequences can be manipulated to create a new promoter 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*. 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; Cramer *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; Cramer *et al.* (1998) *Nature* 391:288-291; and U.S. Patent Nos. 5,605,793 and 5,837,458.

The nucleotide sequences of the embodiments can be used to isolate corresponding sequences from other organisms, particularly other plants, for example, 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 sequence set forth herein. Sequences isolated based on their sequence identity to the entire Cyclo1 promoter sequence set forth herein or to fragments thereof are encompassed by the embodiments. The promoter regions of the embodiments may be isolated from any plant, including, but not limited to, corn (*Zea mays*), *Brassica* (*Brassica napus*, *Brassica rapa ssp.*), alfalfa (*Medicago sativa*), rice (*Oryza sativa*), rye (*Secale cereale*), sorghum (*Sorghum bicolor*, *Sorghum vulgare*), sunflower (*Helianthus annuus*), wheat (*Triticum aestivum*), soybean (*Glycine max*), tobacco (*Nicotiana tabacum*), potato (*Solanum tuberosum*), peanut (*Arachis hypogaea*), cotton (*Gossypium hirsutum*), sweet

potato (*Ipomoea batatas*), cassava (*Manihot esculenta*), coffee (*Cofea 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
5 (*Mangifera indica*), olive (*Olea europaea*), oats, safflower, barley, vegetables,
ornamentals, and conifers.

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 of interest. Methods for designing PCR primers and
10 PCR cloning are generally known in the art and are disclosed in Sambrook, *supra*.
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,
15 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
20 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 ³²P, or any other detectable marker. Thus, for example, probes for
25 hybridization can be made by labeling synthetic oligonucleotides based on the
Cyclo1 promoter sequence of the embodiments. 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, *supra*.

For example, the entire Cyclo1 promoter sequence disclosed herein, or one
30 or more portions thereof, may be used as a probe capable of specifically
hybridizing to corresponding Cyclo1 promoter sequences. To achieve specific
hybridization under a variety of conditions, such probes include sequences that
are unique among Cyclo1 promoter sequences and are at least about 10

nucleotides in length, and generally at least about 20 nucleotides in length. Such probes may be used to amplify corresponding Cyclo1 promoter sequences from a chosen plant by PCR. This technique may be used to isolate additional coding sequences from a desired plant or as a diagnostic assay to determine the presence of coding sequences in a plant. Hybridization techniques include hybridization screening of plated DNA libraries (either plaques or colonies; see, for example, Sambrook *supra*).

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, often 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 final wash in 0.1X SSC at 60 to 65 °C for at least 30 minutes. 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 thermal melting point (T_m) can be approximated from the equation of Meinkoth and Wahl (1984) *Anal. Biochem.* 138:267-284: $T_m = 81.5\text{ }^\circ\text{C} + 16.6 (\log M) + 0.41 (\%GC) - 0.61 (\% \text{ form}) - 500/L$; where M is the molarity of monovalent cations, %GC is the percentage of guanosine and cytosine nucleotides in the DNA, % 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 $\geq 90\%$ identity are sought, the T_m can be decreased $10\text{ }^\circ\text{C}$. Generally, stringent conditions are selected to be about $5\text{ }^\circ\text{C}$ lower than the 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 $4\text{ }^\circ\text{C}$ lower than the T_m ; moderately stringent conditions can utilize a hybridization and/or wash at 6, 7, 8, 9, or $10\text{ }^\circ\text{C}$ lower than the T_m ; low stringency conditions can utilize a hybridization and/or wash at 11, 12, 13, 14, 15, or $20\text{ }^\circ\text{C}$ lower than the 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 $45\text{ }^\circ\text{C}$ (aqueous solution) or $32\text{ }^\circ\text{C}$ (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), hereinafter "Ausubel". See also Sambrook *supra*.

Thus, isolated sequences that have root-preferred promoter activity and which hybridize under stringent conditions to the Cyclo1 promoter sequences disclosed herein, or to fragments thereof, are encompassed by the embodiments.

In general, sequences that have promoter activity and hybridize to the promoter sequences disclosed herein will be at least 40% to 50% homologous, about 60% to 70% homologous, and even about 80%, 85%, 90%, 95% to 98% homologous or more with the disclosed sequences. That is, the sequence
5 similarity of sequences may range, sharing at least about 40% to 50%, about 60% to 70%, and even about 80%, 85%, 90%, 95% to 98% sequence similarity.

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
10 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.

15 (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.
20 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.

25 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. Preferred, 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)
30 *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
5 program (available from Intelligenetics, Mountain View, California); the ALIGN program (Version 2.0); the ALIGN PLUS program (Version 3.0, copyright 1997); and GAP, BESTFIT, BLAST, FASTA, and TFASTA in the Wisconsin Genetics Software Package of Genetics Computer Group, Version 10 (available from
10 Accelrys, 9685 Scranton Road, San Diego, CA, 92121, USA). The scoring matrix used in Version 10 of the Wisconsin Genetics Software Package is BLOSUM62 (see Henikoff and Henikoff (1989) *Proc. Natl. Acad. Sci. USA* 89:10915).

Alignments using these programs can be performed using the default parameters. The CLUSTAL program is well described by Higgins *et al.* (1988)
15 *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 and the ALIGN PLUS programs are 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
20 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
25 the embodiments. 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 embodiments. 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-
30 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 the web site for the National Center for Biotechnology Information on the world wide web. Alignment may also be performed manually by inspection.

Unless otherwise stated, sequence identity/similarity values provided herein
5 refer to the value obtained using the GAP program with default parameters, or
any equivalent program. By "equivalent program" is intended any sequence
comparison program that, for any two sequences in question, generates an
alignment having identical nucleotide or amino acid residue matches and an
identical percent sequence identity when compared to the corresponding
10 alignment generated by GAP.

The GAP program uses the algorithm of Needleman and Wunsch (1970)
supra, to find the alignment of two complete sequences that maximizes the
number of matches and minimizes the number of gaps. GAP considers all
possible alignments and gap positions and creates the alignment with the largest
15 number of matched bases and the fewest gaps. It allows for the provision of a gap
creation penalty and a gap extension penalty in units of matched bases. GAP
must make a profit of gap creation penalty number of matches for each gap it
inserts. If a gap extension penalty greater than zero is chosen, GAP must, in
addition, make a profit for each gap inserted of the length of the gap times the gap
20 extension penalty. Default gap creation penalty values and gap extension penalty
values in Version 10 of the Wisconsin Genetics Software Package for protein
sequences are 8 and 2, respectively. For nucleotide sequences the default gap
creation penalty is 50 while the default gap extension penalty is 3. The gap
creation and gap extension penalties can be expressed as an integer selected
25 from the group of integers consisting of from 0 to 200. Thus, for example, the gap
creation and gap extension penalties can be 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20,
25, 30, 35, 40, 45, 50, 55, 60, 65 or greater.

(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
30 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, at least 80%, at least 90%, and 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%, 70%, 80%, 90%, or 95%.

Another indication that nucleotide sequences are substantially identical is if two molecules hybridize to each other under stringent conditions. Generally, 5 stringent conditions are selected to be about 5 °C lower than the 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 lower than the T_m , depending upon the desired degree of stringency as otherwise qualified herein. Nucleic acids that do not hybridize to each other under stringent 10 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 15 with the polypeptide encoded by the second nucleic acid.

The Cyclo1 promoter sequence disclosed herein, as well as variants and fragments thereof, are useful for genetic engineering of plants, e.g. for the production of a transformed or transgenic plant, to express a phenotype of interest. As used herein, the terms "transformed plant" and "transgenic plant" 20 refer to a plant that comprises within its genome a heterologous polynucleotide. Generally, the heterologous polynucleotide is stably integrated within the genome of a transgenic or transformed plant such that the polynucleotide is passed on to successive generations. The heterologous polynucleotide may be integrated into the genome alone or as part of a recombinant DNA construct. It is to be 25 understood that as used herein the term "transgenic" includes any cell, cell line, callus, tissue, plant part, or plant the genotype of which has been altered by the presence of heterologous nucleic acid including those transgenics initially so altered as well as those created by sexual crosses or asexual propagation from the initial transgenic. The term "transgenic" as used herein does not encompass 30 the alteration of the genome (chromosomal or extra-chromosomal) by conventional plant breeding methods or by naturally occurring events such as random cross-fertilization, non-recombinant viral infection, non-recombinant bacterial transformation, non-recombinant transposition, or spontaneous mutation.

A transgenic "event" is produced by transformation of plant cells with a heterologous DNA construct, including a nucleic acid expression cassette that comprises a transgene of interest, the regeneration of a population of plants resulting from the insertion of the transgene into the genome of the plant, and
5 selection of a particular plant characterized by insertion into a particular genome location. An event is characterized phenotypically by the expression of the transgene. At the genetic level, an event is part of the genetic makeup of a plant. The term "event" also refers to progeny produced by a sexual outcross between the transformant and another variety that include the heterologous DNA.

10 As used herein, the term "plant" includes reference to whole plants, plant organs (e.g., leaves, stems, roots, etc.), seeds, plant cells, and progeny of same. Parts of transgenic plants are to be understood within the scope of the embodiments to comprise, for example, plant cells, protoplasts, tissues, callus, embryos as well as flowers, stems, fruits, ovules, leaves, or roots originating in
15 transgenic plants or their progeny previously transformed with a DNA molecule of the embodiments, and therefore consisting at least in part of transgenic cells, are also encompassed by the embodiments.

As used herein, the term "plant cell" includes, without limitation, seeds, suspension cultures, embryos, meristematic regions, callus tissue, leaves, roots,
20 shoots, gametophytes, sporophytes, pollen, and microspores. The class of plants that can be used in the methods of the embodiments is generally as broad as the class of higher plants amenable to transformation techniques, including both monocotyledonous and dicotyledonous plants.

The promoter sequences and methods disclosed herein are useful in
25 regulating expression of any heterologous nucleotide sequence in a host plant. Thus, the heterologous nucleotide sequence operably linked to the promoters disclosed herein may be a structural gene encoding a protein of interest. Genes of interest are reflective of the commercial markets and interests of those involved in the development of the crop. Crops and markets of interest change, and as
30 developing nations open up world markets, new crops and technologies will emerge also. In addition, as our understanding of agronomic traits and characteristics such as yield and heterosis increase, the choice of genes for transformation will change accordingly. General categories of genes of interest for

the embodiments include, for example, those genes involved in information, such as zinc fingers, those involved in communication, such as kinases, and those involved in housekeeping, such as heat shock proteins. More specific categories of transgenes, for example, include genes encoding proteins conferring resistance to abiotic stress, such as drought, temperature, salinity, and toxins such as pesticides and herbicides, or to biotic stress, such as attacks by fungi, viruses, bacteria, insects, and nematodes, and development of diseases associated with these organisms. Various changes in phenotype are of interest including modifying expression of a gene in a plant root, altering a plant's pathogen or insect defense mechanism, increasing the plant's tolerance to herbicides, altering root development to respond to environmental stress, and the like. The results can be achieved by providing expression of heterologous or increased expression of endogenous products in plants. Alternatively, the results can be achieved by providing for a reduction of expression of one or more endogenous products, particularly enzymes, transporters, or cofactors, or affecting nutrients uptake in the plant. These changes result in a change in phenotype of the transformed plant. It is recognized that any gene of interest can be operably linked to the promoter sequences of the embodiments and expressed in a plant root.

A DNA construct comprising one of these genes of interest can be used with transformation techniques, such as those described below, to create disease or insect resistance in susceptible plant phenotypes or to enhance disease or insect resistance in resistant plant phenotypes. Accordingly, the embodiments encompass methods that are directed to protecting plants against fungal pathogens, bacteria, viruses, nematodes, insects, and the like. By "disease resistance" or "insect resistance" is intended that the plants avoid the harmful symptoms that are the outcome of the plant-pathogen interactions.

Disease resistance and insect resistance genes such as lysozymes, cecropins, maganins, or thionins for antibacterial protection, or the pathogenesis-related (PR) proteins such as glucanases and chitinases for anti-fungal protection, or *Bacillus thuringiensis* endotoxins, protease inhibitors, collagenases, lectins, and glycosidases for controlling nematodes or insects are all examples of useful gene products.

Pathogens of the embodiments include, but are not limited to, viruses or viroids, bacteria, insects, nematodes, fungi, and the like. Viruses include tobacco or cucumber mosaic virus, ringspot virus, necrosis virus, maize dwarf mosaic virus, etc. Nematodes include parasitic nematodes such as root knot, cyst, and
5 lesion nematodes, etc.

Insect resistance genes may encode resistance to pests that have great yield drag such as rootworm, cutworm, European corn borer, and the like. Such genes include, for example, *Bacillus thuringiensis* toxic protein genes (U.S. Pat. Nos. 5,366,892; 5,747,450; 5,737,514; 5,723,756; 5,593,881; and Geiser *et al.*
10 (1986) *Gene* 48:109); lectins (Van Damme *et al.* (1994) *Plant Mol. Biol.* 24:825); and the like.

Genes encoding disease resistance traits include detoxification genes, such as against fumonisin (U.S. Patent No. 5,792,931) avirulence (*avr*) and disease resistance (*R*) genes (Jones *et al.* (1994) *Science* 266:789; Martin *et al.* (1993)
15 *Science* 262:1432; Mindrinos *et al.* (1994) *Cell* 78:1089); and the like.

Herbicide resistance traits may be introduced into plants by genes coding for resistance to herbicides that act to inhibit the action of acetolactate synthase (ALS), in particular the sulfonylurea-type herbicides (*e.g.*, the acetolactate
20 synthase (ALS) gene containing mutations leading to such resistance, in particular the S4 and/or Hra mutations), genes coding for resistance to herbicides that act to inhibit action of glutamine synthase, such as phosphinothricin or basta (*e.g.*, the *bar* gene), or other such genes known in the art. The *bar* gene encodes resistance to the herbicide basta, the *nptII* gene encodes resistance to the antibiotics kanamycin and geneticin, and the ALS gene encodes resistance to the
25 herbicide chlorsulfuron.

Glyphosate resistance is imparted by mutant 5-enol pyruvylshikimate-3-phosphate synthase (EPSPS) and *aroA* genes. See, for example, U.S. Patent No. 4,940,835 to Shah *et al.*, which discloses the nucleotide sequence of a form of EPSPS which can confer glyphosate resistance. U.S. Patent No. 5,627,061 also
30 describes genes encoding EPSPS enzymes. See also U.S. Patent Nos. 6,248,876; 6,040,497; 5,804,425; 5,633,435; 5,145,783; 4,971,908; 5,312,910; 5,188,642; 4,940,835; 5,866,775; 6,225,114; 6,130,366; 5,310,667; 4,535,060; 4,769,061; 5,633,448; 5,510,471; RE 36,449; RE 37,287; and 5,491,288; and

international publications WO 97/04103; WO 97/04114; WO 00/66746; WO 01/66704; WO 00/66747 and WO 00/66748, which are incorporated herein by reference for this purpose. Glyphosate resistance is also imparted to plants that express a gene that encodes a glyphosate oxido-reductase enzyme as described more fully in U.S. Patent Nos. 5,776,760 and 5,463,175, which are incorporated
5 herein by reference for this purpose. In addition glyphosate resistance can be imparted to plants by the over-expression of genes encoding glyphosate N-acetyltransferase. See, for example, U.S. Patent Application Serial Nos. 10/004,357; and 10/427,692.

10 Sterility genes can also be encoded in a DNA construct and provide an alternative to physical detasseling. Examples of genes used in such ways include male tissue-preferred genes and genes with male sterility phenotypes such as QM, described in U.S. Patent No. 5,583,210. Other genes include kinases and those encoding compounds toxic to either male or female gametophytic
15 development.

Commercial traits can also be encoded on a gene or genes that could increase for example, starch for ethanol production, or provide expression of proteins. Another important commercial use of transformed plants is the production of polymers and bioplastics such as described in U.S. Patent No.
20 5,602,321. Genes such as β -Ketothiolase, PHBase (polyhydroxybutyrate synthase), and acetoacetyl-CoA reductase (see Schubert *et al.* (1988) *J. Bacteriol.* 170:5837-5847) facilitate expression of polyhydroxyalkanoates (PHAs).

Agronomically important traits that affect quality of grain, such as levels and types of oils, saturated and unsaturated, quality and quantity of essential amino
25 acids, levels of cellulose, starch, and protein content can be genetically altered using the methods of the embodiments. Modifications include increasing content of oleic acid, saturated and unsaturated oils, increasing levels of lysine and sulfur, providing essential amino acids, and modifying starch. Hordothionin protein modifications in corn are described in U.S. Patent Nos. 5,990,389; 5,885,801;
30 5,885,802 and 5,703,049; herein incorporated by reference. Another example is lysine and/or sulfur rich seed protein encoded by the soybean 2S albumin described in U.S. Patent No. 5,850,016, and the chymotrypsin inhibitor from

barley, (Williamson *et al.* (1987) *Eur. J. Biochem.* 165:99-106), the disclosures of which are herein incorporated by reference.

Exogenous products include plant enzymes and products as well as those from other sources including prokaryotes and other eukaryotes. Such products
5 include enzymes, cofactors, hormones, and the like.

Examples of other applicable genes and their associated phenotype include the gene that encodes viral coat protein and/or RNA, or other viral or plant genes that confer viral resistance; genes that confer fungal resistance; genes that confer insect resistance; genes that promote yield improvement; and genes that provide
10 for resistance to stress, such as dehydration resulting from heat and salinity, toxic metal or trace elements, or the like.

In other embodiments, the Cyclo1 promoter sequences are operably linked to genes of interest that improve plant growth or increase crop yields under high plant density conditions. For example, the Cyclo1 promoter may be operably
15 linked to nucleotide sequences expressing agronomically important genes that result in improved primary or lateral root systems. Such genes include, but are not limited to, nutrient/water transporters and growth inducers. Examples of such genes, include but are not limited to, maize plasma membrane H⁺-ATPase (MHA2) (Frias *et al.* (1996) *Plant Cell* 8:1533-44); AKT1, a component of the
20 potassium uptake apparatus in *Arabidopsis* (Spalding *et al.* (1999) *J. Gen. Physiol.* 113:909-18); RML genes, which activate cell division cycle in the root apical cells (Cheng *et al.* (1995) *Plant Physiol.* 108:881); maize glutamine synthetase genes (Sukanya *et al.* (1994) *Plant Mol. Biol.* 26:1935-46); and hemoglobin (Duff *et al.* (1997) *J. Biol. Chem.* 27:16749-16752; Arredondo-Peter *et al.* (1997) *Plant*
25 *Physiol.* 115:1259-1266; Arredondo-Peter *et al.* (1997) *Plant Physiol.* 114:493-500 and references cited therein). The Cyclo1 promoter may also be useful in expressing antisense nucleotide sequences of genes that negatively affect root development under high-planting density conditions.

"RNAi" refers to a series of related techniques to reduce the expression of
30 genes (See for example U.S. Patent No. 6,506,559). Older techniques referred to by other names are now thought to rely on the same mechanism, but are given different names in the literature. These include "antisense inhibition," the production of antisense RNA transcripts capable of suppressing the expression of

the target protein, and "co-suppression" or "sense-suppression," which refer to the production of sense RNA transcripts capable of suppressing the expression of identical or substantially similar foreign or endogenous genes (U.S. Patent No. 5,231,020, incorporated herein by reference). Such techniques rely on the use of
5 constructs resulting in the accumulation of double stranded RNA with one strand complementary to the target gene to be silenced. The Cyclo1 promoter sequence of the embodiments, and its related biologically active fragments or variants disclosed herein, may be used to drive expression of constructs that will result in RNA interference including microRNAs and siRNAs.

10 The heterologous nucleotide sequence operably linked to the Cyclo1 promoter and its related biologically active fragments or variants disclosed herein may be an antisense sequence for a targeted gene. The terminology "antisense DNA nucleotide sequence" is intended to mean a sequence that is in inverse orientation to the 5'-to-3' normal orientation of that nucleotide sequence. When
15 delivered into a plant cell, expression of the antisense DNA sequence prevents normal expression of the DNA nucleotide sequence for the targeted gene. The antisense nucleotide sequence encodes an RNA transcript that is complementary to and capable of hybridizing to the endogenous messenger RNA (mRNA) produced by transcription of the DNA nucleotide sequence for the targeted gene.
20 In this case, production of the native protein encoded by the targeted gene is inhibited to achieve a desired phenotypic response. 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, for example, 70%, 80%, or 85% sequence identity to the corresponding
25 antisense 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. Thus, the promoter sequences disclosed herein may be operably linked to antisense DNA sequences to reduce or inhibit expression of a native
30 protein in the plant.

In one embodiment, DNA constructs will comprise a transcriptional initiation region comprising one of the promoter nucleotide sequences disclosed herein, or variants or fragments thereof, operably linked to a heterologous nucleotide

sequence whose expression is to be controlled by the promoter of the
embodiments. Such a DNA construct is provided with a plurality of restriction sites
for insertion of the nucleotide sequence to be under the transcriptional regulation
of the regulatory regions. The DNA construct may additionally contain selectable
5 marker genes.

The DNA construct will include in the 5'-3' direction of transcription, a
transcriptional initiation region (i.e., a root-preferred promoter of the
embodiments), translational initiation region, a heterologous nucleotide sequence
of interest, a translational termination region and, optionally, a transcriptional
10 termination region functional in the host organism. The regulatory regions (i.e.,
promoters, transcriptional regulatory regions, and translational termination
regions) and/or the polynucleotide of the embodiments may be native/analogous
to the host cell or to each other. Alternatively, the regulatory regions and/or the
polynucleotide of the embodiments may be heterologous to the host cell or to each
15 other. As used herein, "heterologous" in reference to a sequence is a sequence
that originates from a foreign species, or, if from the same species, is substantially
modified from its native form in composition and/or genomic locus by deliberate
human intervention. For example, a promoter operably linked to a heterologous
polynucleotide is from a species different from the species from which the
20 polynucleotide was derived, or, if from the same/analogous species, one or both
are substantially modified from their original form and/or genomic locus, or the
promoter is not the native promoter for the operably linked polynucleotide.

The optionally included termination region may be native with the
transcriptional initiation region, may be native with the operably linked
25 polynucleotide of interest, may be native with the plant host, or may be derived
from another source (i.e., foreign or heterologous) to the promoter, the
polynucleotide of interest, the host, or any combination thereof. Convenient
termination regions are available from the Ti-plasmid of *A. tumefaciens*, such as
the octopine synthase and nopaline synthase termination regions. See also
30 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 Acids*

Res. 15:9627-9639. In particular embodiments, the potato protease inhibitor II gene (PinII) terminator is used. See, for example, Keil *et al.* (1986) *Nucl. Acids Res.* 14:5641-5650; and An *et al.* (1989) *Plant Cell* 1:115-122, herein incorporated by reference in their entirety.

5 The DNA construct comprising a promoter sequence of the embodiments operably linked to a heterologous nucleotide sequence may also contain at least one additional nucleotide sequence for a gene to be cotransformed into the organism. Alternatively, the additional sequence(s) can be provided on another DNA construct.

10 Where appropriate, the heterologous nucleotide sequence whose expression is to be under the control of the promoter sequence of the embodiments and any additional nucleotide sequence(s) may be optimized for increased expression in the transformed plant. That is, these nucleotide sequences can be synthesized using plant preferred codons for improved
15 expression. Methods are available in the art for synthesizing plant-preferred nucleotide sequences. 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
20 in a cellular host. These include elimination of sequences encoding exon-intron splice site signals, spurious polyadenylation signals, transposon-like repeats, and other such well-characterized sequences that may be deleterious to gene expression. The G-C content of the heterologous nucleotide sequence may be adjusted to levels average for a given cellular host, as calculated by reference to
25 known genes expressed in the host cell. When possible, the sequence is modified to avoid predicted hairpin secondary mRNA structures.

 The DNA constructs may additionally contain 5' leader sequences. Such leader sequences can act to enhance translation. Translation leaders are known in the art and include: picornavirus leaders, for example, EMCV leader
30 (Encephalomyocarditis 5' noncoding region) (Elroy-Stein *et al.* (1989) *Proc. Nat. Acad. Sci. USA* 86:6126-6130); potyvirus leaders, for example, TEV leader (Tobacco Etch Virus) (Allison *et al.* (1986) *Virology* 154:9-20); MDMV leader (Maize Dwarf Mosaic Virus); 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) *Molecular Biology of RNA*, pages 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 Physiology* 84:965-968. Other methods known to enhance translation and/or mRNA stability can also be utilized, for example, introns, such as the maize Ubiquitin intron (Christensen and Quail (1996) *Transgenic Res.* 5:213-218; Christensen *et al.* (1992) *Plant Molecular Biology* 18:675-689) or the maize Adhl intron (Kyojuka *et al.* (1991) *Mol. Gen. Genet.* 228:40-48; Kyojuka *et al.* (1990) *Maydica* 35:353-357), and the like.

The DNA constructs of the embodiments can also include further enhancers, either translation or transcription enhancers, as may be required. These enhancer regions are well known to persons skilled in the art, and can include the ATG initiation codon and adjacent sequences. The initiation codon must be in phase with the reading frame of the coding sequence to ensure translation of the entire sequence. The translation control signals and initiation codons can be from a variety of origins, both natural and synthetic. Translational initiation regions may be provided from the source of the transcriptional initiation region, or from the structural gene. The sequence can also be derived from the regulatory element selected to express the gene, and can be specifically modified so as to increase translation of the mRNA. It is recognized that to increase transcription levels enhancers may be utilized in combination with the promoter regions of the embodiments. Enhancers are known in the art and include the SV40 enhancer region, the 35S enhancer element, and the like.

In preparing the DNA construct, 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. Restriction sites may be added or removed, superfluous DNA may be removed, or other modifications of the like may be made to the sequences of the embodiments. For this purpose, *in vitro*

mutagenesis, primer repair, restriction, annealing, re-substitutions, for example, transitions and transversions, may be involved.

Reporter genes or selectable marker genes may be included in the DNA constructs. Examples of suitable reporter genes known in the art can be found in, for example, Jefferson *et al.* (1991) in *Plant Molecular Biology Manual*, ed. Gelvin *et al.* (Kluwer Academic Publishers), pp. 1-33; DeWet *et al.* (1987) *Mol. Cell. Biol.* 7:725-737; Goff *et al.* (1990) *EMBO J.* 9:2517-2522; Kain *et al.* (1995) *BioTechniques* 19:650-655; and Chiu *et al.* (1996) *Current Biology* 6:325-330.

Selectable marker genes for selection of transformed cells or tissues can include genes that confer antibiotic resistance or resistance to herbicides. Examples of suitable selectable marker genes include, but are not limited to, genes encoding resistance to chloramphenicol (Herrera Estrella *et al.* (1983) *EMBO J.* 2:987-992); methotrexate (Herrera Estrella *et al.* (1983) *Nature* 303:209-213; Meijer *et al.* (1991) *Plant Mol. Biol.* 16:807-820); hygromycin (Waldron *et al.* (1985) *Plant Mol. Biol.* 5:103-108; Zhijian *et al.* (1995) *Plant Science* 108:219-227); streptomycin (Jones *et al.* (1987) *Mol. Gen. Genet.* 210:86-91); spectinomycin (Bretagne-Sagnard *et al.* (1996) *Transgenic Res.* 5:131-137); bleomycin (Hille *et al.* (1990) *Plant Mol. Biol.* 7:171-176); sulfonamide (Guerineau *et al.* (1990) *Plant Mol. Biol.* 15:127-136); bromoxynil (Stalker *et al.* (1988) *Science* 242:419-423); glyphosate (Shaw *et al.* (1986) *Science* 233:478-481); phosphinothricin (DeBlock *et al.* (1987) *EMBO J.* 6:2513-2518).

Other genes that could serve utility in the recovery of transgenic events but might not be required in the final product would include, but are not limited to, examples such as GUS (b-glucuronidase; Jefferson (1987) *Plant Mol. Biol. Rep.* 5:387), GFP (green fluorescence protein; Chalfie *et al.* (1994) *Science* 263:802), luciferase (Riggs *et al.* (1987) *Nucleic Acids Res.* 15(19):8115 and Luehrsen *et al.* (1992) *Methods Enzymol.* 216:397-414), and the maize genes encoding for anthocyanin production (Ludwig *et al.* (1990) *Science* 247:449).

The nucleic acid molecules of the embodiments are useful in methods directed to expressing a nucleotide sequence in a plant. This may be accomplished by transforming a plant cell of interest with a DNA construct comprising a promoter identified herein, operably linked to a heterologous nucleotide sequence, and regenerating a stably transformed plant from said plant

cell. The methods of the embodiments are also directed to expressing a nucleotide sequence in a plant. Those methods comprise transforming a plant cell with a DNA construct comprising a promoter identified herein that initiates transcription in a plant cell in a root-preferred manner, operably linked to a
5 heterologous nucleotide sequence, and regenerating a transformed plant from said plant cell.

The DNA construct comprising the particular promoter sequence of the embodiments operably linked to a nucleotide sequence of interest can be used to transform any plant. In this manner, genetically modified, i.e. transgenic or
10 transformed, plants, plant cells, plant tissue, seed, root, and the like can be obtained.

Plant species suitable for the embodiments include, but are 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
15 (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*), peanut
20 (*Arachis hypogaea*), cotton (*Gossypium barbadense*, *Gossypium hirsutum*), sweet potato (*Ipomoea batatas*), cassava (*Manihot esculenta*), coffee (*Cofea* 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*),
25 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.,
30 *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 embodiments 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*). For example, plants of the embodiments may be crop plants (for example, corn, alfalfa, sunflower, *Brassica*, soybean, cotton, safflower, peanut, sorghum, wheat, millet, tobacco, etc.). This invention is, for example, suitable for any member of the monocot plant family including, but not limited to, maize, rice, barley, oats, wheat, sorghum, rye, sugarcane, pineapple, yams, onion, banana, coconut, and dates.

As used herein, "vector" refers to a DNA molecule such as a plasmid, cosmid, or bacterial phage for introducing a nucleotide construct, for example, a DNA construct, into a host cell. Cloning vectors typically contain one or a small number of restriction endonuclease recognition sites at which foreign DNA sequences can be inserted in a determinable fashion without loss of essential biological function of the vector, as well as a marker gene that is suitable for use in the identification and selection of cells transformed with the cloning vector. Marker genes typically include genes that provide tetracycline resistance, hygromycin resistance, or ampicillin resistance.

The methods of the embodiments involve introducing a nucleotide construct into a plant. By "introducing" is intended presenting to the plant the nucleotide construct in such a manner that the construct gains access to the interior of a cell of the plant. The methods of the embodiments do not depend on a particular method for introducing a nucleotide construct to a plant, only that the nucleotide construct gains access to the interior of at least one cell of the plant. Methods for introducing nucleotide constructs into plants are known in the art including, but not

limited to, stable transformation methods, transient transformation methods, and virus-mediated methods.

By "stable transformation" is intended that the nucleotide construct introduced into a plant integrates into the genome of the plant and is capable of
5 being inherited by progeny thereof. By "transient transformation" is intended that a nucleotide construct introduced into a plant does not integrate into the genome of the plant.

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

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
20 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 (U.S. Patent Nos. 5,981,840 and 5,563,055), direct gene transfer (Paszkowski *et al.* (1984) *EMBO J.* 3:2717-2722), and ballistic particle acceleration (see, for example, U.S. Patent Nos. 4,945,050;
25 5,879,918; 5,886,244; 5,932,782; Tomes *et al.* (1995) 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.*
30 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); U.S. Patent Nos. 5,240,855; 5,322,783 and 5,324,646; Klein *et al.* (1988) *Plant Physiol.* 91:440-444 (maize); Fromm *et al.* (1990) *Biotechnology* 8:833-839 (maize); Hooykaas-Van Slogteren *et al.* (1984) *Nature (London)* 311:763-764; 5 U.S. Patent No. 5,736,369 (cereals); Bytebier *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:5345-5349 (Liliaceae); De Wet *et al.* (1985) in *The Experimental Manipulation of Ovule Tissues*, ed. Chapman *et al.* (Longman, New York), pp. 197-209 (pollen); Kaeppler *et al.* (1990) *Plant Cell Reports* 9:415-418 and Kaeppler *et al.* (1992) *Theor. Appl. Genet.* 84:560-566 (whisker-mediated 10 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.

15 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 expression of the desired phenotypic characteristic identified. Two 20 or more generations may be grown to ensure that expression of the desired phenotypic characteristic is stably maintained and inherited and then seeds harvested to ensure expression of the desired phenotypic characteristic has been achieved. Thus as used herein, "transformed seeds" refers to seeds that contain the nucleotide construct stably integrated into the plant genome.

25 There are a variety of methods for the regeneration of plants from plant tissue. The particular method of regeneration will depend on the starting plant tissue and the particular plant species to be regenerated. The regeneration, development and cultivation of plants from single plant protoplast transformants or from various transformed explants is well known in the art (Weissbach and 30 Weissbach, (1988) In.: *Methods for Plant Molecular Biology*, (Eds.), Academic Press, Inc., San Diego, CA). This regeneration and growth process typically includes the steps of selection of transformed cells, culturing those individualized cells through the usual stages of embryonic development through the rooted

plantlet stage. Transgenic embryos and seeds are similarly regenerated. The resulting transgenic rooted shoots are thereafter planted in an appropriate plant growth medium such as soil. Preferably, the regenerated plants are self-pollinated to provide homozygous transgenic plants. Otherwise, pollen obtained from the regenerated plants is crossed to seed-grown plants of agronomically important lines. Conversely, pollen from plants of these important lines is used to pollinate regenerated plants. A transgenic plant of the embodiments containing a desired polypeptide is cultivated using methods well known to one skilled in the art.

The embodiments provide compositions for screening compounds that modulate expression within plants. The vectors, cells, and plants can be used for screening candidate molecules for agonists and antagonists of the Cyclo1 promoter. For example, a reporter gene can be operably linked to a Cyclo1 promoter and expressed as a transgene in a plant. Compounds to be tested are added and reporter gene expression is measured to determine the effect on promoter activity.

Embodiments of the invention are drawn to compositions and methods for impacting plant diseases and insect pests, particularly plant pests. More specifically, the isolated nucleic acids of the embodiments, and fragments and variants thereof, comprise nucleotide sequences that encode pesticidal polypeptides (*e.g.*, proteins). The disclosed pesticidal proteins are biologically active (*e.g.*, pesticidal) against insect pests.

Other embodiments include compositions comprising isolated nucleic acids, and fragments and variants thereof, that encode pesticidal polypeptides, DNA constructs comprising nucleotide sequences of embodiments of the invention, isolated pesticidal proteins, and pesticidal compositions. Other embodiments further provide plants and microorganisms transformed with these novel nucleic acids, and methods involving the use of such nucleic acids, pesticidal compositions, transformed organisms, and products thereof in impacting insect pests.

The nucleic acids and nucleotide sequences described herein may be used to transform any organism to produce the encoded pesticidal proteins. Methods are provided that involve the use of such transformed organisms to impact or control plant pests. The nucleic acids and nucleotide sequences may also be

used to transform organelles such as chloroplasts (McBride *et al.* (1995) *Biotechnology* 13:362-365; Kota *et al.* (1999) *Proc. Natl. Acad. Sci. U S A* 96: 1840-1845).

5 Some embodiments further provide fragments and variants of the naturally occurring coding sequences that also encode biologically active (*e.g.*, pesticidal) polypeptides. These nucleotide sequences find direct use in methods for impacting pests. Accordingly, embodiments of the invention provide new approaches for impacting insect pests that do not depend on the use of traditional, synthetic chemical pesticides. Some embodiments involve the discovery of
10 naturally-occurring, biodegradable pesticides and the genes that encode them.

Embodiments of the invention also encompass nucleic acid sequences that have been optimized for expression by the cells of a particular organism, for example nucleic acid sequences that have been back-translated (*i.e.*, reverse translated) using plant-preferred codons based on the amino acid sequence of a
15 polypeptide having enhanced pesticidal activity. Further embodiments provide mutations which confer improved or altered properties on polypeptides comprising them. Such mutations may be utilized with any background sequence so long as the provided toxin exhibits altered or improved pesticidal activity.

Embodiments of the present invention provide, *inter alia*, compositions and
20 methods for modulating the total level of polypeptides and/or altering their ratios in a plant. As used herein, the term "modulation" is intended to mean an increase or decrease in a particular character, quality, substance, or response. The compositions comprise nucleotide and amino acid sequences from various plant species.

25 As used herein, the terms "plant pathogen" or "plant pest" refer to any organism that can cause harm to a plant. A plant can be harmed by an inhibition or slowing of the growth of a plant, by damage to the tissues of a plant, by a weakening of the immune system of a plant, by a reduction in the resistance of a plant to abiotic stresses, by a premature death of the plant, and the like.

30 As used herein, the terms "disease resistance" or "pathogen resistance" are intended to mean that the organisms avoid the disease symptoms that are the outcome of organism-pathogen interactions. That is, pathogens are prevented

from causing diseases and the associated disease symptoms, or alternatively, the disease symptoms caused by the pathogen are minimized or lessened.

As used herein, the terms "pesticidal activity" and "insecticidal activity" are used synonymously to refer to activity of an organism or a substance (such as, for example, a protein) that can be measured by but is not limited to pest mortality, pest weight loss, pest repellency, and other behavioral and physical changes of a pest after feeding and exposure for an appropriate length of time. In this manner, pest
5 pesticidal activity impacts at least one measurable parameter of pest fitness. Accordingly, "pesticidal activity" and "insecticidal activity" include, but are not limited to, damage caused by plant insect pests. For example "pesticidal proteins" are proteins that display pesticidal activity by themselves or in combination with other proteins. Endotoxins are pesticidal proteins. Other examples of pesticidal proteins include, e.g., pentin-1 (see U.S. Pat. Nos. 6,057,491 and 6,339,144).
10

A "pesticidal agent" will act similarly to suppress, control, and/or kill an
15 invading pathogen.

An "insecticidal composition" is intended to mean that the compositions of embodiments have activity against plant insect pests, and thus are capable of suppressing, controlling, and/or killing the invading insect. An insecticidal composition of the embodiments will reduce the symptoms resulting from insect
20 challenge by at least about 5% to about 50%, at least about 10% to about 60%, at least about 30% to about 70%, at least about 40% to about 80%, or at least about 50% to about 90% or greater. Hence, the methods of the embodiments of the invention can be utilized to protect organisms, particularly plants, from invading insects.

Assays that measure insecticidal activity are commonly known in the art, such as insect-feeding bioassays. See, for example, Marrone *et al.* (1985) *J. Econ. Entomol.* 78:290-293 and Czapla and Lang (1990) *J. Econ. Entomol.* 83:2480-2485, herein incorporated by reference. The preferred developmental stage for testing for pesticidal activity is larvae or immature forms of these above
25 mentioned insect pests. The insects may be reared in total darkness at from about 20°C to about 30°C and from about 30% to about 70% relative humidity. Methods of rearing insect larvae and performing bioassays are well known to one of ordinary skill in the art.
30

A wide variety of bioassay techniques are known to one skilled in the art. General procedures include addition of the experimental compound or organism to the diet source in an enclosed container. Pesticidal activity can be measured by, but is not limited to, changes in mortality, weight loss, attraction, repellency and
5 other behavioral and physical changes after feeding and exposure for an appropriate length of time. Bioassays described herein can be used with any feeding insect pest in the larval or adult stage.

The compositions of the embodiments can be used in a variety of methods whereby the protein products can be expressed in crop plants to function as
10 insecticidal proteins. The compositions of the embodiments may be expressed in a crop plant such as maize or soybean to function as an insecticidal agent. Expression of the proteins of the embodiments can also be altered, resulting in changes or modulation of the level, tissue, or timing of expression in order to achieve enhanced insect resistance.

15 The coding sequence for the cyclotide can be used in combination with a promoter that is introduced into a crop plant. In one embodiment, a high-level expressing constitutive promoter may be utilized and would result in high levels of expression of the cyclotide. In other embodiments, the coding sequence may be operably linked to a tissue-preferred promoter to direct the expression to a plant
20 tissue known to be susceptible to an insect. Likewise, manipulation of the timing of expression may be utilized. For example, by judicious choice of a promoter, expression can be enhanced early in plant growth to prime the plant to be responsive to insect attack. Likewise, pathogen inducible promoters can be used wherein expression of the cyclotide is turned on in the presence of the insect. If
25 desired, a transit peptide can be utilized to direct cellular localization of the protein product. In this manner, the native transit peptide or a heterologous transit peptide can be used. However, it is recognized that both extracellular expression and intracellular expression are encompassed by the methods of the
30 embodiments.

Sequences of the embodiments, as discussed in more detail below, encompass coding sequences, antisense sequences, and fragments and variants thereof. Expression of the sequences of the embodiments can be used to modulate or regulate the expression of corresponding cyclotide proteins.

The compositions and methods of the embodiments can be used for enhancing resistance to plant insect pests. The method involves stably transforming a plant with a nucleotide sequence capable of modulating the plant insect defense system operably linked with a promoter capable of driving
5 expression of a gene in a plant cell. By "enhancing resistance" increasing the tolerance of the plant to insects is intended. That is, the cyclotide may slow or prevent insect infection and/or spread.

Consequently, an isolated polynucleotide comprising a nucleotide sequence of at least 30 contiguous nucleotides derived from the nucleotide
10 sequence of SEQ ID NO: 11 and its complement may be used in methods of selecting an isolated polynucleotide that affects the expression of a plant cyclotide polypeptide in a host cell. For example, an isolated polynucleotide comprising at least 30, at least 40, at least 50, at least 60 or at least any number of nucleotides up to the full length of SEQ ID NO:11. A method of selecting an isolated
15 polynucleotide that affects the level of expression of a polypeptide in a virus or in a host cell (eukaryotic, such as plant or yeast, prokaryotic such as bacterial) may comprise the steps of: constructing an isolated polynucleotide or an isolated chimeric gene; introducing the isolated polynucleotide or the isolated chimeric gene into a host cell; measuring the level of a polypeptide or enzyme activity in the
20 host cell containing the isolated polynucleotide; and comparing the level of a polypeptide or enzyme activity in the host cell containing the isolated polynucleotide with the level of a polypeptide or enzyme activity in a host cell that does not contain the isolated polynucleotide.

Genes encoding other plant cyclotides, either as cDNAs or genomic DNAs,
25 could be isolated directly by using all or a portion of the instant nucleic acid fragments as DNA hybridization probes to screen libraries from any desired plant employing methodology well known to those skilled in the art. Specific oligonucleotide probes based upon the instant nucleic acid sequences can be designed and synthesized by methods known in the art (Sambrook *et al.* (1989),
30 *supra*). Moreover, the entire sequences can be used directly to synthesize DNA probes by methods known to the skilled artisan such as random primer DNA labeling, nick translation, or end-labeling techniques, or RNA probes using available *in vitro* transcription systems. In addition, specific primers can be

designed and used to amplify a part or all of the instant sequences. The resulting amplification products can be labeled directly during amplification reactions or labeled after amplification reactions, and used as probes to isolate full length cDNA or genomic fragments under conditions of appropriate stringency.

5 In addition, two short segments of the instant nucleic acid fragments may be used in PCR protocols to amplify longer nucleic acid fragments encoding homologous genes from DNA or RNA. PCR may also be performed on a library of cloned nucleic acid fragments wherein the sequence of one primer is derived from the instant nucleic acid fragments, and the sequence of the other primer
10 takes advantage of the presence of the polyadenylic acid tracts to the 3' end of the mRNA precursor encoding plant genes. Alternatively, the second primer sequence may be based upon sequences derived from the cloning vector. For example, the skilled artisan can follow the RACE protocol (Frohman *et al.* (1988) *Proc. Natl. Acad. Sci. USA* 85:8998-9002) to generate cDNAs by using PCR to
15 amplify copies of the region between a single point in the transcript and the 3' or 5' end. Primers oriented in the 3' and 5' directions can be designed from the instant sequences. Using commercially available 3' RACE or 5' RACE systems (BRL), specific 3' or 5' cDNA fragments can be isolated (Ohara *et al.* (1989) *Proc. Natl. Acad. Sci. USA* 86:5673-5677; Loh *et al.* (1989) *Science* 243:217-220). Products
20 generated by the 3' and 5' RACE procedures can be combined to generate full-length cDNAs (Frohman and Martin (1989) *Techniques* 1:165). Consequently, a polynucleotide comprising a nucleotide sequence of at least 60 (or at least 40, or at least 30) contiguous nucleotides derived from the nucleotide sequence set forth in SEQ ID NO: 11 and its complement, may be used in such methods to obtain a
25 nucleic acid fragment encoding a substantial portion of an amino acid sequence of a polypeptide of the embodiments.

Embodiments of the invention relate to a method of obtaining a nucleic acid fragment encoding a substantial portion of a cyclotide polypeptide comprising the steps of: synthesizing an oligonucleotide primer comprising a nucleotide sequence
30 of at least 10, at least 20, or at least 30 or more contiguous nucleotides derived from the nucleotide sequence set forth in SEQ ID NO:11 and its complement; and amplifying a nucleic acid fragment using the oligonucleotide primer. The amplified

nucleic acid fragment preferably will encode a portion of a plant cyclotide polypeptide.

Availability of the instant nucleotide and deduced amino acid sequences facilitates immunological screening of cDNA expression libraries. Synthetic peptides representing portions of the cyclotide amino acid sequences may be synthesized. These peptides can be used to immunize animals to produce polyclonal or monoclonal antibodies with specificity for peptides or proteins comprising the amino acid sequences. These antibodies can then be used to screen cDNA expression libraries to isolate full-length cDNA clones of interest (Lerner (1984) *Adv. Immunol.* 36:1-34; Sambrook *et al.* (1989) *supra*).

Fragments of a nucleotide sequence may encode protein fragments that retain the biological activity of the native protein and hence have cyclotide activity, for example, insecticidal activity, and thereby affect responses to pathogens. Alternatively, fragments of a nucleotide sequence that are useful as hybridization probes generally do not encode protein fragments retaining biological activity. Thus, fragments of a nucleotide sequence may range from at least about 20 nucleotides, about 50 nucleotides, about 100 nucleotides, and up to the full-length nucleotide sequence encoding the proteins of the embodiments.

A fragment of a cyclotide nucleotide sequence that encodes a biologically active portion of a cyclotide protein of the embodiments will encode at least 10, 15, 25, 30, 50, 100, contiguous amino acids, or up to the total number of amino acids present in a full-length protein of the embodiments. Fragments of a cyclotide nucleotide sequence that are useful as hybridization probes for PCR primers generally need not encode a biologically active portion of a cyclotide protein.

Thus, a fragment of a cyclotide nucleotide sequence may encode a biologically active portion of a cyclotide protein, or it may be a fragment that can be used as a hybridization probe or PCR primer using methods disclosed herein. A biologically active portion of a cyclotide protein can be prepared by isolating a portion of one of the cyclotide nucleotide sequences of the embodiments, expressing the encoded portion of the cyclotide protein (e.g., by recombinant expression *in vitro*), and assessing the activity of the encoded portion of the cyclotide protein. Nucleic acid molecules that are fragments of a cyclotide nucleotide sequence comprise at least 16, 20, 30, 40, 50, 75, 100, 150, 200, 250,

or 300 nucleotides, or up to the number of nucleotides present in a full-length cyclotide nucleotide sequence disclosed herein.

By "variants" substantially similar sequences are intended. 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 cyclotide polypeptides of the embodiments. Naturally occurring allelic variants such as these can be identified with the use of well-known molecular biology techniques, such as, for example, with polymerase chain reaction (PCR) and hybridization techniques as outlined herein. 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 cyclotide protein. Generally, variants of a particular nucleotide sequence of the embodiments will have at least about 50%, 60%, 65%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 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" 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 is intended. Variant proteins encompassed by the embodiments are biologically active, that is they continue to possess the desired biological activity of the native protein, that is, cyclotide activity as described herein, for example, insecticidal activity. Such variants may result from, for example, genetic polymorphism or from human manipulation. Biologically active variants of a native cyclotide protein of the embodiments will have at least about 40%, 50%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, and including 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 the native protein 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 polypeptides of the embodiments may be altered in various ways including amino acid substitutions, deletions, truncations, and insertions. Novel proteins having properties of interest may be created by combining elements and fragments of proteins of the embodiments with other proteins as well. Methods for such manipulations are generally known in the art. For example, amino acid sequence variants of the cyclotide 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; U.S. 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 embodiments include both naturally occurring sequences as well as mutant forms. Likewise, the proteins of the embodiments encompass naturally occurring proteins as well as variations and modified forms thereof. Such variants will continue to possess the desired cyclotide activity (for example, insecticidal activity) or defense response activity. Obviously, 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 Publication No. 0 075 444 B1).

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. Biological activity of the variant polypeptides of the embodiments can be assayed by any method known in the art, such as those already discussed and referenced elsewhere in this application.

Variant nucleotide sequences and proteins also encompass sequences and
5 proteins derived from a mutagenic and recombinogenic procedure such as DNA
shuffling. With such a procedure, one or more different cyclotide coding
sequences can be manipulated to create a new cyclotide protein possessing the
desired properties. In this manner, libraries of recombinant polynucleotides are
generated from a population of related sequence polynucleotides comprising
10 sequence regions that have substantial sequence identity and can be
homologously recombined in vitro or in vivo. 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; Cramer *et al.* (1997)
Nature Biotech. 15:436-438; Moore *et al.* (1997) *J. Mol. Biol.* 272:336-347; Zhang
15 *et al.* (1997) *Proc. Natl. Acad. Sci. USA* 94:4504-4509; Cramer *et al.* (1998)
Nature 391:288-291; and U.S. Patent Nos. 5,605,793 and 5,837,458.

“Codon degeneracy” refers to divergence in the genetic code permitting
variation of the nucleotide sequence without affecting the amino acid sequence of
an encoded polypeptide. Accordingly, the instant invention relates to any nucleic
20 acid fragment comprising a nucleotide sequence that encodes all or a substantial
portion of the amino acid sequences set forth herein. The skilled artisan is well
aware of the “codon-bias” exhibited by a specific host cell in usage of nucleotide
codons to specify a given amino acid. Therefore, when synthesizing a nucleic
acid fragment for improved expression in a host cell, it is desirable to design the
25 nucleic acid fragment such that its frequency of codon usage approaches the
frequency of preferred codon usage of the host cell. Determination of preferred
codons can be based on a survey of genes derived from the host cell where
sequence information is available. For example, the codon frequency tables
available on the world wide web at Kazusa.or.jp/codon/ may be used to determine
30 preferred codons for a variety of organisms. See also Campbell and Gowri (1990)
Plant Physiol. 92:1-11; Murray *et al.* (1989) *Nucleic Acids Res.* 17:477-498, U.S.
Patent Nos. 5,380,831 and 5,436,391; and the information found on the world

wide web at agron.missouri.edu/mnl/77/10simmons.html; herein incorporated by reference.

“Synthetic nucleic acid fragments” can be assembled from oligonucleotide building blocks that are chemically synthesized using procedures known to those skilled in the art. These building blocks are ligated and annealed to form larger nucleic acid fragments which may then be enzymatically assembled to construct the entire desired nucleic acid fragment. “Chemically synthesized”, as related to a nucleic acid fragment, means that the component nucleotides were assembled *in vitro*. Manual chemical synthesis of nucleic acid fragments may be accomplished using well established procedures, or automated chemical synthesis can be performed using one of a number of commercially available machines. Accordingly, the nucleic acid fragments can be tailored for optimal gene expression based on optimization of nucleotide sequence to reflect the codon bias of the host cell. The skilled artisan appreciates the likelihood of successful gene expression if codon usage is biased towards those codons favored by the host.

“Gene” refers to a nucleic acid fragment that expresses a specific protein, including regulatory sequences preceding (5' non-coding sequences) and following (3' non-coding sequences) the coding sequence. “Native gene” refers to a gene as found in nature with its own regulatory sequences. “Chimeric gene” refers any gene that is not a native gene, comprising regulatory and coding sequences that are not found together in nature. Accordingly, a chimeric gene may comprise regulatory sequences and coding sequences that are derived from different sources, or regulatory sequences and coding sequences derived from the same source, but arranged in a manner different than that found in nature.

“Endogenous gene” refers to a native gene in its natural location in the genome of an organism. A “foreign” gene refers to a gene not normally found in the host organism, but that is introduced into the host organism by gene transfer. Foreign genes can comprise native genes inserted into a non-native organism, or chimeric genes. A “transgene” is a gene that has been introduced into the genome by a transformation procedure.

“Synthetic genes” can be assembled from oligonucleotide building blocks that are chemically synthesized using procedures known to those skilled in the art. These building blocks are ligated and annealed to form gene segments which are

then enzymatically assembled to construct the entire gene. "Chemically synthesized", as related to a sequence of DNA, means that the component nucleotides were assembled *in vitro*. Manual chemical synthesis of DNA may be accomplished using well established procedures, or automated chemical
5 synthesis can be performed using one of a number of commercially available machines. Accordingly, the genes can be tailored for optimal gene expression based on optimization of the nucleotide sequence to reflect the codon bias of the host cell. The skilled artisan appreciates the likelihood of successful gene expression if codon usage is biased towards those codons favored by the host.
10 Determination of preferred codons can be based on a survey of genes derived from the host cell where sequence information is available.

"Coding sequence" refers to a nucleotide sequence that codes for a specific amino acid sequence. As used herein, the terms "encoding" or "encoded" when used in the context of a specified nucleic acid mean that the nucleic acid
15 comprises the requisite information to guide translation of the nucleotide sequence into a specified protein. The information by which a protein is encoded is specified by the use of codons. A nucleic acid encoding a protein may comprise non-translated sequences (e.g., introns) within translated regions of the nucleic acid or may lack such intervening non-translated sequences (e.g., as in
20 cDNA).

"Regulatory sequences" refer to nucleotide sequences located upstream (5' non-coding sequences), within, or downstream (3' non-coding sequences) of a coding sequence, and which influence the transcription, RNA processing or stability, or translation of the associated coding sequence. Regulatory sequences
25 may include promoters, translation leader sequences, introns, and polyadenylation recognition sequences.

A "protein" or "polypeptide" is a chain of amino acids arranged in a specific order determined by the coding sequence in a polynucleotide encoding the polypeptide.

30 "Altered levels" or "altered expression" refers to the production of gene product(s) in transgenic organisms in amounts or proportions that differ from that of normal or non-transformed organisms.

“Null mutant” refers here to a host cell that either lacks the expression of a certain polypeptide or expresses a polypeptide which is inactive or does not have any detectable expected enzymatic function.

In nature, some polypeptides are produced as complex precursors which, in addition to targeting labels such as the signal peptides discussed elsewhere in this application, also contain other fragments of peptides which are removed (processed) at some point during protein maturation, resulting in a mature form of the polypeptide that is different from the primary translation product (aside from the removal of the signal peptide). The following terms are of relevance. “Mature protein”, “preproprotein” or “prepropeptide” refer to a post-translationally processed polypeptide; i.e., one from which any pre- or propeptides present in the primary translation product have been removed. “Precursor protein” refers to the primary product of translation of mRNA; i.e., with pre- and propeptides still present. Pre- and propeptides may be, but are not limited to, intracellular localization signals. The form of the translation product with only the signal peptide removed but not further processing yet is called a “propeptide” or “proprotein”. The fragments to be removed may themselves are also referred to as “propeptides”. The skilled artisan will need to determine, depending on the species in which the proteins are being expressed and the desired intracellular location, if higher expression levels might be obtained by using a gene construct encoding just the mature form of the protein, the mature form with a signal peptide, or the proprotein (i.e., a form including propeptides) with a signal peptide. For optimal expression in plants or fungi, the pre- and propeptide sequences may be needed. The propeptides may play a role in aiding correct peptide folding.

A “chloroplast transit peptide” is an amino acid sequence that is translated in conjunction with a protein and directs the protein to the chloroplast or other plastid types present in the cell in which the protein is made. “Chloroplast transit sequence” refers to a nucleotide sequence that encodes a chloroplast transit peptide. A “signal peptide” is an amino acid sequence that is translated in conjunction with a protein and directs the protein to the secretory system (see Chrispeels (1991) *Ann. Rev. Plant Phys. Plant Mol. Biol.* 42:21-53). If the protein is to be directed to a vacuole, a vacuolar targeting signal can further be added, or if to the endoplasmic reticulum, an endoplasmic reticulum retention signal may be

added. If the protein is to be directed to the nucleus, any signal peptide present should be removed and instead a nuclear localization signal included (see Raikhel (1992) *Plant Phys.* 100:1627-1632).

5 “Transformation” refers to the transfer of a nucleic acid fragment into the genome of a host organism, resulting in genetically stable inheritance. Host organisms containing the transformed nucleic acid fragments are referred to as “transgenic” organisms. Examples of methods of plant transformation include *Agrobacterium*-mediated transformation (De Blaere *et al.* (1987) *Meth. Enzymol.* 143:277) and particle-accelerated or “gene gun” transformation technology (Klein *et al.* (1987) *Nature* (London) 327:70-73; U.S. Patent No. 4,945,050, incorporated 10 herein by reference). Additional transformation methods are disclosed below. Thus, isolated polynucleotides of the embodiments can be incorporated into recombinant constructs, typically DNA constructs, capable of introduction into and replication in a host cell. Such a construct can be a vector that includes a 15 replication system and sequences that are capable of transcription and translation of a polypeptide-encoding sequence in a given host cell. A number of vectors suitable for stable transfection of plant cells or for the establishment of transgenic plants have been described in, e.g., Pouwels *et al.*, (1985; Supp. 1987) *Cloning Vectors: A Laboratory Manual*, Weissbach and Weissbach (1989) *Methods for* 20 *Plant Molecular Biology*, (Academic Press, New York); and Flevin *et al.*, (1990) *Plant Molecular Biology Manual*, (Kluwer Academic Publishers). Typically, plant expression vectors include, for example, one or more cloned plant genes under the transcriptional control of 5' and 3' regulatory sequences and a dominant selectable marker. Such plant expression vectors also can contain a promoter 25 regulatory region (e.g., a regulatory region controlling inducible or constitutive, environmentally- or developmentally-regulated, or cell- or tissue-specific expression), a transcription initiation start site, a ribosome binding site, an RNA processing signal, a transcription termination site, and/or a polyadenylation signal.

Standard recombinant DNA and molecular cloning techniques used herein 30 are well known in the art and are described more fully in Sambrook *et al.* (1989) *supra*.

In another embodiment, this invention concerns viruses and host cells comprising either the chimeric genes of the embodiments as described herein or

an isolated polynucleotide of the embodiments as described herein. Examples of host cells that can be used to practice the embodiments include, but are not limited to, yeast, bacterial, fungal, insect, amphibian, mammalian, and plant cells.

As used herein, "host cell" refers to a cell which comprises a heterologous nucleic acid sequence of the embodiments. Host cells may be prokaryotic cells such as *E. coli*, or eukaryotic cells such as yeast, fungal, insect, amphibian, mammalian or plant cells. Host plant cells include monocotyledonous or dicotyledonous plant cells. One example of a monocotyledonous host cell is a maize host cell. One example of a dicotyledonous host cell is a soybean host cell.

Overexpression of the proteins of the instant invention may be accomplished by first constructing a chimeric gene in which the coding region is operably linked to a promoter capable of directing expression of a gene in the desired tissues at the desired stage of development. The chimeric gene may comprise promoter sequences and translation leader sequences derived from the same genes. 3' non-coding sequences encoding transcription termination signals may also be provided. The instant chimeric gene may also comprise one or more introns in order to facilitate gene expression.

The cyclotide sequences of the embodiments are provided in DNA constructs for expression in the plant of interest. The cassette will include 5' and 3' regulatory sequences operably linked to a cyclotide sequence of the embodiments. 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 DNA constructs.

Such a DNA construct is provided with a plurality of restriction sites for insertion of the cyclotide sequence to be under the transcriptional regulation of the regulatory regions. The DNA construct may additionally contain selectable marker genes.

In specific embodiments, methods for increasing pathogen resistance in a plant comprise stably transforming a plant with a DNA construct comprising an antipathogenic nucleotide sequence of the embodiments operably linked to a 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, examples of promoters include constitutive and pathogen-inducible promoters.

A number of promoters can be used in the practice of the embodiments. The promoters can be selected based on the desired outcome. That is, the
5 nucleic acids can be combined with constitutive, tissue-preferred, or other promoters for expression in the host cell of interest. Such constitutive promoters include, for example, the core promoter of the Rsyn7 promoter and other constitutive promoters disclosed in WO 99/43838 and U.S. Patent No. 6,072,050; the core CaMV 35S promoter (Odell *et al.* (1985) *Nature* 313:810-812); rice actin
10 (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, those disclosed in U.S.
15 Patent Nos. 5,608,149; 5,608,144; 5,604,121; 5,569,597; 5,466,785; 5,399,680; 5,268,463; 5,608,142; and 6,177,611, herein incorporated by reference.

Generally, it will be beneficial to express the gene from an inducible promoter, for example from a pathogen-inducible promoter. Such promoters include those from pathogenesis-related proteins (PR proteins), which are induced
20 following infection by a pathogen; e.g., PR proteins, SAR proteins, beta-1,3-glucanase, chitinase, etc. See, for example, Redolfi *et al.* (1983) *Neth. J. Plant Pathol.* 89:245-254; Uknes *et al.* (1992) *Plant Cell* 4:645-656; and Van Loon (1985) *Plant Mol. Virol.* 4:111-116. See also WO 99/43819 published September 9, 1999, herein incorporated by reference.

Of interest are promoters that are expressed locally at or near the site of
25 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*
30 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).

5 Additionally, as pathogens find entry into plants through wounds or insect damage, a wound-inducible promoter may be used in the constructions of the
embodiments. 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*, U.S. Patent No. 5,428,148; *win1* and *win2* (Stanford *et al.* (1989) *Mol. Gen. Genet.* 215:200-208);
10 *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
15 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
20 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
25 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.

30 Tissue-preferred promoters can be utilized to target enhanced cyclotide 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 are known in the art. See, for example, Yamamoto *et al.* (1997) *Plant J.* 12(2):255-265; Kwon *et al.* (1994) *Plant Physiol.* 105:357-67; Yamamoto *et al.* (1994) *Plant Cell Physiol.* 35(5):773-778; Gotor *et al.* (1993) *Plant J.* 3:509-18; Orozco *et al.* (1993) *Plant Mol. Biol.* 23(6):1129-1138; and Matsuoka *et al.* (1993) *Proc. Natl. Acad. Sci. USA* 90(20):9586-9590.

"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); milps (myo-inositol-1-phosphate synthase); and celA (cellulose synthase) (see WO 00/11177, herein incorporated by reference). 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. See also WO 00/12733, where seed-preferred promoters from end1 and end2 genes are disclosed; herein incorporated by reference.

The method of transformation/transfection is not critical to the instant invention. Various methods of transformation or transfection are currently available. As newer methods are available to transform crops or other host cells they may be used with the instant invention. Accordingly, a wide variety of methods have been developed to insert a DNA sequence into the genome of a

host cell to obtain the transcription and/or translation of the sequence to effect phenotypic changes in the organism. The nucleic acid fragments of the instant invention may be used to create transgenic plants in which the disclosed plant cyclotides are present at higher or lower levels than normal or in cell types or developmental stages in which they are not normally found. This would have the effect of altering the level of disease (e.g., fungal) and pathogen resistance in those cells. Thus, any method, which provides for effective transformation/transfection may be employed.

Prokaryotic cells may be used as hosts for expression. Prokaryotes most frequently are represented by various strains of *E. coli*; however, other microbial strains may also be used. Commonly used prokaryotic control sequences which are defined herein to include promoters for transcription initiation, optionally with an operator, along with ribosome binding sequences, include such commonly used promoters as the beta lactamase (penicillinase) and lactose (lac) promoter systems (Chang *et al.* (1977) *Nature* 198:1056), the tryptophan (trp) promoter system (Goeddel *et al.* (1980) *Nucleic Acids Res.* 8:4057) and the lambda derived PL promoter and N-gene ribosome binding site (Simatake and Rosenberg (1981) *Nature* 292:128). Examples of selection markers for *E. coli* include, for example, genes specifying resistance to ampicillin, tetracycline, or chloramphenicol.

The vector is selected to allow introduction into the appropriate host cell. Bacterial vectors are typically of plasmid or phage origin. Appropriate bacterial cells are infected with phage vector particles or transfected with naked phage vector DNA. If a plasmid vector is used, the bacterial cells are transfected with the plasmid vector DNA. Expression systems for expressing a protein of the embodiments are available using *Bacillus* sp. and *Salmonella* (Palva *et al.* (1983) *Gene* 22:229-235 and Mosbach *et al.* (1983) *Nature* 302:543-545).

A variety of eukaryotic expression systems such as yeast, insect cell lines, plant and mammalian cells, are known to those of skill in the art. As explained briefly below, a polynucleotide of the embodiments can be expressed in these eukaryotic systems. In some embodiments, transformed/transfected plant cells, as discussed *infra*, are employed as expression systems for production of the proteins of the instant invention. Such antimicrobial proteins can be used for any application including coating surfaces to target microbes as described further *infra*.

Synthesis of heterologous nucleotide sequences in yeast is well known. Sherman, *et al.* (1982) *Methods in Yeast Genetics* (Cold Spring Harbor Laboratory) is a well recognized work describing the various methods available to produce proteins in yeast. Two widely utilized yeasts for production of eukaryotic proteins are *Saccharomyces cerevisiae* and *Pichia pastoris*. Vectors, strains, and protocols for expression in *Saccharomyces* and *Pichia* are known in the art and available from commercial suppliers (e.g., Invitrogen). Suitable vectors usually have expression control sequences, such as promoters, including 3-phosphoglycerate kinase or alcohol oxidase, and an origin of replication, termination sequences and the like, as desired.

A protein of the embodiments, once expressed, can be isolated from yeast by lysing the cells and applying standard protein isolation techniques to the lysates. The monitoring of the purification process can be accomplished by using Western blot techniques, radioimmunoassay, or other standard immunoassay techniques.

The sequences of the embodiments can also be ligated to various expression vectors for use in transfecting cell cultures of, for instance, mammalian, insect, or plant origin. Illustrative cell cultures useful for the production of the peptides are mammalian cells. A number of suitable host cell lines capable of expressing intact proteins have been developed in the art, and include the HEK293, BHK21, and CHO cell lines. Expression vectors for these cells can include expression control sequences, such as an origin of replication, a promoter (e.g. the CMV promoter, a HSV tk promoter or pgk (phosphoglycerate kinase) promoter), an enhancer (Queen *et al.* (1986) *Immunol. Rev.* 89:49), and necessary processing information sites, such as ribosome binding sites, RNA splice sites, polyadenylation sites (e.g., an SV40 large T Ag poly A addition site), and transcriptional terminator sequences. Other animal cells useful for production of proteins of the embodiments are available, for instance, from the American Type Culture Collection.

Appropriate vectors for expressing proteins of the embodiments in insect cells are usually derived from the SF9 baculovirus. Suitable insect cell lines include mosquito larvae, silkworm, armyworm, moth and *Drosophila* cell lines such

as a Schneider cell line (See, Schneider (1987) *J. Embryol. Exp. Morphol.* 27:353-365).

As with yeast, when higher animal or plant host cells are employed, polyadenylation or transcription terminator sequences are typically incorporated
5 into the vector. An example of a terminator sequence is the polyadenylation sequence from the bovine growth hormone gene. Sequences for accurate splicing of the transcript may also be included. An example of a splicing sequence is the VP1 intron from SV40 (Sprague *et al.* (1983) *J. Virol.* 45:773-781). Additionally, gene sequences to control replication in the host cell may be incorporated into the
10 vector such as those found in bovine papilloma virus type-vectors. Saveria-Campo (1985) "Bovine Papilloma Virus DNA a Eukaryotic Cloning Vector," in *DNA Cloning Vol. II: A Practical Approach*, ed. D.M. Glover (IRL Press, Arlington, Virginia), pp. 213-238.

Animal and lower eukaryotic (e.g., yeast) host cells are competent or
15 rendered competent for transfection by various means. There are several well-known methods of introducing DNA into animal cells. These include: calcium phosphate precipitation, fusion of the recipient cells with bacterial protoplasts containing the DNA, treatment of the recipient cells with liposomes containing the DNA, DEAE dextrin, electroporation, biolistics, and micro-injection of the DNA
20 directly into the cells. The transfected cells are cultured by means well known in the art. Kuchler (1997) *Biochemical Methods in Cell Culture and Virology* (Dowden, Hutchinson and Ross, Inc.).

Plasmid vectors comprising the instant isolated polynucleotide (or chimeric gene) may be constructed. The choice of plasmid vector is dependent upon the
25 method that will be used to transform host plants. The skilled artisan is well aware of the genetic elements that must be present on the plasmid vector in order to successfully transform, select and propagate host cells containing the chimeric gene. The skilled artisan will also recognize that different independent transformation events will result in different levels and patterns of expression
30 (Jones *et al.* (1985) *EMBO J.* 4:2411-2418; De Almeida *et al.* (1989) *Mol. Gen. Genetics* 218:78-86), and thus that multiple events must be screened in order to obtain lines displaying the desired expression level and pattern. Such screening

may be accomplished by Southern analysis of DNA, northern analysis of mRNA expression, western analysis of protein expression, or phenotypic analysis.

For some applications it may be useful to direct the instant polypeptides to different cellular compartments, or to facilitate its secretion from the cell. It is thus
5 envisioned that the chimeric gene described above may be further supplemented by directing the coding sequence to encode the instant polypeptides with appropriate intracellular targeting sequences such as transit sequences (Keegstra (1989) *Cell* 56:247-253), signal sequences or sequences encoding endoplasmic reticulum (ER) localization (Chrispeels (1991) *supra*), or nuclear localization
10 signals (Raikhel (1992) *supra*) with or without removing targeting sequences that are already present.

Unlike the promoter, which acts at the transcriptional level, such targeting information is part of the initial translation product. The location of the protein in different compartments of the cell may make it more efficacious or make it
15 interfere less with the functions of the cell. For example, one may produce a protein preceded by a signal peptide, which directs the translation product into the ER, by including in the chimeric construct sequences encoding a signal peptide (such sequences may also be called the "signal sequence"). The signal sequence used could be that associated with the gene encoding the polypeptide, or it may
20 be taken from another gene. There are many signal peptides described in the literature, and they are largely interchangeable (Raikhel N, Chrispeels MJ (2000) Protein sorting and vesicle traffic. *In* B Buchanan, W Gruissem, R Jones, eds, Biochemistry and Molecular Biology of Plants. American Society of Plant Physiologists, Rockville, MD, pp 160-201, herein incorporated by reference). The
25 addition of a signal peptide will result in the translation product entering the ER (in the process of which the signal peptide itself is removed from the polypeptide), but the final intracellular location of the protein depends on other factors, which may be manipulated to result in localization most appropriate for the pathogen and cell type.

30 The default pathway, that is, the pathway taken by the polypeptide if no other targeting labels are included, results in secretion of the polypeptide across the cell membrane (Raikhel and Chrispeels, (2000) *supra*). This will leave the peptide between the cell membrane and cell wall, which will often be a suitable location.

Other pathogens may be more effectively combated by locating the peptide within the cell. This can be accomplished, for example, by adding an ER retention signal encoding sequence to the sequence to the gene. Methods and sequences for doing this are described in Raikhel and Chrispeels (2000) *supra*; for example, adding sequences encoding the amino acids K, D, E and L in that order, or variations thereof described in the literature, to the end of the protein coding portion of the polypeptide will accomplish this. Alternatively, the use of vacuolar targeting labels such as those described by Raikhel and Chrispeels (2000) *supra*, in addition to a signal peptide will result in localization of the peptide in a vacuolar structure. Use of a plastid transit peptide encoding sequence instead of a signal peptide encoding sequence will result in localization of the polypeptide in the plastid of the cell type chosen. One of skill in the art could also envision localizing the polypeptide in other cellular compartments by addition of suitable targeting information. While the references cited give examples of each of these, the list is not exhaustive and more targeting signals of use may be discovered in the future.

It may also be desirable to reduce or eliminate expression of genes encoding the instant polypeptides in plants for some applications. In order to accomplish this, a chimeric gene designed for co-suppression of the instant polypeptide can be constructed by linking a gene or gene fragment encoding that polypeptide to plant promoter sequences. Alternatively, a chimeric gene designed to express antisense RNA for all or part of the instant nucleic acid fragment can be constructed by linking the gene or gene fragment in reverse orientation to plant promoter sequences. Either the co-suppression or antisense chimeric genes could be introduced into plants via transformation wherein expression of the corresponding endogenous genes are reduced or eliminated.

Molecular genetic solutions to the generation of plants with altered gene expression have a decided advantage over more traditional plant breeding approaches. Changes in plant phenotypes can be produced by specifically inhibiting expression of one or more genes by antisense inhibition or cosuppression (U.S. Patent Nos. 5,190,931, 5,107,065 and 5,283,323). An antisense or cosuppression construct would act as a dominant negative regulator of gene activity. While conventional mutations can yield negative regulation of gene activity these effects are most likely recessive. The dominant negative

regulation available with a transgenic approach may be advantageous from a breeding perspective. In addition, the ability to restrict the expression of a specific phenotype to the reproductive tissues of the plant by the use of tissue specific promoters may confer agronomic advantages relative to conventional mutations
5 which may have an effect in all tissues in which a mutant gene is ordinarily expressed.

The person skilled in the art will know that special considerations are associated with the use of antisense or cosuppression technologies in order to reduce expression of particular genes. For example, the proper level of
10 expression of sense or antisense genes may require the use of different chimeric genes utilizing different regulatory elements known to the skilled artisan. Once transgenic plants are obtained by one of the methods described above, it will be necessary to screen individual transgenics for those that most effectively display the desired phenotype. Accordingly, the skilled artisan will develop methods for
15 screening large numbers of transformants. The nature of these screens will generally be chosen on practical grounds. For example, one can screen by looking for changes in gene expression by using antibodies specific for the protein encoded by the gene being suppressed, or one could establish assays that specifically measure enzyme activity. A preferred method will be one that allows
20 large numbers of samples to be processed rapidly, since it will be expected that a large number of transformants will be negative for the desired phenotype.

The instant polypeptides are useful in methods for impacting a plant pathogen comprising introducing into a plant or cell thereof at least one nucleotide construct comprising a nucleotide sequence of the embodiments operably linked
25 to a promoter that drives expression of an operably linked sequence in plant cells, wherein said nucleotide sequence is selected from the group consisting of: a nucleotide sequence set forth in SEQ ID NO: 11; a nucleotide sequence that encodes a polypeptide having the amino acid sequence set forth in SEQ ID NOs: 12 or 13; a nucleotide sequence characterized by at least 85% sequence identity
30 to the nucleotide sequences set forth in SEQ ID NO: 11; a nucleotide sequence characterized by at least 90% sequence identity to the nucleotide sequence set forth in SEQ ID NO: 11; a nucleotide sequence characterized by at least 95%

sequence identity to the nucleotide sequence set forth in SEQ ID NO: 11; and a nucleotide sequence that comprises the complement of any one of the above.

The instant polypeptides (or portions thereof) may be produced in heterologous host cells, particularly in the cells of microbial hosts, and can be used to prepare antibodies to the these proteins by methods well known to those skilled in the art. The antibodies are useful for detecting the polypeptides of the instant invention *in situ* in cells or *in vitro* in cell extracts. Polyclonal cyclotide antibodies can be prepared by immunizing a suitable subject (e.g., rabbit, goat, mouse, or other mammal) with a cyclotide agent immunogen. The anti-cyclotide antibody titer in the immunized subject can be monitored over time by standard techniques, such as with an enzyme linked immunosorbent assay (ELISA) using immobilized antimicrobial polypeptides. At an appropriate time after immunization, e.g., when the anti-cyclotide agent antibody titers are highest, antibody-producing cells can be obtained from the subject and used to prepare monoclonal antibodies by standard techniques, such as the hybridoma technique originally described by Kohler and Milstein (1975) *Nature* 256:495-497, the human B cell hybridoma technique (Kozbor *et al.* (1983) *Immunol. Today* 4:72), the EBV-hybridoma technique (Cole *et al.* (1985) in *Monoclonal Antibodies and Cancer Therapy*, ed. Reisfeld and Sell (Alan R. Liss, Inc., New York, NY), pp. 77-96) or trioma techniques. The technology for producing hybridomas is well known (see generally Coligan *et al.*, eds. (1994) *Current Protocols in Immunology* (John Wiley & Sons, Inc., New York, NY); Galfre *et al.* (1977) *Nature* 266:55052; Kenneth (1980) in *Monoclonal Antibodies: A New Dimension In Biological Analyses* (Plenum Publishing Corp., New York); and Lerner (1981) *Yale J. Biol. Med.* 54:387-402).

Alternative to preparing monoclonal antibody-secreting hybridomas, a monoclonal anti-cyclotide antibody can be identified and isolated by screening a recombinant combinatorial immunoglobulin library (e.g., an antibody phage display library) with a cyclotide to thereby isolate immunoglobulin library members that bind the defensive agent. Kits for generating and screening phage display libraries are commercially available (e.g., the Pharmacia Recombinant Phage Antibody System, Catalog No. 27-9400-01; and the Stratagene SurfZAP™ Phage Display Kit, Catalog No. 240612). Additionally, examples of methods and

reagents particularly amenable for use in generating and screening an antibody display library can be found in, for example, U.S. Patent No. 5,223,409; PCT Publication Nos. WO 92/18619; WO 91/17271; WO 92/20791; WO 92/15679; 93/01288; WO 92/01047; 92/09690; and 90/02809; Fuchs *et al.* (1991)

5 *Bio/Technology* 9:1370-1372; Hay *et al.* (1992) *Hum. Antibod. Hybridomas* 3:81-85; Huse *et al.* (1989) *Science* 246:1275-1281; Griffiths *et al.* (1993) *EMBO J.* 12:725-734. The antibodies can be used to identify homologs of the cyclotides of the embodiments.

10 All or a substantial portion of the polynucleotides of the instant invention may also be used as probes for genetically and physically mapping the genes that they are a part of, and as markers for traits linked to those genes. Such information may be useful in plant breeding in order to develop lines with desired phenotypes. For example, the instant nucleic acid fragments may be used as restriction fragment length polymorphism (RFLP) markers. Southern blots (Sambrook *et al.*
15 (1989) *supra*) of restriction-digested plant genomic DNA may be probed with the nucleic acid fragments of the instant invention. The resulting banding patterns may then be subjected to genetic analyses using computer programs such as MapMaker (Lander *et al.* (1987) *Genomics* 1:174-181) in order to construct a genetic map. In addition, the nucleic acid fragments of the instant invention may
20 be used to probe Southern blots containing restriction endonuclease-treated genomic DNAs of a set of individuals representing parent and progeny of a defined genetic cross. Segregation of the DNA polymorphisms is noted and used to calculate the position of the instant nucleic acid sequence in the genetic map previously obtained using this population (Botstein *et al.* (1980) *Am. J. Hum.*
25 *Genet.* 32:314-331).

The production and use of plant gene-derived probes for use in genetic mapping is described in Bernatzky and Tanksley (1986) *Plant Mol. Biol. Reporter* 4:37-41. Numerous publications describe genetic mapping of specific cDNA clones using the methodology outlined above or variations thereof. For example,
30 F2 intercross populations, backcross populations, randomly mated populations, near isogenic lines, and other sets of individuals may be used for mapping. Such methodologies are well known to those skilled in the art.

Nucleic acid probes derived from the instant nucleic acid sequences may also be used for physical mapping (i.e., placement of sequences on physical maps; see Hoheisel *et al.* in: *Nonmammalian Genomic Analysis: A Practical Guide*, Academic Press, New York), 1996, pp. 319-346, and references cited
5 therein).

In another embodiment, nucleic acid probes derived from the instant nucleic acid sequences may be used in direct fluorescence in situ hybridization (FISH) mapping (Trask (1991) *Trends Genet.* 7:149-154). Although current methods of FISH mapping favor use of large clones (several to several hundred KB; see Laan
10 *et al.* (1995) *Genome Res.* 5:13-20), improvements in sensitivity may allow performance of FISH mapping using shorter probes.

A variety of nucleic acid amplification-based methods of genetic and physical mapping may be carried out using the instant nucleic acid sequences. Examples include allele-specific amplification (Kazazian (1989) *J. Lab. Clin. Med.* 11:95-96),
15 polymorphism of PCR-amplified fragments (CAPS; Sheffield *et al.* (1993) *Genomics* 16:325-332), allele-specific ligation (Landegren *et al.* (1988) *Science* 241:1077-1080), nucleotide extension reactions (Sokolov (1990) *Nucleic Acid Res.* 18:3671), Radiation Hybrid Mapping (Walter *et al.* (1997) *Nat. Genet.* 7:22-28) and Happy Mapping (Dear and Cook (1989) *Nucleic Acid Res.*
20 17:6795-6807). For these methods, the sequence of a nucleic acid fragment is used to design and produce primer pairs for use in the amplification reaction or in primer extension reactions. The design of such primers is well known to those skilled in the art. In methods employing PCR-based genetic mapping, it may be necessary to identify DNA sequence differences between the parents of the
25 mapping cross in the region corresponding to the instant nucleic acid sequence. This, however, is generally not necessary for mapping methods.

Loss of function mutant phenotypes may be identified for the instant cDNA clones either by targeted gene disruption protocols or by identifying specific mutants for these genes contained in a maize population carrying mutations in all
30 possible genes (Ballinger and Benzer (1989) *Proc. Natl. Acad. Sci USA* 86:9402-9406; Koes *et al.* (1995) *Proc. Natl. Acad. Sci. USA* 92:8149-8153; Bensen *et al.* (1995) *Plant Cell* 7:75-84). The latter approach may be accomplished in two ways. First, short segments of the instant nucleic acid

fragments may be used in polymerase chain reaction protocols in conjunction with a mutation tag sequence primer on DNAs prepared from a population of plants in which Mutator transposons or some other mutation-causing DNA element has been introduced (see Bensen, *supra*). The amplification of a specific DNA
5 fragment with these primers indicates the insertion of the mutation tag element in or near the plant gene encoding the instant polypeptide. Alternatively, the instant nucleic acid fragment may be used as a hybridization probe against PCR amplification products generated from the mutation population using the mutation tag sequence primer in conjunction with an arbitrary genomic site primer, such as
10 that for a restriction enzyme site-anchored synthetic adaptor. With either method, a plant containing a mutation in the endogenous gene encoding the instant polypeptide can be identified and obtained. This mutant plant can then be used to determine or confirm the natural function of the instant polypeptides disclosed herein.

15 The methods of the embodiments can be used with other methods available in the art for enhancing insect resistance in plants. For example, embodiments of the invention encompass any one of a variety of second nucleotide sequences being utilized such that, when expressed in a plant, they help to increase the resistance of a plant to insect pests. It is recognized that
20 such second nucleotide sequences may be used in either the sense or antisense orientation depending on the desired outcome.

The methods of the embodiments can be used with other methods available in the art for enhancing disease and pathogen resistance in plants. Similarly, the antimicrobial compositions described herein may be used alone or in combination
25 with other nucleotide sequences, polypeptides, or agents to protect against plant diseases and pathogens. Although any one of a variety of second nucleotide sequences may be utilized, specific 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.

30 Proteins, peptides, and lysozymes that naturally occur in insects (Jaynes *et al.* (1987) *Bioassays* 6:263-270), plants (Broekaert *et al.* (1997) *Critical Reviews in Plant Sciences* 16:297-323), animals (Vunnam *et al.* (1997) *J. Peptide Res.* 49:59-66), and humans (Mitra and Zang (1994) *Plant Physiol.* 106:977-981; Nakajima *et*

al. (1997) *Plant Cell Reports* 16:674-679) are also a potential source of plant pathogen resistance (Ko, K. (2000) on the world wide web at Scisoc.org/feature/BioTechnology/antimicrobial.html). Examples of such plant resistance-conferring sequences include those encoding sunflower rhoGTPase-
5 Activating Protein (rhoGAP), lipoxygenase (LOX), Alcohol Dehydrogenase (ADH), and Sclerotinia-Inducible Protein-1 (SCIP-1) described in U.S. Patent Application Serial No. 09/714,767, herein incorporated by reference. These nucleotide sequences enhance plant disease resistance through the modulation of development, developmental pathways, and the plant pathogen defense system.
10 It is recognized that such second nucleotide sequences may be used in either the sense or antisense orientation depending on the desired outcome.

In another embodiment, the cyclotides comprise isolated polypeptides. The cyclotides of the embodiments find use in the decontamination of plant pathogens during the processing of grain for animal or human food consumption; during the
15 processing of feedstuffs, and during the processing of plant material for silage. In this embodiment, the cyclotides of the embodiments are presented to grain, plant material for silage, or a contaminated food crop, or during an appropriate stage of the processing procedure, in amounts effective for antimicrobial activity. The compositions can be applied to the environment of a plant pathogen by, for
20 example, spraying, atomizing, dusting, scattering, coating or pouring, introducing into or on the soil, introducing into irrigation water, by seed treatment, or dusting at a time when the plant pathogen has begun to appear or before the appearance of pests as a protective measure. It is recognized that any means that bring the cyclotide polypeptides in contact with the plant pathogen can be used in the
25 practice of the embodiments.

Additionally, the compositions can be used in formulations used for their antimicrobial activities. Methods are provided for controlling plant pathogens comprising applying a decontaminating amount of a polypeptide or composition of the embodiments to the environment of the plant pathogen. The polypeptides of
30 the embodiments can be formulated with an acceptable carrier into a composition(s) that is, for example, a suspension, a solution, an emulsion, a dusting powder, a dispersible granule, a wettable powder, an emulsifiable

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

Such compositions disclosed above may be obtained by the addition of a surface-active agent, an inert carrier, a preservative, a humectant, a feeding
5 stimulant, an attractant, an encapsulating agent, a binder, an emulsifier, a dye, a UV protectant, a buffer, a flow agent or fertilizers, micronutrient donors or other preparations that influence plant growth. One or more agrochemicals including, but not limited to, herbicides, insecticides, fungicides, bactericides, nematocides, molluscicides, acaricides, plant growth regulators, harvest aids, and fertilizers, can
10 be combined with carriers, surfactants, or adjuvants customarily employed in the art of formulation or other components to facilitate product handling and application for particular target mycotoxins. Suitable carriers and adjuvants can be solid or liquid and correspond to the substances ordinarily employed in formulation technology, e.g., natural or regenerated mineral substances, solvents,
15 dispersants, wetting agents, tackifiers, binders, or fertilizers. The active ingredients of the embodiments are normally applied in the form of compositions and can be applied to the crop area or plant to be treated, simultaneously or in succession, with other compounds. In some embodiments, methods of applying an active ingredient of the embodiments or an agrochemical composition of the
20 embodiments (which contains at least one of the proteins of the embodiments) are foliar application, seed coating, and soil application.

Suitable surface-active agents include, but are not limited to, anionic compounds such as a carboxylate of, for example, a metal; a carboxylate of a long chain fatty acid; an N-acylsarcosinate; mono or di-esters of phosphoric acid
25 with fatty alcohol ethoxylates or salts of such esters; fatty alcohol sulfates such as sodium dodecyl sulfate, sodium octadecyl sulfate, or sodium cetyl sulfate; ethoxylated fatty alcohol sulfates; ethoxylated alkylphenol sulfates; lignin sulfonates; petroleum sulfonates; alkyl aryl sulfonates such as alkyl-benzene sulfonates or lower alkyl-naphthalene sulfonates, e.g., butyl-naphthalene sulfonate;
30 salts of sulfonated naphthalene-formaldehyde condensates; salts of sulfonated phenol-formaldehyde condensates; more complex sulfonates such as the amide sulfonates, e.g., the sulfonated condensation product of oleic acid and N-methyl taurine; or the dialkyl sulfosuccinates, e.g., the sodium sulfonate or dioctyl

succinate. Non-ionic agents include condensation products of fatty acid esters, fatty alcohols, fatty acid amides or fatty-alkyl- or alkenyl-substituted phenols with ethylene oxide, fatty esters of polyhydric alcohol ethers, e.g., sorbitan fatty acid esters, condensation products of such esters with ethylene oxide, e.g.

5 polyoxyethylene sorbitar fatty acid esters, block copolymers of ethylene oxide and propylene oxide, acetylenic glycols such as 2, 4, 7, 9-tetraethyl-5-decyn-4, 7-diol, or ethoxylated acetylenic glycols. Examples of a cationic surface-active agent include, for instance, an aliphatic mono-, di-, or polyamine such as an acetate, naphthenate, or oleate; or oxygen-containing amine such as an amine oxide of
10 polyoxyethylene alkylamine; an amide-linked amine prepared by the condensation of a carboxylic acid with a di- or polyamine; or a quaternary ammonium salt.

Examples of inert materials include, but are not limited to, inorganic minerals such as kaolin, phyllosilicates, carbonates, sulfates, phosphates, or botanical materials such as cork, powdered corncobs, peanut hulls, rice hulls, and
15 walnut shells.

The compositions of the embodiments can be in a suitable form for direct application or as a concentrate of a primary composition, which requires dilution with a suitable quantity of water or other diluent before application. The decontaminating concentration will vary depending upon the nature of the
20 particular formulation, specifically, whether it is a concentrate or to be used directly.

In a further embodiment, the compositions, as well as the polypeptides of the embodiments can be treated prior to formulation to prolong the activity when applied to the environment of a plant pathogen as long as the pretreatment is not deleterious to the activity. Such treatment can be by chemical and/or physical
25 means as long as the treatment does not deleteriously affect the properties of the composition(s). Examples of chemical reagents include, but are not limited to, halogenating agents; aldehydes such as formaldehyde and glutaraldehyde; anti-infectives, such as zephiran chloride; alcohols, such as isopropanol and ethanol;
30 and histological fixatives, such as Bouin's fixative and Helly's fixative (see, for example, Humason (1967) *Animal Tissue Techniques* (W.H. Freeman and Co.)).

In one embodiment, the compositions comprise a microbe having stably integrated the nucleotide sequence of a cyclotide agent. The resulting microbes

can be processed and used as a microbial spray. Any suitable microorganism can be used for this purpose. See, for example, Gaertner *et al.* (1993) in *Advanced Engineered Pesticides*, Kim (Ed.). In another embodiment, the nucleotide sequences are introduced into microorganisms that multiply on plants (epiphytes) to deliver the cyclotides to potential target crops. Epiphytes can be, for example, gram-positive or gram-negative bacteria.

It is further recognized that whole, i.e., unlysed, cells of the transformed microorganism can be treated with reagents that prolong the activity of the polypeptide produced in the microorganism when the microorganism is applied to the environment of a target plant. A secretion signal sequence may be used in combination with the gene of interest such that the resulting enzyme is secreted outside the microorganism for presentation to the target plant.

In this manner, a gene encoding a cyclotide agent of the embodiments may be introduced via a suitable vector into a microbial host, and said transformed host applied to the environment, plants, or animals. Microorganism hosts that are known to occupy the "phytosphere" (phylloplane, phyllosphere, rhizosphere, and/or rhizoplane) of one or more crops of interest may be selected for transformation. These microorganisms are selected so as to be capable of successfully competing in the particular environment with the wild-type microorganisms, to provide for stable maintenance and expression of the gene expressing the detoxifying polypeptide, and to provide improved protection of the proteins of the embodiments from environmental degradation and inactivation.

Such microorganisms include bacteria, algae, and fungi. Illustrative prokaryotes, both Gram-negative and -positive, include *Enterobacteriaceae*, such as *Escherichia*, *Erwinia*, *Shigella*, *Salmonella*, and *Proteus*; *Bacillaceae*; *Rhizobiaceae*, such as *Rhizobium*; *Spirillaceae*, such as photobacterium, *Zymomonas*, *Serratia*, *Aeromonas*, *Vibrio*, *Desulfovibrio*, *Spirillum*; *Lactobacillaceae*; *Pseudomonadaceae*, such as *Pseudomonas* and *Acetobacter*; *Azotobacteraceae*; and *Nitrobacteraceae*. Among eukaryotes are fungi, such as *Phycomycetes* and *Ascomycetes*, which includes yeast, such as *Saccharomyces* and *Schizosaccharomyces*; and *Basidiomycetes* yeast, such as *Rhodotorula*, *Aureobasidium*, *Sporobolomyces*, and the like.

Of particular interest are microorganisms, such as bacteria, e.g., *Pseudomonas*, *Erwinia*, *Serratia*, *Klebsiella*, *Xanthomonas*, *Streptomyces*, *Rhizobium*, *Rhodopseudomonas*, *Methylius*, *Agrobacterium*, *Acetobacter*, *Lactobacillus*, *Arthrobacter*, *Azotobacter*, *Leuconostoc*, and *Alcaligenes*; fungi, particularly yeast, e.g., *Saccharomyces*, *Pichia*, *Cryptococcus*, *Kluyveromyces*, *Sporobolomyces*, *Rhodotorula*, *Aureobasidium*, and *Gliocladium*. Of particular interest are such phytosphere bacterial species as *Pseudomonas syringae*, *Pseudomonas fluorescens*, *Serratia marcescens*, *Acetobacter xylinum*, *Agrobacteria*, *Rhodopseudomonas spheroides*, *Xanthomonas campestris*, *Rhizobium melioli*, *Alcaligenes entrophus*, *Clavibacter xyli*, and *Azotobacter vinlandii*; and phytosphere yeast species such as *Rhodotorula rubra*, *R. glutinis*, *R. marina*, *R. aurantiaca*, *Cryptococcus albidus*, *C. diffluens*, *C. laurentii*, *Saccharomyces rosei*, *S. pretoriensis*, *S. cerevisiae*, *Sporobolomyces roseus*, *S. odorus*, *Kluyveromyces veronae*, and *Aureobasidium pullulans*.

The cyclotides of the embodiments can be used for any application including coating surfaces to target microbes. In this manner, target microbes include human pathogens or microorganisms. Surfaces that might be coated with the cyclotides of the embodiments include carpets and sterile medical facilities. Polymer bound polypeptides of the embodiments may be used to coat surfaces. Methods for incorporating compositions with antimicrobial properties into polymers are known in the art. See U.S. Patent No.5,847,047 herein incorporated by reference.

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

25

EXPERIMENTAL

The embodiments is further defined in the following Examples, in which parts and percentages are by weight and degrees are Celsius, unless otherwise stated. Techniques in molecular biology were typically performed as described in Ausubel or Sambrook, *supra*. It should be understood that these Examples, while indicating certain embodiments of the invention, are given by way of illustration only. From the above discussion and these Examples, one skilled in the art can ascertain the essential characteristics of this invention, and without departing from

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

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

Example 1: Expression Pattern of the Cyclo1 Gene

10 Evidence that the Cyclo1 gene is expressed in a root preferred manner was obtained using Lynx Massively Parallel Signature Sequencing technology (MPSS) (see Brenner S, *et al.* (2000) *Nature Biotechnology* 18:630-634, Brenner S *et al.* (2000) *Proc Natl Acad Sci USA* 97:1665-1670). This technology involves the generation of 17 base signature tags from mRNA samples that have been reverse
15 transcribed. The tags are simultaneously sequenced and assigned to genes or ESTs. The abundance of these tags is given a numerical value that is normalized to parts per million (PPM) which then allows the tag expression, or tag abundance, to be compared across different tissues. Thus, the MPSS platform can be used to determine the expression pattern of a particular gene and its expression level in
20 different tissues.

Searching the Lynx MPSS database identified signature tags that were root preferred. One of these tags corresponded to the Cyclo1 gene. The distribution of the Cyclo1 tag across different tissues represented in the database is shown in Figure 1. The combination of the quantitative, spatial and temporal characteristics
25 suggested that the Cyclo1 gene promoter is a suitable candidate to drive transgene expression in a root preferred manner in plants, such as maize. Such transgenes can include insecticidal genes but may also include other biotic and abiotic stress-resistance genes (drought, salt, cold, etc), agronomic trait genes and siRNAs and microRNAs.

**Example 2: 5' RACE to map the transcriptional start site
for the Cyclo1 mRNA.**

5' RACE was used to map the transcriptional start site for the Cyclo1
5 mRNA and was performed according to the protocol provided with the 5' RACE
System for the Rapid Amplification of cDNA Ends (Invitrogen, Carlsbad, CA) using
RNA isolated from roots of maize inbred B73 plants. Gene-specific primers (GSP)
were designed to the 3' UTR (using the EST sequence) as indicated below. GSP1
(SEQ ID NO: 2), located on the complementary strand from +284 to +314 relative
10 to the translation start site (ATG) of Cyclo1, was used for first strand synthesis.
Following RNase treatment, reaction clean-up and dC-tailing of the cDNA, PCR
amplification was performed with GSP2 (SEQ ID NO:3), also located on the
complementary DNA strand between +257 to +287 relative to the ATG, in
combination with the Abridged anchor primer (supplied in the kit).

15 PCR was performed using Pfx DNA Polymerase (Invitrogen). The PCR
conditions used were: 1 cycle of 93 °C for 2 minutes; followed by 35 cycles of the
combination of 93 °C for 15 seconds, 60 °C for 30 seconds, and 68 °C for 2
minutes; and completed by 1 cycle of 68 °C for 10 minutes. A second round of
PCR was conducted using GSP3 (SEQ ID NO: 4) with the kit supplied AUAP
20 primer using the first round PCR products as templates. Second round PCR was
performed with Platinum Taq DNA polymerase (Invitrogen) using the following
PCR conditions: 1 cycle of 94 °C for 2 minutes; followed by 35 cycles of 94 °C for
30 seconds, 55 °C for 30 seconds, and 72 °C for 2 minutes; and completed by 1
cycle of 72 °C for 5 minutes.

25 The 5' RACE products were TOPO-cloned into pCR4.0 (Invitrogen) and
sequenced using M13 fwd and rev primers. The 5' RACE sequence identified the
transcriptional start site for the cyclo1 mRNA to be located – 137 bp relative to the
cyclo1 translational (ATG) start site.

30 **Example 3: Isolation of the Promoter for the Cyclo1 Gene**

Sequence upstream of the translational start site of the Cyclo1 gene was
isolated using the Universal Genome Walker kit (BD Biosciences Clontech, Palo
Alto, CA).

Genomic libraries were constructed from maize B73 genomic DNA according to the kit protocol. Gene-specific primers (GSPs) were designed to the complement of the EST clone sequence corresponding to +138 to +171 bp (GSP7; SEQ ID NO: 5) and to +108 to +141 (GSP8; SEQ ID NO: 6) relative to the
5 Cyclo1 translational start site. Primary PCR reactions using primers GSP7 and AP1 and secondary reactions using GSP8 and AP2 (AP1 and AP2 are supplied in the kit) were performed using the Advantage cDNA PCR Kit (BD BioSciences Clontech). Touchdown PCR was conducted under the following conditions for both primary and secondary PCR reactions: 20 cycles of 94 °C for 15 seconds and
10 65 °C for 4 minutes – with each successive cycle being reduced by 0.5 °C for the annealing and extension reaction (e.g. 65 °C, 64.5 °C, 64 °C etc); followed by 15 cycles of 94 °C for 15 seconds and 55 °C for 4 minutes (35 cycles in total). The secondary PCR reactions used the products from the primary PCR as a template.

Genome walking (GW) PCR products from the secondary reaction were
15 TOPO cloned into pCR4.0 (Invitrogen) and sequenced. This set of GW reactions resulted in 480 bp of sequence upstream of the Cyclo1 translational start site. To obtain additional 5' flanking sequence, a second round of GW was performed using the primers GSP 9 (SEQ ID NO: 7) and GSP10 (SEQ ID NO: 8). These primers complement the newly obtained upstream sequence and are located –
20 364 to –334 and –397 to –364 bp relative to the Cyclo1 ATG, respectively. PCR was performed as described above and the final PCR products were TOPO cloned into pCR4.0 (Invitrogen) and sequenced. The combination of the two rounds of GW resulted in a total of 1140 bp of sequence upstream of the Cyclo1 translational start site.

25 The Cyclo1 5' flanking region was PCR amplified from genomic maize B73 DNA using the Advantage cDNA kit (BD BioSciences Clontech) and the primers BamC1 (SEQ ID NO: 9) and XhoC1 (SEQ ID NO: 10). BamHI and XhoI restriction enzyme recognition sites were added to the 5' end of the primers to facilitate subsequent sub-cloning. PCR conditions for amplification were: 1 cycle of 94 °C
30 for 5 minutes; 40 cycles of 94 °C for 30 seconds, 50 °C for 30 seconds, and 72 °C for 30 seconds; followed by 1 cycle of 72°C for 7 minutes. TOPO cloning into pCR4.0 (Invitrogen) and sequence verification showed that (-)1140 bp relative to the ATG of Cyclo1 had been obtained.

Example 4: Cyclo1 Promoter Sequence Analysis

Analysis of the Cyclo1 promoter sequence indicated the presence of some motifs of interest.

5 The TATA box was identified and is indicated in Figure 1. It is located at positions 976 through 981 of SEQ ID NO: 1.

The "ATATT" motif, previously identified as being present in other promoters with root specific expression (Elmayan & Tepfer (1995) *Transgenic Research* 4, 388-396), was identified in the Cyclo1 promoter in eight separate
10 locations. These locations can be seen in Figure 1.

The transcriptional start site (TSS) was mapped by 5' RACE (See Example 2) and is indicated at position 1004 in Figure 1.

15 Example 5: Promoter Activity of Cyclo1

Transient particle bombardment assays were performed to demonstrate that the 1140 bp Cyclo1 polynucleotide sequence functions as a promoter. These assays provided a rapid assessment of whether the DNA sequence is able to direct gene expression.

20 The 1140 bp genomic PCR fragment (see Example 3) was cloned into an expression vector in front of the B-glucuronidase (GUS) gene. Biolistic bombardment of 3-day-old maize seedlings with this expression cassette resulted in numerous GUS staining foci on the coleoptile (>20 foci/coleoptile). The level of staining was moderate compared to a control containing the strong, constitutive
25 promoter, *Ubi-1* directing GUS expression. However, the results indicated that the 1140 bp Cyclo1 DNA is able to function as a promoter.

Materials and Methods Utilized for the Biolistic Transient Root Expression Assay

30 B73 seeds were placed along the edge of a sheet of germination paper previously soaked in a 7% sucrose solution. An additional sheet of germination paper identical in size to the first was soaked in 7% sucrose and used to overlay the seeds. The germination paper – kernel – germination paper sandwich was subsequently rolled and placed into a beaker of 7% sucrose solution, such that the solution would wick up the paper to the kernels at the top of the roll. This process

allowed for straight root growth. Kernels were permitted to germinate and develop for 2-3 days in the dark at 27-28 °C. Prior to bombardment, the outer sheath covering the coleoptile was removed and then the seedlings were placed in a sterile petri dish (60 mm) on a layer of Whatman #1 filter paper moistened with 1 mL of H₂O. Two seedlings per plate were arranged in opposite orientations and anchored to the filter paper with a 0.5% agarose solution.

DNA/gold particle mixtures were prepared for bombardment in the following method. Sixty mg of 0.6 – 1.0 micron gold particles were pre-washed with ethanol, rinsed with sterile distilled H₂O, and resuspended in a total of 1 mL of sterile H₂O. 50 µL aliquots of gold particle suspension were stored in siliconized Eppendorf tubes at room temperature. DNA was precipitated onto the surface of the gold particles by combining, in order, 50 µL of pre-washed 0.6 µM gold particles, 5-10 µg of test DNA, 50 µL 2.5 M CaCl₂ and 25 µL of 0.1 M spermidine. The solution was immediately vortexed for 3 minutes and centrifuged briefly to pellet the DNA/gold particles. The DNA/gold was washed once with 500 µL of 100% ethanol and suspended in a final volume of 50 µL of 100% ethanol. The DNA/gold solution was incubated at –20 °C for at least 60 minutes prior to applying 6 µL of the DNA/gold mixture onto each Mylar™ macrocarrier.

Seedlings prepared as indicated above were bombarded twice using the PDS-1000/He gun at 1100 psi under 27-28 inches of Hg vacuum. The distance between macrocarrier and stopping screen was between 6-8 cm. Plates were incubated in sealed containers for 24 hours in the dark at 27-28 °C following bombardment.

After 18 - 24 hours of incubation the bombarded seedlings were assayed for transient GUS expression. Seedlings were immersed in 10 -15 mL of GUS assay buffer containing 100 mM NaH₂PO₄-H₂O (pH 7.0), 10 mM EDTA, 0.5 mM K₄Fe(CN)₆-3H₂O, 0.1% Triton X-100 and 2 mM 5-bromo-4-chloro-3-indoyl glucuronide. The tissues were incubated in the dark for 24 hours at 37 °C. Replacing the GUS staining solution with 100% ethanol stopped the assay. GUS expression/staining was visualized under a microscope.

Example 6: Expression Pattern of Cyclo1

Stably transformed plants were created to allow for a more detailed characterization of promoter activity, including spatial and quantitative regulation of the promoter. The 1140 bp Cyclo1 promoter (SEQ ID NO: 1) was operably linked to the GUS gene (abbreviated as Cyclo1: GUS) which allowed for the detection of promoter activity by histochemically staining plant tissues for GUS enzymatic activity. The pattern of GUS staining reflects where the promoter is active.

The spatial expression pattern directed by the Cyclo1 promoter showed a root preference (Tables 1 & 2). GUS expression was detected predominantly in the roots of young plants growing on nutrient agarose and in the roots of greenhouse-grown plants when they reached V5-V6 stage (5-6 collared leaves) in development. These results supported the MPSS results in Example 1 and indicated that the 1140 bp 5' flanking region was sufficient for root preferred expression.

Table 1: MUG Assay Results For The Cyclo1 Promoter

	Leaf	Root Tip	Mature Region
Cyclo1 (-1140)	0	48	764
untransformed (negative control)	0	0	0

Values given are median values, as nmole MU/mg total protein/hr

MU= 4-methyl umbelliferone

MUG = 4-methyl umbelliferyl-B-D-glucuronide

Table 2: Plant Expression Results For The Cyclo1 Promoter

	V5 – V6		R1 – R2			Mature
	Root	Leaf	Pollen	Silk	Tassel	Kernel
Cyclo1 (-1140)	+++++	++	-	-(+)	++	+
untransformed (negative control)	-	-	-	-	-	-

Further analysis of the greenhouse plants showed that the expression was found mostly in the mature areas of the nodal root and not in the rapidly dividing and expanding regions of the root tips. This expression pattern was also observed in the root tips of lateral roots, even in lateral roots emerging from mature areas that stained heavily for GUS.

Quantitative analysis of GUS expression supported the spatial results above and showed that expression levels were highest in mature regions of the root and lower in the root tips. Analysis of the results on a per plant basis indicated that a few of the plants did have some low level of GUS expression. This was supported by histochemical analysis which showed expression was mostly confined to the leaf veins.

Plants were allowed to develop to the reproductive stage of R1 (denoted by silking and pollen shed). At this stage, attempts were made to detect Cyclo1-directed GUS expression in pollen, tassels, and silks. No expression was observed histochemically in pollen from any of the plants. Similar results were obtained when silks were scored. No expression was detected with the exception of a few silk strands from 2 plants. In tassels, analysis revealed that approximately half of the plants stained for GUS activity. The staining intensity tended to be low, and in comparison to the level of staining in the roots, expression in the tassels was below the level in roots.

Kernels were also histochemically stained after mature ears were harvested from the plants. Results showed that approximately half of the plants assayed had kernels that did not demonstrate any expression. In those that showed GUS activity, the staining was mostly confined to the abscission layer and to an area of the scutellum granular layer near the abscission layer.

Histochemical Staining Of Plant Tissues For GUS Activity

Detection of GUS activity was accomplished by placing tissue from transformed plants into 48-well, 12-well or 6-well plates containing 0.5 to 5 mL GUS assay buffer (assay buffer recipe described in Example 5). Plates were placed under house vacuum for 10 min, and then incubated overnight at 37 °C. Tissue was cleared of pigmentation with 1 to 3 successive 12-hour incubations in

100% ethanol at room temperature. The tissues were stored in 70% ethanol at 4 °C.

Example 7: *Agrobacterium*-Mediated Transformation of Maize and Regeneration of Transgenic Plants

For *Agrobacterium*-mediated transformation of maize with a promoter sequence of the embodiments, the method of Zhao was employed (U.S. Patent No. 5,981,840, (hereinafter the '840 patent) and PCT patent publication WO98/32326; the contents of which are hereby incorporated by reference).

Agrobacterium were grown on a master plate of 800 medium and cultured at 28°C in the dark for 3 days, and thereafter stored at 4°C for up to one month. Working plates of *Agrobacterium* were grown on 810 medium plates and incubated in the dark at 28°C for one to two days.

Briefly, embryos were dissected from fresh, sterilized corn ears and kept in 561Q medium until all required embryos were collected. Embryos were then contacted with an *Agrobacterium* suspension prepared from the working plate, in which the *Agrobacterium* contained a plasmid comprising the promoter sequence of the embodiments. The embryos were co-cultivated with the *Agrobacterium* on 562P plates, with the embryos placed axis down on the plates, as per the '840 patent protocol.

After one week on 562P medium, the embryos were transferred to 563O medium. The embryos were subcultured on fresh 563O medium at 2 week intervals and incubation was continued under the same conditions. Callus events began to appear after 6 to 8 weeks on selection.

After the calli had reached the appropriate size, the calli were cultured on regeneration (288W) medium and kept in the dark for 2-3 weeks to initiate plant regeneration. Following somatic embryo maturation, well-developed somatic embryos were transferred to medium for germination (272V) and transferred to a lighted culture room. Approximately 7-10 days later, developing plantlets were transferred to 272V hormone-free medium in tubes for 7-10 days until plantlets were well established. Plants were 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.

Media used in *Agrobacterium*-mediated transformation and regeneration of
5 transgenic maize plants:

561Q medium 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, 68.5 g/L sucrose, 36.0 g/L glucose, 1.5 mg/L 2,4-D, and 0.69 g/L L-proline (brought to volume with dl H₂O following adjustment to pH 5.2 with KOH); 2.0 g/L Gelrite™
10 (added after bringing to volume with dl H₂O); and 8.5 mg/L silver nitrate (added after sterilizing the medium and cooling to room temperature).

800 medium comprises 50.0 mL/L stock solution A and 850 mL dl H₂O, and brought to volume minus 100 mL/L with dl H₂O, after which is added 9.0 g of phytagar. After sterilizing and cooling, 50.0 mL/L stock solution B is added, along
15 with 5.0 g of glucose and 2.0 mL of a 50 mg/mL stock solution of spectinomycin. Stock solution A comprises 60.0 g of dibasic K₂HPO₄ and 20.0 g of monobasic sodium phosphate, dissolved in 950 mL of water, adjusted to pH 7.0 with KOH, and brought to 1.0 L volume with dl H₂O. Stock solution B comprises 20.0 g
20 NH₄Cl, 6.0 g MgSO₄•7H₂O, 3.0 g potassium chloride, 0.2 g CaCl₂, and 0.05 g of FeSO₄•7H₂O, all brought to volume with dl H₂O, sterilized, and cooled.

810 medium comprises 5.0 g yeast extract (Difco), 10.0 g peptone (Difco), 5.0 g NaCl, dissolved in dl H₂O, and brought to volume after adjusting pH to 6.8. 15.0 g of bacto-agar is then added, the solution is sterilized and cooled, and 1.0 mL of a 50 mg/mL stock solution of spectinomycin is added.

562P medium 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 dl H₂O following adjustment to pH 5.8 with KOH); 3.0 g/L Gelrite™ (added after bringing to volume with dl
25 H₂O); and 0.85 mg/L silver nitrate and 1.0 mL of a 100mM stock of acetosyringone
30 (both added after sterilizing the medium and cooling to room temperature).

563O medium 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, 1.5 mg/L 2,4-D, 0.69 g L-proline, and 0.5 g MES buffer (brought to

volume with dl H₂O following adjustment to pH 5.8 with KOH). Then, 6.0 g/L Ultrapure™ agar-agar (EM Science) is added and the medium is sterilized and cooled. Subsequently, 0.85 mg/L silver nitrate, 3.0 mL of a 1 mg/mL stock of Bialaphos, and 2.0 mL of a 50 mg/mL stock of carbenicillin are added.

5 288 W 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 H₂O) (Murashige and Skoog (1962) *Physiol. Plant.* 15:473), 100 mg/L myo-inositol, 0.5 mg/L zeatin, and 60 g/L sucrose, which is then brought to volume with polished D-I
10 H₂O after adjusting to pH 5.6. Following, 6.0 g/L of Ultrapure™ agar-agar (EM Science) is added and the medium is sterilized and cooled. Subsequently, 1.0 mL/L of 0.1 mM abscisic acid; 1.0 mg/L indoleacetic acid and 3.0 mg/L Bialaphos are added, along with 2.0 mL of a 50 mg/mL stock of carbenicillin.

Hormone-free medium (272V) comprises 4.3 g/L MS salts (GIBCO 11117-
15 074), 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 dl H₂O), 0.1 g/L myo-inositol, and 40.0 g/L sucrose (brought to volume with polished dl H₂O after adjusting pH to 5.6); and 6 g/L Bacto-agar (added after bringing to volume with polished dl H₂O), sterilized and cooled to
20 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
25 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
30 of the appended claims.

THAT WHICH IS CLAIMED:

1. An isolated nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:
 - 5 a) a nucleotide sequence comprising the sequence set forth in SEQ ID NO:1 or a complement thereof;
 - b) a nucleotide sequence comprising the plant promoter sequences of the plasmids deposited as Patent Deposit No: NRRL B-30794, or a complement thereof;
 - 10 c) a nucleotide sequence comprising at least 20 contiguous nucleotides of the sequence set forth in SEQ ID NO:1, wherein said sequence initiates transcription in a plant cell; and
 - d) a nucleotide sequence comprising a sequence having at least 95% sequence identity to the sequence set forth in SEQ ID NO:1, wherein said
15 sequence initiates transcription in the plant cell.
2. A DNA construct comprising a nucleotide sequence of claim 1 operably linked to a heterologous nucleotide sequence of interest.
3. A vector comprising the DNA construct of claim 2.
- 20 4. A plant cell having stably incorporated into its genome the DNA construct of claim 2.
 5. The plant cell of claim 4, wherein said plant cell is from a monocot.
 6. The plant cell of claim 5, wherein said monocot is maize.
 7. The plant cell of claim 4, wherein said plant cell is from a dicot.
 - 25 8. A plant having stably incorporated into its genome the DNA construct of claim 2.
 9. The plant of claim 8, wherein said plant is a monocot.
 10. The plant of claim 9, wherein said monocot is maize.

11. The plant of claim 8, wherein said plant is a dicot.

12. A transgenic seed of the plant of any one of claims 8 to 11, wherein the seed comprise the DNA construct of claim 2.

13. The plant of claim 8, wherein the heterologous nucleotide sequence
5 of interest encodes a gene product that confers herbicide, salt, cold, drought, pathogen, or insect resistance.

14. A method for expressing a nucleotide sequence in a plant, said method comprising introducing into a plant a DNA construct, said DNA construct comprising a promoter and operably linked to said promoter a heterologous
10 nucleotide sequence of interest, wherein said promoter comprises a nucleotide sequence selected from the group consisting of:

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

b) a nucleotide sequence comprising the plant promoter
15 sequences of the plasmids designated as Patent Deposit No: NRRL B-30794;

c) a nucleotide sequence comprising at least 20 contiguous nucleotides of the sequence set forth in SEQ ID NO:1, wherein said nucleotide sequence initiates transcription in said plant; and

d) a nucleotide sequence comprising a sequence having at least
20 95% sequence identity to the sequence set forth in SEQ ID NO:1 or SEQ ID NO:2, wherein said nucleotide sequence initiates transcription in a plant cell.

15. The method of claim 14, wherein said heterologous nucleotide sequence of interest is selectively expressed in the root.

16. A method for expressing a nucleotide sequence in a plant cell, said
25 method comprising introducing into a plant cell a DNA construct comprising a promoter operably linked to a heterologous nucleotide sequence of interest, wherein said promoter comprises a nucleotide sequence selected from the group consisting of:

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

b) a nucleotide sequence comprising the plant promoter sequences of the plasmids designated as Patent Deposit No: NRRL B-30794;

c) a nucleotide sequence comprising at least 20 contiguous nucleotides of the sequence set forth in SEQ ID NO:1, wherein said nucleotide sequence initiates transcription in said plant cell; and

d) a nucleotide sequence comprising a sequence having at least 95% sequence identity to the sequence set forth in SEQ ID NO:1, wherein said nucleotide sequence initiates transcription in a plant cell.

10 17. A method for selectively expressing a nucleotide sequence in a plant root, said method comprising introducing into a plant cell a DNA construct, and regenerating a transformed plant from said plant cell, said DNA construct comprising a promoter and a heterologous nucleotide sequence operably linked to said promoter, wherein said promoter comprises a nucleotide sequence selected

15 from the group consisting of:

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

b) a nucleotide sequence comprising the plant promoter sequences of the plasmids deposited as Patent Deposit No: NRRL B-30794;

20 c) a nucleotide sequence comprising at least 20 contiguous nucleotides of the sequence set forth in SEQ ID NO:1, wherein said sequence initiates transcription in a plant root cell; and

d) a nucleotide sequence comprising a sequence having at least 95% sequence identity to the sequence set forth in SEQ ID NO:1, wherein

25 said sequence initiates transcription in a plant root cell.

18. The method of claim 17, wherein expression of said heterologous nucleotide sequence alters the phenotype of said plant.

19. The method of any one of claims 14 or 17, wherein the plant is a

30 monocot.

20. The method of claim 16, wherein the plant cell is from a monocot.

21. The method of any one of claims 19 or 20, wherein the monocot is maize.
22. The method of any one of claims 14 or 17, wherein the plant is a dicot.
- 5 23. The method of claim 16, wherein the plant cells from a dicot.
24. The method of any one of claims 14, 16, or 17, wherein the heterologous nucleotide sequence encodes a gene product that confers herbicide, salt, pathogen, or insect resistance.
- 10 25. An isolated nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:
- (a) a nucleotide sequence set forth in SEQ ID NO: 11;
 - (b) a nucleotide sequence that encodes a polypeptide having the amino acid sequence set forth in SEQ ID NO: 12 or 13;
 - (c) a nucleotide sequence that comprises the full length complement
- 15 of (a) or (b).
26. The nucleic acid molecule according to claim 25, wherein the nucleotide sequence is optimized for expression in a plant.
- 20 27. A DNA construct comprising a nucleotide sequence of claim 25, wherein said nucleotide sequence is operably linked to a promoter that drives expression in a host cell.
28. An expression cassette comprising the DNA construct of claim 27.
- 25 29. A host cell having stably incorporated into its genome at least one DNA construct of claim 27.
30. The host cell of claim 29, wherein said host cell is a plant cell.
- 30

31. A plant having stably incorporated into its genome the DNA construct of claim 27.

5

32. The plant according to claim 31, wherein said plant is a monocot.

33. The plant according to claim 32, wherein said monocot is maize.

34. The plant according to claim 31, wherein said plant is a dicot.

10

35. The plant according to claim 34, wherein said dicot is soybean.

36. Transformed seed of the plant of any one of claims 31-35, wherein the seed comprise the DNA construct of claim 27.

15

37. An isolated cyclotide comprising the amino acid sequence set forth in SEQ ID NOs: 12 or 13.

38. The isolated cyclotide of claim 37, wherein said cyclotide is characterized by insecticidal activity against at least one plant pest.

20

39. A method for impacting a plant pest comprising introducing into a plant or cell thereof at least one nucleotide construct comprising a nucleotide sequence operably linked to a promoter that drives expression of a coding sequence in plant cells, wherein said nucleotide sequence is selected from the group consisting of:

25

(a) a nucleotide sequence set forth in SEQ ID NO: 11;

(b) a nucleotide sequence that encodes a polypeptide having the amino acid sequence set forth in SEQ ID NO: 12 or 13;

(c) a nucleotide sequence that comprises the full length complement of (a) or (b).

30

40. The method of claim 39, wherein said plant produces a polypeptide characterized by insecticidal activity against at least one plant pest.

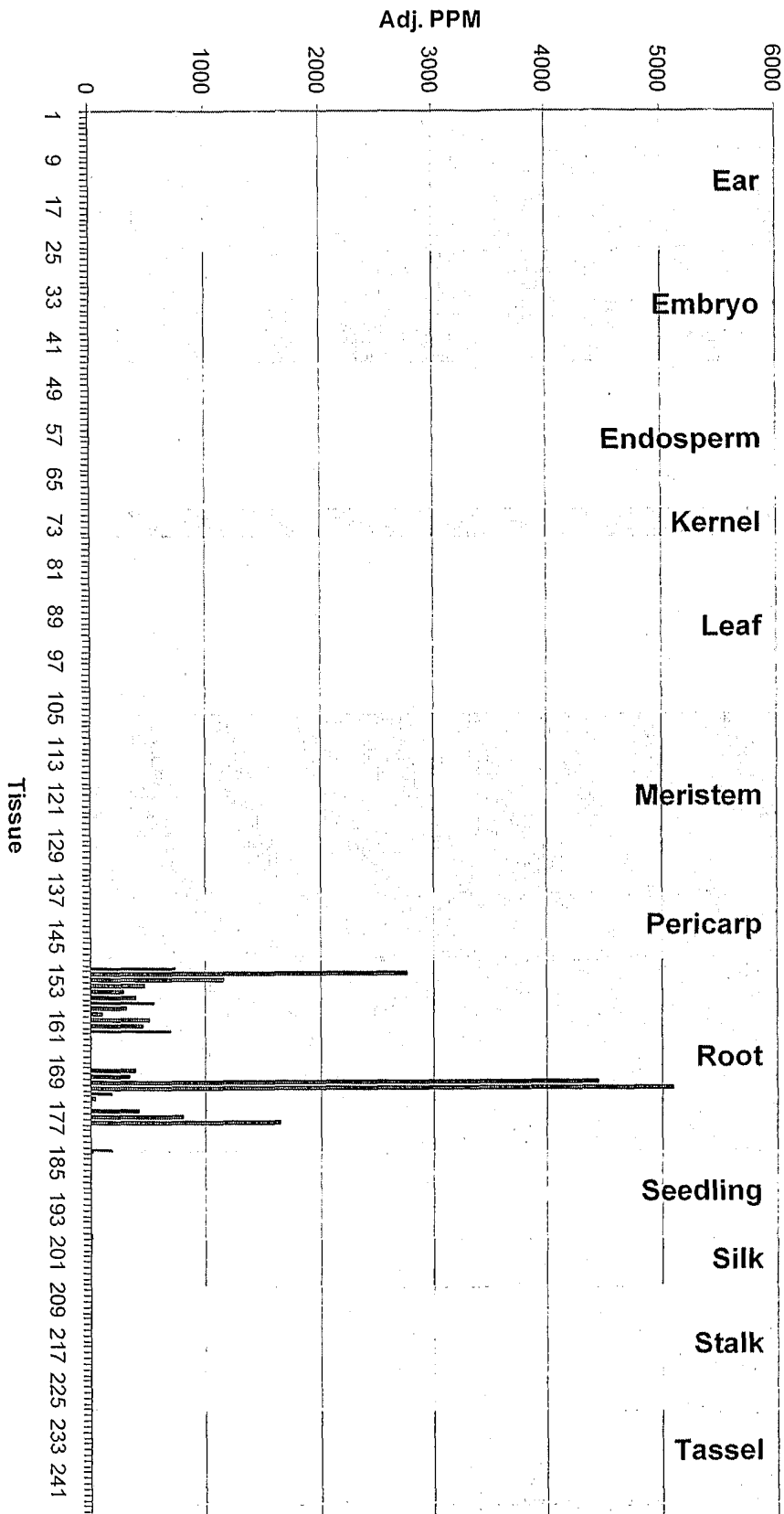


Figure 1: Cyclo1 Expression Profile

Figure 2

Cyclo1 Promoter Sequence with 5' UTR

```

1 CCTCCTCGCC TTGTGGCTCC TCCTGAACCA CCTGCTCTTC TCCTGTGGGG GGGTGTGAGA
61 CAGCAAGGGT GAGCTCACAC ATGATCATAG CTCAACAAGT TGTGGGGAAC CAGTGGACAT
121 GAACTCACAA AGGTGGGAGT TCATGTGATG GTTCCTCTAG ATGCTCAACT TGTTCATTA
181 TATTACGCAA TTGCTCCGAC GCTTCATCAA TTAATCCTCC TTAGTGATAT TAAATAACGG
241 AATAATATTA GAGAAATAAA CAATAATCTA AGACATTAGC GCATAAAGAT GTGACAAAAT
301 GATTGAGTCT GGTCAATATTA CCCTCCTTCA TCCTTTATTG CATAAAAAGAT TGTAGTTTAC
361 ACCTTCGGCT TTACAAAGGA GAGCTCGAAG GTAATATTAC AGCTTCGAAG GCGGAGTGAT
421 TTGATTCTCC CTGTTCAAA AAGCGAGATC TCTTCATATC ATTGTGCCTC TATTTATAGT
481 AACCAAGTAC AATTTCATAT GAAATTACAA ACATGCTCAT GGACATGATA ATTCCAGTGC
541 ACATCCAACC CTGCTTGATA CAAAACATGC TCATAATCAT GATGATTCAA GTGCACATCC
601 ACCCTGCTCG ATACAACAGT TGGCGACCTG GTGTGAGAGT CAGACCAGAC GGGCTTTCAC
661 AATCGCCATG CATGTCATTC TCTCGTGGTC CACGTGTTTA TTAATATTGC CATTAATTGG
721 AGGGAAATAA AATCAACAAG AATAGCTTAT TGATGAGTCA TATATTATGA ATACATCTTA
781 TCATCTTACC AAACAAAAAC ATATGACCGT CGATGACCTG AAAGTAGACT ATTCGGGATC
841 TGCAATGATC TGCTTGFAA TATTAATTTG CACATCACGC CATTGCATGC ACATCGGCGT
901 GGCATTATT AATTTGGATT GGACGAAAAA TCAACCAGAG GCGGTCACCC TTTTGCTAGT
961 TGGCCTTGTA ATACTTATAA TAATTATCCG TATAGTCTAG TACGTACGGG ACATCACGCC
1021 ATTGCCCTGTT GTGTATAAAA AGCGAGCATG AGCTAGGTTG GGAGCAGAGC AAAGCGTAGT
1081 CATCACCTGT GTCTAGGTTG GGAGCAAAGC AAAGAGAGAG AGGAAAGCTA GCTAGCTAGC

```

The **TATA box** is shown in bold, underlined text.

ATATT – Eight ATATT root motifs are shown in underlined text.

The transcriptional start site (TSS) mapped by 5' RACE is indicated at position 1004 by a "G" in a box.

Figure 3

Multiple Sequence Alignment

	1		32	
Cycloviolacin H1	~GIPCGESCV	YIPC.LTAAI	GC SKSKVCY	RN (SEQ ID NO: 16)
Circulin_A	~GIPCGESCV	WIPC.ISAAL	GC SKKNVCY	RN (SEQ ID NO: 15)
Circulin_B	GVI PC GESCV	FIPC.ISTILL	GC SKKNVCY	RN (SEQ ID NO: 14)
Cyclopsychoctrin A	~SIPCGESCV	FIPCTV TALL	GC SKSKVCY	KN (SEQ ID NO: 18)
Cycloviolacin O1	~GIPCAESCV	YIPCTV TALL	GC SSNRVCY	N~ (SEQ ID NO: 17)
Cyclol_sequence	GAISCGESCV	IIPC.VSTILL	GC RCE NK L CV	K ~ (SEQ ID NO: 13)

SEQUENCE LISTING

<110> Pioneer Hi-Bred International, Inc.
E. I. DuPont de Nemours & Co.

<120> MAIZE CYCLO1 GENE AND PROMOTER

<130> 1880-PCT

<150> 60/643,720

<151> 2005-01-13

<160> 18

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 1140

<212> DNA

<213> Zea mays

<400> 1

```

cctcctcggc ttgtggctcc tctgaacca cctgctcttc tctgtgggg ggggtgtgaga 60
cagcaagggg gagctcacac atgatcatag ctcaacaagt tgtggggaac cagtggacat 120
gaactcacia aggtggggagt tcatgtgatg gttcctctag atgctcaact tgttgcatta 180
tattacgcaa ttgctccgac gttcatcaa ttaatcctcc ttagtgatat taaataacgg 240
aataatatta gagaaataaa caataatcta agacattagc gcataaagat gtgacaaaat 300
gattgagtct ggtcatatta ccctccttca tcttttattg cataaaagat tgtagtgttac 360
acctcgggct ttacaaaagga gagctcgaag gtaatattac agcttcgaag gcgagtgat 420
ttgattctcc cttgttcaaa aagcgagatc tctcatatc attgtgcctc tatttatagt 480
aaccaagtac aatctcatat gaaattaca acatgctcat ggacatgata attccagtgc 540
acatccaacc ctgcttgata caaaacatgc tcataatcat gatgattcaa gtgcacatcc 600
accctgctcg atacaacagt tggcgacctg gtgtgagagt cagaccagac gggctttcac 660
aatcgccatg catgtcattc tctcgtggtc cacgtgttta ttaatattgc cattaattgg 720
agggaaataa aatcaacaag aatagcttat tgatgagtca tatattatga atacatctta 780
tcatcttacc aaacaaaaac atatgacctg cgatgacctg aaactagact attcgggatc 840
tgcaatgatc tgcttgtaaa tattaatttg cacatcacgc cattgcatgc acatcggcgt 900
gggcattatt aatttgattt ggacgaaaaa tcaaccagag ggcgtcacc ttttgctagt 960
tggccttgta atacttataa taattatccg tatagtctag tacgtacggg acatcacgcc 1020
attgctgtgt gtgtataaaa agcgagcatg agctagggtt ggagcagagc aaagcgtagt 1080
catcacctgt gtctagggtt ggagcaaacg aaagagagag aggaaagcta gctagctagc 1140

```

<210> 2

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide Primer GSP1

<400> 2

tcacacgtaa agggatcaga gatatatggc

30

<210> 3

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide Primer GSP2

<400> 3

ggcctcgtcc cgactgtagc agcggcgtca

30

<210> 4

<211> 36

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide Primer GSP3

<400> 4

tcacgagttt atttcacgca gagtttgttt tcacat

36

<210> 5

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide Primer GSP7

<400> 5

ggagatggct ccatctggat ggccaaactc cct

33

<210> 6

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide Primer GSP8

<400> 6

cctcgcgatg aagtccaggc caaacacggg gga

33

<210> 7

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide Primer GSP9

<400> 7

tcatatgttt ttgtttggta agatgataag

30

<210> 8

<211> 32

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide Primer GSP10

<400> 8

atgtattcat aatatatgac tcatcaataa gc

32

```

<210> 9
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide Primer BamC1

<400> 9
ggatccctcc tcgccttggtg gtcctcctg aaccac 36

<210> 10
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide Primer XhoC1

<400> 10
ctcgaggcta gctagctagc tttcctctct ctcttt 36

<210> 11
<211> 519
<212> DNA
<213> Zea mays

<220>
<221> CDS
<222> (70)...(321)
<223> coding region for cyclol1 gene

<400> 11
aagcgtagtc atcacctgtg tctaggttgg gagcaaagca aagagagaga ggaaagctag 60
ctagctagc atg gag agt ggc agc aag aag gtc gca gcc ggt gtg ttg gtg 111
      Met Glu Ser Gly Ser Lys Lys Val Ala Ala Gly Val Leu Val
          1             5             10

ctg ctg ctt ctc cag ctg atg gta gct ccg acg acg gcg aca gcc cgc 159
Leu Leu Leu Leu Gln Leu Met Val Ala Pro Thr Thr Ala Thr Ala Arg
 15             20             25             30

ctc ctg cag gcg gat acc tcc ccc gtg ttt ggc ctg gac ttc atc gcg 207
Leu Leu Gln Ala Asp Thr Ser Pro Val Phe Gly Leu Asp Phe Ile Ala
          35             40             45

agg gag ttt ggc cat cca gat gga gcc atc tcc tgt ggt gaa tca tgt 255
Arg Glu Phe Gly His Pro Asp Gly Ala Ile Ser Cys Gly Glu Ser Cys
          50             55             60

gtc atc ata cca tgc gtt tcg aca ctc ttg gga tgc cga tgt gaa aac 303
Val Ile Ile Pro Cys Val Ser Thr Leu Leu Gly Cys Arg Cys Glu Asn
          65             70             75

aaa ctc tgc gtg aaa taa actcgtgacg ccgctgctac agtcgggacg 351
Lys Leu Cys Val Lys *
      80

aggccatata tctctgatcc ctttacgtgt gaagaaccag ctccatcttt gtaacaaatt 411
tgtgttggtg gaaacaaaaa gacgggtggtg tttaacttat tgtgaaaggt gatttaataa 471

```

taatatataa taaacgggtac gttggatgtg tgtactcaaa aaaaaaaaa

519

<210> 12
 <211> 83
 <212> PRT
 <213> Zea mays

<400> 12
 Met Glu Ser Gly Ser Lys Lys Val Ala Ala Gly Val Leu Val Leu Leu
 1 5 10 15
 Leu Leu Gln Leu Met Val Ala Pro Thr Thr Ala Thr Ala Arg Leu Leu
 20 25 30
 Gln Ala Asp Thr Ser Pro Val Phe Gly Leu Asp Phe Ile Ala Arg Glu
 35 40 45
 Phe Gly His Pro Asp Gly Ala Ile Ser Cys Gly Glu Ser Cys Val Ile
 50 55 60
 Ile Pro Cys Val Ser Thr Leu Leu Gly Cys Arg Cys Glu Asn Lys Leu
 65 70 75 80
 Cys Val Lys

<210> 13
 <211> 30
 <212> PRT
 <213> Zea mays

<220>
 <221> PEPTIDE
 <222> (0)...(0)
 <223> Mature Cyclol peptide

<400> 13
 Gly Ala Ile Ser Cys Gly Glu Ser Cys Val Ile Ile Pro Cys Val Ser
 1 5 10 15
 Thr Leu Leu Gly Cys Arg Cys Glu Asn Lys Leu Cys Val Lys
 20 25 30

<210> 14
 <211> 31
 <212> PRT
 <213> Chassalia parviflora

<220>
 <221> PEPTIDE
 <222> (0)...(0)
 <223> Circulin B Peptide

<400> 14
 Gly Val Ile Pro Cys Gly Glu Ser Cys Val Phe Ile Pro Cys Ile Ser
 1 5 10 15
 Thr Leu Leu Gly Cys Ser Cys Lys Asn Lys Val Cys Tyr Arg Asn
 20 25 30

<210> 15
 <211> 30
 <212> PRT
 <213> Chassalia parviflora

<220>
 <221> PEPTIDE
 <222> (0)...(0)
 <223> Circulin A peptide

<400> 15
 Gly Ile Pro Cys Gly Glu Ser Cys Val Trp Ile Pro Cys Ile Ser Ala
 1 5 10 15
 Ala Leu Gly Cys Ser Cys Lys Asn Lys Val Cys Tyr Arg Asn
 20 25 30

<210> 16
 <211> 30
 <212> PRT
 <213> Viola hederacea

<220>
 <221> PEPTIDE
 <222> (0)...(0)
 <223> Cycloviolacin H1 peptide

<400> 16
 Gly Ile Pro Cys Gly Glu Ser Cys Val Tyr Ile Pro Cys Leu Thr Ser
 1 5 10 15
 Ala Ile Gly Cys Ser Cys Lys Ser Lys Val Cys Tyr Arg Asn
 20 25 30

<210> 17
 <211> 30
 <212> PRT
 <213> Viola odorata

<220>
 <221> PEPTIDE
 <222> (0)...(0)
 <223> Cycloviolacin O1 peptide

<400> 17
 Gly Ile Pro Cys Ala Glu Ser Cys Val Tyr Ile Pro Cys Thr Val Thr
 1 5 10 15
 Ala Leu Leu Gly Cys Ser Cys Ser Asn Arg Val Cys Tyr Asn
 20 25 30

<210> 18
 <211> 31
 <212> PRT
 <213> Psychotria longipes

<220>
 <221> PEPTIDE
 <222> (0)...(0)
 <223> Cyclopsychotride A peptide

<400> 18
 Ser Ile Pro Cys Gly Glu Ser Cys Val Phe Ile Pro Cys Thr Val Thr
 1 5 10 15
 Ala Leu Leu Gly Cys Ser Cys Lys Ser Lys Val Cys Tyr Lys Asn
 20 25 30

Applicant's or agent's file reference	1880-PCT	International application No.	PCT/US2006/000153
---------------------------------------	----------	-------------------------------	-------------------

**INDICATIONS RELATING TO DEPOSITED MICROORGANISM
OR OTHER BIOLOGICAL MATERIAL**

(PCT Rule 13bis)

A. The indications made below relate to the deposited microorganism or other biological material referred to in the description on page <u>7</u> , line <u>22-28</u> .	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution NRRL Agricultural Research Service Culture Collection	
Address of depositary institution (including postal code and country) 1815 North University Street Peoria, Illinois 61604 United States of America	
Date of deposit 01 December 2004	Accession Number NRRL B-30794
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")	

For receiving Office use only
<input type="checkbox"/> This sheet was received with the international application
Authorized officer

For International Bureau use only
<input type="checkbox"/> This sheet was received by the International Bureau on:
Authorized officer