

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
7 February 2002 (07.02.2002)

PCT

(10) International Publication Number
WO 02/10362 A2

(51) International Patent Classification⁷: **C12N 15/00**

52, CH-4125 Riehen (CH). **GLAZOV, Evgueni, Alexandrovich** [RU/CH]; Florastrasse 11, CH-4057 Basel (CH).

(21) International Application Number: PCT/EP01/08825

(22) International Filing Date: 30 July 2001 (30.07.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/222,202 1 August 2000 (01.08.2000) US

(71) Applicants (*for all designated States except US*):
SYNGENTA PARTICIPATIONS AG [CH/CH]; Schwarzwaldallee 215, CH-4058 Basel (CH). **NOVARTIS FORSCHUNGSSTIFTUNG, ZWEIGNIEDERLASSUNG FRIEDRICH MIESCHER INSTITUTE FOR BIOMEDICAL RESEARCH** [CH/CH]; Maulbeerstrasse 66, CH-4058 Basel (CH).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): **LEVIN, Joshua, Zvi** [US/US]; Syngenta, 3054 Cornwallis Road, Research Triangle Park, NC 27709 (US). **PHILLIPS, Kenneth, Lyon** [US/US]; 2713 Spencer Street, Durham, NC 27705 (US). **BUDZISZEWSKI, Gregory, Joseph** [US/US]; Syngenta, 3054 Cornwallis Road, Research Triangle Park, NC 27709 (US). **MEINS, Frederick, Jr.** [US/CH]; Leimgrubenweg

(74) Agent: **BASTIAN, Werner**; c/o Syngenta Participations AG, Intellectual Property, P.O. Box, CH-4002 Basel (CH).

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

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

Published:

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

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.

(54) Title: METHODS OF CONTROLLING GENE EXPRESSION AND GENE SILENCING

(57) Abstract: The present invention relates to methods to regulate gene expression in plants. In particular, manipulation of the expression in a plant cell of a nucleotide sequence encoding a polypeptide comprising a 3'-5' exonuclease domain is disclosed. More stable and predictable expression is thus obtained. The present invention also relates to method of increasing or decreasing port-transcriptional silencing. The invention further relates to novel nucleic acid molecules comprising nucleotide sequences encoding polypeptides comprising a 3'-5' exonuclease domain.



WO 02/10362 A2

METHODS OF CONTROLLING GENE EXPRESSION AND GENE SILENCING

The present invention relates to the field of molecular biology, in particular to the regulation of gene expression in plants and to gene silencing. The present invention also relates to novel isolated nucleic acid molecules comprising nucleotide sequences encoding novel polypeptides comprising a 3'-5' exonuclease domain.

Developments in the techniques of molecular biology and transformation have allowed the production of transgenic plants with various desirable traits. However, in some transgenic lines, the loss of expression of previously active genes has been observed in response to developmental, environmental or unknown signals. This phenomenon is commonly referred to as gene silencing. It occurs at a frequency higher than that of mutations, yet is markedly stable during somatic transmission. Chromosomal position or structure of the affected loci are factors determining the frequency and strength of gene silencing, and inactivation seems to preferentially affect genes present in multiple copies and is thought to be a consequence of sequence redundancy. While post-transcriptional silencing seems to mainly involve the formation of aberrant RNA molecules and is occasionally, but not necessarily, accompanied by DNA methylation, silencing that interferes with transcription initiation is more strictly correlated with hypermethylation of the DNA and possibly with alteration of chromatin structure at the silent loci. Currently, posttranscriptional gene silencing (PTGS) generally refers to the epigenetic inactivation of gene expression resulting from the specific degradation of mRNAs derived from genes with transcribed regions similar in sequence (Meins (2000) *Plant Mol. Biol.* 43: 261-273).

There have been attempts to understand the mechanism of gene silencing in plants. For example, in *Arabidopsis* two lines with independent mutant loci *egs1* and *egs2* were isolated from a M2 population by direct screening for silencing of an *Agrobacterium rhizogenes rolB* transgene (Dehio and Schell (1994) *PNAS* 91:5538-42). The *egs1* mutation appears to lead to the inactivation of this *rolB* transgene, and consequently, the wild type *egs1* allele may actively prevent silencing. Other mutants affected in post-transcriptional gene silencing (*sgs1* and *sgs2*, for suppressor of gene silencing) have been described in Elmayan et al. (1998) *Plant Cell* 10:1747-58. In this case, mutant plants carried a recessive monogenic mutation that appears to be involved in the release of silencing. In yet another report, disruption of a gene called MOM released transcriptional silencing of methylated genes (Amedeo et al. (2000) *Nature* 405:203-206). Although promising, these results are still preliminary.

Recently, five RecQ -like proteins have been isolated and characterized from *Arabidopsis thaliana* (Hartung et al. (2000) Nucleic Acids Research 28:4275-4282). These proteins are proposed to be involved in processes linked to DNA replication, DNA recombination and gene silencing.

The cellular functions involved in the switch from active to inactive genes are still not known, and tools allowing one skilled in the art to manipulate this phenomenon are lacking. One such enzyme that is proposed to be involved are exonucleases. A recent review of exoribonuclease superfamilies analyzed the structure and phylogenetic distribution of known exoribonucleases (Zuo et al. (2001) Nucleic Acid Res. 29:1017-1026). The authors grouped the exoribonucleases into six superfamilies and various subfamilies. The article furthered proposed common motifs to be used to characterize newly-discovered enzymes. In the production of transgenic plants with improved characteristics large numbers of independent transgenic lines have to be tested through several generations to ensure that they are not affected by gene silencing. This is time-consuming and very expensive. There is therefore a long-felt but unfulfilled need for novel methods allowing one to effectively and predictably control gene silencing in plant cells in order to obtain plants with improved properties in a cost-effective manner.

There is also a need in the field of functional genomics to provide cells or plants having no or insignificant levels of gene silencing so that analysis of gene functions can be performed more efficiently. By inhibiting or removing expression of genes responsible for gene silencing, the expression of genes of interest in functional genomics may be analyzed without the interference of gene silencing.

There is further a need in the field for increased gene silencing in cells or plants for more stringent control of gene expression or resistance to pathogens, in particular, viral pathogens.

The present invention addresses the need for methods to reproducibly and predictably manipulate gene expression in a plant cell. In particular, the present invention addresses the need for stable and predictable expression of a nucleotide sequence in a plant cell. According to the present invention, this is achieved by manipulating the expression in a plant cell of a nucleotide sequence encoding a polypeptide comprising a 3'-5' exonuclease domain. The present invention therefore provides a clear advantage over the prior art by reducing the number of transgenic lines which have to be screened until a suitable line is

selected, and by providing stable and better controlled expression of a nucleotide sequence in the plant cell.

In one aspect, the present invention encompasses novel methods for controlling gene silencing in a plant cell. The present invention encompasses the suppression of gene silencing or the increase in gene silencing in plants. In a preferred embodiment, this is achieved by altering the expression in the plant cell of a nucleotide sequence encoding a polypeptide comprising a 3'-5' exonuclease domain. In another embodiment, the nucleotide molecule encodes a polypeptide comprising exonuclease activity, preferably having 3'-5' RNA exonuclease activity. Preferably, the polypeptide comprises a 3'-5' exonuclease domain. More preferably, the 3'-5' exonuclease domain is an RNase D related domain. In another preferred embodiment, the polypeptide is identical or substantially similar to SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:22, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:24, SEQ ID NO:36 or SEQ ID NO:38. Preferably, the nucleotide sequence is identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37. Most preferably, the nucleotide sequence is identical or substantially similar to SEQ ID NO:23. In another embodiment, the invention provides novel isolated and substantially purified polypeptides comprising, consisting of or having an amino acid sequence identical or substantially similar to SEQ ID NO:24.

In a preferred embodiment, the expression of a nucleotide sequence encoding a polypeptide comprising a 3'-5' exonuclease domain is altered by altering its transcription or translation. Reduced expression is for example obtained by expressing in the plant cell a nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, or SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37 in sense orientation, or a portion thereof; or expressing in the plant cell a nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37 in anti-sense orientation, or a portion thereof; or expressing in the plant cell a sense RNA of a nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37 or a portion thereof, and an anti-sense RNA of said nucleotide sequence identical or

substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37 or a portion thereof, wherein said sense and said anti-sense RNAs are capable of forming a double-stranded RNA molecule; or expressing in said plant cell a ribozyme capable of specifically cleaving a messenger RNA transcript encoded by a nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:23; SEQ ID NO:35 or SEQ ID NO:37 or modifying by homologous recombination in said plant cell at least one chromosomal copy of a nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37 or of a regulatory region thereof; or expressing in said plant cell a zinc finger protein that is capable of binding to a nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37 or to a regulatory region thereof; or introducing into said plant cell a chimeric oligonucleotide that is capable of modifying at least one chromosomal copy of a nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37 or a regulatory region thereof. Preferably, the expression of the sequence is altered by insertional mutagenesis, point mutation or deletion mutagenesis of the genomic copy of a nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37 or a regulatory region thereof. Alternatively, the sequence has a mutation due to rearrangement of the gene sequence.

Increased expression of a polypeptide comprising a 3'-5' exonuclease domain is also within the scope of the present invention and is, for example, obtained by over-expressing in the plant cell a nucleotide sequence of the present invention.

In a further aspect, the present invention encompasses methods to alter the expression of a nucleotide sequence of interest in a plant cell, and methods to stabilize the expression of a nucleotide sequence of interest in a plant cell. In a preferred embodiment, the nucleotide sequence of interest is a heterologous nucleotide sequence. In another preferred embodiment, the nucleotide sequence of interest is an endogenous nucleotide sequence of the plant cell. The expression of a nucleotide sequence of interest is preferably altered by

altering the expression in the plant cell of a nucleotide sequence encoding a polypeptide comprising a 3'-5' exonuclease domain as described above. The plant cell with altered expression of a nucleotide sequence encoding a polypeptide comprising a 3'-5' exonuclease domain also comprises the nucleotide sequence of interest, or a portion thereof, or a reverse complement of either. In a preferred embodiment, the nucleotide sequence of interest, or a portion thereof, or a reverse complement thereof is introduced into plant cell with altered expression of a nucleotide sequence encoding a polypeptide comprising a 3'-5' exonuclease domain.

In a preferred embodiment, the nucleotide sequence of interest is derived from a pathogen of a plant, or is substantially similar thereto. A pathogen is, for example but not limited to, a viral, fungal or bacterial pathogen of a plant. Preferably the pathogen is a viral pathogen. Therefore, it is a further aspect of the present invention to provide for methods to control a pathogen comprising the steps of obtaining a plant cell with altered expression of a nucleotide sequence that encodes a polypeptide comprising a 3'-5' exonuclease domain as described above and wherein the plant cell further comprises a nucleotide sequence identical or substantially similar to a nucleotide sequence derived from the pathogen.

The present invention also encompasses a recombinant nucleic acid molecule comprising a nucleotide sequence that encodes a polypeptide comprising a 3'-5' exonuclease domain as described above, or a reverse complement thereof, or complement thereof.

The present invention also encompasses an expression cassette comprising a nucleic acid molecule of the present invention comprising a nucleotide sequence encoding a polypeptide comprising a 3'-5' exonuclease domain, or complement thereof. Preferably, the expression cassette comprises a nucleic acid molecule comprising a nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, or SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37.

The present invention also relates to a vector comprising the nucleic acid molecules of the present invention encoding a polypeptide comprising a 3'-5' exonuclease domain and/or activity. Preferably, the vector further comprises a promoter operably linked to the nucleic acid molecule of the present invention. More preferably, the vector further comprises a promoter and terminator, each operably linked to the nucleic acid molecule of the invention. Further, the present invention encompasses a plant cell (or a plant comprising such a plant cell) comprising a nucleic acid, recombinant nucleic acid molecule, an expression cassette or vector of the present invention encoding a polypeptide comprising a 3'-5' exonuclease

domain. The invention also provides progeny of the plant cells or plants described above, seeds, and parts of such a plant of the present invention, and the progeny thereof.

In yet a further aspect, the present invention also provides for methods to identify a compound that is capable of interacting with a polypeptide comprising a 3'-5' exonuclease domain as described above. Preferably, the compound is capable of altering the activity of said polypeptide. The compound can alter the activity of the polypeptide by increasing or decreasing the polypeptide exonuclease or gene silencing activity. In a preferred embodiment, such compound is a nucleic acid molecule, such as an aptamer, or a small-molecule ligand. In another preferred embodiment, such compound is applied to a plant or a plant cell, and such application results in the alteration of the activity of a polypeptide comprising a 3'-5' exonuclease domain in the plant or plant cell. Application of such a compound results in a more stable and predictable expression of a nucleotide sequence of interest in a plant cell or plant.

Thus, through an alteration of the expression of a nucleic acid molecule of the invention, the *stable and predictable expression of a nucleic acid molecule of interest in a plant cell*, the present invention provides a great advantage over current methods for the manipulation of gene expression in plant cells and plants. Current methods of transformation require extensive screening and testing of a large number of plants to identify a plant that stably and predictably expresses a nucleotide sequence of interest. Suppressing or decreasing expression of the nucleic acid molecule of the present invention results in decreased levels of post transcriptional gene silencing and improved expression of genes of interest. Therefore, the present invention allows for the production of improved plants, particularly improved commercial varieties, in a more timely and cost-effective manner.

The present invention thus provides:

An isolated nucleic acid molecule comprising a nucleotide sequence encoding a polypeptide comprising a 3'-5' exonuclease domain, and wherein the polypeptide is identical or substantially similar to an amino acid sequence of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:22, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:24, SEQ ID NO:36 or SEQ ID NO:38 or complements thereof. Preferably, the polypeptide is identical or substantially similar to SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6 or SEQ ID NO:22. More preferably, the polypeptide is identical or substantially similar to SEQ ID NO:2 or SEQ ID NO:24. In another preferred embodiment, the nucleotide sequence is identical or substantially similar to a nucleotide sequence of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11,

SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37. Preferably, the nucleotide sequence is identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, or SEQ ID NO:23. More preferably, the nucleotide sequence is substantially similar to SEQ ID NO:1. Most preferably, the nucleotide sequence is identical or substantially similar to SEQ ID NO:23. In another preferred embodiment, the 3'-5' exonuclease domain preferably comprises an RNase D related domain. Preferably, the polypeptide comprises 3'-5' exonuclease activity, and most preferably, 3'-5' RNA exonuclease activity. In yet another preferred embodiment, the nucleotide sequence is derived from a plant.

The present invention further provides an isolated recombinant nucleic acid molecule comprising a nucleotide sequence encoding a polypeptide encoded by the amino acid sequence identical or substantially similar to SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:22, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:24, SEQ ID NO:36 or SEQ ID NO:38, or complements thereof. More preferably, the recombinant nucleic acid molecules comprise the nucleotide sequence of SEQ ID NO:23 or complement thereof. The recombinant nucleic acid molecule is operatively linked to a promoter functional in a cell. Preferably, the promoter is functional in a plant cell.

Preferably, the nucleotide sequence of the present invention is in sense orientation in the nucleic acid molecule or in anti-sense orientation in the recombinant nucleic acid molecule. In yet another preferred embodiment, the polypeptide does not encode or comprise a helicase domain.

The present invention further provides:

An isolated and substantially purified polypeptide comprising an amino acid sequence identical or substantially similar to SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:22, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:24, SEQ ID NO:36, or SEQ ID NO:38. Preferably, the polypeptide comprises the amino acid sequence of SEQ ID NO:24. Alternatively, the polypeptide consists of the amino acid sequence of SEQ ID NO:24.

The present invention further provides:

An expression cassette comprising a nucleic acid or DNA molecule of the present invention. Preferably, the expression cassette further comprises a promoter and terminator. More preferably, the promoter is a constitutive promoter, an inducible promoter, a tissue-specific promoter or a developmentally-regulated promoter.

A vector comprising the nucleic acid molecules of the present invention.

A cell comprising the nucleic acid or recombinant nucleic acid molecule of the present invention, and a cell comprising the expression cassette of the present invention. Preferably, the cell is a plant cell. In a preferred embodiment, the nucleotide sequence of the present invention is expressed in said plant cell. In another preferred embodiment, the expression cassette promoter is a constitutive promoter, an inducible promoter, a tissue-specific promoter or a developmentally-regulated promoter. In another preferred embodiment, the expression cassette or recombinant nucleic acid molecule is stably integrated in the genome of the plant cell. In yet another preferred embodiment, the plant cell comprises an endogenous nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, or SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37. Preferably, the endogenous nucleotide sequence is identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, or SEQ ID NO:23. More preferably, the endogenous nucleotide sequence is identical or substantially similar to SEQ ID NO:1. Most preferably, the nucleotide sequence is identical or substantially similar to SEQ ID NO:23. Preferably, the expression of said endogenous nucleotide sequence in said plant cell is altered.

In a further preferred embodiment, the plant cell or plant comprises a nucleic acid molecule, or recombinant nucleic acid molecule, or expression cassette or vector of the present invention and further comprises a nucleic acid molecule comprising a nucleotide sequence of interest, wherein the expression of said nucleotide sequence of interest in said plant cell is altered as compared to the expression of said nucleotide sequence of interest in a plant cell lacking said nucleic acid molecule of the present invention. In another embodiment, the nucleotide sequence of interest is operably linked to a promoter. In yet another embodiment, the nucleotide sequence of interest is in an expression cassette.

The invention further provides a plant comprising the plant cell, and progeny and seeds from the plant comprising a nucleic acid sequence of the present invention.

The present invention further provides:

A plant cell comprising an endogenous nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37 and wherein said plant cell comprises a mutation in said endogenous nucleotide sequence, or in a regulatory region thereof. Preferably, said mutation is due to the insertion of a nucleic

acid molecule into said endogenous nucleotide sequence or into a regulatory region thereof, wherein the expression of said endogenous nucleotide sequence in said plant is altered. Preferably, the endogenous nucleotide sequence is identical or substantially similar to SEQ ID NO:1. Most preferably, the nucleotide sequence is as described or substantially similar to SEQ ID NO:23. Preferably, the insertion of a nucleic acid molecule comprises one T-DNA border region or a transposable element. An advantage of the invention is that the expression of said endogenous nucleotide sequence in said plant cell is reduced. In another preferred embodiment, the mutation is due to a deletion or rearrangement. In yet another embodiment, the mutation is due to a point mutation.

Preferably, the plant cell further comprises an expression cassette comprising a nucleotide sequence of interest, wherein the expression of said nucleotide sequence of interest in said plant cell is stabilized or increased as compared to the expression of said nucleotide sequence of interest in a plant cell lacking said nucleic acid molecule of the present invention. In another preferred embodiment, the expression of said endogenous nucleotide sequence described above in said plant cell is increased.

In yet another preferred embodiment, the plant cell further comprises an expression cassette comprising a nucleotide sequence of interest, wherein the expression of said nucleotide sequence of interest in said plant cell is decreased as compared to the expression of said nucleotide sequence of interest in a plant cell lacking said nucleic acid molecule of the present invention.

Further provided is a plant comprising the plant cell comprising the above-described nucleic acid molecules or expression cassettes, or recombinant nucleic acid molecules.

The present invention further provides:

A plant cell or plant transgenically expressing a sense RNA molecule of a nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37 and an anti-sense RNA molecule of said nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37 wherein said sense and said anti-sense RNA molecules are capable of forming a double-stranded RNA molecule. An advantage of the invention is that the expression in said plant cell of an endogenous nucleotide sequence of said plant cell that is substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, SEQ ID

NO:9, SEQ ID NO:11, SEQ ID NO:13, or SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37 is reduced.

In another preferred embodiment, the plant cell further comprises an expression cassette comprising a second nucleotide sequence, wherein the expression of said second nucleotide sequence in said plant cell is stabilized or increased as compared to the expression of said second nucleotide sequence in a plant cell that is not expressing said sense and said anti-sense RNA molecules.

A plant, seed or progeny thereof comprising the plant cell comprising the sense and antisense constructs as described above is also provided.

The present invention further provides:

A method for altering the expression of an endogenous nucleotide sequence that is identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37 in a plant cell or plant comprising the step of: altering the transcription or translation of said endogenous nucleotide sequence in said plant cell or plant.

In a preferred embodiment, wherein altering the transcription or translation of said endogenous nucleotide sequence in said plant cell or plant comprises the step of:

a) expressing in said plant cell a nucleotide sequence identical or substantially similar to

SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37 or a portion thereof, in sense orientation; or

b) expressing in said plant cell a nucleotide sequence identical or substantially similar to

SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37 or a portion thereof, in anti-sense orientation; or

c) expressing in said plant cell a sense RNA of a nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37, or a portion thereof, and an anti-sense RNA of said nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37, or a

portion thereof, wherein said sense and said anti-sense RNAs are capable of forming a double-stranded RNA molecule; or

d) expressing in said plant cell a ribozyme capable of specifically cleaving a messenger RNA transcript encoded by a nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37; or

e) modifying by homologous recombination in said plant cell at least one chromosomal copy of the nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37 or of a regulatory region thereof; or

f) expressing in said plant cell a zinc finger protein that is capable of binding to a nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37, or to a regulatory region thereof; or

g) introducing into said plant cell a chimeric oligonucleotide that is capable of modifying at least one chromosomal copy of the nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37 or a regulatory region thereof.

The present invention further provides:

A method for altering the expression of an endogenous nucleotide sequence that is as described or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37 in a plant cell comprising introducing into said plant cell a means for altering the transcription or translation of said endogenous nucleotide sequence in said plant cell.

The present invention further provides:

A method for altering the expression of a nucleotide sequence of interest in a plant cell or plant comprising the steps of:

a) altering the expression in said plant cell or plant of an endogenous nucleotide sequence of said plant cell that is identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37; and

b) introducing into said plant cell a nucleic acid molecule comprising said nucleotide sequence of interest, wherein the expression of said nucleotide sequence of interest in said plant cell or plant is altered.

In a preferred embodiment, said step a) comprises:

a) expressing in said plant cell a nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37, or a portion thereof, in sense orientation; or

b) expressing in said plant cell a nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37, or a portion thereof, in anti-sense orientation; or

c) expressing in said plant cell a sense RNA of a nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37, or a portion thereof, and an anti-sense RNA of said nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37, or a portion thereof, wherein said sense and said anti-sense RNAs are capable of forming a double-stranded RNA molecule; or

d) expressing in said plant cell a ribozyme capable of specifically cleaving a messenger RNA transcript encoded by a nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37 or

e) modifying by homologous recombination in said plant cell at least one chromosomal copy of the nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37, or of a regulatory region thereof; or

f) expressing in said plant cell a zinc finger protein that is capable of binding to a nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37, or to a regulatory region thereof; or

g) introducing into said plant cell a chimeric oligonucleotide that is capable of modifying at least one chromosomal copy of the nucleotide identical or sequence substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37 or a regulatory region thereof.

The present invention further provides:

A method for altering, increasing or stabilizing the expression of a nucleotide sequence of interest in a plant cell comprising the steps of:

- a) obtaining a plant cell comprising an expression cassette of the present invention expressing the nucleotide sequence of the present invention; and
- b) introducing into said plant cell a nucleic acid molecule comprising said nucleotide sequence of interest, wherein the expression of said nucleotide sequence of interest in said plant cell is altered, increased or stabilized as compared to the expression of said nucleotide sequence of interest in a plant cell lacking said expression cassette.

Alternatively, the expression of said nucleotide sequence of interest in said plant cell is reduced or increased. Preferably, the nucleotide sequence of interest is identical or substantially similar to an endogenous nucleotide sequence of said plant cell.

The present invention further provides:

A method for stabilizing the expression of a nucleotide sequence of interest in a plant cell comprising:

- a) altering the expression in a plant cell of an endogenous nucleotide sequence of said plant cell that encodes a polypeptide comprising a 3'-5' exonuclease domain, and wherein said polypeptide is identical or substantially similar to an amino acid sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:22, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:24, SEQ ID NO:36 and SEQ ID NO:38; and
- b) introducing into said plant cell a nucleotide sequence of interest, wherein the expression of said nucleotide sequence of interest in said plant cell is stabilized.

Preferably, the polypeptide has 3'-5' RNA exonuclease activity. Preferably, the 3'-5' exonuclease domain is an RNase D related domain. Preferably, the endogenous nucleotide sequence is identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37.

Preferably, the expression of said endogenous nucleotide sequence is altered by:

a) expressing in said plant cell a nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, or SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37, or a portion thereof, in sense orientation; or

b) expressing in said plant cell a nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37, or a portion thereof, in anti-sense orientation; or

c) expressing in said plant cell a sense RNA of a nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37, or a portion thereof, and an anti-sense RNA of said nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37, or a portion thereof, wherein said sense and said anti-sense RNAs are capable of forming a double-stranded RNA molecule; or

d) expressing in said plant cell a ribozyme capable of specifically cleaving a messenger RNA transcript encoded by a nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37; or

e) expressing in said plant cell an aptamer specifically directed to a polypeptide identical or substantially similar to SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:22, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:24, SEQ ID NO:36 or SEQ ID NO:38; or

f) modifying by homologous recombination in said plant cell at least one chromosomal copy of the nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37, or of a regulatory region thereof; or

g) expressing in said plant cell a zinc finger protein that is capable of binding to a nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ

ID NO:5, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37, or to a regulatory region thereof; or

h) introducing into said plant cell a chimeric oligonucleotide that is capable of modifying at least one chromosomal copy of the nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, or SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37, or a regulatory region thereof.

Preferably, the expression in a plant cell of said endogenous nucleotide sequence that encodes a polypeptide comprising a 3'-5' exonuclease domain is reduced.

The present invention further provides:

A method for identifying a compound capable of interacting with a polypeptide comprising a 3'-5' exonuclease domain comprising:

a) combining a polypeptide comprising the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:22, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:24, SEQ ID NO:36 or SEQ ID NO:38, or a homolog thereof, and a compound to be tested for the ability to interact with said polypeptide, under conditions conducive to interaction; and

b) selecting a compound from step (a) that is capable of interacting with said polypeptide.

Preferably, the polypeptide is encoded by a nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37.

The present invention further provides:

A compound identifiable by the method disclosed immediately above. Preferably, the compound is capable of altering the activity of said polypeptide. More preferably, the compound is capable of decreasing or increasing gene silencing activity of the polypeptide.

DESCRIPTION OF THE SEQUENCE LISTING

SEQ ID NO:1	nucleotide sequence corresponding to GenPept accession CAB36851
SEQ ID NO:2	GenPept accession CAB36851
SEQ ID NO:3	nucleotide sequence corresponding to GenPept accession AAD25623
SEQ ID NO:4	GenPept accession AAD25623
SEQ ID NO:5	nucleotide sequence corresponding to GenPept accession AAC69936

SEQ ID NO:6	GenPept accession AAC69936
SEQ ID NO:7	nucleotide sequence corresponding to GenPept accession AAC42241
SEQ ID NO:8	GenPept accession AAC42241
SEQ ID NO:9	nucleotide sequence corresponding to GenPept accession AAD26968
SEQ ID NO:10	GenPept accession AAD26968
SEQ ID NO:11	nucleotide sequence corresponding to GenPept accession AAC25931
SEQ ID NO:12	GenPept accession AAC25931
SEQ ID NO:13	nucleotide sequence corresponding to GenPept accession AAF98185
SEQ ID NO:14	GenPept accession AAF98185
SEQ ID NO:15	nucleotide sequence corresponding to GenPept accession CAA80137
SEQ ID NO:16	GenPept accession CAA80137
SEQ ID NO:17	nucleotide sequence corresponding to GenPept accession AAF06162
SEQ ID NO:18	GenPept accession AAF06162
SEQ ID NO:19	Oligonucleotide 3' specific primer
SEQ ID NO:20	Oligonucleotide pD991 primer
SEQ ID NO:21	corrected nucleotide sequence corresponding to corrected GenPept accession AAC42241
SEQ ID NO:22	corrected GenPept accession AAC42241
SEQ ID NO:23	nucleotide sequence of cDNA encoding a polypeptide comprising a RNase D related domain from <i>Arabidopsis thaliana</i>
SEQ ID NO:24	amino acid sequence of polypeptide comprising a RNase D related domain from <i>Arabidopsis thaliana</i>
SEQ ID NO:25	oligonucleotide T-DNA specific primer LB1
SEQ ID NO:26	oligonucleotide T-DNA specific primer LB2
SEQ ID NO:27	oligonucleotide T-DNA specific primer LB3
SEQ ID NO:28	oligonucleotide arbitrary degenerate primer AD3
SEQ ID NO:29	oligonucleotide primer 36851TD#3
SEQ ID NO:30	gene-specific oligonucleotide primer L22F4F
SEQ ID NO:31	gene-specific oligonucleotide primer F22L4R
SEQ ID NO:32	oligonucleotide primer AtWRN CDS F
SEQ ID NO:33	oligonucleotide primer AtWRN-RT-R
SEQ ID NO:34	oligonucleotide primer AtWRN CDS R
SEQ ID NO:35	nucleotide sequence corresponding to GenPept accession AAG50917
SEQ ID NO:36	GenPept accession AAG50917

SEQ ID NO:37 nucleotide sequence corresponding to GenPept accession BAB11227

SEQ ID NO:38 GenPept accession BAB11227

DEFINITIONS

For clarity, certain terms used in the specification are defined and used as follows:

Alter: to “alter” the expression of a nucleotide sequence in a plant cell means that the level of expression of the nucleotide sequence in a plant cell after applying a method of the present invention is different from its expression in the cell before applying the method. In a preferred embodiment, to alter expression means that the expression of the nucleotide sequence in the plant is reduced after applying a method of the present invention as compared to before applying the method. The term “Reduced” means herein lower, preferably significantly lower, more preferably the expression of the nucleotide sequence is not detectable. In another preferred embodiment, to alter expression means that the expression of the nucleotide sequence in the plant is increased after applying a method of the present invention as compared to before applying the method.

Antiparallel: “Antiparallel” refers herein to two nucleotide sequences paired through hydrogen bonds between complementary base residues with phosphodiester bonds running in the 5'-3' direction in one nucleotide sequence and in the 3'-5' direction in the other nucleotide sequence.

Complementary: “Complementary” refers to two nucleotide sequences which comprise antiparallel nucleotide sequences capable of pairing with one another upon formation of hydrogen bonds between the complementary base residues in the antiparallel nucleotide sequences.

DNA shuffling: “DNA shuffling” is a method to rapidly, easily and efficiently introduce mutations or rearrangements, preferably randomly, in a DNA molecule or to generate exchanges of DNA sequences between two or more DNA molecules, preferably randomly. The DNA molecule resulting from DNA shuffling is a shuffled DNA molecule that is a non-naturally occurring DNA molecule derived from at least one template DNA molecule. The shuffled DNA encodes an enzyme modified with respect to the enzyme encoded by the template DNA, and preferably has an altered biological activity with respect to the enzyme encoded by the template DNA.

Double-stranded RNA: A “double-stranded RNA (dsRNA)” molecule comprises a sense RNA fragment of a nucleotide sequence and an antisense RNA fragment of the nucleotide sequence, which both comprise nucleotide sequences complementary to one

another, thereby allowing the sense and antisense RNA fragments to pair and form a double-stranded RNA molecule.

Endogenous: An "endogenous" nucleotide sequence refers to a nucleotide sequence which is present in the genome of the untransformed plant cell.

Essential: An "essential" gene is a gene encoding a protein such as e.g. a biosynthetic enzyme, receptor, signal transduction protein, structural gene product, or transport protein that is essential to the growth or survival of the plant.

Expression: "Expression" refers to the transcription and/or translation of a nucleotide sequence, for example an endogenous gene or a heterologous gene, in a cell. In the case of antisense constructs, for example, expression may refer to the transcription of the antisense DNA only.

Expression cassette: "Expression cassette" as used herein means a DNA sequence capable of directing expression of a particular nucleotide sequence in an appropriate host cell, comprising a promoter functional in the plant cell into which it will be introduced, operatively linked to the nucleotide sequence of interest which is operatively linked to termination signals. It also typically comprises sequences required for proper translation of the nucleotide sequence. The coding region usually codes for a protein of interest but may also code for a functional RNA of interest, for example antisense RNA or a nontranslated RNA, in the sense or antisense direction. The expression cassette comprising the nucleotide sequence of interest may be chimeric, meaning that at least one of its components is heterologous with respect to at least one of its other components. The expression cassette may also be one which is naturally occurring but has been obtained in a recombinant form useful for heterologous expression. Typically, however, the expression cassette is heterologous with respect to the host, i.e., the particular DNA sequence of the expression cassette does not occur naturally in the host cell and must have been introduced into the host cell or an ancestor of the host cell by a transformation event. The expression of the nucleotide sequence in the expression cassette may be under the control of a constitutive promoter or of an inducible promoter which initiates transcription only when the host cell is exposed to some particular external stimulus. In the case of a multicellular organism, such as a plant, the promoter can also be specific to a particular tissue or organ or stage of development.

Heterologous DNA Sequence: a DNA sequence not naturally associated with a host cell into which it is introduced, including non-naturally occurring multiple copies of an endogenous DNA sequence.

Homologous DNA Sequence: a DNA sequence naturally associated with a host cell.

Isogenic: plants which are genetically identical, except that they may differ by the presence or absence of a heterologous DNA sequence.

Isolated: in the context of the present invention, an isolated DNA molecule or an isolated enzyme is a DNA molecule or enzyme which, by the hand of man, exists apart from its native environment and is therefore not a product of nature. An isolated DNA molecule or enzyme may exist in a purified form or may exist in a non-native environment such as, for example, in a transgenic host cell.

Mature protein: protein which is normally targeted to a cellular organelle, such as a chloroplast, and from which the transit peptide has been removed.

Minimal Promoter: promoter elements, particularly a TATA element, that are inactive or that have greatly reduced promoter activity in the absence of upstream activation. In the presence of a suitable transcription factor, the minimal promoter functions to permit transcription.

"Nucleic Acids" and "Nucleotides" refer to naturally occurring or synthetic or artificial nucleic acid or nucleotides.

Plant: A "plant" refers to any plant or part of a plant at any stage of development. Therein are also included cuttings, cell or tissue cultures and seeds. As used in conjunction with the present invention, the term "plant tissue" includes, but is not limited to, whole plants, plant cells, plant organs, plant seeds, protoplasts, callus, cell cultures, and any groups of plant cells organized into structural and/or functional units.

Pre-protein: protein which is normally targeted to a cellular organelle, such as a chloroplast, and still comprising its transit peptide.

Significant Increase or Decrease: an increase or decrease, for example in enzymatic activity or in gene expression, that is larger than the margin of error inherent in the measurement technique, preferably an increase or decrease by about 2-fold or greater of the activity of the control enzyme or expression in the control cell, more preferably an increase or decrease by about 5-fold or greater, and most preferably an increase or decrease by about 10-fold or greater.

Stabilize: to "stabilize" the expression of a nucleotide sequence in a plant cell means that the level of expression of the nucleotide sequence after applying a method of the present invention is approximately the same in cells from the same tissue in different plants from the same generation or throughout multiple generations when the plants are grown under the same or comparable conditions.

In its broadest sense, the term "substantially similar", when used herein with respect to a nucleotide sequence, means a nucleotide sequence corresponding to a reference nucleotide sequence, wherein the corresponding sequence encodes a polypeptide having substantially the same structure and function as the polypeptide encoded by the reference nucleotide sequence, e.g. where only changes in amino acids not affecting the polypeptide function occur. Desirably, the substantially similar nucleotide sequence encodes the polypeptide encoded by the reference nucleotide sequence. The term "substantially similar" is specifically intended to include nucleotide sequences wherein the sequence has been modified to optimize expression in particular cells. The percentage of identity between the amino acid sequence encoded by the substantially similar nucleotide sequence and the reference nucleotide sequence is desirably at least 24%, more desirably at least 30%, more desirably at least 45%, preferably at least 60%, more preferably at least 75%, still more preferably at least 90%, yet still more preferably at least 95%, yet still more preferably at least 99%. Sequence comparisons are carried out using default GAP analysis with the University of Wisconsin GCG, SEQWEB application of GAP, based on the algorithm of Needleman and Wunsch (Needleman and Wunsch (1970) J Mol. Biol. 48: 443-453). A nucleotide sequence "substantially similar" to reference nucleotide sequence hybridizes to the reference nucleotide sequence in 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO₄, 1 mM EDTA at 50°C with washing in 2X SSC, 0.1% SDS at 50°C, more desirably in 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO₄, 1 mM EDTA at 50°C with washing in 1X SSC, 0.1% SDS at 50°C, more desirably still in 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO₄, 1 mM EDTA at 50°C with washing in 0.5X SSC, 0.1% SDS at 50°C, preferably in 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO₄, 1 mM EDTA at 50°C with washing in 0.1X SSC, 0.1% SDS at 50°C, more preferably in 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO₄, 1 mM EDTA at 50°C with washing in 0.1X SSC, 0.1% SDS at 65°C. Homologs of the nucleotide sequence include nucleotide sequences that encode an amino acid sequence that is at least 24% identical, more preferably at least 35% identical, yet more preferably at least 50% identical, yet more preferably at least 65% identical to the reference amino acid sequence, as measured using the parameters described above, wherein the amino acid sequence encoded by the homolog has the biological activity of a 3'-5' exonuclease. More preferably, the homolog has the biological activity of a 3'-5' RNA exonuclease. In another preferred embodiment, a homolog of the nucleotide sequence encodes an amino acid sequence that comprises a 3'-5' exonuclease domain.

The term "substantially similar", when used herein with respect to a polypeptide, means a protein corresponding to a reference polypeptide, wherein the polypeptide has substantially the same structure and function as the reference protein, e.g. where only changes in amino acids sequence not affecting the polypeptide function occur. When used for a polypeptide or an amino acid sequence the percentage of identity between the substantially similar and the reference polypeptide or amino acid sequence desirably is at least 24%, more desirably at least 30%, more desirably at least 45%, preferably at least 60%, more preferably at least 75%, still more preferably at least 90%, yet still more preferably at least 95%, yet still more preferably at least 99%, using default GAP analysis parameters as described above. Homologs are amino acid sequences that are at least 24% identical, more preferably at least 35% identical, yet more preferably at least 50% identical, yet more preferably at least 65% identical to the reference polypeptide or amino acid sequence, as measured using the parameters described above, wherein the amino acid sequence encoded by the homolog has the biological activity of a 3'-5' exonuclease. More preferably, the homolog has the biological activity of a 3'-5' RNA exonuclease. In another preferred embodiment, a homolog of the nucleotide sequence encodes an amino acid sequence that comprises a 3'-5' exonuclease domain.

Target gene: A "target gene" is any gene in a plant cell. For example, a target gene is a gene of known function or is a gene whose function is unknown, but whose total or partial nucleotide sequence is known. Alternatively, the function of a target gene and its nucleotide sequence are both unknown. A target gene is a native gene of the plant cell or is a heterologous gene which has previously been introduced into the plant cell or a parent cell of said plant cell, for example by genetic transformation. A heterologous target gene is stably integrated in the genome of the plant cell or is present in the plant cell as an extrachromosomal molecule, e.g. as an autonomously replicating extrachromosomal molecule.

Transformation: a process for introducing heterologous nucleic acid molecule into a cell, tissue, or plant. Transformed cells, tissues, or plants are understood to encompass not only the end product of a transformation process, but also transgenic progeny thereof.

Transgenic: transformed, preferably stably transformed, with a recombinant DNA molecule that preferably comprises a suitable promoter operatively linked to a DNA sequence of interest.

The ability to reproducibly and predictably manipulate gene expression in plants is an important consideration for the production of novel commercial varieties with improved properties. New traits are often introduced into plant cells by transgenic methods but their expression is sometimes subject to variations between individual plants or between different generations. This phenomenon is referred to as gene silencing and the selection of lines not affected by gene silencing requires substantial efforts and is expensive. In other applications, it is desired to reduce or eliminate the expression of a particular endogenous gene in a plant cell, but with current methods it is often difficult to achieve this routinely in a stable and reproducible manner. Therefore, it is an object of the present invention to provide novel methods that address these needs and allow stabilizing or altering the expression of a nucleotide sequence of interest in a plant cell in a predictable and stable manner. According to the present invention, this is preferably achieved by altering the expression in a plant cell of a nucleotide sequence encoding a polypeptide having 3'-5' exonuclease domain.

I. Nucleotide sequences encoding a polypeptide comprising a 3'-5' exonuclease domain
In one aspect, the present invention provides for nucleic acid molecules having nucleotide sequence encoding polypeptides comprising a 3'-5' exonuclease domain. Preferably, the 3'-5' exonuclease domain is a RNase D related domain. The present invention also provides nucleic acid molecules comprising a nucleotide sequence encoding a polypeptide comprising 3'-5' exonuclease activity. Preferably, the polypeptide has 3'-5' RNA exonuclease activity. In yet another preferred embodiment, the nucleotide sequence is isolated from a plant, preferably from a monocotyledonous plant or a dicotyledonous plant. Preferably, the plants are, but not limited to, corn, rice, wheat, soybean, cotton, sunflower, *Brassica* spp., canola, tomato, potato, *Solanaceae* spp. or sugar beets. More preferably, the nucleic acid molecules are isolated from *Arabidopsis thaliana*.

A 3'-5' exonuclease domain typically comprises three subdomains designated as exo I, exo II and exo III (Moser et al. (1997) Nucl. Acids Res. 25:5110-5118, incorporated herein by reference in its entirety). These motifs are clustered around the active site and contain four negatively charged residues that serve as ligands for the two metal ions necessary for catalysis in addition to a catalytically active tyrosine. Typically, a 3'-5' exonuclease domain is approximately 140 amino acids long. 3'-5' exonuclease domains are for example found in DNA polymerases where they are sometimes referred to as the 3'-5' exodeoxyribonuclease (or proofreading) domains.

3'-5' exonuclease domains are also found in the RNase D family of polypeptides, that includes for example the *E. coli* ribonuclease (RNase D), the *S. cerevisiae* Rrp6p protein and the human Werner syndrome protein (see Mian (1997) Nucleic Acids Research 25:3187-3195, incorporated herein by reference in its entirety). Such domains are referred to as RNase D related domains. An alignment of polypeptides comprising an RNase D related domain is shown in Mian (1997). RNase D related domains and proofreading domains appear to be similar.

The inventors of the present invention are the first to screen for plant nucleotide sequences encoding a polypeptide comprising a 3'-5' exonuclease domain, and to successfully identify such nucleotide sequences. This is carried out according to the methods disclosed in Example 1. The amino acid sequences and corresponding nucleotide sequences identified using the method and algorithms disclosed in Example 1 are set forth in SEQ ID NO:1-14, and 35-38 and briefly described as follows. An amino acid sequence predicted from a genomic sequence from *Arabidopsis thaliana* is found in GenBank under accession #CAB36851 and is set forth in SEQ ID NO:2. The corresponding nucleotide sequence is found in BAC F18A5, GenBank accession number AL035528.2. An amino acid sequence predicted from a genomic sequence from *Arabidopsis thaliana* is found in GenBank under accession #AAD25623 and is set forth in SEQ ID NO:4. The corresponding nucleotide sequence is found in BAC F20D21, GenBank accession number AC005287.4. An amino acid sequence predicted from a genomic sequence from *Arabidopsis thaliana* is found in GenBank under accession #AAC69936 and is set forth in SEQ ID NO:6. The corresponding nucleotide sequence is found in *Arabidopsis thaliana* chromosome II section 181 of 255, GenBank accession number AC005700.2. An amino acid sequence predicted from a genomic sequence from *Arabidopsis thaliana* is found in GenBank under accession #AAC42241 and is set forth in SEQ ID NO:8. The corresponding nucleotide sequence is found in *Arabidopsis thaliana* chromosome II section 145 of 255, GenBank accession number AC005395.2. An amino acid sequence predicted from a genomic sequence from *Arabidopsis thaliana* is found in GenBank under accession #AAD26968 and is set forth in SEQ ID NO:10. The corresponding nucleotide sequence is found in *Arabidopsis thaliana* chromosome II section 197 of 255, GenBank accession number AC007135.7. An amino acid sequence predicted from a genomic sequence from *Arabidopsis thaliana* is found in GenBank under accession #AAC25931 and is set forth in SEQ ID NO:12. The corresponding nucleotide sequence is found in *Arabidopsis thaliana* chromosome II section 182 of 255, GenBank accession number AC004681.2. An amino acid sequence predicted

from a genomic sequence from *Arabidopsis thaliana* is found in GenBank under accession #AAF98185 and is set forth in SEQ ID NO:14. The corresponding nucleotide sequence is found in BAC F17F8, GenBank accession number AC000107.2. An amino acid sequence predicted from a genomic sequence from *Arabidopsis thaliana* is found in GenBank under accession # AAG50917 and is set forth in SEQ ID NO:36. The corresponding nucleotide sequence is found in BAC F14G9, GenBank accession number AC069159. An amino acid sequence predicted from a genomic sequence from *Arabidopsis thaliana* is found in GenBank under accession # BAB11227 and is set forth in SEQ ID NO:38. The corresponding nucleotide sequence is found in BAC K16H17, GenBank accession number AB016884.

The inventors of the present invention also discovered that the 5' end of GenPept accession AAC42241 is missing due to incorrect annotation, and that GenPept accession AAC42241 lacks the exo I motif of the 3'-5' exonuclease domain. The amino acid sequence comprising the entire 3'-5' exonuclease domain (including exo I) is disclosed for the first time in the instant application and is set forth in SEQ ID NO:22. The corresponding nucleotide sequence is set forth in SEQ ID NO:21.

Further, the present invention provides for nucleic acid molecules encoding a full length nucleotide sequence encoding a polypeptide comprising a 3'-5' exonuclease domain of SEQ ID NO:24 as was cloned from *Arabidopsis thaliana* as set forth in Examples 2-3. The invention also provides a nucleic acid molecule comprising or having the sequence identical or substantially similar to the nucleotide sequence of SEQ ID NO:23 or complements thereof. The inventors of the present invention predicted a 3'-5' exonuclease domain between about amino acid positions 129 and 287 in the amino acid sequence set forth in SEQ ID NO:2. The inventors of the present invention also predicted that the amino acid sequence between about amino acid positions 136 and 271 in SEQ ID NO:4 is comprised in a 3'-5' exonuclease domain, that the amino acid sequence between about amino acid positions 76 and 210 in SEQ ID NO:6 is comprised in a 3'-5' exonuclease domain, that the amino acid sequence between about amino acid positions 46 and 199 in SEQ ID NO:22 is comprised in a 3'-5' exonuclease domain, that the amino acid sequence between about amino acid positions 57 and 193 in SEQ ID NO:10 is comprised in a 3'-5' exonuclease domain, that the amino acid sequence between about amino acid positions 66 and 202 in SEQ ID NO:12 is comprised in a 3'-5' exonuclease domain, that the amino acid sequence between about amino acid positions 380 and 538 in SEQ ID NO:36 is comprised in a 3'-5' exonuclease domain, that the amino acid sequence between about amino acid positions 30

and 193 in SEQ ID NO:38 is comprised in a 3'-5' exonuclease domain. The inventors of the present invention also predict that the amino acid sequence between about amino acid positions 129 and 282 in SEQ ID NO:24 comprises a 3'-5' exonuclease domain.

Preferably, the nucleotide sequence of the present invention encode a polypeptide comprising a 3'-5' exonuclease domain. In another aspect of the invention, the nucleotide sequence encodes a polypeptide comprising at least one 3'-5' exonuclease domain. In yet another embodiment, the nucleotide sequence encodes a polypeptide comprising more than one 3'-5' exonuclease domain.

Thus, the present invention discloses a nucleotide sequence encoding a polypeptide identical or substantially similar to SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:22, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:36 or SEQ ID NO:38. Preferably, the polypeptide is identical or substantially similar to SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, or SEQ ID NO:22. More preferably, the polypeptide is identical or substantially similar to SEQ ID NO:2. Most preferably, the polypeptide is identical or substantially similar to the amino acid sequence of SEQ ID NO:24.

Preferably, the nucleotide sequence is identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11 SEQ ID NO:13, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37. More preferably, the nucleotide sequence is substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, or SEQ ID NO:21. Yet more preferably, the nucleotide sequence is identical or substantially similar to SEQ ID NO:1. Most preferably, the nucleotide sequence is identical or substantially similar to SEQ ID NO:23.

The inventors of the present invention are also the first to predict and demonstrate that a nucleotide sequence of the present invention is involved in gene silencing, and to use such nucleotide sequences to alter or stabilize the expression of a nucleotide sequence of interest in a cell as set forth in Example 5. The nucleotide sequences of the present invention are useful to alter or stabilize the expression of another nucleotide sequence of interest in a plant cell.

Based on Applicants' disclosure of the present invention, nucleotide sequences encoding polypeptides identical or substantially similar to SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:24, SEQ ID NO:36 or SEQ ID NO:38 are isolated, preferably from the genome of any desired plant. For example, all or part of the nucleotide sequence set forth in SEQ ID NO:1 is used as a probe that selectively hybridizes to other nucleotide sequences present in a population of cloned

genomic DNA fragments or cDNA fragments (i.e. genomic or cDNA libraries) from a chosen source organism. Such techniques include hybridization screening of plated DNA libraries (either plaques or colonies; see, e.g. Sambrook *et al.*, "Molecular Cloning", eds., Cold Spring Harbor Laboratory Press. (1989)) and amplification by PCR using oligonucleotide primers corresponding to sequence domains conserved among such polypeptides (see, e.g. Innis *et al.*, "PCR Protocols, a Guide to Methods and Applications", Academic Press (1990)). For example, oligonucleotide primers corresponding to a portion of a 3'-5' exonuclease domain are used. These methods are particularly well suited to the isolation of nucleotide sequences from organisms closely related to the organism from which the probe sequence is derived. Isolation of such a nucleic acid molecule of the present invention, in particular SEQ ID NO:23, is described in Example 7.

The isolated nucleotide sequences taught by the present invention are manipulated according to standard genetic engineering techniques to suit any desired purpose. For example, they may be used as a probe capable of specifically hybridizing to coding sequences and messenger RNAs. To achieve specific hybridization under a variety of conditions, such probes include preferably at least 10 nucleotides in length, preferably at least 20 nucleotides in length, and most preferably at least 50 nucleotides in length. Such probes are used to amplify and analyze nucleotide sequences from a chosen organism via PCR.

Specific hybridization probes also are used to map the location of these native genes in the genome of a chosen plant using standard techniques based on the selective hybridization of the probe to genomic sequences. These techniques include, but are not limited to, identification of DNA polymorphisms identified or contained within the probe sequence, and use of such polymorphisms to follow segregation of the gene relative to other markers of known map position in a mapping population derived from self fertilization of a hybrid of two polymorphic parental lines (see e.g. Helentjaris *et al.*, *Plant Mol. Biol.* 5: 109 (1985); Sommer *et al. BioTechniques* 12:82 (1992); D'Ovidio *et al.*, *Plant Mol. Biol.* 15: 169 (1990)). Mapping of genes in this manner is contemplated to be particularly useful for breeding purposes. For instance, by knowing the genetic map position of a mutant gene, flanking DNA markers are identified from a reference genetic map (see, e.g., Helentjaris, *Trends Genet.* 3: 217 (1987)). During introgression of the herbicide resistance trait into a new breeding line, these markers are used to monitor the extent of linked flanking chromosomal DNA still present in the recurrent parent after each round of back-crossing. Specific hybridization probes also are used to quantify levels of mRNA in a plant using standard

techniques such as Northern blot analysis.

In another aspect of the present invention, a nucleotide sequence encoding a polypeptide comprising a 3'-5' exonuclease domain is inserted in a recombinant nucleic acid molecule. The recombinant nucleic acid molecule is preferably operatively linked to a promoter. More preferably, the promoter is functional in a plant cell. Recombinant nucleic acid molecules can be introduced into plant cells by genetic transformation, as described for example in further detail infra.

The present invention also provides for expression cassettes comprising a promoter operably linked to a nucleic acid molecule encoding polypeptides comprising 3'-5' exonuclease domains described above and a terminator. The expression cassettes of the present invention may further comprise an enhancer.

In another aspect, the present invention provides vectors comprising the nucleic acid molecules encoding polypeptides comprising 3'-5' exonuclease domains described above. Also, the vectors further comprise a promoter and terminator operationally linked to the nucleic acid molecule of the present invention. Plasmid and viral vectors known to those skilled in the art of molecular biology further comprising the nucleic acid molecules of the present invention are encompassed by the invention.

II. Methods for altering the expression of a polypeptide having 3'-5' exonuclease domain in a cell

The inventors of the present invention are the first to discover that nucleotide sequences of the present invention are useful to manipulate or alter gene expression or post-transcriptional gene silencing (PTGS). Preferably, gene expression or PTGS is manipulated or altered in plant cells. Thus, one object of the present invention is to alter the expression in a plant cell of a nucleotide sequence of said plant said that encodes a polypeptide comprising a 3'-5' exonuclease domain and/or activity.

As described in Examples 5 and 6, decreasing or preventing expression of a nucleic acid molecule encoding a polypeptide comprising a 3'-5' exonuclease domain, causes a decrease or eliminates detectable levels of PTGS. The levels of PTGS are determined by measuring the levels of expression of a GFP reporter gene. Replacement of such a sequence of the present invention, restores PTGS activity in the plant.

Additionally, overexpression of a nucleic acid molecule of the present invention encoding a polypeptide comprising a 3'-5' exonuclease domain increases or supplements levels of PTGS.

The present invention provides a number of methods for altering the expression of a nucleic acid molecule encoding a polypeptide comprising a 3'-5' exonuclease domain. These methods allow for the decrease or increase in the level of expression of the nucleic acid molecule encoding polypeptide comprising a 3'-5' exonuclease domain which in turn, produces alteration of expression of nucleic acid molecules or genes of interest.

For example, the alteration in expression of the nucleic acid molecule of the present invention is achieved in one of the following ways:

(1) "Sense" Suppression

Alteration of the expression of a nucleotide sequence of the present invention, preferably reduction of its expression, is obtained by "sense" suppression (referenced in e.g. Jorgensen et al. (1996) Plant Mol. Biol. 31, 957-973). In this case, the entirety or a portion of a nucleotide sequence of the present invention is comprised in a DNA molecule. The DNA molecule is preferably operatively linked to a promoter functional in a cell comprising the target gene, preferably a plant cell, and introduced into the cell, in which the nucleotide sequence is expressible. The nucleotide sequence is inserted in the DNA molecule in the "sense orientation", meaning that the coding strand of the nucleotide sequence can be transcribed. In a preferred embodiment, the nucleotide sequence is fully translatable and all the genetic information comprised in the nucleotide sequence, or portion thereof, is translated into a polypeptide. In another preferred embodiment, the nucleotide sequence is partially translatable and a short peptide is translated. In a preferred embodiment, this is achieved by inserting at least one premature stop codon in the nucleotide sequence, which bring translation to a halt. In another more preferred embodiment, the nucleotide sequence is transcribed but no translation product is being made. This is usually achieved by removing the start codon, e.g. the "ATG", of the polypeptide encoded by the nucleotide sequence. In a further preferred embodiment, the DNA molecule comprising the nucleotide sequence, or a portion thereof, is stably integrated in the genome of the plant cell. In another preferred embodiment, the DNA molecule comprising the nucleotide sequence, or a portion thereof, is comprised in an extrachromosomally replicating molecule.

In transgenic plants containing one of the DNA molecules described immediately above, the expression of the nucleotide sequence corresponding to the nucleotide sequence

comprised in the DNA molecule is preferably reduced. Preferably, the nucleotide sequence in the DNA molecule is at least 70% identical to the nucleotide sequence the expression of which is reduced, more preferably it is at least 80% identical, yet more preferably at least 90% identical, yet more preferably at least 95% identical, yet more preferably at least 99% identical.

(2) "Anti-sense" Suppression

In another preferred embodiment, the alteration of the expression of a nucleotide sequence of the present invention, preferably the reduction of its expression is obtained by "anti-sense" suppression. The entirety or a portion of a nucleotide sequence of the present invention is comprised in a DNA molecule. The DNA molecule is preferably operatively linked to a promoter functional in a plant cell, and introduced in a plant cell, in which the nucleotide sequence is expressible. The nucleotide sequence is inserted in the DNA molecule in the "anti-sense orientation", meaning that the reverse complement (also called sometimes non-coding strand) of the nucleotide sequence can be transcribed. In a preferred embodiment, the DNA molecule comprising the nucleotide sequence, or a portion thereof, is stably integrated in the genome of the plant cell. In another preferred embodiment the DNA molecule comprising the nucleotide sequence, or a portion thereof, is comprised in an extrachromosomally replicating molecule. Several publications describing this approach are cited for further illustration (Green, P. J. et al., Ann. Rev. Biochem. 55:569-597 (1986); van der Krol, A. R. et al, Antisense Nuc. Acids & Proteins, pp. 125-141 (1991); Abel, P. P. et al., Proc. Natl. Acad. Sci. USA 86:6949-6952 (1989); Ecker, J. R. et al., Proc. Natl. Acad. Sci. USA 83:5372-5376 (Aug. 1986)).

In transgenic plants containing one of the DNA molecules described immediately above, the expression of the nucleotide sequence corresponding to the nucleotide sequence comprised in the DNA molecule is preferably reduced. Preferably, the nucleotide sequence in the DNA molecule is at least 70% identical to the nucleotide sequence the expression of which is reduced, more preferably it is at least 80% identical, yet more preferably at least 90% identical, yet more preferably at least 95% identical, yet more preferably at least 99% identical.

(3) Homologous Recombination

In another preferred embodiment, at least one genomic copy corresponding to a nucleotide sequence of the present invention is modified in the genome of the plant by homologous recombination as further illustrated in Paszkowski et al., EMBO Journal 7:4021-26 (1988).

This technique uses the property of homologous sequences to recognize each other and to exchange nucleotide sequences between each by a process known in the art as homologous recombination. Homologous recombination can occur between the chromosomal copy of a nucleotide sequence in a cell and an incoming copy of the nucleotide sequence introduced in the cell by transformation. Specific modifications are thus accurately introduced in the chromosomal copy of the nucleotide sequence. In one embodiment, the regulatory elements of the nucleotide sequence of the present invention are modified. Such regulatory elements are easily obtainable by screening a genomic library using the nucleotide sequence of the present invention, or a portion thereof, as a probe. The existing regulatory elements are replaced by different regulatory elements, thus altering expression of the nucleotide sequence, or they are mutated or deleted, thus abolishing the expression of the nucleotide sequence. In another embodiment, the nucleotide sequence is modified by deletion of a part of the nucleotide sequence or the entire nucleotide sequence, or by mutation. Expression of a mutated polypeptide in a plant cell is also contemplated in the present invention. More recent refinements of this technique to disrupt endogenous plant genes have been described (Kempin et al., *Nature* 389:802-803 (1997) and Miao and Lam, *Plant J.*, 7:359-365 (1995).

In another preferred embodiment, a mutation in the chromosomal copy of a nucleotide sequence is introduced by transforming a cell with a chimeric oligonucleotide composed of a contiguous stretch of RNA and DNA residues in a duplex conformation with double hairpin caps on the ends. An additional feature of the oligonucleotide is for example the presence of 2'-O-methylation at the RNA residues. The RNA/DNA sequence is designed to align with the sequence of a chromosomal copy of a nucleotide sequence of the present invention and to contain the desired nucleotide change. For example, this technique is further illustrated in US patent 5,501,967 and Zhu et al. (1999) *Proc. Natl. Acad. Sci. USA* 96: 8768-8773.

(4) Ribozymes

In a further embodiment, the RNA coding for a polypeptide of the present invention is cleaved by a catalytic RNA, or ribozyme, specific for such RNA. The ribozyme is expressed in transgenic plants and results in reduced amounts of RNA coding for the polypeptide of the present invention in plant cells, thus leading to reduced amounts of polypeptide accumulated in the cells. This method is further illustrated in US patent 4,987,071.

(5) Dominant-Negative Mutants

In another preferred embodiment, the activity of the polypeptide encoded by the nucleotide sequences of this invention is changed. This is achieved by expression of dominant negative mutants of the proteins in transgenic plants, leading to the loss of activity of the endogenous protein.

(6) Aptamers

In a further embodiment, the activity of polypeptide of the present invention is inhibited by expressing in transgenic plants nucleic acid ligands, so-called aptamers, which specifically bind to the protein. Aptamers are preferentially obtained by the SELEX (Systematic Evolution of Ligands by EXponential Enrichment) method. In the SELEX method, a candidate mixture of single stranded nucleic acids having regions of randomized sequence is contacted with the protein and those nucleic acids having an increased affinity to the target are partitioned from the remainder of the candidate mixture. The partitioned nucleic acids are amplified to yield a ligand enriched mixture. After several iterations a nucleic acid with optimal affinity to the polypeptide is obtained and is used for expression in transgenic plants. This method is further illustrated in US patent 5,270,163.

(7) Zinc finger proteins

A zinc finger protein that binds a nucleotide sequence of the present invention or to its regulatory region is also used to alter expression of the nucleotide sequence. Preferably, transcription of the nucleotide sequence is reduced or increased. Zinc finger proteins are for example described in Beerli et al. (1998) *PNAS* 95:14628-14633, or in WO 95/19431, WO 98/54311, or WO 96/06166, all incorporated herein by reference in their entirety.

(8) dsRNA

Alteration of the expression of a nucleotide sequence of the present invention is also obtained by dsRNA interference as described for example in WO 99/32619, WO 99/53050 or WO 99/61631, all incorporated herein by reference in their entirety.

(9) Insertion of a DNA molecule (Insertional mutagenesis)

In another preferred embodiment, a DNA molecule is inserted into a chromosomal copy of a nucleotide sequence of the present invention, or into a regulatory region thereof. Preferably, such DNA molecule comprises a transposable element capable of transposition in a plant cell, such as e.g. Ac/Ds, Em/Spm, mutator. Alternatively, the DNA molecule comprises a T-DNA border of an *Agrobacterium* T-DNA. The DNA molecule may also

comprise a recombinase or integrase recognition site which can be used to remove part of the DNA molecule from the chromosome of the plant cell. An example of this method is set forth in Example 2. Methods of insertional mutagenesis using T-DNA, transposons, oligonucleotides or other methods known to those skilled in the art are also encompassed. Methods of using T-DNA and transposon for insertional mutagenesis are described in Winkler et al. (1989) *Methods Mol. Biol.* 82:129-136 and Martienssen (1998) *PNAS* 95:2021-2026, incorporated herein by reference in their entireties.

(10) Deletion mutagenesis

In yet another embodiment, a mutation of a nucleic acid molecule of the present invention is created in the genomic copy of the sequence in the cell or plant by deletion of a portion of the nucleotide sequence or regulator sequence. Methods of deletion mutagenesis are known to those skilled in the art. See, for example, Miao et al, (1995) *Plant J.* 7:359. In yet another embodiment, this deletion is created at random in a large population of plants by chemical mutagenesis or irradiation and a plant with a deletion in a gene of the present invention is isolated by forward or reverse genetics. Irradiation with fast neutrons or gamma rays is known to cause deletion mutations in plants (Silverstone et al, (1998) *Plant Cell*, 10:155-169; Bruggemann et al., (1996) *Plant J.*, 10:755-760; Redei and Koncz in *Methods in Arabidopsis Research*, World Scientific Press (1992), pp. 16-82). Deletion mutations in a gene of the present invention can be recovered in a reverse genetics strategy using PCR with pooled sets of genomic DNAs as has been shown in *C. elegans* (Liu et al., (1999), *Genome Research*, 9:859-867.). A forward genetics strategy would involve mutagenesis of a line displaying PTGS followed by screening the M2 progeny for the absence of PTGS. Among these mutants would be expected to be some that disrupt a gene of the present invention. This could be assessed by Southern blot or PCR for a gene of the present invention with genomic DNA from these mutants.

(11) Overexpression in a plant cell

In yet another preferred embodiment, a nucleotide sequence of the present invention encoding a polypeptide comprising a 3'-5' exonuclease domain and/or activity in a plant cell is overexpressed. Examples of nucleic acid molecules and expression cassettes for overexpression of a nucleic acid molecule of the present invention are described infra (see Examples 8-10). Methods known to those skilled in the art of over-expression of nucleic acid molecules are also encompassed by the present invention.

In a preferred embodiment, the expression of the nucleotide sequence of the present invention is altered in every cell of a plant. This is for example obtained through homologous recombination or by insertion in the chromosome. This is also for example obtained by expressing a sense or antisense RNA, zinc finger protein or ribozyme under the control of a promoter capable of expressing the sense or antisense RNA, zinc finger protein or ribozyme in every cell of a plant. Constitutive expression, inducible, tissue-specific or developmentally-regulated expression are also within the scope of the present invention and result in a constitutive, inducible, tissue-specific or developmentally-regulated alteration of the expression of a nucleotide sequence of the present invention in the plant cell. Constructs for expression of the sense or antisense RNA, zinc finger protein or ribozyme, or for overexpression of a nucleotide sequence of the present invention, are prepared and transformed into a plant cell according to the teachings of the present invention, e.g. as described infra.

III. Methods for manipulating the expression of a nucleotide sequence of interest in a plant cell

In another aspect of the present invention, a plant cell with altered expression of a nucleotide sequence of the present invention and as described above is used to alter or stabilize the expression of a nucleotide sequence of interest in a plant cell.

In a preferred embodiment, manipulation of the expression of a heterologous nucleotide sequence of interest is desired. In this case, the heterologous nucleotide sequence is introduced into an expression cassette. The heterologous nucleotide sequence is preferably introduced into a plant cell with altered expression of a nucleotide sequence encoding a polypeptide comprising a 3'-5' exonuclease domain and/or activity. In a preferred embodiment, a plant cell with reduced expression of a nucleotide sequence encoding a polypeptide comprising a 3'-5' exonuclease domain and/or activity is used to stabilize or to increase the expression of the nucleotide sequence of interest. Alternatively, a plant cell with increased expression of a nucleotide sequence encoding a polypeptide comprising a 3'-5' exonuclease domain and/or activity is preferably used to reduce the expression of the nucleotide sequence of interest. Constitutive, inducible, tissue-specific or developmentally-regulated alteration of the nucleotide sequence of interest is preferably obtained by using a plant cell with constitutive, inducible, tissue-specific or developmentally-

regulated alteration of the expression of the nucleotide sequence encoding a polypeptide comprising a 3'-5' exonuclease domain.

In another preferred embodiment, the expression of an endogenous nucleotide sequence in a plant cell is manipulated using the present invention. In this case, a nucleotide sequence identical or substantially similar to the endogenous nucleotide sequence, or a reverse complement thereof, is introduced into a plant cell with altered expression of a nucleotide sequence of the present invention. In a preferred embodiment, a plant cell with increased expression of a nucleotide sequence of the present invention is preferably used to reduce the expression of the endogenous nucleotide sequence of interest.

Alternatively, a plant cell with reduced expression of a nucleotide sequence of the present invention is used to increase the expression of the nucleotide sequence of interest.

Constitutive, inducible, tissue-specific or developmentally-regulated alteration of the endogenous nucleotide sequence is preferably obtained by using a plant cell with constitutive, inducible, tissue-specific or developmentally-regulated alteration of the expression of its nucleotide sequence encoding a polypeptide comprising a 3'-5' exonuclease domain. Any portion of the endogenous nucleotide sequence is used. For example, if the nucleotide sequence comprises a coding region, the entire coding region or a portion thereof is used. Alternatively, a portion of the regulatory regions is used, preferably a transcribed portion of the regulatory region. Such portion is introduced into a recombinant nucleic acid molecule which is preferably introduced into an expression cassette or vector and transformed into a plant cell with altered expression of a nucleotide sequence of the present invention. Preferably, a nucleotide sequence used at least 70% identical to the endogenous nucleotide sequence, more preferably it is at least 80% identical, yet more preferably at least 90% identical, yet more preferably at least 95% identical, yet more preferably at least 99% identical.

A heterologous nucleotide sequence encodes for example, but not limited to, a polypeptide involved in waxy starch, herbicide tolerance, resistance for bacterial, fungal, or viral disease, insect resistance, enhanced nutritional quality, improved performance in an industrial process, altered reproductive capability, such as male sterility or male fertility, yield stability and yield enhancement. Using the present invention, such traits are stably and reproducibly expressed in a plant cell. Examples of endogenous nucleotide sequences of interest whose expression in a plant cell is altered using the present invention are found for example in WO 99/53050.

In another preferred embodiment, the nucleotide sequence of interest is derived from a pathogen of a plant, preferably a viral pathogen. Therefore, it is a further aspect of the present invention to provide for methods to control a pathogen. Preferably, a plant cell with altered expression of a nucleotide sequence that encodes a polypeptide comprising a 3'-5' exonuclease domain is obtained as described above. Preferably, the plant cell further comprises a nucleotide sequence substantially similar to a nucleotide sequence derived from the pathogen. Preferably, increased expression of the nucleotide sequence that encodes a polypeptide comprising a 3'-5' exonuclease domain results in increased gene silencing in the plant cell and increased resistance or tolerance to the pathogen.

IV. Plant Transformation Technology

Nucleotide sequences of the present invention can be incorporated in plant or bacterial cells using conventional recombinant DNA technology. Generally, this involves inserting a nucleotide sequence of the present invention into an expression system to which the nucleotide sequence is heterologous (i.e., not normally present) using standard cloning procedures known in the art. The vector contains the necessary elements for the transcription and translation of the inserted protein-coding sequences in a host cell containing the vector. A large number of vector systems known in the art can be used, such as plasmids, bacteriophage viruses and other modified viruses. The components of the expression system optionally are modified to increase expression. For example, truncated sequences, nucleotide substitutions or other modifications optionally are employed. Expression systems known in the art are used to transform virtually any crop plant cell under suitable conditions. Transformed cells are regenerated into whole plants.

A. Requirements for Construction of Plant Expression Cassettes

Gene sequences intended for expression in transgenic plants are first operatively linked to a suitable promoter expressible in plants. Such expression cassettes optionally comprise further sequences required or selected for the expression of the transgene. Such sequences include, but are not restricted to, transcription terminators, extraneous sequences to enhance expression such as introns, vital sequences, and sequences intended for the targeting of the gene product to specific organelles and cell compartments. These expression cassettes are easily transferred to the plant transformation vectors described *infra*. The following is a description of various components of typical expression cassettes.

1. Promoters

The selection of the promoter used determines the spatial and temporal expression pattern of the transgene in the transgenic plant. Selected promoters will express transgenes in specific cell types (such as leaf epidermal cells, mesophyll cells, root cortex cells) or in specific tissues or organs (roots, leaves or flowers, for example) and the selection will reflect the desired location of accumulation of the gene product. Alternatively, the selected promoter may drive expression of the gene under various inducing conditions. Promoters vary in their strength, i.e., ability to promote transcription. Depending upon the host cell system utilized, any one of a number of suitable promoters known in the art can be used. For example, for constitutive expression, the CaMV 35S promoter, the rice actin promoter, or the ubiquitin promoter may be used. For regulatable expression, the chemically inducible PR-1 promoter from tobacco or *Arabidopsis* may be used (see, e.g., U.S. Patent No. 5,689,044).

2. Transcriptional Terminators

A variety of transcriptional terminators are available for use in expression cassettes. These are responsible for the termination of transcription beyond the transgene and its correct polyadenylation. Appropriate transcriptional terminators are those that are known to function in plants and include the CaMV 35S terminator, the *tml* terminator, the nopaline synthase terminator and the pea *rbcS* E9 terminator. These can be used in both monocotyledons and dicotyledons.

3. Sequences for the Enhancement or Regulation of Expression

Numerous sequences are known to enhance gene expression from within the transcriptional unit and these sequences can be used in conjunction with the genes of this invention to increase their expression in transgenic plants. For example, various intron sequences such as introns of the maize *Adhl* gene have been shown to enhance expression, particularly in monocotyledonous cells. In addition, a number of non-translated leader sequences derived from viruses also are known to enhance expression, and these are particularly effective in dicotyledonous cells.

4. Coding Sequence Optimization

The coding sequence of the selected gene optionally is genetically engineered by altering

the coding sequence for optimal expression in the crop species of interest. Methods for modifying coding sequences to achieve optimal expression in a particular crop species are well known (see, *e.g.* Perlak *et al.*, *Proc. Natl. Acad. Sci. USA* 88: 3324 (1991); and Koziel *et al.*, *Bio/technol.* 11: 194 (1993); Fennoy and Bailey-Serres. *Nucl. Acids Res.* 21: 5294-5300 (1993). Methods for modifying coding sequences by taking into account codon usage in plant genes and in higher plants, green algae, and cyanobacteria are well known (see table 4 in: Murray *et al.* *Nucl. Acids Res.* 17: 477-498 (1989); Campbell and Gowri *Plant Physiol.* 92: 1-11(1990).

5. Targeting of the Gene Product Within the Cell

Various mechanisms for targeting gene products are known to exist in plants and the sequences controlling the functioning of these mechanisms have been characterized in some detail. For example, the targeting of gene products to the chloroplast is controlled by a signal sequence found at the amino terminal end of various proteins which is cleaved during chloroplast import to yield the mature protein (*e.g.* Comai *et al.* *J. Biol. Chem.* 263: 15104-15109 (1988)). Other gene products are localized to other organelles such as the mitochondrion and the peroxisome (*e.g.* Unger *et al.* *Plant Molec. Biol.* 13: 411-418 (1989)). The cDNAs encoding these products are manipulated to effect the targeting of heterologous gene products to these organelles. In addition, sequences have been characterized which cause the targeting of gene products to other cell compartments. Amino terminal sequences are responsible for targeting to the ER, the apoplast, and extracellular secretion from aleurone cells (Koehler & Ho, *Plant Cell* 2: 769-783 (1990)). Additionally, amino terminal sequences in conjunction with carboxy terminal sequences are responsible for vacuolar targeting of gene products (Shinshi *et al.* *Plant Molec. Biol.* 14: 357-368 (1990)). By the fusion of the appropriate targeting sequences described above to transgene sequences of interest one skilled in the art is able to direct the transgene product to any organelle or cell compartment.

B. Construction of Plant Transformation Vectors

Numerous transformation vectors available for plant transformation are known to those of ordinary skill in the plant transformation arts, and the genes pertinent to this invention are used in conjunction with any such vectors. The selection of vector will depend upon the preferred transformation technique and the target species for transformation. For certain target species, different antibiotic or herbicide selection markers may be preferred.

Selection markers used routinely in transformation include the *nptII* gene, which confers resistance to kanamycin and related antibiotics (Messing & Vierra. *Gene* 19: 259-268 (1982); Bevan et al., *Nature* 304:184-187 (1983)), the *bar* gene, which confers resistance to the herbicide phosphinothricin (White et al., *Nucl. Acids Res* 18: 1062 (1990), Spencer et al. *Theor. Appl. Genet* 79: 625-631 (1990)), the *hph* gene, which confers resistance to the antibiotic hygromycin (Blochinger & Diggelmann, *Mol Cell Biol* 4: 2929-2931), and the *dhfr* gene, which confers resistance to methotrexate (Bourouis et al., *EMBO J.* 2(7): 1099-1104 (1983)), and the EPSPS gene, which confers resistance to glyphosate (U.S. Patent Nos. 4,940,935 and 5,188,642).

1. Vectors Suitable for *Agrobacterium* Transformation

Many vectors are available for transformation using *Agrobacterium tumefaciens*. These typically carry at least one T-DNA border sequence and include vectors such as pBIN19 (Bevan, *Nucl. Acids Res.* (1984)). Typical vectors suitable for *Agrobacterium* transformation include the binary vectors pCIB200 and pCIB2001, as well as the binary vector pCIB10 and hygromycin selection derivatives thereof. (See, for example, U.S. Patent No. 5,639,949).

2. Vectors Suitable for non-*Agrobacterium* Transformation

Transformation without the use of *Agrobacterium tumefaciens* circumvents the requirement for T-DNA sequences in the chosen transformation vector and consequently vectors lacking these sequences can be utilized in addition to vectors such as the ones described above which contain T-DNA sequences. Transformation techniques that do not rely on *Agrobacterium* include transformation via particle bombardment, protoplast uptake (e.g. PEG and electroporation) and microinjection. The choice of vector depends largely on the preferred selection for the species being transformed. Typical vectors suitable for non-*Agrobacterium* transformation include pCIB3064, pSOG19, and pSOG35. (See, for example, U.S. Patent No. 5,639,949).

C. Transformation Techniques

Once the coding sequence of interest has been cloned into an expression system, it is transformed into a plant cell. Methods for transformation and regeneration of plants are well known in the art. For example, Ti plasmid vectors have been utilized for the delivery of foreign DNA, as well as direct DNA uptake, liposomes, electroporation, micro-injection, and microprojectiles. In addition, bacteria from the genus *Agrobacterium* can be utilized to

transform plant cells.

Although a nucleotide sequence of the present invention can be inserted into any plant cell falling within these broad classes, it is particularly useful in crop plant cells, such as rice, wheat, barley, rye, corn, potato, carrot, sweet potato, sugar beet, bean, pea, chicory, lettuce, cabbage, cauliflower, broccoli, turnip, radish, spinach, asparagus, onion, garlic, eggplant, pepper, celery, carrot, squash, pumpkin, zucchini, cucumber, apple, pear, quince, melon, plum, cherry, peach, nectarine, apricot, strawberry, grape, raspberry, blackberry, pineapple, avocado, papaya, mango, banana, soybean, tobacco, tomato, sorghum and sugarcane.

Transformation techniques for dicotyledons are well known in the art and include *Agrobacterium*-based techniques and techniques that do not require *Agrobacterium*. Non-*Agrobacterium* techniques involve the uptake of exogenous genetic material directly by protoplasts or cells. This can be accomplished by PEG or electroporation mediated uptake, particle bombardment-mediated delivery, or microinjection. In each case the transformed cells are regenerated to whole plants using standard techniques known in the art.

Transformation of most monocotyledon species has now also become routine. Preferred techniques include direct gene transfer into protoplasts using PEG or electroporation techniques, particle bombardment into callus tissue, as well as *Agrobacterium*-mediated transformation.

D. Plastid Transformation

In another preferred embodiment, a nucleotide sequence of the present invention is directly transformed into the plastid genome. Plastid expression, in which genes are inserted by homologous recombination into the several thousand copies of the circular plastid genome present in each plant cell, takes advantage of the enormous copy number advantage over nuclear-expressed genes to permit expression levels that can readily exceed 10% of the total soluble plant protein. In a preferred embodiment, the nucleotide sequence is inserted into a plastid targeting vector and transformed into the plastid genome of a desired plant host. Plants homoplasmic for plastid genomes containing the nucleotide sequence are obtained, and are preferentially capable of high expression of the nucleotide sequence.

Plastid transformation technology is for example extensively described in U.S. Patent Nos. 5,451,513, 5,545,817, 5,545,818, and 5,877,462 in PCT application no. WO 95/16783 and WO 97/32977, and in McBride *et al.* (1994) Proc. Natl. Acad. Sci. USA 91, 7301-7305, all incorporated herein by reference in their entirety. The basic technique for plastid

transformation involves introducing regions of cloned plastid DNA flanking a selectable marker together with the nucleotide sequence into a suitable target tissue, e.g., using biolistics or protoplast transformation (e.g., calcium chloride or PEG mediated transformation). The 1 to 1.5 kb flanking regions, termed targeting sequences, facilitate homologous recombination with the plastid genome and thus allow the replacement or modification of specific regions of the plastome. Initially, point mutations in the chloroplast 16S rRNA and rps12 genes conferring resistance to spectinomycin and/or streptomycin are utilized as selectable markers for transformation (Svab, Z., Hajdukiewicz, P., and Maliga, P. (1990) *Proc. Natl. Acad. Sci. USA* 87, 8526-8530; Staub, J. M., and Maliga, P. (1992) *Plant Cell* 4, 39-45). The presence of cloning sites between these markers allowed creation of a plastid targeting vector for introduction of foreign genes (Staub, J.M., and Maliga, P. (1993) *EMBO J.* 12, 601-606). Substantial increases in transformation frequency are obtained by replacement of the recessive rRNA or r-protein antibiotic resistance genes with a dominant selectable marker, the bacterial *aadA* gene encoding the spectinomycin-detoxifying enzyme aminoglycoside-3'-adenyltransferase (Svab, Z., and Maliga, P. (1993) *Proc. Natl. Acad. Sci. USA* 90, 913-917). Other selectable markers useful for plastid transformation are known in the art and encompassed within the scope of the invention.

V. Recombinant Production of Polypeptides and Uses Thereof

In a further aspect, the present invention discloses the use of a nucleotide sequence of the present invention to recombinantly produce a polypeptide having 3'-5' exonuclease activity. For recombinant production of a polypeptide in a host organism, a nucleotide sequence of the present invention is inserted into an expression cassette designed for the chosen host and introduced into the host where it is recombinantly produced. The choice of specific regulatory sequences such as promoter, signal sequence, 5' and 3' untranslated sequences, and enhancer appropriate for the chosen host is within the level of skill of the routineer in the art. The resultant molecule, containing the individual elements operably linked in proper reading frame, is inserted into a vector capable of being transformed into the host cell. Suitable expression vectors and methods for recombinant production of proteins are well known for host organisms such as *E. coli*, yeast, and insect cells (see, e.g., Luckow and Summers, *Bio/Technol.* 6: 47 (1988)). Specific examples include plasmids such as pBluescript (Stratagene, La Jolla, CA), pFLAG

(International Biotechnologies, Inc., New Haven, CT), pTrcHis (Invitrogen, La Jolla, CA), and baculovirus expression vectors, e.g., those derived from the genome of *Autographica californica* nuclear polyhedrosis virus (AcMNPV). A preferred baculovirus/insect system is pVI11392/Sf21 cells (Invitrogen, La Jolla, CA).

Recombinantly produced polypeptide is isolated and purified using a variety of standard techniques. The actual techniques used varies depending upon the host organism used, whether the enzyme is designed for secretion, and other such factors. Such techniques are well known to the skilled artisan (*see, e.g.* chapter 16 of Ausubel, F. *et al.*, "Current Protocols in Molecular Biology", pub. by John Wiley & Sons, Inc. (1994)).

Recombinantly produced polypeptides are useful for a variety of purposes. For example, they are used in assays to screen for chemicals that interact with the polypeptide or that alter the activity of the polypeptide.

VI. Method to assay a compound that interact with a polypeptide of the present invention

In another aspect of the present invention, assays to identify a compound that interacts with a polypeptide comprising a 3'-5' exonuclease domain are disclosed. In a preferred embodiment, such a compound is capable of altering the activity of the polypeptide. Preferably, the compound is capable of inhibiting or stimulating the activity of the polypeptide. Preferably, such compound is applied to a plant or a plant cell, and, as a result, the activity of the polypeptide in the plant or plant cell is altered. In such plant or plant cell, the expression of a nucleotide sequence of interest and as described above is altered. The present invention thus further discloses methods to alter the expression of a nucleotide sequence of interest in a plant or plant cell comprising applying to said plant or plant cell a compound capable of inhibiting the activity of a nucleotide sequence of said plant or plant cell that encodes a polypeptide comprising a 3'-5' exonuclease domain. In a preferred embodiment, the nucleotide sequence of interest is a heterologous or an endogenous nucleotide sequence. Preferably, the plant cell comprises the heterologous nucleotide sequence as described above in section II. Preferably, the plant cell comprises a nucleotide sequence identical or substantially similar to the endogenous nucleotide sequence as described above in section II.

1. *In Vitro* Inhibitor Assays: Discovery of compounds that Interacts with a polypeptide of the present invention

Three methods (fluorescence correlation spectroscopy, surface-enhanced laser desorption/ionization, and biacore technologies) that can detect interactions between a polypeptide and a compound are described below.

Fluorescence Correlation Spectroscopy (FCS) theory was developed in 1972 but it is only in recent years that the technology to perform FCS became available (Madge et al. (1972) Phys. Rev. Lett., 29: 705-708; Maiti et al. (1997) Proc. Natl. Acad. Sci. USA, 94: 11753-11757). FCS measures the average diffusion rate of a fluorescent molecule within a small sample volume. The sample size can be as low as 10^3 fluorescent molecules and the sample volume as low as the cytoplasm of a single bacterium. The diffusion rate is a function of the mass of the molecule and decreases as the mass increases. FCS can therefore be applied to protein-ligand interaction analysis by measuring the change in mass and therefore in diffusion rate of a molecule upon binding. In a typical experiment, the target to be analyzed is expressed as a recombinant protein with a sequence tag, such as a poly-histidine sequence, inserted at the N or C-terminus. The expression takes place in *E. coli*, yeast or insect cells. The protein is purified by chromatography. For example, the poly-histidine tag can be used to bind the expressed protein to a metal chelate column such as Ni²⁺ chelated on iminodiacetic acid agarose. The protein is then labeled with a fluorescent tag such as carboxytetramethylrhodamine or BODIPY® (Molecular Probes, Eugene, OR). The protein is then exposed in solution to the potential ligand, and its diffusion rate is determined by FCS using instrumentation available from Carl Zeiss, Inc. (Thornwood, NY). Ligand binding is determined by changes in the diffusion rate of the protein.

Surface-Enhanced Laser Desorption/Ionization (SELDI) was invented by Hutchens and Yip during the late 1980's (Hutchens and Yip (1993) Rapid Commun. Mass Spectrom. 7: 576-580). When coupled to a time-of-flight mass spectrometer (TOF), SELDI provides a mean to rapidly analyze molecules retained on a chip. It can be applied to ligand-protein interaction analysis by covalently binding the target protein on the chip and analyze by MS the small molecules that bind to this protein (Worrall et al. (1998) Anal. Biochem. 70: 750-756). In a typical experiment, the target to be analyzed is expressed as described for FCS. The purified protein is then used in the assay without further preparation. It is bound to the SELDI chip either by utilizing the poly-histidine tag or by other interaction such as ion exchange or hydrophobic interaction. The chip thus prepared is then exposed to the

potential ligand via, for example, a delivery system capable to pipet the ligands in a sequential manner (autosampler). The chip is then submitted to washes of increasing stringency, for example a series of washes with buffer solutions containing an increasing ionic strength. After each wash, the bound material is analyzed by submitting the chip to SELDI-TOF. Ligands that specifically bind the target will be identified by the stringency of the wash needed to elute them.

Biacore relies on changes in the refractive index at the surface layer upon binding of a ligand to a protein immobilized on the layer. In this system, a collection of small ligands is injected sequentially in a 2-5 microlitre cell with the immobilized protein. Binding is detected by surface plasmon resonance (SPR) by recording laser light refracting from the surface. In general, the refractive index change for a given change of mass concentration at the surface layer, is practically the same for all proteins and peptides, allowing a single method to be applicable for any protein (Liedberg et al. (1983) *Sensors Actuators* 4: 299-304; Malmquist (1993) *Nature*, 361: 186-187). In a typical experiment, the target to be analyzed is expressed as described for FCS. The purified protein is then used in the assay without further preparation. It is bound to the Biacore chip either by utilizing the poly-histidine tag or by other interaction such as ion exchange or hydrophobic interaction. The chip thus prepared is then exposed to the potential ligand via the delivery system incorporated in the instruments sold by Biacore (Uppsala, Sweden) to pipet the ligands in a sequential manner (autosampler). The SPR signal on the chip is recorded and changes in the refractive index indicate an interaction between the immobilized target and the ligand. Analysis of the signal kinetics on rate and off rate allows the discrimination between non-specific and specific interaction.

2. *In Vivo* Inhibitor Assay

In another embodiment, an *in vivo* screening assay for compounds altering the activity of a polypeptide encoded by a nucleotide sequence of the present invention uses transgenic plants, plant tissue, plant seeds or plant cells capable of overexpressing a nucleotide sequence of the present invention.

A chemical is then applied to the transgenic plants, plant tissue, plant seeds or plant cells and to the isogenic non-transgenic plants, plant tissue, plant seeds or plant cells, and gene silencing in the transgenic and non-transformed plants, plant tissue, plant seeds or plant cells is determined after application of the chemical and compared.

VI. Assays for testing the alteration of gene silencing

Several methods are described to test for the alteration of gene silencing in a plant cell.

A. Introduction of a marker gene in a plant cell and analysis of its expression

A marker gene is introduced into wild-type lines and into lines with potentially altered gene silencing. An alteration in gene silencing is detected as a difference in the T1 progeny in the number of lines exhibiting low levels of marker activity vs. high levels of marker activity. Lines with high levels of marker activity are not likely to be silenced, whereas lines with low levels of or without activity are likely to be silenced. Choices for a non-endogenous marker gene include luciferase, green fluorescent protein (GFP), or beta-glucuronidase (GUS). Assay methods for each of these markers have been described (Ishitani et al. (1997) *Plant Cell*, 9:1935-1949; Cutler et al. (2000) *Proc. Natl. Acad. Sci. USA* 97: 3718-3723; Jefferson et al. (1989) *EMBO J.*, 6:3901- 3907).

B. Analysis of the expression of an endogenous gene

This assay method is similar to the one above, except that an endogenous gene is used in place of a marker gene. The expression of the endogenous gene is measured in wild-type lines and in lines with potentially altered gene silencing. Both types of lines further comprise a transgenic "silencing" construct used to silence the endogenous gene. Such "silencing" construct for example comprises a promoter directing the transcription of the endogenous gene in a sense orientation, or an antisense orientation, or in both an antisense and a sense orientation in the same transcript. The promoter is for example constitutive, like ACTIN2 (An et al., 1996, *Plant J.*, 10:107-121), or inducible, like PR1 (see e.g. US Patent 5,614,395), or activatable by a hybrid transcription factor (Guyer et al., 1998, *Genetics* 149:633-639). In plants with the "silencing" transgene, the level of gene silencing is assessed by analyzing alterations in the function of the endogenous gene, for example appearance of a mutant phenotype, relative to plants without the transgene. By comparing the range of phenotypes observed in the T1 progeny of these plants, it is determined whether the original lines have altered gene silencing capabilities. Endogenous genes that are used include for example: APETALA1, which has a mutant phenotype in which petals are absent and sepals are converted to leaves with axillary flowers (Bowman et al., 1989, *Plant Cell*, 1:37-52), GLABROUS1, which has a mutant phenotype in which the number of trichomes on leaves is greatly reduced (Oppenheimer et al, 1991, *Cell* 67:483-493), and NIM1 (also known as NPR1), which has a mutant phenotype in some ecotypes in which

Peronospora isolates become infectious and SAR genes such as PR1 are not induced (for example Ryals et al. (1997) *Plant Cell* 9:425-39). Induction of PR1 can be detected by Northern or RT-PCR.

C. Analysis of the expression of a characterized silenced transgene

To determine whether a given line alters gene silencing, introduction of a characterized silenced (either post transcriptionally or transcriptionally) gene is accomplished by crossing the line in question with a line with a characterized silenced gene and examining the effects in the F1 and F2 progeny. For a line with a characterized silenced gene, the experiment measures changes in the levels of expression of this gene in the mutant backgrounds. For recessive mutations that might alter gene silencing, it is necessary to compare F2 progeny homozygous, heterozygous, and wild type for the mutant allele for differences in expression levels of the silenced gene. For dominant mutations that might alter gene silencing, it is possible to compare F1 progeny heterozygous and wild type for the mutant allele for differences in expression levels of the silenced gene. A line with a constitutive promoter and a marker gene is used in such experiments.

VII. Assay for 3'-5' exonuclease activity

Assays are available to test for 3'-5' exonuclease activity in the polypeptides encoded by the nucleotide molecules and sequences of the present invention. Assays for 3'-5' exonuclease activity are set forth in Kamath-Loeb et al. (1998) *J. Biol. Chem.* 273:34145-50, Huang et al., (1998) *Nat. Genet.* 20:114-6, and Suzuki et al. (1999) *Nucleic Acids Res.* 27:2361-8, each incorporated by reference in their entireties. Briefly, the polypeptide or protein is incubated with radioactively labeled DNA oligomers. After incubation, the reaction products are analyzed by polyacrylamide gel electrophoresis.

VIII. Polypeptides encoded by the nucleic acid molecules.

The present invention provides polypeptides encoded by the nucleic acid molecules of the invention and variants thereof. These polypeptides are exemplified by those encoded by the nucleotide sequences of SEQ ID NOS: 2, 4, 6, 22, 18, 12, 14, 24, 36 and 38; polypeptides encoded by nucleic acid sequences having at least 70% sequence similarity to the sequences of SEQ ID NOS: 1, 3, 5, 21, 9, 11, 13, 23, 35 or 37 and variants and mutants thereof. Preferably, the isolated and substantially purified polypeptides are identical or substantially similar to the amino acid sequence of SEQ ID NO:24.

The polypeptides of the invention may be altered in various ways including amino acid substitutions, deletions, truncations, and insertions. Methods for such manipulations are generally known in the art. Methods for mutagenesis and nucleotide sequence alterations are well known in the art. See, for example, Kunkel, Proc. Natl. Acad. Sci. USA, 82:488, (1985); Kunkel et al., Methods in Enzymol., 154:367 (1987); US Patent No. 4,873,192; Walker and Gaastra, eds. (1983) Techniques in Molecular Biology (MacMillan Publishing Company, New York) and the references cited therein. Guidance as to appropriate amino acid substitutions that do not affect biological activity of the protein of interest may be found in the model of Dayhoff et al. (1978) Atlas of Protein Sequence and Structure (Natl. Biomed. Res. Found., Washington, D.C.), herein incorporated by reference. Conservative substitutions may be preferred.

The proteins of the invention encompass both naturally occurring polypeptides as well as variants and modified forms thereof. Obviously, the mutations that will be made in the DNA encoding the mutation must not place the sequence out of reading frame and preferably will not create complementary regions that could produce secondary mRNA structure. See, EP Patent Application Publication No. 75,444.

The invention will be further described by reference to the following detailed examples. These examples are provided for purposes of illustration only, and are not intended to be limiting unless otherwise specified.

EXAMPLES

Example 1: Identification of polypeptides comprising a 3'-5' exonuclease domain

Method 1

Using the MEME and MotifSearch programs of GCG SEQWEB (version 1.1, University of Wisconsin), seven Arabidopsis polypeptide sequences potentially containing RNase D-related motifs are identified. MEME starts with a set of unaligned polypeptide sequences and identifies common motifs. Then, these motifs are used to create gapless profiles that can be used as input to MotifSearch to search other sequences for these motifs.

First, the *C. elegans* *mut-7* gene (ZK1098.8, GenPept accession CAA80137) is used in a BLASTP search to identify related Arabidopsis polypeptide sequences. One polypeptide sequence is identified (GenPept accession CAB36851, SEQ ID NO:2). Second, sequences of several of the proteins in Branch B of Figure 4 of Moser et al. (1997) (Nucl. Acids Res. 25:5110-5118) are used together with the Arabidopsis predicted polypeptide sequence

(GenPept accession CAB36851, SEQ ID NO:2) to identify common motifs with the MEME program. These protein sequences include: entire *C. elegans mut-7* (GenPept accession CAA80137, SEQ ID NO:16, corresponding nucleotide sequence SEQ ID NO:15), C-terminus (amino acid positions 428 to end) of *C. elegans mut-7*-related protein (ZK1098.3, GenPept accession CAA80141), C-terminus (amino acid positions 291 to end) of *H. sapiens* 100 kDa nucleolar Polymyositis Scleroderma autoantigen (PMSC100, GenPept accession CAA46904), C-terminus (amino acid positions 216 to end) of *S. cerevisiae* RRP6 (GenPept accession NP_014643), N-terminus (amino acid positions 1 to 333) of *H. sapiens* Werner syndrome protein (WRN, GenPept accession AAF06162, SEQ ID NO:18, corresponding nucleotide sequence SEQ ID NO:17), entire *E. coli* RNase D (SwissProt accession P09155), and C-terminus (amino acid positions 546 to end) of *D. melanogaster Egalitarian* (EGL, GenPept accession AAB49975, and entire phage phi-C31 hypothetical protein 11 (GenPept accession CAA53907). Truncated versions of some proteins are used to allow identification of RNase D related motifs in polypeptides with other sequence regions or motifs. Third, five MEME motifs are identified. Fourth, MotifSearch is used to search GenPept Plant division for sequences containing these motifs and seven Arabidopsis polypeptide sequences are identified. The GenPept accessions for these sequences are listed from lowest to highest P-value from the MotifSearch program:

CAB36851 (SEQ ID NO:2, corresponding nucleotide sequence SEQ ID NO:1),
 AAC69936 (SEQ ID NO:6, corresponding nucleotide sequence SEQ ID NO:5),
 AAD25623 (SEQ ID NO:4, corresponding nucleotide sequence SEQ ID NO:3),
 AAD26968 (SEQ ID NO:10, corresponding nucleotide sequence SEQ ID NO:9),
 AAC25931 (SEQ ID NO:12, corresponding nucleotide sequence SEQ ID NO:11),
 AAC42241 (SEQ ID NO:8, corresponding nucleotide sequence SEQ ID NO:7),
 AAF98185 (SEQ ID NO:14, corresponding nucleotide sequence SEQ ID NO:13).

A lower value has greater probability of being significantly different from random.

The inventors of the present invention also discovered that the 5' end of GenPept accession AAC42241 is missing due to incorrect annotation, and that GenPept accession AAC42241 lacks the exo I motif of the 3'-5' exonuclease domain. The amino acid sequence comprising the entire 3'-5' exonuclease domain (including exo I) is disclosed for the first time in the instant application and is set forth in SEQ ID NO:22. The corresponding nucleotide sequence is set forth in SEQ ID NO:21.

Method 2

The *C. elegans* *mut-7* protein contains a 3'-5' exonuclease domain. The HMMsearch (hidden Markov model) program (Eddy, S.R. (1996) *Curr. Opin. Struct. Biol.* 6:361-365) is used to search the GenPept plant division for protein sequences with the 3'-5' exonuclease profile, which is found in the Pfam database (A. Bateman, et al. (2000) *Nucleic Acids Research*, 28:263-266, incorporated herein by reference in its entirety). Pfam is a database of multiple alignments of protein domains or conserved protein regions. These alignments represent some evolutionary conserved structure that has implications for the protein's function. Profile HMMs built from the Pfam alignments are used for automatically recognizing that new proteins belong to an existing protein family, even if the sequence similarity is weak. Five Arabidopsis polypeptide sequences are identified. The GenPept accessions for these sequences are listed from lowest to highest E-value from the HMMsearch program: AAD25623, AAC69936, CAB36851, AAC42241 and AAD26968. A lower value has greater probability of being significantly different from random.

The 3'-5' exonuclease domain consists of three sequence motifs termed Exo I, Exo II, and Exo III (Moser et al. (1997) *Nucl. Acids Res.* 25:5110-5118). These motifs are clustered around the active site and contain four negatively charged amino acids that serve as ligands for the two metal ions necessary for catalysis in addition to a catalytically active tyrosine. The presence of these amino acids in SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:22, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16, and SEQ ID NO:18, and their position in the corresponding amino acid sequences is indicated in Table 1 below. The positions of the exo I, exo II, and exo III motifs in these amino acid sequences is shown in Table 2.

Method 3

Additional Arabidopsis genes that encode proteins with 3'-5' exonuclease domains are identified starting from the experimentally determined cDNA sequence for the gene encoded by SEQ ID NO:1. This protein sequence (SEQ ID NO:24) is identified as described in Example 6 (see below). A BLASTp search (Altschul, S.F. et al. (1997) *Nucleic Acids Res.* 25:3389-3402) with SEQ ID NO:24 as the query reveals that there are several predicted Arabidopsis proteins with E values less than 1E-03. From this group of Arabidopsis proteins, a HMM search is performed with each protein as a query compared to the HMM database of PFAM models (version 6.4) on a TimeLogic DeCypher machine. Two Arabidopsis proteins (AAG50917 (SEQ ID NO:36) and BAB11227 (SEQ ID NO:38)) are

identified that were not identified by methods 1 and 2 earlier. The nucleotide sequences corresponding to AAG50917 and BAB11227 are SEQ ID NO:35 and SEQ ID NO:37, respectively. These proteins have E values from the HMM search that are below 1E-01 and likely to be significant. For comparison, the E-values for all the Arabidopsis proteins identified in Example 1 are shown in the third column in Tables 1 and 2 as recalculated with the same HMM program, PFAM model, and computer server.

Table 1:

Accession #	E-value HMM	E-value HMM**	exo I	exo II	exo III
AAD25623 (SEQ ID NO:4)	4.6E-54	2.7E-67	D140, E142	D199	Y264, D268
AAC69936 (SEQ ID NO:6)	1.5E-44	5.8E-82	D80, E82	D138	Y203, D207
CAB36851 (SEQ ID NO:2)	2.0E-04	9.8E-12	D133, E135	D194	Y263, D267
AAC42241* (SEQ ID NO:22)	1.5E-01	7.6E-27	D50, E52	D108	A192, D196
AAD26968 (SEQ ID NO:10)	5.1E+00	-	D61, Q63	D118	Q186, D190
AAC25931 (SEQ ID NO:12)	-	-	G70, Q72	D127	Q195, D199
AAF98185 (SEQ ID NO:14)	-	-	-	-?	Y60, R64
AAG50917 (SEQ ID NO:36)	-	6.6E-11	D384, E386	D449	Y531, D535
BAB11227 (SEQ ID NO:38)	-	2.4E-14	D34, E36	D104	Y186, D190
CAA80137 (SEQ ID NO:16)	-	-	D435, E437	D503	Y585, D589
AAF06162 (SEQ ID NO:18)	-	-	D82, E84	D143	Y212, D216

*: using corrected sequence because 5' end is missing due to incorrect annotation, so that exo 1 may be present.

** : HMM PFAM analysis done as per Method 3 of Example 1

exo I, II, & III motifs defined as in Figure 6 of Mian (1997) *Nucleic Acids Res* 25:3187

the letters in the exo I, II, & III columns stand for the standard single letter amino acid code, and the numbers stand for the position of the amino acid in the polypeptide.

Table 2:

Accession #	E-value HMM	E-value HMM**	exo I	exo II	exo III
AAD25623 (SEQ ID NO:4)	4.6E-54	2.7E-67	136-145	191-206	261-271
AAC69936 (SEQ ID NO:6)	1.5E-44	5.8E-82	76-85	130-135	200-210
CAB36851 (SEQ ID NO:2)	2.0E-04	9.8E-12	129-138	186-201	260-270
AAC42241* (SEQ ID NO:22)	1.5E-01	7.6E-27	46-55	100-115	189-199
AAD26968 (SEQ ID NO:10)	5.1E+00	-	57-66	110-125	183-193
AAC25931 (SEQ ID NO:12)	-	-	66-75	119-134	192-202
AAF98185 (SEQ ID NO:14)	-	-	-	-	57-67
AAG50917 (SEQ ID NO:36)		6.6E-11	380-389	441-456	528-538
BAB11227 (SEQ ID NO:38)		2.4E-14	30-39	96-111	183-193
CAA80137 (SEQ ID NO:16)	-	-	431-440	495-510	582-592
AAF06162 (SEQ ID NO:18)	-	-	78-87	135-150	209-219

*: using corrected sequence because 5' end is missing due to incorrect annotation, so that exo 1 may be present.

** : HMM PFAM analysis done as per Method 3 of Example 1

Example 2: Insertion mutagenesis in a nucleotide sequence encoding a polypeptide comprising a RNase D related domain

Insertion mutagenesis facilitates direct reverse genetic screens by providing a physical link to the gene of interest. In plants both T-DNA and transposon insertion mutagens have been utilized as insertion mutagens (Winkler et al. (1998) *Methods Mol Biol.* 82:129-136, Martienssen (1998) *PNAS* 95:2021-2026). T-DNA insertions within any given gene can be detected by polymerase chain reaction (PCR) methods utilizing one gene specific primer and one T-DNA specific primer (Winkler et al. (1998) *Plant Physiol.* 3:743-750, and Krysan

et al. (1999) *Plant Cell*, 11:2283-2290). Specific PCR product is formed only when a T-DNA element has inserted either within or close to the gene of interest. Due to the exponential nature of PCR amplification, it is possible to screen many thousands of independently transformed *Arabidopsis* mutants by sample pooling (Krysan et al., 1999). Once a T-DNA pool is identified with an insertion in the gene of interest, the process of isolating a single plant with that insertion requires de-convolution of the pool architecture.

To assess the function of a polypeptide encoded by the nucleotide sequence set forth in SEQ ID NO:1, a pool of ~60,480 independent tagged *Arabidopsis* lines (Krysan et al., 1999) is screened by PCR utilizing pairs of primers corresponding to the T-DNA left border and the SEQ ID NO:1 3'-specific region. The SEQ ID NO:1 3' specific primer (5'- cga cat gat ctg ata cat cgt tat gcc att -3', SEQ ID NO:19) corresponds to position 96817-96790 on BAC F18A5, GenBank accession number AL035528.2. The left border primer from *A. tumefaciens* T-DNA vector pD991 is represented by SEQ ID NO:20 (5'- cat ttt ata ata acg ctg cgg aca tct ac -3'). (Krysan et al., 1999). One specific PCR product is identified, isolated and designated S11.13. Sequencing of the PCR-amplified fragment reveals a T-DNA insertion 26 bp 5' of the predicted CDS region of SEQ ID NO:1. De-convolution of pool architecture as described (Krysan et al., 1999) leads to the identification of seven individual lines containing the specific T-DNA element, designated S11.13-8, S11.13-13, S11.13-34, S11.13-38, S11.13-41, S11.13-44, S11.13-48. PCR is subsequently utilized for genotyping individual lines. All of the lines are heterozygous for the insertion, except S11.13-34 is homozygous for the insertion. No visible phenotype is observed in line S11.13-34 at the embryo and seedling stages.

Example 3: Analysis of the expression of a characterized silenced transgene in *Arabidopsis* line S11.13-34

Line S11.13-34 (see Example 2 above) is crossed with line L1, which has been shown to have a post-transcriptionally silenced GUS transgene (Elmayan et al. (1998) *Plant Cell* 10:1747-1758). Individual F1 progeny with a silenced GUS transgene are allowed to self fertilize. About 100 F2 progeny from individual F1 plants are grown and tested for GUS activity. The genotype of each F2 plant with respect to the T-DNA insertion in the RNase D related domain (RDRD) gene is determined by PCR as described in Example 2. Similarly, the presence of the GUS transgene is determined by PCR for each plant. Levels of GUS

activity in plants homozygous for the insertion in the RNase D related domain gene are compared to plants heterozygous for the insertion RDRD gene and wild-type plants.

Example 4: The *Arabidopsis thaliana* transgenic lines 8Z-2 and 5 exhibit post-transcriptional silencing of a green-fluorescent protein reporter gene *Agrobacterium*-mediated transformation as described by Bechtold (Methods in Molecular Biology 82: 259-266, 1998) is used to obtain transgenic *Arabidopsis thaliana* ecotype Columbia plants exhibiting PTGS. The Ti-plasmid used contains a chimeric green fluorescent protein (GFP) (Reichel et al. (1996) PNAS 93: 5888-93) reporter gene regulated by a duplicated cauliflower mosaic virus (CaMV) 35S RNA promoter and transcriptional terminator (Goodall and Filipowicz (1989) Cell 58: 473-483) in the binary vector pBIN19 (Bevan (1984) Nucl. Acids Res. 12: 8711-8721). The plasmid is designated p35S-GFP.

To evaluate PTGS in the resultant 35S-GFP transformants, GFP expression is monitored in transgenic plants by GFP excitation with UV light (approximate range of wavelengths 390 to 480 nm). Selection of transgenic lines showing PTGS is based on absence of GFP expression in mature plants that showed normal GFP expression in earlier stages of plant development. Based on this criterion, two lines designated as 8Z-2 and 5, which are homozygous for the T-DNA insertion, show PTGS associated with greatly reduced GFP-mRNA levels detected by RNA blot hybridization as described by Sambrook et al. (Molecular Cloning, 2nd edition, 1989). Line 8Z-2 shows PTGS in approximately 90-96% of sibling plants. Line 5 shows PTGS in approximately 30-50% of sibling plants.

DNA blot hybridization as described by Sambrook et al. (Molecular Cloning, 2nd edition, 1989) reveals that post-transcriptionally silenced line 8Z-2 carries two copies of T-DNA. Further analysis based on polymerase chain reaction (PCR) and utilizing combinations of T-DNA specific primers (Kumar and Fladung (2000) BioTechniques 28: 1128-1137) shows that these two copies are arranged in a direct tandem repeat. Similarly, line 5 is shown to carry one full-length T-DNA and a second, truncated T-DNA copy arranged in an inverted tandem repeat. The genomic position of the T-DNA copies in line 8Z-2 is determined to be chromosome I, BAC F22L4, gene #11 by thermal asymmetric interlaced polymerase chain reaction (TAIL-PCR) (Liu et al. (1995) Plant Journal 8: 457-463) using the T-DNA specific primers LB1 (5'-ttc gga acc acc atc aaa cag g-3', SEQ ID NO:25), LB2 (5'-ttg ctg caa ctc tct cag ggc c-3', SEQ ID NO:26), and LB3 (5'-tca gct ggc gcc cgt ctc act-3', SEQ ID NO:27) and the degenerate primer AD3 (5'-wgt gna gwa nca nag a-3').

where W=A/T and N=G/A/T/C, SEQ ID NO:28). The genomic position of the T-DNA copies in line 5 is determined to be linked to BAC F22L4 on chromosome 1.

Example 5: Analysis of the expression of the silenced 8Z-2 transgene in Arabidopsis line S11.13-34

The line 8Z-2 (see Example 4 above) is crossed with the line S11.13-34 (see Example 2 above) and the resultant F1 generation plants are allowed to self-fertilize to obtain the F2 generation. Approximately 60 F2 plants are grown and tested for a presence of T-DNA insertion in the RDRD gene derived from the S11.13-34 parental line and for the 35S-GFP T-DNA derived from the 8Z-2 parental line. The presence of the T-DNA insertion in the RDRD gene is demonstrated as described in Example 2. Plants homozygous for this T-DNA insertion are then checked for homozygosity by PCR using the 3' specific primer (SEQ ID NO:19) and 36851TD#3 (5'-gct ccg ccc aca taa ttc aaa caa cac-3', SEQ ID NO:29). These primers span a region of genomic DNA including the insertion site such that only the wild-type copy of DNA results in amplification of a genomic fragment. A similar strategy is used to screen for lines homozygous for the 35S-GFP T-DNA. First, the presence of the 35S-GFP T-DNA is demonstrated by using the T-DNA-specific PCR primer LB1 and the gene-specific PCR primer L22F4F (5'- ttc gaa aac att acc tcc gat c-3', SEQ ID NO:30). Second, plants carrying the 35S-GFP T-DNA are tested for homozygosity by using the gene-specific primers L22F4F and F22L4R (5'-ggc ttt tgc att tgg tat cta cta g-3', SEQ ID NO:31). The plants homozygous for both the S11.13-34 and 8Z-2 transgenes and plants homozygous for the 8Z-2 transgene but with no S11.13-34 transgene are allowed to self fertilize to obtain F3 generation plants. These plants and the parental line 8Z-2 are scored for incidence of PTGS based on GFP fluorescence as described in Example 4. The results summarized in Table 3 show that PTGS of the 35S-GFP transgene is lost in plants with a T-DNA insert interrupting the region encoding a polypeptide comprising an RNase D-related domain.

Table 3: The Incidence of 35S-GFP PTGS in S11.13-34 x 8Z-2 Hybrids

Line tested	Description	Comments	% Plants exhibiting PTGS	Total number of plants scored

Parental 8Z-2	Homozygous for the 35S-GFP transgene. No S11.13-34 T-DNA insert	Gene encoding the RNase D domain related protein is expressed.	90	40
Outcrossed F3	F3 plants derived from the S11.13-34 x 8Z-2 F3 hybrid homozygous for the 35S-GFP transgene and the S11.13-34 T-DNA crossed out	Gene encoding the RNase D domain related protein is expressed.	88	33
Homozygous F3	F3 plants derived from the S11.13-34 x 8Z-2 F3 hybrid homozygous for the 35S-GFP and S11.13-34 T-DNAs	Gene encoding the RNase D domain related protein with T-DNA insertion. mRNA is not expressed.	0	36

Example 6: Expression of RNase D domain related protein mRNAs in *Arabidopsis* line S11.13-34

The accumulation of RDRD gene mRNA is measured by RT-PCR. The primers AtWRN CDS F (5'-atg tca tcg tca aat tgg atc gac g-3', SEQ ID NO:32) and AtWRN-RT_R (5'-cgc tta tca acc tca gta gca gtc ttg-3', SEQ ID NO:33) are designed to amplify a 329 bp fragment spanning a 5' fragment of the coding sequence. The fragment of predicted length is detected in RNA sample prepared from wild-type *Arabidopsis* plants. Neither this predicted fragment nor any other sequences are detected in RNA samples prepared from the S11.13-34 mutant. This indicates that RDRD mRNA is expressed in wild-type plants, but not in the homozygous RDRD mutant S11.13-34.

Example 7: Identification of a cDNA sequence encoding a polypeptide comprising a RNase D related domain

The gene encoding SEQ ID NO:1, which is also known as AT4g13870 located on *Arabidopsis thaliana* chromosome IV contig fragment 37 (ATCHRIV37, GenBank accession AL161537),

encodes a polypeptide comprising a RNase D related domain. The cDNA for this gene is isolated as follows. 5' and 3' RACE primers are designed based on the exon/intron boundaries in the gene model in ATCHRV37. 5' and 3' RACE is performed (GeneRacer kit, Clontech). The resulting amplicons are TA-cloned (Original TA-Cloning kit, Invitrogen) and sequenced. The elucidated cDNA sequence (SEQ ID NO:23) differs from the sequence predicted in the GenBank annotation, thus identifying the actual open reading frame. SEQ ID NO:24 contains the protein sequence predicted from a translation of bases 42 to 905 of this cDNA. Analysis of the cDNA sequence from this gene reveals a high degree of similarity to an *Arabidopsis thaliana* mRNA for an exonuclease named "wrnexo" (GenBank accession AJ404476). The cloned cDNA sequence is nearly identical to that of wrnexo. The two sequences are likely to derive from the same gene. The difference between the two sequences is noted in 9 extra bases, present in the cloned cDNA encoding a polypeptide comprising a RNase D related domain (bases 830 to 838 of SEQ ID NO:23) but absent in wrnexo.

Example 8: Overexpression of a nucleotide sequence encoding a polypeptide comprising a RNase D related domain

A transgenic construct designed to overexpress a polypeptide comprising a RNase D related domain is introduced into a transgenic line comprising a second transgene. A suitable line expresses the second transgene at a high level with no silencing or without complete silencing preferably with less than half the plants showing silencing or with the silenced plants showing silencing to levels greater than 50% of the average levels of all the plants. The transgenic construct is created by expressing the GUS marker gene (GenBank accession S69414), using the strong constitutive ACT2 promoter (GenBank accession U41998), with the CaMV 35S transcriptional terminator (nucleotides 2868 to 2938 in pJG304 (Guyer et al., 1998, Genetics 149:633-639)) in a binary T-DNA vector. This construct is introduced into *Arabidopsis* via *Agrobacterium*-mediated transformation. T2 plants from a single T1 plant expressing high levels of GUS activity are examined for silencing.

These T2 plants, or their progeny, are also transformed with one of two constructs. One construct allows overexpression of the RNase D related domain coding sequence (bases 42 to 908 of SEQ ID NO:23) with a strong promoter and a transcriptional terminator different from those used in the construct described above. The other construct is a control that is essentially the same as the RNase D related domain construct, except that in place of a RNase D related domain protein, a marker gene, such as luciferase or GFP is overexpressed or no gene is overexpressed. These two binary vector constructs have a selectable marker that differs from

the GUS construct, so that they can be used to superinfect with a second T-DNA construct. When each of these constructs is transformed into the T2 plants described above, the level of GUS expression is determined for the doubly-transformed T1 progeny. Those T1 plants overexpressing the RNase D related domain protein are expected to have lower levels of GUS expression due to increased silencing. If a difference is not detected in those T1 plants, lines homozygous for the RNase D related domain overexpression construct can be produced in the T2 generation and examined.

Alternatively, a nucleotide sequence set forth in any one of SEQ ID NO:3, SEQ ID NO:5, SEQ NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:35 or SEQ ID NO:37 is included in a construct as described above and is used for overexpression of a polypeptide comprising a 3'-5' exonuclease domain.

Example 9: Complementation of the PTGS deficiency of line S11.13-34 by overexpression of a nucleotide sequence encoding RDRD confirms that polypeptide comprising an RNase D related domain is required for PTGS

A construction designed to overexpress a polypeptide comprising an RNase D related domain introduced into *Arabidopsis* plants as described in Example 4. The coding sequence comprising an RNase D related domain is amplified by RT-PCR from RNA prepared from *Arabidopsis* leaves using the primers AtWRN CDS F (SEQ ID NO:32) and AtWRN CDS R (5'-tta tga gcc act gac a atc agg-3') (SEQ ID NO:34). This RDRD coding sequence was placed under the regulation of the strong, constitutive UBQ3 gene promoter (BAC F15A17, GenBank accession AL163002) in binary vector pCAMBIA-1380 (GenBank accession AF234301). The resultant RDRD expression vector is designated pRDP1.

For complementation studies, RDP1 transformants obtained by transformation of wild-type *Arabidopsis* plants with the vector pRDP1 are allowed to self-fertilize. The resultant T1 generation plants are tested for the hygromycin resistance phenotype to detect the presence of the RDP1 T-DNA. The hygromycin-resistant plants are then allowed to self-fertilize and the resultant T2 generation is scored for hygromycin-resistance to identify homozygous transformants with T-DNA inserts at a single locus. Homozygous RDP1 plants are crossed with the double-homozygous F3 generation 8Z-2 S11.13-34 transformants described in Example 5 to obtain the F1 generation. F1 plants are allowed to self-fertilize and the resultant F2 generation plants resistant to both kanamycin and hygromycin are allowed to self-fertilize to obtain the F3 generation. The F3 plants are screened for antibiotic resistance to identify plants homozygous for the RDP1, 35S-GFP, and S11.13-34 T-DNAs. These triple-homozygous lines, the

homozygous parent 8Z-2 S11.13-34 line, and the homozygous 8Z-2 35S-GFP line are screened for PTGS of the 35S-GFP transgene.

Example 10: Overexpression of RDRD in *Arabidopsis* promotes PTGS

Homozygous RDP1 transformants (see Example 9) are crossed with PTGS lines 8Z-2 and 5 to obtain F3 generation plants homozygous for both the 35S-GFP and RDP1 transgenes by using the methods described in Example 9. Assaying these homozygous lines for GFP expression as described in Example 4 shows that there is an increase in the fraction of plants exhibiting PTGS among the 8Z-2 RDP and 5 RDP plants compared with the original 8Z-2 and 5 lines, respectively. Therefore, overexpression of additional copies of a RDRD transgene promotes PTGS of lines showing less than a 100% incidence of PTGS.

The present invention is not to be limited in scope by the specific embodiments described which are intended as single illustrations of individual aspects of the invention, and any constructs, nucleic acid sequences or transformed plants which are functionally equivalent are within the scope of this invention. Indeed, various modifications of the invention in addition to those shown herein will become apparent to those skilled in the art from the foregoing description and accompanying drawings. Such modifications are intended to fall within the scope of the appended claims.

Various patents and references are cited within the present specification, all of which are incorporated by reference in their entireties.

What is claimed is:

1. An isolated nucleic acid molecule comprising a nucleotide sequence encoding a polypeptide comprising a 3'-5' exonuclease domain, and wherein said polypeptide is identical or substantially similar to an amino acid sequence of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:22, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:24, SEQ ID NO:36 or SEQ ID NO:38.
2. The isolated nucleic acid molecule according to claim 1, wherein said 3'-5' exonuclease domain is an RNase D related domain.
3. The isolated nucleic acid molecule according to claim 1 or 2, wherein said polypeptide has 3'-5' exonuclease activity.
4. An isolated nucleic acid molecule comprising a nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, or SEQ ID NO:13, SEQ ID NO:23 SEQ ID NO:35 or SEQ ID NO:37.
5. An isolated nucleic acid molecule comprising a nucleotide sequence of SEQ ID NO:23 or complements thereof.
6. The isolated nucleic acid molecule according to any one of claims 1 to 5, wherein the nucleotide sequence is obtained or derived from a plant.
7. An isolated recombinant nucleic acid molecule comprising a nucleic acid molecule of any one of claims 1 to 6 or complement thereof operatively linked to a promoter functional in a cell.
8. The isolated recombinant nucleic acid molecule according to claim 7, wherein the promoter is functional in a plant cell.
9. The isolated recombinant nucleic acid molecule according to claim 7, wherein the nucleic acid molecule of any one of claims 1 to 6 is in sense orientation.

10. The isolated recombinant nucleic acid molecule according to claim 7, wherein the nucleic acid molecule of any one of claims 1 to 6 is in anti-sense orientation.
11. An expression cassette comprising a nucleic acid molecule comprising a nucleotide sequence encoding a polypeptide comprising a 3'-5' exonuclease domain, and wherein said polypeptide is identical or substantially similar to an amino acid sequence of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:22, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:24, SEQ ID NO:36 or SEQ ID NO:38, a promoter, and a terminator.
12. The expression cassette according to claim 11, wherein said promoter is a constitutive promoter, an inducible promoter, a tissue-specific promoter or a developmentally-regulated promoter.
13. A vector comprising the nucleic acid molecule of any one of claims 1 to 6.
14. An isolated and substantially purified polypeptide comprising the amino acid sequence SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:22, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:24, SEQ ID NO:36 or SEQ ID NO:38.
15. An isolated and substantially purified polypeptide comprising the amino acid sequence of SEQ ID NO:24.
16. A cell comprising the nucleic acid molecule of any one of claims 1 to 6.
17. A cell comprising the expression cassette according to claim 11 or 12.
18. The cell according to claim 17, wherein the cell is a plant cell.
19. The cell of claim 17 or 18, wherein the nucleic acid molecule comprising the nucleotide sequence encoding a polypeptide comprising a 3'-5' exonuclease domain is expressed in the cell.

20. The plant cell according to claim 18 or 19, further comprising an endogenous nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37.
21. The plant cell according to claim 20, wherein the expression of said endogenous nucleotide sequence in said plant cell is altered.
22. The plant cell according to claim 19, wherein said plant cell further comprises a nucleotide sequence of interest, wherein the expression of the nucleotide sequence of interest in the plant cell is altered as compared to the expression of the nucleotide sequence of interest in a plant cell lacking the expression cassette.
23. A plant cell comprising an endogenous nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37, and wherein said plant cell comprises a mutation in said endogenous nucleotide sequence, or in a regulatory region thereof.
24. A plant cell of claim 23, wherein the mutation is a deletion or rearrangement.
25. A plant cell of claim 23, wherein the mutation is a point mutation.
26. The plant cell of claim 23, wherein the mutation is due to an insertion of a nucleic acid molecule.
27. The plant cell according to claim 23, wherein the insertion of a nucleic acid molecule comprises one T-DNA border region.
28. The plant cell according to claim 23, wherein the insertion comprises a transposable element.

29. The plant cell according to claim 23, wherein the expression of said endogenous nucleotide sequence in said plant cell is reduced.
30. The plant cell according to claim 23, wherein the expression of said endogenous nucleotide sequence in said plant cell is increased.
31. The plant cell according to claim 18, wherein said plant cell further comprises a second expression cassette comprising a nucleic acid molecule of interest, wherein the expression of the nucleic acid molecule of interest in said plant cell is stabilized or increased as compared to the expression of nucleic acid molecule of interest in a plant cell lacking said the first expression cassette.
32. The plant cell according to claim 18, further comprising a second expression cassette comprising a nucleic acid molecule of interest, wherein the expression of said nucleic acid molecule of interest in said plant cell is decreased as compared to the expression of said nucleic acid molecule of interest in a plant cell lacking said the first expression cassette.
33. A plant cell transgenically expressing a sense RNA molecule of a nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37, and an anti-sense RNA molecule of said nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37, wherein said sense and said anti-sense RNA molecules are capable of forming a double-stranded RNA molecule.
34. The plant cell according to claim 33, wherein the expression of an endogenous nucleotide sequence of said plant cell that is identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37 is reduced.
35. The plant cell according to claim 34, further comprising an expression cassette comprising a nucleotide sequence of interest, wherein the expression of said nucleotide

sequence of interest in said plant cell is stabilized or increased as compared to the expression of said nucleotide sequence of interest in a plant cell that is not expressing said sense and said anti-sense RNA molecules.

36. A plant comprising the expression cassette of claim 11 or 12, or progeny thereof, or seeds thereof.
37. A plant comprising the plant cell of any one of claims 18 to 35, or progeny thereof, or seeds thereof.
38. A method for altering the expression in a plant cell or plant of an endogenous nucleotide sequence encoding a polypeptide comprising a 3'-5' exonuclease domain, wherein said polypeptide is identical or substantially similar to SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:22, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:24, SEQ ID NO:36 or SEQ ID NO:38 comprising the step of:
altering the transcription or translation of said endogenous nucleotide sequence in said plant cell or plant.
39. The method according to claim 38, wherein altering the transcription or translation of said endogenous nucleotide sequence in the plant cell or plant comprises the step of:
- a) expressing in said plant cell a nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37, or a portion thereof, in sense orientation; or
 - b) expressing in said plant cell a nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37, or a portion thereof, in anti-sense orientation; or
 - c) expressing in said plant cell a sense RNA of a nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37, or a portion thereof, and an anti-sense RNA of said nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:23,

- SEQ ID NO:35 or SEQ ID NO:37, or a portion thereof, wherein said sense and said anti-sense RNAs are capable of forming a double-stranded RNA molecule; or
- d) expressing in said plant cell a ribozyme capable of specifically cleaving a messenger RNA transcript encoded by a nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37; or
- e) modifying by homologous recombination in said plant cell at least one chromosomal copy of the nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37 or of a regulatory region thereof; or
- f) expressing in said plant cell a zinc finger protein that is capable of binding to a nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37 or to a regulatory region thereof; or
- g) introducing into said plant cell a chimeric oligonucleotide that is capable of modifying at least one chromosomal copy of the nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37 or a regulatory region thereof.
40. A method for altering the expression of an endogenous nucleotide sequence that is identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37 in a plant cell or plant comprising the step of introducing into said plant cell a means for altering the transcription or translation of said endogenous nucleotide sequence in said plant cell.
41. A method for altering the expression of a nucleotide sequence of interest in a plant cell or plant comprising the steps of:
- a) altering the expression in said plant cell or plant of an endogenous nucleotide sequence of said plant cell that is identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37; and

- b) introducing into said plant cell or plant a nucleic acid molecule comprising said nucleotide sequence of interest, wherein the expression of said nucleotide sequence of interest in said plant cell or plant is altered.
42. The method according to claim 41, wherein step a) comprises:
- a) expressing in said plant cell or plant a nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37, or a portion thereof, in sense orientation; or
 - b) expressing in said plant cell or plant a nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37, or a portion thereof, in anti-sense orientation; or
 - c) expressing in said plant cell or plant a sense RNA of a nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37, or a portion thereof, and an anti-sense RNA of said nucleotide sequence substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37 or a portion thereof, wherein said sense and said anti-sense RNAs are capable of forming a double-stranded RNA molecule; or
 - d) expressing in said plant cell or plant a ribozyme capable of specifically cleaving a messenger RNA transcript encoded by a nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37; or
 - e) modifying by homologous recombination in said plant cell or plant at least one chromosomal copy of the nucleotide sequence substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37 or of a regulatory region thereof; or

- f) expressing in said plant cell or plant a zinc finger protein that is capable of binding to a nucleotide sequence substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37 or to a regulatory region thereof; or
 - g) introducing into said plant cell or plant a chimeric oligonucleotide that is capable of modifying at least one chromosomal copy of the nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37 or a regulatory region thereof.
43. A method for altering or stabilizing the expression of a nucleotide sequence of interest in a plant cell or plant comprising the steps of:
- a) obtaining a plant cell comprising a first expression cassette according to claim 11 or 12 or plant thereof; and
 - b) introducing into said plant cell or plant a second nucleic acid molecule comprising said nucleotide sequence of interest, wherein the expression of said nucleotide sequence of interest in said plant cell is altered or stabilized as compared to the expression of said nucleotide sequence of interest in a plant cell or plant lacking said first expression cassette.
44. The method of claim 43, wherein the expression of the nucleotide sequence of interest is increased.
45. The method of claim 43, wherein the expression of the nucleotide sequence of interest is reduced.
46. A method for stabilizing the expression of a nucleotide sequence of interest in a plant cell or plant comprising the steps of:
- a) obtaining a plant cell or plant having altered expression in a plant cell of an endogenous nucleotide sequence of said plant cell or plant that encodes a polypeptide comprising a 3'-5' exonuclease domain, and wherein said polypeptide is identical or substantially similar to an amino acid sequence of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:22, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:24, SEQ ID NO:36 or SEQ ID NO:38; and

- b) introducing into said plant cell or plant a nucleotide sequence of interest, wherein the expression of said nucleotide sequence of interest in said plant cell is stabilized as compared to the expression of said nucleotide sequence of interest in a plant cell or plant lacking said first expression cassette.

47. The method according to claim 46, wherein said endogenous nucleotide sequence is identical or substantially similar to a nucleotide sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37.
48. The method according to claim 46, wherein the expression of said endogenous nucleotide sequence is altered by:
- a) expressing in said plant cell a nucleotide sequence substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37, or a portion thereof, in sense orientation; or
 - b) expressing in said plant cell a nucleotide sequence substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37, or a portion thereof, in anti-sense orientation; or
 - c) expressing in said plant cell a sense RNA of a nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37 or a portion thereof, and an anti-sense RNA of said nucleotide sequence substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37 or a portion thereof, wherein said sense and said anti-sense RNAs are capable of forming a double-stranded RNA molecule; or
 - d) expressing in said plant cell a ribozyme capable of specifically cleaving a messenger RNA transcript encoded by a nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37 or

- e) expressing in said plant cell an aptamer specifically directed to a polypeptide substantially similar to SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:22, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:24, SEQ ID NO:36 or SEQ ID NO:38 or
 - f) modifying by homologous recombination in said plant cell at least one chromosomal copy of the nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37 or of a regulatory region thereof; or
 - g) expressing in said plant cell a zinc finger protein that is capable of binding to a nucleotide sequence substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37 or to a regulatory region thereof; or
 - h) introducing into said plant cell a chimeric oligonucleotide that is capable of modifying at least one chromosomal copy of the nucleotide sequence identical or substantially similar to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:21, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:23, SEQ ID NO:35 or SEQ ID NO:37 or a regulatory region thereof.
49. The method according to claim 48, wherein the expression in a plant cell of said endogenous nucleotide sequence is reduced.
50. A method for identifying a compound capable of interacting with a polypeptide comprising a 3'-5' exonuclease domain comprising:
- a) combining a polypeptide comprising the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:22, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:24, SEQ ID NO:36 or SEQ ID NO:38 or a homolog thereof, and a compound to be tested for the ability to interact with said polypeptide, under conditions conducive to interaction; and
 - b) selecting a compound from step (a) that is capable of interacting with said polypeptide.
51. A compound identifiable by the method of claim 50.

52. A compound identifiable by the method of claim 50, wherein said compound is capable of altering the activity of said polypeptide.

SEQUENCE LISTING

<110> Syngenta Participations AG

<120> Methods of Controlling Gene Expression and Gene Silencing

<130> PB/5-31481A

<140>

<141>

<150> US 60/222202

<151> 2000-08-01

<160> 38

<170> PatentIn Ver. 2.1

<210> 1

<211> 942

<212> DNA

<213> Arabidopsis thaliana

<400> 1

```

atgtcatcgt caaattggat cgacgacgct tttacagagg aagagcttct cgctatcgac 60
gccatcgaag cttcctacaa tttctcccggt tcttcttctt cttcttcctc tgctgctccg 120
accgtacaag ctacaacctc cgtccatggc cacgaggagg atccaaatca aatccccaat 180
aatatccgtc gccaatggcc tcgttccatc acttcttcta catcttataa acgatttcct 240
ctctcccggt gccgagctag gaattttcca gcaatgaggt ttggtggtag gattttgtat 300
agcaagactg ctactgaggt tgataagcga gcaatgcagc ttattaaagt tcttgatacc 360
aagagagatg aatctggaat agcttttgtt ggcttggata ttgagtggag accaagtttt 420
agaaaagggt ttctcccggt gaagggtgcg actgtccaga tatgtgtaga tagtaattat 480
tgtgatgtta tgcataattt tcattctggt atccctcaaa gtctccaaca tcttattgaa 540
gattcaacac ttgtaaagggt aggtattgga attgatgggt actctgtgaa gcttttccat 600
gactatggag ttagtatcaa agatgttgag gatctttcag atttagccaa ccaaaaaatt 660
ggtggagata aaaaatgggg ccttgcttca ctaactgaga cacttgtttg caaagagctc 720
ctgaagccaa acagaatcag gcttgggaac tgggagtttt atcctctgtc aaagcagcag 780
ttacaatacg cagcaacgga tgcttatgct tcatggcatc tttacaagggt aacaacaacg 840
aaaaaccatc ttctcacact caacgacctt gaagcaaaaa tctcacatcg ttctaattat 900
aatactgtta cttgtcgaaa acctggaggt tatcttcggt ga 942

```

<210> 2

<211> 313

<212> PRT

<213> Arabidopsis thaliana

<400> 2

```

Met Ser Ser Ser Asn Trp Ile Asp Asp Ala Phe Thr Glu Glu Glu Leu
 1             5             10             15

Leu Ala Ile Asp Ala Ile Glu Ala Ser Tyr Asn Phe Ser Arg Ser Ser
          20             25             30

Ser Ser Ser Ser Ser Ala Ala Pro Thr Val Gln Ala Thr Thr Ser Val
 35             40             45

```

His Gly His Glu Glu Asp Pro Asn Gln Ile Pro Asn Asn Ile Arg Arg
 50 55 60
 Gln Leu Pro Arg Ser Ile Thr Ser Ser Thr Ser Tyr Lys Arg Phe Pro
 65 70 75 80
 Leu Ser Arg Cys Arg Ala Arg Asn Phe Pro Ala Met Arg Phe Gly Gly
 85 90 95
 Arg Ile Leu Tyr Ser Lys Thr Ala Thr Glu Val Asp Lys Arg Ala Met
 100 105 110
 Gln Leu Ile Lys Val Leu Asp Thr Lys Arg Asp Glu Ser Gly Ile Ala
 115 120 125
 Phe Val Gly Leu Asp Ile Glu Trp Arg Pro Ser Phe Arg Lys Gly Val
 130 135 140
 Leu Pro Gly Lys Val Ala Thr Val Gln Ile Cys Val Asp Ser Asn Tyr
 145 150 155 160
 Cys Asp Val Met His Ile Phe His Ser Gly Ile Pro Gln Ser Leu Gln
 165 170 175
 His Leu Ile Glu Asp Ser Thr Leu Val Lys Val Gly Ile Gly Ile Asp
 180 185 190
 Gly Asp Ser Val Lys Leu Phe His Asp Tyr Gly Val Ser Ile Lys Asp
 195 200 205
 Val Glu Asp Leu Ser Asp Leu Ala Asn Gln Lys Ile Gly Gly Asp Lys
 210 215 220
 Lys Trp Gly Leu Ala Ser Leu Thr Glu Thr Leu Val Cys Lys Glu Leu
 225 230 235 240
 Leu Lys Pro Asn Arg Ile Arg Leu Gly Asn Trp Glu Phe Tyr Pro Leu
 245 250 255
 Ser Lys Gln Gln Leu Gln Tyr Ala Ala Thr Asp Ala Tyr Ala Ser Trp
 260 265 270
 His Leu Tyr Lys Val Thr Thr Thr Lys Asn His Leu Leu Thr Leu Asn
 275 280 285
 Asp Leu Glu Ala Lys Ile Ser His Arg Ser Asn Tyr Asn Thr Val Thr
 290 295 300
 Cys Arg Lys Pro Gly Gly Tyr Leu Arg
 305 310

<210> 3

<211> 1929

<212> DNA

<213> *Arabidopsis thaliana*

<400> 3

```

atgagatttg atgatcccat ggatgagttc aagaggaatc gaaagatgga ggaagattcg 60
aagaaggtaa tcgatgtgaa agtggctgag agtgataagg gattcgcgaa atttggcaag 120
gcagaggttc cgtttcatat accgacgtta acgaagcctc aagaggagta taagattttg 180
gtagacaatg ctaataatcc ttttgagcat gttttgttgg agaagagtga agacgggtctt 240
cggttcattc atccactgga ggaactatct gtgatggact ttgttgatag aaatctaagt 300
gagatgagac ctgttaagcc tctcccattg gaagagactc cattcaagct agttgaagaa 360
gtcaaagatc ttgaggactt agctgctgca ttgcaaagtg ttgaagagtt tgctgtcgat 420
ctggagcata atcagtatag aacttttcaa ggattaacat gcttgatgca aatctctact 480
agaaccgagg attatattgt tgatatattc aagctttggg atcacattgg tccttatcta 540
agggaaactct tcaaagaccc taaaaagaaa aaggtaatcc atggagcaga tcgagatatt 600
at ttggcttc aacgggactt tggcatttat gtctgcaatc tttttgacac aggacaggct 660
tcaaggggtgc taaagctgga gagaaatagt ctggaatttc ttctgaagca ttattgtgga 720
gttgctgcaa acaaagaata ccaaaaagca gactggagaa taagaccctt tccagatgta 780
atgaaaagat atgctagaga agatacacat tatcttttgt acatttatga tgtaatgcga 840
atggagttgc acacaatggc aaaggaagat gagcaatctg actctcctct ggtagagggtg 900
tacaagcgca gttatgacgt gtgcatgcaa ctatatgaaa aagagctttg gactagggat 960
tcatatcttc acgttttatgg ggttcagaca ggtaatctca atgcggttca actttccatt 1020
gttgcgctgc aggggctttg tgaatggcgg gatcggattg cacgcgcaga tgatgagagc 1080
accggttatg tattgccaaa taaaactctt tttgacatag ccaaggagat gccaatgtgt 1140
gttgccagct tgcgccgttt gttgaagtca aagcttcctt acctcgagcg taattttgac 1200
gcagtgatca gtgtcatcag acgatcaatg caaatgcag cggcattcga gccagttggt 1260
caatctttga aagataggcg tcctgaaaca gtgggtgaaa tgaatataga acctagatt 1320
gagaaaacag acacaggagc ttcagcgtct tctctgagtc tggagaaggt ttgtgtggat 1380
gattcaaaga aacaaagcag tggttttgga gttttgccgt taaagaggaa gttggaaagt 1440
gacaaaacgg tgggtgaaaa gaatatcgaa cctaagattg agaaaacagg cacagaagct 1500
tcagcttctt ctctgagttc gaagaagggt tgtgtggatg attcaaagaa acaaagcagt 1560
ggttttggag ttttgctgtc aaagaggaag tttgaaagtg ataacaagaa gttgcaggta 1620
aaagaagagg tcaaagtgtc caagtccaag ccagataagg taatcatagt ggtggatgat 1680
gatgatgatg atgatgatga tgagtcttat gaacagagca cgaaagccgc tgatgctttg 1740
gacagagttt cggaaacgcc ttcgaaggga tcaccatcgt tgactcaaaa gccgaagaca 1800
tgtaatacag aggttattgt gtttagcgat gatgatgact cggaaagcag agaagatgaa 1860
gacatgcgta ggagaagtga gaaacatagg agattcatga atatgaaacg tggctttctt 1920
aacatttag 1929

```

<210> 4

<211> 642

<212> PRT

<213> *Arabidopsis thaliana*

<400> 4

```

Met Arg Phe Asp Asp Pro Met Asp Glu Phe Lys Arg Asn Arg Lys Met
 1             5             10            15

Glu Glu Asp Ser Lys Lys Val Ile Asp Val Lys Val Ala Glu Ser Asp
      20             25            30

Lys Gly Phe Ala Lys Phe Gly Lys Ala Glu Val Pro Phe His Ile Pro
      35             40            45

Thr Leu Thr Lys Pro Gln Glu Glu Tyr Lys Ile Leu Val Asp Asn Ala
      50             55            60

```

Asn Asn Pro Phe Glu His Val Leu Leu Glu Lys Ser Glu Asp Gly Leu
 65 70 75 80
 Arg Phe Ile His Pro Leu Glu Glu Leu Ser Val Met Asp Phe Val Asp
 85 90 95
 Arg Asn Leu Ser Glu Met Arg Pro Val Lys Pro Leu Pro Leu Glu Glu
 100 105 110
 Thr Pro Phe Lys Leu Val Glu Glu Val Lys Asp Leu Glu Asp Leu Ala
 115 120 125
 Ala Ala Leu Gln Ser Val Glu Glu Phe Ala Val Asp Leu Glu His Asn
 130 135 140
 Gln Tyr Arg Thr Phe Gln Gly Leu Thr Cys Leu Met Gln Ile Ser Thr
 145 150 155 160
 Arg Thr Glu Asp Tyr Ile Val Asp Ile Phe Lys Leu Trp Asp His Ile
 165 170 175
 Gly Pro Tyr Leu Arg Glu Leu Phe Lys Asp Pro Lys Lys Lys Lys Val
 180 185 190
 Ile His Gly Ala Asp Arg Asp Ile Ile Trp Leu Gln Arg Asp Phe Gly
 195 200 205
 Ile Tyr Val Cys Asn Leu Phe Asp Thr Gly Gln Ala Ser Arg Val Leu
 210 215 220
 Lys Leu Glu Arg Asn Ser Leu Glu Phe Leu Leu Lys His Tyr Cys Gly
 225 230 235 240
 Val Ala Ala Asn Lys Glu Tyr Gln Lys Ala Asp Trp Arg Ile Arg Pro
 245 250 255
 Leu Pro Asp Val Met Lys Arg Tyr Ala Arg Glu Asp Thr His Tyr Leu
 260 265 270
 Leu Tyr Ile Tyr Asp Val Met Arg Met Glu Leu His Thr Met Ala Lys
 275 280 285
 Glu Asp Glu Gln Ser Asp Ser Pro Leu Val Glu Val Tyr Lys Arg Ser
 290 295 300
 Tyr Asp Val Cys Met Gln Leu Tyr Glu Lys Glu Leu Trp Thr Arg Asp
 305 310 315 320
 Ser Tyr Leu His Val Tyr Gly Val Gln Thr Gly Asn Leu Asn Ala Val
 325 330 335
 Gln Leu Ser Ile Val Ala Leu Gln Gly Leu Cys Glu Trp Arg Asp Arg
 340 345 350
 Ile Ala Arg Ala Asp Asp Glu Ser Thr Gly Tyr Val Leu Pro Asn Lys
 355 360 365

Thr Leu Phe Asp Ile Ala Lys Glu Met Pro Ile Val Val Ala Gln Leu
 370 375 380
 Arg Arg Leu Leu Lys Ser Lys Leu Pro Tyr Leu Glu Arg Asn Phe Asp
 385 390 395 400
 Ala Val Ile Ser Val Ile Arg Arg Ser Met Gln Asn Ala Ala Ala Phe
 405 410 415
 Glu Pro Val Val Gln Ser Leu Lys Asp Arg Arg Pro Glu Thr Val Val
 420 425 430
 Glu Met Asn Ile Glu Pro Lys Ile Glu Lys Thr Asp Thr Gly Ala Ser
 435 440 445
 Ala Ser Ser Leu Ser Leu Glu Lys Val Cys Val Asp Asp Ser Lys Lys
 450 455 460
 Gln Ser Ser Gly Phe Gly Val Leu Pro Leu Lys Arg Lys Leu Glu Ser
 465 470 475 480
 Asp Lys Thr Val Val Glu Lys Asn Ile Glu Pro Lys Ile Glu Lys Thr
 485 490 495
 Gly Thr Glu Ala Ser Ala Ser Ser Leu Ser Ser Lys Lys Val Cys Val
 500 505 510
 Asp Asp Ser Lys Lys Gln Ser Ser Gly Phe Gly Val Leu Leu Ser Lys
 515 520 525
 Arg Lys Phe Glu Ser Asp Asn Lys Lys Leu Gln Val Lys Glu Glu Val
 530 535 540
 Lys Val Ser Lys Ser Lys Pro Asp Lys Val Ile Ile Val Val Asp Asp
 545 550 555 560
 Asp Asp Asp Asp Asp Asp Asp Glu Ser Tyr Glu Gln Ser Thr Lys Ala
 565 570 575
 Ala Asp Ala Leu Asp Arg Val Ser Glu Thr Pro Ser Lys Gly Ser Pro
 580 585 590
 Ser Leu Thr Gln Lys Pro Lys Thr Cys Asn Thr Glu Val Ile Val Leu
 595 600 605
 Asp Asp Asp Asp Asp Ser Glu Ser Arg Glu Asp Glu Asp Met Arg Arg
 610 615 620
 Arg Ser Glu Lys His Arg Arg Phe Met Asn Met Lys Arg Gly Phe Leu
 625 630 635 640
 Asn Ile

<210> 5
 <211> 714
 <212> DNA
 <213> *Arabidopsis thaliana*

<400> 5
 atgaatttgc attttggattt ttggtgtttt atatttgaaa ctaatgcaga gaaaccttcg 60
 aatgggtcatc catatgaaac tgagatcact gttttgtag agaatcctca gattgagttt 120
 ggatttttga gaggagagtg ttcattggaa atgagtgtt catatgtgtg gggtgagaca 180
 gagtgcgagt taaaggaact tgcagaaata ttagcaaaag aacaagttt tgcgggttgac 240
 actgagcagc atagtttgcg gtcgtttctt ggtttactg ctctaattca gattttctaca 300
 catgaggaag acttttttggg ggacacaatt gcgttacatg atgtaatgag tattcttcgt 360
 cctgttttct ctgatcctaa tatttgaag gtgtttcacg gggctgacaa cgatgttatc 420
 tggcttcaaa gagacttcca tatatatgtt gttaatatgt ttgatactgc caaggcatgt 480
 gaagtgttgt caaagcctca acgatcactg gcatacttac ttgagacagt atgtggagtg 540
 gctactaaca aattgctgca gcgtgaagat tggagacagc gtcctctgtc cgaagagatg 600
 gtgcgatatg ctagaacaga tgcacactat ctgctttata ttgcagatag tttgacaact 660
 gaactcaaac aattagccac tggtaggcac ctttgctatg gagaaacatt ttag 714

<210> 6
 <211> 237
 <212> PRT
 <213> *Arabidopsis thaliana*

<400> 6
 Met Asn Leu His Phe Asp Phe Trp Cys Phe Ile Phe Glu Thr Asn Ala
 1 5 10 15
 Glu Lys Pro Ser Asn Gly His Pro Tyr Glu Thr Glu Ile Thr Val Leu
 20 25 30
 Leu Glu Asn Pro Gln Ile Glu Phe Gly Phe Leu Arg Gly Glu Cys Ser
 35 40 45
 Leu Glu Met Ser Asp Ser Tyr Val Trp Val Glu Thr Glu Ser Gln Leu
 50 55 60
 Lys Glu Leu Ala Glu Ile Leu Ala Lys Glu Gln Val Phe Ala Val Asp
 65 70 75 80
 Thr Glu Gln His Ser Leu Arg Ser Phe Leu Gly Phe Thr Ala Leu Ile
 85 90 95
 Gln Ile Ser Thr His Glu Glu Asp Phe Leu Val Asp Thr Ile Ala Leu
 100 105 110
 His Asp Val Met Ser Ile Leu Arg Pro Val Phe Ser Asp Pro Asn Ile
 115 120 125
 Cys Lys Val Phe His Gly Ala Asp Asn Asp Val Ile Trp Leu Gln Arg
 130 135 140
 Asp Phe His Ile Tyr Val Val Asn Met Phe Asp Thr Ala Lys Ala Cys
 145 150 155 160

Glu Val Leu Ser Lys Pro Gln Arg Ser Leu Ala Tyr Leu Leu Glu Thr
 165 170 175

Val Cys Gly Val Ala Thr Asn Lys Leu Leu Gln Arg Glu Asp Trp Arg
 180 185 190

Gln Arg Pro Leu Ser Glu Glu Met Val Arg Tyr Ala Arg Thr Asp Ala
 195 200 205

His Tyr Leu Leu Tyr Ile Ala Asp Ser Leu Thr Thr Glu Leu Lys Gln
 210 215 220

Leu Ala Thr Gly Arg His Leu Cys Tyr Gly Glu Thr Phe
 225 230 235

<210> 7

<211> 849

<212> DNA

<213> Arabidopsis thaliana

<400> 7

```

atgcagattg cattctctaa tgcaatatac ttggttgatg tcatcgaagg tggagaggtg 60
attatgaaag cgtgtaagcc tgcactcgag tctaattaca tcacgaaagt tattcacgat 120
tgcaagcgtg acagtgaggc tctatacttc cagtttgagg taagattgca caatgttggtg 180
gacactcaga ttgcttattc tctgattgaa gaacaagaag ggcggaggag acctctagat 240
gattacatat cgtttgtttc actccttgct gatccacgtt actgcggtat atcctatgaa 300
gagaaagaag aagttcgagt tctcatgcgc caggacccaa agttttggac atacaggcct 360
atgactgagc tcatgatccg cgcagctgct gatgatgtcc gcttccttct gtatctctat 420
cacaaaatga tgggaaagct aatcagcgg tcaactatggc atcttgcaat tctgtgtgct 480
ttgtactgtc ggtgtctctg ctgcatgaat gatgctgatt ttgctgattg gccaacggtt 540
cctccaattc cagttttcct cgtaagggtc gtatatgctg tagagacaaa gaaaaaaga 600
cgggtgacat tagcttcgat tgggttactg attgtagttg gacttttaaa tgtggcagat 660
aacctgaagt cagaagatca atgtcttgaa gaagagatcc tgtcagtgct tgatgttcca 720
ccaggaaaga tgggacgtgt gattggaagg aaaggagcat cgatcctcgc cattaaggaa 780
gcttgcaacg cggaaattct aattggaggg gcaaagggtc cacctgataa ggtagtctt 840
attccatag                                     849

```

<210> 8

<211> 282

<212> PRT

<213> Arabidopsis thaliana

<400> 8

Met Gln Ile Ala Phe Ser Asn Ala Ile Tyr Leu Val Asp Val Ile Glu
 1 5 10 15

Gly Gly Glu Val Ile Met Lys Ala Cys Lys Pro Ala Leu Glu Ser Asn
 20 25 30

Tyr Ile Thr Lys Val Ile His Asp Cys Lys Arg Asp Ser Glu Ala Leu
 35 40 45

Tyr Phe Gln Phe Gly Ile Arg Leu His Asn Val Val Asp Thr Gln Ile
 50 55 60

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

<210> 9

<211> 720

<212> DNA

<213> Arabidopsis thaliana

<400> 9

atggctagga tcagaagaag aatccaaaag cgccatatcc acgaaaaccg ctacatcgat 60
 ttctttggag aacgtttgat cgtcacggtc actcatacta cctcaaccat ccgccgttgg 120
 attcatagca tccgtttctt cagccgtctt cgctcctcac accctctcgt tgttggactc 180
 gacgtccaat ggacacccgg tggttccgat cctccaccgg atattctcca actatgtgtt 240
 ggtaaccgct gtctcatcat ccagttgtct cactgtaaac gcattcctga ggtccttcga 300
 agtttcttgg aagatgagac aatcactttt gtcggcgtct ggaacagcca agaccagggc 360

aagctcgaaa gattccgcca tcagttggag atatggagac ttctagacat aaggcactat 420
 ctgcctacga ggctcctcaa tagttcgttt gagaagattg tagaggagtg tttgggggtac 480
 aagggagtga ggaaagataa ggagatatgt atgagtaatt ggggtgctcg tagcctttcc 540
 catgatcaga ttgttcaggc gtcagatgat gtctatgttt gctgcaagct cgggtgtaaag 600
 gaatgtatct ggaaagagcg ctgcaatgtt aaagaacgta tctggaaaga gagctcgaat 660
 gttaaggaac atatctggaa agagagctcg aaactttatt ttgttgggggt atgtttctga 720

<210> 10

<211> 239

<212> PRT

<213> Arabidopsis thaliana

<400> 10

Met Ala Arg Ile Arg Arg Arg Ile Gln Lys Arg His Ile His Glu Asn
 1 5 10 15

Arg Tyr Ile Asp Phe Phe Gly Glu Arg Leu Ile Val Thr Val Thr His
 20 25 30

Thr Thr Ser Thr Ile Arg Arg Trp Ile His Ser Ile Arg Phe Phe Ser
 35 40 45

Arg Leu Arg Ser Ser His Pro Leu Val Val Gly Leu Asp Val Gln Trp
 50 55 60

Thr Pro Gly Gly Ser Asp Pro Pro Pro Asp Ile Leu Gln Leu Cys Val
 65 70 75 80

Gly Asn Arg Cys Leu Ile Ile Gln Leu Ser His Cys Lys Arg Ile Pro
 85 90 95

Glu Val Leu Arg Ser Phe Leu Glu Asp Glu Thr Ile Thr Phe Val Gly
 100 105 110

Val Trp Asn Ser Gln Asp Gln Gly Lys Leu Glu Arg Phe Arg His Gln
 115 120 125

Leu Glu Ile Trp Arg Leu Leu Asp Ile Arg His Tyr Leu Pro Thr Arg
 130 135 140

Leu Leu Asn Ser Ser Phe Glu Lys Ile Val Glu Glu Cys Leu Gly Tyr
 145 150 155 160

Lys Gly Val Arg Lys Asp Lys Glu Ile Cys Met Ser Asn Trp Gly Ala
 165 170 175

Arg Ser Leu Ser His Asp Gln Ile Val Gln Ala Ser Asp Asp Val Tyr
 180 185 190

Val Cys Cys Lys Leu Gly Val Lys Glu Cys Ile Trp Lys Glu Arg Ser
 195 200 205

Asn Val Lys Glu Arg Ile Trp Lys Glu Ser Ser Asn Val Lys Glu His
 210 215 220

Ile Trp Lys Glu Ser Ser Lys Leu Tyr Phe Val Gly Val Cys Phe
 225 230 235

<210> 11
 <211> 654
 <212> DNA
 <213> Arabidopsis thaliana

<400> 11
 atgaagagag gtatcaaaca tctatgtttc aatggcttca cgggctactc atcacttcat 60
 catcattatc atgaacacca cgtcgacttc tttggagaac gtttgatcgt cacagtcact 120
 catactccct cagtgatacg tcgatggatc cacagtatcc gcttcgtcag ccgtcttcgc 180
 ttatcacacc ctctagttgt cggacttggc gttcaatgga caccocgtgg ttccgatcct 240
 ccaccggata ttctccaact atgtgttggc actcgtctgc tcatcattca gttgtctcac 300
 tgtaagtacg tccccgacgt ccttagaagt ttcttgggaag atcagacaat cacttttgtc 360
 ggcgtatgga acagccaaga caaggacaag ctcgagagat tccaccatca gttggatata 420
 tggagacttg tccacataag gcactatctc catccgttgc tcttgagtag ctcgtttgag 480
 acgattgtga aggtgtatatt ggggcatgaa ggagtgcga aagataagga gttatgtatg 540
 agtaattggg gtgctcgtag cctctctcat gatcagatag tacaagcgtc acatgatgtc 600
 tatgtttgct gcaagctcgg tgtaaggaa cgtctctgga aaatgggagc ttaa 654

<210> 12
 <211> 217
 <212> PRT
 <213> Arabidopsis thaliana

<400> 12
 Met Lys Arg Gly Ile Lys His Leu Cys Phe Asn Gly Phe Thr Gly Tyr
 1 5 10 15
 Ser Ser Leu His His His Tyr His Glu His His Val Asp Phe Phe Gly
 20 25 30
 Glu Arg Leu Ile Val Thr Val Thr His Thr Pro Ser Val Ile Arg Arg
 35 40 45
 Trp Ile His Ser Ile Arg Phe Val Ser Arg Leu Arg Leu Ser His Pro
 50 55 60
 Leu Val Val Gly Leu Gly Val Gln Trp Thr Pro Arg Gly Ser Asp Pro
 65 70 75 80
 Pro Pro Asp Ile Leu Gln Leu Cys Val Gly Thr Arg Cys Leu Ile Ile
 85 90 95
 Gln Leu Ser His Cys Lys Tyr Val Pro Asp Val Leu Arg Ser Phe Leu
 100 105 110
 Glu Asp Gln Thr Ile Thr Phe Val Gly Val Trp Asn Ser Gln Asp Lys
 115 120 125
 Asp Lys Leu Glu Arg Phe His His Gln Leu Asp Ile Trp Arg Leu Val
 130 135 140

His Ile Arg His Tyr Leu His Pro Leu Leu Leu Ser Ser Ser Phe Glu
145 150 155 160

Thr Ile Val Lys Val Tyr Leu Gly His Glu Gly Val Thr Lys Asp Lys
165 170 175

Glu Leu Cys Met Ser Asn Trp Gly Ala Arg Ser Leu Ser His Asp Gln
180 185 190

Ile Val Gln Ala Ser His Asp Val Tyr Val Cys Cys Lys Leu Gly Val
195 200 205

Lys Glu Arg Leu Trp Lys Met Gly Ala
210 215

<210> 13

<211> 261

<212> DNA

<213> Arabidopsis thaliana

<400> 13

atgatcaagt cgatcgagag ctttattgct cgttatgttt tccaagctac attatacaca 60
atctggtgcg aacgaaatgg cagaaggcat ggagagagag aggccgaacc agccgctctt 120
ctaataaaaat tgattgataa gcaagccatc tttcctcaat cagagcaact gatgatctac 180
gctatgacaa gagcctccaa atctggttg catcgagtgg ttgaaatcaa accatctata 240
gttcatgtgt ttaaccttta a 261

<210> 14

<211> 86

<212> PRT

<213> Arabidopsis thaliana

<400> 14

Met Ile Lys Ser Ile Glu Ser Phe Ile Ala Arg Tyr Val Phe Gln Ala
1 5 10 15

Thr Leu Tyr Thr Ile Trp Cys Glu Arg Asn Gly Arg Arg His Gly Glu
20 25 30

Arg Glu Ala Glu Pro Ala Ala Leu Leu Ile Lys Leu Ile Asp Lys Gln
35 40 45

Ala Ile Phe Pro Gln Ser Glu Gln Leu Met Ile Tyr Ala Met Thr Arg
50 55 60

Ala Ser Lys Ser Gly Leu His Arg Val Val Glu Ile Lys Pro Ser Ile
65 70 75 80

Val His Val Phe Asn Leu
85

<210> 15

<211> 2733

<212> DNA

<213> C. elegans

<400> 15

```

atggaagaag aaccgtacaa aagaaagcta acgaaagccg agaaaaaggc aaaatatcga 60
acagactacg ccgaaccggt gaagtcacgc cgggaagttc tgaaagcaat tatgaatgga 120
cccagatctg agcgggagag aaaagttcga gccaaaaatc gagaattttt caacgaggac 180
tatagatcag gagtcaacat ctacggaatg gctgtggata tgatgaaagc gatgccggat 240
agaggaaaaa catcggggaca aagtttggcg gtttggatat tggaggattt tggagtttgg 300
ttaaaagagt cgggacagga gacggagctc agacagaaat atctgactgg aactattcaa 360
ataaacgcct tagatgtgtg cacaattgga caaaagcagc ttctcagtga aatcttcgat 420
atcaccaaag agaaattcac tgaggacatt acacagttgc tagatgctgc catcaagaaa 480
caagacttct ccgttgctgc agatatggct attcagtaca atcttctacg ggatcatcat 540
ttcgaacatc ttgttcttcc attaatgcta tctggcaaag atcaaacggc ttataaattg 600
ataagtaaca atgagaggat gcaacagcag cttgtagagt tttttgatcg aatggttgga 660
atctcagtgg ttgccgttga agagatgctg aaaccctaca aggaaaccaa aattatgacg 720
attcctatgg agaaattgac gggaaaaacg ttggacaaac tcatttcgac gattatcaac 780
aaaaatactc acgaatacaa cttctccagg gaattgtcga agttcgccaa aaaccactca 840
cagaatggga atctgaaggc attgaagttt aatacagtg aacgatacga gaagggaaaa 900
tccgatgaca actatttcca gcataatggtt gaaactttta ccaaagccga agatgttcgt 960
gaacctatit ttgtttactt gtggagctca aatgacaccg agaaacaaat agatgccatc 1020
tgctttgcta tctacttagg aatcgctagt tccagcagct atcaactgcc gaatgttatg 1080
agggacttct ttcgacaacc tgattcgaag cttagagaag caaaagaact tctagtgaga 1140
agaaaaacac tgcaagttcc tctaaatggc gaacaattat tctgatttga gaatgagcga 1200
agaactcaaa tccacatggt gaaaactgaa tctgagatga attacttatg ttccgagatc 1260
aatcactaa ggcagcagcc agcacctgtt tacgttggat tctgattctga gtggaaaccg 1320
tcgaatctta cagctgttca tgactcgaaa attgctatca tacaattgtt cttcaaaaat 1380
tgtgtatggc ttgtggattg cgtagaatta gaaaaggcaa atatggcaga tgactggtgg 1440
caaaagttcg catctcgatt gttcggagat tctcctgtaa aagtcgtagg atttgatatg 1500
aggaacgatc tggatgcaat ggctacaatc ccagcactga agtcatccat gaagatagaa 1560
gataccaaaa atgcattcga tctgaagcga ttagcagaga atgtttgcga tatcgacatg 1620
gaaattttag agctgccaaa gaagactttc aaattggcag atttgacaca ttatctactg 1680
ggattggagc tgcacaaaac tgaacaatgc agtaactggc aatgtcgtcc tctgcgaaaa 1740
aaacaaattg tgtacgcagc attggacgca gttgtcgtgg tggaaacatt caagaaaatc 1800
ttgtcgattg tagaggagaa aaacaaggac gcagatatcg agaagattgt cagagaatca 1860
aatgtaatgg ctccgaaaaa agacaaagga cacaaatcgt accgcaagct gaaaactatt 1920
ccatggcttg agctctacga tatcttgcca agccatcgta atcctacgag atcaccacag 1980
cgaccacacg acattaaagt tattgttgac acaatgctga ttggatttgg aaagaatctg 2040
aggagagttg gaattgatgt tattcttcca aaagacgtga gcgacttccg aaagtacctg 2100
aaggaaattg aacgagttgg cggcgagcat ctacgtcata taatcacagt gccatcgaaa 2160
agttacgaag ccttgaaaat ggattatgat aattatacaa ttgcaattcc ggaactcaat 2220
aacatgtctc ccgtagatca gcttattgag tttttcgacc tgttcaacgt agatattcgt 2280
ccggaagacg tatatcctcg atgcactgaa tgcaattctc ggcttcaaat taaattcccc 2340
ggaccagttt tacatttctt gcaccaatac tgtgtcatcc atgtgcaaaa tgtttatcgt 2400
gcgatatga gcgaatttcc actggaagaa tgggtggaatc gtatgcttca tatcaatcca 2460
gacgactacg acggagtaaa agtggaaatg tgcgcacat ctccaacaag caagtggatc 2520
gtggcaactg ttcccacagg atgcctacat attacgcgac aaactgcact tcacaccaat 2580
ctgccagatg gaattgaagt tcgaatccac aaagtgcctg atgacgagtt caagcgtcga 2640
aatctcagct tctatgtgtg tggagaatgc ggtacgggtg cttgtgatgg tctgtggcaat 2700
caagcgtcgg agagcaccag ccaggaatgt tga 2733

```

<210> 16

<211> 910

<212> PRT

<213> C. elegans

<400> 16

```

Met Glu Glu Glu Pro Tyr Lys Arg Lys Leu Thr Lys Ala Glu Lys Lys
 1              5              10              15

Ala Lys Tyr Arg Thr Asp Tyr Ala Glu Pro Leu Lys Ser Arg Arg Glu
          20              25              30

Val Leu Lys Ala Ile Met Asn Gly Pro Glu Ser Glu Arg Glu Arg Lys
          35              40              45

Val Arg Ala Lys Asn Arg Glu Phe Phe Asn Glu Asp Tyr Arg Ser Gly
          50              55              60

Val Asn Ile Tyr Gly Met Ala Val Asp Met Met Lys Ala Met Pro Asp
          65              70              75              80

Arg Gly Lys Thr Ser Gly Gln Ser Leu Ala Val Trp Tyr Leu Glu Asp
          85              90              95

Phe Gly Val Trp Leu Lys Glu Ser Gly Gln Glu Thr Glu Leu Arg Gln
          100              105              110

Lys Tyr Leu Thr Gly Thr Ile Gln Ile Asn Ala Leu Asp Val Cys Thr
          115              120              125

Ile Gly Gln Lys Gln Leu Leu Ser Glu Ile Phe Asp Ile Thr Lys Glu
          130              135              140

Lys Phe Thr Glu Asp Ile Thr Gln Leu Leu Asp Ala Ala Ile Lys Lys
          145              150              155              160

Gln Asp Phe Ser Val Ala Ala Asp Met Ala Ile Gln Tyr Asn Leu Leu
          165              170              175

Arg Asp His His Phe Glu His Leu Val Leu Pro Leu Met Leu Ser Gly
          180              185              190

Lys Asp Gln Thr Ala Tyr Lys Leu Ile Ser Asn Asn Glu Arg Met Gln
          195              200              205

Gln Gln Leu Val Glu Phe Phe Asp Arg Met Val Gly Ile Ser Val Val
          210              215              220

Ala Val Glu Glu Met Leu Lys Pro Tyr Lys Glu Thr Lys Ile Met Thr
          225              230              235              240

Ile Pro Met Glu Lys Leu Thr Gly Lys Thr Leu Asp Lys Leu Ile Ser
          245              250              255

Thr Ile Ile Asn Lys Asn Thr His Glu Tyr Asn Phe Ser Arg Glu Leu
          260              265              270

Ser Lys Phe Ala Lys Asn His Ser Gln Asn Gly Asn Leu Lys Ala Leu
          275              280              285

```

Lys Phe Asn Ile Ser Glu Arg Tyr Glu Lys Gly Lys Ser Asp Asp Asn
 290 295 300
 Tyr Phe Gln His Met Val Glu Thr Phe Thr Lys Ala Glu Asp Val Arg
 305 310 315 320
 Glu Pro Ile Leu Phe Tyr Leu Trp Ser Ser Asn Asp Thr Glu Lys Gln
 325 330 335
 Ile Asp Ala Ile Cys Phe Ala Ile Tyr Leu Gly Ile Ala Ser Ser Ser
 340 345 350
 Ser Tyr Gln Leu Pro Asn Val Met Arg Asp Phe Phe Arg Gln Pro Asp
 355 360 365
 Ser Lys Leu Arg Glu Ala Lys Glu Leu Leu Val Arg Arg Lys Thr Leu
 370 375 380
 Gln Val Pro Leu Asn Gly Glu Gln Leu Phe Val Phe Glu Asn Glu Arg
 385 390 395 400
 Arg Thr Gln Ile His Met Val Lys Thr Glu Ser Glu Met Asn Tyr Leu
 405 410 415
 Cys Ser Glu Ile Lys Ser Leu Ser Asp Glu Pro Ala Pro Val Tyr Val
 420 425 430
 Gly Phe Asp Ser Glu Trp Lys Pro Ser Asn Leu Thr Ala Val His Asp
 435 440 445
 Ser Lys Ile Ala Ile Ile Gln Leu Phe Phe Lys Asn Cys Val Trp Leu
 450 455 460
 Val Asp Cys Val Glu Leu Glu Lys Ala Asn Met Ala Asp Asp Trp Trp
 465 470 475 480
 Gln Lys Phe Ala Ser Arg Leu Phe Gly Asp Ser Pro Val Lys Val Val
 485 490 495
 Gly Phe Asp Met Arg Asn Asp Leu Asp Ala Met Ala Thr Ile Pro Ala
 500 505 510
 Leu Lys Ser Ser Met Lys Ile Glu Asp Thr Lys Asn Ala Phe Asp Leu
 515 520 525
 Lys Arg Leu Ala Glu Asn Val Cys Asp Ile Asp Met Glu Ile Leu Glu
 530 535 540
 Leu Pro Lys Lys Thr Phe Lys Leu Ala Asp Leu Thr His Tyr Leu Leu
 545 550 555 560
 Gly Leu Glu Leu Asp Lys Thr Glu Gln Cys Ser Asn Trp Gln Cys Arg
 565 570 575
 Pro Leu Arg Lys Lys Gln Ile Val Tyr Ala Ala Leu Asp Ala Val Val
 580 585 590

Val Val Glu Thr Phe Lys Lys Ile Leu Ser Ile Val Glu Glu Lys Asn
 595 600 605
 Lys Asp Ala Asp Ile Glu Lys Ile Val Arg Glu Ser Asn Val Met Ala
 610 615 620
 Pro Lys Lys Asp Lys Gly His Lys Ser Tyr Arg Lys Leu Lys Thr Ile
 625 630 635 640
 Pro Trp Leu Glu Leu Tyr Asp Ile Leu Arg Ser His Arg Asn Pro Thr
 645 650 655
 Arg Ser Pro Gln Arg Pro His Asp Ile Lys Val Ile Val Asp Thr Met
 660 665 670
 Leu Ile Gly Phe Gly Lys Asn Leu Arg Arg Val Gly Ile Asp Val Ile
 675 680 685
 Leu Pro Lys Asp Val Ser Asp Phe Arg Lys Tyr Leu Lys Glu Ile Glu
 690 695 700
 Arg Val Gly Gly Glu His Leu Arg His Ile Ile Thr Val Pro Ser Lys
 705 710 715 720
 Ser Tyr Glu Ala Leu Lys Met Asp Tyr Asp Asn Tyr Thr Ile Ala Ile
 725 730 735
 Pro Glu Leu Asn Asn Met Ser Pro Val Asp Gln Leu Ile Glu Phe Phe
 740 745 750
 Asp Leu Phe Asn Val Asp Ile Arg Pro Glu Asp Val Tyr Pro Arg Cys
 755 760 765
 Thr Glu Cys Asn Ser Arg Leu Gln Ile Lys Phe Pro Gly Pro Val Leu
 770 775 780
 His Phe Leu His Gln Tyr Cys Val Ile His Val Gln Asn Val Tyr Arg
 785 790 795 800
 Ala Asp Met Ser Glu Phe Pro Leu Glu Glu Trp Trp Asn Arg Met Leu
 805 810 815
 His Ile Asn Pro Asp Asp Tyr Asp Gly Val Lys Val Glu Met Ser Arg
 820 825 830
 Pro Ser Pro Thr Ser Lys Trp Ile Val Ala Thr Val Pro Thr Gly Cys
 835 840 845
 Leu His Ile Thr Arg Gln Thr Ala Leu His Thr Asn Leu Pro Asp Gly
 850 855 860
 Ile Glu Val Arg Ile His Lys Val Pro Asp Asp Glu Phe Lys Arg Arg
 865 870 875 880
 Asn Leu Ser Phe Tyr Val Cys Gly Glu Cys Gly Thr Val Ala Cys Asp

										885											890											895
										Gly Arg Gly Asn Gln Ala Ser Glu Ser Thr Ser Gln Glu Cys																						
										900											905											910
<210> 17																																
<211> 4299																																
<212> DNA																																
<213> Homo sapiens																																
<400> 17																																
atgagtga	aaaaattg	aacaactg	cagcagcg	aatgtcct	atggatga	60																										
gtgcaga	aaagatgt	tgtagaag	agaaaggc	gtgttcgg	gagtgttt	120																										
gaagatga	cccccttc	agaattca	ggatccat	tgtatagt	cgatgcta	180																										
gattgtct	tctgtcag	agatattag	atgagtct	cagatggg	tgtggtgg	240																										
tttgacat	agtggcc	attataca	agagggaa	ttggcaa	tgactaat	300																										
cagttgtg	tttctgag	caaagtta	ttgtccac	tttcttcc	gtcagttt	360																										
cccaggga	taaaaatg	gcttgaaa	aaagcagt	aaaaggc	tgtagga	420																										
gaaggaga	agtggaa	tctacgtg	tttgatat	aattgaag	ttttgtgg	480																										
ttgacaga	ttgccaat	aaagctga	tgtacaga	cctggagc	taacagtc	540																										
gttaaac	tcttaggt	acagctct	aaagaca	ctatccgc	tagcaatt	600																										
agtaaat	ctctcact	ggaccag	ctgtatgc	ccactgat	ttatgctg	660																										
tttattat	accgaa	agagattt	gatgata	tgcaaagg	tgctata	720																										
aaagagga	aaatccta	tagcgaca	aacaaac	tgacttca	ctctgagg	780																										
gtgatgga	tggtctag	tcttccta	gctttcag	aattggaa	cccacgg	840																										
gtttctat	tactaaag	tatttcaga	aattctat	cactgagg	gatgata	900																										
gggtctac	acattgag	tgaactga	cccagcaa	atttaaac	attatcct	960																										
gaagattc	ctactggg	agtacaac	aaacaaat	gagaacat	agtttta	1020																										
cacgttga	atgaaac	ggaccaac	cttgatca	tagctaa	tgatggag	1080																										
gatgtact	gaaataa	ggaacgaa	gaagatgg	ttgaagat	agtagaac	1140																										
aacaaatt	aagagaat	ggaaagag	tgtttgat	cgtagata	tacagaac	1200																										
gaactcca	ttttgga	gcagtctc	gaagaata	ttagtgat	tgcttata	1260																										
tctactga	atttatct	caatgata	gaaaacga	cgctctat	aattgaga	1320																										
gatgaaga	tagaaatg	gatgctta	catttatc	ccaatgata	tgaaaacg	1380																										
acgtccta	taattgag	tgatgaag	ttagaaat	agatgctta	gtctttag	1440																										
aacctcaa	gtggcacg	agaacca	cattctaa	gcttaaaa	ggaaagaa	1500																										
ctgggtct	ctactaa	agaagaag	gatgatga	atgaagct	tgaagggg	1560																										
gaagatga	ataaggac	tttgtggc	gcaccca	aagagcaa	tacttgcc	1620																										
aagatgta	ttggccat	cagtttta	ccagttca	ggaaagt	tcattcag	1680																										
ttagaaga	gaagagata	tggtgtgt	atggcaac	gatatgg	gagtttgt	1740																										
ttccagta	cacctgtt	tgtaggca	attggcct	ttatctct	ccttattt	1800																										
ctgatgga	accaagt	acagctta	atgtcca	tcccagct	cttccttg	1860																										
tcagcac	cagaaaat	tctaacag	attaaatt	gtaaata	gattgtat	1920																										
gtaactcc	aatactgt	aggtaac	ggcctgct	agcaact	ggctgat	1980																										
ggtatcac	tcattgct	ggatgagg	cactgtat	ctgagtgg	gcatgatt	2040																										
agggattc	tcaggaag	gggctcca	aagacag	tgccaat	tccaatcg	2100																										
gcacttact	ctactgca	ttcttcaa	cggaagac	ttgtacgt	cttaaatc	2160																										
agaaatcc	agatcac	tactggtt	gatcgacc	acctgtat	agaagtta	2220																										
cgaaaaac	ggaatat	tcaggat	cagccatt	ttgtcaaa	aagttccc	2280																										
tggaat	aaggtcca	aatcatct	tgtccttc	gaaaaat	acaacaag	2340																										
acaggtga	ttaggaa	taatctat	tgtggaac	accatgcg	catgagtt	2400																										
agcacagg	aagacatt	tcataggt	gtaagaga	aaattcag	tgtcatag	2460																										
accatagc	ttggaatg	cattaata	gctgacat	gccaagtc	tcattacg	2520																										
gctccta	acatgga	atattatc	gagattgg	gagctgg	tgatggac	2580																										
caaagttc	gtcacgtc	ctgggctc	gcagacat	acttaaat	gcaccttc	2640																										

```

actgagatac gtaatgagaa gtttcgatta tacaatttaa agatgatggc aaagatggaa 2700
aaatatcttc attctagcag atgtaggaga caaatcatct tgtctcattt tgaggacaaa 2760
caagtacaaa aagcctcctt gggaattatg ggaactgaaa aatgctgtga taattgcagg 2820
tccagattgg atcattgcta ttccatggat gactcagagg atacatcctg ggacttttgt 2880
ccacaagcat ttaagctttt gtctgctgtg gacatcttag gcgaaaaatt tgggaattggg 2940
cttccaattt tatttctccg aggatctaata tctcagcgtc ttgccgatca atatcgagg 3000
cacagtttat ttggcactgg caaggatcaa acagagagtt ggtggaaggc tttttcccg 3060
cagctgatca ctgaggggatt cttggtagaa gtttctcggg ataacaaatt tatgaagatt 3120
tgcgccctta cgaaaaagg tagaaattgg cttcataaag ctaatacaga atctcagagc 3180
ctcatccttc aagctaataga agaatttgtt ccaaagaagt ttcttctgcc tagttcgaaa 3240
actgtatctt cgggcaccaa agagcattgt tataatcaag taccagttga attaagtaca 3300
gagaagaagt ctaacttggg gaagtatat tcttataaac catgtgataa gatttcttct 3360
gggagtaaca tttctaaaaa aagtatcatg gtacagtcac cagaaaaagc ttacagttcc 3420
tcacagcctg ttatttcggc acaagagcag gagactcaga ttgtgttata tggcaaattg 3480
gtagaagcta ggcagaaaca tgccaataaa atggatgttc cccagctat tctggcaaca 3540
aacaagatac tgggtggatat ggccaaaatg agaccaacta cgggtgaaaa cgtaaaaagg 3600
attgatgggtg tttctgaagg caaagctgcc atgttggccc ctctgttggg agtcatcaaa 3660
catttctgcc aaacaaatag tgttcagaca gacctctttt caagtacaaa acctcaagaa 3720
gaacagaaga cgagtctggg agcaaaaaat aaaatatgca cactttcaca gtctatggc 3780
atcacatact ctttattcca agaaaagaag atgcctttga agagcatagc tgagagcagg 3840
attctgcctc tcatgacaat tggcatgcac ttatcccaag cggtgaaagc tggctgcccc 3900
cttgatttgg agcgagcagg cctgactcca gaggttcaga agattattgc tgatgttatc 3960
cgaaaccctc ccgtcaactc agatatgagt aaaattagcc taatcagaat gtagttcct 4020
gaaaacattg acacgtacct tatccacatg gcaattgaga tccttaaaaca tggctcctgac 4080
agcggacttc aaccttcatg tgatgtcaac aaaaggagat gttttcccg tttctgaagag 4140
atctgttcaa gttctaagag aagcaaggaa gaagtaggca tcaatactga gacttcatct 4200
gcagagagaa agagacgatt acctgtgtgg tttgccaaag gaagtgatac cagcaagaaa 4260
ttaatggaca aaacgaaaag gggaggtctt tttagttaa 4299

```

<210> 18

<211> 1432

<212> PRT

<213> Homo sapiens

<400> 18

```

Met Ser Glu Lys Lys Leu Glu Thr Thr Ala Gln Gln Arg Lys Cys Pro
 1             5             10             15

Glu Trp Met Asn Val Gln Asn Lys Arg Cys Ala Val Glu Glu Arg Lys
      20             25             30

Ala Cys Val Arg Lys Ser Val Phe Glu Asp Asp Leu Pro Phe Leu Glu
      35             40             45

Phe Thr Gly Ser Ile Val Tyr Ser Tyr Asp Ala Ser Asp Cys Ser Phe
      50             55             60

Leu Ser Glu Asp Ile Ser Met Ser Leu Ser Asp Gly Asp Val Val Gly
      65             70             75             80

Phe Asp Met Glu Trp Pro Pro Leu Tyr Asn Arg Gly Lys Leu Gly Lys
      85             90             95

Val Ala Leu Ile Gln Leu Cys Val Ser Glu Ser Lys Cys Tyr Leu Phe
      100            105            110

```

His Val Ser Ser Met Ser Val Phe Pro Gln Gly Leu Lys Met Leu Leu
 115 120 125
 Glu Asn Lys Ala Val Lys Lys Ala Gly Val Gly Ile Glu Gly Asp Gln
 130 135 140
 Trp Lys Leu Leu Arg Asp Phe Asp Ile Lys Leu Lys Asn Phe Val Glu
 145 150 155 160
 Leu Thr Asp Val Ala Asn Lys Lys Leu Lys Cys Thr Glu Thr Trp Ser
 165 170 175
 Leu Asn Ser Leu Val Lys His Leu Leu Gly Lys Gln Leu Leu Lys Asp
 180 185 190
 Lys Ser Ile Arg Cys Ser Asn Trp Ser Lys Phe Pro Leu Thr Glu Asp
 195 200 205
 Gln Lys Leu Tyr Ala Ala Thr Asp Ala Tyr Ala Gly Phe Ile Ile Tyr
 210 215 220
 Arg Asn Leu Glu Ile Leu Asp Asp Thr Val Gln Arg Phe Ala Ile Asn
 225 230 235 240
 Lys Glu Glu Glu Ile Leu Leu Ser Asp Met Asn Lys Gln Leu Thr Ser
 245 250 255
 Ile Ser Glu Glu Val Met Asp Leu Ala Lys His Leu Pro His Ala Phe
 260 265 270
 Ser Lys Leu Glu Asn Pro Arg Arg Val Ser Ile Leu Leu Lys Asp Ile
 275 280 285
 Ser Glu Asn Leu Tyr Ser Leu Arg Arg Met Ile Ile Gly Ser Thr Asn
 290 295 300
 Ile Glu Thr Glu Leu Arg Pro Ser Asn Asn Leu Asn Leu Leu Ser Phe
 305 310 315 320
 Glu Asp Ser Thr Thr Gly Gly Val Gln Gln Lys Gln Ile Arg Glu His
 325 330 335
 Glu Val Leu Ile His Val Glu Asp Glu Thr Trp Asp Pro Thr Leu Asp
 340 345 350
 His Leu Ala Lys His Asp Gly Glu Asp Val Leu Gly Asn Lys Val Glu
 355 360 365
 Arg Lys Glu Asp Gly Phe Glu Asp Gly Val Glu Asp Asn Lys Leu Lys
 370 375 380
 Glu Asn Met Glu Arg Ala Cys Leu Met Ser Leu Asp Ile Thr Glu His
 385 390 395 400
 Glu Leu Gln Ile Leu Glu Gln Gln Ser Gln Glu Glu Tyr Leu Ser Asp

				405					410					415			
Ile	Ala	Tyr	Lys	Ser	Thr	Glu	His	Leu	Ser	Pro	Asn	Asp	Asn	Glu	Asn		
			420					425					430				
Asp	Thr	Ser	Tyr	Val	Ile	Glu	Ser	Asp	Glu	Asp	Leu	Glu	Met	Glu	Met		
		435					440					445					
Leu	Lys	His	Leu	Ser	Pro	Asn	Asp	Asn	Glu	Asn	Asp	Thr	Ser	Tyr	Val		
	450					455					460						
Ile	Glu	Ser	Asp	Glu	Asp	Leu	Glu	Met	Glu	Met	Leu	Lys	Ser	Leu	Glu		
465					470				475						480		
Asn	Leu	Asn	Ser	Gly	Thr	Val	Glu	Pro	Thr	His	Ser	Lys	Cys	Leu	Lys		
			485						490					495			
Met	Glu	Arg	Asn	Leu	Gly	Leu	Pro	Thr	Lys	Glu	Glu	Glu	Glu	Asp	Asp		
			500					505					510				
Glu	Asn	Glu	Ala	Asn	Glu	Gly	Glu	Glu	Asp	Asp	Asp	Lys	Asp	Phe	Leu		
	515						520					525					
Trp	Pro	Ala	Pro	Asn	Glu	Glu	Gln	Val	Thr	Cys	Leu	Lys	Met	Tyr	Phe		
	530					535					540						
Gly	His	Ser	Ser	Phe	Lys	Pro	Val	Gln	Trp	Lys	Val	Ile	His	Ser	Val		
545					550				555						560		
Leu	Glu	Glu	Arg	Arg	Asp	Asn	Val	Ala	Val	Met	Ala	Thr	Gly	Tyr	Gly		
			565					570						575			
Lys	Ser	Leu	Cys	Phe	Gln	Tyr	Pro	Pro	Val	Tyr	Val	Gly	Lys	Ile	Gly		
		580						585					590				
Leu	Val	Ile	Ser	Pro	Leu	Ile	Ser	Leu	Met	Glu	Asp	Gln	Val	Leu	Gln		
	595						600					605					
Leu	Lys	Met	Ser	Asn	Ile	Pro	Ala	Cys	Phe	Leu	Gly	Ser	Ala	Gln	Ser		
	610					615					620						
Glu	Asn	Val	Leu	Thr	Asp	Ile	Lys	Leu	Gly	Lys	Tyr	Arg	Ile	Val	Tyr		
625					630				635						640		
Val	Thr	Pro	Glu	Tyr	Cys	Ser	Gly	Asn	Met	Gly	Leu	Leu	Gln	Gln	Leu		
			645						650					655			
Glu	Ala	Asp	Ile	Gly	Ile	Thr	Leu	Ile	Ala	Val	Asp	Glu	Ala	His	Cys		
		660						665					670				
Ile	Ser	Glu	Trp	Gly	His	Asp	Phe	Arg	Asp	Ser	Phe	Arg	Lys	Leu	Gly		
	675						680					685					
Ser	Leu	Lys	Thr	Ala	Leu	Pro	Met	Val	Pro	Ile	Val	Ala	Leu	Thr	Ala		
	690					695					700						

Thr Ala Ser Ser Ser Ile Arg Glu Asp Ile Val Arg Cys Leu Asn Leu
 705 710 715 720
 Arg Asn Pro Gln Ile Thr Cys Thr Gly Phe Asp Arg Pro Asn Leu Tyr
 725 730 735
 Leu Glu Val Arg Arg Lys Thr Gly Asn Ile Leu Gln Asp Leu Gln Pro
 740 745 750
 Phe Leu Val Lys Thr Ser Ser His Trp Glu Phe Glu Gly Pro Thr Ile
 755 760 765
 Ile Tyr Cys Pro Ser Arg Lys Met Thr Gln Gln Val Thr Gly Glu Leu
 770 775 780
 Arg Lys Leu Asn Leu Ser Cys Gly Thr Tyr His Ala Gly Met Ser Phe
 785 790 795 800
 Ser Thr Arg Lys Asp Ile His His Arg Phe Val Arg Asp Glu Ile Gln
 805 810 815
 Cys Val Ile Ala Thr Ile Ala Phe Gly Met Gly Ile Asn Lys Ala Asp
 820 825 830
 Ile Arg Gln Val Ile His Tyr Gly Ala Pro Lys Asp Met Glu Ser Tyr
 835 840 845
 Tyr Gln Glu Ile Gly Arg Ala Gly Arg Asp Gly Leu Gln Ser Ser Cys
 850 855 860
 His Val Leu Trp Ala Pro Ala Asp Ile Asn Leu Asn Arg His Leu Leu
 865 870 875 880
 Thr Glu Ile Arg Asn Glu Lys Phe Arg Leu Tyr Lys Leu Lys Met Met
 885 890 895
 Ala Lys Met Glu Lys Tyr Leu His Ser Ser Arg Cys Arg Arg Gln Ile
 900 905 910
 Ile Leu Ser His Phe Glu Asp Lys Gln Val Gln Lys Ala Ser Leu Gly
 915 920 925
 Ile Met Gly Thr Glu Lys Cys Cys Asp Asn Cys Arg Ser Arg Leu Asp
 930 935 940
 His Cys Tyr Ser Met Asp Asp Ser Glu Asp Thr Ser Trp Asp Phe Gly
 945 950 955 960
 Pro Gln Ala Phe Lys Leu Leu Ser Ala Val Asp Ile Leu Gly Glu Lys
 965 970 975
 Phe Gly Ile Gly Leu Pro Ile Leu Phe Leu Arg Gly Ser Asn Ser Gln
 980 985 990
 Arg Leu Ala Asp Gln Tyr Arg Arg His Ser Leu Phe Gly Thr Gly Lys
 995 1000 1005

Asp Gln Thr Glu Ser Trp Trp Lys Ala Phe Ser Arg Gln Leu Ile Thr
 1010 1015 1020
 Glu Gly Phe Leu Val Glu Val Ser Arg Tyr Asn Lys Phe Met Lys Ile
 1025 1030 1035 1040
 Cys Ala Leu Thr Lys Lys Gly Arg Asn Trp Leu His Lys Ala Asn Thr
 1045 1050 1055
 Glu Ser Gln Ser Leu Ile Leu Gln Ala Asn Glu Glu Leu Cys Pro Lys
 1060 1065 1070
 Lys Phe Leu Leu Pro Ser Ser Lys Thr Val Ser Ser Gly Thr Lys Glu
 1075 1080 1085
 His Cys Tyr Asn Gln Val Pro Val Glu Leu Ser Thr Glu Lys Lys Ser
 1090 1095 1100
 Asn Leu Glu Lys Leu Tyr Ser Tyr Lys Pro Cys Asp Lys Ile Ser Ser
 1105 1110 1115 1120
 Gly Ser Asn Ile Ser Lys Lys Ser Ile Met Val Gln Ser Pro Glu Lys
 1125 1130 1135
 Ala Tyr Ser Ser Ser Gln Pro Val Ile Ser Ala Gln Glu Gln Glu Thr
 1140 1145 1150
 Gln Ile Val Leu Tyr Gly Lys Leu Val Glu Ala Arg Gln Lys His Ala
 1155 1160 1165
 Asn Lys Met Asp Val Pro Pro Ala Ile Leu Ala Thr Asn Lys Ile Leu
 1170 1175 1180
 Val Asp Met Ala Lys Met Arg Pro Thr Thr Val Glu Asn Val Lys Arg
 1185 1190 1195 1200
 Ile Asp Gly Val Ser Glu Gly Lys Ala Ala Met Leu Ala Pro Leu Leu
 1205 1210 1215
 Glu Val Ile Lys His Phe Cys Gln Thr Asn Ser Val Gln Thr Asp Leu
 1220 1225 1230
 Phe Ser Ser Thr Lys Pro Gln Glu Glu Gln Lys Thr Ser Leu Val Ala
 1235 1240 1245
 Lys Asn Lys Ile Cys Thr Leu Ser Gln Ser Met Ala Ile Thr Tyr Ser
 1250 1255 1260
 Leu Phe Gln Glu Lys Lys Met Pro Leu Lys Ser Ile Ala Glu Ser Arg
 1265 1270 1275 1280
 Ile Leu Pro Leu Met Thr Ile Gly Met His Leu Ser Gln Ala Val Lys
 1285 1290 1295
 Ala Gly Cys Pro Leu Asp Leu Glu Arg Ala Gly Leu Thr Pro Glu Val

1300	1305	1310
Gln Lys Ile Ile Ala Asp Val Ile Arg Asn Pro Pro Val Asn Ser Asp		
1315	1320	1325
Met Ser Lys Ile Ser Leu Ile Arg Met Leu Val Pro Glu Asn Ile Asp		
1330	1335	1340
Thr Tyr Leu Ile His Met Ala Ile Glu Ile Leu Lys His Gly Pro Asp		
1345	1350	1355
Ser Gly Leu Gln Pro Ser Cys Asp Val Asn Lys Arg Arg Cys Phe Pro		
1365	1370	1375
Gly Ser Glu Glu Ile Cys Ser Ser Ser Lys Arg Ser Lys Glu Glu Val		
1380	1385	1390
Gly Ile Asn Thr Glu Thr Ser Ser Ala Glu Arg Lys Arg Arg Leu Pro		
1395	1400	1405
Val Trp Phe Ala Lys Gly Ser Asp Thr Ser Lys Lys Leu Met Asp Lys		
1410	1415	1420
Thr Lys Arg Gly Gly Leu Phe Ser		
1425	1430	

<210> 19
 <211> 30
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence:
 Oligonucleotide

<400> 19
 cgacatgatc tgatacatcg ttatgccatt

30

<210> 20
 <211> 29
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence:
 Oligonucleotide

<400> 20
 cattttataa taacgctgcg gacatctac

29

<210> 21
 <211> 1041
 <212> DNA

<213> *Arabidopsis thaliana*

<400> 21

```

atgtttgagt ttttcgcttc aggaggaagg tcgccgacac aagaagctaa tgagccacca 60
gttccgattt acattgtgac ggatccgttt caacttcctg ctgatttcct aaacccttct 120
cctgaaaaga aattggttat cggttttgac tgtgaggggtg ttgacctctg ccgacatggg 180
aaactttgta tcatgcagat tgcattctct aatgcaatat acttggttga tgtcatcgaa 240
gggtggagagg tgattatgaa agcgtgtaag cctgcactcg agtctaatta catcacgaaa 300
gttattcacg attgcaagcg tgacagttag gctctatact tccagtttgg gataagattg 360
cacaatgttg tggacactca gattgcttat tctctgattg aagaacaaga agggcggagg 420
agacctctag atgattacat atcgtttggt tcaactcctg ctgatccacg ttactgcggt 480
atatcctatg aagagaaaga agaagttcga gttctcatgc gccaggaccc aaagttttgg 540
acatacaggc ctatgactga gctcatgacg cgcgcagctg ctgatgatgt ccgcttcctt 600
ctgtatctct atcacaaaat gatgggaaag ctaaatacagc ggtcactatg gcattcttga 660
gttcgtgggtg ctttgtactg tcggtgtctc tgetgcatga atgatgctga ttttgctgat 720
tggtccaaccg ttcttccaat tccagttttt ctcgttaagg tcgtatatgc tgtagagaca 780
aagaaaaaaa gacgggtgac attagcttcg attgggttac tgattgtagt tggactttta 840
aatgtggcag ataacctgaa gtcagaagat caatgtcttg aagaagagat cctgtcagtg 900
cttgatgttc caccaggaaa gatgggacgt gtgattggaa ggaaaggagc atcgatcctc 960
gccattaagg aagcttgcaa cgcggaaatt ctaattggag gggcaaaggg tccacctgat 1020
aaggtagtgc ttattccata g                                     1041

```

<210> 22

<211> 346

<212> PRT

<213> *Arabidopsis thaliana*

<400> 22

```

Met Phe Glu Phe Phe Ala Ser Gly Gly Arg Ser Pro Thr Gln Glu Ala
 1              5              10              15

Asn Glu Pro Pro Val Pro Ile Tyr Ile Val Thr Asp Pro Phe Gln Leu
      20              25              30

Pro Ala Asp Phe Leu Asn Pro Ser Pro Glu Lys Lys Leu Val Ile Gly
      35              40              45

Phe Asp Cys Glu Gly Val Asp Leu Cys Arg His Gly Lys Leu Cys Ile
      50              55              60

Met Gln Ile Ala Phe Ser Asn Ala Ile Tyr Leu Val Asp Val Ile Glu
      65              70              75              80

Gly Gly Glu Val Ile Met Lys Ala Cys Lys Pro Ala Leu Glu Ser Asn
      85              90              95

Tyr Ile Thr Lys Val Ile His Asp Cys Lys Arg Asp Ser Glu Ala Leu
      100             105             110

Tyr Phe Gln Phe Gly Ile Arg Leu His Asn Val Val Asp Thr Gln Ile
      115             120             125

Ala Tyr Ser Leu Ile Glu Glu Gln Glu Gly Arg Arg Arg Pro Leu Asp
      130             135             140

```

Asp Tyr Ile Ser Phe Val Ser Leu Leu Ala Asp Pro Arg Tyr Cys Gly
 145 150 155 160
 Ile Ser Tyr Glu Glu Lys Glu Glu Val Arg Val Leu Met Arg Gln Asp
 165 170 175
 Pro Lys Phe Trp Thr Tyr Arg Pro Met Thr Glu Leu Met Ile Arg Ala
 180 185 190
 Ala Ala Asp Asp Val Arg Phe Leu Leu Tyr Leu Tyr His Lys Met Met
 195 200 205
 Gly Lys Leu Asn Gln Arg Ser Leu Trp His Leu Ala Val Arg Gly Ala
 210 215 220
 Leu Tyr Cys Arg Cys Leu Cys Cys Met Asn Asp Ala Asp Phe Ala Asp
 225 230 235 240
 Trp Pro Thr Val Pro Pro Ile Pro Val Phe Leu Val Lys Val Val Tyr
 245 250 255
 Ala Val Glu Thr Lys Lys Lys Arg Arg Val Thr Leu Ala Ser Ile Gly
 260 265 270
 Leu Leu Ile Val Val Gly Leu Leu Asn Val Ala Asp Asn Leu Lys Ser
 275 280 285
 Glu Asp Gln Cys Leu Glu Glu Glu Ile Leu Ser Val Leu Asp Val Pro
 290 295 300
 Pro Gly Lys Met Gly Arg Val Ile Gly Arg Lys Gly Ala Ser Ile Leu
 305 310 315 320
 Ala Ile Lys Glu Ala Cys Asn Ala Glu Ile Leu Ile Gly Gly Ala Lys
 325 330 335
 Gly Pro Pro Asp Lys Val Ser Leu Ile Pro
 340 345

<210> 23

<211> 1049

<212> DNA

<213> Arabidopsis thaliana

<400> 23

```

accaaagcat taatttttat ttttttgttt cagtaaaaga aatgtcatcg tcaaattgga      60
tcgacgacgc ttttacagag gaagagcttc tcgctatcga cgccatcgaa gcttcctaca      120
atttctcccg ttcttcttct tcttcttctt ctgctgctcc gaccgtacaa gctacaacct      180
ccgtccatgg ccacgaggag gatccaaatc aaatccccaa taatatccgt cgccaattgc      240
ctcgttccat cacttcttct acatcttata aacgatttcc tctctcccgt tgccgagcta      300
ggaatttttc agcaatgagg tttggttgta ggattttgta tagcaagact gctactgagg      360
ttgataagcg agcaatgcag cttattaaag ttcttgatac caagagagat gaatctggaa      420
tagcttttgt tggcttgat attgagtgga gaccaagttt tagaaaagggt gttctcccg      480
ggaaggttgc gactgtccag atatgtgtag atagtaatta ttgtgatgtt atgcatattt      540
ttcattctgg tatccctcaa agtctccaac atcttattga agattcaaca cttgtaaagg      600

```

```

taggtattgg aattgatggg gactctgtga agcttttcca tgactatgga gttagtatca 660
aagatgttga ggatctttca gatttagcca accaaaaaat tgggtggagat aaaaaatggg 720
gccttgccct actaactgag acacttggtt gcaaagagct cctgaagcca aacagaatca 780
ggcttgggaa ctgggagttt tctctctgtg caaagcagca gttacaatac gcagcaacgg 840
atgcttatgc ttcattggcat ctttacaagg ttcttaagga ccttcctgat gctgtcagtg 900
gctcataacg tgaaggagga agcttaaagg ttagcctata accccaagag ttagcatcaa 960
atgatatgat acacctaata tagtcaagta gatgcaattc ttgtgaatat tgtatctagt 1020
tctggtccct ttaaccgtcc agaaactag 1049

```

<210> 24

<211> 288

<212> PRT

<213> Arabidopsis thaliana

<400> 24

```

Met Ser Ser Ser Asn Trp Ile Asp Asp Ala Phe Thr Glu Glu Glu Leu
1          5          10          15

```

```

Leu Ala Ile Asp Ala Ile Glu Ala Ser Tyr Asn Phe Ser Arg Ser Ser
          20          25          30

```

```

Ser Ser Ser Ser Ser Ala Ala Pro Thr Val Gln Ala Thr Thr Ser Val
          35          40          45

```

```

His Gly His Glu Glu Asp Pro Asn Gln Ile Pro Asn Asn Ile Arg Arg
          50          55          60

```

```

Gln Leu Pro Arg Ser Ile Thr Ser Ser Thr Ser Tyr Lys Arg Phe Pro
65          70          75          80

```

```

Leu Ser Arg Cys Arg Ala Arg Asn Phe Pro Ala Met Arg Phe Gly Gly
          85          90          95

```

```

Arg Ile Leu Tyr Ser Lys Thr Ala Thr Glu Val Asp Lys Arg Ala Met
          100          105          110

```

```

Gln Leu Ile Lys Val Leu Asp Thr Lys Arg Asp Glu Ser Gly Ile Ala
          115          120          125

```

```

Phe Val Gly Leu Asp Ile Glu Trp Arg Pro Ser Phe Arg Lys Gly Val
          130          135          140

```

```

Leu Pro Gly Lys Val Ala Thr Val Gln Ile Cys Val Asp Ser Asn Tyr
          145          150          155          160

```

```

Cys Asp Val Met His Ile Phe His Ser Gly Ile Pro Gln Ser Leu Gln
          165          170          175

```

```

His Leu Ile Glu Asp Ser Thr Leu Val Lys Val Gly Ile Gly Ile Asp
          180          185          190

```

```

Gly Asp Ser Val Lys Leu Phe His Asp Tyr Gly Val Ser Ile Lys Asp
          195          200          205

```

```

Val Glu Asp Leu Ser Asp Leu Ala Asn Gln Lys Ile Gly Gly Asp Lys

```

210	215	220
Lys Trp Gly Leu Ala Ser Leu Thr Glu Thr Leu Val Cys Lys Glu Leu		
225	230	235 240
Leu Lys Pro Asn Arg Ile Arg Leu Gly Asn Trp Glu Phe Tyr Pro Leu		
	245	250 255
Ser Lys Gln Gln Leu Gln Tyr Ala Ala Thr Asp Ala Tyr Ala Ser Trp		
	260	265 270
His Leu Tyr Lys Val Leu Lys Asp Leu Pro Asp Ala Val Ser Gly Ser		
	275	280 285

<210> 25

<211> 22

<212> DNA

<213> Description of Artificial Sequence:
Oligonucleotide

<400> 25

ttcgggaacca ccatcaaaca gg

22

<210> 26

<211> 22

<212> DNA

<213> Description of Artificial Sequence:
Oligonucleotide

<400> 26

ttgctgcaac tctctcaggg cc

22

<210> 27

<211> 21

<212> DNA

<213> Description of Artificial Sequence:
Oligonucleotide

<400> 27

tcagctgttg cccgtctcac t

21

<210> 28

<211> 16

<212> DNA

<213> Description of Artificial Sequence:
Oligonucleotide
Other n=a, c, g, or t

<400> 28

wgtgnagwan canaga

16

<210> 29

<211> 27
 <212> DNA
 <213> Description of Artificial Sequence:
 Oligonucleotide

<400> 29
 gctccgccca cataattcaa acaacac 27

<210> 30
 <211> 22
 <212> DNA
 <213> Description of Artificial Sequence:
 Oligonucleotide

<400> 30
 ttcgaaaaca ttacctccga tc 22

<210> 31
 <211> 25
 <212> DNA
 <213> Description of Artificial Sequence:
 Oligonucleotide

<400> 31
 ggcttttgca ttggtatct actag 25

<210> 32
 <211> 25
 <212> DNA
 <213> Description of Artificial Sequence:
 Oligonucleotide

<400> 32
 atgtcatcgt caaattggat cgacg 25

<210> 33
 <211> 27
 <212> DNA
 <213> Description of Artificial Sequence:
 Oligonucleotide

<400> 33
 cgcttatcaa cctcagtagc agtcttg 27

<210> 34
 <211> 24
 <212> DNA
 <213> Description of Artificial Sequence:
 Oligonucleotide

<400> 34
 ttatgagcca ctgacagcat cagg 24

<210> 35

<211> 1749

<212> DNA

<213> *Arabidopsis thaliana*

<400> 35

ATGGGTTTGG ATTCTAAAGA AGCTGATTTG GAGGTAATAA GAGATGAGAA ATCTGAAGCA	60
AACACTGTGT GTTTACATGC GTTTTTCAGAT TTAACCTATG TGTCTCCTGT TGTGTTCTTA	120
TACCTACTCA AAGAATGCTA TAAACATGGT AGCTTGAAGG CAACAAAAAA GTTCCAAGCT	180
TTACAGTATC AAGTTCATCG AGTTCTAGCT AATAAACCTC AACCAGGACC TGCTACTTTC	240
ATTATTAAIT GTCTCACTTT ACTTCCITTA TTTGGGGTAT ATGGTGAAGG CTTTAGTCAT	300
TTAGTTATAT CAGCTCTTCG CCGCTTCTTT AAAACAGTAT CTGAACCAAC TAGTGAAGAA	360
GATATTTGTT TGGCGAGAAA GCTAGCTGCT CAGTTCTTCC TTGCTACTGT TGGTGGATCT	420
TTAACTTATG ATGAGAAGGT TATGGTGCAT ACTCTTAGAG TGTTTGATGT GAGGTTAACT	480
AGTATCGATG AAGCCTTGTC TATCTCGGAA GTTTGGCAGA GATATGGGTT TGCTTGTTGA	540
AATGCGTTTC TGAACAATA CATTTCTGAC TTGATCAAGT CGAAATCTTT CATGACGGCT	600
GTGACTCTGT TAGAGCATTT CTCTTTCCGT TTCCCTGGAG AAACCTTTCT TCAACAAATG	660
GTTGAGGATA AAAATTTCCA AGCTGCAGAG AGATGGGCTA CCTTCATGGG AAGGCCAAGT	720
TTATGCAITC TTGTTCAAGA GTATGGCTCA AGGAATATGC TAAAGCAGGC CTATAATATC	780
ATAAATAAGA ACTATCTACA GCATGACTTT CCCGAATTGT ATCACAAGTG TAAAGAAAAGT	840
GCTCTGAAGG TTCTAGCAGA AAAAGCATGT TGGGATGTTG CTGAAATTAA GACAAAAGGT	900
GATAGACAGC TTCTGAAGTA TCTGGTATAC TTGGCAGTGG AAGCTGGATA CTTGGAGAAG	960
GTTGATGAAC TGTGCGATCG ATATTACATT CAAGGGCTGC CAAAAGCACG AGAGGCTGAG	1020
GTTGCTTTTG TTGAAAAAAG CTTTCTGCGT CTCAACGATC TAGCTGTAGA AGATGTAGTT	1080
TGGGTTGATG AAGTCAACGA GTTGAGAAAA GCAACTTCTT TTCTTGAAGG ATGTAGAGTT	1140
GTGGGTATTG ACTGTGAATG GAAACCTAAT TATATTAAAG GCAGTAAACA GAACAAGGTT	1200
TCAATCATGC AAATTGATC TGATACCAA ATTTTCATAT TGGACTTGAT AAAGCTTTAC	1260
AATGACGCCT CTGAAATTCT GGACAACCTG CTTAGTCACA TTTTGCAATC GAAGAGTACA	1320
TTAAAGCTCG TCTCTCTGAC TGAGGATTAC CCTGATCATA AATTATCCTC AGGTTACAAT	1380
TTTCAATGTG ACATCAAGCA GTTGGCGCTT TCATATGGGG ATTTGAAATG TTTGAGCGA	1440
TACGACATGT TGCTAGACAT TCAAAATGTT TTTAATGAAC CATTTGGTGG TTTAGCAGGA	1500
CTAACGAAGA AAATATTGGG AGTGTCTTTG AACAAAACAA GACGCAATAG CGACTGGGAA	1560
CAAAGGCCCT TGAGCCAGAA TCAGCTTGAG TATGCTGCTC TTGATGCTGC AGTGTGATT	1620
CACATATTTT GCCATGTTTG CGATCATCCT CCACATGACA GTAGTTCAGA GACAACCCAG	1680
TGGAAATCTC ACATTGTAAG TACCTCTTAT AAAAGCCCTT ATCTTTCATC TGATAATTCA	1740
AGACGATAA	1749

<210> 36

<211> 582

<212> PRT

<213> Arabidopsis thaliana

<400> 36

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

275	280	285
Ala Cys Trp Asp Val	Ala Glu Ile Lys Thr Lys Gly Asp Arg Gln Leu	
290	295	300
Leu Lys Tyr Leu Val	Tyr Leu Ala Val Glu Ala Gly Tyr Leu Glu Lys	
305	310	315
Val Asp Glu Leu Cys	Asp Arg Tyr Ser Leu Gln Gly Leu Pro Lys Ala	
325	330	335
Arg Glu Ala Glu Val	Ala Phe Val Glu Lys Ser Phe Leu Arg Leu Asn	
340	345	350
Asp Leu Ala Val Glu	Asp Val Val Trp Val Asp Glu Val Asn Glu Leu	
355	360	365
Arg Lys Ala Thr Ser	Phe Leu Glu Gly Cys Arg Val Val Gly Ile Asp	
370	375	380
Cys Glu Trp Lys Pro	Asn Tyr Ile Lys Gly Ser Lys Gln Asn Lys Val	
385	390	395
Ser Ile Met Gln Ile	Gly Ser Asp Thr Lys Ile Phe Ile Leu Asp Leu	
405	410	415
Ile Lys Leu Tyr Asn	Asp Ala Ser Glu Ile Leu Asp Asn Cys Leu Ser	
420	425	430
His Ile Leu Gln Ser	Lys Ser Thr Leu Lys Leu Val Ser Leu Thr Glu	
435	440	445
Asp Tyr Pro Asp His	Lys Leu Ser Ser Gly Tyr Asn Phe Gln Cys Asp	
450	455	460
Ile Lys Gln Leu Ala	Leu Ser Tyr Gly Asp Leu Lys Cys Phe Glu Arg	
465	470	475
Tyr Asp Met Leu Leu	Asp Ile Gln Asn Val Phe Asn Glu Pro Phe Gly	
485	490	495
Gly Leu Ala Gly Leu	Thr Lys Lys Ile Leu Gly Val Ser Leu Asn Lys	
500	505	510
Thr Arg Arg Asn Ser	Asp Trp Glu Gln Arg Pro Leu Ser Gln Asn Gln	
515	520	525
Leu Glu Tyr Ala Ala	Leu Asp Ala Ala Val Leu Ile His Ile Phe Arg	
530	535	540
His Val Arg Asp His	Pro Pro His Asp Ser Ser Ser Glu Thr Thr Gln	
545	550	555
Trp Lys Ser His Ile	Val Ser Thr Ser Tyr Lys Ser Pro Tyr Leu Ser	
565	570	575

Ser Asp Asn Ser Arg Arg
580

<210> 37
<211> 1518
<212> DNA
<213> *Arabidopsis thaliana*

<400> 37
ATGGAGACCA ATCTAAAGAT CTATCTAGTT TCATCCACCG ACTCGTCCGA GTTCACTCAC 60
CTGAAATGGT CTTTCACTCG TTCTACGATC ATCGCCTTAG ACGCCGAATG GAAGCCACAA 120
CACTCCAATA CGTCGTTCGT TCCGACCGTC ACTCTCCTCC AAGTCGCATG CCGACTCAGT 180
CACGCCACGG ATGTCTCCGA TGTCTTCCTC ATTGATTGTA GTTCGATTCA TCTTCCATCG 240
GTTTGGGAGC TGTTGAATGA TATGTTCTGT TCGCCGGATG TTCTGAAACT AGGGTTTCGG 300
TTTAAACAGG ATTTGGTTTA CTTGTCTTCG ACATTTACTC AACATGGATG TGAAGGTGGA 360
TTCCAAGAGG TGAAACAATA CTTGGATATT ACAAGCATAT ACAATTATCT GCAACATAAG 420
CGGTTTGGGA GAAAGGCGCC AAAGGATATC AAGAGCTTGG CTGCTATATG TAAGGAAATG 480
CTGGACATCT CTCTCTCAAA GGAAC TTCAA TGTAGTGATT GGT CATATCG TCTCTTACA 540
GAAGACAGA AACTATACGC TGCCACAGAT GCTCACTGCC TGCTCCAGAT ATTCGATGTA 600
TTTGAGGCGC ATCTTGTTGA AGGAATCACA GTGCAAGATC TTAGAGTGAT AAATGTTGGC 660
TTACAAGAAA TTCTGACTGA ATCGGACTAT AGCAGTAAGA TTGTCACAGT CAAACTTTGC 720
AAGGCTACAG ATGTAATCAG ATCAATGTCG GAAAATGGTC AAAACATAGC CAATGGAGTG 780
GTTCCAAGAA AAACGACACT AAACACGATG CCAATGGATG AGAATTTGTT GAAGATTGTC 840
AGGAAGTTTG GAGAACGGAT CCTGTTGAAG GAGTCTGATC TTCTACCAA GAACTTAAG 900
AAGAAAACAA GAAGACGTGT CGCCTCAAGC ACTATGAACA CAAATAAGCA GTTGGTCTGT 960
TCTGCGGACT GGCAAGGTCC ACCGCCATGG GACTCATCTT TAGGCGGTGA TGGCTGCCCT 1020
AAATTTCTAT TGGATGTGAT GGTGAAGGT TTGGCGAAAC ATCTACGTTG TGTGGGGATT 1080
GATGCTGCAA TCCACACTC AAAGAAGCCG GATTCAAGGG AGTTGCTTGA TCAAGCATTC 1140
AAAGAGAACA GAGTTCTATT AACAAGAGAT ACAAATTGT TGAGACACCA GGATTTGGCA 1200
AAGCATCAAA TATATCGAGT AAAGAGTCTT CTTAAAAATG AGCAGCTACT TGAGGTGATA 1260
GAGACTTTCC AGCTAAAGAT CAGCGGAAAT CAGCTGATGT CCAGATGTAC GAAGTGCAAT 1320
GGGAAATTTA TTCAGAAACC TCTAAGCATT GAAGAAGCTA TTGAAGCAGC AAAGGGTTTC 1380
CAAAGAATAC CCAACTGCTT ATTTAACAAA AATTTAGAGT TTTGGCAGTG CATGAACTGC 1440
CATCAACTAT ACTGGGAGGG AACTCAGTAT CATAACGCAG TCCAGAAGTT CATGGAAGTA 1500
TGCAAGTTGA GTGAGTGA 1518

<210> 38
<211> 505
<212> PRT
<213> *Arabidopsis thaliana*

<400> 38

```

Met Glu Thr Asn Leu Lys Ile Tyr Leu Val Ser Ser Thr Asp Ser Ser
1           5           10           15

Glu Phe Thr His Leu Lys Trp Ser Phe Thr Arg Ser Thr Ile Ile Ala
          20           25           30

Leu Asp Ala Glu Trp Lys Pro Gln His Ser Asn Thr Ser Ser Phe Pro
          35           40           45

Thr Val Thr Leu Leu Gln Val Ala Cys Arg Leu Ser His Ala Thr Asp
          50           55           60

Val Ser Asp Val Phe Leu Ile Asp Leu Ser Ser Ile His Leu Pro Ser
65           70           75           80

Val Trp Glu Leu Leu Asn Asp Met Phe Val Ser Pro Asp Val Leu Lys
          85           90           95

Leu Gly Phe Arg Phe Lys Gln Asp Leu Val Tyr Leu Ser Ser Thr Phe
          100          105          110

Thr Gln His Gly Cys Glu Gly Gly Phe Gln Glu Val Lys Gln Tyr Leu
          115          120          125

Asp Ile Thr Ser Ile Tyr Asn Tyr Leu Gln His Lys Arg Phe Gly Arg
          130          135          140

Lys Ala Pro Lys Asp Ile Lys Ser Leu Ala Ala Ile Cys Lys Glu Met
145          150          155          160

Leu Asp Ile Ser Leu Ser Lys Glu Leu Gln Cys Ser Asp Trp Ser Tyr
          165          170          175

Arg Pro Leu Thr Glu Glu Gln Lys Leu Tyr Ala Ala Thr Asp Ala His
          180          185          190

Cys Leu Leu Gln Ile Phe Asp Val Phe Glu Ala His Leu Val Glu Gly
          195          200          205

Ile Thr Val Gln Asp Leu Arg Val Ile Asn Val Gly Leu Gln Glu Ile
          210          215          220

Leu Thr Glu Ser Asp Tyr Ser Ser Lys Ile Val Thr Val Lys Leu Cys
225          230          235          240

Lys Ala Thr Asp Val Ile Arg Ser Met Ser Glu Asn Gly Gln Asn Ile
          245          250          255

Ala Asn Gly Val Val Pro Arg Lys Thr Thr Leu Asn Thr Met Pro Met
          260          265          270

Asp Glu Asn Leu Leu Lys Ile Val Arg Lys Phe Gly Glu Arg Ile Leu
          275          280          285

```

Leu Lys Glu Ser Asp Leu Leu Pro Lys Lys Leu Lys Lys Lys Thr Arg
 290 295 300
 Arg Arg Val Ala Ser Ser Thr Met Asn Thr Asn Lys Gln Leu Val Cys
 305 310 315 320
 Ser Ala Asp Trp Gln Gly Pro Pro Pro Trp Asp Ser Ser Leu Gly Gly
 325 330 335
 Asp Gly Cys Pro Lys Phe Leu Leu Asp Val Met Val Glu Gly Leu Ala
 340 345 350
 Lys His Leu Arg Cys Val Gly Ile Asp Ala Ala Ile Pro His Ser Lys
 355 360 365
 Lys Pro Asp Ser Arg Glu Leu Leu Asp Gln Ala Phe Lys Glu Asn Arg
 370 375 380
 Val Leu Leu Thr Arg Asp Thr Lys Leu Leu Arg His Gln Asp Leu Ala
 385 390 395 400
 Lys His Gln Ile Tyr Arg Val Lys Ser Leu Leu Lys Asn Glu Gln Leu
 405 410 415
 Leu Glu Val Ile Glu Thr Phe Gln Leu Lys Ile Ser Gly Asn Gln Leu
 420 425 430
 Met Ser Arg Cys Thr Lys Cys Asn Gly Lys Phe Ile Gln Lys Pro Leu
 435 440 445
 Ser Ile Glu Glu Ala Ile Glu Ala Ala Lys Gly Phe Gln Arg Ile Pro
 450 455 460
 Asn Cys Leu Phe Asn Lys Asn Leu Glu Phe Trp Gln Cys Met Asn Cys
 465 470 475 480
 His Gln Leu Tyr Trp Glu Gly Thr Gln Tyr His Asn Ala Val Gln Lys
 485 490 495
 Phe Met Glu Val Cys Lys Leu Ser Glu
 500 505